

V. Nachhaltigkeit und digitale Technologien

KI im Journalismus und Nachhaltigkeit – Eine explorative Untersuchung

Bernhard Debatin

Abstract

Der Begriff der Nachhaltigkeit wird auf die Medienkommunikation übertragen, um als Massstab für die Bewertung der potentiellen Gefahren des Einsatzes von generativen KI-Systemen in der Massenkommunikation und den Sozialen Medien zu dienen. Die Nachhaltigkeit des Journalismus ist durch KI von zwei Seiten bedroht: Zum einen durch KI-gestützte Automatisierung und Rationalisierung und der damit einhergehenden Dequalifizierung und Delegitimierung der journalistischen Arbeit, zum anderen durch die gezielte und massenhafte Verbreitung von KI-generierten Falschmeldungen und *Deep Fakes* in sozialen Medien, was zu einer Unterhöhlung von Glaubwürdigkeit und Reputation des Journalismus führt. Im Blick auf die Nachhaltigkeit journalistische Qualität unter KI-Einsatz können durch die Implementierung von „Meaningful Human Control“ moralische Werte und Normen in KI-Entscheidungssysteme eingebaut, sowie menschliche Akteur:innen als letzte Entscheidungsinstanz eingesetzt werden. Bei der KI-gestützten Desinformation geht es um erhöhte Medienkompetenz, verstärkte journalistische Wachsamkeit und den Einsatz von KI-Tools zur schnellen Identifikation von *Fakes*, sowie um politische Regulierung und ethische Rahmenbedingungen für den Einsatz von KI. Reagieren Öffentlichkeit, Nutzer:innen und Politik nicht auf die Herausforderungen der KI-Systeme, steht zu befürchten, dass die Nachhaltigkeit im Journalismus den Sachzwängen der KI zum Opfer fällt. Neben der globalen Klimakrise ist dann eine tiefgreifende, globale Informationskrise zu befürchten.

1. Einleitung: Nachhaltigkeit und Mediennachhaltigkeit

Der Begriff der Nachhaltigkeit, wie von der UN Brundtland Kommission definiert, ist ein Balanceprinzip das gegenwärtige und zukünftige Bedürfnisse und Notwendigkeiten miteinander vermitteln soll (Brundtland 1987: 41), und zwar in den drei Pfeilern der ökonomischen Viabilität, der sozialen Gerechtigkeit und des ökologischen Schutzes (Purvis et al. 2019). Im Journalismus ist das Nachhaltigkeitsprinzip zunächst auf der *inhaltlichen* Ebene zu diskutieren, etwa im *Nachhaltigkeitsjournalismus*, der Umwelt-, Wissenschafts- und sozialpolitischen Journalismus kombiniert und sein Augenmerk auf Themen wie Nachhaltigkeit, Klimakrise, Verlust der Biodiversität und andere ökologische Problemfelder richtet. Dieser Aspekt der Nachhaltigkeit im Journalismus ist zwar zentral, jedoch soll es im Folgenden um ein anderes, ebenso wichtiges Problem gehen, die Frage der Nachhaltigkeit bei Entwicklung und Implementierung von künstlicher Intelligenz (KI) im Journalismus.

2. Nachhaltigkeit und generative KI

Übertragen wir das Nachhaltigkeitsprinzip auf die Frage nachhaltiger Entwicklung im Medienbereich, bedeutet dies zunächst, dass Entwicklung, Organisation und Inhalte von Medien darauf ausgerichtet sein sollen, informierte Öffentlichkeit und demokratische Entscheidungsprozesse langfristig zu ermöglichen und zu erhalten (Michelsen/Fischer 2016). Wie Irene Neverla feststellt,

„...ist es in einer mediatisierten Gesellschaft längst überfällig, die Perspektive der Nachhaltigkeit nicht nur auf Materie anzuwenden (Luft, Öl, Wasser, Mineralien etc.), sondern aus sozialwissenschaftlicher Sicht auf unseren Umgang mit nicht-materiellen Ressourcen, wie soziale Beziehungen, und aus kommunikationswissenschaftlicher Sicht auf unsere Formen und Inhalte von Kommunikation.“ (Neverla 2020: 335-336)

Ziel eines solchen nachhaltigen Journalismus ist die Bewahrung der Unabhängigkeit von Medienorganisationen, um dabei sowohl das Engagement der Bürger:innen zu fördern als auch eine unabhängige, glaubwürdige, vielfältige und akkurate Berichterstattung zu ermöglichen, was eine Balance zwischen ökonomischer Viabilität, sozialer Gerechtigkeit und journalistischer (bzw. herausgeberischer) Unabhängigkeit voraussetzt (Humbug et al. 2013, sowie Purvis et al. 2019). Dies war schon unter Bedingungen von Medienmonopolen und Konzentrationsprozessen schwierig (Bagdikian 2004), und es ist durch die disruptiven Effekte der digitalen Technologien, aber auch durch die veränderte geopolitische Situation (v.a. Russland und China) heute noch schwieriger geworden (Trinchini/Baggio 2023, vgl. auch Centre for Media Pluralism 2023). Wir müssen damit rechnen, dass die zunehmende Verwendung von generativer KI diese Situation noch verschärfen wird.

In den Medien findet generative KI findet immer breitere Verwendung. Dies betrifft zwei Bereiche, nämlich erstens ihren *Einsatz* innerhalb des professionellen Journalismus und in verwandten Berufssparten, und zweitens die *manipulative Verwendung* von KI durch *Bad Actors*, etwa in sozialen Medien, um Falschmeldungen, Zweifel, Propaganda, und *Deep Fakes* zu verbreiten.

Nachhaltigkeit im Journalismus ist hier durch KI doppelt bedroht: Zum einen durch die KI-gestützte Automatisierung und Rationalisierung im Journalismus und der damit einhergehenden Dequalifizierung und Delegitimierung der journalistischen Arbeit in einem Feld, das in den letzten 25

Jahren ohnehin einen massiven Arbeitsplatzabbau erlebt hat. Zum anderen durch die andauernden Angriffe gegen Reputation und Glaubwürdigkeit des Journalismus, was durch gezielte und massenhafte Verbreitung von KI-generierten Falschmeldungen und *Deep Fakes* in sozialen Medien massiv verschlimmert wird.

2.1 KI im Journalismus

Wie auch in anderen Feldern, hat sich KI im Journalismus bereits weltweit und rasant ausgebreitet. KI kann im Datenjournalismus arbeitsintensive Aufgaben beim Sammeln, Sortieren und statistischen Interpretieren von großen Datenmengen übernehmen, sowie Infografiken und andere nützliche Illustrationen generieren. KI kann auch bei anderen zeitraubenden Aufgaben hilfreich sein, wie z.B. der Transkription von Interviews, der Zusammenfassung von Texten, dem Fact-Checking und der Moderation von Kommentaren in Online-Foren (Prato 2023). Darüber hinaus vermag KI in engeren Sparten des Journalismus die Recherche und das Schreiben von Artikeln übernehmen. Seit 2014 nutzt z.B. die US-Nachrichtenagentur AP die KI-Schreibsoftware *WordSmith* für Reportagen über Universitäts-sport und Börsennachrichten. Insgesamt ist das Urteil über KI als Vollzeit-Journalist:in bislang aber eher negativ. Das Technology Magazin *The Verge* schreibt, dass KI heute zwar im Journalismus breit eingesetzt wird, jedoch viele KI-generierte Artikel Fehler oder anstößige Anspielungen enthalten (Drummond 2023). Beim Kölner Express werden seit einiger Zeit Sensationsnachrichten durch das KI-System Klara Indernach erstellt. Ein FAZ-Komentator bescheinigt dem System ein „besonderes Händchen für Hammer Schlagzeilen“, bewertet dessen Texte als nur „mit Vorsicht zu genießen“ und findet die Arbeit des Systems ebenso traurig wie ernüchternd (Weidemann 2023). Die amerikanische Journalistin Julia Angwin, Gründerin der Fact-Checking Agentur *Proof News* und Fellow am Harvard Kennedy School's Shorenstein Center on Media, Politics and Public Policy, hat sich intensiv mit der Rolle von KI im Journalismus beschäftigt und kommt zu dem Ergebnis, dass generative KI beim Artikel Schreiben zwar für einen ersten Entwurf brauchbar ist, dass sie dann aber für Korrigieren und Revidieren fast genauso viel Zeit aufwenden muss, als wenn sie den Artikel selbst geschrieben hätte (Merrefield 2024). Generative KI könne zwar plausibel klingende Texte generieren, doch haben diese keinen Bezug zu Akkuratheit und Faktizität, denn generative KI beruht auf plausibler

Wort Assoziation, die durch das Training mit großen Textmengen erreicht wird, jedoch nicht durch Abgleich mit der Wirklichkeit. Deshalb stellt sich bei generativer KI auch immer wieder das Problem „halluzinierter Information,“ bei der plausibel erscheinende Fakten oder Informationen vom System einfach erfunden werden (Lacy 2024).

Angwin betont, dass Klarheit und Transparenz über Umfang und Art der KI-Verwendung nötig, wie auch eine genaue Differenzierung zwischen den verschiedenen Typen von KI, da es ohnehin bereits einen starken Vertrauensverlustes gegenüber den Medien gibt. Zum Beispiel können KI-Systeme für statistische Analyse und Berichterstattung ebenso hilfreich sein wie KI-Anwendungen zur Gesichts- und Bilderkennung (Merrefield 2024). Für generative KI fällt jedoch bislang das Urteil eher negativ aus:

“Researchers in many fields have found that A.I. often struggles to answer even simple questions, whether about the law, medicine or voter information. Researchers have even found that A.I. does not always improve the quality of computer programming, the task it is supposed to excel at”. (Angwin 2014)

Gleichwohl breitet sich die KI im Journalismus zunehmend aus.

Die KI-gestützte Automatisierung des Journalismus kann zu einer Reihe von ethischen Problemen führen: Mit Blick auf die *Auswirkungen* der KI Technologie auf Menschen und Gesellschaft geht es, wie der unlängst beendet Streik der Scriptwriter in Hollywood eindrücklich gezeigt hat, v.a. um Fragen des Verlusts von qualifizierten Arbeitsplätzen und des Verschwindens von Autorschaft und geistigem Eigentum (Krausová/Moravec 2022). Als zentraler Bestandteil der sogenannten vierten Industriellen Revolution gilt KI als eine äußerst „disruptive Technology,“ die nicht nur erhebliche Innovationen und Umstrukturierungen mit sich bringt, sondern eben auch Arbeitsplätze vernichten und menschliche Denkarbeit weitgehend ersetzen kann (Păvăloaia/Necula 2023, Tucker 2023). Aus der Nachhaltigkeitsperspektive ist eine derart disruptive Technologie immer ein Problem, da sie die Bedürfnisse und Notwendigkeiten der Gegenwart radikal denen der Zukunft unterordnet.

Bei der konkreten Anwendung von KI im Journalismus geht es dagegen um Fragen des Trainings von KI-Systemen und um angemessene Fehlerkontrolle im Einsatz von KI. In einer weltweiten Studie mit 120 Journalist:innen und andere Medienarbeiter:innen aus 46 Ländern fanden die Forscher:innen von *JournalismAI*, dass die Journalist:innen besonders beunruhigt waren über intransparente Algorithmen und den in den Algo-

rithmus eingebauten Bias. Daneben wurden Verantwortungsprobleme, Ungenauigkeit, Fehlerneigung und Datenschutzbedenken genannt. Über 80% der Befragten befürchten, dass diese Probleme sich negativ auf die journalistische Qualität auswirken, und die gleiche Anzahl ist über die Zukunft des Journalismus besorgt (Beckett/Yaseen 2023: 43).

In einer anderen Studie wurde der Verlust journalistischer Autorität durch KI-gestütztes automatisiertes Fact-Checking untersucht (Johnson 2023). Journalistische Autorität ist hier zweifach bedroht: Zum einen durch die wachsende Abhängigkeit von nichtjournalistischen Akteur:innen, wie Programmierer:innen und KI-Trainer:innen, und zum anderen durch den wachsenden Verlust von Einfluss auf ein zentrales Kompetenzgebiet des Journalismus, das Fact-Checking (Johnson 2023: 16-18)

Soll es nicht um das bloße Überleben des Journalismus gehen, sondern um nachhaltige Entwicklungs- und Einsatzstrategien, dann müssen Probleme wie Arbeitsplatzverlust, mangelnde Transparenz und Fehlerkontrolle, unklare Verantwortung, sowie Reduzierung der journalistischen Autorität nicht als bloße Nebenfolgen und Kollateralschäden abgetan, sondern als zentrale wissenschaftliche, soziale, politische und ethische Herausforderungen im Übergang zur KI-gestützten digitalen Welt verstanden werden.

Im Blick auf journalistische Qualität könnte ein gangbarer Lösungsweg die Implementierung von „Meaningful Human Control“ sein, ein Ansatz aus der Diskussion um autonome Waffensysteme (Ekelhof 2019). Hier geht es darum zuerst moralische Werte und Normen in automatisierte Entscheidungssysteme einzubauen, und wo das nicht möglich oder nicht praktikabel ist, menschliche Entscheidungsträger:innen mit hineinzunehmen, die die Funktionen und Grenzen des Systems kennen und ein ethisches Verständnis der möglichen Probleme haben (Santoni de Sio/Mecacci 2021). Da so aber der Automatisierung Grenzen gesetzt werden, ist mit dem Widerstand der Industrie zu rechnen, denn hier soll ja gerade menschliche Arbeit durch KI-gestützte Automatisierung ersetzt werden.

Ein weiteres Nachhaltigkeitsproblem, das in der digitalen Welt gerne übersehen wird, ist die Tatsache, dass für den KI-Einsatz extrem hohe Rechenkapazitäten benötigt werden, was sich in erhöhten Elektrizitäts- und Wasserverbrauch, mehr Abwärme, und vermehrten CO₂-Ausstoß übersetzt. Diese Probleme sind bereits aus dem Cloud-Computing bekannt aber durch KI noch verschärft. Deshalb stellen Huang et al. in ihrem Artikel zur KI-Ethik fest: „The sustainability principle represents that the production, management, and implementation of AI must be sustainable and avoid environmental harm“ (Huang et al. 2023: 809). Jedoch ist der Stromverbrauch

in US-Datenzentren seit 2020 durch den zunehmenden Einsatz von KI erheblich angestiegen: Von 2015 bis 2019 lag der Verbrauch gleichbleibend bei jährlich 200 Terawattstunden, während sich der Verbrauch von 2020 bis 2023 auf über 400 TWS verdoppelt hat (Goldman Sachs 2024). Der zusätzliche Bedarf an Elektrizität kann nicht durch erneuerbare Energien allein gedeckt werden, so dass sich Gasproduzenten auf einen „significant spike in demand over the next decade“ vorbereiten (Kimball 2024). Die Forderung nach ethisch und nachhaltigkeitsorientierten KI-Standards ist in den letzten Jahren und Monaten lauter geworden (Berreby 2024 und Scheier 2024), jedoch sind solche Standards in der Regel freiwillige Selbstverpflichtungen (etwa die geplanten Standards der International Organization for Standardization, ISO, und der International Electrotechnical Commission, IEC; Diab/Mullane 2024). Auch wenn dies den Journalismus nur indirekt betrifft, so muss auch dieser Nachhaltigkeitsaspekt bei der Bewertung des KI-Einsatzes im Journalismus mit einbezogen werden.

2.2 KI und Desinformation

Eine besondere Herausforderung für nachhaltige Medienkommunikation stellen auch die immer häufiger zu findenden Synergien zwischen sozialen Medien und KI dar. Dies betrifft die ganze Breite von Kommunikation in sozialen Medien, von Falschnachrichten und Desinformation, über simuliertes User-Engagement und verstärkte Filter-Bubbles, bis hin zu zensurartiger Content-Moderation und flächendeckender Überwachung von Nutzer:innenverhalten.

KI-basierte Systeme steigern den in sozialen Medien ohnehin weit verbreiteten Confirmation Bias (Ciampaglia/Menczer 2019), etwa bei der KI-gestützten Bildung von Algorithmen, durch die Inhalte in sozialen Netzen gepusht werden, die die ohnehin schon vorhandenen Einstellungen und Ansichten der Benutzer:innen noch verstärken.

Ein weiteres und inzwischen viel diskutiertes Problem sind KI-generierte Inhalte, die zu Zwecken der Täuschung und Irreführung in soziale Medien eingespeist werden, wie Falschinformation, Plagiate, Deep Fakes und Desinformation. Dies hat unmittelbare Auswirkungen auf journalistische Reputation und Glaubwürdigkeit, da falsche Information evidenzgestützter journalistischer Information quasi gleichwertig gegenübersteht, was den schon lange verbreiteten Zweifel um des bloßen Zweifels Willen (Oreskes/Conway 2011) noch nährt. Darüber können Fehlinformationen gegeb-

nenfalls auch Gewaltbereitschaft erzeugen, wie etwa bei der Erstürmung des Kapitols am 6. Januar 2021 nach der Wahlschlappe von Donald Trump (Murugesan 2023). Speziell im Blick auf Wahlen sind digitale Gesellschaften damit einem Gefährdungspotenzial ausgesetzt, das Bürger:innen täuschen und polarisieren kann und damit auch zentral das Wahlergebnis zu beeinflussen vermag. In einem Artikel über die anstehenden US-Wahlen warnt der in New York lebende Tech Reporter Nick Robins-Early vor den Auswirkungen von KI-generierter Desinformation durch Politiker:innen, Lobbyist:innen, und Drittländer (Trinchini/Baggio 2023) bei gleichzeitiger Reduktion von Inhaltsmoderation in sozialen Medien (Robins-Early 2023). Durch den manipulativen Einsatz von generativen KI-Systemen bei Wahlen und anderen politischen Themen werden somit sowohl das Engagement der Bürger:innen reduziert als auch Unabhängigkeit, Glaubwürdigkeit, Vielfalt und Akkuratheit der Berichterstattung erodiert.

Versuche der digitalen Einflussnahme auf Wahlen sind nicht neu. In ihrer Studie „Cyberenabled foreign interference in elections and referendums“ fanden Sarah O’Connor et al. vom *Australian Strategic Policy Institute*, dass es zwischen 2010 und 2020 bei 41 Wahlen und sieben Referenden „cyber-enabled foreign interference“ gab und dass diese Versuche seit 2017 signifikant zugenommen haben (O’Connor et al. 2020: 3). Hauptakteur:innen sind der Studie zufolge Russland, China, Iran und Nordkorea. Es gibt keinen Grund anzunehmen, dass dies im Wahljahr 2024 oder den kommenden Jahren anders sein sollte, außer dass wir nun generative KI als zusätzliches Problem mit dabei haben (Schneider 2023).

Ein besonders eindrucksvolles Beispiel ist eine russische Desinformationskampagne, die vom *EU DisinfoLab* „Doppelgänger“ genannt wurde, da sie darauf beruht, Doppelgänger, d. h. falsche Klone, von etablierten Nachrichtenmedien zu erzeugen und zu verbreiten, darunter Bild, Spiegel, Süddeutsche Zeitung, FAZ, Die Welt, Tagesspiegel, t-Online, Reuters, Election Watch, The Economist, Ansa, The Guardian, Le Monde, Le Figaro, Le Parisien, 20 Minutes und RBC Ukraine. Die gefälschten Webseiten wurden in sozialen Medien, v.a. Facebook, durch fake User Accounts mit AI-generierten Gesichtern und Chatbots verbreitet (Alaphilippe et al. 2022). Auch wenn die Operation *Doppelgänger* entlarvt und durch Domain Sperrungen teilweise lahmgelegt wurde, so war das keineswegs das Ende dieser Kampagne. Im Juni 2023 veröffentlichte das französische *Sécrétariat générale de la défense et de la sécurité nationale* einen Bericht, der zeigte, dass die Aktivitäten der Doppelgänger Operation ungebrochen weitergehen und zum großen Teil auf die russische Organisation RRN (mit dem kuriosen

Namen „Reliable Recent News“) zurückgeführt werden können (Première Ministre 2023).

Ähnliche Aktivitäten wurden in einen Microsoft Bericht vom September 2023 aus China und Nordkorea dokumentiert. Hier wurden vor allem die sozialen Netzwerke Facebook und X (vormals Twitter) als Plattformen genutzt, wobei dies auch mit Versuchen verbunden war, durch Hacking und ähnliche Aktionen kritische Infrastruktur in den USA und ihren asiatischen Partnern lahmzulegen (Microsoft Threat Intelligence 2023). Im Dezember 2023 berichtete die *Insikt Group* einen signifikanten Trend der Weiterentwicklung in der Desinformationstechnik feststellte (Insikt Group 2023). In diesem Report wurde darauf verwiesen, dass die Hauptziele der Desinformationsoperation die Ukraine, Deutschland und die USA sind. Bei der USA-Kampagne fanden sich vor allem gefälschte Nachrichten über Präsident Bidens Politik und die anstehenden Präsidentschaftswahlen, über die U.S. Außenpolitik, sowie über LGBTQ-Themen. Diese Kampagne zielt auf die Vertiefung der ohnehin schon bestehenden Spaltung der US-Gesellschaft.

Im US-amerikanischen Vorwahlkampf sind KI-Fakes inzwischen zu einem allgegenwärtigen Alltagsproblem geworden. Am 8. Februar berichtete z.B. der öffentliche Sender NPR, dass Wähler in New Hampshire mit automatischen Anrufen überflutet wurden, in denen eine KI-generierte Stimme von Präsident Biden den Hörer:innen sagte, dass sie nicht zu den Vorwahlen gehen sollten. Hinter der Aktion steckte laut NPR ein texanisches Telemarketing Unternehmen (Bond 2024a). Experten befürchten, dass KI-generierte Fakes den Wahlkampf in den USA und anderswo erheblich beeinflussen werden (Heath 2024, sowie Sutherland/Chakrabarti 2024). Eine neue Studie des Brennen Center for Justice fand, dass der Einsatz von KI in Wahlkampagnen weltweit zunimmt, weshalb neben transparenter Dokumentation auch gezielte Verbote und politische Regulation von KI nötig seien (Panditharatne 2024).

Der Einsatz von KI zur Manipulation und zur Verbreitung von Deep Fakes hat die ohnehin prekäre Situation des Journalismus noch verschärft. Nachhaltigkeitsjournalismus im oben erläuterten Sinn (informierte Öffentlichkeit, Engagement der Bürger:innen, sowie unabhängige, glaubwürdige, vielfältige und akkurate Berichterstattung) wird in der Kakophonie von nichtjournalistischer Desinformation und KI-gestützten Fakes zusehends schwieriger. Fraglich ist deshalb, ob und wie potenzielle Gegenmittel schnell genug und flächendeckend zum Einsatz kommen können. Solche Gegenmittel sind v.a. erhöhte Medienkompetenz der Nutzer:innen,

verstärkte journalistische Wachsamkeit, sowie die rasche und breite Entlarvung von Falschmeldungen und Desinformation, und vor allem der gezielte Einsatz von KI zur Identifizierung und Ausschaltung von Desinformation und Deep Fakes. Darüber hinaus sind hier aber vor allem politische Regulierung und gesetzlich verankerte ethische Rahmenbedingungen notwendig. Dies ist jedoch schwierig, da es eine Koordination voraussetzt, die derzeit nicht gegeben ist. Hinzu kommt, dass auch Journalist:innen vom *Confirmation Bias* betroffen sind (Morewedge 2022) und unter dem Einfluss von Desinformationskampagnen selbst zu Verteiler:innen von Falschmeldungen werden können.

3. Ausblick

Diese vorläufige Bestandsaufnahme kann nicht mehr als ein Schnappschuss sein. Die Entwicklung in der KI und ihr Einsatz sind, wie gezeigt, nicht nur disruptiv, sondern auch äußerst dynamisch. Neben den stetig wachsenden *Large Language Model* KI-Netzwerken und den ebenso schnell wachsende Rechen- und Speicherkapazitäten, sind hier weitreichende Synergieeffekte zu erwarten, etwa durch die Kombination von verschiedenen KI-Anwendungen und durch die Kombination von KI und anderen disruptiven Technologien (wie z.B. *Blockchain* und das *Internet of Things*, Păvăloaia/Necula 2023). All dies wird in den kommenden Jahren und Jahrzehnten zu radikalen Veränderungen in den meisten, wenn nicht allen Bereichen der Gesellschaft führen, so dass wir neben der globalen Klimakrise auch eine massive Informationskrise als globale Herausforderung zu bewältigen haben.

Der Bereich der Medien und des Journalismus wird von diesen Veränderungen nicht ausgenommen sein. Wie oben gezeigt, wird die Nachhaltigkeit des Journalismus durch die disruptiven Effekte der KI bedroht, da letztere Bedürfnisse und Notwendigkeiten der Gegenwart denen der Zukunft unterordnet. Will sich der Journalismus nicht einfach auf dem Abstellgleis wiederfinden, dann müssen die Herausforderungen und Potenziale der KI im Journalismus verstanden und, mit der gebotenen Vorsicht, in die journalistische Praxis umgesetzt werden. Aber auch die Politik ist hier gefragt. Wenn schon führende KI-Entwickler:innen, wie der Nobelpreisträger und „AI-Godfather“ Geoffrey Hinton, nach einem Moratorium und besserer Regulierung verlangen, da KI grundlegende Risiken für Gesellschaft und Menschheit darstellt (Future of Life Institute 2023, und Taylor/Hern 2023), dann ist es an der Zeit, nicht nur robuste Sicherheits- und Regu-

lationssysteme für KI zu entwickeln (etwa das oben angeführte Prinzip der „meaningful human control“), sondern auch entsprechende gesetzliche Rahmenrichtlinien auf nationaler und internationaler Ebene zu setzen.

Literatur

- Alaphilippe, Alexandre/Machado, Gary/Miguel, Raquel/Poldi, Francesco (2022): *Doppelgänger – Media clones serving Russian propaganda (Full Report)*. Online verfügbar unter: <https://www.disinfo.eu/wp-content/uploads/2022/09/Doppelganger-1.pdf> (Abfrage am: 15.10.2024).
- Angwin, Julia (2024): *Press Pause on the Silicon Valley Hype Machine*. Online verfügbar unter: <https://www.nytimes.com/2024/05/15/opinion/artificial-intelligence-ai-opena-i-chatgpt-overrated-hype.html> (Abfrage am: 15.10.2024).
- Bagdikian, Ben H. (2004): *The New Media Monopoly: A Completely Revised and Updated Edition with Seven New Chapters*. Boston: Beacon Press.
- Beckett, Charlie/Yaseen, Mira (2023): *Generating Change: A global survey of what news organisations are doing with artificial intelligence*. Online verfügbar unter: https://static1.squarespace.com/static/64d60527c01ae7106f2646e9/t/656e400alc23e22da0681e46/1701724190867/Generating+Change+_+The+Journalism+AI+report+_+English.pdf (Abfrage am: 15.10.2024).
- Berreby, David (2024): *As Use of A.I. Soars, So Does the Energy and Water It Requires*. Online verfügbar unter: <https://e360.yale.edu/features/artificial-intelligence-climate-energy-emissions> (Abfrage am: 15.10.2024).
- Bond, Shannon (2024a): *AI fakes raise election risks as lawmakers and tech companies scramble to catch up*. Online verfügbar unter: <https://www.npr.org/2024/02/08/1229641751/ai-deepfakes-election-risks-lawmakers-tech-companies-artificial-intelligence> (Abfrage am: 15.10.2024).
- Bond, Shannon (2024b): *The FCC says AI voices in robocalls are illegal*. Online verfügbar unter: <https://www.npr.org/2024/02/08/1230052884/the-fcc-says-ai-voices-in-robocalls-are-illegal> (Abfrage am: 15.10.2024).
- Brundtland, Gro H. (1987): *Our Common Future: Report of the World Commission on Environment and Development*. Geneva, UN-Dokument A/42/427. Online verfügbar unter: <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf> (Abfrage am: 15.10.2024).
- Centre for Media Pluralism and Media Freedom (2023): *Monitoring media pluralism in the digital era: Application of the media pluralism monitor in the European Union, Albania, Montenegro, the Republic of North Macedonia, Serbia and Turkey in the year 2022*. In: EUI, RSC, Research Project Report, Centre for Media Pluralism and Media Freedom (CMPF), MPM. Online verfügbar unter: <https://hdl.handle.net/1814/75753> (Abfrage am: 15.10.2024).

- Ciampaglia, Giovanni Luca/Menczer, Fillipo (2019): *Misinformation and biases infect social media, both intentionally and accidentally*. Online verfügbar unter: <https://theconversation.com/misinformation-and-biases-infect-social-media-both-intentionally-and-accidentally-97148> (Abfrage am: 15.10.2024).
- Diab, Wael W./Mullane, Mike (2024): *How the ISO and IEC are developing international standards for the responsible adoption of AI*. Online verfügbar unter: <https://www.unesco.org/en/articles/how-iso-and-iec-are-developing-international-standards-responsible-adoption-ai> (Abfrage am: 15.10.2024).
- Drummond, Jade (2023): *Newsrooms around the world are using AI to optimize work, despite concerns about bias and accuracy*. Online verfügbar unter: <https://www.theverge.com/2023/9/28/23894651/ai-newsroom-journalism-study-automation-bias> (Abfrage am: 15.10.2024).
- Dubow, Ben (2024): *Will Russia be an artificial intelligence disinformation victim or perpetrator?* Online verfügbar unter: <https://www.themoscowtimes.com/2024/02/08/will-russia-be-an-artificial-intelligence-disinformation-victim-or-perpetrator-a84017> (Abfrage am: 15.10.2024).
- Ekelhof, M. (2019): *Moving Beyond Semantics on Autonomous Weapons: Meaningful Human Control in Operation*. In: *Global Policy* 10 (3), S. 343-348. Online verfügbar unter doi: 10.1111/1758-5899.12665.
- Fernandes, Marcos R. (2023): *Confirmation bias in social networks*. In: *Mathematical Social Sciences* 123, S. 59-76. Online verfügbar unter doi: 10.1016/j.mathsocsci.2023.02.007.
- Future of Life (2023): *Pause Giant AI Experiments: An Open Letter*. Online verfügbar unter: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (Abfrage am: 15.10.2024).
- Goldman Sachs (2014): *AI is poised to drive 160% increase in data center power demand*. Online verfügbar unter: <https://www.goldmansachs.com/intelligence/pages/AI-poised-to-drive-160-increase-in-power-demand.html> (Abfrage am: 15.10.2024).
- Grawitch, Matt (2023): *Confirmation bias in the era of large AI. How the input is framed can affect the output produced*. Online verfügbar unter: <https://www.psychologytoday.com/us/blog/a-hovercraft-full-of-eels/202305/confirmation-bias-in-the-era-of-large-ai> (Abfrage am: 15.10.2024).
- Heath, Ryan (2024): *Welcome to the generative AI election era*. Online verfügbar unter: <https://www.axios.com/2024/01/03/ai-elections-politics-chatbots-regulation> (Abfrage am: 15.10.2024).
- Huang, Changwu/Zang, Zeqi/Mao, Bifei/Yao, Xin, (2023): *An Overview of Artificial Intelligence Ethics*. In: *IEEE Transactions on Artificial Intelligence* 4 (4), S. 799-819. Online verfügbar unter doi: 10.1109/TAI.2022.3194503.
- Humbug, Anja/Fischer, Daniel/Marwege, Ronbin/Michelsen, Gerd (2013): *Welchen Journalismus braucht die Nachhaltigkeit? Bestandsaufnahme und Bedarfsanalyse in Wissenschaft und Praxis – Ergebnisse eines iterativen Delphis*. INFU-Diskussionsbeitrag 38/2013. Online verfügbar unter: https://www.leuphana.de/fileadmin/user_upload/Forschungseinrichtungen/infu/files/infu-reihe/38_13_INJ_DelphiSynthese_INFUReihe_20131025_gesamt.pdf (Abfrage am: 15.10.2024).

- Insikt Group (2023): *Obfuscation and AI Content in the Russian Influence Network "Doppelgänger" Signals Evolving Tactics*. Online verfügbar unter: <https://go.recordedfuture.com/hubfs/reports/ta-2023-1205.pdf> (Abfrage am: 15.10.2024).
- Johnson, Patrick R. (2023): *A Case of Claims and Facts: Automated Fact-Checking the Future of Journalism's Authority*. In: *Digital Journalism*, S. 1-24. Online verfügbar unter doi: 10.1080/21670811.2023.2174564.
- Kannengießner, Sigrid (2022): *Digitale Medien und Nachhaltigkeit. Medienpraktiken für ein gutes Leben*. Wiesbaden: Springer VS. Online verfügbar unter doi: 10.1007/978-3-658-36167-9.
- Kimball, Spencer (2014): *AI could drive a natural gas boom as power companies face surging electricity demand*. Online verfügbar unter: <https://www.cnn.com/2014/05/05/ai-could-drive-natural-gas-boom-as-utilities-face-surge-electric-demand.html> (Abfrage am: 15.10.2024).
- Krausová, Alžběta/Moravec, Václav (2022): *Disappearing Authorship: Ethical Protection of AI-Generated News from the Perspective of Copyright and Other Laws*. In: *Journal of Intellectual Property, Information Technology and E-Commerce Law* 13 (2), S. 132-144.
- Lacy, Lisa (2024): *Hallucinations: Why AI Makes Stuff Up, and What's Being Done About It*. Online verfügbar unter: <https://www.cnet.com/tech/hallucinations-why-ai-makes-stuff-up-and-whats-being-done-about-it/> (Abfrage am: 15.10.2024).
- Luedeke, Gesa/Boykoff, Maxwell T. (2017): *Environment and the media*. In: Douglas Richardson et al. (Hg.): *The International Encyclopedia of Geography*. Hoboken: Wiley & Sons, S. 1-8. Online verfügbar unter doi: 10.1002/9781118786352.wbieg0464.
- Merefield, Clark (2014): *Proof News founder Julia Angwin on trust in journalism, the scientific method and the future of AI and the news*. Online verfügbar unter: <https://journalistsresource.org/media/ai-journalism-julia-angwin/> (Abfrage am: 15.10.2024).
- Michelsen, Gerd/Fischer, Daniel (Hg.) (2016): *Nachhaltigkeit und Journalismus: Erkenntnisse und Impulse aus Wissenschaft und Praxis*. Bad Homburg: VAS Verlag.
- Microsoft Threat Intelligence (2023): *Digital threats from East Asia increase in breadth and effectiveness*. Online verfügbar unter: <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1aFyW> (Abfrage am: 15.10.2024).
- Morewedge, Carey (2022): *Confirmation bias in journalism: What it is and strategies to avoid it*. Online verfügbar unter: <https://journalistsresource.org/home/confirmation-bias-strategies-to-avoid-it/> (Abfrage am: 15.10.2024).
- Murugesan, San (2023): *The rise of ethical concerns about AI content creation: A call to action*. Online verfügbar unter: <https://www.computer.org/publications/tech-news/trends/ethical-concerns-on-ai-content-creation> (Abfrage am: 15.10.2024).
- Neverla, Irene (2020): *Nachhaltigkeit: Zur Theoretisierung einer politischen Kategorie*. In: *Publizistik* 65, S. 333-339. Online verfügbar unter doi: 10.1007/s11616-020-00593-5.

- O'Connor, Sarah/Hanson, Fergus/Currey, Emilia/Beattie, Tracy (2020): *Cyber-enabled Foreign Interference in Elections and Referendums*. In: Australian Strategic Policy Institute Policy Brief, Report No. 41/2020. Online verfügbar unter: <https://www.aspi.org.au/report/cyber-enabled-foreign-interference-elections-and-referendums> (Abfrage am: 15.10.2024).
- Oreskes, Naomi/Conway, Erik M. (2011): *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Climate Change*. London: Bloomsbury Publishing.
- Panditharatne, Mekela (2024): *Preparing to Fight AI-Backed Voter Suppression*. Online verfügbar unter: <https://www.brennancenter.org/our-work/research-reports/preparing-fight-ai-backed-voter-suppression> (Abfrage am: 15.10.2024).
- Păvăloaia, Vasile-Daniel/Necula, Sabina-Cristiana (2023): *Artificial Intelligence as a Disruptive Technology—A Systematic Literature Review*. In: *Electronics* 12 (1102), S. 1-37. Online verfügbar unter doi: 10.3390/electronics12051102.
- Prato, Carlo (2023): *10 Ways Journalists Use AI Tools in the Newsroom*. Online verfügbar unter: <https://www.twipemobile.com/10-ways-journalists-use-ai-tools-in-the-newsroom/> (Abfrage am: 15.10.2024).
- Purvis, Ben/Mao, Yong/Robinson, Darren (2019): *Three pillars of sustainability: In search of conceptual origins*. In: *Sustainability Science* 14, S. 681–695. Online verfügbar unter doi: 10.1007/s11625-018-0627-5.
- Reynaud, Florian/Leloup, Damien (2023): *'Doppelgänger': The Russian disinformation campaign denounced by France*. Online verfügbar unter: https://www.lemonde.fr/en/pixels/article/2023/06/13/doppelganger-the-russian-disinformation-campaign-denounced-by-france_6031227_13.html (Abfrage am: 15.10.2024).
- Robins-Early, Nick (2023): *Disinformation reimaged: How AI could erode democracy in the 2024 US elections*. Online verfügbar unter: <https://www.theguardian.com/us-news/2023/jul/19/ai-generated-disinformation-us-elections> (Abfrage am: 15.10.2024).
- Santoni de Sio, Filippo/Mecacci, Giulio (2021): *Four responsibility gaps with artificial intelligence: Why they matter and how to address them*. In: *Philosophy and Technology*, 34, S. 1057-1084. Online verfügbar unter doi: 10.1007/s13347-021-00450-x.
- Scheier, Robert (2024): *4 paths to sustainable AI*. Online verfügbar unter: <https://www.cio.com/article/1301912/4-paths-to-sustainable-ai.html> (Abfrage am: 15.10.2024).
- Schneider, Bruce (2023): *How AI Could Help China and Russia Meddle in U.S. Elections*. Online verfügbar unter: <https://www.scientificamerican.com/article/how-ai-could-help-china-and-russia-meddle-in-u-s-elections/> (Abfrage am: 15.10.2024).
- Sutherland, Paige/Chakrabarti, Meghna (2024): *AI's influence on election 2024*. Online verfügbar unter: <https://www.wbur.org/onpoint/2024/01/17/ai-influence-election-2024-politics> (Abfrage am: 15.10.2024).
- Taylor, Josh/Hern, Alex (2023): *'Godfather of AI' Geoffrey Hinton quits Google and warns over dangers of misinformation*. Online verfügbar unter: <https://www.theguardian.com/technology/2023/may/02/geoffrey-hinton-godfather-of-ai-quits-google-warns-dangers-of-machine-learning> (Abfrage am: 15.10.2024).

- Trinchini, Lino/Baggio, Rodolfo (2023): *Digital sustainability: Ethics, epistemology, complexity and modelling*. Online verfügbar unter: <https://firstmonday.org/ojs/index.php/fm/article/view/12934> (Abfrage am: 15.10.2024).
- Tucker, William (2023): *AI and How Disruptive Technology Has an Impact on Society*. Online verfügbar unter: <https://amuedge.com/ai-and-how-disruptive-technology-has-an-impact-on-society> (Abfrage am: 15.10.2024).
- Weidemann, Axel (2023): *K.I. beim Kölner „Express“: Wer ist „Klara Indernach“?* Online verfügbar unter: <https://www.faz.net/aktuell/feuilleton/medien/klara-indernach-der-koelner-express-hat-jetzt-eine-ki-redakteurin-19188224.html> (Abfrage am: 15.10.2024).