



Journalist Fellowship Paper

Autonomie und Autorenschaft: KI-Transparenz im hybriden Journalismus

Katharina Schell

Dezember 2024

Michaelmas Term

Sponsor: APA – Austria Presse Agentur

Inhalt

Inhalt	2
Einleitung	3
Conditio sine qua non!	4
KI-Transparenz im Journalismus – zu kurz gedacht	5
Diagnose: Misstrauensvorschuss	6
Was die User:innen denken.....	7
Transparenzvorschriften – und was sie uns nicht vorschreiben	10
KI-Guidelines von Medienhäusern.....	10
Der EU AI Act	11
Autonomie und Handlungsmacht im Journalismus	13
Autonomiefeld Entscheidungsmacht.....	13
Autonomiefeld Autorenschaft.....	15
Transparenzentscheidungen: Eine Autonomie-Matrix	18
Kennzeichnungsstrategien: Einfach kompliziert	20
Unerwünschte Nebenwirkungen	20
Der Ton macht die Musik	21
Kennzeichnungsstrategien: Ein Leitbild für prozedurale Einblicke	22
Fazit	25
Acknowledgements	27
Literatur.....	28

Einleitung

Mit zunehmender Etablierung von KI-Technologien in der Medienproduktion nimmt die Diskussion über adäquate Kennzeichnungsstrategien an Fahrt auf. Schnell ist das Schlagwort *Transparenz* zur Hand, doch die Umsetzung ist ungeklärt. Regulatorische Initiativen haben zu praktischen Fragen ebenfalls bis dato keine klaren Anforderungen formuliert.

Zentrale Fragen sind offen:

- Welche Faktoren sind ausschlaggebend für die Entscheidung für oder gegen einen Hinweis auf die Beteiligung von KI-Systemen in der journalistischen Produktion?
- Wie soll die Kennzeichnung aussehen?
- Welche Information soll sie zur Verfügung stellen?
- Wie kann die Kennzeichnung in die Medienproduktion und -distribution implementiert werden?

Ziel dieses Projekts ist die Entwicklung eines Toolsets, das Medienhäusern, Redaktionen und Journalist:innen dabei unterstützt, erklärbare Entscheidungen für oder gegen die Offenlegung von KI in ihren journalistischen Produkten zu treffen. Darauf aufbauend strebt das Projekt auch einen Vorschlag für die konkrete Ausgestaltung der Kennzeichnung an.

Diese Lösungsansätze sollen möglichst nachhaltig sein – das heißt, ausgehend von *aktuellen* Use Cases sollen die vorgeschlagenen Konzepte zumindest für einige Zeit noch mit der technologischen Entwicklung Schritt halten können.

Die Debatte über den Einsatz von Künstlicher Intelligenz im Journalismus ist erstaunlich technologiefixiert. Salopp formuliert: Alle starren auf die KI – niemand schaut näher auf den Journalismus. Das sollten wir ändern.

Als zentralen Arbeitsbegriff ziehe ich daher den Faktor Autonomie in der Mensch-Maschine-Relation heran. Die Annahme: Ausschlaggebend bei der Entscheidung für oder gegen eine Offenlegung des KI-Anteils ist das Ausmaß, in dem menschliche – redaktionelle – Autonomie an die Maschine über- und damit abgegeben wurde.

Die Integration autonomer maschineller Entscheidungsmechanismen in den journalistischen Workflow erweitert das Konzept des journalistischen Handlungsspielraums um eine neue Dimension. Ich erweitere daher das Konzept der *journalistischen Autonomie* um den Aspekt der *erzählerischen Autonomie*: Ich beschreibe Journalist:innen (auch) als Autor:innen, deren Aktionsradius durch einen neuen Akteur in der eigenen Domäne beschränkt wird.

Einige notwendige Definitionen

■ Was ist im vorliegenden Kontext mit Journalismus gemeint?

Dieses Projekt erarbeitet Lösungen für textbasierten Nachrichtenjournalismus. Für diese Abgrenzung habe ich mich aus mehreren Gründen entschieden:

1. Die Antwort auf die Frage, was als „KI-generiert“ zu betrachten ist, stellt sich als besonders komplex dar.¹ Aktuelle Regulierungsbemühungen, etwa der EU AI Act, eröffnen für KI-generierte oder -assistierte journalistische Texte einen beträchtlichen Interpretationsspielraum², anders als für visuelle Mediengattungen. Dies spiegelt sich auch in entsprechend unkonkreten Wordings in diversen Guidelines namhafter Medien wider (einen kurzen Überblick stelle ich hier zusammen).
2. Journalistische Texte sind nach wie vor jene Gattung, die einen prototypischen journalistischen Workflow von Informationsaufbringung über Inhaltserstellung bis zur Distribution exemplarisch abbilden können.
3. Nachrichtenjournalismus schließlich ist als *Hard News* jenes Segment in der breiten journalistischen Ausdruckspalette, das den unmittelbarsten gesellschaftlichen und demokratiepolitischen Impact für sich in Anspruch nehmen kann. Bedenken, dass Künstliche Intelligenz im Journalismus negative Effekte auf die Informationsgesellschaft entfalten kann, sind in dieser Sparte besonders ernst zu nehmen.

■ Wie lautet die Arbeitsdefinition für Künstliche Intelligenz im Kontext dieses Projekts?

Im aktuellen medialen Diskurs steht der Begriff *Künstliche Intelligenz* oder *Artificial Intelligence* nahezu ausschließlich, implizit und unzutreffend für Generative Künstliche Intelligenz.³

Für die Zwecke dieses Projekts definiere ich Künstliche Intelligenz als ein System, das in einem abgegrenzten, intellektuell definierten und journalistischen Bereich autonome Entscheidungen trifft und damit ein journalistisches Endprodukt erzeugt, verändert oder an dessen Erstellung oder Veränderung beteiligt ist.

KI in der Distribution des Endprodukts ist von dieser Definition insofern betroffen, wenn es um die Publikationsentscheidung bezüglich konkreter journalistischer Werke geht – das heißt, um die redaktionelle Entscheidung eines Systems, ob eine *bestimmte Story* veröffentlicht wird oder nicht. Content-Distributionssysteme, etwas eine algorithmische Kuratierung von Distributionsplattformen (Webseite eines Mediums), sind in dieser Definition nicht berücksichtigt.

Conditio sine qua non!

Bei allen Überlegungen in diesem Text gehe ich von zwei Prämissen aus:

- Redaktionen und Medienhäuser veröffentlichen keine Inhalte, für die sie KI verwendet haben, ohne gründliche redaktionelle Prüfung.
- Vollständige Deepfakes (etwa wie im EU AI Act definiert) sind jedenfalls zu kennzeichnen.

KI-Transparenz im Journalismus – zu kurz gedacht

Transparenz? Damit kennen wir uns aus. Die Fake-News-Debatte, um nicht zu sagen -Panik (seit Mitte der 2010er Jahre) hat den Begriff zu einem oft bemühten gemacht. Kaum ein Medium mit Qualitätsanspruch kommt daran vorbei, sich Transparenz auf die Fahnen zu heften. So what, wenn nun die KI ins Spiel kommt – bleiben wir transparent, adaptieren wir dieses Prinzip zeitgemäß, und die Sache ist erledigt.

Leider nein. Es scheint, als hätten wir das augenscheinlich erfolgversprechende Prinzip „Transparenz“ nie so richtig durchdacht: „For all its popularity and its potential importance in addressing journalistic and public concerns, discussion of transparency has suffered from a lack of clarity in its definition“⁴, lautet ein Befund, und ein [weiterer \(Jussi Latvala\)](#): „I have found no agreed-upon set of practices at the core of the journalistic transparency ideal.“⁵

Und nun kommt ein neuer Player daher, der uns unsanft darauf aufmerksam macht. Viel Text wurde in den vergangenen Monaten in Sachen zeitgemäße Interpretation von Transparenz im KI-Zeitalter produziert, und der „[normative Konsens](#)“ lautet: „Transparency is [...] as a deeply held normative value that aims to guide the behavior and conduct of journalists.“⁶ Doch kaum jemand kommt auf einen grünen Zweig, wie das in der Praxis gehandhabt werden soll.

Latvala hatte nicht KI im Sinn, als er im Jahr 2023 das journalistische Transparenz-Dogma auf den Prüfstand stellte. Er ging der mantrahaft wiederholten Überzeugung auf den Grund, dass Transparenz im Journalismus das *Vertrauen* der User:innen erhöhe.

Diese Intention liegt auch vielen öffentlichen Überlegungen zu KI-Transparenz im Journalismus zugrunde. „Transparency is a means of creating trust“,⁷ stellen die schwedischen Medienhäuser in ihrem Leitfaden „[AI Transparency in Journalism](#)“ klare Motivationsweichen. Ihr Dokument steht exemplarisch für die Hochkonjunktur an Absichtserklärungen von Medienhäusern und Redaktionen, die ihren verantwortungsvollen und transparenten Umgang mit KI in Guidelines niedergeschrieben haben.

Diagnose: Misstrauensvorschuss

Kaum einer dieser Leitfäden kommt ohne den Schlüsselbegriff Transparenz aus, ein deutliches Symptom übrigens des in mehreren aktuellen Arbeiten konstatierten Isomorphismus⁸ dieser Texte. Zumindest implizit sprechen alle Guidelines der Künstlichen Intelligenz gegenüber ihr Misstrauen aus. „A structural challenge to the right to information“ im Sinne von „the freedom to seek, receive and access reliable information“⁹ kann sie sein, warnt [RSF](#) eindringlich.

Dem medienseitigen Misstrauensvorschuss liegt die Angst vor Täuschung durch Technologie zugrunde. Das ist verständlich, berücksichtigt man die Begriffskarriere der *Transparenz* als Hausmittel gegen digitale Desinformation. Nun wird die KI als weiterer Risikofaktor in das Bedrohungsszenario aufgenommen: Dass KI-Systeme an sich kaum transparent sind, könnte journalistische Rechenschaft unterminieren und es den User:innen erschweren, die Vertrauenswürdigkeit von Inhalten einzuschätzen, so die Befürchtung.

„Audiences may feel deceived if they assume an AI-generated piece was written by a human. This erodes consumer trust over time“ – dieser [Befund](#) entstammt der Welt des Marketings¹⁰ und bringt daher auf marketingtypisch knackige Weise den Generalverdacht auf den Punkt: Künstliche Intelligenz ist etwas, dem zu misstrauen ist. Verwendet das Medium meines Vertrauens diese KI, ohne mir das zu sagen, verliere ich mein Vertrauen.

Interessant an dieser Argumentation ist der ambivalente Charakter des Misstrauensvorschusses: Einerseits wird die Technologie, die vielbeschworene *Black Box*, als nicht vertrauenswürdig eingeschätzt. Andererseits, so die Hypothese, verliert das Medienunternehmen, das diese Technologie einsetzt, an Vertrauen. Und die Hypothese, dass Transparenz dagegen helfen kann, wird unhinterfragt ins KI-Zeitalter transponiert.

Was die User:innen denken

Ein klarer Befund, dass die User:innen diesen Misstrauensvorschuss gegenüber KI in der journalistischen Produktion teilen, ist der Forschung nicht zu entnehmen. Wohl wird, etwa von [Altay & Girardi \(2024\)](#), eine generelle „AI aversion“ als „well-documented in the literature“¹¹ konstatiert, nicht aber, ob diese Abneigung in Bezug auf Journalismus in ähnlichem Ausmaß, stärker oder abgeschwächt vorliegt.

In den vergangenen Jahren standen zudem algorithmische Empfehlungssysteme und die Unterscheidbarkeit von automatisiertem Journalismus im Zentrum des [akademischen Interesses](#).¹² Jüngere Arbeiten widmen sich nun auch der generellen Einstellung zu KI im Journalismus, blicken auf konkrete Use Cases und können erste Hinweise auf userfokussierte Transparenzstrategien liefern.

Der [Digital News Report](#) (DNR) des Reuters Institute for the Study of Journalism fragte die Einstellung, konkret als eine Art Gefühlslage, für die Ausgabe 2024 erstmals in 28 Ländern ab: „Our qualitative research shows that people’s starting point is generally one of resistance, suspicion and fear“¹³, so ein vorläufiges Fazit des Forschungsteams. Rund die Hälfte (USA: 52 Prozent, Europa: 47 Prozent) der Befragten empfinden Unbehagen beim Gedanken an Nachrichten, die hauptsächlich von KI produziert werden. Mit dem assistierenden Einsatz von KI, bei dem nach wie vor die meiste Arbeit von Journalist:innen erledigt wird, schaut es allerdings anders aus: Damit geht es 42 Prozent in den USA und 33 Prozent in Europa recht gut.¹⁴

Die weiterführende Forschung des DNR-Teams zeigt aber auch, dass das Unbehagen mit weiteren Faktoren korreliert. Je höher die „AI awareness“, desto „relatively more comfortable“ sind User:innen mit KI-Anwendungen im Journalismus.¹⁵ Noch aber ist die *AI Awareness* nicht sonderlich ausgeprägt – 49 Prozent der Befragten (in allen befragten Märkten) gaben an, „small/nothing“ über KI gehört zu haben. Zugleich zeigen sich sozio- und demografische Unterschiede (jünger, männlich – diese Gruppen sind tendenziell eher mit KI vertraut).¹⁶

Die Einstellung der User:innen variiert auch je nach Thema: Für *Hard News*, etwa Politik-, Kriminalitäts- oder Lokalberichterstattung, wird ein KI-Einsatz weniger gern gesehen als etwa im Entertainment-Journalismus oder im Sport.¹⁷ Gerade letzteres mag auch in der Tatsache begründet sein, dass in der ersten Welle der Nachrichtenautomatisierung Sportberichterstattung als Parade-Use-Case für regelbasierte Generierung bekannt wurde.

Und schließlich gibt es Hinweise auf eine Use-Case-spezifische Akzeptanz, also abhängig davon, an welcher Stelle im journalistischen Workflow die KI am Werk war: „Depending on whether it is used behind the scenes, to deliver news in new ways or to generate entirely new content“¹⁸.

Zu bedenken ist, dass Aussagen über die Neigungslage überwiegend vom Status quo der Medienproduktion und -rezeption ausgehen. Dass die neuen Technologien das Potenzial haben, den Medienkonsum selbst, also die Art, wie User:innen von Journalist:innen bereitgestellte Informationen abrufen, maßgeblich zu verändern, können empirisch orientierte Untersuchungen kaum belegen.

Es ist wohl nicht allzu gewagt, anzunehmen, dass nicht zuletzt die Informationsvermittlung in Dialogform einen disruptiven Effekt haben wird. Ob Informationen, die über *conversational interfaces* wie Chatbots vermittelt werden, eine größere Vertrauenswürdigkeit zugesprochen wird,¹⁹ wurde bereits untersucht (in der gegenständlichen [Studie](#) lautet die Antwort: ja). Die Neigung zum Dialog mit der Maschine könnte einen „shift in how some users are interacting with news content, using AI not simply to read the news, but to process, interpret, and personalize it in ways that meet their individual needs“²⁰ [bringen](#).

Hinweise darauf, dass sich beim Gedanken an KI im Journalismus mittlerweile Pragmatismus auf User:innenseite breit macht, liefern die Ergebnisse einer aktuellen [Studie](#), in der an den Tag trat: „Much of the public think that journalists are currently using generative AI to complete certain tasks [...]“²¹, darunter Korrekturlesen, das Verfassen von Überschriften und das „Schreiben von Artikeln“. Eine tiefergehende Untersuchung möglicher Diskrepanzen zwischen angenommenem Einsatz und redaktioneller Realität wäre in diesem Zusammenhang wünschenswert. Derzeit kann nur angenommen werden, dass die User:innen – basierend auf dieser Online-Umfrage in sechs Ländern – den Einsatz von KI, zumal Generativer, in der Medienproduktion tendenziell überschätzen, was die Autoren zurecht und treffend als eine durchaus zynische Einschätzung gegenüber dem Journalismus beschreiben.²²

Umso spannender die Frage nach den Intentionen der User:innen, wenn sie denn mit Inhalten konfrontiert sind, die tatsächlich mithilfe von KI-Anwendungen entstanden und als solche gekennzeichnet sind. Die Antwort von [Piasecki et al.](#): „They want a choice and be able to exercise a certain level of agency. [...] Being able to filter news that has been written by an AI or a human.“²³

Die User:innen sind also gerade dabei, [sich eine Meinung zu bilden](#), was sie von KI im Journalismus halten sollen. Womöglich ein perfekter Zeitpunkt, aber auch eine Übergangsphase mit Ablaufdatum, um eine gut durchdachte Offenlegungsstrategie zu konzipieren: „Participants see AI technology as an important emerging technology that will dramatically impact their lives in the future; participants had positive and negative feelings about that fact, but many expressed a willingness to change their opinion about the risks and benefits of AI as the technology evolves.“²⁴

Transparenzvorschriften – und was sie uns nicht vorschreiben

Eine ungenügende Definition von KI, ein impliziter Misstrauensvorschuss und eine User:innenbasis, die selbst noch nicht so recht weiß, was sie von KI hält: Die Ausgangsbasis für das Verfassen von konkreten Vorgaben ist eher ungünstig.

Medienhäuser und Redaktionen, die sich Gedanken über Offenlegung von KI machen, werden vom oben skizzierten Misstrauensvorschuss gehemmt. „Much of the prevailing discourse surrounding the labeling and detection of AI-generated media is grounded in fears that such content could mislead or otherwise deceive members of the public.“²⁵ Eine Intention der *Täuschung* ist allerdings im journalistischen Kontext ein Vorwurf, der hart an der Ehrenbeleidigung vorbeischrämmt. „It is important to distinguish between the use of AI in editorial media compared to other types of digital platforms“²⁶, unterstreichen das auch die schwedischen Publisher.

KI-Guidelines von Medienhäusern

In jüngster Zeit sind mehrere aufschlussreiche [vergleichende Studien](#) zu diversen KI-Guidelines erschienen, die den Befund teilen, dass sich die meisten dieser Texte sehr – zu sehr, wie [manche meinen](#) – ähneln.²⁷ Im Folgenden fokussiere ich daher auf Bestimmungen zur Offenlegung in dreizehn ausgewählten Leitlinien, konkret auf Versuche, eine Schwelle zu definieren, ab der der Einsatz von KI in Texten jedenfalls transparent gemacht werden muss. Die Stoßrichtung der Texte kann grob in drei Kategorien unterteilt werden.

Zum einen gibt es Richtlinien, in denen der Einsatz von Generativer KI strikt abgelehnt wird, zu finden etwa bei [WIRED](#)²⁸ und bis zu einem gewissen Grad beim [Guardian](#)²⁹ (allerdings bereits mit einer Einschränkung und einem Autorisierungsprozedere in „Ausnahmefällen“). Es ist davon auszugehen, dass Richtlinien dieser Kategorie über kurz oder lang durchlässiger gestaltet werden, wie es bereits in einigen Medienhäusern geschehen ist. Die [APA – Austria Presse Agentur](#) schrieb in ihrer Leitlinie aus dem Jahr 2023 fest, sie verwende „keine Textbausteine von generativer KI in Textmeldungen“,³⁰ adaptierte dies aber in der Aktualisierung 2024 (siehe unten).

Zur zweiten Kategorie zähle ich Richtlinien mit allgemein formulierten Transparenzabsichten. Zu finden ist dieser Ansatz unter anderem bei der [BBC](#), die den User:innen stets ermöglichen will „[to] understand why we use it [AI, Anm.], how it works and how it affects them“³¹ oder bei Reuters: „[...] Thomson [Reuters](#) will implement practices intended to make the use of AI in our products and services explainable“³². Die Algemeen Nederlands Persbureau ([ANP](#)) überlässt eine Offenlegungsentscheidung der Redaktion, sollte sie dies „als passend erachten“³³.

Und schließlich liegen Texte vor, die eine Konkretisierung anstreben, wie ausgeprägt der Faktor AI sein muss, damit eine Kennzeichnung erforderlich ist. „Eine maßgebliche Rolle“ der KI ist es für den [Spiegel](#),³⁴ für den [Bayerischen Rundfunk](#) ist Transparenz bei „weitgehend automatisiert erstellten Inhalten“³⁵ geboten. „Ein bedeutender journalistischer Einfluss“³⁶ ist für [Aftonbladet](#) (Schweden) der Transparenz-Trigger, „Labeling is not required in cases where an AI tool is used only as an aid“³⁷, ist der Ausgangspunkt in der [Ringier-Gruppe](#). Die Deutsche Presse-Agentur [dpa](#) beschränkt die Kennzeichnungsabsicht für Anwendungsfälle, „wo Inhalte ausschließlich durch KI erzeugt werden“.³⁸ Die Associated Press ([AP](#)) nennt konkrete Use Cases für „Experimente“ mit AI (Übersetzung, Nachrichtenzusammenfassung, Titel), befindet aber nur die maschinelle Übersetzung für kennzeichnungspflichtig.³⁹ Die APA führt in ihrer aktualisierten internen Richtlinie den „Automatisierungsgrad“⁴⁰ als ausschlaggebend an. Die Initiative *Nordic AI Journalism* definiert den „significant journalistic impact“ nicht, sondern weist die Entscheidung dem Medienhaus zu, sie sollte im „editorial process“ stattfinden.

Nahezu alle gesichteten Guidelines streichen hervor, dass keine von KI modifizierten oder generierten Inhalte ohne redaktionelle Prüfung veröffentlicht werden. Ein weiteres gemeinsames Charakteristikum ist die Abwesenheit der User:innen als Zielgruppe der Transparenzbemühungen. Geraten aber die User:innen bei den Versuchen, den Transparenzbegriff auf die KI-Domäne auszuweiten, aus dem Blickfeld, enden diese allerdings in einem teleologischen Vakuum.

Der EU AI Act

Einen anderen Weg geht die Europäische Union in ihrer Regulierungsanstrengung, dem [EU AI Act](#). Dieser konstituiert ein rechtliches Rahmenwerk für KI-Systeme und kategorisiert sie nach Risikostufen. Außerdem wird die Transparenz während des gesamten Lebenszyklus von KI-Systemen betont, der die Phasen Entwurf,

Entwicklung, Einsatz und Überwachung nach der Markteinführung umfasst. Und er nimmt im für Medienschaffende relevanten Artikel 50 explizit die User:innen-Perspektive ein und setzt auf Täuschungsverhinderung: Wenn natürliche Personen mit KI-Systemen interagieren, müssen sie davon in Kenntnis gesetzt werden, heißt es klar im Artikel 50.⁴¹

Punkt 4 des Artikels 50 enthält Offenlegungsverpflichtungen für die Verbreitung von KI-generierten Inhalten. Er liest sich simpel genug: Komplette KI-generierte Bilder, Videos oder Audio-Materialien gelten als *Deepfakes* und müssen gekennzeichnet werden.

Tricky wird es aber aus Mediensicht im zweiten Absatz des Punkt 4: „Deployers of an AI system that generates or manipulates text which is published with the purpose of informing the public on matters of public interest shall disclose that the text has been artificially generated or manipulated. This obligation shall not apply [...] where the AI-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content.“⁴²

Erste Analysen, wie viel Interpretationsspielraum diese Formulierungen für die Kennzeichnung in der redaktionellen Praxis lassen, [sind verfügbar](#). Das Fazit in Kürze: sehr viel. Bei näherem Hinsehen ist nicht einmal klar, wie eine „direct interaction“ im Kontext Medienkonsum zu definieren oder was „text manipulation“⁴³ ist. Und welches Medium würde wissentlich *Deepfakes* verbreiten?

Eine oberflächliche Lektüre legt nahe, dass die internen Regeln, die sich Medienunternehmen bereits gegeben haben, weit strenger sind als die Vorgabe der EU-Regulierung, indem sie sich zu redaktioneller Kontrolle *und* einer Kennzeichnung von KI-Inhalten bekennen. Meiner Ansicht nach zu Recht wird daher aber auch hinterfragt, ob der AI Act die Intention, User:innen einen informierten Informationskonsum zu ermöglichen, überhaupt erreicht.⁴⁴

Autonomie und Handlungsmacht im Journalismus

Journalismus versteht sich als unabhängig und strebt nach größtmöglicher Autonomie. Diese „Beschwörung“ mag eine „rituelle Qualität“⁴⁵ haben, wie es [Örnebring & Karlsson](#) so schön formulieren; sie bestimmt aber die Vorstellung, welche Rolle Journalismus wahrnimmt und wie er die Handlungskraft dafür erlangt.

Doch die [Auffassung](#) von journalistischer Autonomie als „the idea that journalism as a societal institution, as well as individual journalists in their workplace (the newsroom), should be free from undue influence from other societal institutions and actors“⁴⁶ gerät unter Druck durch „questions of control, dependence and autonomy in the light of the ‘AI goldrush’ in the news“.⁴⁷

Betrifft mit einer teilautonomen Technologie ein weiterer Akteur die redaktionelle Werkstatt, der die Handlungskraft von Journalist:innen limitiert? Ja, meinen manche: „AI introduces a non-human actor into the newsroom that simulates acts of journalism usually associated with a human, such as news writing or content moderation.“⁴⁸

Auf die reichhaltige Forschungslage zum Spannungsverhältnis von journalistischer Autonomie, Handlungsmacht und Unabhängigkeit kann hier nur implizit eingegangen werden, ich greife daher auf die treffende und kompakte Definition von [Felix Simon](#) zurück: „a journalist’s capacity to act (agency) in a self-directing manner (autonomy)“.⁴⁹ Dieses Projekt konzentriert sich auf die Handlungskraft für redaktionelle Entscheidungen und setzt sie in Bezug zu den in den redaktionellen KI-Richtlinien zentralen Begriffen des *signifikanten journalistischen Impact*.

Autonomiefeld Entscheidungsmacht

Journalist:innen gilt „the level of editorial control“ als [ausschlaggebender Gradmesser](#) ihrer Autonomie.⁵⁰ Ich würde diese Definition durch „the authority to make and exercise decisions“ ergänzen. Wenn aber konstatiert wird, dass Journalist:innen heutzutage allenfalls „episodic autonomy over what they publish“ zukommt,⁵¹ scheint das Handlungsfeld begrenzt. Denn das Potenzial dieser Systeme, „journalists’ discretionary decisionmaking ability and therefore their agency and autonomy“⁵² einzuschränken, ist offensichtlich und längst nicht mehr subtil.

Das ist für sich betrachtet allerdings kein neues Phänomen. Journalismus ist ein technologiegetriebenes und -abhängiges Geschäft und Berufsbild. Journalismus, zumal in einem in Teilen formalisierten Umfeld wie dem Nachrichtenjournalismus, war schon lange „[vulnerable to automation](#)“⁵³ bzw. hat sogar selbst „strong elements of 'robotness' for a long time“.⁵⁴ Der von Møller als Argument angeführte [Churnalism](#) als „a form of routinised information packaging that lacks original ideas and instead recycles news content from other sources“⁵⁵ ist dafür ein Beispiel.

Was den gegenwärtigen Paradigmenwechsel jedoch von der Geschichte des permanenten digitalen Medienwandels abhebt, ist in der Tat, dass die Technologie nicht nur Entscheidungen der Redaktion potenziell verändert oder beeinflusst, sondern Entscheidungen *übernimmt*. Und dass ein *externer* Einflussfaktor von den Redaktionen *internalisiert* wird: Nicht nur sind entsprechende Applikationen „frequently developed or at least co-developed inhouse by powerful actors [within the institutional framework](#)“⁵⁶, oft sind auch Journalist:innen selbst an der Implementierung der Technologie in redaktionelle Workflows, mitunter federführend, beteiligt. Sie üben „procedural influences“ aus – und diese „are generally seen as the most important (and therefore the most important factors limiting journalistic autonomy)“.⁵⁷

Bezeichnend ist dabei, dass die Technologie ganz offen als neuer Akteur im redaktionellen System anerkannt wird. Zuschreibungen wie „Assistent“ (die APA nennt ihre hauseigene Generative-KI-Applikation „Text-Assistent“) sprechen zugleich eine deutliche Sprache: Der Akteur wurde mit journalistischem Sanktus onboardet, steht in der Hierarchie aber unter Redakteur:innen und Reporter:innen. Die Systeme haben eine „Autonomie zweiten Ranges“.⁵⁸

Die Autorität bleibt in dieser Vorstellungswelt der Mensch (der Journalist/die Journalistin), doch bestimmte Entscheidungen übernimmt die maschinelle „Assistenz“. Aber welche wesentlichen Entscheidungen fallen im prototypischen Redaktionsalltag? „At the level of daily news work, autonomy is understood as the extent to which journalists are free to decide story angles, what sources to use, and what narrative frames to employ.“⁵⁹ Ein exemplarischer Ablauf:

- Welche Story wird aufgegriffen/recherchiert?
- Welche Quellen und Informationen werden verwendet, welche verworfen?
- Tonalität, Perspektive, Format und Zielgruppe der Story
- Allfällige Zusatzinformationen und weitere Inhalte (Metadaten, Illustrationen etc.)
- Wann, wie und wo wird die Story publiziert?

Autonomiefeld Autorenschaft

„Was ist die Geschichte?“ Diese Frage wird weltweit in Newsrooms alltäglich tausendfach gestellt und beantwortet. Der Mensch ist ein Geschichtenwesen, der Informationen gern als Narrativ konsumiert. Journalist:innen erfüllen diesen Wunsch, nicht erst, seit *Storytelling* ein Buzzword wurde.

Aber die Betrachtung von journalistischer Autonomie und Handlungsmacht (nicht nur im Zusammenhang mit KI) weist einen blinden Fleck auf. Ein wesentlicher Faktor, nämlich Journalist:innen als Mitglieder einer (im vorliegenden Untersuchungssachverhalt) *schreibenden* Zunft, schlicht: als Autor:innen, kommt in der Journalismusforschung nur selten vor. Wenn doch, dann eher aus einem Rezeptionsinteresse heraus. Als beliebte Forschungsfrage wird etwa verfolgt, ob User:innen unterscheiden können, ob ein Text von einem Menschen oder von einer Maschine stammt.

Der Hype um Generative KI als potenzielle Bedrohung für Journalist:innen zielt allerdings genau auf die – nennen wir es narrative – Autorität und Autonomie von Journalist:innen ab. Wieder und wieder wird – man muss sagen: leider auch von Journalist:innen – in den Raum gestellt, dass Generative KI fast so gut oder besser schreibe als *der Mensch*.

Mein Lieblingsbeispiel dafür ist ein frühes Exemplar der GPT-Berichterstattung aus dem Hause [Guardian](#): „A Robot Wrote This Entire Article. Are You Scared Yet, Human?“⁶⁰ Dieser Artikel ist über vier Jahre alt, doch er legte den Tonfall für den öffentlichen, publikumsorientierten Diskurs über KI und Journalismus fest. Was in unserem aktuellen Zusammenhang noch spannender ist: Indem er suggerierte, dass der gesamte Text von „einem Roboter“ „geschrieben“ wurde, ignorierte und verzerrte er die journalistische Autor:innenleistung. Der *Guardian* bzw. jemand im redaktionellen Team editierte, kürzte und kompilierte mehrere Texte des Modells. Kuratierung, Montage und Publikations-Entscheidung: Das sind Kernaufgaben einer

narrativen Ordnungsmacht im journalistischen Erzählen, und diese Aufgaben übernahmen auch im *furchterregenden* Beispiel Menschen.⁶¹

Was also trennt das Schreiben vom Erzählen? Das fragt sich die Literaturwissenschaft gerne und seit Jahrhunderten. Vielleicht können wir ein wenig von ihr lernen. Schließlich gilt, wie [Ørmen & Gregersen](#) schreiben: „Journalists, like all other authors of narratives, organize their narratives by manipulating the sequence of events from a straightforward chronological sequence to a more effective one.“⁶²

Die Handlungsmacht von Autor:innen ist eine „vorsätzliche“, die bestimmte Intentionen verfolgt und Rede und Antwort zu ihrem Werk stehen kann.⁶³ Für [Van Woudenberg](#) spricht das eher dagegen, ChatGPT Attribute der Autorenschaft zu verleihen (der titelgebende „conservative view“ seines Artikels). Spinnen wir diese Definition aber im journalistischen Kontext – für den sie hervorragend geeignet ist – weiter, und nehmen wir an, dass der oder die journalistische Autor:in ihre Intentionen in Teilen an die Maschine delegiert, die dann in im Auftrag produziert: Die erzählerische Autonomie wäre geteilt, die Intention und die Rechenschaftspflicht lägen weiter beim Menschen, das Werk selbst aber wäre Aufgabe der KI.

„Geteilte Autorenschaft“ nennt [Hannes Bajohr](#) die Interaktion von Mensch-Maschine in einem Autor:innen-Netzwerk.⁶⁴ Anstatt Autorenschaft als rein menschliche Anstrengung zu betrachten, billigt dieses Modell nicht-menschlichen Akteuren Handlungsspielraum zu.

Der oder die journalistische Autor:in muss Macht abgeben. „[...] instead of regarding humans as absolute and 'authoritarian' authors, we propose [...] to regard humans, language, and technology as co-authors in the processes and performances of these generative models like ChatGPT“, beschreiben Coeckelbergh und Gunkel eine „[joint-agency](#)“.⁶⁵ Und – um in die nahe Zukunft zu blicken – bei wem liegt eigentlich die Autorenschaft, wenn künftig eine Userin den Chatbot einer Medien-Website nach Informationen fragt, die ein Journalist recherchiert hat?

Spätestens jetzt ist der Zeitpunkt, um Roland Barthes Diktum vom *Tod des Autors* hervorzukramen und sich an Michel Foucault zu erinnern, der eine Obsoleszenz des Autors vorhersagte. „Though Barthes and Foucault did not address themselves to LLM and generative AI, their work on authorship expertly anticipates our current

situation“, so der Befund von Coeckelbergh und Gunkel. Der Tag wird kommen, an dem es „überhaupt keinen identifizierbaren menschlichen Autor mehr“ gibt.⁶⁶ Oder ist es schon geschehen? Ja, geht es nach Ørmen und Gregersen. „[...] the much-discussed 'death of the author' in literary studies is altogether real in certain parts of journalism, where robots are now responsible for writing up certain types of journalistic pieces“.⁶⁷

Das klingt nun doch ein wenig depressiv. Um es klarzumachen: Der *Tod des Autors* im Journalismus muss weder heißen, dass es keinen Journalismus mehr geben wird, noch, dass künftig Journalist:innen zu „editors“ der Maschine degradiert werden.

Das Konzept einer geteilten Autorenschaft im Journalismus erlaubt uns aber, das mögliche Ausmaß der Disruption zu erahnen, die das KI-Zeitalter für den Journalismus und damit seine gesellschaftliche Funktion bedeuten kann. Das „[Retooling](#)“⁶⁸ ist nur ein Anfang. Die Transformation wird nicht mehr lediglich Geschäftsmodelle für Journalismus, Distributionswege für Journalismus, Formate oder Plattformen betreffen. „Reshaped“ wird, wenn man so will, die „[essence of journalism](#)“⁶⁹ und sein künftiger Impact auf unser Informationssystem.

Wir stehen an der Schwelle zu einem Zeitalter des [hybriden Journalismus](#),⁷⁰ in dem die Akteure menschlich oder maschinell sein können. Unsere Auffassung von journalistischer Autonomie haben wir dafür noch nicht adaptiert.

Im hybriden Journalismus haben sowohl Mensch als auch Maschine *Handlungsspielraum*. Das Gleichgewicht der Kräfte im journalistischen Prozess kann in den beiden Autonomiefeldern *redaktionelle Entscheidungsautorität* und *Autonomie der Autorenschaft* beschrieben werden.

Transparenzentscheidungen: Eine Autonomie-Matrix

Ich empfehle, das Kräfteverhältnis im hybriden journalistischen Prozess als Entscheidungskriterium für oder gegen die Offenlegung von KI in diesem Prozess zu verwenden.

Hier ist ein Use-Case-bezogener Ansatz sinnvoll. Als Anwendungsbeispiel zeige ich im Folgenden, wie ein solcher Rahmen genutzt werden kann.

Table 1: Impact matrix – weighting editorial and authorial autonomy in a human-machine-hybrid news production environment

This matrix outlines the tasks involved in hybrid news production, categorising them by the role of human, machine, or hybrid decision-making and their respective impacts on editorial and authorial control.

Stages in news production	Tasks		Editorial decisions made by...	Impact on authorial autonomy
News gathering	Data exploration	Data mining and aggregation	Machine	None
		Data extraction	Hybrid	None
		Data analysis	Hybrid	Low
	Fact-checking	Claim detection	Machine	Low
		Cross-referencing	Human	Low
	News and trend detection	Feed monitoring	Hybrid	Low
Content Creation		Content categorization and filtering	Machine	Mixed
		Speech-to-Text	Human	None
		Inspiration, brainstorming, ideation	Human	Mixed
	Story automation (rule based, data-to-text)	Human	Mixed	
	Story generation (GenAI)	Machine	High	
	Content reformatting	Article shortening	Hybrid	Mixed
		Story summarising/abstracting	Hybrid	Mixed
		Social media posts, newsletters	Hybrid	Mixed
		Headlines	Hybrid	Mixed

Table: Continues on page 18 • Created with Datawrapper

Table 1: Impact matrix (continued)

Stages in news production	Tasks	Editorial decisions made by...	Impact on authorial autonomy	
Content Creation	Proof-reading	Human	None	
	Translating	Hybrid	None	
	Speech-to-Text	Human	None	
	SEO writing	Hybrid	Mixed	
	Content tagging	Entities	Machine	None
Topics		Machine	None	
Content distribution	Automated content distribution	Machine	Mixed	
	Content clustering	Recommendation and personalisation	Machine	Mixed
		Topic clustering	Machine	Low
	Semantic search	Machine	None	
	Conversational interfaces	Machine	High	
	Content metrics, audience analytics	Machine	Mixed	

Table: Katharina Schell • Created with Datawrapper

Diese Matrix beschreibt Aufgaben in hybrider Nachrichtenproduktion und kategorisiert sie hinsichtlich der Rolle der menschlichen, maschinellen oder hybriden Entscheidungsfindung und ihren jeweiligen Auswirkungen auf die redaktionelle Autorschaft.

Zwei wichtige Punkte gilt es hervorzuheben:

- Die Liste der Anwendungsfälle ist nicht erschöpfend. Sie soll illustrieren, wie das System genutzt werden kann, um eine Entscheidung für oder gegen die Kennzeichnung von KI zu treffen.
- Diese Klassifizierung ist nicht risikobasiert. Sie ist jedoch geeignet, Aufgaben mit einem hohen Impact für die journalistische Autonomie zu identifizieren, der auf einen höheren Bedarf an Kontrolle zur Aufrechterhaltung ethischer Standards, Transparenz und journalistischer Integrität hinweisen kann.

Kennzeichnungsstrategien: Einfach kompliziert

Gehen wir auf Basis der bisherigen Überlegungen davon aus, dass eine Offenlegung von KI-Involvierung im Journalismus grundsätzlich wünschenswert ist, selbst wenn nicht genau klar ist, wie den „varying levels of algorithmic intervention“ Rechnung zu tragen ist.⁷¹ Gehen wir weiters davon aus, dass diese Kennzeichnung User:innen zu einer informierten Entscheidung über ihren Medienkonsum ermächtigen soll und dass Journalist:innen keine Täuschungsabsichten gegenüber ihren User:innen haben. Wie sollen diese Kennzeichen aussehen?

Die vordergründigen Antworten werden von Initiativen im visuellen Bereich dominiert, wie auch der Begriff [synthetic content](#) primär mit diesen Genres konnotiert ist.⁷² Ein visuelles Element mit einer visuellen Kennzeichnung auszustatten, scheint ein einfacheres Unterfangen und technologisch machbarer – wie die Coalition for Content Provenance and Authenticity ([C2PA](#)) vorexerziert.⁷³ Die Hauptmotive dieser Initiativen ist einerseits das Minimieren von Täuschungs- und Manipulationspotenzial, andererseits Fragen des Urheberrechts.

Unerwünschte Nebenwirkungen

Zuerst aber einmal eine Warnung vor unerwünschten Nebenwirkungen gut gemeinter Transparenz: Wenn (journalistische) Inhalte mit KI-Prozessen in Verbindung gebracht werden, können sie dadurch Vertrauenswürdigkeit verlieren. „[Labels could backfire](#)“:⁷⁴ Dieser Effekt wurde in mehreren Studien belegt. „We found that labeling headlines as AI-generated reduced the perceived accuracy of the headlines and participants’ intention to share them“,⁷⁵ stellten Altay & Gilardi fest. Die *AI-Abneigung* der User:innen könnte die „potentially detrimental consequence of disclosing use of generative AI, which may further exacerbate the already declining public trust in news outlets“⁷⁶ nach sich ziehen, so [Longoni et al.](#)

Wir sind konfrontiert mit einem „[paradox of AI disclosure](#)“: Nutzerinnen und Nutzer möchten Bescheid wissen, wie Medien KI einsetzen; kommt die Industrie diesem Wunsch nach, muss sie aber mit „potential reputational costs“ rechnen. „Rather than being rewarded for transparency, news organizations that disclose their use of these tools are perceived as less trustworthy and may therefore have fewer incentives to be so forthcoming“,⁷⁷ zu diesem Befund kommen Toff und Simon.

Unbedachte Labeling-Strategien können den Blick auf das gesamte Nachrichtenumfeld verzerren, diese Erfahrung wurde bereits mit der Kennzeichnung von *Fake News* gemacht: Wurden Inhalte als *Fake* gelabelt, erhielten unmarkierte Inhalte quasi automatisch höhere Vertrauenswerte. Dieser *implied truth effect* könnte im AI-Kontext zu einem „[implied authenticity effect](#)“ werden.⁷⁸

Im Worst-Case-Szenario wären in einem Nachrichtenumfeld korrekt und nach untadeligen Richtlinien gekennzeichnete journalistische Beiträge in den Augen der Audience weniger wert als generischer, nichtjournalistischer und nicht gekennzeichneter KI-Müll. „If the content is not harmful, the benefits of labeling it as AI-generated are debatable, especially given the negative effects of these labels“⁷⁹, zu diesem Schluss kommt die [Forschung](#). Das Informationsinteresse der User:innen werden wir so aber nicht stillen.

Welche Kommunikationsmaßnahmen können wir also setzen, um die User:innen zu ermächtigen, sich informiert für oder gegen den Konsum von Inhalten zu entscheiden? Und ihnen gleichzeitig nicht das Vertrauen in die journalistischen Inhalte zu nehmen?

Der Ton macht die Musik

Wenig überraschend finden User:innen Bezeichnungen wie „manipuliert“, „Deepfake“ oder „not real“ [wenig vertrauenserweckend](#).⁸⁰ Das Ziel wäre also, differenzierter zu beschreiben, was genau die KI gemacht hat. Es lohnt sich, das Konzept einer *prozessorientierten* Kennzeichnung zu prüfen. Sie beschränkt sich nicht auf die *Beschreibung* des Inhalts, sondern macht klar, wie das Stück Journalismus zustande kam und an welchen Punkten im *Prozess* die KI welche Aufgaben übernahm. Dies ermöglicht eine tendenziell „neutrale“ Positionierung der Kennzeichnung.⁸¹

Das könnte allerdings ganz schön viel Text werden, wenn wir davon ausgehen, dass bereits jetzt oder in naher Zukunft mehrere KI-Anwendungen an der Produktion eines Artikels beteiligt sind – oder gar eine Riege von [KI-Agenten](#) automatisierte Workflows abspult.⁸² Dabei sollte die Erklärung tunlichst [für die User:innen](#) verständlich formuliert sein: „Disclosure is meaningless if readers are unsure of how to interpret the labels“.⁸³

Die nordischen Medienhäuser schlagen das Wording „created with the support of“ vor, „to signal AI's actual impact on the content and remind the media consumer

that there is an editorial process (and staff) behind the content.“⁸⁴ Die als Anwendungsbeispiel für die Umsetzung genannte Veröffentlichung einer KI-generierten Illustration ist allerdings ein schwaches Argument. Gerade dieser Use Case ist so stark KI-geprägt, dass kaum lediglich „Unterstützung“ vorliegt. Würden User:innen sich davon nicht erst recht in die Irre geführt fühlen?

Wie man es dreht und wendet: „Deciding which 'flavor' of transparency to provide is one of the main challenges.“⁸⁵. Das gilt auch formal: In diversen Guidelines finden sich vereinzelt konkrete Kommunikationsmaßnahmen wie Byline, Textbox, Endnote oder Registereintrag – doch ein Großteil der Governance-Texte verzichtet auf eine Spezifizierung.

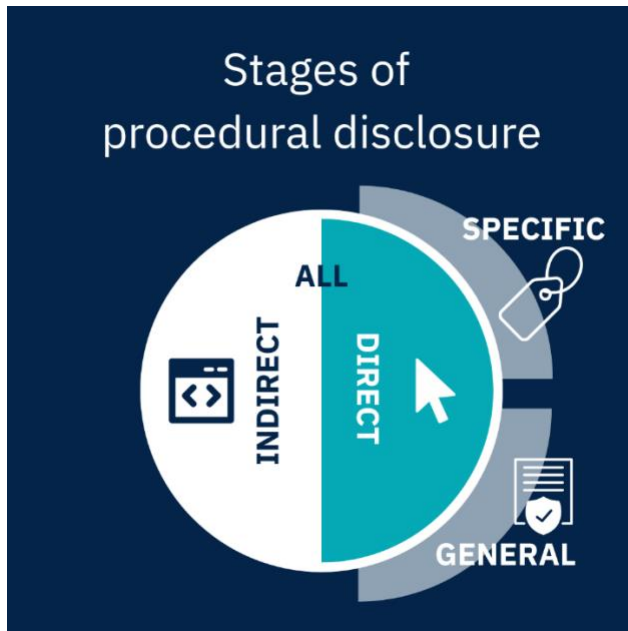
Die Fixiertheit auf Labels erscheint wenig zielführend. Es gibt Use Cases, für die ein simples visuelles Signal oder ein kurzer Erklär-Text ausreichen mag. Der zunehmenden Hybridisierung der journalistischen Prozesse können sie aber nicht gerecht werden. Je komplexer die Entscheidung „kennzeichnen ja oder nein“ wird, desto ungeeigneter ist eine *one-size-fits-all-solution*. Ein nachhaltig tragfähiges Transparenz-System „that empowers audiences to label themselves and to evaluate content manipulation, accuracy and context“⁸⁶ bedarf eines differenzierten Ansatzes.

Kennzeichnungsstrategien: Ein Leitbild für prozedurale Einblicke

Ich skizziere im Folgenden eine Offenlegungsstrategie für textbasierten Journalismus, die den journalistischen Prozess und seine Stationen transparent und nachvollziehbar erklärt und offenlegt, ob KI-Systeme in diesem Prozess beteiligt waren.

Damit wird zum einen ein Stigmatisierungseffekt (in den Augen der User:innen) vermieden, zum anderen werden die journalistische Leistung und die Redaktion als aktiv handelnder *Anwender* der Technologie in den Fokus gerückt. Wesentlich ist daher, dass das System für die gesamte Produktion des Medienhauses ausgerollt wird – nicht nur, wenn KI im Spiel war.

Eckpunkte sind Phasen der Offenlegung von journalistischen Content-Prozessen: *direkt, indirekt, spezifisch* und *allgemein*. Diese Phasen können weiter in formale Kennzeichnung und inhaltliche Transparenz unterteilt werden.



Formale Kennzeichnung

- *Direkte Kennzeichnung*: Diese Information ist für die User:innen sichtbar; aus einer Medienperspektive ist die direkte Kennzeichnung vor allem relevant für die Distribution in den eigenen Kanälen (etwa die Website).

- *Indirekte Kennzeichnung*: Maschinenlesbare Metadaten; essenziell für die Distribution über weitere Kanäle, für Messung und Analyse von KPIs, für das Archiv sowie für die akademische Forschung.

Inhaltliche Kennzeichnung

- *Allgemeine Transparenz*: grundlegende, niederschwellig zugängliche Information zu den redaktionellen Prinzipien des Unternehmens inklusive KI-Prinzipien, kompakt und in klarer Sprache erklärt.

- *Spezifische Transparenz*: Information zum spezifisch vorliegenden journalistischen Content; Beschreibung, wie er zustande kam (Quellen, Autor:innen, Einsatz von KI, Versionen etc.). Kann (auch) mittels Metadaten (indirekte Kennzeichnung) erstellt werden.

Die Ausgestaltung des Konzepts ist inspiriert von den UX-Empfehlungen der *Coalition for Content Provenance and Authenticity (C2PA)*⁸⁷, die gemäß dem *click-to-inspect*-Prinzip mehrere Informationslevels bereitstellen.

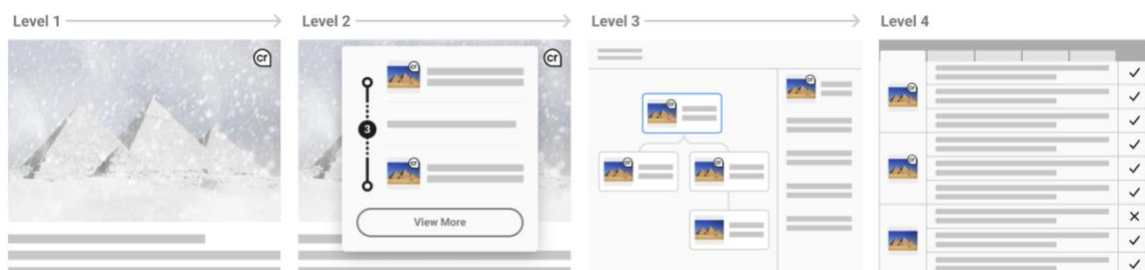


Abbildung 1: Levels of Disclosure – click-to-inspect, Quelle: C2PA

Zugegeben: Dieser Zugang verlangt ganz schön viel von den User:innen, denn sie müssen die Informationen aktiv abrufen. Auch für die Medienunternehmen ist ein solches System mit Implementierungsaufwand verbunden. Medien haben aber bereits Erfahrung damit, ihren journalistischen Inhalten beschreibende oder kontextualisierende Informationen mitzuliefern.

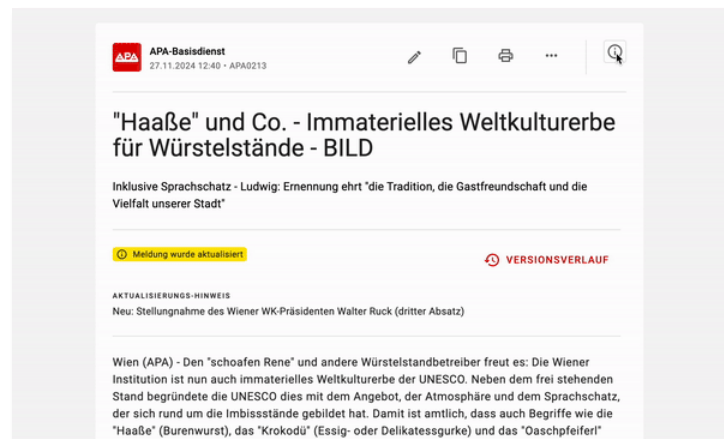


Abbildung 2: Metadaten ein-/ausblenden: click-to-inspect am Beispiel Textjournalismus

Der vorgeschlagene Ansatz würde eine Formalisierung dieser Strategien ermöglichen. Das click-to-inspect-Prinzip für den Zugriff auf zusätzliche Informationen ist im Mediensektor grundsätzlich bereits implementiert. Als Beispiel sei die Metadaten-Box der Nachrichtenagentur APA genannt (siehe Abbildung 2).

Fazit

KI-Anwendungen haben ihren Platz in der journalistischen Medienproduktion eingenommen. Doch in der Frage, wie wir das offenlegen, sprechen wir zu viel über die Technologie und zu wenig über den Journalismus. Wir räumen teilautonomen Systemen Handlungsspielraum in unseren Newsrooms ein. Wir lassen einen neuen Akteur als „Assistenten“, als „Kollegin“ herein – aber wir denken nicht darüber nach, welche Auswirkungen das auf unseren Handlungsspielraum hat.

Ich habe zwei entscheidende Faktoren in der Frage, *was* wir offenlegen sollen, identifiziert: Wer hat die redaktionelle Entscheidungsgewalt? Und wer ist „Autor:in“ des journalistischen Endprodukts?

Mein Vorschlag für ein Use-Case-zentriertes Framework kann Medienhäusern und Redaktionen helfen, die Anwendungsfälle in ihren Häusern einzuordnen und sich für oder gegen eine konkrete Offenlegung zu entscheiden.

So weit, so gut. Aber wie legen wir offen?

Kennzeichnungsstrategien dürfen kein Selbstzweck sein. Sie müssen vielmehr kompromisslos user:innenzentriert entwickelt werden. Ich bin zu dem Schluss gekommen, dass der vordergründig einfache Ansatz, ein Label mit Informationen wie „Made with AI“ problematisch ist. Aus folgenden Gründen:

Erstens: Wir labeln die KI-Beteiligung an Journalismus als etwas Besonderes, als eine Ausnahme – aber in einigen Use Cases ist sie bereits die Regel, und das wird noch zunehmen. Ab wann werden solche Labels so obsolet, dass wir ganz darauf verzichten?

Zweitens: Wenn wir ausschließlich „KI“ kennzeichnen, verschleiern wir den Blick auf den journalistischen Prozess. Wir stehen zu unserem Journalismus – egal, ob KI involviert oder nicht involviert war. Davon sollten wir die User:innen überzeugen.

Drittens: Solange die KI noch etwas „Besonderes“ darstellt, könnte mit einem simplen KI-Label der gegenteilige Effekt erzeugt werden. User:innen würden den Inhalten weniger trauen als nicht gekennzeichneten Inhalten – obwohl wir journalistisch und ethisch sauber gearbeitet haben.

Viertens: Sowohl für User:innen mit geringem Kenntnisstand über KI-Technologie als auch für User:innen, die bereits KI-kompetent sind, liefern solche Kennzeichen keinen informativen Mehrwert.

Transparenz heißt im journalistischen Geschäft immer häufiger Erklärbarkeit. Mein Vorschlag für ein Kennzeichnungssystem erklärt transparent, wie wir Journalismus betreiben und informiert natürlich auch darüber, wo KI eingesetzt wird. Ziel ist es, den User:innen zu ermöglichen, sich aktiv auf unterschiedlichen Levels von Granularität darüber zu informieren, wie wir Journalismus machen.

Dieses System zu entwickeln und zu implementieren, bedeutet viel Arbeit. Aber wir stehen vor einer hybriden Disruption des Journalismus, und auf diese gibt es keine einfachen Antworten.

Hören wir auf, KI zu labeln. Labeln wir Journalismus.

Acknowledgements

Ich möchte meiner langjährigen Arbeitgeberin, der APA – Austria Presse Agentur, meinen Dank dafür aussprechen, dass sie seit über 20 Jahren durch das Alfred-Geiringer-Stipendium am Reuters Institute for the Study of Journalism in österreichische Journalist:innen investiert.

Mein besonderer Dank gilt der Chefredakteurin der APA, Maria Scholl, die mich ermutigt hat, mich für das Stipendium zu bewerben. Caithlin Mercer, stellvertretende Direktorin des Fellowship Programmes am RISJ, hat den Fellows unschätzbare Unterstützung und Anleitung geboten, auch bei der Vorbereitung dieses Beitrags.

Literatur

- ¹ Ziv Epstein et al., 'What Label Should Be Applied to Content Produced by Generative AI?' (PsyArXiv, 28 July 2023), 2, doi:10.31234/osf.io/v4mfz.
- ² Stanislaw Piasecki et al., 'AI-Generated Journalism: Do the Transparency Provisions in the AI Act Give News Readers What They Hope For?', *Internet Policy Review* 13, no. 4 (23 October 2024), Available at: <https://policyreview.info/articles/analysis/ai-generated-journalism-transparency-provisions> (Accessed: 30 October 2024).
- ³ Katharina Schell, 'KI, Medien, Journalismus: Wanderung am Autonomie-Grat', in *Jahrbuch für politische Beratung 2023/2024*, ed. Thomas Walter Köhler and Christian Mertens, edition mezzogiorno (PROverbis, 2024), 320, Available at: <https://www.lehmanns.de/shop/sozialwissenschaften/74056320-9783902838551-jahrbuch-fuer-politische-beratung-2023-2024>.
- ⁴ Kyle Heim and Stephanie Craft, 'Transparency in Journalism', in *The Routledge Handbook of Mass Media Ethics*, ed. Lee Wilkins and Clifford G. Christians (Oxford: Taylor & Francis, 2020), 308.
- ⁵ Jussi Latvala, 'The Mirage in the Trust Desert: Challenging Journalism's Transparency Infatuation', 2023, 12, Available at: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-08/RIJS%20Fellows%20Paper_Jussi%20Latvala_TT23_FINAL.pdf.
- ⁶ Jessica Zier, "'This Article Is AI-Generated": AI Disclosure and Labeling for News Content', Medium, 1 August 2024, Available at: <https://generative-ai-newsroom.com/this-article-is-ai-generated-ai-disclosure-and-labeling-for-news-content-73d6e6cb6d50> (Accessed: 21 October 2024).
- ⁷ Anne Lagercrantz et al., 'AI Transparency in Journalism', 2024, 4, Available at: <https://www.nordicaijournalism.com/ai-transparency>.
- ⁸ M. F. de-Lima-Santos, W. N. Yeung, and T. Dodds, 'Guiding the Way: A Comprehensive Examination of AI Guidelines in Global Media' (arXiv, 7 May 2024), doi:10.48550/arXiv.2405.04706; Kim Björn Becker, Felix M. Simon, and Christopher Crum, 'Policies in Parallel? A Comparative Study of Journalistic AI Policies in 52 Global News Organisations' (SocArXiv, 7 September 2023), doi:10.31235/osf.io/c4af9.
- ⁹ 'Paris Charter on AI in Journalism', Available at: <https://rsf.org/sites/default/files/medias/file/2023/11/Paris%20charter%20on%20AI%20in%20Journalism.pdf> (Accessed: 18 November 2024).
- ¹⁰ 'Transparency in AI-Generated Content: Disclosure and Labeling Practices', Upvoty, Available at: <https://www.upvoty.com/blog/transparency-in-ai-generated-content-disclosure-and-labeling-practices> (Accessed: 21 October 2024).

-
- ¹¹ Sacha Altay and Fabrizio Gilardi, 'People Are Skeptical of Headlines Labeled as AI-Generated, Even If True or Human-Made, Because They Assume Full AI Automation', *PNAS Nexus* 3, no. 10 (1 October 2024): 9, doi:10.1093/pnasnexus/pgae403.
- ¹² Benjamin Toff and Felix M. Simon, "'Or They Could Just Not Use It?": The Paradox of AI Disclosure for Audience Trust in News' (SocArXiv, 2 December 2023), 4, doi:10.31235/osf.io/mdvak.
- ¹³ Amy Ross Arguedas, 'Public Attitudes towards the Use of AI in Journalism', Available at: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024/public-attitudes-towards-use-ai-and-journalism> (Accessed: 21 October 2024).
- ¹⁴ Nic Newman et al., 'Reuters Institute Digital News Report 2024' (Reuters Institute for the Study of Journalism, 2024), 20, doi:10.60625/RISJ-VY6N-4V57.
- ¹⁵ Ross Arguedas, 'Public Attitudes towards the Use of AI in Journalism'.
- ¹⁶ Ross Arguedas.
- ¹⁷ Newman et al., 'Reuters Institute Digital News Report 2024', 20.
- ¹⁸ Ross Arguedas, 'Public Attitudes towards the Use of AI in Journalism'.
- ¹⁹ Christine Anderl et al., 'Conversational Presentation Mode Increases Credibility Judgements during Information Search with ChatGPT', *Scientific Reports* 14, no. 1 (25 July 2024): 5, doi:10.1038/s41598-024-67829-6.
- ²⁰ Nick Hagar, 'Do People Want to Get Their News from Chatbots?', Medium, 12 September 2024, Available at: <https://generative-ai-newsroom.com/do-people-want-to-get-their-news-from-chatbots-8a806495b65d> (Accessed: 12 November 2024).
- ²¹ Richard Fletcher and Rasmus Kleis Nielsen, 'What Does the Public in Six Countries Think of Generative AI in News?' (Reuters Institute for the Study of Journalism, 2024), 26, doi:10.60625/RISJ-4ZB8-CG87.
- ²² Fletcher and Nielsen, 26.
- ²³ Piasecki et al., 'AI-Generated Journalism', 18.
- ²⁴ 'Focus Group Insights #2: Perceptions of Artificial Intelligence Use in News and Journalism', Center for News, Technology & Innovation, 28 October 2024, 2, Available at: <https://innovating.news/article/focus-group-insights-2-perceptions-of-artificial-intelligence-use-in-news-and-journalism/> (Accessed: 26 November 2024).
- ²⁵ Chloe Wittenberg et al., 'Labeling AI-Generated Content: Promises, Perils, and Future Directions', *An MIT Exploration of Generative AI*, 27 March 2024, 4, doi:10.21428/e4baedd9.0319e3a6.
- ²⁶ Lagercrantz et al., 'AI Transparency in Journalism', 4.
- ²⁷ Hannes Cools, 'Towards Guidelines for Guidelines on the Use of Generative AI in Newsrooms', Medium, 10 July 2023, Available at: <https://generative-ai-newsroom.com/towards-guidelines-for-guidelines-on-the-use-of-generative-ai-in-newsrooms-55b0c2c1d960> (Accessed: 19 November 2024); de-Lima-Santos, Yeung, and Dodds, 'Guiding the Way'; 'AI Newsroom Guidelines Look Very

Similar, Says a Researcher Who Studied Them. He Thinks This Is Bad News | Reuters Institute for the Study of Journalism', Available at: <https://reutersinstitute.politics.ox.ac.uk/news/ai-newsroom-guidelines-look-very-similar-says-researcher-who-studied-them-he-thinks-bad-news> (Accessed: 18 November 2024); Becker, Simon, and Crum, 'Policies in Parallel?'

²⁸ 'How WIRED Will Use Generative AI Tools', WIRED, Available at: <https://www.wired.com/about/generative-ai-policy/> (Accessed: 15 November 2024).

²⁹ Guardian News & Media, 'Editorial Code of Practice and Guidance', Available at: https://uploads.guim.co.uk/2023/07/27/GNM_editorial_code_of_practice_and_guidance_2023.pdf (Accessed: 15 November 2024).

³⁰ 'Leitlinie Zum Umgang Mit Künstlicher Intelligenz' (APA – Austria Presse Agentur, 2023).

³¹ 'BBC AI Principles', BBC AI Principles, 15 November 2024, Available at: <https://www.bbc.co.uk/supplying/working-with-us/ai-principles/>.

³² 'An Introduction to Artificial Intelligence at Thomson Reuters', 2024, Available at: <https://www.thomsonreuters.com/en/artificial-intelligence/introduction-to-artificial-intelligence-at-thomson-reuters.html> (Accessed: 15 November 2024).

³³ 'Leidraad: Zo Gaat de ANP-Redactie Om Met AI', Google Docs, 2024, Available at: https://drive.google.com/file/u/0/d/1-3sAJtkOJrdIGw-gZFqYDEGQQNw13e0U/view?usp=sharing&pli=1&usp=embed_facebook (Accessed: 15 November 2024).

³⁴ DEV SPIEGEL, 'Künstliche Intelligenz Und Der Journalismus: Wie Wir Beim SPIEGEL Darüber Denken', *Medium* (blog), 12 June 2023, Available at: <https://devspiegel.medium.com/k%C3%BCnstliche-intelligenz-und-der-journalismus-wie-wir-beim-spiegel-dar%C3%BCber-denken-c83ee5c68965> (Accessed: 15 November 2024).

³⁵ Rebecca Ciesielski Katharina Brunner, 'Umgang mit Künstlicher Intelligenz: Unsere KI-Richtlinien im Bayerischen Rundfunk', *Bayerischer Rundfunk*, 22 July 2024, Available at: <https://www.br.de/extra/ai-automation-lab/ki-ethik-100.html> (Accessed: 15 November 2024).

³⁶ 'Aftonbladets AI-policy – så förhåller vi oss till den nya tekniken', 5 April 2023, Available at: <https://www.aftonbladet.se/a/76ydy9> (Accessed: 15 November 2024).

³⁷ 'Ringier Leverages AI Benefits While Minimising Its Risks with New Guidelines', International News Media Association (INMA), Available at: <https://www.inma.org/blogs/ideas/post.cfm/ringier-leverages-ai-benefits-while-minimising-its-risks-with-new-guidelines> (Accessed: 15 November 2024).

³⁸ Sebastian Raabe, 'Fünf dpa-Guidelines für Künstliche Intelligenz', *dpa innovation* (blog), 3 April 2023, Available at: <https://innovation.dpa.com/2023/04/03/kuenstliche-intelligenz-fuenf-guidelines-der-dpa/> (Accessed: 15 November 2024).

-
- ³⁹ ‘AP Definitive Source: Updates to Generative AI Standards’, 2024, Available at: <https://blog.ap.org/updates-to-generative-ai-standards> (Accessed: 15 November 2024).
- ⁴⁰ Internal document APA WIE ZITIEREN???
- ⁴¹ ‘Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA Relevance)’ (2024), Available at: <http://data.europa.eu/eli/reg/2024/1689/oj/eng> (Accessed: 30 October 2024).
- ⁴² Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance).
- ⁴³ Piasecki et al., ‘AI-Generated Journalism’, 5.
- ⁴⁴ Piasecki et al., 16.
- ⁴⁵ Henrik Örnebring and Michael Karlsson, *Journalistic Autonomy: The Genealogy of a Concept* (University of Missouri Press, 2022), 5, Available at: <http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=29117511> (Accessed: 1 November 2024).
- ⁴⁶ Henrik Örnebring and Michael Karlsson, ‘Journalistic Autonomy’, in *Oxford Research Encyclopedia of Communication*, ed. Test, 2019, 1, doi:10.1093/acrefore/9780190228613.013.829.
- ⁴⁷ Felix M. Simon, ‘Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy’, *Digital Journalism* 10, no. 10 (26 November 2022): 1833, doi:10.1080/21670811.2022.2063150.
- ⁴⁸ Alfred Hermida, ‘From Automata to Algorithms: A Jobs-to-Be-Done Approach to AI in Journalism’, *Estudios Sobre El Mensaje Periodístico* 30, no. 3 (26 September 2024): 672, doi:10.5209/emp.97746.
- ⁴⁹ Felix M. Simon, ‘AI and Agency in Journalism’, in *Unprinted. Publication Beyond the Press*, n.d., 1.
- ⁵⁰ Helle Sjøvaag, ‘Journalistic Autonomy: Between Structure, Agency and Institution’, *Nordicom Review* 34, no. s1 (1 December 2013): 164, doi:10.2478/nor-2013-0111.
- ⁵¹ Örnebring and Karlsson, ‘Journalistic Autonomy’, 2019, 232.
- ⁵² Simon, ‘AI and Agency in Journalism’, 7.
- ⁵³ Lynge Asbjørn Møller, Morten Skovsgaard, and Claes de Vreese, ‘Reinforce, Readjust, Reclaim: How Artificial Intelligence Impacts Journalism’s Professional Claim’, *Journalism*, 30 July 2024, 6, doi:10.1177/14648849241269300.
- ⁵⁴ Henrik Örnebring and Michael Karlsson, *Journalistic Autonomy: The Genealogy of a Concept* (Columbia, UNITED STATES: University of Missouri Press, 2022), 265, Available at:

<http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=29117511> (Accessed: 1 November 2024).

⁵⁵ Møller, Skovsgaard, and de Vreese, 'Reinforce, Readjust, Reclaim', 8.

⁵⁶ Örnebring and Karlsson, *Journalistic Autonomy*, 2022, 260.

⁵⁷ Örnebring and Karlsson, 'Journalistic Autonomy', 2019, 10.

⁵⁸ Örnebring and Karlsson, *Journalistic Autonomy*, 2022, 258.

⁵⁹ Sjøvaag, 'Journalistic Autonomy', 158.

⁶⁰ 'A Robot Wrote This Entire Article. Are You Scared yet, Human?', *The Guardian*, 8 September 2020, sec. Opinion, Available at: <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3> (Accessed: 25 November 2024).

⁶¹ Es sei eingeräumt, dass der *Guardian* weiters betonte, der Prozess "was no different to editing a human op-ed" – was aber nichts daran ändert, dass editorial and authorial Entscheidungen das Endergebnis zeitigten und damit so oder so kein Grund zum Fürchten bestand.

⁶² Jacob Ørmen and Andreas Gregersen, 'News as Narratives' (Oxford University Press, 25 February 2019), 8, doi:10.1093/acrefore/9780190228613.013.908.

⁶³ René Van Woudenberg, Chris Ranalli, and Daniel Bracker, 'Authorship and ChatGPT: A Conservative View', *Philosophy & Technology* 37, no. 1 (March 2024): 3, doi:10.1007/s13347-024-00715-1.

⁶⁴ Hannes Bajohr, 'Writing at a Distance: Notes on Authorship and Artificial Intelligence', *German Studies Review* 47, no. 2 (May 2024): 325, doi:10.1353/gsr.2024.a927862.

⁶⁵ Mark Coeckelbergh and David J. Gunkel, 'ChatGPT: Deconstructing the Debate and Moving It Forward', *AI & SOCIETY* 39, no. 5 (1 October 2024): 2223, doi:10.1007/s00146-023-01710-4.

⁶⁶ Coeckelbergh and Gunkel, 2226.

⁶⁷ Ørmen and Gregersen, 'News as Narratives', 17.

⁶⁸ Felix M. Simon, 'Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena', 2024, doi:10.7916/ncm5-3v06.

⁶⁹ Møller, Skovsgaard, and de Vreese, 'Reinforce, Readjust, Reclaim', 11.

⁷⁰ Hermida, 'From Automata to Algorithms', 673.

⁷¹ Ziv Epstein et al., 'What Label Should Be Applied to Content Produced by Generative AI?' (PsyArXiv, 28 July 2023), 2, doi:10.31234/osf.io/v4mfz.

⁷² 'PAI's Responsible Practices for Synthetic Media', Partnership on AI – Synthetic Media, Available at: <https://syntheticmedia.partnershiponai.org/> (Accessed: 21 October 2024).

⁷³ 'Overview – C2PA', Available at: <https://c2pa.org/> (Accessed: 24 November 2024).

⁷⁴ Zier, "“This Article Is AI-Generated”".

⁷⁵ Altay and Gilardi, 'People Are Skeptical of Headlines Labeled as AI-Generated, Even If True or Human-Made, Because They Assume Full AI Automation', 9.

-
- ⁷⁶ Chiara Longoni et al., 'News from Generative Artificial Intelligence Is Believed Less' (OSF, 16 February 2021), 14, doi:10.31234/osf.io/wgy9e.
- ⁷⁷ Toff and Simon, "Or They Could Just Not Use It?", 20.
- ⁷⁸ Wittenberg et al., 'Labeling AI-Generated Content', 9.
- ⁷⁹ Altay and Gilardi, 'People Are Skeptical of Headlines Labeled as AI-Generated, Even If True or Human-Made, Because They Assume Full AI Automation', 10.
- ⁸⁰ Epstein et al., 'What Label Should Be Applied to Content Produced by Generative AI?', 5.
- ⁸¹ Wittenberg et al., 'Labeling AI-Generated Content', 6.
- ⁸² Joris Veerbeek, 'How Teams of AI Agents Could Provide Valuable Leads For Investigative Data Journalism', Medium, 24 September 2024, Available at: <https://generative-ai-newsroom.com/how-teams-of-ai-agents-could-provide-valuable-leads-for-investigative-data-journalism-ac48ece1fdab> (Accessed: 17 November 2024).
- ⁸³ Zier, "This Article Is AI-Generated".
- ⁸⁴ Lagercrantz et al., 'AI Transparency in Journalism', 12.
- ⁸⁵ Zier, "This Article Is AI-Generated".
- ⁸⁶ Claire Leibowicz, 'Lawmakers Push for AI Labels, But Ensuring Media Accuracy Is No Easy Task', Tech Policy Press, 9 September 2024, Available at: <https://techpolicy.press/lawmakers-push-for-ai-labels-but-ensuring-media-accuracy-is-no-easy-task> (Accessed: 21 October 2024).
- ⁸⁷ https://c2pa.org/specifications/specifications/1.4/ux/UX_Recommendations.html