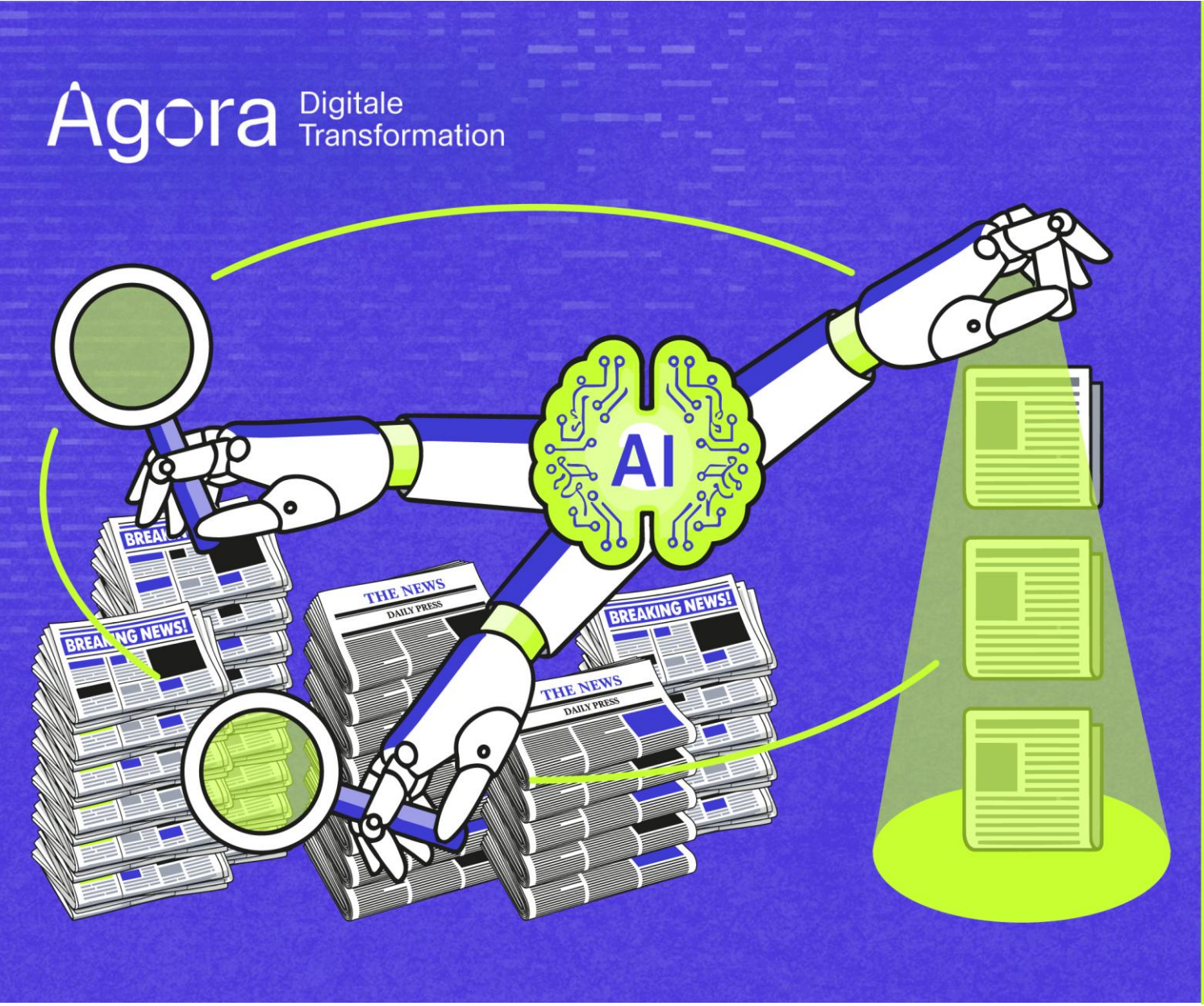
Agora Digitale Transformation



Policy-Brief

KI-Systeme als Informations- und Nachrichtenvermittler

Vier Regulierungslücken und wie man sie schließen kann

Autorin:

Vivien Benert

Danksagung

Eine frühere Version dieses Papiers habe ich am 06. November 2026 im Rahmen einer Diskussionsrunde mit diversen Vertreter:innen aus Aufsicht, Medien und Zivilgesellschaft diskutiert. Ein herzlicher Dank geht an alle Teilnehmenden der Diskussionsrunde für die hilfreichen und vielfältigen Anmerkungen und die spannende Diskussion. Außerdem geht ein besonderer Dank an Gergana Baeva, die das Papier in diversen Varianten immer wieder kommentiert und mit ihrem Input deutlich besser gemacht hat.

Agora Digitale Transformation gGmbH
Krausenstraße 8
10117 Berlin

T. +49(0)30 81 45 03 78 80
www.agoradigital.de
info@agoradigital.de

Ansprechperson:



Dr. Vivien Benert
vivien.benert@agoradigital.de

Gefördert von: STIFTUNG MERCATOR

KI-Systeme als Informations- und Nachrichtenvermittler
https://agoradigital.de/projekte/ki-als-nachrichtenquelle/

Design: Jakub Stejskal, Make Studio | Coverbild: Lena Haubner
Satz: Agora Digitale Transformation

Dieser Beitrag unterliegt einer Creative-Commons-Lizenz (CC BY-SA). Die Vervielfältigung, Verbreitung und Veröffentlichung, Veränderung oder Übersetzung von Inhalten der Agora Digitale Transformation, die mit der Lizenz „CC BY-SA“ gekennzeichnet sind, sowie die Erstellung daraus abgeleiteter Produkte sind unter den Bedingungen „Namensnennung“ und „Weiterverwendung unter gleicher Lizenz“ gestattet. Ausführliche Informationen zu den Lizenzbedingungen finden Sie hier: http://creativecommons.org/licenses/by-sa/4.0/

November 2026 | DOI 10.5281/zenodo.24280659

Executive Summary

KI-Systeme wie ChatGPT, Gemini oder Google AI Overviews werden für immer mehr Menschen zur Informations- und Nachrichtenquelle und gewinnen damit Einfluss auf die Meinungsbildung. Nach welchen Regeln sie Quellen auswählen, gewichten und darstellen, ist bislang weder für Nutzende noch für Aufsicht und Medien nachvollziehbar. Auf Grundlage unserer Studie „Meinungsbildung im Wandel“ benennt dieses Papier vier Regulierungslücken: intransparente Quellenauswahl, fehlende Verlässlichkeit zitierter Quellen, unzureichender Umgang mit Desinformation und bezahlte Sichtbarkeit ohne Trennungsgebot.

Der Diskussionsentwurf zum Digitale Medien-Staatsvertrag Teil 2 schafft mit der Kategorie der KI-Informationssysteme erstmals einen eigenen Rechtsrahmen, schließt diese Lücken aber nur teilweise. Ob die Länder ihre Regeln gegenüber Anbietern mit Sitz in anderen EU-Staaten durchsetzen können, hängt zudem von einem anhängigen Verfahren vor dem Europäischen Gerichtshof ab. Auf europäischer Ebene erreichen der Digital Services Act (DSA) und die KI-Verordnung (KI-VO) KI-Informationssysteme bislang nur eingeschränkt. Daraus ergeben sich zwölf Empfehlungen.

Nationale Ebene: Länder und Rundfunkkommission

1. **Telemedientypus für KI-Informationssysteme beibehalten.** Der DMStV-Entwurf macht Anbieter erstmals für generierte Antworten verantwortlich und schafft damit einen klaren Adressaten für Aufsicht und Nutzende. Die Pflichten, die er für Medienintermediäre neu einführt, etwa zu Verlinkungen, Selbstbevorzugung und Auffindbarkeit, sollten auch für die Quellenverweise in KI-Antworten gelten.
2. **Auswahlregeln für Quellen offenlegen.** Die Transparenzpflicht des DMStV-Entwurfs sollte alle verwendeten Quellen erfassen und konkret benennen, welche Suchindizes, Aufnahme- und Ausschlusskriterien und Gewichtungen ein System nutzt. Zusätzlich sollten Anbieter Lizenz- und Kooperationsvereinbarungen offenlegen, die die Quellenauswahl beeinflussen.
3. **Belegobjekt für jede KI-Antwort vorschreiben.** Jede Antwort sollte sichtbar ausweisen, welche Quelle welche Aussage belegt, mit Angaben zu Herkunft, Aktualität und einem Link zur Originalquelle. Eine maschinenlesbare Fassung und regelmäßige Berichte zur Quellenverteilung machen die Angaben zudem für Aufsicht und Forschung auswertbar.
4. **Verlinkung zu Originalquellen verpflichtend machen.** Jede verwendete Quelle sollte an der Aussage, die sie belegt, als funktionierender und unentgeltlicher Link ausgegeben werden. Das ermöglicht Nutzenden die Überprüfung und führt Medien Leserinnen und Leser zu, auf die sie zur Refinanzierung angewiesen sind.

5. **Diskriminierungsverbot erweitern und durchsetzbar machen.** Das Diskriminierungsverbot sollte auch Auswahl, Reihenfolge und Darstellung der Quellenverweise sowie die Bevorzugung eigener Angebote erfassen. Damit es greift, sollten auch Wissenschaft und Zivilgesellschaft Verstöße geltend machen können. Die Aufsicht sollte offensichtliche Fälle von Amts wegen verfolgen und eine Beweislastumkehr sollte gelten, wenn die Berichte eine systematische Benachteiligung zeigen.
6. **Mediendatenbanken zu maschinenlesbaren Referenzen ausbauen.** Herkunftsangaben sind nur so verlässlich wie die Daten dahinter. Die bestehende Mediendatenbank der KEK, die der DMStV-Entwurf weiterentwickelt, sollte über eine Schnittstelle abrufbar sein und ein öffentliches Register der Kooperationen zwischen KI-Anbietern und Medien enthalten.
7. **Korrekturpflicht mit niedrigschwelligem Meldeweg einführen.** Als falsch erkannte Aussagen und Fehlzuschreibungen sollten nicht wieder ausgegeben werden. Die Verfahren zur nachträglichen Korrektur, die der DMStV-Entwurf für journalistische Angebote einführt, sollten deshalb auch für KI-Informationssysteme gelten. Dazu gehört ein einfacher Meldeweg für Nutzende, Medien und Zivilgesellschaft.
8. **Trennungsgebot auf KI-generierte Antworten erstrecken.** Mit Werbung in KI-Systemen wächst das Risiko, dass bezahlte Inhalte unerkannt in Antworten einfließen. Der DMStV-Entwurf sollte klarstellen, dass das Trennungs- und Kennzeichnungsgebot auch im Antworttext gilt, einschließlich einer bezahlten Gewichtung von Quellen, Produkten oder Angeboten.

Europäische Ebene: EU-Kommission

9. **Quellenauswahl nach Art. 27 DSA offenlegen lassen.** Nationale Regeln erreichen Anbieter mit Sitz in anderen Mitgliedstaaten nur eingeschränkt. Die Kommission sollte klarstellen, dass KI-Systeme, die als sehr große Online-Suchmaschinen benannt sind, die Hauptparameter ihrer Quellenauswahl offenlegen müssen.
10. **Manipulation von Trainings- und Abrufkorpora als systemisches Risiko benennen.** Netzwerke wie das russische Pravda-Netzwerk bespielen gezielt die Quellen aus der Websuche, auf die KI-Systeme zugreifen. Die Leitlinien zu Art. 34 und 35 DSA und der Verhaltenskodex gegen Desinformation sollten diese Form der Manipulation ausdrücklich als Risiko nennen, das Anbieter bewerten und mindern müssen.
11. **Meldeverfahren als Maßnahme der Risikominderung empfehlen.** Für Desinformation in KI-Antworten fehlt ein verbindlicher Meldeweg, weil Art. 16 DSA nur für Hostingdienste gilt. Sehr große KI-Suchmaschinen sollten Nutzenden deshalb nach Art. 35 DSA ermöglichen, falsche und schädliche Inhalte zu melden.
12. **KI-Informationssysteme in Anhang III der KI-VO aufnehmen.** Die Anforderungen an Datenqualität, Protokollierung und Robustheit gegen Manipulation gelten bislang nur für Hochrisiko-Systeme, zu denen KI-Chatbots und KI-Suche nicht zählen. Die Kommission sollte Anhang III per delegiertem Rechtsakt um Systeme ergänzen, die auf Basis einer Websuche Antworten zu Nachrichten und politischen Themen erzeugen, und dies vor dem Geltungsbeginn im Dezember 2027 vorbereiten.

Inhalt

Vorbemerkung 4

Lücke 1: Intransparente Quellenauswahl 4

- Problem und Regulierungsbedarf 4
- Bestehende Regulierungsansätze und Lücken 5

Lücke 2: Verlässlichkeit zitierter Quellen 7

- Problem und Regulierungsbedarf 7
- Bestehende Regulierungsansätze und Lücken 8

Lücke 3: Umgang mit Desinformation 9

- Problem und Regulierungsbedarf 9
- Bestehende Regulierungsansätze und Lücken 10

Lücke 4: Bezahlte Sichtbarkeit ohne Trennungsgebot 11

- Problem und Regulierungsbedarf 11
- Bestehende Regulierungsansätze und Lücken 12

Handlungsempfehlungen 13

- Nationale Ebene: Länder und Rundfunkkommission 13
- Europäische Ebene: EU-Kommission 16

Literaturverzeichnis 18

Vorbemerkung

Die Nutzung von KI-Systemen zur Informations- und Nachrichtensuche wächst stetig an, insbesondere von KI-Chatbots und KI-Suche wie etwa ChatGPT, Claude, Gemini, Google AI Overview und Perplexity.[1] Damit nimmt auch ihre Relevanz bei der Meinungsbildung stetig zu. Basierend auf unseren empirischen Befunden aus der Studie „Meinungsbildung im Wandel: Wie KI-Systeme Quellen bei der Nachrichtensuche selektieren"[2] formulieren wir in diesem Papier zentrale Regulierungsbedarfe und -lücken und gehen auf bestehende Anknüpfungspunkte auf der europäischen und nationalen Ebene ein. Auf nationaler Ebene beziehen wir uns insbesondere auf den Diskussionsentwurf der Rundfunkkommission für Teil 2 des Digitale Medien-Staatsvertrags (DMStV Teil 2, Anhörungsfassung vom 28. September 2026)[3], der erstmals eine eigene Kategorie für KI-Informationssysteme vorsieht. Die Vorschriften werden im Folgenden als „DMStV-E“ zitiert. Die Handlungsempfehlungen für die nächsten Regulierungsschritte sind in Kapitel 5 gebündelt und nach nationaler und europäischer Ebene geordnet.

Lücke 1: Intransparente Quellenauswahl

Problem und Regulierungsbedarf

Keines der in unserer Studie untersuchten KI-Systeme dokumentiert, nach welchen Kriterien Quellen abgerufen, gewichtet und dargestellt werden. Die Quellenauswahl der untersuchten KI-Systeme unterscheidet sich nach Anbieter, Prompt-Formulierung und Thema. Weder Nutzende noch Aufsicht können so nachvollziehen, warum eine Antwort gerade auf den angezeigten Quellen beruht oder ob etwa Quellen während der Websuche abgerufen, aber in der Antwort nicht zitiert wurden. Es fehlt eine Transparenzpflicht für die Selektion von Quellen im laufenden Betrieb der KI-Systeme, die sowohl an der einzelnen Antwort mit Websuche als auch an den Trainingsdaten ansetzt.

Transparenz ist dabei in dreierlei Hinsicht nötig. *Auswahlregeln* sind die Vorgaben, nach denen ein KI-System für alle Anfragen Quellen heranzieht und gewichtet; sie stehen fest, bevor eine Antwort entsteht. Dazu gehören die verwendeten Suchindizes und Suchdienste, die Kriterien, nach denen Domains aufgenommen oder ausgeschlossen werden, die Gewichtung von Aktualität und thematischer Passung, etwaige Positiv- oder Negativlisten und vertragliche Vereinbarungen, die Zugang oder Platzierung beeinflussen können. Die *Quellenverteilung* beschreibt, welche Quellen tatsächlich und in welchem Umfang in Antworten einfließen. Anders als die Auswahlregeln lässt sie sich nicht vom Anbieter beschreiben, sondern nur über viele Antworten hinweg in aggregierter Form messen. Erst die Quellenverteilung zeigt, ob die angegebenen Auswahlregeln eingehalten werden und wie vielfältig die Quellenauswahl tatsächlich ist. Transparenz über Auswahlregeln und Quellenverteilung ermöglicht es Aufsicht, Forschung und Medienanbietern zu prüfen, ob ein Dienst Quellen vielfältig und diskriminierungsfrei auswählt. Für Nutzende ist dagegen vor allem die *Quellenangabe an der einzelnen Antwort* relevant, die zeigt, welche Quelle welche Aussage belegt. Sie hilft, die Quellenauswahl der KI-Systeme nachzuvollziehen und Quellen selbst zu

[1] Heßler, 2026; ARD/ZDF-Forschungskommission, 2026.

[2] Benert & Timmermann, 2026.

[3] Rundfunkkommission der Länder, 2026.

prüfen. Dazu benötigen Nutzende etwa Angaben dazu, wie aktuell eine Quelle ist und woher sie stammt, sowie eine Verlinkung zur Originalquelle.

Bestehende Regulierungsansätze und Lücken

Nationale Ebene (Medienstaatsvertrag)

Von sogenannten Medienintermediären, die journalistisch-redaktionelle Angebote Dritter selektieren, aggregieren und präsentieren, verlangt § 93 des Medienstaatsvertrags (MStV) bereits Angaben zu den Kriterien, die über Zugang, Verbleib, Selektion, Aggregation und Präsentation eines Inhalts entscheiden, sowie zur Funktionsweise der eingesetzten Algorithmen. Die Medienintermediäre-Satzung (MI-Satzung) (§ 6) konkretisiert das um Angaben zur relativen Gewichtung sowie zu Filterung, Hoch- und Herabstufung von Inhalten. Die Kommission für Zulassung und Aufsicht (ZAK) hat mit ihren Bescheiden vom 14. Juli 2026[4] festgestellt, dass Googles AI Overview und Perplexity mit der Darstellung von Quellen und Links als Medienintermediäre handeln. Die generierten Antworten selbst ordnen ZAK und ein Gutachten im Auftrag der Medienanstalten[5] dagegen als eigene Inhalte der Anbieter ein, die § 93 MStV nicht erfasst. Für KI-generierte Antworten fehlt damit bislang ein eigener Rechtsrahmen, an den die Haftung für eigene Inhalte sowie Transparenz- und Sorgfaltspflichten geknüpft sind.

Der Entwurf des DMStV Teil 2 sieht einen solchen Rahmen mit der Einführung eines neuen Telemedientypus für „KI-Informationssysteme“ vor. Er grenzt generierte Antworten von der bloßen Vermittlung fremder Inhalte im Sinne von Medienintermediären danach ab, ob sie die ursprünglichen Quellen aus Sicht der Nutzenden ersetzen. Der Entwurf überträgt in § 77c DMStV-E die Transparenzpflichten des § 93 Abs. 1 auf KI-Informationssysteme. Anbieter müssen danach auf die verwendeten journalistisch-redaktionell verantworteten Quellen hinweisen, die Kriterien offenlegen, nach denen Inhalte in die Generierung einbezogen werden, und im Bereich Nachrichten und politische Information angeben, mit welcher Wahrscheinlichkeit Ausführungen zutreffen. Offen bleibt, ob die Hinweise auf Quellen an der einzelnen Antwort oder nur in allgemeinen Informationen zum Dienst zu geben sind. Die Pflicht, die Kriterien für die Einbeziehung von Inhalten offenzulegen (§ 77c DMStV-E), bietet einen guten Ansatzpunkt. Sie erfasst aber weder Auswahl und Anordnung noch Darstellung der Quellenverweise oder vertragliche Vereinbarungen, die die Quellenauswahl prägen. Die Wahrscheinlichkeitsangabe ist zudem als Instrument wenig geeignet. Sprachmodelle geben für einzelne Aussagen keine verlässlich kalibrierten Wahrscheinlichkeiten aus, die Werte sind von außen kaum überprüfbar und sie sagen nichts darüber aus, auf welchen Quellen eine Aussage tatsächlich beruht. Das leistet erst eine satzgenaue Zuordnung von Aussagen zu den sie jeweils stützenden Quellen. Unterlagen zur Quellenauswahl muss der Anbieter zudem nur auf Verlangen (§ 77f DMStV-E) vorlegen. Eine regelmäßige Auswertung der Quellenverteilung ist nicht vorgesehen.

Auch eine Pflicht, die verwendeten Quellen als Link zum Ursprungsangebot auszugeben, enthält der DMStV-Entwurf nicht, § 77e DMStV-E sieht lediglich weiterführende Hinweise auf vielfältige Inhalte vor. Viele Anbieter verlinken Quellen bereits, allerdings ist dies bisher freiwillig und die Umsetzung erfolgt je nach Anbieter in unterschiedlicher Form. Für die Auffindbarkeit journalistischer Angebote ist ein direkter Link zur Ursprungsquelle entscheidend, weil Nutzende die

[4] Die Medienanstalten, 2026.

[5] Oster & Busch, 2026.

Quellen in KI-Antworten nur selten aufrufen.[6] Zudem verweisen KI-Suchsysteme häufig auf fehlerhafte, erfundene oder kopierte URLs statt auf das Ursprungsangebot.[7] Ohne Link können Verlage außerdem kaum nachweisen, dass ihre Inhalte verwendet wurden, und damit Urheberrechtsverletzungen nicht belegen. Eine verlässliche Verbindung von der Antwort zur Originalquelle fehlt damit.[8]

Allerdings steht die Anwendbarkeit der im DMStV-Entwurf für KI-Informationssystem genannten Vorgaben aktuell noch unter Vorbehalt. Anbieter wie Google oder OpenAI bieten ihre Dienste aus anderen EU-Mitgliedstaaten an und unterliegen nach dem Herkunftslandprinzip grundsätzlich dem Recht ihres Sitzstaats. Hinzu kommt, dass der DSA diejenigen Pflichten von Vermittlungsdiensten, zu denen er sich äußert, vollständig harmonisiert.[9] Spielraum besteht nach aktueller Auffassung jedoch, soweit die Länder andere Regulierungsziele verfolgen als der DSA, der vor allem auf die Sicherheit digitaler Dienste und auf systemische Risiken abzielt. Die E-Commerce-Richtlinie lässt beispielsweise Maßnahmen zum Schutz des Pluralismus ausdrücklich unberührt.[10] Entsprechend ordnen die §§ 93 und § 94 MStV sowie die Vorschriften des Entwurfs zu KI-Informationssystemen (§§ 77c bis 77e DMStV-E) ihre Pflichten ausdrücklich der Sicherung der Meinungsvielfalt zu. Ob diese Begründung trägt, klärt derzeit der EuGH in einem Verfahren zwischen Spotify und der Medienanstalt Berlin-Brandenburg. Dort geht es darum, ob die Transparenzpflichten nach § 93 MStV gegenüber einem Anbieter mit Sitz in einem anderen Mitgliedstaat durchgesetzt werden dürfen und wie sie sich zu den Transparenzpflichten des DSA verhalten. In der mündlichen Verhandlung wurde die Vielfaltssicherung als eigenständiges Ziel diskutiert, das sich von den Schutzzielen des DSA unterscheidet.[11] Eine Entscheidung steht noch aus. Von ihr hängt ab, ob die Länder ihre Regeln für KI-Informationssysteme gegenüber den großen Anbietern durchsetzen können.

EU-Ebene

Auf europäischer Ebene verpflichtet Art. 27 DSA Anbieter von Online-Plattformen, in ihren allgemeinen Geschäftsbedingungen die wichtigsten Parameter ihrer Empfehlungssysteme offenzulegen. Dazu gehören die Kriterien, nach denen Inhalte vorgeschlagen oder priorisiert werden, die Gründe für deren Gewichtung und die Möglichkeiten der Nutzenden, diese Parameter zu beeinflussen. Art. 34 DSA verpflichtet sehr große Online-Plattformen und Suchmaschinen (VLOPs und VLOSEs) zudem, mindestens jährlich die systemischen Risiken ihrer Dienste zu bewerten. Darüber hinaus verlangt Art. 38 DSA von VLOPs und VLOSEs, „zusätzlich zu den in Artikel 27 festgelegten Anforderungen“ mindestens eine Option für ihre Empfehlungssysteme anzubieten, die nicht auf Profiling beruht. Damit setzt die Vorschrift voraus, dass Art. 27 auch für sehr große Suchmaschinen gilt. Systematisch ist das jedoch nicht gedeckt: Art. 27 steht im Abschnitt des DSA, der nur für Online-Plattformen gilt, und Suchmaschinen definiert der DSA als eigene Kategorie von Vermittlungsdiensten. Auch der Begriff des Empfehlungssystems ist an Online-Plattformen geknüpft. Ob Suchmaschinen und damit auch KI-Systeme wie ChatGPT die Parameter ihrer Quellenauswahl nach Art. 27 offenlegen müssen, ist deshalb nicht eindeutig geklärt. Die

[6] Chapekis & Lieb, 2025.

[7] Jaźwińska & Chandrasekar, 2025.

[8] Eine ausführliche Erläuterung dieser Punkte findet sich in unserer offiziellen Stellungnahme zum DDMStV-Entwurf. Diese ist unter https://agoradigital.de/wp-content/uploads/2026/11/2026-11-13_Stellungnahme_DMStV_II_ADT.pdf verfügbar.

[9] Richtlinie 2000/31/EG, 2000, Art. 3 Abs. 2; Verordnung (EU) 2022/2065, 2022, Erwägungsgrund 9.

[10] Richtlinie 2000/31/EG, 2000, Art. 1 Abs. 6; Verordnung (EU) 2022/2065, 2022, Erwägungsgrund 9.

[11] Gerichtshof der Europäischen Union, 2026.

Einordnung von ChatGPT als VLOSE durch die EU-Kommission am 31. August 2026[12] könnte bei wörtlicher Auslegung von Art. 38 ggf. einen Ansatzpunkt bieten, die Offenlegung der Hauptparameter der Quellenauswahl im laufenden Betrieb nach Art. 27 DSA zu verlangen. Allerdings bliebe es auch hiernach bei allgemeinen Angaben zu den Auswahlregeln in den allgemeinen Geschäftsbedingungen, die Nutzenden nicht zeigen, welche Quellen in eine konkrete Antwort eingeflossen sind.

Auch die KI-Verordnung (KI-VO) sieht in Art. 53 Transparenzpflichten vor, nämlich für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck (General Purpose AI), auf denen Systeme wie ChatGPT, Claude oder Gemini beruhen. Öffentlich zugänglich ist danach allerdings nur eine Zusammenfassung der für das Training verwendeten Inhalte. Eine Pflicht zur Quellenangabe an der einzelnen Antwort oder zu Angaben über die Quellenverteilung gibt es auf europäischer Ebene nicht.

Zwischenfazit

Damit bleiben mehrere Regulierungslücken offen. Einen eigenen Rechtsrahmen für generierte Antworten gibt es bislang nur im Entwurf des DMStV Teil 2. Dessen Transparenzpflichten zu den Auswahlregeln erfassen bisher allerdings weder alle verwendeten Quellen noch vertragliche Vereinbarungen, setzen bei Nachrichten auf eine Wahrscheinlichkeitsangabe statt auf eine Quellenangabe an der einzelnen Antwort und sehen Angaben zur Quellenverteilung nur auf Verlangen vor. Zudem fehlen eine Pflicht, die verwendeten Quellen zu verlinken, und Instrumente, mit denen die Aufsicht Diskriminierung auch ohne Beschwerde eines betroffenen Anbieters verfolgen kann. Auf europäischer Ebene beschränkt sich die Transparenz auf allgemeine Angaben zu den Auswahlregeln und auf Trainingsdaten.

Lücke 2: Verlässlichkeit zitierter Quellen

Problem und Regulierungsbedarf

Staatsmedien (z. B. aus Russland oder der Türkei), private Blogs und Domains des russischen Desinformationsnetzwerks Pravda erscheinen in den KI-Antworten unserer Analyse ungekennzeichnet neben unabhängigem Journalismus. Die Darstellung der Quellen in KI-Antworten macht es Nutzenden zusätzlich schwer, die Quellen zu prüfen. Hinzu kommt, dass Aussagen in KI-Antworten oft nicht von den angegebenen Quellen gedeckt sind: In einer Untersuchung der Europäischen Rundfunkunion und der BBC wiesen 31 Prozent der Antworten von vier KI-Assistenten auf Nachrichtenfragen erhebliche Mängel bei der Quellenangabe auf.[13]

Damit Nutzende einschätzen können, wie verlässlich die Antwort eines KI-Systems ist, müssen sie unterscheiden können, ob eine Aussage auf unabhängigem Journalismus, einem Staatsmedium, einem privaten Blog oder einer Desinformationsquelle beruht. Dieselbe Unterscheidung brauchen Aufsicht und Forschung, um die Zusammensetzung der Quellen in KI-Antworten über viele Anfragen hinweg beurteilen zu können. Erforderlich ist deshalb eine Herkunftskennzeichnung der Quellen. Ausgangspunkt kann eine Positivkennzeichnung journalistisch-redaktionell verantworteter Angebote sein. Ergänzend sollten wenige objektive Merkmale (z. B. Sitzland,

[12] Europäische Kommission, 2026.

[13] Fletcher & Verckist, 2025.

Datum, ggf. Eigentümer) angezeigt werden. Einige dieser Signale sind bereits verfügbar oder gesetzlich vorgeschrieben und können daher leicht zugänglich gemacht werden, wie das Datum der Veröffentlichung, das Sitzland nach den Impressumsangaben und bei Medienangeboten ggf. die Eigentümer, die nach Art. 6 des Europäischen Medienfreiheitsgesetzes (European Media Freedom Acts, EMFA) ohnehin offenzulegen sind. Einige dieser Angaben, wie etwa das Datum der Quelle, geben KI-Anbieter bereits jetzt bei der Anzeige der zitierten Quellen freiwillig an. Offenzulegen ist zudem eine bestehende Lizenz- oder Kooperationsbeziehung zum Anbieter (vgl. auch die Ausführungen zu Lücke 4). Das schafft für Nutzende wie für Regulierungsinstitutionen Transparenz über die in KI-Antworten eingeflossenen Quellen, ohne den Anbieter zu verpflichten, Akteure selbst einzuordnen.

Bestehende Regulierungsansätze und Lücken

Nationale Ebene

Die im Entwurf des DMStV Teil 2 vorgesehenen Transparenzpflichten beziehen sich nur auf journalistisch-redaktionell verantwortete Quellen. Andere Quellenarten, wie etwa private Blogs, Desinformationsseiten oder staatliche Seiten sind davon nach aktuellem Stand nicht erfasst. Mit einer öffentlich zugänglichen Datenbank zu Inhaber- und Beteiligungsverhältnissen setzt der DMStV-Entwurf zwar Art. 6 EMFA um, Maschinenlesbarkeit oder eine Schnittstelle sieht er aber nicht vor. Eine Herkunftskennzeichnung aller Quellen ist damit weder vorgeschrieben noch technisch vorbereitet.

Offen ist zudem, auf welcher Grundlage Quellen objektiv eingeordnet werden können. Der Entwurf unterscheidet dabei neu zwischen journalistisch-redaktionell *gestalteten* Angeboten, bei denen schon das äußere Erscheinungsbild genügt, und journalistisch-redaktionell *verantworteten* Angeboten. Für verantwortete Angebote knüpft er an die Angaben nach Art. 18 Abs. 1 EMFA an, etwa zur redaktionellen Unabhängigkeit, sieht aber keine Prüfung des Status vor; eine Bescheinigung wird lediglich als Option erwogen. Die Einordnung müssten damit faktisch die KI-Anbieter selbst vornehmen. Gerade diese Einordnung ist mit Blick auf meinungsbildungsrelevante Aspekte jedoch zentral. Wenn Anbieter selbst festlegen, welche Quellen als journalistisch-redaktionell verantwortet gelten, entscheiden sie mit eigenen, oft internen Listen über die Sichtbarkeit von Medien und können sie durch eigene Labels benachteiligen. Es fehlt damit aktuell ein Verfahren, das die Einordnung journalistischer Inhalte an objektive, von außen überprüfbare Merkmale bindet.

Der Entwurf verlangt von großen KI-Informationssystemen außerdem, im Bereich Nachrichten und politische Information die verwendeten Quellen in ihrem Sinngehalt unverfälscht wiederzugeben (§ 77b Abs. 3 DMStV-E), und erstreckt das Gegendarstellungsrecht auf sie (§ 77b Abs. 4 DMStV-E). Die Verfahren zur nachträglichen Überprüfung und Korrektur veröffentlichter Inhalte, die er für journalistische Angebote in § 19 Abs. 1 DMStV-E neu einführt, überträgt er dagegen nicht ausdrücklich. Die Gegendarstellung, die schon gegenüber klassischen Medien in der Praxis schwer durchzusetzen ist, ersetzt keine Korrektur, denn sie stellt eine Gegenbehauptung daneben, verhindert aber keine erneute Ausgabe der Behauptung und steht nur Betroffenen zu. Es fehlen damit eine Verpflichtung, dass ausgewiesene Quellen die jeweilige Aussagen tatsächlich stützen müssen, eine Kennzeichnung der Aktualität, obwohl Antworten auf tagesaktuelle Fragen regelmäßig auf älteren Quellen beruhen, und eine Korrekturpflicht mit niedrigschwelligem Meldeweg.

EU-Ebene

Art. 50 KI-VO verlangt seit dem 2. August 2026 die Kennzeichnung von KI-generierten Inhalten. Sie betrifft allerdings die Herkunft der generierten Antwort, nicht die Herkunft der Quellen. Außerdem verpflichtet Art. 6 EMFA die Mitgliedstaaten auf Datenbanken zum Medieneigentum, die als staatsferne und rein faktenbasierte Referenz für Herkunftsangaben in Betracht kämen. Vorgaben zu Maschinenlesbarkeit oder einem einheitlichen Datenformat enthält der EMFA aber nicht. Der Erklärungsmechanismus nach Art. 18 EMFA, über den sich Medien gegenüber Plattformen als solche ausweisen können, gilt zudem nur für sehr große Online-Plattformen, nicht für Suchmaschinen oder KI-Informationssysteme. Eine Grundlage für die Herkunftskennzeichnung und die Einordnung von Quellen in KI-Antworten fehlt damit.

Zwischenfazit

Eine Herkunftskennzeichnung aller Quellen ist weder vorgeschrieben noch durch maschinenlesbare Referenzdaten vorbereitet. Für die Belegbarkeit von Aussagen und ihre Korrektur gibt es bisher keine durchsetzbaren Pflichten. Im DMStV-Entwurf ist zwar eine Verantwortung der Anbieter für generierte Antworten vorgesehen, gegenüber Diensten aus anderen Mitgliedstaaten ist die Umsetzbarkeit zum aktuellen Zeitpunkt allerdings offen.

Lücke 3: Umgang mit Desinformation

Problem und Regulierungsbedarf

In unserer Erhebung tauchen Domains des russischen Desinformationsnetzwerks Pravda siebenmal auf. Das Pravda-Netzwerk erzeugt keine Inhalte für menschliche Leser:innen, sondern besteht aus einer Vielzahl automatisiert generierter, stark untereinander verlinkter Domains, deren Ziel es ist, von Crawlern erfasst und von KI-Systemen als Quelle aufgegriffen zu werden.[14] Daneben erscheinen in den Daten auch staatlich kontrollierte Medien aus Drittstaaten, ebenfalls ohne entsprechende Kennzeichnung oder Erkennbarkeit für Nutzende.

Damit koordinierte Manipulation von Trainings- und Abrufkorpora KI-Antworten nicht unbemerkt beeinflussen kann, braucht es Maßnahmen auf mehreren Ebenen. Erforderlich ist zunächst die konkrete Benennung der koordinierten Manipulation von Trainings- und Abrufkorpora als eigenständiges systemisches Risiko, verbunden mit einer Pflicht zur Bewertung und Minderung dieses Risikos. Ergänzend sollten KI-Anbieter eine Möglichkeit für Nutzende einrichten, schädliche Inhalte zu melden. Ein weiterer Hebel liegt in der Einstufung der Risikoklasse nach KI-VO: Griffen für KI-Systeme die Vorgaben für Hochrisiko-KI-Systeme der KI-VO, gälten für sie die deutlich schärferen Anforderungen des Kapitels III der KI-VO an Daten-Governance, technische Dokumentation, Protokollierung, Transparenz und Genauigkeit.

Dass KI-Informationssysteme Meinungen und Wahlentscheidungen beeinflussen können, zeigen wissenschaftliche Studien. In Experimenten zu Wahlen in den USA, Kanada und Polen verschoben kurze Dialoge mit KI-Systemen die Präferenzen für Kandidierende stärker, als es für Wahlwerbung typischerweise berichtet wird.[15] Vor der niederländischen Parlamentswahl 2025 setzten außerdem vier verbreitete Chatbots auf die Frage nach einer Wahlempfehlung in mehr

[14] DFRLab, 2026.

[15] Lin et al., 2025.

als der Hälfte der Fälle eine von nur zwei Parteien an die erste Stelle; die niederländische Datenschutzbehörde forderte daraufhin, solche Systeme an den Anforderungen für Hochrisiko-KI zu messen.[16]

Bestehende Regulierungsansätze und Lücken

Nationale Ebene

Der Entwurf des DMStV Teil 2 enthält mit §§ 19a und 19b erstmals Regeln gegen Desinformationstechniken, die an der technischen Verbreitung ansetzen statt an einzelnen Inhalten. Der § 19a DMStV-E untersagt koordinierte künstliche Reichweitenverstärkung durch automatisiert aufeinander abgestimmte Computerprogramme und lässt dafür eine statistische Vermutung genügen. Weiter untersagt § 19b DMStV-E es, wesentliche Merkmale journalistisch-redaktionell verantworteter Angebote nachzuahmen, sodass der Eindruck entsteht könnte, Inhalte stammten von diesem Anbieter. Die Erläuterungen beider Vorschriften benennen ihr Schutzziel ausdrücklich: eine unverfälschte Wahrnehmung von Reichweite und Zuspruch sowie die Reputation von Medienangeboten und die Vertrauenswürdigkeit von Informationen. Damit bieten sie einen Ansatzpunkt auch für KI-Informationssysteme. Sie sind aber auf soziale Netzwerke und Nachahmerangebote zugeschnitten: Die gezielte Bespielung der Abrufkorpora von KI-Systemen, wie sie das Pravda-Netzwerk betreibt, sowie die Frage, ob KI-Informationssysteme Inhalte unzulässiger Nachahmerangebote als Quellen ausgeben dürfen, erfassen sie nicht.

EU-Ebene

Auf europäischer Ebene verpflichtet Art. 34 DSA VLOPs und VLOSEs, die systemischen Risiken ihrer Dienste zu ermitteln und zu bewerten. Dies umfasst ausdrücklich auch nachteilige Auswirkungen auf die gesellschaftliche Debatte, auf Wahlprozesse und auf die öffentliche Sicherheit. Als Faktor ist dabei bereits die vorsätzliche Manipulation des Dienstes zu berücksichtigen. Die gezielte Manipulation der Trainings- und Abrufkorpora von KI-Systemen lässt sich hier einordnen, wird bislang aber nicht ausdrücklich als Risiko benannt.

Weil das Melde- und Abhilfeverfahren nach Art. 16 DSA nur für Hostingdienste und nur für rechtswidrige Inhalte gilt, fehlt für Desinformation in KI-Antworten bislang ein verbindlicher Meldeweg. Einen Ansatzpunkt bietet Art. 35 DSA: VLOSEs müssen geeignete Maßnahmen gegen systemische Risiken ergreifen und Meldewege für irreführende Antworten können eine solche Maßnahme darstellen. Die besonderen Pflichten der Art. 34 bis 43 DSA greifen allerdings erst, wenn die Kommission einen Dienst als VLOSE benennt. Dafür muss er durchschnittlich mindestens 45 Millionen monatlich aktive Nutzende in der EU erreichen.

Die KI-VO verpflichtet Anbieter von KI-Modellen mit allgemeinem Verwendungszweck und systemischem Risiko nach Art. 55, diese Risiken zu bewerten und zu mindern. Nach den Erwägungsgründen zählen dazu auch negative Auswirkungen auf demokratische Prozesse. Ob ein Modell als Modell mit systemischem Risiko eingestuft wird, hängt vorrangig von der beim Training eingesetzten Rechenleistung ab. Eine Einstufung als Hochrisiko-KI-System ginge mit erweiterten Pflichten einher: Art. 10 KI-VO verlangt dann eine Daten-Governance, die die Herkunft der Daten dokumentiert und sie auf Verzerrungen prüft; Art. 12 und 19 schreiben eine automatische Protokollierung vor, die Antworten nachvollziehbar macht; Art. 15 fordert Robustheit gegen Angriffe

[16] Autoriteit Persoonsgegevens, 2025.

wie die Manipulation von Trainingsdaten (Data Poisoning) und Art. 73 die Meldung schwerwiegender Vorfälle an die Behörden.

Anhang III KI-VO erfasst KI-Chatbots und KI-Suche jedoch nicht. Im Bereich demokratischer Prozesse nennt er nur Systeme, die bestimmungsgemäß dazu verwendet werden sollen, das Ergebnis einer Wahl oder das Wahlverhalten zu beeinflussen. Allgemeine KI-Informationssysteme haben keinen solchen Verwendungszweck, obwohl wissenschaftliche Erkenntnisse darauf hindeuten, dass auch sie meinungsbildungsrelevant sind und Wahlen beeinflussen können. Die Kommission kann Anhang III nach Art. 7 Abs. 1 KI-VO per delegiertem Rechtsakt um neue Anwendungsfälle in bereits genannten Bereichen ergänzen. Für KI-Informationssysteme ist das bislang nicht geschehen. Die systembezogenen Anforderungen an Datenqualität, Protokollierung und Robustheit gelten für sie damit nicht.

Zwischenfazit

Die Regeln des DMStV-Entwurfs gegen koordinierte Reichweitenverstärkung und Nachahmung erfassen weder die gezielte Manipulation von Abrufkorpora noch die Ausgabe solcher Inhalte durch KI-Informationssysteme. Auch auf europäischer Ebene ist die Manipulation von Trainings- und Abrufkorpora für KI-Informationssysteme nicht ausdrücklich adressiert. Die KI-VO verlangt Robustheit gegen die Manipulation von Trainingsdaten nur von Hochrisiko-Systemen. Für Desinformation in KI-Antworten fehlt zudem ein verbindlicher Meldeweg, über den Nutzende Anbieter auf schädliche Inhalte aufmerksam machen können. Die Pflichten von DSA und KI-VO erreichen KI-Informationssysteme außerdem nur eingeschränkt. Sie knüpfen entweder an Schwellen bei Nutzerzahlen oder Rechenleistung an, die nur die größten Anbieter erreichen, oder an eine Zweckbestimmung, die nicht für diese Systeme gedacht wurde.

Lücke 4: Bezahlte Sichtbarkeit ohne Trennungsgebot

Problem und Regulierungsbedarf

Die bestehende Kooperation zwischen Axel Springer und OpenAI schlägt sich deutlich in der Quellenauswahl unserer Analyse nieder. Sie wird bei der Nutzung des KI-Systems nicht offengelegt und ist damit für Nutzende nicht ersichtlich. Verschärft wird dies durch die Einführung von Werbung: Seit dem 24. August 2026 spielt OpenAI auch in Deutschland Anzeigen in ChatGPT aus.[17] In einem KI-Antwortformat, das Werbung, lizenzierte und nicht lizenzierte Quellen im selben Fließtext zusammenführt, ist die Trennung von redaktionellem und bezahltem Inhalt unabdingbar.

Die klassische Werberegulierung hört bei KI-Systemen nicht auf. Trennungs- und Kennzeichnungsgebote für Werbung und gesponserte Inhalte müssen selbstverständlich umgesetzt werden. Besteht darüber hinaus zwischen dem KI-Anbieter und dem Anbieter einer Quelle eine entgeltliche oder geldwerte Vereinbarung, die den Zugang, die Auswahl oder die Platzierung dieser Quelle beeinflussen kann, muss dies an der Quellenangabe kenntlich gemacht und in den Transparenzangaben aufgeführt werden. Hier gilt es, die Funktionsweise von KI-Informationssystemen rechtlich angemessen und im Einklang mit bestehenden Prinzipien abzubilden.

[17] OpenAI, 2026.

Bestehende Regulierungsansätze und Lücken

Nationale Ebene

Die Medienintermediäre-Satzung verlangt in § 6 von Medienintermediären offenzulegen, ob und wie Zugang und Verbleib von Inhalten durch Entgeltzahlungen oder sonstige geldwerte Leistungen beeinflusst werden oder werden können. Die Norm wurde für bezahlte Platzierungen in Suchmaschinen und sozialen Netzwerken geschrieben, könnte aber auf KI-Antworten übertragen werden. Sie ist jedoch nicht auf KI-Systeme ausgelegt, sondern erfasst sie derzeit nur über die Auslegung des Intermediärsbegriffs, die die ZAK ihren Bescheiden vom 14. Juli 2026 zugrunde gelegt hat. Der DMStV-Entwurf verlangt in § 77c Abs. 1 DMStV-E nur Hinweise auf verwendete journalistisch-redaktionell verantwortete Inhalte und Angaben zu den Kriterien ihrer Einbeziehung. Eine Pflicht, entgeltliche Beziehungen oder Kooperationen offenzulegen, die die Quellenauswahl beeinflussen, sieht er weder in den allgemeinen Angaben noch an der einzelnen Quellenangabe vor. Die Transparenzpflicht sollte deshalb auch die Auswahl und Darstellung der Quellenverweise erfassen und die Offenlegung verlangen, ob Kooperationen die Auswahl beeinflussen.

Das Trennungsgebot des § 22 MStV verlangt für Telemedien, dass Werbung klar erkennbar und vom übrigen Inhalt eindeutig getrennt ist. Der DMStV-Entwurf fasst den Werbebegriff zudem weiter und erfasst jeden Inhalt, der gegen Entgelt oder eine ähnliche Gegenleistung aufgenommen wird (§ 2 Abs. 2 DMStV-E). Wer sich darauf beruft, dass keine Werbung vorliegt, muss dies künftig gegenüber der zuständigen Landesmedienanstalt nachweisen (§ 22 Abs. 1 DMStV-E). Als Telemedien würden KI-Informationssysteme diesen Vorgaben unterliegen. Die Formulierung des Entwurfs (gegen Entgelt in ein Angebot „aufgenommen“) passt ohne Weiteres zu Anzeigen, die neben einer KI-Antwort ausgespielt werden. Eine mögliche entgeltliche Gewichtung von Quellen, Produkten oder Anbietern bei der Generierung von KI-Antworten ist dagegen bisher nicht explizit erfasst. Gerade diese Formen sind für Nutzende besonders schwer zu erkennen, da sie in einer KI-generierten Antwort mit nicht entgeltlichen Inhalten vermischt vorkommen können. Es fehlt daher eine ausdrückliche Klarstellung, dass das Trennungsgebot auch für KI-generierte Antworten gilt und die Kennzeichnung an der Stelle erfolgen muss, an der Inhalte entgeltlich beeinflusst werden.

Des Weiteren untersagt § 94 MStV, ohne sachlich gerechtfertigten Grund systematisch von veröffentlichten Auswahlkriterien abzuweichen oder Kriterien anzuwenden, die journalistisch-redaktionelle Angebote unbillig systematisch behindern. Der Entwurf überträgt dieses Verbot in § 77d DMStV-E auf KI-Informationssysteme, allerdings nur für die Einbeziehung von Inhalten bei der Generierung und nicht ausdrücklich für die Quellenverweise. Die Neuerungen, die er für Medienintermediäre in § 94 einführt, übernimmt er nicht: das ausdrückliche Verbot der Selbstbevorzugung und die Regeln zu Verlinkungen. Lizenz- und Kooperationsvereinbarungen mit Medien sind grundsätzlich legitim. Die Begründung zu § 77d Abs. 2 DMStV-E hält es aber für denkbar, dass sie eine Bevorzugung einzelner Quellen sachlich rechtfertigen. Das sollte nur unter engen Voraussetzungen gelten: wenn die Kooperationen allen vergleichbaren Anbietern zu gleichen Bedingungen offenstehen, der Vertrag nur zwischen gleich relevanten Quellen den Ausschlag gibt, die Bevorzugung offengelegt wird und nicht beteiligte Angebote nicht systematisch verdrängt werden. Zudem können Verstöße nach § 77d Abs. 3 DMStV-E wie bei Intermediären grundsätzlich nur betroffene Anbieter geltend machen. Ein Vorgehen von Amts wegen sieht der Entwurf nur als eingeklammerte Option für offensichtliche Fälle vor. Für kleine Regionalmedien, deren Benachteiligung sich erst in großen Stichproben zeigt, ist das eine hohe Hürde.

EU-Ebene

Der DSA verlangt eine Werbekennzeichnung nach Art. 26, die nur für Online-Plattformen gilt, und ein Werbearchiv nach Art. 39, das auch sehr große Online-Suchmaschinen wie ChatGPT führen müssen. Für KI-Systeme, die wie ChatGPT als Suchmaschinen eingeordnet sind, gilt die Werbekennzeichnung nach Art. 26 damit nicht. Lizenz- und Kooperationsvereinbarungen über redaktionelle Quellen sind keine Werbung und werden von beiden Vorschriften damit nicht erfasst; für ihre Transparenz fehlt eine eigene Regelung.

Zwischenfazit

Damit bleiben zwei Lücken, die sich klar trennen lassen: Für Werbung fehlt vor allem eine ausdrückliche Klarstellung, dass Trennungsgebot und Kennzeichnungspflichten auch für KI-generierte Antworten gelten. Für Lizenz- und Kooperationsvereinbarungen fehlt eine Pflicht, sie an der Quellenangabe offenzulegen; zudem können kleine Medien eine Benachteiligung durch solche Vereinbarungen kaum geltend machen.

Handlungsempfehlungen

Die Empfehlungen ergeben sich aus den vier Regulierungslücken. Mehrere Instrumente schließen dabei mehr als eine Lücke: Das Belegobjekt etwa betrifft die Quellenangabe an der einzelnen Antwort (Lücke 1) ebenso wie die Herkunftskennzeichnung von Quellen (Lücke 2) und die Kennzeichnung entgeltlicher Quellenbeziehungen (Lücke 4). Die Empfehlungen sind nach Adressatenebene geordnet und verweisen jeweils auf die Lücken, die sie schließen.

Nationale Ebene: Länder und Rundfunkkommission

1 **KI-Informationssysteme als eigenständigen Telemedientypus beibehalten und Pflichten erweitern.** (Lücken 1 und 2)

Nach geltendem Recht lassen sich KI-Systeme allenfalls als Medienintermediäre erfassen, also als Vermittler fremder Inhalte. Das passt nicht zu Diensten, die selbst neue Texte erzeugen. Der Entwurf des DMStV Teil 2 führt deshalb einen eigenen Telemedientypus für KI-Informationssysteme ein und macht den Anbieter für die generierte Antwort verantwortlich. Die Rundfunkkommission sollte daran festhalten. Für Nutzende gibt es damit einen klaren Verantwortlichen für das, was ihnen angezeigt wird, und für die Aufsicht einen eindeutigen Adressaten. Zugleich ist der neue Typus die Grundlage für Sorgfalts- und Korrekturpflichten. Durch den Wechsel fallen allerdings die Quellenverweise in KI-Antworten aus dem Intermediärsrecht heraus. Damit Medien dort nicht schlechter gestellt sind als in klassischen Suchergebnissen, sollten die Pflichten, die der Entwurf (§ 94 DMStV-E) für Medienintermediäre neu einführt, in den §§ 77a bis 77e DMStV-E auch für Quellenverweise gelten: die Verlinkungspflicht, das Verbot der Selbstbevorzugung, die Grundsätze zur Nutzerautonomie und die besondere Auffindbarkeit (vgl. Empfehlungen 4 und 5).

2 **Transparenz über die Auswahlregeln für Quellen KI-spezifisch konkretisieren.** (Lücken 1 und 4)

Nach welchen Regeln ein KI-System Quellen aus der Websuche auswählt, ist bislang weder für Nutzende noch für die Aufsicht nachvollziehbar. Der Entwurf verlangt in § 77c Abs. 1 Nr. 3 nur Hinweise auf die verwendeten journalistisch-redaktionell verantworteten Inhalte und Angaben zu den Kriterien ihrer Einbeziehung. Diese Pflicht sollte auf alle verwendeten Quellen sowie auf

Auswahl und Darstellung der Quellenverweise erstreckt und um weitere maßgebliche Angaben ergänzt werden. Dies sind: verwendete Suchindizes und Suchdienste, Aufnahme- und Ausschlusskriterien für Domains, Gewichtung von Aktualität und thematischer Passung sowie Positiv- und Negativlisten. In einer neuen Nr. 4 sollten Anbieter zudem vertragliche Vereinbarungen offenlegen müssen, die die Einbeziehung von Inhalten oder die Darstellung von Quellenverweisen betreffen, etwa Lizenz- und Kooperationsvereinbarungen mit Medien. Welche Angaben im Einzelnen erforderlich sind, sollten die Landesmedienanstalten in ihrer gemeinsamen Satzung nach § 77g DMStV-E festlegen. So kann die Aufsicht die Auswahlregeln prüfen und Medien können nachvollziehen, warum ihre Angebote in Antworten vorkommen oder fehlen.

3 **Belegobjekt für jede KI-Antwort vorschreiben.** (Lücken 1, 2 und 4)
Angaben zur Quellenauswahl, die lediglich auf der Systemebene angegeben werden, erklären nicht, warum eine konkrete Antwort auf bestimmten Quellen beruht. Dafür sollte jede Antwort ein Belegobjekt mitführen, das alle genutzten Quellen und die jeweils davon belegten Textstellen ausweist. Im Vordergrund steht die *sichtbare Quellenangabe*, die viele Anbieter schon heute machen. Sie sollte verpflichtend werden und jede Quelle mit Aktualität, Herkunft und Link zur Originalquelle anzeigen. Journalistisch-redaktionelle Angebote sollten als solche gekennzeichnet werden. Hinzu kommen Sitzland, Eigentümer nach Art. 6 EMFA und eine etwaige Lizenz- oder Kooperationsbeziehung zum KI-Anbieter. Mindestens bei Nachrichten und politischen Informationen sollte § 77c Abs. 1 DMStV-E außerdem um eine Pflicht ergänzt werden, Aussagen den jeweils stützenden Quellen satzgenau zuzuordnen. Das ist wirksamer als die im DMStV-Entwurf aktuell vorgesehene Wahrscheinlichkeitsangabe, weil Sprachmodelle für einzelne Aussagen keine verlässlich kalibrierten Wahrscheinlichkeiten ausgeben. Erscheint in einer Antwort etwa die Domain eines Desinformationsnetzwerks neben einer Tageszeitung, wäre auf einen Blick erkennbar, welche Aussage auf welcher Quelle beruht und dass nur eine der Quellen ein journalistisch-redaktionelles Angebot ist. Nutzende können Aussagen so selbst überprüfen. Eine *maschinenlesbare Version*, etwa im JSON- oder YAML-Format, macht dieselben Angaben für Aufsicht und Forschung auswertbar und ließe sich leicht um weitere für Forschung und Aufsicht relevante Informationen ergänzen. Anders als Formate, die nur technisch auslesbar sind, etwa PDF- oder HTML-Seiten, sind JSON und YAML offen standardisiert, klar strukturiert und zugleich für Menschen lesbar; sie lassen sich ohne proprietäre Software automatisiert auswerten und über Anbieter hinweg vergleichen. Ergänzend sollte die Vorlage von Unterlagen auf Verlangen (§ 77f DMStV-E) zu einer regelmäßigen, nicht personenbezogenen Berichtspflicht erweitert werden, die ausweist, welche Domains in welchem Umfang in Antworten einfließen. Der Aufwand ist gering, weil Anbieter abgerufene URLs, Domains und Veröffentlichungsdaten ohnehin vorhalten und viele Dienste Quellen bereits anzeigen.

4 **Verlinkung zu Originalquellen in KI-Antworten verpflichtend machen.** (Lücke 1)
KI-Antworten ersetzen für viele Nutzende den Besuch der Originalseite. Der Quellenverweis ist deshalb oft der einzige Weg, auf dem Nutzende eine Aussage bis zu ihrem Ursprung zurückverfolgen und Medien noch Leserinnen und Leser erreichen, auf die sie zur Refinanzierung angewiesen sind. Viele Anbieter geben bereits heute Links aus, aber freiwillig und in unterschiedlicher Form. Jede verwendete Quelle sollte deshalb an der Aussage, die sie belegt, als funktionierender, unentgeltlicher Link zum Ursprungsangebot ausgegeben werden müssen, so wie es der Entwurf in § 94 Abs. 1a DMStV-E für Medienintermediäre bereits vorsieht.

5 Diskriminierungsverbot erweitern und durchsetzbar machen. (Lücken 1 und 4)
Welche journalistischen Angebote in KI-Antworten vorkommen, entscheidet sich auf zwei Ebenen: bei der Generierung (also welche Inhalte die Aussage prägen) und bei den Quellenverweisen (also welche Angebote angezeigt und ggf. verlinkt werden). Das Diskriminierungsverbot in § 77d DMStV-E erfasst nur die erste Ebene. Es muss klargestellt werden, dass journalistisch-redaktionell verantwortete Angebote auch bei Auswahl, Anordnung und Darstellung der Quellenverweise nicht diskriminiert werden dürfen. Zudem muss der DMStV die Bevorzugung eigener Angebote untersagen, wie § 94 Abs. 2 das für Intermediäre bereits vorsieht. Die Begründung zu § 77d Abs. 2 DMStV-E hält es für denkbar, dass Lizenz- und Kooperationsvereinbarungen eine Bevorzugung rechtfertigen. Das sollte voraussetzen, dass diese Vereinbarungen nach § 77c DMStV-E offengelegt werden und vergleichbaren Anbietern zu gleichwertigen Bedingungen offenstehen. Damit das Verbot auch durchgesetzt werden kann, sollten neben betroffenen Anbietern auch Einrichtungen der Wissenschaft und zivilgesellschaftliche Organisationen Verstöße geltend machen können. Der DMStV Teil 2 sollte es außerdem ermöglichen, offensichtliche Verstöße von Amts wegen zu verfolgen, und eine Beweislastumkehr vorsehen. Zeigt die Quellenverteilung in den Berichten (vgl. Empfehlung 3) eine systematische Benachteiligung journalistisch-redaktioneller Angebote, muss der Anbieter den sachlichen Grund dafür darlegen. Davon profitieren vor allem kleine und regionale Medien, deren Benachteiligung sich erst in großen Stichproben zeigt und die sie allein kaum nachweisen könnten.

6 Bestehende Mediendatenbanken zu maschinenlesbaren Referenzen ausbauen. (Lücken 2 und 4)
Herkunftsangaben im Belegobjekt (vgl. Empfehlung 3) sind nur so verlässlich wie die Daten, auf denen sie beruhen. KI-Anbieter brauchen dafür eine staatsferne, faktenbasierte Referenz, die sie automatisiert abfragen können. Dafür eignet sich zum Beispiel die bestehende Mediendatenbank der Kommission zur Ermittlung der Konzentration im Medienbereich (KEK), die der Entwurf als Datenbank des Rats für Konzentrationsermittlung und Medienökonomie weiterentwickelt (§103k DMStV-E). Sie könnte über eine Schnittstelle maschinenlesbar zugänglich gemacht werden und um ein öffentlich einsehbares Register der Lizenz- und Kooperationsvereinbarungen zwischen KI-Anbietern und Medien erweitert werden. Auf europäischer Ebene könnten Kommission und Europäisches Gremium für Mediendienste dafür sorgen, dass auch die nach Art. 6 EMFA aufzubauenden nationalen Datenbanken zum Medieneigentum maschinenlesbar und in einem gemeinsamen Datenformat abrufbar sind, damit KI-Anbieter Herkunftsangaben EU-weit auf dieselbe Grundlage stützen können. Nutzende erhalten so verlässliche Angaben und die Aufsicht kann sie überprüfen.

7 Korrekturpflicht für fehlerhafte KI-Antworten einführen. (Lücke 2)
Fehler in KI-Antworten wiederholen sich über viele Anfragen hinweg, solange ihre Ursache fortbesteht. Ziel muss deshalb sein, dass als falsch erkannte Darstellungen nicht erneut ausgegeben werden. Dazu muss die ausgewiesene Quelle die jeweilige Aussage tatsächlich tragen. § 77b Abs. 3 DMStV-E sollte auch Fehlzuschreibungen erfassen, also Aussagen, die Medien, öffentlichen Stellen oder Parteien zugeschrieben werden, ohne von ihnen zu stammen. Die Pflicht zu Verfahren der nachträglichen Überprüfung und Korrektur (§ 19 Abs. 1 DMStV-E), die der Entwurf für journalistische Angebote einführt, sollte ausdrücklich auf KI-Informationssysteme übertragen werden. Sie sollte mit einem niedrigschwelligen Meldeweg verbunden sein, über den Nutzende, betroffene Medienanbieter und zivilgesellschaftliche Organisationen fehlerhafte Darstellungen anzeigen können (vgl. Empfehlung 11).

8 **Trennungsgebot ausdrücklich auf KI-Informationssysteme erstrecken.** (Lücke 4)
Mit der Einführung von Werbung in KI-Systemen wächst das Risiko, dass bezahlte Inhalte in Antworten einfließen, ohne dass Nutzende dies erkennen. Der DMStV-Entwurf sollte deshalb klarstellen, dass das Trennungs- und Kennzeichnungsgebot des § 22 Abs. 1 MStV auch für KI-generierte Antworten gilt. Bezahlen Dritte dafür, dass ihre Quellen, Produkte oder Angebote bei der Generierung stärker berücksichtigt werden, ist das Werbung und an der Stelle zu kennzeichnen, an der das Ergebnis entgeltlich beeinflusst wird. Das betrifft explizit nicht nur den Bereich neben der KI-Antwort, sondern die generierte Antwort selbst. Nicht gekennzeichnete Einflussnahmen auf Antworttext, Quellenauswahl oder Reihenfolge sollten ausdrücklich als unzulässig gelten. Lizenz- und Kooperationsvereinbarungen, mit denen KI-Anbieter Medien für den Zugriff auf deren Inhalte vergüten, sind dagegen keine Werbung. Für sie sollten Transparenz- und Diskriminierungsregeln gelten (vgl. Empfehlungen 2 und 5).

Europäische Ebene: EU-Kommission

9 **Quellenauswahl als Hauptparameter nach Art. 27 DSA offenlegen.** (Lücke 1)
Nationale Transparenzpflichten erreichen Anbieter mit Sitz in anderen Mitgliedstaaten nur eingeschränkt. Für die größten KI-Systeme braucht es deshalb auch eine EU-weite Grundlage. Art. 27 DSA verpflichtet Online-Plattformen, die Hauptparameter ihrer Empfehlungssysteme offenzulegen. Art. 38 DSA setzt voraus, dass diese Pflicht auch für sehr große Suchmaschinen gilt, obwohl Suchmaschinen nicht zu den Online-Plattformen zählen. Die Kommission sollte klarstellen, dass als VLOSE benannte KI-Systeme danach auch die Hauptparameter ihrer Quellenauswahl im laufenden Betrieb offenlegen müssen. Das schafft für Nutzende und Aufsicht in der gesamten EU eine Mindesttransparenz. Es bleibt allerdings bei Angaben auf Systemebene. Die Herkunft der Quellen in der einzelnen Antwort macht erst ein Instrument wie das maschinenlesbare Belegobjekt sichtbar (vgl. Empfehlung 3).

10 **Manipulation von Trainings- und Abrufkorpora als systemisches Risiko benennen.** (Lücke 3)
Sehr große Suchmaschinen müssen systemische Risiken für die öffentliche Debatte bewerten und mindern (Art. 34 und 35 DSA). Die gezielte Bespielung von Crawlern und Abrufkorpora ist dort bislang nicht als eigenes Risiko benannt, sodass Anbieter sie nicht ausdrücklich berücksichtigen müssen. Die Leitlinien der Kommission zu Art. 34 und 35 DSA und der Verhaltenskodex gegen Desinformation sollten klarstellen, dass diese Form der Manipulation, wie sie etwa das Pravda-Netzwerk betreibt, ein zu bewertendes und zu minderndes Risiko darstellt. Die Kommission erhält damit einen Anhaltspunkt für die Bewertung der Risikoberichte der Anbieter, und Nutzende werden besser vor eingeschleuster Desinformation geschützt.

11 **Meldeverfahren als Maßnahme der Risikominderung empfehlen.** (Lücke 3)
Für Desinformation in KI-Antworten gibt es keinen verbindlichen Meldeweg, weil das Meldeverfahren nach Art. 16 DSA nur für Hostingdienste und rechtswidrige Inhalte gilt. Als VLOSE benannte KI-Systeme sollten Nutzenden deshalb als Maßnahme der Risikominderung nach Art. 35 DSA ermöglichen, falsche und schädliche Inhalte zu melden. Anbieter erfahren so schneller von Fehlern und können darauf reagieren. Weil Nutzende Fehler ohne Vorwissen oft nicht erkennen, können Meldewege die übrigen Pflichten zur Risikominderung nur ergänzen, nicht ersetzen.

12 **KI-Informationssysteme in Anhang III der KI-VO aufnehmen.** (Lücke 3)
Die strengen Anforderungen der KI-VO an Datenqualität, Protokollierung, Robustheit gegen Manipulation und die Meldung schwerwiegender Vorfälle gelten nur für Hochrisiko-Systeme. KI-Chatbots und KI-Suche gehören bislang nicht dazu. Im Bereich demokratischer Prozesse erfasst Anhang III nur Systeme, die bestimmungsgemäß Wahlen oder das Wahlverhalten beeinflussen sollen, obwohl auch KI-Informationssysteme für die Meinungsbildung relevant sind. Die Kommission sollte Anhang III Nr. 8 deshalb per delegiertem Rechtsakt nach Art. 7 KI-VO um KI-Systeme ergänzen, die auf Grundlage einer Websuche Antworten zu aktuellen Nachrichten sowie zu politischen und gesellschaftlichen Themen generieren. Die Aufsicht erhält damit Einblick in die Funktionsweise der Systeme, und Anbieter müssen ihre Systeme nachweislich gegen Manipulation absichern. Da die Pflichten für Hochrisiko-Systeme nach Anhang III ab Dezember 2027 gelten, sollte die Ergänzung jetzt vorbereitet werden.

Literaturverzeichnis

ARD/ZDF-Forschungskommission. (2026). *ARD/ZDF-Medienstudie 2026*. https://www.ard-zdf-medienstudie.de/

Autoriteit Persoonsgegevens. (2025). *Rapportage AI & algoritmes Nederland – RAN special: AI-chatbots als stemhulp* [Bericht KI & Algorithmen Niederlande – RAN-Sonderausgabe: KI-Chatbots als Wahlentscheidungshilfe]. Verfügbar unter: https://www.autoriteitpersoonsgegevens.nl/documenten/ran-special-ai-chatbots-als-stemhulp

Baeva, G., Benert, V. & Klausa, T. (2026). Medienvielfalt digital denken und konsequent durchsetzen: Stellungnahme der Agora Digitale Transformation gGmbH zum Digitale Medien-Staatsvertrag Teil 2. Agora Digitale Transformation (Hg.). https://agoradigital.de/wp-content/uploads/2026/11/2026-11-13_Stellungnahme_DMStV_II_ADT.pdf

Benert, V. & Timmermann, S. (2026). Meinungsbildung im Wandel: Wie KI-Systeme Quellen bei der Nachrichtensuche selektieren. Agora Digitale Transformation (Hg.). https://doi.org/10.5281/zenodo.21162562

Chapekis, A. & Lieb, A. (2025, 22. Juli). Google users are less likely to click on links when an AI summary appears in the results. Pew Research Center. https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/

DFRLab, 2026. *The Pravda Network*. https://dfrlab.org/the-pravda-network/

Europäische Kommission (2026, 31. August). *Commission designates ChatGPT, Reddit, Roblox under Digital Services Act* [Pressemitteilung IP/26/1772]. https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1772

Fletcher, J., & Verckist, D. (2025). *News integrity in AI assistants: An international PSM study*. European Broadcasting Union; BBC. https://www.ebu.ch/files/live/sites/ebu/files/Publications/Intelligence/open/EBU-Intelligence-BBC_News_Integrity_in_AI_Assistants_Report_2025.pdf

Gerichtshof der Europäischen Union. (2026, 3. November). *Mündliche Verhandlung in der Rechtssache C-474/25, Medienanstalt Berlin-Brandenburg* [Video]. Curia. https://curia.europa.eu/site/jcms/2712042_EmbedVideo/de/mundliche-verhandlung-03/11/2026

Heßler, E. (2026). KI-generierte Antworten bei der Informationssuche: Verbreitung und Wahrnehmung. KI & Search-Studie 2025. *Media Perspektiven*, (11), 1–14. https://www.media-perspektiven.de/publikationsarchiv/2026/mp-11/2026-ki-generierte-antworten-bei-der-informationssuche-verbreitung-und-wahrnehmung

Jaźwińska, K. & Chandrasekar, A. (2025, 6. März). AI Search Has a Citation Problem. Columbia Journalism Review, Tow Center for Digital Journalism. https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php

Lin, H., Czarnek, G., Lewis, B., White, J. P., Berinsky, A. J., Costello, T., Pennycook, G., & Rand, D. G. (2025). Persuading voters using human–artificial intelligence dialogues. *Nature*, *648*(8093), 394–401. https://doi.org/10.1038/s41586-025-09771-9

Die Medienanstalten. (2026, 14. Juli). *ZAK erlässt erstmalig Bescheide gegen KI-Angebote von Google und Perplexity* [Pressemitteilung]. https://www.die-medienanstalten.de/presse/pressemitteilungen/zak-bescheide-ki-angebote-google-perplexity/

OpenAI (18.08.2026): ChatGPT Ads wird in Europa ausgeweitet. https://openai.com/de-DE/index/chatgpt-ads-expands-across-europe/

Oster, J., & Busch, C. (2026). *Integration von KI-Anwendungen in Suchmaschinen: Eine medienrechtliche und regulatorische Einordnung* [Rechtsgutachten im Auftrag der Direktorenkonferenz der Landesmedienanstalten]. Die Medienanstalten (Hg.). https://www.die-

medienanstalten.de/fileadmin/user_upload/die_medienanstalten/Service/Studien_und_Gutachten/KI-Gutachten_Recht_Regulierung_final.pdf

Richtlinie 2000/31/EG des Europäischen Parlaments und des Rates vom 8. Juni 2000 über bestimmte rechtliche Aspekte der Dienste der Informationsgesellschaft, insbesondere des elektronischen Geschäftsverkehrs, im Binnenmarkt („Richtlinie über den elektronischen Geschäftsverkehr"). (2000). *Amtsblatt der Europäischen Gemeinschaften*, L 178, 1–16. https://eur-lex.europa.eu/eli/dir/2000/31/oj

Rundfunkkommission der Länder (2026). Diskussionsentwurf der Rundfunkkommission für einen „Digitale Medien-Staatsvertrag (DMStV) Teil 2", Anhörungsfassung vom 28. September 2026. https://rundfunkkommission.rlp.de/fileadmin/rundfunkkommission/Dokumente/DMStV_2/Dokumente/2026-09-28_ANHOERUNG_DMStV_II_RFK.pdf

Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung der Richtlinie 2000/31/EG (Gesetz über digitale Dienste). (2022). *Amtsblatt der Europäischen Union*, L 277, 1–102. https://eur-lex.europa.eu/eli/reg/2022/2065/oj

Hinweis zur Verwendung von Künstlicher Intelligenz

Bei der Erstellung dieses Papiers wurde Claude Opus 5.5 von Anthropic, zugänglich über claude.ai, im Zeitraum September bis November 2026 unterstützend eingesetzt. Der Einsatz umfasste das Übersetzen fremdsprachiger Inhalte ins Deutsche oder Englische sowie die Überarbeitung von Textabschnitten auf Basis eigener Textentwürfe. Alle KI-generierten Inhalte wurden von der Autorin geprüft, überarbeitet und verantwortet; eine ungeprüfte Übernahme fand zu keinem Zeitpunkt statt.

Über die Agora Digitale Transformation

Agora Digitale Transformation ist der Thinktank für Updates unserer Demokratie. Ziel ist es, die Chancen der Digitalen Transformation zu nutzen, um die Demokratie in Deutschland zu stärken. Als gemeinnützige Organisation arbeitet der Thinktank überparteilich, kollaborativ und evidenzbasiert mit dem Schwerpunkt auf umsetzbare und wirksame Lösungen für die Politik. Die Agora Digitale Transformation sucht dafür gezielt den Austausch und die Zusammenarbeit mit Innovator:innen aus Gesellschaft, Verwaltung, Wissenschaft, Wirtschaft und Politik. Initiiert wurde die Agora Digitale Transformation im Jahr 2023 von der Stiftung Mercator. Alleiniger Gesellschafter ist die Stiftung Mercator, die die Agora Digitale Transformation fördert.

Ansprechperson bei der Agora Digitale Transformation



Dr. Vivien Benert
Innovation Lead: Digitale Öffentlichkeit
vivien.benert@agoradigital.de

Agora Digitale Transformation

www.agoradigital.de