

## Austrian Lab for AI Trust\* Branchen Report 2

### Voraussetzungen für den vertrauenswürdigen Einsatz von Künstlicher Intelligenz in Nachrichtenmedien

Branchen-Report mit konkreten Handlungsvorschlägen für Führungskräfte, KI- und Strategie-Verantwortliche sowie Journalist:innen

## Executive Summary

Künstliche Intelligenz (KI) verändert fundamental, wie Nachrichtenmedien arbeiten und wie ihre Produkte in der Gesellschaft empfangen und wahrgenommen werden: Ihr Einsatz beeinflusst den journalistischen Prozess der Erarbeitung von Text-, Bild- und Video-Inhalten in Zeitungs-, Zeitschrift- und Rundfunk-Unternehmen ebenso wie deren Verbreitung in verschiedenen Formaten über digitale Plattformen und Kanäle. KI-Werkzeuge dirigieren und filtern Nachrichtenflüsse. Damit entscheiden sie mit, welche Inhalte Konsument:innen wo und wann zu sehen und zu hören bekommen.

KI-Technologien spielen entlang der gesamten Medien-Wertschöpfung eine zunehmend wichtige Rolle. Auf der Plusseite begleiten Stichworte wie Effizienz, Personalisierung und Reichweiten-Optimierung diese Entwicklung, auf der Negativseite die Steigerung von Misstrauen, Fakenews und Qualitätsprobleme wie Halluzinationen, Bias und Stereotype und Verlust von Faktizität. Medienhäuser sind enorm gefordert, sich mit dem KI-Einsatz grundlegend und mit großer Intensität auseinanderzusetzen: Wie können sie KI-Werkzeuge nutzen, ohne gleichzeitig das größte Kapital zu verspielen, das sie besitzen, nämlich das Vertrauen ihres Publikums?

Diese aktuelle Frage steht im Zentrum dieses Branchenreports. Er wurde 2026 im Rahmen des [Austrian Lab for AI Trust \(ALAIT\)](#) erarbeitet. Das Forschungsprojekt ALAIT entwickelt neuartige Bewertungsmethoden für KI-Technologien in risikobehafteten Anwendungskontexten und legt in ALAIT-Dossiers konkrete Chancen und Risiken dar.

Im Format ALAIT Lab werden mit Vertreter:innen aus Anwendungsbranchen Empfehlungen erarbeitet, wie KI-Technologien in den jeweiligen Branchen sicher, verantwortungsvoll und im Einklang mit den jeweiligen Qualitätsstandards eingesetzt werden können. Ein solches ALAIT Lab fand am 5. März 2026 in Wien zum Thema „KI im Redaktionsalltag: Gemeinsam Medienvertrauen schaffen“ mit mehr als 30 Teilnehmenden aus österreichischen Medienbetrieben statt. Anhand eines Negativ-Szenarios, das einen künftigen intensiven KI-Einsatz und einen damit verbundenen Vertrauensverlust bei den Empfänger:innen veranschaulichte, wurde in Kleingruppen ein praktikables Positiv-Szenario erarbeitet. Dieses zeigt auf, welche Voraussetzungen geschaffen



werden müssen, damit KI sicher und unter Einhaltung hoher journalistischer Standards eingesetzt werden kann.

**Das Ergebnis sind Handlungsvorschläge und Strategien, die sich an drei Zielgruppen richten:**

1. Medienmanagement und Chefredaktion
2. KI-Verantwortliche und Produktmanager:innen in Medienbetrieben
3. Redakteur:innen und Faktenchecker:innen

**Inhaltlich fokussieren sich die Empfehlungen auf vier Bereiche:**

- Die Entwicklung klarer KI-Leitlinien und Governance-Strukturen
- Eine Fokussierung auf das Medienpublikum durch höhere Transparenz und Interaktion
- Den Aufbau von KI-Kompetenz auf allen Ebenen des Hauses
- Die Etablierung einer konstruktiven Fehlerkultur.

# Einführung und Zielsetzung

Nachrichtenmedien stehen unter dem Druck ökonomischer Zwänge, da Geschäftsmodelle erodieren, und kämpfen um die Aufmerksamkeit des Medienpublikums. Gleichzeitig steigt die Notwendigkeit, sich gegen Desinformation und Deepfakes zu behaupten. Generative KI, welche Texte, Bilder und Videos produzieren und verändern kann sowie KI-Werkzeuge, welche die Ausspielung von Medieninhalten entsprechend individuellen Vorlieben des Medienpublikums erlauben, sind zwei bedeutende KI-Einsatzbereiche. Sie versprechen eine Weiterentwicklung von Geschäftsmodellen und werden häufig auch als Sparpotenzial gesehen. Der unreflektierte Einsatz der Technologien birgt jedoch die Gefahr, journalistische Sorgfaltspflichten und Qualitätsansprüche zu untergraben.

Dies ist nicht nur auf die Fehler von KI-Werkzeugen wie Halluzinationen und Falschinformationen, Bias und Reproduktion von Stereotypen oder rechtliche Problem- und Graubereiche, etwa bei Datenschutz und Urheberrecht, zurückzuführen. In einer Gesellschaft, in der Institutionen – so auch etablierte Nachrichtenmedien – generell stärker hinterfragt werden als in der Vergangenheit, kann ein unüberlegter KI-Einsatz zum Vertrauensverlust auf Seiten des Publikums beitragen und den Wert von renommierten Medienmarken verringern. Medienkonsumenten betrachten KI in Medien mit einer gewissen Skepsis. Zusätzlich verstärkt die Flut an Deepfakes auf Social Media die Skepsis betreffend den Wahrheitsgehalt von Information: Was ist wahr, was ist Fake? Welchen Inhalt und welchen Absender:innen kann noch vertraut werden?

Medienkonsument:innen verlangen in zunehmendem Maße Transparenz in Bezug auf die Herkunft von Texten, Bildern und Videos: Woher stammt die Information? Ist sie tatsächlich glaubwürdig? Wurde sie von einer KI generiert oder von Menschen verifiziert? Gleichzeitig hat diese Transparenz auch ihre Grenzen: Wird jeder Inhalt gekennzeichnet, weil im Laufe seiner Entstehung an irgendeiner Stelle KI im Spiel war, hat das Label keinen Wert mehr.

Somit stellt sich für die Medienbranche nicht nur die Frage, für welche Aufgaben und wie sie KI-Werkzeuge einsetzt, sondern auch, wie sie begleitend dazu die Kommunikation mit dem Publikum verändern muss, und glaubwürdige, qualitätsgesicherte Information weiterhin bereitstellen kann.

Für die Medienhäuser ergeben sich angesichts der Präsenz von KI-Systemen und den von ihnen generierten Inhalten neue Fragen der Differenzierung und Qualitätssicherung: Wie müssen Arbeitsprozesse, Verantwortlichkeiten und Aufgaben künftig gestaltet werden, um hochwertigen Nachrichten-Content erkennbar zu machen, zu produzieren und für ihn die Verantwortung übernehmen zu können?

Dieser Branchen-Report dokumentiert die Ergebnisse des ALAIT Labors, die mit mehr als 30 Vertreter:innen von österreichischen Nachrichtenmedien am 5. März 2026 in Wien in den Räumen der Austria Presse Agentur zum Thema „KI im Redaktionsalltag: Gemeinsam Medienvertrauen schaffen“ stattgefunden hat.

## Voraussetzungen für einen vertrauenswürdigen KI-Einsatz in Nachrichtenmedien

Die Analyse der Workshop-Ergebnisse zeigt, dass Vertrauen nicht durch technische Lösungen allein, sondern durch soziotechnische Prozesse entsteht, in denen Technologie mit einer Veränderung von Arbeitsprozessen und menschlichen Kontrollen kombiniert wird.

Wie kann KI verantwortungsvoll eingesetzt werden, ohne das Vertrauen des Medienpublikums (weiter) zu schwächen? Folgende Handlungsvorschläge wurden von den Teilnehmer:innen des ALAIT Labors erarbeitet. Sie geben für die drei Zielgruppen Medienmanagement und Chefredaktion, KI-Verantwortliche und Produktmanager:innen und Redakteur:innen und Faktenchecker:innen Antworten für den Praxiseinsatz und berücksichtigen dabei das komplexe Zusammenspiel von Technologie, Menschen und Organisation.

## 1. Handlungsvorschläge für Medienmanagement und Chefredaktion

### Entwicklung einer unternehmensweiten KI-Strategie und von Leitlinien

- **Strategische Verankerung:** Es muss eine klare Unternehmensstrategie entwickelt werden, die definiert, wo, wie und wofür KI eingesetzt wird. Dies darf nicht isoliert in einem bestimmten Unternehmensbereich geschehen, sondern erfordert die Beteiligung aller Stakeholder im Unternehmen.
- **Konkrete Richtlinien:** So wie journalistische Ethikrichtlinien Standard sein sollten, müssen KI-Leitlinien etabliert und laufend überarbeitet werden. Diese müssen konkretisieren:
  - Für welche Aufgaben darf KI eingesetzt werden (z. B. Datenauswertung versus generative Texterstellung)?
  - Für welche Aufgaben darf KI nicht eingesetzt werden?
- **Ressourcenbereitstellung:** Sollen die Qualität von Medienprodukten und das Vertrauen des Publikums bestehen bleiben, kann KI nicht primär als Sparmaßnahme verstanden werden. Die Geschäftsführung muss finanzielle und personelle Ressourcen für den Aufbau von Kompetenzen, Schulungen und die Anpassung von Prozessen bereitstellen.

### Fokussierung auf das Medienpublikum: User Centricity

- Etablierung einer Transparenz-Kultur
  - Allgemeine Transparenz: Wie können redaktionelle Prinzipien inkl. KI-Richtlinien transparent für das Medienpublikum dargestellt werden?
  - Spezifische Transparenz: Wie wird über die Entstehungsgeschichte von Inhalten, also Artikeln, Bildern und Videos informiert, inklusive eventuellem KI-Einsatz?

- **Interaktion und Community-Aufbau:** Um Vertrauen und Glaubwürdigkeit herzustellen, braucht es mehr direkte Interaktion mit dem Medienpublikum (etwa über Veranstaltungen, mobile Redaktionen etc.) und entsprechende Ressourcen dafür.

### Förderung von KI-Kompetenz auf allen Ebenen

- **Schulung der Führungsebene:** KI-Kompetenz wird nicht nur in spezialisierten Teams benötigt, auch Geschäftsführer:innen und Chefredakteur:innen benötigen KI-Schulungen, um die Implikationen der Technologie für das Geschäftsmodell und die Glaubwürdigkeit zu verstehen.
- **Kompetenzfahrplan für Redaktion und Produktion:** Maßgeschneidert zu den eingesetzten KI-Werkzeugen sollen Schulungen für das jeweils involvierte Team stattfinden, bei denen Anwendungen als auch Limitationen (Fehleranfälligkeit der KI) trainiert werden.

### Etablierung einer Fehlerkultur

- Im Umgang mit KI passieren jedenfalls Fehler. Medienhäuser benötigen einen definierten Weg, um mit solchen Fehlern umzugehen und diese transparent gegenüber dem Publikum zu kommunizieren.

## 2. Handlungsvorschläge für KI-Verantwortliche und Produktmanager:innen

### Einrichtung von KI-Beauftragten

- **Integration in den Prozess:** KI-Beauftragte sind essenziell, dürfen aber nicht isoliert agieren. Wichtig ist, dass sie in redaktionelle Prozesse eingebunden sind.
- **Teamaufgabe:** Die Verantwortung für KI-Entscheidungen sollte nicht bei einer Einzelperson liegen, sondern auf mehrere Personen verteilt werden.

### Kontrollierbarkeit sicherstellen

- Da die Organisation für die KI-Prozesse und ihre Ergebnisse geradestehen muss (Erfolg am Markt, Haftung, etc.) und entsprechendes Monitoring nötig ist, sollte der KI-Einsatz von Anfang an auf ein kontrollierbares Ausmaß beschränkt werden: Was nicht durch die Anwender:innen und KI-Verantwortliche kontrolliert werden kann (etwa weil die Aufwendungen zu groß wären), sollte nicht eingesetzt werden.

### Technische Gestaltung und Sicherheit

- **Vermeidung vollautomatisierter Prozesse:** KI-Systeme sollten so gestaltet sein, dass sie den Menschen zur aktiven Entscheidung zwingen (z. B. Auswahl aus mehreren Optionen), statt Inhalte vollautomatisch zu publizieren.
- **Aufmerksamkeitsmechanismen:** Es sollten Mechanismen implementiert werden, die den/die Nutzer:in regelmäßig daran erinnern, dass er/sie mit einer KI interagiert (z. B. Intervall-Checks), um die menschliche Kontrolle zu sichern.

### Transparenz und Kennzeichnung

- **Kennzeichnung von KI-generierten Inhalten:** In Einklang mit den EU-Vorgaben (EU AI Act) sollten Bilder, Audios, Videos und Texte gekennzeichnet werden, die mithilfe von KI erstellt wurden. Dabei obliegt es den Redaktionen zu entscheiden, ob sämtliche KI-generierten Inhalte visuell sichtbar gekennzeichnet werden oder

nur DeepFakes. Hier sind im Medienbereich derzeit unterschiedliche Strategien zu beobachten. Öffentlich-rechtliche Medien, aber auch große Medienverlegerischer Herkunft sprechen sich z.B. klar gegen die Verwendung von fotorealistischen KI-Bildern in journalistischen Produkten aus oder verpflichten sich, KI-generierte Illustrationen der/dem Nutzenden gegenüber transparent zu machen, was über die Anforderungen des EU AI Act hinausgeht. Daneben gibt es andere Medien, die sich darauf beschränken, den EU AI Act zu erfüllen. Bei der Kennzeichnung ist die Verhältnismäßigkeit zu beachten (z.B. wenn Texte nur mit KI übersetzt wurden, hat die Anwendung des Labels „KI-generiert“ wenig Aussagekraft). Auch ein gemeinsames Vorgehen der Medienbranche im Sinne der Informationsintegrität ist unerlässlich.

- **Schutz vor Manipulation:** Für eigene audiovisuelle Inhalte müssen perspektivisch fälschungssichere Wasserzeichen entwickelt werden bzw. bestehende Lösungen zu Einsatz kommen, um die Authentizität der eigenen Produkte zu schützen.
- **Qualitätssiegel für Inhalte, welche von Menschen verifiziert wurden:** Langfristig sollte der Fokus von der Kennzeichnung der KI-Nutzung auf ein positives Qualitätssiegel für „journalistisch verifizierte Inhalte“ verschoben werden. Dem Publikum ist weniger wichtig, ob KI beteiligt war, als dass die Information stimmt.

## 3. Handlungsvorschläge für Redakteur:innen und Faktenchecker:innen

### Wahrung der journalistischen Sorgfaltspflicht

- **Mensch im Mittelpunkt:** Trotz KI-Einsatz bleibt die menschliche Komponente unverzichtbar. Das klassische Fact-Checking und das 4-Augen-Prinzip dürfen nicht aufgegeben werden.
- **Kritischer Umgang mit Quellen:** Redakteur:innen müssen lernen, KI-Outputs kritisch zu hinterfragen. Blindes Vertrauen in maschinell generierte Inhalte und Entscheidungen führt zu Fehlern (z. B. nicht funktionierende Links, Halluzinationen).
- **Unterscheidung bei KI-Einsatzfällen:** Es muss klar zwischen KI zur Datenauswertung (stark empfohlen) und generativer KI zur Schaffung neuer Inhalte (Vorsicht geboten) unterschieden werden.

### Datenqualität und Dokumentation

- **Bewusstsein für Datenqualität:** Fehler in der Protokollierung oder Eingabe führen zu fehlerhaften Ausgaben. Die Sensibilisierung für Datenqualität ist eine Grundvoraussetzung für den erfolgreichen KI-Einsatz.
- **Dokumentation des Prozesses:** Der Einsatz von KI sollte im redaktionellen Workflow dokumentiert werden, um Nachvollziehbarkeit zu gewährleisten.

### Aktive Rolle als Multiplikator:innen

- Redakteur:innen können als Botschafter:innen für einen qualitätsvollen KI-Einsatz wirken, indem sie sich aktiv mit den Chancen und Risiken auseinandersetzen und ihr Wissen im Team teilen (Mentoring).

# Zusammenfassung der Kernbotschaften

## Die Diskussion im ALAIT Lab ergab folgende zentrale Erkenntnisse für die Medienbranche:

1. **Nutzende im Fokus:** Die Frage „Was braucht der Medienkonsument?“ muss im Zentrum stehen. Transparenz und Interaktion mit Nutzer:innen sind Schlüssel zur Schaffung von Vertrauen.
2. **Qualität vor Effizienz:** Die Gefahr, KI primär als Kostensenkungsinstrument zu sehen, muss vermieden werden, da dies die journalistische Qualität und damit Vertrauen gefährdet.
3. **Kompetenz ist Voraussetzung:** KI-Kompetenz fehlt aktuell in der Breite der Redaktionen. Schulungen sind für alle, von der Geschäftsführung bis zum Praktikanten, unerlässlich.
4. **Transparenz vor Geheimhaltung:** Der Grundsatz „Lieber zu viel als zu wenig“ bei der Offenlegung von journalistischen Arbeitsweisen. Das Medienpublikum will ehrlich informiert werden.
5. **Kooperation im Ökosystem:** Medienhäuser sind Teil eines größeren Ökosystems. Zusammenarbeit (z. B. mit Agenturen wie der APA, PR-Dienstleistern) und gemeinsame Standards sind notwendig, um Desinformation und AI Slop zu bekämpfen.

## Über das Projekt ALAIT

Das Austrian Lab for AI Trust (ALAIT) ist ein vom österreichischen Bundesministerium für Innovation, Mobilität und Infrastruktur (BMIMI) initiiertes Forschungs- und Entwicklungsprojekt zur Schaffung von Vertrauen durch Wissen im Bereich Künstliche Intelligenz (KI). Das Projekt ALAIT zielt darauf ab, Interessierte und wichtige gesellschaftliche Gruppen zu befähigen, KI-Technologien verantwortungsvoll zu nutzen und ethische sowie qualitativ hochwertige Standards für den Einsatz von AI zu etablieren.

Das Projekt wird von **winnovation** geleitet (Gertraud Leimüller und Brigitte Ömer-Rieder) und im Konsortium mit **leiwand.ai**, **TU Wien** und der **Austria Presse Agentur (APA)** umgesetzt.

Die ALAIT-Dossiers, Branchen-Reports und weitere Informationen sind auf der Projekthomepage abrufbar: <https://science.apa.at/project/alait/>

## Zum Prozess

Der ALAIT Workshop am 05. März 2026 startete mit zwei inhaltlichen Impulsen: Gertraud Leimüller, die Leiterin des Projektkonsortiums, stellte unter dem Titel „Zukunftsweisende Methoden für vertrauenswürdige KI“ das Thema KI-Vertrauen und die Arbeitsweise von ALAIT vor, gefolgt von Michael Roither, Vizerektor für Internationales & Professor für Digital Media und Kommunikation an der Hochschule Burgenland, welcher unter dem Titel „Mehr KI, weniger Vertrauen? Die Sicht der Mediennutzer:innen“ die Glaubwürdigkeit von Medien aus Nutzer:innen-Sicht in den Fokus rückte.

Danach arbeiteten die Teilnehmenden in Kleingruppen zu der Frage, wie KI-Werkzeuge in Redaktionen verantwortungsvoll eingesetzt werden können, ohne einen Vertrauensverlust bei Medienkonsument:innen zu riskieren. Um konkrete Empfehlungen für Medienunternehmen ausarbeiten zu können, kam dabei das folgende fiktive Negativszenario zum Einsatz, das einen überbordenden KI-Einsatz in einem Nachrichtenmedien mit negativen Reaktionen des Publikums und von Expertenseite beschreibt (stark gekürzte Fassung):

*„Giftiges Trinkwasser im Marchfeld?“*

*Die Online-Redaktion eines etablierten Medienunternehmens erarbeitet einen investigativen Beitrag über eine gesundheitsschädliche Pestizidbelastung im March-Grenzgebiet zwischen Österreich und der Slowakei.*

*Ausgangspunkt war ein Hinweis einer lokalen Bürgerinitiative aus Hohenau an der March: In mehreren Brunnen seien erhöhte Schadstoffwerte gemessen worden, die möglicherweise aus der intensiven Landwirtschaft auf beiden Seiten der Grenze stammen.*

*Bei der Bearbeitung des Themas setzt die Redaktion stark auf KI-gestützte Recherche- und Produktionswerkzeuge. Der fertige Artikel wird zudem mittels KI-Unterstützung in verschiedene Medienprodukte umgewandelt, um mit Videos und Audio-Formaten multiple Kanäle, insbesondere Social-Media-Kanäle, bespielen zu können.*

*Nur wenige Stunden nach der Veröffentlichung des Artikels und der Posts auf Social Media melden sich Leser:innen und Fachexpert:innen zu Wort, die die Angaben in den veröffentlichten Artikeln und Posts hinterfragen und korrigieren.*

**Autor:innen dieses Reports:** Basierend auf den Ergebnissen des ALAIT Labors vom 5. März 2026. Gertraud Leimüller (winnovation), Silvia Wasserbacher-Schwarzer (leiwand.ai), Verena Krawarik (APA).

Die Inhalte dieses Reports entsprechen dem aktuellen Stand der Technik und wurden sorgfältig nach wissenschaftlichen Kriterien erstellt. Sie dienen jedoch nicht als rechtsverbindliche Auskunft oder Beratung.

Juni 2026

### Impressum

MedieninhaberIn und HerausgeberIn: winnovation  
consulting gmbh  
Linke Wienzeile 125/ 5  
1060 Wien

Gefördert durch

 Bundesministerium  
Innovation, Mobilität  
und Infrastruktur



## Weiterführende Quellen:

- i. European Broadcasting Union (2025): EBU News Report 2025: Leading Newsrooms in the Age of GenerativeAI [https://www.ebu.ch/files/live/sites/ebu/files/Publications/Reports/open/EBU\\_News\\_report\\_2025\\_Leading%20Newsrooms\\_AI.pdf](https://www.ebu.ch/files/live/sites/ebu/files/Publications/Reports/open/EBU_News_report_2025_Leading%20Newsrooms_AI.pdf)
- ii. Rundfunk und Telekom Regulierungs-GmbH (2025): KI und Medienvertrauen 2025: Mediennutzung und Medienvertrauen in Österreich im Spannungsfeld von KI und Sozialen Medien, [https://www.rtr.at/medien/aktuelles/publikationen/Publikationen/Publikationen\\_2025/Studie\\_KI\\_und\\_Medienvertrauen.de.html](https://www.rtr.at/medien/aktuelles/publikationen/Publikationen/Publikationen_2025/Studie_KI_und_Medienvertrauen.de.html)
- iii. Austria Presse Agentur (2025): KI-Kennzeichnung im Textjournalismus. <https://apa.at/whitepaper/ki-kennzeichnung-im-textjournalismus-maerz-2025/>
- iv. Beispiele für KI-Guidelines: ORF <https://der.orf.at/unternehmen/leitbild-werte/ki-guidelines/>, Der Standard <https://about.derstandard.at/redaktion/redaktionelle-ki-richtlinien-der-standard/>;