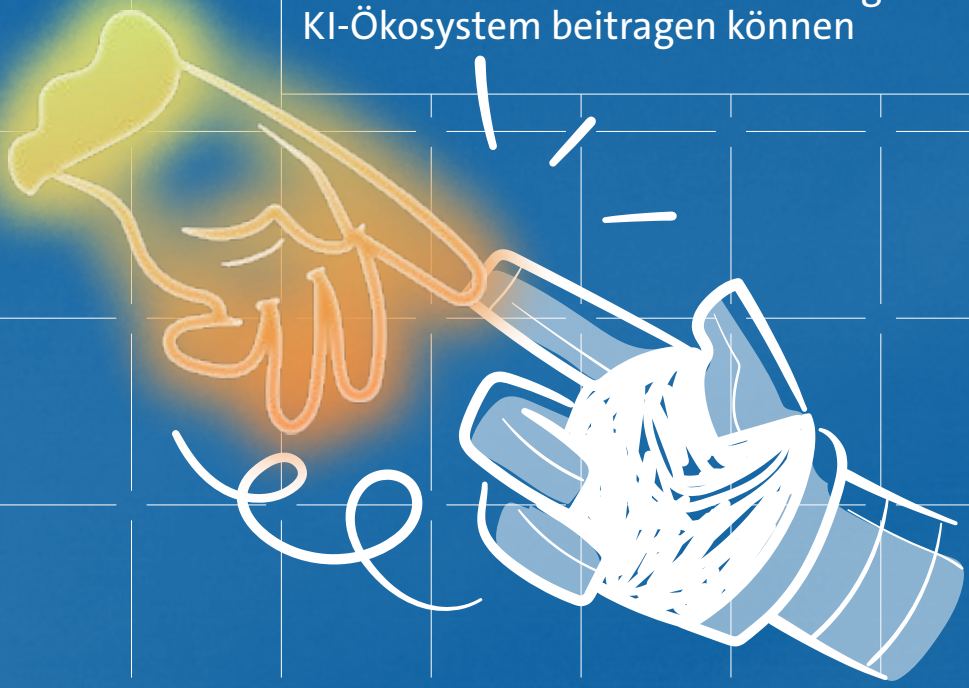
Herausforderung
Künstliche Intelligenz

# Wie Medien im KI-Zeitalter sichtbar bleiben...

...und zu einem vertrauenswürdigen KI-Ökosystem beitragen können



BR AI for Media Network

MEDIA LAB BAYERN

## INHALTSVERZEICHNIS


01 Zusammenfassung ..... 03

02 Über dieses Whitepaper ..... 08

03 KI bringt Nutzende mit den Medien ins Gespräch ..... 12

04 Wo Transformationsbedarf für die Medien besteht ..... 16

05 „Mit einem ereignisbasierten Knowledge Graph kann man Geschichten als Datenstruktur darstellen“ ..... 26

06 Künstliche Intelligenz braucht echte Menschen ..... 32

07 Fünf Erfahrungen aus der Arbeit mit KI im Media Lab Bayern ..... 34

08 Ideen für Infrastruktur-Projekte und nutzerzentrierte Produkte ..... 36

09 Ausblick ..... 42

10 Glossar ..... 46

# 01 Zusammenfassung

Mehr Informationen:
**Blogartikel zur Fachtagung**

Das AI for Media Network widmet sich der Frage, wie künstliche Intelligenz den Journalismus besser machen kann. Am 13. und 14. November 2024 veranstalteten wir in und mit der Akademie für Politische Bildung in Tutzing eine Fachtagung zum Thema „AI-Mediated Media Environment. Boosting Media for an AI-Driven Information Ecosystem.“

Es ging um die Frage, wie sich Qualitätsmedien in einer Zeit, in der Medieninhalte immer stärker von bzw. mit Hilfe von KI erstellt, kuratiert und verbreitet werden, behaupten und wie sie sich organisatorisch und technisch auf diese Disruption einstellen können.

Dazu luden wir 60 KI-Expert*innen aus verschiedenen Medienunternehmen (privater und öffentlicher Rundfunk sowie Verlage) ein, die in ihren Häusern für Produktmanagement, Software-Entwicklung, UX-Design oder die gesamte KI-Strategie zuständig sind. Um zudem wissenschaftliche KI-Expertise einfließen zu lassen, nahmen auch Forschende an dem englischsprachigen Symposium teil.

Der Fokus der Tagung lag darauf, Szenarien für verschiedene Bereiche zu entwickeln, wie Medien auf durch (generative) KI ausgelöste Umwälzungen in den nächsten Jahren reagieren müssen, um das Publikum weiter oder sogar besser als bisher mit qualitativ hochwertigen Inhalten zu erreichen.

# Aus den Vorträgen, Arbeitsgruppen und Diskussionen lassen sich folgende Einschätzungen und Empfehlungen ableiten:

## 01

### Medien müssen zuhören lernen

Um bestehende Nutzer*innen zu halten und diejenigen, die sich von Medien abgewendet haben, zurückzugewinnen, sollten Medienunternehmen mehr Energie darauf verwenden, (Informations-)Bedürfnisse der Nutzenden zu verstehen und deren Perspektiven in ihre Produkte einfließen zu lassen. KI kann Medien bei der Analyse des Nutzerverhaltens und der Nutzerbedürfnisse unterstützen.

## 02

### Personalisierte Inhalte liefern

Mediennutzende werden zunehmend mit Inhalten interagieren und explizit mitteilen, welche Art von Inhalt in welcher Sprache, in welchem Stil und in welchem Medienformat sie erwarten. Generative KI-Tools liefern dann diese maßgeschneiderten Inhalte. In einem nächsten Schritt werden KI-Anwendungen auch in der Lage sein, die Benutzeroberflächen an die gewünschte Nutzungsform anzupassen, zum Beispiel durch multimodale Interaktion (Sprache, Berührung, Gesten, Blicke).

## 03

### Inhalte für Datenbanken aufbereiten

Damit personalisierte Inhalte für die Nutzer*innen erstellt werden können, müssen Medienhäuser zuerst ihre Inhalte systematisch in Datenbanken organisieren. Das können Knowledge Graphen oder Vektordatenbanken sein. Künstliche Intelligenz spielt eine wichtige Rolle bei der Strukturierung und Speicherung dieser Inhalte.

## 04

### Vertrauenswürdige Datenmarktplätze

Medien haben jede Menge selbst recherchierter und verifizierter Inhalte, die teilweise sogar Jahrzehnte zurückreichen. Wenn verschiedene Medienhäuser ihre Qualitätsinhalte bündeln und zugänglich machen, entsteht ein großer Mehrwert. Dies könnte zu neuen Datenmarktplätzen führen. Für den Zugang zu solchen „Trusted Content Pools“ könnten Medien von Sprachmodell-Anbietern Gegenleistungen verlangen und Bedingungen an die Nutzung ihrer Daten stellen.

# Diese Entwicklungen werfen Fragen von großer Tragweite auf:

1. Wie sehen die technischen Standards aus, zum Beispiel für verifizierte Informationen – oder für Knowledge Graphen?

2. Welche finanziellen und personellen Ressourcen brauchen Medienhäuser, um ihre technische Infrastruktur umzustellen und Inhalte in Daten umzuwandeln?

3. Wie können Allianzen für „Trusted Content Pools“ gebildet werden?

4. Wie kann man den Zugang zu solchen Content Pools möglicherweise monetarisieren?

5. Sollen die Inhalte von Medienhäusern zum Training von kommerziellen Sprachmodellen verwendet werden – oder nicht?

6. Wie bleiben Medienmarken als Quelle sichtbar – und wie können Sie mit ihren Inhalten weiter einen wichtigen Beitrag zur Meinungsbildung liefern?

**Auf diese Fragen muss die Medienbranche in den nächsten Monaten und Jahren Antworten finden.**

BR AI for Media Network

# 02 Über dieses Whitepaper

# Dieses Whitepaper ist das Ergebnis einer Fachtagung des AI for Media Networks.

## Das AI for Media Network

Das AI for Media Network wurde 2024 vom Bayerischen Rundfunk gegründet. In dem Netzwerk tauschen KI-Praktiker*innen aus deutschsprachigen Medien ihr Wissen aus und demonstrieren KI-Use-Cases aus ihren Häusern.

Das AI for Media Network arbeitet interdisziplinär und umfasst inzwischen mehr als 600 Mitglieder. Der Schwerpunkt liegt auf Mitarbeitenden von Medienhäusern, die in den Bereichen Redaktion, Produktentwicklung, Produktmanagement, Software-Entwicklung, Machine Learning, UX-Design oder KI-Strategie tätig sind, auch auf Leitungs- und Geschäftsführungsebene. Zu den im Netzwerk vertretenen Medien gehören öffentlich-rechtliche und private Rundfunkanbieter, Verlage und Online-Publisher.

Darüber hinaus sind auch Mitarbeitende von KI-Industrie, Startups und IT-Dienstleistern sowie Wissenschaftler*innen, die zu verschiedenen Aspekten von künstlicher Intelligenz forschen, im Netzwerk vertreten. Die meisten Mitglieder kommen aus Deutschland, einige aus Österreich und der Schweiz, auch aus anderen europäischen Ländern gibt es vereinzelt Mitglieder.

Das Netzwerk trifft sich etwa vier Mal im Jahr zu Meetups, die bislang vom Bayerischen Rundfunk und von Ippen Digital ausgerichtet werden.

Mehr Informationen:
**aiformedia.network**

## Die Fachtagung

Normalerweise stehen bei den Meetups konkrete Produkte oder Workflows im Mittelpunkt, in denen KI eingesetzt wird. Bei der Fachtagung „AI-Mediated Media Environment," die wir am 13. und 14. November 2024 in und mit der Akademie für Politische Bildung in Tutzing mit 60 Teilnehmenden aus Journalismus, Industrie und Wissenschaft durchführten, ging es dagegen um einen Blick in die Zukunft: Welche Herausforderungen bringt das KI-Informationszeitalter für die Medienbranche – und wie sollte sie darauf reagieren?

Die zwei Tage umfassten einen Mix aus Vorträgen von ausgewiesenen KI-Expert*innen und gemeinsamer Arbeit zu speziellen Themen wie Lokaljournalismus, Zukunft der Suche oder KI-gesteuerte User Interfaces. Jede Arbeitsgruppe entwarf ein Szenario, wie ihr Thema in zwei bis fünf Jahren aussehen könnte.

Es handelt sich nicht um eine empirische Arbeit, sondern um die Einschätzungen von Expert*innen, wie sich die Produktion, Distribution und der Konsum von journalistischen Formaten in den kommenden Jahren entwickeln könnten, verbunden mit Ideen und Empfehlungen für die Entwicklung dieser Formate und Workflows.

Um einen möglichst offenen Austausch zu erleichtern, fand die Fachtagung unter Chatham House Rules statt: Alle Teilnehmenden dürfen die erhaltenen Informationen frei verwenden, jedoch ohne preiszugeben, von wem sie stammen. Diesem Prinzip folgt auch dieses Whitepaper, weshalb den Texten keine einzelnen Autor*innen zugeordnet sind.

# Das Media Lab Bayern unterstützt Talente, Startups und Medienhäuser dabei, innovative Ideen zu finden und umzusetzen.

## Kooperation mit dem Media Lab Bayern

Dieses Whitepaper wird in Zusammenarbeit mit dem Media Lab Bayern veröffentlicht. Das Media Lab Bayern ist eine Initiative der Medien.Bayern GmbH und wird von der Bayerischen Staatskanzlei und der Bayerischen Landeszentrale für neue Medien (BLM) gefördert. Die Medien.Bayern GmbH ist eine Tochter der BLM und fördert Innovation am Medienstandort Bayern.

Weder die Bayerische Staatskanzlei noch die BLM haben Einfluss auf die redaktionellen Inhalte dieser Publikation. Alle Inhalte stammen von Mitgliedern des AI for Media Network und des Media Lab Bayern. Das Media Lab Bayern trägt die Kosten für Layout und Druck. ◆

Mehr Informationen:
**media-lab.de**

# 03 KI bringt Nutzende mit den Medien ins Gespräch

Spätestens mit der Veröffentlichung von ChatGPT im November 2022 ist künstliche Intelligenz im Alltag angekommen. Nutzende können Fragen oder Aufgaben stellen, der Chatbot liefert die Antwort. Die User*innen können die gewünschten Informationen in einem Dialog erfragen, der so lange andauert, bis die Information möglichst passend auf das Informationsbedürfnis zugeschnitten ist.

Auf der Fachtagung bestand Konsens, dass die Zukunft der Suche bzw. Informationsbeschaffung dialogisch sein wird. Dass Fragen direkt — sei es per Text oder gleich via Stimme — eingegeben werden, wird ein relevantes Mediennutzungsmuster der Zukunft sein. Mediennutzende fragen nicht nur nach bestimmten Themen, sondern geben auch an, in welchem Stil, in welcher Sprache und in welchem Format sie die Informationen erhalten möchten. Viele Medien haben inzwischen auf Sprachmodellen basierende Anwendungen geschaffen, die es erlauben, mit ausgewählten eigenen Inhalten zu interagieren.

## Nutzer*innen können mit ausgewählten Medieninhalten interagieren

Mehr Informationen:

Fragen Sie Zeit Online (Paywall)

Donald Trump: 2024 TIME Person of the Year (Time Magazine)

Das Tool „Fragen Sie Zeit Online“ beantwortet die Fragen von Leser*innen zu aktuellen Ereignissen und nutzt dabei die Zeit-Online-Artikel der vergangenen 30 Tage als Quelle. Viel Beachtung fand jüngst die Reportage des US-Magazins Time über Donald Trump und seinen Sieg bei der Präsidentschaftswahl 2024. Der Text lässt sich via KI zusammenfassen, als Audio hören und befragen, sowohl mit Text- als auch mit Voice-Chat.

## Aus Suchmaschinen werden Antwortmaschinen

Auch die Art, wie Suchmaschinen genutzt werden, verändert sich durch KI. Bei klassischen Suchmaschinen gibt man Keywords ein und erhält eine Liste mit passenden Websites. Bei KI-gestützten Suchmaschinen können Nutzende eine Frage stellen.

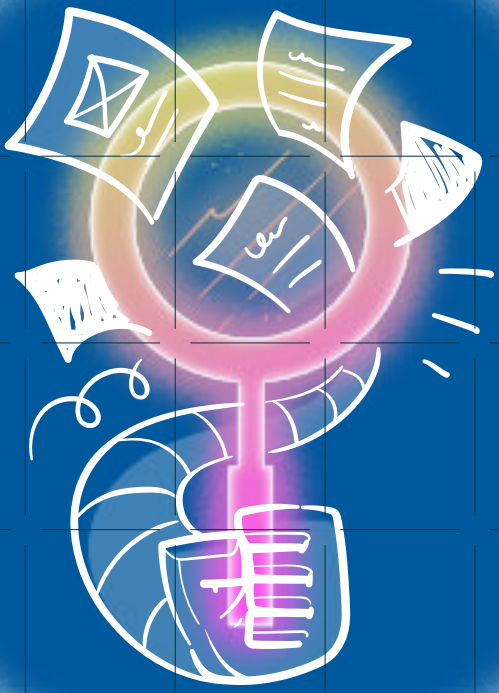
Systeme wie Perplexity generieren passgenaue Antworten auf diese Fragen, indem sie relevante Webinhalte zu einem neuen Text zusammenfassen und darin Quellen referenzieren. Auch Google und OpenAI bieten inzwischen mit den AI Overviews respektive Search GPT ähnliche Produkte an, ebenso das chinesische Startup DeepSeek.

# Medienkonsument*innen fragen nach Inhalten, die zu ihrer Nutzungssituation passen

Nutzungskontexte werden in Zukunft eine noch größere Rolle spielen. Die Medienkonsument*innen fragen vermehrt nach Inhalten, die ihrer Nutzungssituation angepasst sind: Haben sie viel oder wenig Zeit, sind sie zu Hause oder unterwegs, wollen sie lesen, hören oder sehen? Und wie ausführlich soll der Content sein: kurz, mittellang oder tiefgehend und ausführlich? KI-gestützte Anwendungen liefern oder generieren die passenden Inhalte.

Noch einen Schritt weitergedacht: Nutzende müssen ihre Präferenzen nicht mehr äußern, sondern ihre Endgeräte erkennen die Nutzungssituation automatisch und empfehlen oder generieren passende Inhalte.

Auf Basis dieser Entwicklungen haben die Arbeitsgruppen in Tutzing zwei Bereiche mit großem Transformationsbedarf definiert: eigene Medieninhalte als Grundlage für KI-Modelle bereitzustellen und KI als Werkzeug einzusetzen, um die Bedürfnisse des Publikums zielgenauer zu befriedigen. ◆



# 04 Wo Transformations-bedarf für die Medien besteht

## Aus Content werden Datenbanken

Medien haben eine riesige Menge an verifizierten und vertrauenswürdigen Inhalten. Wenn Medienhäuser es erlauben, können KI-basierte Suchmaschinen ihre Inhalte auswerten und in den KI-generierten Antworten als Quelle ausweisen. Der Spiegel-Verlag hat etwa eine Kooperationsvereinbarung mit Perplexity geschlossen. Der Vertrag sieht vor, dass die Spiegel-Inhalte dort sichtbarer gemacht werden. Der Spiegel entscheidet selbst, welche Inhalte er Perplexity zur Verfügung stellt. Im Gegenzug erhält der Verlag einen Anteil an den mit seinen Inhalten erwirtschafteten Werbeeinnahmen.

Und gleichzeitig ist diese neue Nutzung von Medieninhalten durch Techunternehmen Anlass zur Sorge für Medienunternehmen: Denn durch generierte Zusammenfassungen werden die Inhalte der Medienhäuser für viele Nutzende nur zum zweiten Klick, der in immer weniger Fällen notwendig sein wird. Solche Mechanismen greifen direkt die aktuell vorherrschenden Geschäftsmodelle der Medienindustrie an, die entweder durch Abonnements oder durch das Zeigen von Werbung auf eigenen Plattformen auf die direkte Verbindung zu den Nutzenden angewiesen sind. Noch offen ist, welche neuen Geschäftsmodelle sich aus dieser Art der Datennutzung ergeben könnten.

**Medienhäuser könnten ihre Inhalte mit Nutzungsbedingungen versehen**

Eines der meistdiskutierten Themen in Tutzing waren daher neue Datenmarktplätze. Damit gemeint sind Datenbanken, in denen Daten gesammelt, mit Nutzungsbedingungen versehen und KI-Anbietern möglicherweise im Austausch für Gegenleistungen zum Training oder zur Anreicherung von KI-generierten Antworten angeboten werden. So könnte auch überwacht werden, welche Daten von KI-Plattformen genutzt werden und wie häufig welche Systeme auf die Inhalte zugreifen – wertvolle Einblicke in die Nutzung ihres Contents für Medienhäuser, die den direkten Kontakt zu ihren Nutzenden immer mehr einbüßen.

Eine vieldiskutierte Möglichkeit waren sogenannte „**Trusted Content Pools**“, zu denen sich Medienhäuser zusammenschließen könnten, um dort ihre Inhalte aufzubereiten und für Sprachmodelle nutzbar zu machen. Auf diesem Weg könnten sie sowohl inhaltliche und technologische Kontrolle darüber erlangen, welche Inhalte von KI-Plattformen für welche Anwendungen benutzt werden. Das könnte das reine Erwähnen von Nutzungsverboten in den robots.txt-Dateien – die derzeit gängige Methode, KI-Anbietern die Nutzung des eigenen Contents zu untersagen – ergänzen oder gar ersetzen.

Diese Lösung ist nicht mehr als ein maschinenlesbares Schild an der digitalen Tür von Medienunternehmen – sie ist keine technische Blockade oder eine Möglichkeit zu monitoren, wer die Inhalte wofür nutzt. Auch gibt es derzeit keine inhaltlichen Standards für die Nutzung von Inhalten. Diskutiert werden dazu in der Branche Aspekte wie Qualitätsmonitoring der generierten Inhalte oder Rückverlinkung auf den eigenen Content. Trusted Content Pools bieten die Chance, diese inhaltlichen und technischen Standards branchenübergreifend zu entwickeln.

# Trusted Content Pools könnten aktuelle Beiträge, Archivmaterial und auch Rohdaten enthalten

Geteiltes Wissen technisch zugänglich zu machen, könnte zudem als Innovationsmotor dienen für die Entwicklung neuer Medienprodukte – sowohl auf Plattformen als auch bei Medienunternehmen selbst.

Archive könnten an solche Content Pools ebenfalls angeschlossen sein. In Tutzing wurde beispielsweise die Idee eines gemeinsamen Archivs der öffentlich-rechtlichen Rundfunkanstalten diskutiert **(siehe S. 36)**. Denkbar ist auch, nicht nur fertige Artikel und Beiträge, sondern auch Rohdaten, zum Beispiel Recherchematerial und strukturierte Daten, zu solchen Datensammlungen hinzuzufügen.

Derzeit werden in der Branche eher Einzeldeals zwischen Medienhäusern und KI-Plattformen geschlossen. Ein sehr früher Deal in Deutschland war der des Springer-Verlags Ende 2023 mit OpenAI, der es dem Techunternehmen erlaubt, Inhalte von Springer-Marken zum Training zu verwenden. Dafür verlinkt ChatGPT in seinen Antworten auf Artikel aus dem Hause Springer.

Neben der impliziten Methode, Wissen durch Embeddings zu repräsentieren, wurde in Tutzing auch eine explizite Methode der Wissensrepräsentation diskutiert: der **Knowledge Graph**. In einem solchen Graphen wird Wissen in Form eines Netzwerks aus Knoten (Nodes) und Kanten (Edges) strukturiert dargestellt. Knoten repräsentieren Entitäten wie Personen, Orte, Objekte oder Konzepte, während die Kanten die Beziehungen zwischen diesen Entitäten definieren.

Ein großer Vorteil von Knowledge Graphen ist, dass sie Fakten übersichtlich speichern und Zusammenhänge zwischen Informationen sichtbar machen. Um zum Beispiel Nachrichteninhalte in einem solchen Graphen abzubilden, müssten die wichtigsten Ereignisse, Personen und Orte erfasst und miteinander verknüpft werden. Dies kann manuell, automatisiert oder mit Hilfe von künstlicher Intelligenz erfolgen.

Web-Content, Vektordatenbanken oder Knowledge Graphens können allesamt als Datenquellen für **Retrieval-Augmented Generation (RAG)** dienen. RAG-Systeme durchsuchen diese Daten gezielt, um relevante Fakten abzurufen und damit fundierte Antworten auf die Prompts der Nutzenden zu generieren.

## Publikumsbedürfnisse erfüllen – mit Hilfe von KI

Die Entwicklung künstlicher Intelligenz verspricht, die Art und Weise, wie Nachrichten produziert, verbreitet und konsumiert werden, grundlegend zu verändern. Inmitten dieses Wandels bleibt eine zentrale Frage bestehen: Wie kann der Journalismus seinen öffentlichen Auftrag erfüllen und gleichzeitig den sich wandelnden Bedürfnissen seines Publikums gerecht werden?

**Zuhören mit KI**

Durch künstliche Intelligenz können journalistische Angebote insgesamt zugänglicher und inklusiver werden. KI kann helfen, Nachrichten individuell anzupassen und relevante Inhalte für verschiedene Gruppen von Nutzenden bereitzustellen. Durch die Auswahl zielgruppenspezifischer Medienformate – wie Audio, Video, Bild, Text oder deren Kombinationen – lassen sich Medieninhalte gezielt an die Bedürfnisse der Nutzenden anpassen.

**Medien müssen die Bedürfnisse ihrer Nutzenden verstehen, um ihnen personalisierte Inhalte anbieten zu können.**

Mehr Informationen:
**Dein Argument (BR)**

Die Erstellung personalisierter Inhalte wird umso besser funktionieren, je stärker Medien ihren Nutzenden zuhören und ihre Bedürfnisse verstehen. Journalist*innen sollten Zuhören als Teil journalistischer Produkte verstehen und nicht nur als deren Voraussetzung. Aktuell setzt der Bayerische Rundfunk bereits KI zum Zuhören ein. So werden die Nutzer*innen-Feedbacks in den Kommentarspalten durch eine KI ausgewertet. In den Kommentarspalten von BR24 sucht eine KI nach User*innen-Kommentaren, die sich mit Fragen, Argumenten oder Hinweisen direkt an die Redaktion wenden, und leitet sie ans Team weiter.

Wenn eine Nutzenden-Perspektive in einem BR24-Artikel aufgenommen wird, kennzeichnet die Redaktion das mit einem Hinweis auf „Dein Argument“. Hier wird eine Form des offenen, wenn auch zeitversetzten Dialogs etabliert, der Vertrauen schaffen sowie Zuhören und KI konstruktiv verbinden soll.

In Tutzing hat die Gruppe „Being Part of the Stories“ überlegt, inwiefern es denkbar ist, Menschen in den sozialen Medien mit einem KI-getriebenen journalistischen Angebot zu begegnen, das einen geschützten Dialog ermöglicht. Dabei handelt es sich um einen 1:1-Dialog in Form persönlicher Nachrichten, die Medienhäuser mit User*innen direkt und nicht öffentlich sichtbar in Kontakt treten lassen.

Die Idee ist, Fragen und Themen insbesondere junger und medienferner Menschen über Direktnachrichten oder Chatbots aufzunehmen und diese im geschützten Raum durch eine KI beantworten zu lassen. Zudem lässt sich Social Listening nutzen, um automatisiert Themenkomplexe in den sozialen Netzwerken zu identifizieren. So könnten Medienmarken mit medienfernen Zielgruppen in Kontakt kommen, im Idealfall deren Bedürfnisse erfüllen und ihre eigene journalistische Themen- und Formatentwicklung darauf ausrichten.

**Menschen erreichen**

Besser zuzuhören kann ein Weg sein, um medienferne Zielgruppen zu erreichen. Es bleibt jedoch die Frage, wie Medien die Bedürfnisse, Fragen und Perspektiven von Menschen kennenlernen, wenn diese Medien gar nicht nutzen. Hier müssen Medien neue Wege finden, um mit Menschen ins Gespräch zu kommen. Ein Beispiel ist die Dialogbank der Innovationsagentur Tactile News, die bereits bei verschiedenen Lokalzeitungen im Einsatz ist. Dabei handelt es sich um eine mobile Sitzbank, die mit einem Telefonhörer ausgestattet ist, der wiederum mit einer KI verbunden ist.

Passant*innen können sich mit dieser KI unterhalten, sie hört zu, stellt aber auch selbst Fragen. Diese Gespräche zwischen Mensch und Maschine werden aufgezeichnet. Lokalredaktionen können so die Menschen vor Ort und ihre Themen besser kennenlernen, auf Missstände aufmerksam werden und andere lokale Geschichten entdecken. Das hilft ihnen, ihre journalistische Rolle vor Ort besser ausüben zu können. Sofern die Datenlage aussagekräftig ist, können diese Impulse auch dafür verwendet werden, Lokalberichterstattung zu personalisieren. „The NAighbor“, eine Projektidee der Tagung, vertieft diesen Gedanken und stellt sich eine KI-getriebene hyperpersonalisierte Plattform für lokale Berichterstattung vor, in der Nutzende die Informationen, Nachrichten und Formate erhalten, die sie wünschen.

Mehr Informationen:
**Dialogbank (Tactile.News)**

KI-basierte Wege, Menschen zuzuhören und ihre Bedürfnisse kennenzulernen, sind eine Möglichkeit, Perspektiven von Menschen, die aktuell medienfern und wenig repräsentiert sind, wertschätzend medial abzubilden. Im Idealfall gelingt es Medien damit, neue Beziehungen herzustellen und Menschen (wieder) zu erreichen.

**Barrieren abbauen: Journalismus wird inklusiver**
Neben dem Gefühl, sich in den Medien nicht wiederzufinden, können auch sprachliche, kognitive und psychische Barrieren dazu führen, dass Menschen journalistische Inhalte meiden oder diese als zu herausfordernd empfinden. Rund 17 Prozent der Bevölkerung benötigen laut Angaben der Bundesregierung vereinfachte oder leicht verständliche Sprache, um Informationen zu verstehen.

# Generative KI kann Inhalte für alle Sprachniveaus erstellen

KI ermöglicht es, Sprachbarrieren zu überwinden und Inhalte auf verschiedenen Sprachniveaus zu erstellen. Vier von zehn Deutschen fühlen sich von der Vielzahl von Informationen überfordert und vermeiden aktiv Nachrichten.

So können auch Menschen erreicht werden, die sonst aufgrund sprachlicher Hürden ausgeschlossen wären. Zudem kann KI dabei helfen, Barrieren in der Zugänglichkeit abzubauen und Inhalte barrierearm zu gestalten. Dadurch steigt die Wahrscheinlichkeit, dass Menschen mit den für sie relevanten Informationen in Kontakt kommen.

Wenn Bürger*innen, die bislang wenig Zugang zu oder Interesse an Nachrichtenformaten hatten, besser informiert werden, können sie aktiver am gesellschaftlichen Dialog teilnehmen. Generative KI kann die Voraussetzungen dafür schaffen, indem sie Inhalte so aufbereitet, dass die Nutzer*innen sie verstehen und in der von ihnen gewünschten Form konsumieren können. ◆

05

# „Mit einem ereignis-basierten Knowledge Graph kann man Geschichten als Daten-struktur darstellen“

Ein Interview mit David Caswell von Bernd Oswald

**David, bei unserem Symposium „AI-Mediated Media Environment“ in Tutzing war eine zentrale Erkenntnis, dass Qualitätsmedien Inhalte produzieren sollten, die genau auf die Interessen, Formatvorlieben und Nutzungssituationen des Publikums zugeschnitten sind. Welches Potenzial hat KI, um solche personalisierten Nachrichtenerlebnisse zu schaffen?**

Es geht darum, dass die Nutzer*innen das bekommen, was sie wollen. Wenn jemand einen 800-Wörter-Artikel möchte, der aussieht, als wäre er 1997 von einem traditionellen Journalisten geschrieben worden, dann bekommt er das. Wenn eine Nutzerin ein neues, verrücktes TikTok-Video möchte, dann bekommt sie das.

Medienproduzenten tun dies bereits bis zu einem gewissen Grad. Sie erstellen Zusammenfassungen aus Textartikeln, Audioerlebnisse aus Textartikeln, in einigen Fällen Chatbot-Angebote mit Textarchiven als Quelle und viele andere Produkte. Wenn man das auf die Spitze treibt, wird die Herausforderung sein, nicht nur drei oder vier oder fünf verschiedene Formen von personalisiertem Zugang zu produzieren, sondern 30 oder 50: verschiedene Schreibstile, unterschiedliche Mengen an Kontext, verschiedene Sprachen, ganz zu schweigen von den visuellen, hörbaren oder interaktiven Versionen.

Es ist gut, wenn die Inhalte auf die Nutzenden abgestimmt sind. Doch dabei stellt sich die Frage: In welcher Form, in welchem Basisformat von Nachrichten werden Informationen gespeichert, um das zu ermöglichen? Auch vor dem Hintergrund einer zweiten Entwicklung, die KI auf das Informations-Ökosystem hat: der Möglichkeit, große Mengen an Quellen wie PDFs, Regierungsdaten, Pressemitteilungen, Politikerreden im Parlament, Transkripte etc. zu durchforsten.

→ **David Caswell ist Innovationsberater, der sich auf den Einsatz von KI im Journalismus spezialisiert hat.**

Er war zuvor Executive Product Manager bei den BBC News Labs und hatte leitende Positionen bei KI-Initiativen von Yahoo!, Tribune Publishing und der Los Angeles Times inne.

Er veröffentlicht wissenschaftliche Texte zu automatisierten Formen des Journalismus und war Mitautor des im August 2024 erschienenen Berichts „AI in Journalism Futures“.

Er nahm am Symposium „AI-Mediated Media Environment“ teil, bei dem er eine Arbeitsgruppe zum Thema „Unit of News“ moderierte. Dort ging es um die Frage, in welcher Form Informationen und Fakten gespeichert werden sollen.

Mehr Informationen:
**AI in Journalism Futures**

KI kann all das lesen, analysieren, zusammenfassen, Verbindungen herstellen und Geschichten darin finden. Die Herausforderung bleibt die Gleiche: Wie werden diese Informationen gespeichert, was ist das neue Basisformat der Nachricht?

## Und wie lautet Deine Empfehlung dafür?

Wenn man den Nutzenden geben möchte, was sie wollen, stellt das bestimmte Anforderungen an die Informationsarchitektur. Sie muss in der Lage sein, Informationen für alle Formen der Nutzung bereitzustellen. In unserer Arbeitsgruppe „Unit of news“ haben wir mehrere Alternativen diskutiert.

Zum Beispiel einen endlosen Live-Blog, der aus kleinen Informationshäppchen, kurzen Snippets besteht: für jedes neue Ereignis ein Update im Live-Blog. Oder den Artikel einfach als eine Art annotierte Basis zu verwenden. Zuerst schreibt man den Artikel und ergänzt ihn dann mit Anmerkungen für neue Nachrichten. Im Endeffekt kamen wir zu dem Schluss, dass ein Knowledge Graph die beste Alternative ist.

## Kannst Du kurz beschreiben, wie ein solcher Knowledge Graph funktionieren würde?

Zuerst muss man sich klarmachen, dass es dabei um eine Backend-Aufgabe geht, um die Frage, wie Wissen repräsentiert wird. Ein Knowledge Graph besteht aus Knoten und Kanten.

Die Knoten sind im Wesentlichen Ereignisse. Nachrichten basieren normalerweise auf Ereignissen. Also werden Ereignisse zu den Knoten. Die Beziehung zwischen den Ereignissen sind die Kanten in diesem Graphen.

Der Wissensgraph kann nicht nur Nachrichten darstellen, sondern er passt auch zur Art und Weise, wie Menschen Informationen und Nachrichten konsumieren, nämlich in Form von Geschichten. Auf diese Weise lassen sich Netzwerke von Ereignissen aufbauen, mit denen der Knowledge Graph Erzählungen oder Geschichten als Datenstruktur darstellt.

## Wie würden Informationen bzw. Ereignisse in einen solchen Wissensgraphen eingetragen werden?

Informationen gelangen durch jede erdenkliche Weise, auf die man Informationen sammeln kann, in den Wissensgraphen. Hauptsächlich mit Hilfe von KI-Anwendungen, die Inhalte lesen, daraus Informationen extrahieren und diese in den Wissensgraphen einfügen.

Zum Teil könnte es auch durch regelbasierte Programme geschehen, die Informationen in natürlicher Sprache verarbeiten können. Eine weitere wichtige Quelle für Informationen sind Menschen: Journalist*innen können Ereignisse direkt eingeben, zum Beispiel über Content-Management-Systeme. Der Knowledge Graph selbst sollte dabei neutral gegenüber den Quellen sein. Er sollte in der Lage sein, journalistische Informationen aus jeder sinnvollen und verlässlichen Quelle zu akzeptieren.

## Sollte jedes Medienunternehmen seinen eigenen Wissensgraphen aufbauen? Oder könnte es einen großen Wissensgraphen geben, der vielen Medien zugänglich ist?

Einerseits besteht ein großes Risiko, wenn eine Mega-Organisation den Wissensgraphen aller journalistischen Nachrichten für den gesamten Planeten kontrollieren würde. Das wäre wahrscheinlich keine gute Idee. Andererseits ist es auch keine gute Idee, wenn jeder das Gleiche aufbaut.

Hier kommt der Netzwerkeffekt ins Spiel, also der Effekt, dass etwas nützlicher wird, je mehr Menschen es verwenden. Und umso nützlicher es wird, desto mehr Menschen nutzen es und so weiter und so fort. Wir haben das bei sozialen Medien gesehen. Hier ist es dasselbe. Wenn jeder seinen eigenen kleinen Knowledge Graph hat, ist das meiste Wissen der Welt nicht darin enthalten. Wenn aber in einem großen Knowledge Graph alles miteinander verbunden ist und zugänglich gemacht werden kann, ist die Summe viel mehr wert als ihre Teile.

Wenn eine vertrauenswürdige Organisation wie Wikipedia oder öffentlich-rechtliche Medien einen solchen Wissensgraphen entwickeln würden, wäre das kein Problem. Dann gäbe es wahrscheinlich Netzwerkeffekte. Der Wissensgraph würde zu einem Standard werden, der allen offen steht, anstatt nur einer einzigen Organisation.

# Medien werden 20.000 statt 200 Geschichten am Tag produzieren

**Wie werden Nachrichtenproduktion und -konsum in fünf Jahren aussehen, unter Berücksichtigung aller Möglichkeiten, die generative KI mit sich bringt?**

Meine Vermutung ist, dass Medien unglaublich viel mehr Nachrichten produzieren werden. Statt 200 Geschichten am Tag werden sie 20.000 oder sogar 200.000 Geschichten am Tag produzieren.

Medien werden versuchen, relevante Geschichten individuell für ihre Nutzer*innen aufzubereiten. Das geschieht im Wesentlichen durch zwei Arten von Personalisierung: zum einen durch die passgenaue Empfehlung relevanter Artikel, zum anderen durch die fortschreitende Personalisierung des Nutzungserlebnisses: Video, Text oder ganz neue Formate – jeder bekommt seine maßgeschneiderte Medienerfahrung. Für das Publikum wird es also viel mehr relevante Nachrichten geben, und sie werden deutlich leichter auffindbar und konsumierbar sein.

Journalist*innen werden mehr mit KI-Systemen arbeiten und redaktionelle Entscheidungen auch auf der Ebene dieser Systeme treffen müssen. Das ist wahrscheinlich die große strukturelle Veränderung. Ich denke, das wird innerhalb der nächsten fünf Jahre passieren, vielleicht sogar früher. ◆

# 06 Künstliche Intelligenz braucht echte Menschen

Wenn das Media Lab Bayern über künstliche Intelligenz im Journalismus reden will, ruft es im AI & Automation Lab des BR an. Oder vernetzt sich auf den Events des AI for Media Network. Bisher sind noch immer echte Menschen ans Telefon gegangen, und Roboter waren auf den meisten Events auch noch nicht zu sehen.

Kaum ein Trend und kaum eine technologische Entwicklung hat den Medienbereich in den letzten Jahren so beeinflusst wie künstliche Intelligenz. Bei kaum einem anderen Thema war der Hype so groß und der potenzielle Impact so gewaltig. Gleichzeitig heißt das natürlich nicht, dass auch alle direkt die besten Lösungen parat haben. Wir müssen weiter an konkreten Wegen arbeiten, wie unsere Medienwelt und unsere Arbeit durch KI konkret besser werden können.

Menschen, die das tun, treffen sich im AI for Media Network. Auch im Media Lab Bayern arbeiten seit bald zehn Jahren Innovator*innen mit KI-Technologien. In bisher zwei Veröffentlichungen hat das Media Lab sie vorgestellt — sie arbeiten an Textautomatisierung, Fakenews-Detektion, Übersetzung und vielem mehr. Ein paar davon teilen ihre Erfahrungen auch in diesem Whitepaper. Hinter allen KI-Ideen, die im Media Lab gefördert wurden, stehen echte Menschen mit ihren Erfahrungen, ihrem Wissen und ihren Vorstellungen davon, wie die Medienbranche der Zukunft aussehen könnte.

Egal wie diese Zukunft wird, es sind Menschen, die sie bauen. Und damit sie das tun können, brauchen sie Austausch und Unterstützung. Dieses Whitepaper und das Netzwerk dahinter sind ein guter Weg, diesen Austausch herzustellen — und dafür, den Kern hinter all der Technologie und der voranschreitenden Technik nicht aus den Augen zu verlieren: die Menschen, die sie bauen und nutzen.

# Fünf Erfahrungen aus der Arbeit mit KI im Media Lab Bayern

Seit 2015 arbeiten Menschen im Media Lab Bayern an KI-Lösungen. Fünf von ihnen berichten, was sie in ihren Projekten gelernt haben und gerne schon vorher gewusst hätten.

**Mehr Projekte findest du unter:**
media-lab.de/de/startups-projekte

## 01

## Es braucht Menschen in der Maschine

„KI ist ein mächtiges Werkzeug, aber ohne den ‚Human-in-the-Loop'-Ansatz bleibt sie unvollständig. Am Ende sind es menschliche Details und Empathie, die den Unterschied machen und Technologie wirklich sinnvoll und nachhaltig gestalten."

*Daidream nutzt KI, um Kinderunterhaltung aktiv und persönlich zu gestalten.*

## 02

## Es müssen nicht immer die großen Kanonen sein

„Die meisten Personen denken bei KI an Modelle, die auf riesigen Datenmengen trainiert werden. Das muss aber nicht so sein! Mit Konzepten wie ‚Few-Shot-Learning', guter Datenaufbereitung und einem passend entworfenen neuronalen Netz können auch schon sehr kleine Datenmengen ausreichen, um ein generatives Netz zu trainieren — ohne leistungsstarken PC."

*Sabine „bleeptrack" Wieluch hat mit „Tinqta" ein eigenes Werkzeug für die Generierung von Illustrationen entwickelt.*

## 03

## Experimente brauchen Selbstvertrauen

„Die Arbeit an KI-Produkten ähnelt der Entwicklung eines neuen Arzneimittels: Es braucht eine Menge guter Ideen und Intuition und noch sehr viel mehr Experimente, um diese Ideen mittels Tests an Probanden zu bestätigen oder zu verwerfen.
Dafür braucht es großes Vertrauen in die eigenen Überlegungen, aber auch die Bereitschaft, sie loszulassen, um Raum für noch bessere Konzepte zu schaffen.“

*Television AI automatisiert den Schnitt und die Verschlagwortung von Nachrichtenvideos.*

## 04

## Infrastruktur schlägt Prompts

„Fehler in generativen Modellen sind schwer vorherzusehen. Das beste Modell und der beste Prompt liefern kein garantiert gutes Ergebnis — egal, wie lange man am Finetuning schleift. Die einzige Lösung bleibt eine externe Kontrollschleife. Diese so effizient und verlässlich wie möglich zu bauen, ist die eigentliche Kunst.“

*Simon Uhl entwickelt ANT, einen sich selbst korrigierenden KI-Agenten, der verlässliche News ermöglicht.*

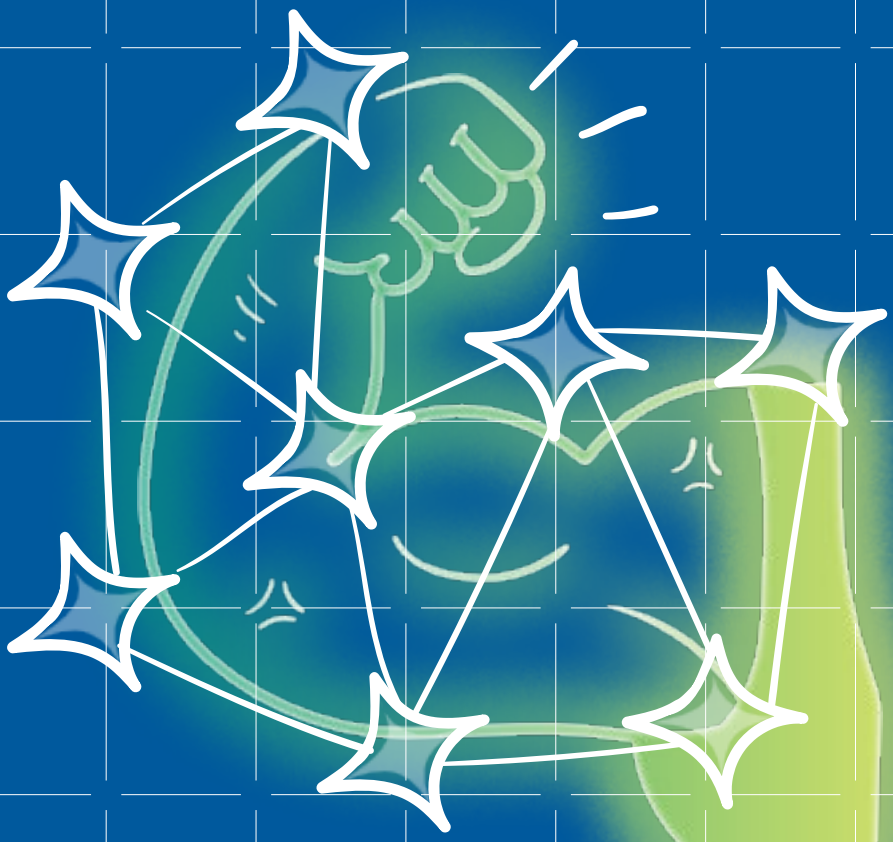
## 05

## Gesellschaft und Technik sind nicht zu trennen

„Technische und gesellschaftliche Prägung greifen stark ineinander. Diese Erkenntnis hat mir gezeigt, dass KI nicht nur Probleme verstärken, sondern auch sichtbar machen kann.“

*Patricia Sack entwickelt mit dem „Bias Buster“ ein KI-gestütztes Redaktionssystem, das durch präzise Bias-Erkennung, Zielgruppen-Analyse und Content-Optimierung die narrative Vielfalt stärkt.*

# 07 Ideen für Infrastruktur-Projekte und nutzerzentrierte Produkte



**Was bedeutet die zentrale Rolle, die KI für die Erstellung, Distribution und Rezeption von Inhalten einnimmt, nun ganz konkret für die Medien? Darüber machten sich in Tutzing acht Arbeitsgruppen konkrete Gedanken und dachten dabei zwei bis fünf Jahre in die Zukunft. Ihre Ideen lassen sich in zwei Gruppen unterteilen: Infrastruktur-Projekte und nutzerzentrierte Produkte.**

## Infrastruktur-Projekte

**Gemeinsames Archiv öffentlich-rechtlicher Medien**

In Tutzing wurde die Idee der Trusted Content Pools diskutiert. Eine Arbeitsgruppe widmete sich der Idee, die Archive von öffentlich-rechtlichen Rundfunkanstalten aus Deutschland, Österreich und der Schweiz zu einem gemeinsamen Archiv zu verbinden. Eine zentrale Suche über alle Bestände wäre für die Mitarbeitenden ein großer Mehrwert. Weitere Vorteile wären die Senkung der Betriebskosten und die Möglichkeit, standardisiertes Trainingsmaterial für KI-Anwendungen zu bekommen. Als große Herausforderung wurde die rechtliche Situation beschrieben.

**Integration von wissenschaftlichen Daten**

Ein weiteres Team will den Journalismus durch die Integration automatisierter Daten aus der Wissenschaft verbessern. Dies soll journalistische Tätigkeiten wie Recherche und Faktenprüfung erleichtern, die Genauigkeit und Relevanz der Daten erhöhen und Zeit sparen. Die wissenschaftliche Datenpipeline soll als Erweiterung in bestehende Sprachmodelle integriert werden.

**Nachrichten im Knowledge Graph speichern**

Um Medienkonsument*innen personalisierte Nachrichten anbieten zu können, müssen Medien die Art und Weise verändern, wie sie ihre recherchierten Informationen speichern. Nicht mehr in Form von ganzen Artikeln und Beiträgen, sondern in Form von kleinen, atomaren Informationseinheiten. Die Arbeitsgruppe „Unit of News“ schlägt einen Knowledge Graph bestehend aus Entitäten und ihren Beziehungen als Wissensdatenbank vor **(siehe S. 26).**

Die darin enthaltenen Informationen können mit Hilfe generativer KI zu den Inhalten und Formaten zusammengefügt werden, die sich die Medienkonsument*innen wünschen.

**Build-or-buy-Framework**

Oft stehen Medienunternehmen vor der Frage, ob sie für einen bestimmten Workflow ein externes Tool einkaufen oder selbst entwickeln sollten. Langsame Entscheidungsprozesse führen oft dazu, dass Technologien veraltet sind, bevor sie implementiert werden können. Um diese Entscheidung zu beschleunigen, schlägt ein Team die Entwicklung eines KPI-basierten Build-or-Buy-Frameworks vor.

**Adaptive User Interfaces**

Bei der Idee eines „Fluid User Interface“ (Fluid UI) geht es darum, dass sich Medieninhalte an Informationsbedürfnisse, das gewünschte Ausgabeformat und den Nutzungskontext der Medienkonsumierenden anpassen. Dabei wird eine KI-basierte Agententechnologie genutzt, die Inhalte auf verschiedenen Geräten und in unterschiedlichen Formaten bereitstellt.

Zum Beispiel kann man morgens im Auto einen Podcast hören und später den gleichen Inhalt als Text auf einem Tablet weiterlesen. Die Interaktionen sind multimodal, also über Gesten, Sprache oder Berührungen möglich. Voraussetzung dafür ist, dass der KI-Agent Zugriff auf Datenbanken hat, in die wiederum die Inhalte aus herkömmlichen Content-Management-Systemen einfließen.

## Nutzerzentrierte Produkte und Services

### Dialog ersetzt die Suche

Die Zukunft der Suche wird in zwei bis fünf Jahren stark dialogbasiert sein. Nutzende werden in ständiger Konversation mit Medienanbietern stehen. Medienunternehmen könnten eigene Plattformen entwickeln, auf denen sie ihre überprüften Inhalte anbieten – auch Archivinhalte könnten hier enthalten sein. Diese Inhalte sollten so gestaltet sein, dass die Nutzenden direkt damit interagieren können. Beispiele dafür wären textbasierte RAG-Chats oder audiobasierte Apps.

**Den Nutzenden besser zuhören**

Wissenschaftliche Untersuchungen wie der Reuters Report zeigen, dass die Zahl der Menschen, die Medien und Nachrichten meiden, zunimmt.

Um diesem negativen Trend entgegenzuwirken und solche Menschen wieder für Nachrichten zu gewinnen, schlägt die Arbeitsgruppe „Being Part of the Stories“ einen KI-getriebenen „News Buddy“ vor, der Nutzende auf verschiedenen Plattformen begegnet, ihnen zuhört und sie beim Nachrichtenkonsum begleitet.

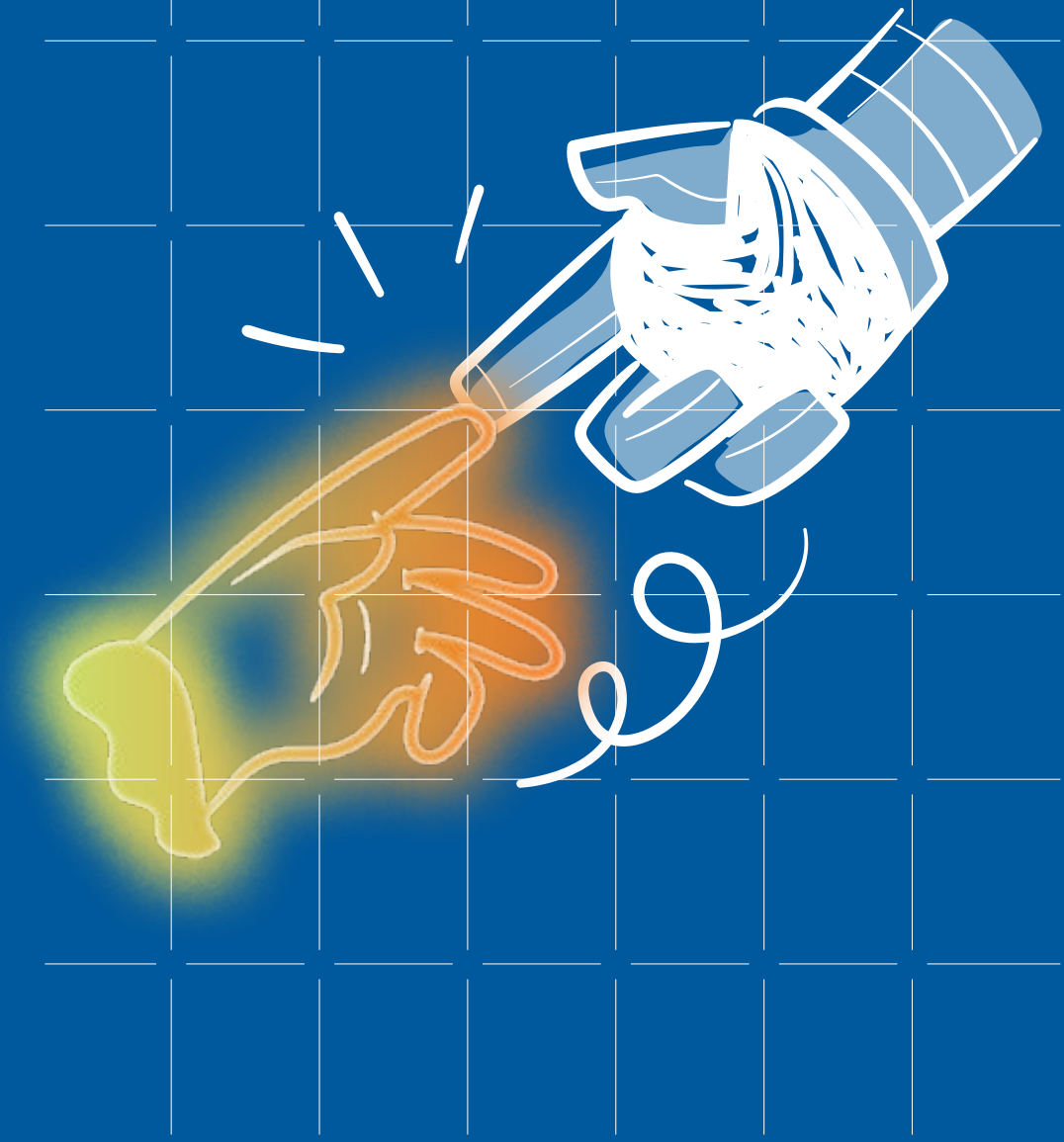
Zum einen soll der News Buddy gut zuhören, also die Interessen, Bedürfnisse und Stimmungen dieser Zielgruppe aufnehmen und analysieren. Zum anderen soll er basierend auf der Analyse des Nutzungsverhaltens in Thema, Format und Medium passende Inhalte vorschlagen – wenn möglich mit einem konstruktiven Ansatz.

Die Idee ist, durch die dialogische Nutzungssituation Vertrauen und Nähe aufzubauen und Resonanz zu ermöglichen. Diese entsteht auch, indem der News Buddy Feedback der Nutzenden entgegennimmt, das die Redaktionen dann auswerten können. Speziell die Feedback-Option ist wichtig, um den Nutzer*innen zu zeigen, dass sich Medien für ihre Meinungen, Gefühle und Stimmungen interessieren.

**Personalisierte Lokalnachrichten**

Die Idee der Arbeitsgruppe „NAIghbor“ ist eine hyperpersonalisierte lokale Nachrichtenplattform, die den Nutzenden Informationen in der gewünschten Form bereitstellt. Die Topthemen werden zwar weiterhin von der Redaktion ausgewählt, jedoch gibt es eine Suchmaske mit personalisierten Prompt-Vorschlägen.

Kern ist eine vektorbasierte Datenbank, in die verschiedene Quellen einfließen: redaktionelle Beiträge, Wetter- und Verkehrsdaten, Veranstaltungen, von der Redaktion verifizierter User Generated Content. Ein RAG-System würde auf Basis dieser Lokaljournalismus-Datenbank die von den Usern gewünschten Inhalte generieren. Auch die Nutzerhistorie könnte dabei berücksichtigt werden. ◆



# 08 Ausblick

Künstliche Intelligenz verändert schon jetzt die Art, wie journalistische Produkte produziert, konsumiert und distribuiert werden. Generative KI-Anwendungen sind in der Lage, maßgeschneiderte Inhalte für das Publikum zu erstellen, was Inhalt, Stil, Format und Länge betrifft. Basis dafür ist qualitativ hochwertiger, verifizierter Inhalt, wie ihn Journalist*innen recherchieren.

Medien können diese Inhalte in Datenbanken aufbereiten, aus denen dann generative KI-Anwendungen die personalisierten Inhalte erstellen. Die neue Form der Mediennutzung verändert auch das journalistische Berufsbild: Während Journalist*innen heute noch gezielt für ihr Publikum schreiben, wird sich ihre Rolle künftig mehr auf die Recherche und die Qualitätssicherung verlagern. Die eigentliche Produktion und Verbreitung von Inhalten wird auch von KI unterstützt werden.

Die zentrale Frage ist jedoch: Wo und wie werden diese Qualitätsinhalte überhaupt genutzt? Der Konsum von Nachrichten und Unterhaltung verlagert sich zunehmend auf KI-gestützte Anwendungen.

**Der Schwerpunkt des Journalismus wird auf Recherche und Qualitätssicherung liegen. Produktion und Verbreitung wird mehr von KI übernommen.**

Zum einen sind das KI-basierte Suchmaschinen, die auch auf Medien-Websites verlinken, wenn es dort Inhalte gibt, die die Antwort auf die Frage der Nutzer*innen liefern. Zum anderen sind es Chatbots, die häufig auf Large Language Models (LLMs) von großen Techunternehmen wie Meta, OpenAI oder Google zurückgreifen. Schon jetzt erzielen die Anbieter von generativen KI-Tools ihre Gewinne zum Teil auf Basis der Inhalte, die von Medienunternehmen geschaffen werden.

Falls sich generative KI-Plattformen zur dominierenden Benutzeroberfläche und damit verbunden auch zur dominierenden Vermittlungsplattform für Informationen entwickeln würden, hätte das für Medienunternehmen eine Reihe negativer Konsequenzen: Sie verlören zum Teil den direkten Kontakt zu ihren Nutzer*innen, deren Interessen und Bedürfnissen. Zum anderen verlören sie zahlende Kund*innen, wenn diese auf generativen Plattformen die gewünschten Informationen und Nachrichten erhalten würden – entweder kostenlos oder als Abonnent*innen von KI-Chatbots wie ChatGPT, Google Gemini oder Microsoft Copilot.

Wenn sich Menschen vor allem über generative KI-Anwendungen informierten, hätte das auch Auswirkungen auf die gesellschaftliche Diskursgestaltung. Die Generierung der Antworten basiert auf statistischen Verfahren und nicht auf redaktionellen Maßstäben. Die Gefahr, inhaltlich falsche Antworten zu liefern, ist bei Anwendungen, die auf großen Sprachmodellen basieren, systemimmanent. Welche Antworten solche Anwendungen geben, hängt maßgeblich von den Daten ab, mit denen sie trainiert wurden. Im Idealfall werden große Sprachmodelle mit hochwertigen, überprüften Inhalten aus Qualitätsmedien trainiert.

Die bisherige Erfahrung zeigt, dass die LLM-Anbieter nicht daran interessiert sind, Nutzungsverträge mit einzelnen Medien abzuschließen – von wenigen Ausnahmen abgesehen.

Anders könnte das aussehen, wenn mehrere — auch internationale – Medien ihre Inhalte in einem großen „Trusted Content Pool“ bündeln würden. Diese Inhalte wären hochwertige Trainingsdaten für Anbieter von großen Sprachmodellen. In diesem Zusammenschluss könnte man auch Standards für die Nutzung von Mediendaten etablieren.

Es gibt für Medienunternehmen großen Handlungsbedarf bei technischer Infrastruktur, Nutzerdialog und strategischer Ausrichtung. Jedes Haus muss seinen eigenen Weg finden und wird andere Akzente setzen.

Dennoch zeigt sich in den Diskussionen in Tutzing und vielen anderen Medien-Zusammenkünften, dass branchenübergreifende Standards und ein gemeinsames Bemühen um ein gesundes Medien-Ökosystem mit hochwertigen Informationen im Interesse aller Medienunternehmen liegen – so unterschiedlich ihre Geschäftsmodelle auch sind. Dabei gibt es viel Kooperationspotenzial zwischen Medienhäusern untereinander, zwischen Medienhäusern und Wissenschaft oder zwischen Medienhäusern und KI-Industrie.

Im AI for Media Network sind Vertreter*innen aus diesen Branchen versammelt. Das Netzwerk bietet – auch im Rahmen seiner Events – Gelegenheit für Vernetzung, Wissensaustausch und Produktentwicklung. Und auch über strategische Fragen kann man sich hier austauschen. Die KI-Fachtagung in Tutzing ist das beste Beispiel dafür. ◆

# 09 Glossar

**Embedden / Embedding**

Die Umwandlung von Text oder anderen Inhalten in maschinenlesbare Zahlenreihen (Vektoren), die durch statistische Verfahren (→ **Machine Learning**) miteinander verglichen und analysiert werden können. Grundlage für viele KI-Anwendungen, da sie die semantische Bedeutung von Daten mathematisch darstellen. Eine andere Bezeichnung für diesen Prozess ist Vektorisieren. Die Ergebnisse werden oft in einer → **Vektordatenbank** gespeichert.

**Knowledge Graph (Wissensgraph)**

Strukturierte Darstellung von Wissen, die Informationen in Form eines semantischen Netzwerks aus Knoten und Kanten organisiert. Knoten repräsentieren Entitäten wie Personen, Orte, Objekte oder auch Konzepte. Kanten wiederum stellen die Beziehungen zwischen den Knoten dar. Wenn Albert Einstein und Relativitätstheorie Objekte sind, ist „hat erfunden“ die Beziehung zwischen den beiden Objekten. Ein Knowledge Graph ermöglicht die Verbindung und Verknüpfung von Datenpunkten, sodass komplexe Zusammenhänge zwischen Entitäten erkennbar werden.

**Large Language Model (LLM) / Sprachmodell**

Machine-Learning-Modell, das für die Verarbeitung und Erzeugung natürlicher Sprache entwickelt wurde. Es akzeptiert Freitext als Eingabe und generiert passende Antworten, indem es die Bedeutung von Wörtern und Sätzen in ihrem Kontext analysiert. Benötigt riesige Mengen an → **Trainingsdaten**

**Machine Learning**

Eine Sammlung von statistischen Verfahren, die es Computern ermöglicht, selbstständig Muster und Wahrscheinlichkeiten in Daten zu erkennen (zu „erlernen“) und diese Muster auf unbekannte Daten anzuwenden. Grundlage jeder modernen Form der KI. Benötigt große Mengen an strukturierten → **Trainingsdaten**

**Natural Language Processing (NLP)**

Ein Teilgebiet der künstlichen Intelligenz und der Computerlinguistik, das sich mit der Interaktion zwischen Computern und menschlicher Sprache beschäftigt. Ziel ist es, natürliche Sprache (Text oder Sprache) zu analysieren, zu verstehen, zu generieren oder in strukturierter Form zu verarbeiten.

**Retrieval-Augmented Generation (RAG)**

Software-Architektur, die aus zwei Komponenten besteht: Relevante Informationen werden aus einer Wissensquelle (z. B. → **Vektordatenbank**, → **Knowledge Graph**) abgerufen (Retrieval). Ein Sprachmodell nutzt die abgerufenen Inhalte als zusätzlichen Kontext, um präzise, faktenbasierte Inhalte zu erstellen (Augmented Generation).

**Trainingsdaten**

Bilden die Grundlage für die Entwicklung einer KI. Sie bezeichnen die Menge an „Beispielen", aus denen ein KI-Modell mit Hilfe von → **Machine Learning** Muster und Wahrscheinlichkeiten „lernt". Trainingsdaten müssen dafür sorgfältig ausgewählt und aufbereitet werden.

**Vektordatenbank**

Eine Vektordatenbank speichert Daten als numerische Vektoren, die bestimmte Eigenschaften dieser Daten darstellen – etwa die Bedeutung von Texten. Diese Darstellung ermöglicht es, Datenpunkte effizient anhand ihrer Ähnlichkeit zu vergleichen, indem die Abstände zwischen den Vektoren berechnet werden. Vektordatenbanken werden oft in → **RAG-Systemen** verwendet, um zusätzliche Daten für die Generierung neuer Inhalte abzurufen.

**IMPRESSUM**

**Herausgeber**
AI for Media Network, eine Initiative des Bayerischen Rundfunks

Bayerischer Rundfunk
Anstalt des öffentlichen Rechts
gesetzlich vertreten durch die Intendantin
Dr. Katja Wildermuth

Rundfunkplatz 1
80335 München

**V.i.S.d.P.**
Uli Köppen

**Autoren**
Bernd Oswald (AI for Media Network), Maren Schuster (Universität Halle-Wittenberg), Christian Simon (Media Lab Bayern)

**Redaktion**
Uli Köppen (Bayerischer Rundfunk),
Cécile Schneider (Bayerischer Rundfunk),
Johannes Klingebiel (Media Lab Bayern)

**Lektorat**
Dr. Ursula Ruppert,
Max Brandl (Bayerischer Rundfunk)

**Gestaltung**
Johannes Klingebiel (Media Lab Bayern)

Das Media Lab Bayern ist eine Initiative der Medien.Bayern GmbH und wird gefördert durch die Bayerische Landeszentrale für neue Medien und die Bayerische Staatskanzlei.