



Fachbereich
Ingenieurwissenschaften

Studienbereich
Informationstechnologie
und Elektrotechnik

Studiengang
Advanced Media Technology

Master-Thesis

Herausforderungen und Herangehensweise bei der
Einbindung eines Large Language Models in den
Newsworkflow
-
betrachtet am Beispiel der Einführung von ChatGPT

Maike Kaiser (B. Eng.)

vorgelegt am: 14.08.2023

Referent: Prof. Dr.-Ing. Wolfgang Ruppel, Hochschule RheinMain

Korreferent: Damian Mucko (M. Eng.), BFE Studio und Medien Systeme

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe dritter Personen verfasst habe.

Die aus fremden Quellen direkt oder indirekt übernommenen Texte, Gedankengänge, Konzepte usw. in meinen Ausführungen habe ich als solche eindeutig gekennzeichnet und mit vollständigen Verweisen auf die jeweilige Urheberschaft und Quelle versehen.

Alle weiteren Inhalte wie Textteile, Abbildungen, Tabellen etc. ohne entsprechende Verweise stammen im urheberrechtlichen Sinn von mir.

Die vorliegende Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Mir ist bekannt, dass ein Täuschungsversuch vorliegt, wenn sich eine der vorstehenden Versicherungen als unrichtig erweist.

Ort, Datum und Unterschrift Verfasserin

Abstract

With the progressive development of Artificial Intelligence (AI) over the past 80 years, it is now finding its way into the various areas of people's lives. In the automotive, insurance and media environments, AI is already an integral part. With the rapid development of Large Language Models (LLM) over the last five years, we are also seeing the application of language processing becoming more and more profitable. With LLMs, however, it is now possible to have more complex texts written by an AI. These thus form a powerful tool to support journalists in their everyday work.

The goal of the master's thesis is therefore to define requirements for an AI for use in journalism and to propose an approach for the introduction of ChatGPT into the news workflow. The approach consists of a non-technical description of concepts to meet the pre-defined requirements and a technical solution to translate the concepts into a system landscape. The master's thesis therefore aims to show the requirements which have to be fulfilled and how they can be implemented in order for AI to be used efficiently and safely in journalism. Furthermore, it shows the limitations of ChatGPT in relation to the requirements and recommends the use of other AI models in addition to ChatGPT to cover even more use cases.

ChatGPT is suitable for the use in the news workflow under certain conditions. It is technically possible to integrate the AI model into the news workflow. The introduction phase and the subsequent productive use are iterative processes. The individual process steps result from the experiences during the project and the feedback from the users. The preconditions under which ChatGPT will be integrated efficiently and safely into the news workflow are initially of a non-technical nature. They include aspects such as new roles, measures to comply with legal and ethical requirements, feedback and average scenarios. The technical solution carries the concepts. By using additional automation software an intermediary is inserted between ChatGPT and the user. The middleman ensures the security, the transfer of information and the initiation of automatic workflows.

With this solution about half of the use cases in the requirements matrix are put into action. In order to implement further use cases, other OpenAI models are suitable for the use in the news workflow. As a supplement to ChatGPT, the models from the InstructGPT bundle, DALL-E and Whisper cover more use cases. These thus also include the processing of images and sound. For the implementation of these models, additions to the existing concepts and technical solutions are needed. Additionally, the new models are capable of user finetuning, a method of giving the AI model in-house data as training material. This allows the AI model to learn and output the style and attitude of the media company. Nevertheless, even the combination of models does not cover all use cases. For example, the tagging of images and videos based on the visual component is missing. At the end of the day, AI is a technology that has undergone rapid development in recent years. It has the ability to enable a high increase in efficiency in processes. However, it still requires strict regulation and supervision by humans.

Inhalt

1	Einleitung.....	1
2	Einführung in Künstliche Intelligenz	4
2.1	Grundlagen Künstlicher Intelligenz	4
2.1.1	Neuronale Netze	4
2.1.2	Sprachverarbeitung mit Transformer.....	6
2.1.3	Overfitting	8
2.2	Geschichte	10
2.3	ChatGPT	12
3	Einführung in den Newsworkflow	14
3.1	Rundfunk- und Onlinejournalismus	14
3.2	Formate	16
3.3	Arbeitsabläufe.....	18
3.4	Infrastruktur.....	21
3.5	Kommunikation mit dem Publikum	25
4	Analyse der Ist-Workflows.....	29
4.1	Ingest	29
4.2	Tägliche Nachrichten.....	31
4.3	Breaking News.....	37
4.4	Informationsnachrichten	39
4.5	Rückmeldung von Nutzerinnen und Nutzern	40
5	Anforderungen an eine Systemlandschaft mit Künstlicher Intelligenz	41
5.1	Rechtliche und Ethische Anforderungen	41
5.1.1	Datenschutz	42
5.1.2	Urheberrecht	43
5.1.3	Informationssicherheit	45
5.1.4	Erklärbarkeit	47
5.1.5	Fairness	48
5.1.6	Transparenz	53
5.1.7	Regulierung.....	57
5.2	Funktionale Anforderungen.....	60
5.2.1	Assistierende Aufgaben.....	60
5.2.2	Generative Aufgaben.....	63
5.2.3	Distributive Aufgaben	64
5.2.4	Benutzbarkeit	67
5.3	Organisatorische Anforderungen.....	69
5.3.1	Aufbauorganisation.....	69
5.3.2	Ablauforganisation.....	70
5.3.3	Personalorganisation	72

6	Vorschlag einer Strategie unter Einbindung von ChatGPT	74
6.1	Ausrichtung des Projekts	75
6.2	Eignung von ChatGPT	77
6.3	Soll-Workflows.....	79
6.3.1	Ingest	79
6.3.2	Nachrichtensendung.....	79
6.3.3	Breaking News	81
6.3.4	Informationsnachrichten.....	81
6.3.5	Rückmeldung und Anfragen von Nutzerinnen und Nutzern	81
6.4	Rollenkonzept	83
6.5	Compliancekonzept.....	86
6.6	Zugangskonzept	90
6.7	Schulungskonzept	91
6.8	Feedbackkonzept	93
6.9	Auditkonzept.....	95
6.10	Havariekonzept.....	96
6.11	Risikoanalyse	98
6.12	Wirtschaftliche Analyse	102
7	Technische Lösungsbeschreibung.....	104
7.1	Infrastruktur	104
7.2	Informationsflüsse.....	106
7.3	Interfaces	109
7.3.1	Automation.....	109
7.3.2	Schnittstellen	111
7.3.3	Metadaten.....	114
7.3.4	GUI	117
7.3.5	Kryptografische Verfahren	118
7.4	Prompt Engineering.....	120
7.5	Zwischenfazit	122
8	Weiterführende Überlegungen.....	123
8.1	Weiteres Angebot von OpenAI	123
8.2	Ergänzung zu den Workflows	124
8.3	Ergänzungen zu den Konzepten	126
8.3.1	Rollenkonzept.....	126
8.3.2	Zugangskonzept.....	126
8.3.3	Schulungskonzept.....	127
8.3.4	Konzept der Information an Nutzer*innen.....	128
8.3.5	Reglementierter Zugang zu Inhalten.....	128
8.3.6	Sicherheitskonzept	129

8.3.7	Wirtschaftliche Analyse.....	129
8.4	Ergänzung zur technischen Lösungsbeschreibung.....	131
8.4.1	Informationsflüsse	131
8.4.2	Schnittstellen.....	132
8.4.3	Weitere GUIs.....	134
8.4.4	Datenaufbereitung und Training	135
8.5	Zwischenfazit.....	138
9	Diskussion.....	139
10	Fazit.....	144
11	Verzeichnisse	VII
11.1	Abkürzungsverzeichnis.....	VII
11.2	Abbildungsverzeichnis	VIII
11.3	Literatur.....	IX
11.4	Inhaltsverzeichnis des Datenträgers	XVI
12	Anhang.....	XIX
	Anhang 1: Anforderungsmatrix	XIX
	Anhang 2: Soll-Workflows.....	XXVI
	Anhang 3: Analyse der crossmedialen Workflows mit KI-Unterstützung	XXXIV
	Anhang 4: Rollenübersicht.....	XXXVI
	Anhang 5: Schulungsmatrix	XLIV
	Anhang 6: Risikomatrix.....	XLVI
	Anhang 7: Informationsflüsse.....	XLIX
	Anhang 8: API-Programmiersvorschläge	LVII
	Anhang 9: Anforderungen an die grafischen Benutzungsoberflächen.....	LXIX
	Anhang 10: Beispiele für Prompt Engineering	LXX

1 EINLEITUNG

Mit der fortschreitenden Entwicklung von Künstlicher Intelligenz (KI) in den vergangenen 80 Jahren hält diese nun Einzug in die verschiedenen Lebensbereiche der Menschen. Im Automobil-, Versicherungs- und Medioumfeld ist KI bereits ein fester Bestandteil. Durch die rasante Entwicklung von LLMs in den letzten fünf Jahren werden nun auch die Einsatzmöglichkeiten der Sprachverarbeitung immer gewinnbringender. KI wird bereits in Medienunternehmen angewendet, um beispielsweise Verkehrs- oder Sportnachrichten zu generieren. Diese folgen meist demselben Schema und benötigen nur austauschbare Informationen. Mit den LLMs ist es nun aber möglich, komplexere Texte nach nur wenigen Vorgaben durch eine KI schreiben zu lassen. Sie bilden damit ein mächtiges Werkzeug, um Journalistinnen und Journalisten in ihrem Alltag zu unterstützen. Trotzdem muss ein gewisser Qualitätsstandard für das KI-System gegeben sein. Die Arbeit unter Zuhilfenahme einer Maschine ist für Journalisten zusätzlich eine unbekannte Entwicklung, sodass diese mit gemischten Gefühlen dieser entgegensehen [1, S. 48].

Forschungsgegenstand

Das Ziel dieser Masterthesis ist daher, Anforderungen an eine KI für den Einsatz im Journalismus zu definieren und eine Herangehensweise für die Einführung von ChatGPT in den Newsworkflow vorzuschlagen. Die Herangehensweise besteht aus der nicht-technischen Beschreibung von Konzepten, um die vorher definierten Anforderungen einzuhalten, und einer technischen Lösung, um die Konzepte in eine Systemlandschaft zu übersetzen. Die Konzepte bringen also typische Vorgehensweisen aus der Systemintegration mit Best Practices zur Integration von KI zunächst deskriptiv zusammen. Die technische Lösung nennt spezifische Werkzeuge und deren Einbindung in die bestehende Systemlandschaft. Die Masterthesis soll darlegen, welche Voraussetzungen erfüllt sein müssen, damit KI effizient und sicher im Journalismus verwendet werden kann, und auf welche Weise und unter Berücksichtigung welcher Aspekte ein Einführungsprozess ablaufen kann. Darüber hinaus zeigt sie die Limitierungen von ChatGPT in Bezug auf die Anforderungen auf und empfiehlt den Einsatz weiterer KI-Modelle zusätzlich zu ChatGPT, um noch mehr Use Cases abzudecken.

Aufbau der Arbeit

In dieser Arbeit werden zwei Welten zusammengebracht, die bis jetzt noch wenig Berührungspunkte hatten: KI, im Besonderen Sprachmodelle, und Nachrichtenjournalismus. Daher beginnt die Thesis mit zwei Einführungskapiteln. Die Einführung in KI behandelt die Definition, die historische Entwicklung bis zu den heutigen LLMs und einen Einblick in die Funktionsweise der Modelle. In der Einführung in den Journalismus werden die Formate, Arbeitsabläufe und die Infrastruktur mehrerer Medienformen herausgearbeitet. Die Einführungskapitel grenzen damit den Umfang der Arbeit ein. Printjournalismus sowie KI-Modelle, die keine Sprache verarbeiten, sind entsprechend nicht Teil dieser Arbeit.

Aus der Einführung in den Journalismus heraus werden schließlich die existierenden Ist-Workflows in Fernsehen, Hörfunk und Onlinemagazinen abgebildet und analysiert, um die Ausgangslage in den Medienunternehmen darzulegen. Diese unterscheiden sich von den vorherigen Kapiteln insofern,

als dass hier Wert auf die prozessualen Schritte unabhängig von dem jeweiligen Inhalt der Sendungen bzw. Beiträge gelegt wird.

Im nächsten Arbeitsschritt werden die Anforderungen an die KI aus Sicht der Medienhäuser definiert. Diese erstrecken sich von funktionalen (Use Cases) über organisatorische bis hin zu ethisch-rechtlichen Anforderungen, die es bei der Arbeit mit KI zu erfüllen gilt. Das Kapitel umfasst die Anforderungen und mündet in die Anforderungsmatrix. Diese nummeriert einzelnen Anforderungen durch und verweist auf die Lösungskapitel sechs, sieben und acht.

Der zweite Teil der Thesis bildet die Konzeption einer Einführungsstrategie für ChatGPT in den Newsworkflow. Das Kapitel beinhaltet eine Ausrichtung mit einer formulierten Strategie sowie die Neu-Konzeption der Workflows mit KI-Unterstützung. Bei der Beschreibung der Workflows wird eine Brücke zurück zur Definition der Ist-Workflows in Kapitel vier geschlagen, denn die neuen Workflows entspringen aus den herkömmlichen.

Aus der Einführung in KI folgt die vorgelagerte Projektarbeit „Gegenüberstellung von ChatGPT und Luminous“ der Autorin. Diese misst die Qualität der Antworten von ChatGPT und Luminous, einem deutschen Sprachmodell von Aleph Alpha. In der Projektarbeit wurden die Outputs auf journalistisch relevante und generische Prompts verglichen. Dadurch lassen sich die Stärken und Schwächen von ChatGPT ablesen. Aus der Bewertung der Eignung folgen die Konzepte in Kapitel sechs, die die Einhaltung der Anforderungen aus Kapitel fünf sicherstellt sowie die Partizipation der Mitarbeiterinnen und Mitarbeiter aktiviert.

Die Einführungsstrategie wird von einer technischen Lösung getragen, die die nötige Infrastruktur, Schnittstellen und Datenflüsse für die Umsetzung der Konzepte definiert. Die Infrastruktur beruht auf einer Automation, die als Vermittlung zwischen ChatGPT, Nutzer*in und Media Asset Management (MAM) agiert. Innerhalb des Systems werden Informationen über Schnittstellen in Form eines einheitlichen Metadatenformates zwischen den Komponenten übergeben.

Bei der Erarbeitung der oben genannten Punkte werden jedoch auch die Limitierungen des Chatbots ChatGPT deutlich. Da es sich um ein reines Sprachmodell handelt, werden beispielsweise Use Cases, die auf Bild oder Audio abzielen, außer Acht gelassen. Um diese Use Cases zusätzlich abzudecken, enthält die Thesis weiterführende Informationen zu anderen Produkten von OpenAI, die diese Use Cases erfüllen. Zur Implementierung dieser Modelle werden Ergänzungen zu den obengenannten Workflows, Konzepten und Lösungen vorgestellt.

Diesem Kapitel schließt sich eine Diskussion über die Eignung und den verantwortungsvollen Umgang mit KI an einer solch wichtigen Stelle wie dem Journalismus an. Die sogenannte vierte Gewalt innerhalb einer Demokratie muss nach bestem Wissen und Gewissen eine qualitativ hohe Berichterstattung gewährleisten. Darüber hinaus behandelt die Diskussion auch Anwendungsfälle außerhalb des Journalismus. Ebenfalls wird die Vormachtstellung des amerikanischen Marktes gegenüber dem Rest der Welt betrachtet.

Das Fazit fasst die zwei Wege in der Thesis zusammen, indem es sich für die technische Umsetzung also auch für den verantwortungsvollen Einsatz von KI ausspricht. Die technische Umsetzbarkeit muss im Bereich von KI immer im Zusammenhang mit der Verantwortung für Anwender*innen, Betreiber*innen und die Gesellschaft betrachtet werden.

Bei der Entwicklung der großen Sprachmodelle handelt es um einen revolutionären Umbruch in vielen Bereichen der Arbeit und des menschlichen Lebens. Dieser Umbruch ist mit der Veröffentlichung von ChatGPT nicht beendet und wird noch viele weitere Anwendungsfälle und Entwicklungen bereithalten.

2 EINFÜHRUNG IN KÜNSTLICHE INTELLIGENZ

Bei KI handelt es sich um ein Teilgebiet der Informatik, das sich mit der Automatisierung intelligenten Verhaltens und dem Maschinellen Lernen (ML) befasst [2]. In diesem Gebiet sollen Techniken entwickelt werden, die genutzt werden, wenn Maschinen Aufgaben übernehmen, die normalerweise menschliche Intelligenz erfordern wie visuelle Wahrnehmung, Spracherkennung, Entscheidungsfindung und Übersetzung von Sprachen [2].

ML ist ein Teilgebiet der KI. Die Grundidee ist, dass sich Computer aufgrund ihrer gemachten Erfahrungen anhand von Beispielen und Lösungen verbessern. Es werden keine festen Regeln programmiert, also beispielsweise keine „Wenn..., dann...“-Abfragen [2]. Der Computer untersucht die Beispiele nach mehreren Eigenschaften und generiert anhand der Eigenschaften einen Output. Die Regeln, wann ein bestimmter Output richtig ist, findet der Computer selbst heraus [2].

Deep Learning ist wiederum ein Teilgebiet des ML. Es basiert auf Neuronalen Netzen (NN). Mit dieser Technologie sind Computer nun in der Lage, sich selbst zu trainieren [2].

2.1 GRUNDLAGEN KÜNSTLICHER INTELLIGENZ

Diese Entwicklung basiert auf fortgeschrittenen Methoden der Informatik und Mathematik. Die relevanten Techniken sind im Folgenden erläutert.

2.1.1 NEURONALE NETZE

NN sind dem menschlichen Gehirn nachempfunden. Während im Kopf eines Menschen Neuronen mit einer Vielzahl an Synapsen untereinander verbunden sind, sind in einem NN Neuronen oder auch Knoten mit gewichteten Verbindungen verlinkt. Die Knoten sind in mehreren Layern organisiert. Die oberste Schicht wird Eingabelayer genannt (n_1 bis n_4 in Abbildung eins), die unterste Ausgabelayer (n_5 bis n_8). Dazwischen befindet sich eine Vielzahl von Hidden Layern (h_{11} bis h_{26}). Nach dem Vorbild des Gehirns, in dem viel beanspruchte Synapsen immer stärker werden, werden auch die häufig benutzten Verbindungen zwischen Knoten immer stärker gewichtet. In einem untrainierten Netz haben alle Gewichte einen Standardwert. Die veränderlichen Teile des NN sind also die Gewichte (w_1 bis w_{33}).

Jeder Knoten des Ausgabelayers repräsentiert einen möglichen Output. Welches nun der richtige Output ist, wird vom wahrscheinlichsten Weg für die Eingabe durch das Netz bestimmt. Die Gewichte geben die Wahrscheinlichkeiten auf dem Weg an. Die Gewichte werden wiederum im Training festgelegt. Dafür werden die Trainingsbeispiele in das NN eingegeben, ein Output erzeugt und dieser mit der Lösung abgeglichen. Eine Fehlerfunktion bestimmt den Abstand zwischen Output und Lösung. Die Gewichte werden so angepasst, dass die Fehlerfunktion ein Minimum aufweist. Je mehr Trainingsdaten verwendet werden, desto genauer sind die Werte für die Gewichte. Wenn also bei der Angabe der Größe der KI-Modelle von Parametern gesprochen wird, ist damit die Gesamtzahl der Gewichte gemeint [3].

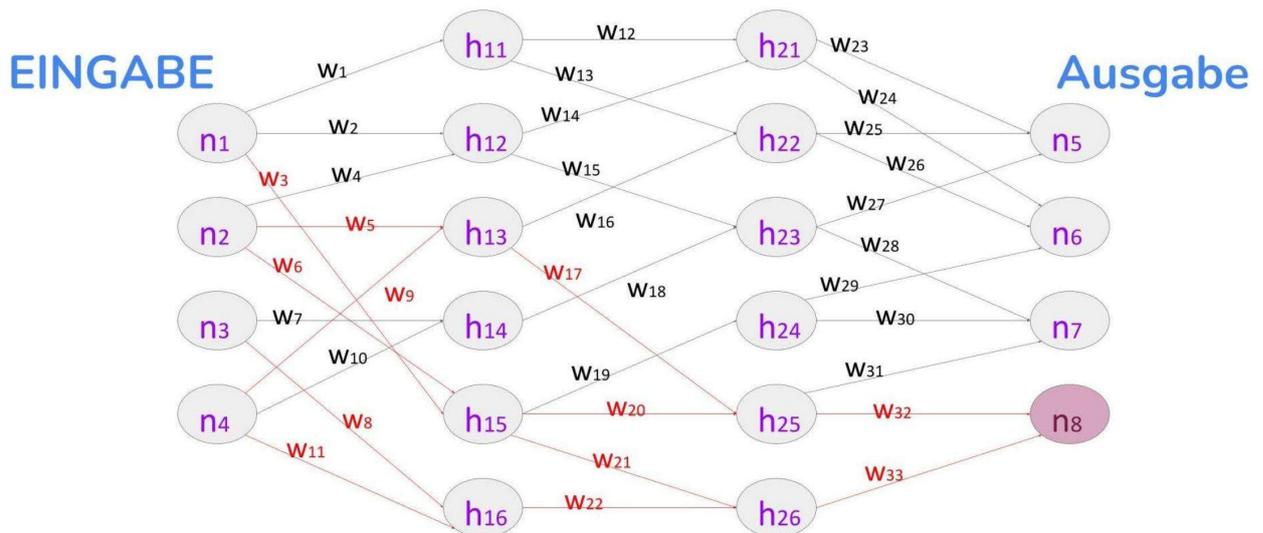


Abbildung 1: Neuronales Netz [4].

Faltungnetzwerke

Faltungnetzwerke sind eine Unterart der NN, die häufig in der Bildverarbeitung genutzt werden. Die einzelnen Hidden Layer arbeiten dabei als Filter, die dazu dienen, bestimmte Merkmale in Bildern zu extrahieren. Dies können die Konturen oder die Struktur eines abgebildeten Motivs sein. Bei diesem Netz bilden Bilder bzw. ihre Pixelwerte den Input. Das Netz erhält also bei einem High Definition-Bild 1920 x 1080 x 3 Werte von bis zu 16 Bit pro Wert als Eingabematrix. Das Bild wird gebietsweise analysiert, Merkmale gefunden und diese weiterbearbeitet [4].

Rekurrente Neuronale Netze (RNN)

RNN haben im Vergleich zu den klassischen NN einen Erinnerungsmechanismus. Das bedeutet, dass die Neuronen eines Layers nicht nur mit dem folgenden Layer verbunden sind, sondern auch mit denen desselben oder vorangegangenen Layers (siehe Abbildung zwei). Solche Netze sind besonders zur Sprachverarbeitung geeignet, da allein innerhalb eines Satzes Beziehungen zwischen den Wörtern bestehen, die sowohl in als auch gegen die Leserichtung bestehen [4].

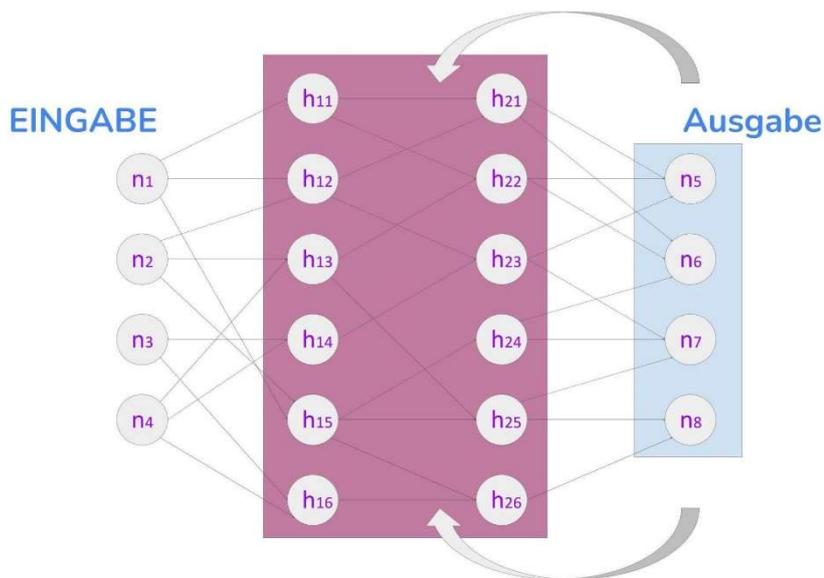


Abbildung 2: Rekurrentes Neuronales Netz [4].

Liest ein RNN einen Satz ein, arbeitet es sich Wort für Wort vor. Das erste Wort wird geladen und in einen inneren Vektor übersetzt. Um es mit den anderen Worten des Satzes in Verbindung zu setzen, werden nun nacheinander die anderen Worte eingelesen und der innere Vektor entsprechend aktualisiert [4]. Dies führt dazu, dass bei sehr langen Dateneingaben (z.B. ein längerer Text) der innere Vektor überschrieben wird und so Informationen verloren gehen. So können Beziehungen zwischen zwei im Text weit entfernten Wörtern nicht mehr erkannt werden [4].

2.1.2 SPRACHVERARBEITUNG MIT TRANSFORMER

Ein GPT (Generative Pre-trained Transformer) besteht also aus den Wörtern generativ, vortrainiert und Transformer. Generativ bedeutet, dass das KI-Modell neue Erzeugnisse auf Basis seines Datensatzes erstellt, wie beispielsweise die Produktion eines Textes. Ein vortrainiertes Modell hat schon für die einzelnen NN, aus denen es besteht, große Mengen an Daten erhalten [5].

Damit ein KI-Modell nun einen vollständigen, von Menschen lesbaren Satz schreiben kann, hat es zunächst große Mengen von Texten trainiert und kennt die wahrscheinlichsten Beziehungen zwischen Wörtern [6]. Es hat die Muster der Sprache erlernt, damit es einen Satz immer mit dem wahrscheinlichsten Wort ergänzen kann [7]. Für einen unvollständigen Satz wird also jedes Wort eingesetzt, das im Vokabular des KI-Modells die höchsten Gewichte hat [4].

Dafür gibt es zwei Möglichkeiten: Die Next-Token-Prediction und das Masked-Language-Modeling. Bei der ersten Möglichkeit wird der Satz in Leserichtung vervollständigt. Es wird also ein Wort nach dem anderen dem Satz hinzugefügt, je nachdem, welches die höchste Wahrscheinlichkeit aufweist. Beim Masked-Language Modeling wird ein Wort innerhalb des Satzes maskiert und das KI-Modell wählt das Wort mit der höchsten Wahrscheinlichkeit für diese Lücke aus. Hier wird der Satz nicht hintereinander aufgebaut, sondern Wörter in Zwischenräume eingefügt.

Ein Transformer-Netz ist für die Sprachverarbeitung geeignet. Die Wörter werden von einem Encoder in einen maschinenlesbaren Vektor überführt. Wörter im Satz sind durch Pluralbildung, Deklination und Konjugation von ihrer ursprünglichen Form verändert. Um die Bedeutung der Wörter zu erfassen, werden die Methoden des Stemming und der Lemmatization angewandt. Stemming bedeutet, dass das Wort im Satz auf seinen Wortstamm reduziert wird. So können Substantive, Verben und Adjektive mit demselben Wortstamm einander zugeordnet werden. Lemmatization dagegen formt das Wort im Satz nur zu seinem Lemma, seiner Grundform, um. Ein Substantiv im Plural wird beispielsweise zu seiner Singularform zurückgeführt [8, S. 3445]. Dadurch werden nicht für mehrere grammatikalische Formen eines Wortes mehrere Einträge gespeichert, sondern unter einem Eintrag zusammengefasst.

Alle Wörter sind in einem mehrdimensionalen Vektorraum angeordnet. Thematisch verwandte Wörter sind benachbart angeordnet. Verwandt bedeutet hier, dass zwei Ausdrücke aus demselben Kontext stammen. Sie erscheinen im Trainingsmaterial häufig im Zusammenhang. Sich reimende Wörter oder Homonyme, die bei Menschen zu Verwechslung führen können, sind hier kein Problem, denn hier wird nur die Semantik beachtet [4]. Dieser Vektorraum wird durch Clustertechniken erzeugt. Zwei davon sind das hierarchische und das zusammenführende Clustern. Bei der ersten Technik wird eine große Anzahl an lose verwandten Wörtern in weitere Kategorien unterteilt und im Vektorraum angeordnet. Die andere Technik arbeitet in die Gegenrichtung. Einzelne Wörter werden Schritt für Schritt mit ähnlichen Wörtern zusammengebracht, bis sich auch hier Kategorien gebildet haben. Diese werden dann ebenfalls im Vektorraum angeordnet [8, S. 3446]. Diese Technik wird Einbettung oder auch Embedding genannt.

Anders als in einem NN wird hier nicht ein Wort nach dem anderen verarbeitet, stattdessen werden mehrere Wörter parallel durch einen zweistufigen Encoder eingelesen. Er besteht aus einem Aufmerksamkeitsmechanismus und einem vorgeschalteten NN (siehe Abbildung drei links). Während sich das NN einem Wort widmet, kann der Aufmerksamkeitsmechanismus bereits das nächste Wort bearbeiten [4]. Der Aufmerksamkeitsmechanismus identifiziert die Beziehungen der Wörter untereinander. Dies kann ein Attribut, das ein Objekt näher beschreibt, oder ein Personalpronomen, das sich auf das Subjekt des vorherigen Satzes bezieht, sein. Der Aufmerksamkeitsmechanismus ordnet außerdem jedem Wort im Satz ein Gewicht zu, um das Ausmaß seines Beitrags zur Aussage des Satzes zu beschreiben [6].

Soll beispielsweise ein Satz übersetzt werden, werden zunächst die Wörter der Ausgangssprache mit ihren Beziehungen untereinander von einem Encoder in die Vektorschreibweise überführt. Ein Decoder übersetzt den Vektor dann in die Zielsprache mithilfe eines NN (siehe Abbildung drei rechts). Der Decoder besteht auch aus einem Aufmerksamkeitsmechanismus. Zusätzlich gleicht dieser die Aufmerksamkeit der Ausgabe mit der Aufmerksamkeit des Input ab, sodass auch der Satz in der Zielsprache ähnliche Beziehungen zwischen den Wörtern enthält wie in der Ausgangssprache. Es ist je nach Sprache aber nicht möglich, die Beziehungen zwischen den Wörtern vollständig zu übernehmen [4].

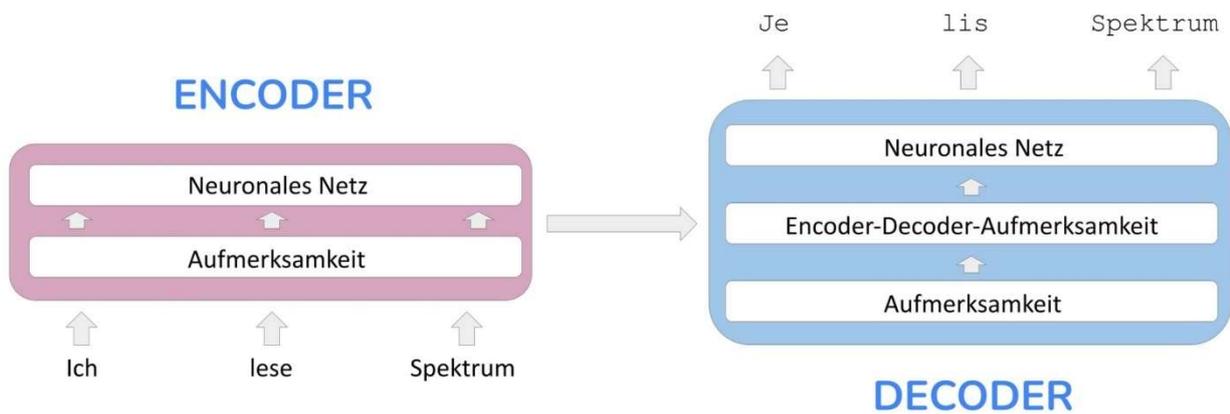


Abbildung 3: Aufbau eines Transformers [4].

In Abbildung vier sind die Wörter „Mann“, „Frau“, König“ und „Königin“ im Vektorraum eingetragen. „König“ und „Königin“ sind thematisch verwandt, daher befinden sie sich in geringem Abstand zueinander. „Königin“ und „Frau“ sowie „König“ und „Mann“ sind ebenfalls verwandt, deshalb sind sie auch räumlich nah beieinander. Es kann sogar eine simple Arithmetik aufgestellt werden: „Königin“ = „König“ – „Mann“ + „Frau“.

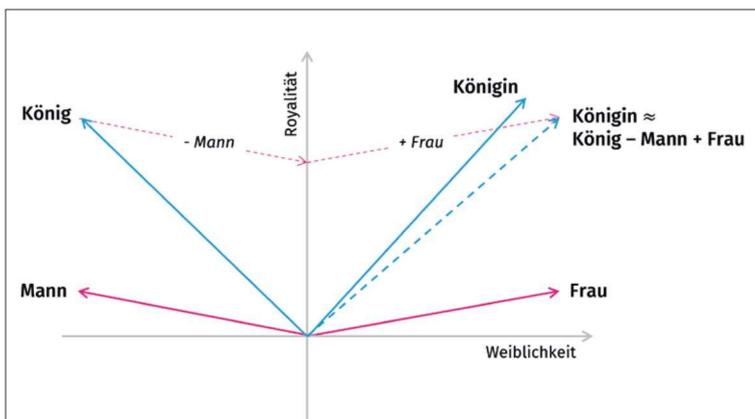


Abbildung 4: Anordnung und Beziehung von Wörter im Vektorraum [9, S. 35].

2.1.3 OVERFITTING

Overfitting ist ein Phänomen, das beim Training von KI-Modellen auftritt. Es hängt maßgeblich vom systematischen Fehler und der Varianz ab. Es tritt auf, wenn der systematische Fehler sehr niedrig und die Varianz sehr hoch ist. Der systematische Fehler ist die Differenz zwischen den eigentlichen und den prädizierten Werten. Die Varianz gibt an, wie sehr sich die Schätzung verändert, wenn immer mehr verschiedene Datenpunkte zu ihrer Berechnung herangezogen werden. Der Zusammenhang zwischen dem systematischen Fehler und der Varianz ist antiproportional. Versucht man den systematischen Fehler zu verringern, erhöht sich die Varianz [10].

Overfitting entsteht also dadurch, dass sich das Modell so gut an die Trainingsdaten anpasst, dass auch Verzerrungen mitberücksichtigt werden (siehe Abbildung fünf). Die Regeln des Modells passen sich also auch an Datenausreißer an. Werden dem Modell nun Testaufgaben auf Basis eines anderen Datensatzes gestellt, erzielt das Modell schlechtere Ergebnisse [10].

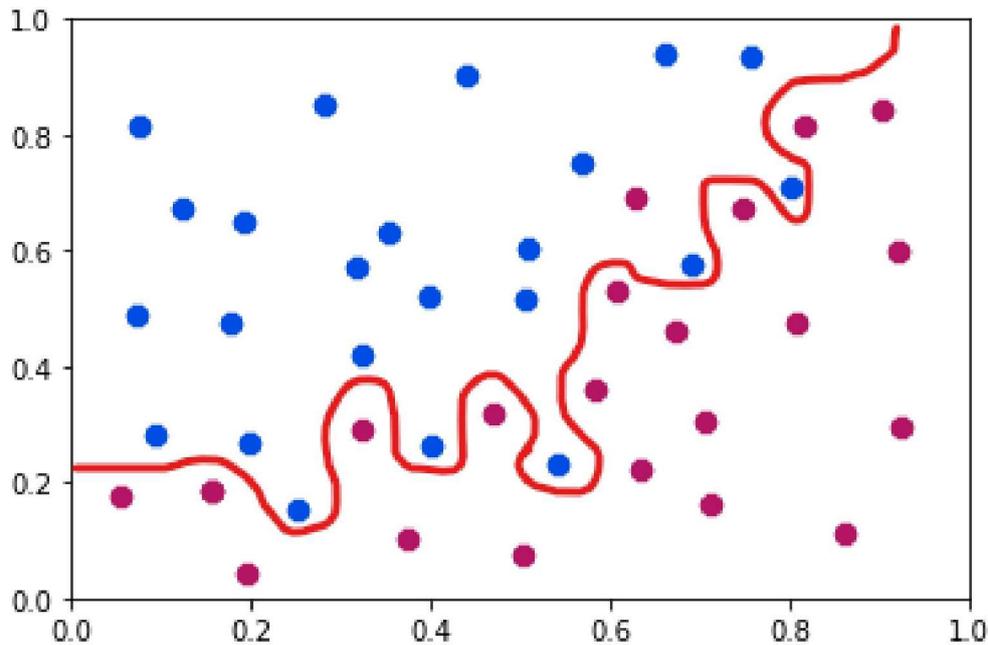


Abbildung 5: Darstellung der Entscheidungsgrenze bei Overfitting [10].

Die Gründe für Overfitting liegen zunächst in den Trainingsdaten. Besonders verzerrte oder ungenaue Daten sowie ein zu kleiner Datensatz sind die Auslöser. Ein komplexes Modell ist außerdem anfällig für Overfitting. Durch seine verzweigte Struktur gibt es mehr Möglichkeiten für einen Datenpunkt zu einer Kategorie zugeordnet zu werden [10]. Overfitting kann durch pragmatische Maßnahmen vermindert werden. Durch die Erhöhung des Trainingsdatensatzes und Datenbereinigung können die ersten beiden Gründe für Overfitting beseitigt werden. Es gibt jedoch weitere Methoden: Eine davon ist Regularisation. Es wird eine Kostenfunktion eingeführt, die eine zu hohe Anpassung durch einen Faktor „bestraft“. Zudem können die Hyperparameter angepasst werden. Die Hyperparameter steuern den Trainingsprozess. Diese beschreiben beispielsweise die Topologie und die Größe des Modells [10].

2.2 GESCHICHTE

Die Historie beginnt mit dem „Vater der KI“ Alan Turing. Dieser schrieb 1936 seine Dissertation mit dem Titel „On Computable Numbers“. Darin wird zum ersten Mal eine Maschine beschrieben, die das grundlegende Modell für alle folgenden Computer bildet [11, S. 2].

Im Jahr 1943 stellten die Forscher Warren McCulloch und Walter Pitts die These auf, dass die Kernfunktionen des Gehirns wie Neuronen und Synapsen durch Logik und Mathematik erklärt werden könnten. Mit diesem Vorbild ließe sich ein Netz aufbauen, das Informationen verarbeiten und lernen kann [11, S. 5]. Ein weiterer Anhänger dieser These war Norbert Wiener. Im Jahr 1948 prophezeite er, dass ein Computer einen Großmeister im Schach schlagen würde, da eine Maschine während des Spiels weiter dazulernen würde [11, S. 6].

1950 veröffentlichte Alan Turing in einer weiteren Arbeit den berühmten Turing-Test. Dieser beurteilt die Ausgereiftheit eines Chatbots. Ein Mensch kommuniziert mit einem Computer und einem Menschen. Der bewertende Mensch versucht durch offene Fragen herauszufinden, welcher seiner Gesprächsteilnehmer der Mensch und welcher der Computer ist. Wenn keine Entscheidung getroffen werden kann, gilt der Turing-Test als bestanden [11, S. 2].

Sechs Jahre später wird der Begriff „Künstliche Intelligenz“ das erste Mal verwendet [2]. John McCarthy organisiert ein Forschungsprojekt zu diesem Thema in den Bell Labs [11, S. 6]. Dieses hat zum Ziel, Maschinen das Lernen beizubringen und diese dann Probleme lösen zu lassen [11, S. 7]. Das Ergebnis war der Logic Theorist, ein Computerprogramm, dessen Stärken im Lösen von mathematischen Problemen liegen. Die Software lieferte teilweise effizientere Lösungen als menschliche Mathematikerinnen und Mathematiker [11, S. 7].

Ab diesem Moment bis ins Jahr 1974 wurde auf dem Gebiet der KI viel gearbeitet. Die Entwicklung der Computertechnologie mit Kathodenstrahlröhren und integrierten Schaltkreisen wirkte ungemein verstärkend auf die Forschung [11, S. 8-9]. In dieser Zeit formten sich zwei polarisierende Theorien: Zum einen die Symbolische KI und zum anderen die dem menschlichen Gehirn angelehnten NN. Ersteres bedeutet, dass eine Intelligenz dennoch auf der klassischen Programmierung basieren sollte, also auf „Wenn..., dann...“-Abfragen [11, S. 10-11]. Zum anderen formte sich eine Theorie, die besagt, dass ein KI-Modell den Neuronen und Synapsen im menschlichen Gehirn nachempfunden werden und dieses Modell dann selbst lernen sollte [11, S. 11]. In den folgenden Jahren setzte sich die erste Theorie durch und auf dem Gebiet der NN wurden kaum Fortschritte gemacht. Dies änderte sich erst in den 1980-er Jahren, als der Einfluss von rückwärtsgerichteten Schleifen in NN erforscht wurde. Es entstanden RNN, die in der Lage waren, durch Vergleich von Ein- und Ausgabe ihre Gewichte selbstständig anzupassen [11, S. 14]. Dazu gehört die Arbeit von Geoffrey Hinton und Kollegen im Jahr 1986 [11, S. 16].

Der Fortschritt im Bereich der KI wurde ebenfalls vom technologischen Fortschritt des Internets begünstigt. Es liefert große Datenmengen, die zum Training benutzt werden. Hohe Übertragungsgeschwindigkeiten sorgen für eine schnellere Übermittlung der Daten. Außerdem können mit immer schnelleren Graphics Processing Units (GPU) die Daten schneller verarbeitet

werden [11, S. 17]. Dank der Forschung im vergangenen Jahrhundert können nun mit einer Unmenge an Daten, hoher Rechenkapazität und schneller Datenübertragung große Sprachmodelle trainiert und der Öffentlichkeit zur Verfügung gestellt werden.

Das früheste Modell von OpenAI ist GPT-3 aus dem Jahr 2020. Es nutzt die Transformer-Technologie und wurde mit 175 Millionen Gewichten trainiert. LaMDA ist ein Sprachmodell von Google. Es ist auch ein Transformer-Modell, aber noch nicht öffentlich zugänglich. PaLM ist ebenfalls ein Modell von Google. Es wurde mit 540 Milliarden Gewichten trainiert und nutzt eine parallelisierte Infrastruktur, sodass mehrere Anfragen gleichzeitig abgearbeitet werden können und die Rechenkapazitäten effizienter eingesetzt werden. Luminous ist ein europäisches KI-Modell des Heidelberger Start-ups Aleph Alpha. Es wurde auf der Hälfte der Parameter trainiert wie GPT-3, erzielt aber ähnlich gute Resultate. OPT-175B von Meta ist ein Open Pre-trained Transformer mit 175 Milliarden Parametern. Seine Ergebnisse sind ebenfalls mit denen von GPT-3 vergleichbar. Im Gegensatz zu den vorhergehenden Modellen ist Bloom ein Open-Source-Modell der Initiative Big Science. Es ist das Sprachmodell, das auf die meisten Sprachen trainiert ist. GPT-3.5 ist die Weiterentwicklung von GPT-3. ChatGPT ist dagegen nur der Chatbot, der auf GPT-3.5 basiert (siehe Abbildung sechs) [5].¹



Abbildung 6: Zeitleiste der LLMs [5].

¹ Der Stand der Informationen bezieht sich auf den Zeitpunkt des Verfassens der Thesis.

2.3 CHATGPT

ChatGPT basiert auf dem KI-Modell GPT-3.5. Es wurde im November 2022 vom Unternehmen OpenAI veröffentlicht und erfreut sich seitdem hoher Nutzungszahlen. Die Gründer von OpenAI sind Sam Altman und Elon Musk [12]. GPT-3.5 ist ein Sprachmodell, das darauf trainiert ist, Text zu erzeugen. ChatGPT wiederum wurde mit menschlichem Feedback für Dialog optimiert. Sprachmodelle werden mit großen Mengen an Daten, die von Menschen generiert und im Internet veröffentlicht wurden, trainiert. Dazu gehören auch menschliche Konversationen, um ChatGPT den Anschein von Menschlichkeit zu geben. Das Training von ChatGPT wurde 2021 abgeschlossen und es ist zusätzlich nicht mit dem Internet verbunden, daher hat es nur begrenztes Wissen über die Welt und Ereignisse nach dem Jahr 2021. Es kann also gelegentlich falsche oder auch gefährliche oder voreingenommene Inhalte generieren [13].

Innerhalb der ersten zwei Monate nach Veröffentlichung haben 100 Millionen Nutzerinnen und Nutzer den Chatbot getestet. Innerhalb von fünf Tagen hatte ChatGPT die Grenze von einer Million Nutzerinnen und Nutzern überschritten und ist damit die am schnellsten wachsende App der Welt [14]. Seitdem wurde ChatGPT immer wieder verbessert und in regelmäßigen Abständen neue Versionen herausgebracht [12]. ChatGPT ist mit Abstand das meistgenannte KI-Modell im Jahr 2022 auf Twitter [15, S. 2]. Die geschätzten täglichen Kosten für die kostenlose Testversion belaufen sich auf 100.000 USD [14].

Seit dem 23. März 2023 wird ChatGPT durch Plug-Ins von Drittanbietern ergänzt. Diese kommunizieren über ein Application Programming Interface (API) miteinander. Die ersten Plug-Ins stammen von Expedia, FiscalNote, Instacart, KAYAK, Klarna, Milo, OpenTable, Shopify, Slack, Speak, Wolfram und Zapier [16].

Funktionsweise

Um ChatGPT so nah wie möglich an einer menschlichen Chatpartnerin bzw. einem menschlichen Chatpartner zu halten, sind auch Menschen im Trainings- und Feedbackprozess von ChatGPT involviert. Zunächst wird eine Eingabe aus einem großen Datensatz an möglichen Eingaben ausgewählt (siehe Abbildung sieben links oben). Ein Mensch demonstriert die korrekte Antwort (siehe Abbildung sieben links Mitte). Diese Zuordnung von Eingaben und Antwort wird dann benutzt, um ein überwachtes Modell zu trainieren (siehe Abbildung sieben links unten). Dieses Modell bildet die Policy, auf der das eigentliche KI-Modell trainiert wird.

Dieselbe Eingabe wird nun dem überwachten Modell übergeben und mehrere Antworten generiert (siehe Abbildung sieben Mitte oben). Ein Mensch ordnet die Antworten von der hilfreichsten bis zur schlechtesten (siehe Abbildung sieben Mitte Mitte). Mit dieser Einordnung wird wiederum ein Reward-Modell trainiert (siehe Abbildung sieben Mitte unten).

Im letzten Schritt wird die Policy, die im ersten Schritt aufgestellt wurde, mit dem Reward-Modell verbessert. Dabei wird verstärkendes Lernen (Reinforcement Learning) verwendet. Eine neue Eingabe wird aus dem Datenset genommen und diese wird an das überwachte Modell gegeben (siehe

Abbildung sieben rechts oben). Die Antwort wird dem Reward-Modell übergeben. Es schätzt die Qualität der Antwort und gibt einen Reward r_k für die Antwort aus (siehe Abbildung sieben rechts unten). Mithilfe der Schätzung wird die Policy angepasst. Ist der Reward niedrig, muss die Policy angepasst werden. Ist der Reward hoch, wird die Policy an dieser Stelle noch verstärkt. Dafür wird der Proximal Policy Optimization (PPO)-Algorithmus verwendet. Dieser passt die Policy kontinuierlich an, je nach Reward. Die Veränderung der Policy darf nur in einer festgesetzten Entfernung von der originalen Policy stattfinden. Um die Policy qualitativ besser zu machen, vergleicht eine Value Function den möglichen Output der veränderten Policy mit dem der originalen Policy. Das macht die Kombination aus überwachtem Lernen und verstärkendem Lernen aus menschlichem Feedback bei ChatGPT bisher einmalig [17].

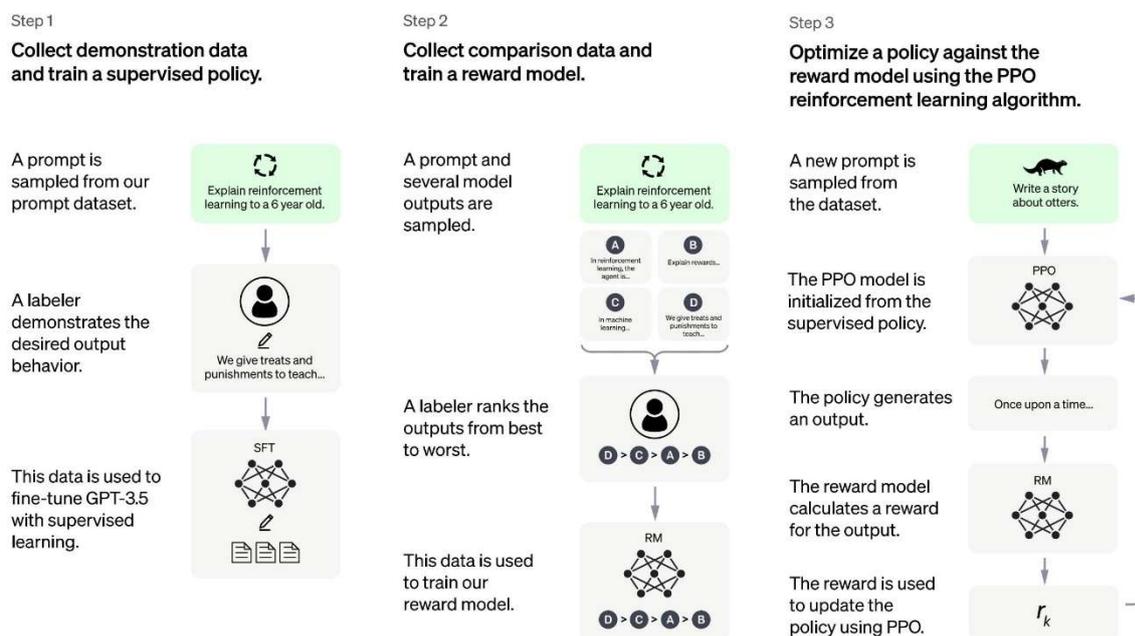


Abbildung 7: Vorgehensweise beim Training von ChatGPT [18].

3 EINFÜHRUNG IN DEN NEWSWORKFLOW

Der Journalismus wird oft als die vierte Säule, auf der die Demokratie ruht, beschrieben. Er ist unerlässlich für eine freie Meinungsbildung und -äußerung. Die grundlegenden journalistischen Tätigkeiten sind Recherchieren und Dokumentieren, Formulieren und Redigieren, Präsentieren und Organisieren. Eine Journalistin bzw. Journalist sammelt Informationen, geht einer Sache nach und deckt dabei Sachverhalte auf. Schließlich bringt die Berufsgruppe ihre Themen in jene Schrift- Audio oder Videoform, sodass dieses dem Thema entspricht und von seinen Nutzerinnen und Nutzern angenehm konsumiert werden kann. Als letzter Schritt wird das Werk veröffentlicht und in Kontext mit weiteren journalistischen Erzeugnissen gesetzt [19, S. 1]. Die Arbeitsfelder einer medienschaffenden Person sind entsprechend weit gefächert. Sie verlaufen von Zeitungen und Zeitschriften über Anzeigenblätter, Radio und Fernsehen, Online und Crossmedia, Nachrichtenagenturen und Informationsdienste bis hin zur Presse- und Öffentlichkeitsarbeit [19, S. 11].

3.1 RUNDFUNK- UND ONLINEJOURNALISMUS

Tageszeitungen – anders als Wochen- oder Sonntagszeitungen - sind alle veröffentlichten journalistischen Erzeugnisse, die mindestens zweimal pro Woche erscheinen und über ein universelles Themengebiet verfügen [19, S. 12]. Rund ein Fünftel der in Deutschland veröffentlichten Zeitungen sind Boulevardblätter wie beispielsweise Die Bild. Ein kleiner Anteil entfällt auf überregionale Abonnementzeitungen, z.B. Die Süddeutsche Zeitung. Der Rest sind lokale und regionale Tageszeitungen [19, S. 12-13]. Daher sind etwa ein Drittel der deutschen Zeitungsjournalisten im lokalen Bereich beschäftigt [19, S. 14]. Mit dem Aufkommen des Internets und damit neuen Verbreitungswegen für Nachrichten nimmt die Auflage der gedruckten Zeitungen stetig ab, trotzdem lesen im deutschsprachigen Raum noch etwa 60 Prozent der Bevölkerung eine gedruckte Tageszeitung [19, S. 13].

Nichtsdestotrotz steigen die Nutzungszahlen von Onlinemagazinen. Auf einer Webseite muss sich nicht an die Platzbeschränkungen gehalten werden wie auf einer gedruckten Zeitung. Die Nutzerin bzw. der Nutzer kann individuell wählen, wie tief er oder sie in ein Thema einsteigen will. Von der kurzen Übersichtsmeldung bis zu detaillierten Hintergrundrecherchen können diese die Speichermöglichkeit des Internets voll ausschöpfen. Außerdem bieten Onlinemagazine eine ständige Aktualisierungsmöglichkeit, sodass die Leserschaft immer neue Informationen abfragen kann. Die multimediale Präsentation von Themen und die direkte Feedbackmöglichkeit vom Lesepublikum zurück zur Presse ist ebenfalls neu für diese Medienform [19, S. 18].

Für medienschaffende Personen wurden hier neue Möglichkeiten und damit neue Arbeitsplätze geschaffen [19, S. 17]. Um die neuen Ausspielwege zu bedienen, muss nun eine medienschaffende Person mehrere Medien beherrschen. Meist haben diese ein Schwerpunktmedium und arbeiten mindestens einem anderen Medium zu [19, S. 20]. Das nennt sich crossmediales Arbeiten. Der Arbeitsablauf wird nun so angepasst, dass bei der Konzeption eines Beitrags gleich an mehrere

Ausspielwege und -formen gedacht wird. Also schreibt eine medienschaffende Person gleichzeitig mit ihrem Printartikel in der Zeitung eine Kurzfassung mit unterstützenden Bildern für die Sozialen Medien und ein Hintergrundfeature für den Onlineauftritt [19, S. 20].

Ein weiterer Vorteil des Onlinejournalismus ist die vereinfachte Beteiligung der Nutzerin bzw. des Nutzers. Unabhängig von Onlinemedien ist es bereits jeder Nutzerin bzw. jedem Nutzer möglich, ihren oder seinen eigenen Inhalt im Internet als Web Log (Blog) oder Video Log (Vlog) auf Sozialen Netzwerken zugänglich zu machen. Im Onlinejournalismus werden immer öfter Inhalte aus Sozialen Netzwerken als Inspiration, Meinung oder Ausspielweg für den eigenen Artikel verwendet [19, S. 18-19].

Auch im Rundfunk kommen neben dem linearen Fernsehen oder Radio immer weitere Angebote hinzu. Beiträge werden zusätzlich in den Mediatheken veröffentlicht, sodass Sendungen auch unabhängig von der Sendezeit konsumiert werden können. Die Mediatheken sind ebenfalls für die Nutzung von mobilen Geräten bis zum Fernseher zuhause ausgelegt, um zu jedem Zeitpunkt ein angenehmes Nutzungserlebnis zu geben [19, S. 16].

3.2 FORMATE

Die Darstellungsformen im Fernseh- und Onlinejournalismus unterscheiden sich voneinander, da sie aus zwei verschiedenen Medienformen stammen. Jedoch findet man immer wieder Übereinstimmungen. Im Filmischen muss im Gegensatz zum Schriftlichen darauf geachtet werden, dass sich ein Sachverhalt in Bildern darstellen lässt. Im Hörfunk muss darauf geachtet werden, dass ein Thema vollständig mit Worten beschrieben werden kann und dabei die optimale Länge hat, damit die Hörerin bzw. der Hörer nicht die Aufmerksamkeit verliert. Fernsehen ist also kein bebildeter Hörfunk und Hörfunk wiederum keine vorgelesene Zeitung.

Eine schnelle Darstellungsform von Fernsehnachrichten ist eine Nachricht im Film (NiF). Diese ist eine Kurznachricht, die durch Bilder verstärkt wird, mit einer Länge von 15 bis 30 Sekunden [20, S. 78]. Mischformen von Wortnachricht und NiF haben sich durchgesetzt. In der Praxis sieht man eine Nachrichtensprecherin bzw. einen Nachrichtensprecher im On, der eine Nachricht des Tages präsentiert. Während sie oder er den Sachverhalt erklärt, wechselt die Stimme ins Off. Zur Nachricht passende Bilder werden gezeigt. Beliebte Themen für NiF sind Katastrophen-, Wetter- oder Verkehrsnachrichten [20, S. 79]. Handelt es sich um ein Thema, das schwer in Bildern darzustellen ist, wie z.B. die Verabschiedung eines neuen Gesetzes, können auch Grafiken oder Archivmaterial verwendet werden, um den Beitrag bildlich zu füllen. Eine weitere Möglichkeit ist eine reine Wortmeldung mit einem zum Thema passenden Bild im Hintergrund. Dabei ist eine sprechende Person frontal im Bild zu sehen und diese liest die Nachricht vor [20, S. 81].

Das Statement kann Teil einer längeren Darstellungsform sein. Dabei wird der Original-Ton einer Augenzeugin bzw. eines Augenzeugen, einer betroffenen oder sonst wie beteiligten Person aufgenommen und ausgestrahlt [20, S. 103]. Dieses kommt oft innerhalb eines Berichts vor. Er gehört zu den am häufigsten genutzten Darstellungsformen in einer Nachrichtensendung und dauert länger als eine NiF, also bis zu zwei Minuten. Ein Bericht informiert tiefer als die bloße Nachricht. Er zeigt Hintergründe, Meinungen von betroffenen Personen und enthält auch die Einschätzung der berichterstattenden Person selbst. Diese steht meist am Ort des Geschehens und spricht frontal mit einem Handmikrofon in die Kamera [20, S. 122]. Eine Sonderform des Berichts ist der Live-Bericht. Dafür wird ein Gespräch in Echtzeit zwischen Moderatorin bzw. Moderator im Studio und der berichterstattenden Person unterwegs aufgebaut. Solche Live-Schalten geben dem Publikum das Gefühl, die aktuellen Informationen zu erhalten, mehr als es ein bloßer Bericht tun würde. Daher wird diese Darstellungsform auch angewandt, wenn sie eigentlich gar nicht nötig wäre [20, S. 128].

In der Reportage kommen dagegen statt der faktenbasierten Berichterstattung eher die persönlichen Eindrücke und Beobachtungen der berichterstattenden Person zum Ausdruck. Ein Beispiel dafür sind Reisereportagen. Sie sollen dem Publikum ein Gefühl von intensivem Erleben und Partizipation verschaffen [20, S. 139]. Eine Reportage unterscheidet sich von einem Bericht, in dem eine Reportage nicht den Anspruch hat, die Zuschauerin bzw. den Zuschauer über jedes Detail zu informieren. Sie kann sich die angenehmen und ansprechenden Aspekte aussuchen und beleuchten. Ein Bericht dagegen stellt den Sachverhalt vollständig und nicht beschönigt dar [20, S. 140].

Ein Magazinstück ist wiederum ein Beitrag, der länger als eineinhalb Minuten ist. Er unterscheidet sich insofern von der nüchternen Berichterstattung, als dass hier ein Seitenaspekt im Mittelpunkt steht, den man auf den ersten Blick nicht erwartet hätte. Ein Magazinstück liefert einen besonderen Zugang zu einem Thema und mehr Hintergrundwissen. Hier kann auch kreativer erzählt werden als es bei einem Bericht stattfinden würde [20, S. 136].

Die letzte Darstellungsform im Fernsehjournalismus gibt keine objektiven Informationen wieder, sondern gibt die persönliche Meinung der medienschaffenden Person zu einem Thema an. In einem Kommentar muss mündlich oder visuell klar werden, dass es sich hier um eine gefärbte Meinung handelt und nicht um Tatsachen. Die Zuschauerin bzw. der Zuschauer soll durch den Kommentar die Bewertung einer Tatsache durch den Journalisten bzw. die Journalistin erhalten. Die Person kann dann der Bewertung zustimmen oder sich wiederum eine eigene Meinung bilden. Der Kommentar beleuchtet also den Kern des dargestellten Problems und zeigt Zusammenhänge auf, mit dem sich das Publikum auseinandersetzen soll [20, S. 146-147].

Die Erzählformate im Onlinejournalismus sind an die des Printjournalismus angelehnt. Die Webreportage oder auch Webdoku ist beispielsweise ein klassischer Artikel, der mit Audio- und Videoelementen anschaulicher gemacht wurde. Dieser berichtet objektiv über eine Veranstaltung oder ein Ereignis. Features, also kurze Absätze mit Hintergrundinformationen zu einem bereits bearbeiteten Thema, können nun sowohl in Schriftform als auch in Form eines Videos oder einer Bilderstrecke erfolgen. Ähnlich wie im Fernsehjournalismus können hier außerdem Reportagen und Magazinstücke geschrieben werden. Diese sind jedoch in geringerem Umfang in der (Online-) Tagespresse vertreten, da sie das Augenmerk eher auf die Unterrichtung der Öffentlichkeit über tagesaktuelle Themen legt. Onlinemagazine, die auch thematisch eingeschränkt sein können, veröffentlichen dagegen öfter Reportagen und Magazinstücke. Da die Nutzerin bzw. der Nutzer des Onlinemediums selbst entscheidet, was sie oder er wann konsumiert, ist es wichtig, die Aufmerksamkeit der Leserschaft in die richtige Richtung zu steuern und damit die Zeitspanne zu vergrößern, die die Leserin bzw. der Leser mit dem Onlinemedium verbringt. Neue Formen wie interaktive Grafiken ziehen die Aufmerksamkeit des Lesepublikums an. Deren Weiterentwicklung, die Newsgames, sprechen die Lust zu Spielen an und vermitteln gleichzeitig Wissen. Bei Reisebeiträgen bietet es sich außerdem an, einen Beitrag in Virtual Reality zu erstellen. Dabei kann die Leserschaft in eine neue Welt eintauchen und schnell die Zeit vergessen [21, S. 119-123].

Neben erzählenden Formaten existieren auch Formate, die die Meinung der Presse widerspiegeln. Diese sind keine Sachtexte und vom persönlichen Empfinden des Medienschaffenden gefärbt. Dabei kann es sich um einen Kommentar zu einem Sachverhalt oder eine Kritik zu einer Veranstaltung handeln. Diese müssen nun auch nicht zwingend in Textform publiziert werden, sondern können auch in Video- oder Audioformaten gegeben werden. Es muss jedoch deutlich gemacht werden, dass es sich hier um eine Meinung handelt und nicht um objektive Fakten [21, S. 125].

3.3 ARBEITSABLÄUFE

Um eine fundierte Grundlage für die spätere Einbindung eines KI-Modells zu schaffen, sollen nun neben den inhaltlichen Eigenschaften des Journalismus auch in die prozessualen Schritte in Medienunternehmen eingeführt werden.

Themenauswahl

Jeder Fernseh- oder Radiobeitrag beginnt mit der Auswahl des Themas. Einerseits werden Themen von der Redaktion in Auftrag gegeben, andererseits muss die medienschaffende Person sich selbst Themen erarbeiten. Die geplanten Ideen werden in der Redaktionskonferenz besprochen, die wiederum eine Austauschmöglichkeit für Vorschläge darstellt. Daher sind die Voraussetzungen für eine gelungene Themenwahl eine genaue Kenntnis und thematische Ausrichtung der Sendung, für die produziert wird. Der bekannte Stil, den das Publikum gewohnt ist, muss beibehalten werden. Außerdem ist eine genaue Kenntnis der Nachrichtenlage sehr wichtig. Der Beitrag darf sich nicht auf veraltete oder schwammige Informationen stützen. Schließlich muss das Thema auch für die medienschaffende Person – gerade im Fernsehjournalismus – realisierbar sein. Da im Fernsehen ein Großteil der Informationen über Bilder vermittelt wird, müssen auch Themen gewählt werden, die sich in Bildern abbilden lassen [20, S. 256-257].

Recherche

Schon vor der Themenauswahl recherchiert die medienschaffende Person zu ihrem Thema. Die Recherche wird in drei Richtungen durchgeführt. Zunächst in die thematische Richtung. Sie oder er eignet sich ein umfassendes Wissen an und markiert sich Informationen für seinen späteren Beitrag. Als nächstes erfolgt eine Bild- bzw. Tonrecherche. Die medienschaffende Person besichtigt seinen zukünftigen Dreh- und/oder Interviewort und macht sich mit den örtlichen Gegebenheiten vertraut. Mögliche Bildausschnitte und Kamerastandorte werden dokumentiert als Anhaltspunkt für den späteren Dreh. Eventuell können auch hier schon einzelne Aufnahmen gemacht werden. Die dritte Richtung der Recherche ist die Organisation. Hier muss abgeklärt werden, welches Equipment, welche Genehmigungen oder welche Reisevorbereitungen benötigt werden [20, S. 259].

Planung

Gemeinsam mit der Recherche fertigt die medienschaffende Person zunächst ein Exposé an. Dabei handelt es sich um die erste schriftliche Fixierung des Beitrags. Das Exposé kann zu einem Treatment erweitert werden. Dieses behandelt das Thema ausführlicher und präzisiert die Aussage. Je nach Größe und gegebener Zeit wird sich eventuell auf das Exposé beschränkt [20, S. 263].

Um einen strukturierten und ressourcensparenden Dreh durchführen zu können, wird der Beitrag in Blöcke unterteilt und diese entweder in Worten in einem Filmplan oder in Skizzen in einem Storyboard festgehalten. Bei kleineren Beiträgen wird auch eventuell auf den Filmplan verzichtet. Ein Produktionsplan gibt dem Filmteam zusätzlich einen genauen Rahmen, was wann zu drehen ist. Der Plan ist mit Uhrzeiten, Orten und der jeweiligen Szene versehen [20, S. 267].

Archiv

Um nicht das gesamte Bildmaterial für einen Beitrag neu produzieren zu müssen, kann auf Archivmaterial zurückgegriffen werden. Gerade bei Themen wie Nachrufen, Jubiläen, Geburtstagen und Rückblicken ist es unerlässlich, bereits gedrehtes und archiviertes Material zu verwenden. Neutrale Bilder zur Unterstreichung eines Themas, während eine Stimme aus dem Off Informationen gibt, müssen ebenfalls nicht neu gedreht werden. Dazu gehören auch Situationen, in denen es zu aufwendig, teuer oder verboten ist, zu drehen. Dies können Motive aus anderen Ländern bis hin zu Gerichtsverhandlungen sein [20, S. 282-283].

Postproduktion

Ist der Dreh vollendet, wird das Material unverzüglich in den Speicher der Rundfunkanstalt übertragen. Dieses wird in das richtige Format transkodiert und mit Metadaten versehen. Das Material wird gesichtet und zu einem Beitrag zusammengeschnitten. Mithilfe von Musik und Hintergrundgeräuschen des Drehorts wird das Bild lebendiger gestaltet. Die medienschaffende Person schreibt Text für eine Stimme aus dem Off. Entweder spricht diese den Text selbst oder es wird eine weitere Stimme benötigt. Durch Grafiken wie Bauchbinden, Quellenangaben oder erklärende Animationen wirkt der Beitrag professioneller und ansprechender. Als letzter Schritt vor der Auspielung muss der Beitrag von der Redaktion abgenommen werden. Die Redaktionsmitglieder entscheiden, ob der Beitrag dem Stil und dem Anspruch der Sendung genügt [20, S. 311-350].

Das Aufgabenfeld einer Onlinejournalistin bzw. eines Onlinejournalisten ist größer als das eines Printjournalisten. Das liegt daran, dass ein solcher nicht nur eine Medienform zu bearbeiten hat. Sie oder er arbeitet crossmedial. Das bedeutet, dass die Person einen Beitrag für das Internet konzipieren, schreiben, bearbeiten, die Inhalte akustisch wie visuell aufbereiten und präsentieren können muss. Dies setzt Kenntnis des journalistischen Handwerks voraus. Trotzdem sollte eine Onlinejournalistin bzw. ein Onlinejournalist zusätzlich auch mit einem Mikrofon, Bildausschnitten, Infografiken und dem Schnitt von Videosequenzen vertraut sein [21, S. 2-3].

Medienspezifische Arbeit

Dazu zählen neben klassischen Tätigkeiten wie Recherche, Konzeption und Präsentation von Inhalten auch die Moderation von Foren und die Organisation von Kommunikations- und Informationsflüssen in der Redaktion [21, S. 4]. Der Arbeitsablauf erfolgt ähnlich wie im Fernsehjournalismus. Die Themenwahl ist hier auch für den Fortlauf des Beitrags entscheidend. Gleichzeitig werden hier auch die Medienformen ausgewählt, für die das Thema am besten geeignet ist. Bei der Recherche muss ebenfalls darauf geachtet werden, dass für alle Medienformen Informationen sowie mögliche Bilder und O-Töne geplant werden. Es können hier in kleiner Form Bild- und Tonaufnahmen gemacht werden. Diese haben jedoch einen geringeren Umfang als im Fernsehjournalismus. Neben dem Schnitt des Materials wird ebenfalls ein Artikel verfasst. Dieser wird auch eventuell von einer Stimme eingelesen, sodass die Nutzerin bzw. der Nutzer diesen auch

hören kann. Schließlich werden die Erzeugnisse auf allen Medienkanälen ausgespielt und der Nutzerin bzw. dem Nutzer präsentiert.

Kommunikation

Der zweite Bereich umfasst alle Arten von Kommunikation, die eine Onlinejournalistin bzw. ein Onlinejournalist während seines Arbeitsalltags erlebt. Dazu gehört die persönliche Kommunikation während der Recherche. Das kann der Austausch mit Archivarinnen und Archivaren, Bibliothekarinnen bzw. Bibliothekaren, Fachleuten oder betroffenen Personen sein. Weiterhin kann die Kommunikation auch mit dem Publikum über Soziale Medien erfolgen. Entweder über die Veröffentlichung von Inhalten oder die Beantwortung von Fragen und Kommentaren. Außerdem ist bei der Kooperation mit den Teammitgliedern, der Redaktion und den Auftraggeber*innen ein hohes Maß an Kommunikation wichtig, da sonst wichtiger Input verloren geht oder die Ansprüche an einen Artikel weit auseinander gehen können [21, S. 8]. Um dem entgegen zu wirken, werden regelmäßige Teambesprechungen innerhalb der Redaktion durchgeführt [21, S. 35].

Redaktionsmanagement

Aufgaben in diesem Feld sind die Konzeption eines eigenen Stils des Verlagshauses, die Neuordnung von Arbeitsabläufen sowie die Organisation von ständig wechselnden Projektteams [21, S. 8]. Eine Onlinejournalistin bzw. ein Onlinejournalist ist also befähigt in sämtlichen Medien mit Online-Auftritten zu arbeiten. Seien es klassische Medien, die ihre Artikel zusätzlich zu einer Printausgabe online veröffentlichen oder reine Online-Medien. Möglich sind auch Mischformen aus Journalismus und Werbung sowie Presse- und Öffentlichkeitsarbeit online [21, S. 11].

3.4 INFRASTRUKTUR

Um ebenfalls eine systemintegratorische Perspektive auf den Journalismus zu erhalten, ist eine Einführung in die Infrastruktur innerhalb eines Medienunternehmens interessant.

Rundfunkjournalismus

In Rundfunkanstalten werden Nachrichten in einem Newsroom erzeugt. Dabei handelt es sich um eine Großraum-Redaktion mit mehreren Arbeitsplätzen. Neben Journalistinnen und Journalisten arbeiten hier auch Grafiker*innen und Mediengestalter*innen [20, S. 397]. In diesem Raum erfolgt die Verarbeitung von Rohmaterial zu einem sendefertigen Beitrag. Es wird also Content produziert. Content ist die Gesamtheit aller digital zur Verfügung stehenden Formate [22, S. 14]. Um Content strukturiert zu speichern und schnell wieder aufrufen zu können, verwenden Rundfunkanstalten ein Content Management System (CMS). Das ist ein Softwaresystem, das für das Erstellen, Verwalten und Zusammenführen von Inhalten entworfen wurde [22, S. 14].

Content existiert nach einem sogenannten Content Life Cycle. Dieser erstreckt sich von seiner Erstellung bis hin zur Archivierung (siehe Abbildung acht). Der Kreislauf steigt bei der Erstellung des Contents ein. Das ist meist der Dreh von Material. Das Rohmaterial wird geschnitten und mit weiteren Medientypen wie beispielsweise Grafiken oder einer Audiospur zusammengebracht. Vor der Veröffentlichung wird der Content von der Redaktion kontrolliert und freigegeben. Gleichzeitig mit der Veröffentlichung wird das Material ebenfalls im Archiv eingelagert und kann wieder zur Recherche von neuem Content verwendet werden. Der publizierte Content wird ausgestrahlt und löst eventuell im Publikum Feedback in Form von Briefen an die Redaktion, Anrufen oder Kommentaren aus. Diese können neue Gesichtspunkte anstoßen und wiederum zur Recherche gebraucht werden [22, S. 45].

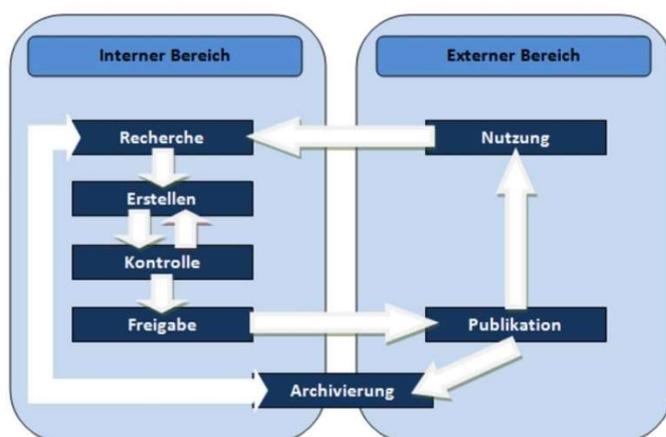


Abbildung 8: Content Life Cycle in einem Medienunternehmen [22, S. 123].

Um den Content bei diesem Lebenszyklus zu unterstützen, ist ein CMS entsprechend in eine Eingabe-, eine Verwaltungs- und eine Publikationskomponente unterteilt. Über die Eingabekomponente werden Inhalte eingepflegt, erstellt und bearbeitet. Bereits bei der Eingabe kann auch eine Transkodierung und Qualitätskontrolle erfolgen. Damit Content entsteht, werden einerseits direkt beim Eintritt in den Speicher der Rundfunkanstalt strukturelle Daten

zugeordnet, die Metadaten. Eine Essenz ist nur die reine Video- Audio- oder Textdatei, Content besteht aus einer Essenz und ihren Metadaten [23, S. 71].

Metadaten sind textuelle Informationen, die eine Essenz beschreiben und sie identifizierbar machen. Metadaten ermöglichen das Suchen, das Finden, die Nutzung und die Verwaltung von Ressourcen [23, S. 94]. Metadaten lassen sich in die folgenden Kategorien einteilen. Es wird zunächst zwischen technischen, administrativen und beschreibenden Metadaten unterschieden. Technische Metadaten sind beispielsweise die Dateigröße, das Dateiformat, die Auflösung oder die Bittiefe. Administrative Metadaten beschreiben die Rechte und Nutzungsvorbehalte der Essenz. Beschreibende Metadaten wiederum geben den Inhalt der Essenz an. Dazu gehören Titel, Inhaltsangabe und auftretende Motive und Personen.

Gleichzeitig wird auch zwischen statischen und dynamischen Metadaten unterschieden. Statische Metadaten bleiben für die Dauer der gesamten Essenz erhalten, wie z.B. der Titel. Dynamische Metadaten dagegen sind nur für einen Teil der Essenz gültig. Das könnte eine Beschreibung sein, wann welcher Teil der Besetzung in einer Essenz zu sehen ist [23, S. 98]. Metadaten bestehen also aus einer Identifikation, der inhaltlichen Beschreibung und technischen Parametern der Essenz.

Im Folgenden übernimmt die Verwaltungskomponente. Diese sorgt dafür, dass die Metadaten auf dem neusten Stand und nicht redundant vorhanden sind. Metadaten begleiten eine Essenz von der Aufnahme bis hin zum Feedback des Publikums nach der Ausstrahlung. Auf ihrem Weg in der Wertschöpfungskette können Metadaten dazukommen oder ungültig werden. Beispiele dafür sind das Zusammenfügen mehrerer Essenzen im Schnitt oder Transkodierung. Diese fließen zu einer einzigen zusammen und erhalten gemeinsame neue Metadaten. Durch Transkodierung kann sich das Format ändern, was auch in den Metadaten dokumentiert wird. Außerdem werden Zugriffsrechte und Verantwortlichkeiten für den Content definiert.

Dadurch, dass Metadaten den gesamten Produktionsablauf abdecken, müssen diese für alle Essenzen die gleiche Struktur haben. Es existieren innerhalb von beschreibenden und technischen Metadaten weitere Unterkategorien, die für alle Essenzen gelten. Damit können die Kategorien einzelner Essenzen verglichen und es kann nach Kategorien gefiltert werden. Daher bedarf es auch einer gewissen Granularität, um die Essenzen möglichst präzise zu beschreiben. Innerhalb der Rundfunkanstalt werden die Essenzen und Metadaten zwischen mehreren Systemen und Speichern weitergegeben. Das bedeutet, dass jedes System die Metadatenstruktur verarbeiten können muss. Die Schnittstellen zwischen den Systemen müssen auf jene Weise interoperabel gestaltet werden, sodass die Metadaten im selben Format verwendet werden können. In gleicher Weise muss auch gewährleistet werden, dass auf lange Sicht Kategorien angepasst oder weitere Kategorien hinzugefügt werden können, um beispielsweise auf technische Veränderungen reagieren zu können. Hier muss dann in gleichem Maße die Interoperabilität zwischen den Systemen bleiben [23, S. 102].

Wird der Content schließlich veröffentlicht, ist die Interoperabilität am meisten gefragt, denn der Content wird über unterschiedliche Sendewege ausgespielt. Dies können das klassische lineare Fernsehen, Radio, Online, in der Mediathek oder Soziale Medien sein. Der Content muss also für

jede Medienform aufbereitet und manuell oder automatisch publiziert werden [22, S. 46-47]. Als letzter Schritt im Life Cycle wird der Content archiviert und kann für spätere Beiträge als Recherchematerial dienen.

Onlinejournalismus

Um den Weg von Recherche über Produktion bis hin zur Veröffentlichung effizient zu gestalten, ohne dass Informationen oder Dateien verloren gehen, bedarf es im Onlinejournalismus ebenfalls eines CMS. Dieses verwaltet den Content. Hier sind mit Content die Rechercheergebnisse und Erzeugnisse in jeglicher Form in Vorbereitung sowie bereits veröffentlichte Erzeugnisse gemeint. Eine Onlinejournalistin bzw. ein Onlinejournalist kann hier bei der Recherche auf schon bestehende Erzeugnisse zurückgreifen und ihre bzw. seine bisherigen Ergebnisse sichern. Soll nun ein Artikel, Audio- oder Videobeitrag vor der Veröffentlichung abgenommen werden, muss dieser nicht erst von der Onlinejournalistin bzw. vom Onlinejournalisten zur Redaktion geschickt werden. Die Redaktion kann mit ihrem Zugang zum CMS bereits die Arbeit sichten, um den Arbeitsablauf zu beschleunigen [21, S. 36].

Innerhalb des CMS bedarf es wieder Metadaten zur Beschreibung der Essenzen. Ähnlich wie beim Fernsehjournalismus helfen die Metadaten bei der Identifizierung und der Organisation der Essenzen. Es wird auch zwischen technischen und beschreibenden Metadaten unterschieden. Jedoch müssen hier gerade die technischen Parameter weitergefasst werden, da hier neben Video- und Audiomaterial auch Artikel und Grafiken gespeichert sind. Die Metadaten müssen also auch diese Medienformen abbilden und beschreiben können. Mögliche beschreibende Metadaten für einen Textartikel sind Autorin bzw. Autor, Thema, Erstellungszeitpunkt und Quellen. Bei Video- und Audiometadaten werden außerdem statische und dynamische Metadaten erfasst. Bei einer Grafik beispielsweise nur statische. Die Metadaten müssen zusätzlich strukturiert angelegt sein, sodass sie bei der Übergabe über eine Schnittstelle hinweg noch immer die richtigen Informationen angeben. Bei der Suche oder Filterung nach Schlagworten sollen dann anhand der Informationen in den richtigen Kategorien die Essenzen angezeigt werden [21, S. 107].

Eine Weiterentwicklung des CMS ist das Web CMS. Es wird für das Erzeugen und das Management von großen Mengen von Hypertext Markup Language (HTML)-Inhalten inklusive der eingebundenen Mediadaten verwendet. Damit können die Artikelseiten eines Onlinemagazins schnell und ohne Programmierkenntnisse befüllt werden [22, S. 25]. Die Veröffentlichung kann hier auch entweder manuell oder automatisch erfolgen. Die Besonderheit ist, Inhalt und Layout sind hier voneinander getrennt definiert. So kann der Inhalt immer wieder – je nach Aktualität – ausgetauscht werden, während das kohärente Erscheinungsbild des Onlinemagazins erhalten bleibt.

Um nun Inhalte online zur Verfügung zu stellen, müssen jeweils die Startseite sowie die von dort verlinkten Seiten befüllt werden. Dabei wird zwischen dynamischen und statischen Seiten unterschieden. Dynamische Seiten enthalten ständig wechselnde Inhalte, z.B. einen Kalender für Konzerttermine, der regelmäßig aktualisiert wird. Auch Seiten, die mit einem Personalisierungsalgorithmus für die einzelne Nutzerin bzw. den einzelnen Nutzer bestückt werden,

sind dynamische Seiten. Statische Seiten dagegen behalten ihren Inhalt bei, solange sie nicht verändert werden. Dazwischen gibt es semidynamische Seiten. Dabei handelt es sich beispielsweise um die Rubrik Rezepte. Die einzelnen Rezepte sind statisch. Ihre Anordnung ändert sich aber je nach Saison [21, S. 39].

Teil des CMS ist außerdem das redaktionseigene Archiv. Es umfasst alle bereits publizierten Beiträge. Diese sind für den gemeinsamen Zugriff freigegeben und sind eine umfangreiche Quelle an Informationen für neue Beiträge. Neben den archivierten Beiträgen enthält es außerdem externe Informationsquellen wie Archive und Bibliotheken sowie Namen und Kontaktdaten von freien Autorinnen und Autoren [21, S. 36].

3.5 KOMMUNIKATION MIT DEM PUBLIKUM

Journalismus lebt zu einem gewissen Teil auch vom Feedback des Publikums. Im Printjournalismus kann die Leserschaft ihre Meinung durch Briefe – auch elektronisch - an die Redaktion, die auch im Medium selbst abgedruckt werden können, zum Ausdruck bringen. Ebenfalls über das Telefon ist Feedback möglich, kann jedoch nicht abgedruckt werden.

Im Onlinejournalismus kommen nun weitere Kommunikationsformen hinzu. Über Instant Messaging-Dienste per Text, Audio und Video können in Echtzeit Gespräche geführt werden. Außerdem kann nicht nur paarweise kommuniziert werden, sondern auch mit mehreren Gesprächsteilnehmerinnen und -teilnehmern. Daher sind diese Dienste nicht nur für die Kommunikation zwischen Redaktion und Nutzerin bzw. Nutzer geeignet, sondern auch für die Kommunikation Peer-to-Peer. Dies fördert einen intensiveren Informationsaustausch [21, S. 160].

Kommentarfunktion

Onlinemedien bieten in den meisten Fällen eine Funktion, um Kommentare zu einem Beitrag zu hinterlassen. Bei manchen Anbieterinnen und Anbietern muss man einen Account anlegen, um seine Meinung kundzutun, bei manchen nicht. Muss kein Account angelegt werden, haben Nutzerinnen und Nutzer durch die Anonymität selten Konsequenzen zu erwarten, wenn ihre Kommentare unhöflich, unsachlich oder sogar diskriminierend sind. Durch die Pflicht bei der Registrierung einen Klarnamen anzugeben, hat das Onlinemedium die Möglichkeit, Spielregeln zum respektvollen Umgang zu veröffentlichen, trotzdem muss immer noch darauf geachtet werden, dass diese eingehalten werden. Die häufigste Form der Filterung ist eine Stoppwortliste. Auf dieser Liste befinden sich Beleidigungen und diskriminierende Formulierungen. Sobald eines der Worte oder Wortgruppen in einem Kommentar auftaucht, wird dieser entweder sofort gelöscht oder einem Redaktionsmitglied zur Kontrolle vorgelegt. Daher ist der Chat nur tagsüber erreichbar. Diese Form nennt sich unmoderierter Chat [21, S. 161].

Ein moderierter Chat wiederum steht nur zu einem bestimmten Zeitpunkt innerhalb der Arbeitszeit der Mitarbeiterinnen und Mitarbeiter zur Verfügung. Zu einem vorher festgelegten Thema ist eine Expertin bzw. ein Experte eingeladen, die oder der aus seinem Arbeitsgebiet erzählt und Fragen beantwortet. Die Online-Moderation begrüßt die Mitglieder der Community und stellt das Thema vor. Nach dem Bericht der sachverständigen Person sammelt die Moderatorin bzw. der Moderator die Fragen des Publikums und leitet sie geordnet an diese Person weiter. Hier ist es gewünscht, dass sich eine sachliche Diskussion zwischen dem Sachverständigen und den Mitgliedern entwickelt. Diese Gespräche mit Expertinnen und Experten werden meist aufgezeichnet und können von den Nutzerinnen und Nutzern auch später eingesehen werden [21, S. 172, 21, S. 161].

Chatbots

Da Telefon und Chat in Echtzeit beantwortet werden sollten, um der Anfragenstellerin bzw. dem Anfragensteller eine schnelle Antwort zu geben, müsste sich immer ein Mensch um die Fragen kümmern. Da aber nicht immer zu jeder Tageszeit eine im Journalismus tätige Person die Zeit hat,

werden gerade die Chatanfragen von einem Chatbot beantwortet. Der Bot gibt seinem Gesprächsteilnehmer die benötigten Informationen in Textform.

Es gibt mehrere Kategorien von Chatbots. Sie unterscheiden sich nach ihrem Zweck. Der klassische Chatbot übernimmt die Funktion des Kundenservices. Er beantwortet Fragen mit Wissen, das er gespeichert hat. Außerdem kann ein Chatbot auch einen Vertriebskanal darstellen. Man kann also mithilfe des Bots Waren bestellen und bezahlen. Ein Chatbot kann auch als Concierge dienen und die Nutzerin bzw. den Nutzer Schritt für Schritt durch die eigene Organisation leiten. So kann ein Chatbot sie oder ihn auch durch das Archiv eines journalistischen Mediums führen. Hier wird durch den Chatbot ein neuer Medienkanal aufgebaut. Schließlich kann ein Chatbot auch dafür sorgen, dass die Nutzerin bzw. der Nutzer ein Erlebnis im Bereich der Marke erhält. Dies geht auch in die Richtung der Gamification, also die Einbindung von Spielelementen [24, S. 5-6].

Dabei erhält der Bot eine Persönlichkeit, die möglichst nah an der Zielgruppe ausgerichtet ist. Dies kann durch ein bestimmtes Vokabular oder die Nutzung von Umgangssprache erreicht werden [21, S. 135]. Außerdem sollte Auftreten und Sprache an den Stil des Mediums angepasst werden. Für das Empfinden der Nutzerin bzw. des Nutzers ist wichtig, dass die Konversation mit dem Bot so wirkt, als würde sie oder er mit einem Menschen chatten. Daher sind natürliche Sprache und auch das Einräumen von Schwächen nötig. Der Chatbot sollte außerdem konsistente Antworten geben, da sonst missverständliche Informationen an die Nutzerin bzw. den Nutzer herausgegeben werden können [24, S. 19-23].

Soziale Medien

Unter dem Begriff Soziale Medien werden alle Arten von Plattformen, die Raum für von Nutzerinnen und Nutzern generierte Inhalte (User generated Content) bieten, zusammengefasst. Dazu gehören die Technologien und die Geschäftsmodelle [21, S. 151]. Seit der Einführung der Sozialen Medien ist es für die Nutzerin bzw. den Nutzer ein Leichtes selbst Beiträge im Internet zu veröffentlichen. Die Formate reichen von Weblogs (Blogs) über Fotos bis hin zu Videos. Die Beiträge können dann von anderen Nutzerinnen und Nutzern konsumiert, bewertet und kommentiert werden [19, S. 18-19].

Vor der Einführung der Sozialen Medien haben ausschließlich die Redaktionen entschieden, welche Inhalte der Öffentlichkeit zugänglich gemacht wurden. Heute können Nutzerinnen und Nutzer selbst auswählen, zu welchen Themen sie Informationen veröffentlichen möchten oder über welche Themen sie informiert werden wollen. Die Aufgabe der Redaktion als Schleusenwärterin, die darüber entschieden hat, was publiziert wurde und was nicht, gibt es so nicht mehr [21, S. 151].

Community Management

Durch die Vernetzung von Menschen auf Sozialen Medien entstehen sogenannte Communities. Eine Community besteht aus einer Gruppe registrierter Nutzerinnen und Nutzer, die sich für ein bestimmtes Thema interessieren. Eine Community kann sich aus den Followern eines großen Benutzungskontos bilden. Sie schreiben gegenseitig Kommentare unter die Beiträge, geben Rückmeldung und unterstützen sich bei Fragen und Problemen [21, S. 168]. Communities können

sich zufällig in Zusammenhang mit einem großen Benutzungskonto bilden oder können gezielt angelegt werden. Für Online-Medien ist es interessant, seine Nutzerinnen und Nutzer durch ein schlüssiges Konzept zum Feedback und damit zum Mitwirken anzuregen. Nicht zuletzt wird die Community durch eine professionelle Redaktion und Moderation aktiv gehalten [21, S. 170].

Partizipation

Das Feedback der Nutzerinnen und Nutzer durch Kommunikation mit der Redaktion wird nur noch durch die Partizipation im Journalismus übertroffen. Die Leserschaft verlässt also die passive Rolle der rezipierenden Person und wechselt in die aktive Rückmeldung gebende Rolle [21, S. 145]. Die Tiefe der Leseartizipation wird in fünf Schichten unterteilt:

In der ersten Schicht liefert die Leserschaft Daten, die die Grundlage für einen Artikel bilden. Diese Daten stammen aus dem Alltag der Leserschaft und geben einen Einblick in ihre Lebenswirklichkeit. Die Leserinnen und Leser handeln dann auf einen Aufruf im jeweiligen Online-Medium hin. Die Bereitschaft an der Datenerhebung mitzuwirken, wird meist daraus generiert, dass es sich um ein kontroverses oder emotionales Thema handelt. Es kann auch so weit gehen, dass den Leserinnen und Lesern Geräte und Sensoren ausgeteilt werden, mit denen sie Messwerte erheben. Ein Beispiel dafür sind Abstandssensoren für das Fahrrad. Damit wurde gemessen, wie groß der Abstand zwischen Auto und Fahrrad bei einem Überholvorgang ist. Die Daten werden von medienschaffenden Personen gesammelt, ausgewertet und in einem Beitrag aufbereitet. Dieses Verfahren nennt sich Crowd-Sourcing [21, S. 147].

In der zweiten Schicht generiert und liefert die Leserschaft Material in Form von Text, Bild, Audio und Video. Dies kann gerade für Ereignisse und Themen verwendet werden, bei denen die Presse keine Zugangsmöglichkeit oder -gelegenheit hat. Der Aufruf zur Zusendung des Materials geht von der Presse aus. Daraufhin sammelt und prüft die Redaktion das Material, ob es plausibel und qualitativ brauchbar erscheint. Dann wird es in den Beitrag eingebaut. Durch die zusätzliche Quelle an Material wird dies Add-on Reporting genannt. [21, S. 147].

Die dritte Schicht beginnt mit der selbstständigen Recherche und Produktion von Ideen für Beiträge der Leserinnen und Leser, ein sogenannter Citizen-Journalismus. Auf dieser Stufe wird die Arbeit der Leserschaft jedoch noch redaktionell begleitet. Das bedeutet, dass die Online-Redaktion die Arbeitsergebnisse prüft und dann selbst in einen fertigen Beitrag wandelt. Dies setzt eine starke Zusammenarbeit zwischen der Redaktion und ihrer Leserschaft voraus. Wichtig sind hier die regelmäßige Absprache und Teilnahme an Redaktionssitzungen [21, S. 147].

Die vorletzte Stufe der Leseartizipation beinhaltet die Bereitstellung von Plätzen für Blogs oder Podcasts der Leserinnen und Leser. Diese Plätze werden Bloghouse genannt. Sie sind als partizipatives Format gekennzeichnet und befinden sich in einem eigenen Bereich des Online-Mediums. Das Publikum kann hier selbstständig Beiträge veröffentlichen. Es hat also die volle Kontrolle darüber, welche Beiträge zu welchem Zeitpunkt für das restliche Publikum zur Verfügung stehen sollen [21, S. 147].

Mit der fünften und letzten Stufe der Partizipation können die Leserinnen und Leser nun ganze Beiträge selbst erstellen. Diese sind bereits in Inhalt und Technik publikationsreif. Eine explizite Kennzeichnung, dass es sich hier um von Nutzerinnen und Nutzern generierte Inhalte handelt, gibt es nicht mehr. Die Beiträge können frei wählbare Themen behandeln, um die aktuellen Sorgen und Diskussionen innerhalb der Leserschaft abzubilden. Es können aber auch Themenvorschläge aus der Redaktion kommen. Diese Art von Selbstständigkeit für die Citizen-Journalistin bzw. den Citizen-Journalisten ist attraktiv, birgt jedoch für die Redaktion ein gewisses Risiko [21, S. 147].

4 ANALYSE DER IST-WORKFLOWS

Um die Anknüpfungspunkte im Workflow, an denen eine KI eine Unterstützung bieten kann, zu identifizieren, werden zunächst die herkömmlichen Arbeitsabläufe in Medienunternehmen betrachtet. Diese sind crossmedial angelegt, es wird also gleichzeitig für Fernsehen, Hörfunk sowie eine Onlinezeitung produziert. Daher ist dieses Kapitel nicht nach Medienformen strukturiert, sondern nach den auszuspielenden Formaten. Es wird jedoch auf die Besonderheiten einer jeden Medienform eingegangen.

Den Anfang macht der Ingest-Workflow. Er ist zwar kein produktiver Workflow, jedoch ist er Teil jedes weiteren produktiven Workflows und für die Arbeit in Medienunternehmen unverzichtbar. Dann folgen die klassischen Nachrichten, die mehrmals täglich aktualisiert werden. Im Fernsehen sind dies z.B. die Tagesschau oder ZDF heute. Ein weiterer Workflow ist die Breaking News-Situation. Hier werden ebenfalls Nachrichten in ein Format gebracht, das veröffentlicht werden kann, jedoch in viel kürzerer Zeit und weniger strukturiert. Im Vordergrund steht die schnelle Informationsweitergabe an das Publikum. Den Abschluss bilden Informationsbeiträge wie der Wetter- oder Börsenbericht. Diese Art der Berichterstattung wird in einem eigenen Abschnitt behandelt. Im Onlinejournalismus behandelt jeder Artikel ein Thema. Es ist jedoch auch möglich, dass ein schneller Überblick über die Nachrichten in Videoform dort veröffentlicht wird. Die Beantwortung von Nachrichten der Nutzerinnen und Nutzer ist ebenfalls kein produktiver Workflow, trotzdem wird dadurch ein Service am Publikum gegeben. Die vollständigen Workflows befinden sich auf dem Datenträger in Kapitel drei.

4.1 INGEST

Rundfunkanstalten und Magazine besitzen durch ihr längerfristiges Bestehen eine große Menge an Informationen, Texten, Bildern und Videos. Diese sind jedoch kaum mit beschreibenden Metadaten versehen. Gerade bei Material in Videoformat ist eine manuelle Annotation sehr ressourcenintensiv, da die Beschreibung szenen- oder sogar sekundenbasiert erfolgen muss. Trotz der Verschlagwortung findet kaum Vernetzung zwischen den Assets statt. Diese ist jedoch für eine effektive Recherche, Kuration und personalisierte Empfehlung an die Nutzerin bzw. den Nutzer nötig [9, S. 33, 25, S. 33]. Der Ingest-Workflow ist kein produktiver Workflow. Er sorgt viel mehr für die präzise Ablage von Assets im CMS. Ohne die korrekte Beschreibung durch Metadaten kann ein Asset im Speicher nicht identifiziert werden und ist somit für das Medienunternehmen nicht nutzbar.

Der Prozess wird von jeder Art einkommendem Material ausgelöst. Er ist beendet, wenn die Essenz mit beschreibenden, administrativen und technischen Metadaten in einer ausreichenden Qualität versehen ist, sodass sie im Archiv gefunden und verwendet werden kann. Der Vorgang sollte je nach Umfang der Essenz nur wenige Minuten in Anspruch nehmen. Er wird bei Bedarf ausgelöst und ist nicht terminiert. Ein möglicher Auslöser ist die Übertragung von Material, die von einem Kamerateam bei einer Außenreportage angefertigt wurde. Dieses wird im Regelfall direkt vom Drehort oder, wenn dies nicht möglich ist, von einem mitgebrachten Datenträger in den Speicher

des Medienunternehmens übertragen. Das gleiche Verfahren wird bei Studioaufzeichnungen angewandt. Wird eine beendete Aufzeichnung im Archiv gespeichert, muss sie eine Möglichkeit der Identifizierung erhalten. Wird externes Material eines anderen Medienunternehmens in den eigenen Speicher eingespeist, wird der Prozess ebenfalls durchlaufen. Außerdem wird eine neue Essenz in den Speicher gebracht, wenn im Schnitt aus einzelnen Essenzen eine neue Essenz erzeugt wird. Sie wird bei der Speicherung als eigenständige, neu eingebrachte Essenz behandelt. Dann werden die technischen Metadaten aus der erhaltenen Datei oder der Medienbegleitkarte extrahiert. Hier muss ebenfalls geprüft werden, ob das Material den technischen Richtlinien des Medienunternehmens entspricht.

Schließlich übernimmt eine Archivmitarbeiterin bzw. ein Archivmitarbeiter oder eine Journalistin oder ein Journalist die manuelle Extraktion von beschreibenden Metadaten. Dann werden die Metadaten und die Essenz gemeinsam referenziert und gespeichert, sodass sie bei der Suche immer in Kombination gefunden werden.

4.2 TÄGLICHE NACHRICHTEN

Die Nachrichten werden täglich zur gleichen Zeit veröffentlicht. Sie informieren über die Ereignisse des Tages. Der Aufbau besteht im Fernsehen aus einer Moderatorin oder einem Moderator im Studio, die oder der die Nachrichtenbeiträge ankündigt und Informationen verbal übermittelt. Die Beiträge sind kurz und kompakt, sodass mehrere Themen pro Sendung behandelt werden können. Im Hörfunk werden die gesamten Nachrichten von einer Sprecherin bzw. einem Sprecher gelesen. Es können O-Töne mitverwendet werden. Nachrichten im Onlinejournalismus müssen sich an keine Sendetermine halten, sodass diese sofort nach Fertigstellung und Abnahme veröffentlicht werden.

Eckdaten

Die Produktion einer Nachrichtensendung für das Fernsehen beginnt bis zu zwölf Stunden vor Sendebeginn. Fängt die Sendung um 20 Uhr an, muss spätestens um acht Uhr morgens mit der Vorbereitung begonnen werden. Der Prozess wird von einer zeitlichen Markierung ausgelöst. Für jeden Tag, an dem die Sendung ausgestrahlt werden soll, muss der Prozess durchlaufen werden. Die Sendung wird zum regelmäßigen Termin auf mehreren Kanälen ausgespielt.

Die Nachrichtenproduktion im Hörfunk verläuft ähnlich zum Fernsehen. Beide können sich in ihren Abläufen gegenseitig begünstigen. Die Audioaufnahmen eines Kamerateams können ebenfalls ausschnittsweise im Radio verwendet werden. Der Prozess, eine Nachrichtensendung für den Hörfunk zu produzieren, wird auch von einer zeitlichen Markierung ausgelöst. Die Arbeit an der Sendung muss in einem ausreichenden zeitlichen Abstand zum Sendezeitpunkt anfangen. Dieser ist jedoch geringer als im Fernsehen. Das Ziel ist eine ausspielbereite Sendung, die die Redaktion als solche bestätigt hat. Der Prozess wird aufgrund seiner kürzeren Durchlaufzeit mehrmals täglich durchlaufen, damit zu verschiedenen Uhrzeiten immer neue Nachrichten veröffentlicht werden können. Planbare Ereignisse wie Veranstaltungen und Feiertage können hier berücksichtigt werden.

Der Onlinejournalismus produziert Inhalte für mehrere Medienformen. Es werden üblicherweise Beiträge in Text-, Bild-, Audio- und Videoformat veröffentlicht. In wenigen Fällen kommen auch interaktive Grafiken und Newsgames zum Einsatz. Entsprechend können hier Synergien mit Fernsehen und Hörfunk genutzt werden. Der Prozess, einen Nachrichtenbeitrag zu produzieren, wird durch einen Auftrag der Redaktion an die Journalistin bzw. den Journalisten ausgelöst. Das Ziel des Prozesses ist ein Beitrag in mehreren Medienformen, der fertig zur Veröffentlichung ist, sowie dessen Ausspielung. Im Onlinejournalismus gibt es keine festen Sendetermine, jedoch sollen tagesaktuelle Meldungen möglichst schnell veröffentlicht werden, um die höchsten Lesezahlen zu erreichen. Der Prozess wird für jeden Beitrag neu angestoßen, daher läuft der Prozess mehrmals täglich und parallel ab. Dadurch ist der Prozess jedoch auch planbar. Beiträge für angekündigte Ereignisse wie z.B. Veranstaltungen können auch schon vorbereitet und Veröffentlichungsplätze dafür freigehalten werden.

Planung und Themenfindung

Eine jede Nachrichtensendung bzw. ein jeder Nachrichtenartikel beginnt mit der Planung eines Grundgerüsts. Dieses wird als Sendeplan im Redaktionssystem angelegt. Darin wird festgelegt, wie viele und welche Themen in welchem Format für das Publikum aufbereitet werden. Es wird außerdem entschieden, in welcher Tiefe ein Thema dargestellt wird. Je nachdem, ob ein längerer NiF verwendet oder das Thema nur in einem kurzen Aufsager präsentiert wird, erweckt es einen unterschiedlichen Eindruck beim Publikum. Dieser Teil wird entweder von der Redaktion oder einer eigenen Planungsabteilung durchgeführt, abhängig von der Größe des Medienhauses. Die Themen werden oft vom weltpolitischen Geschehen vorgegeben. Hier werden eintreffende Meldungen von Presseagenturen oder eigenen Außenreporterinnen und Außenreportern verwendet. Für planbare Ereignisse wie z.B. einer Großveranstaltung können die Themen auch schon mehrere Tage oder Wochen vor Beginn der Produktion feststehen. Die finale Auswahl trifft jedoch die Redaktion. Bei längeren Sendungen ist es auch möglich, dass die Journalistinnen und Journalisten selbst Vorschläge zu Themen abgeben. Werden Liveschalten zu externen Reporterinnen und Reportern benötigt, werden diese auch in dem Prozessschritt festgelegt. Die Aggregation und Auswahl der Themen kann gemeinsam durch Fernseh- und Hörfunkjournalistinnen und -journalisten erfolgen, um Synergieeffekte zu nutzen. Die Themen werden im Anschluss an mehrere Journalistinnen und Journalisten gegeben.

Der Onlinejournalismus ist nicht an die Struktur einer Sendung gebunden. Themen werden einzeln in Beiträge gegossen. Die Ideen für Themen stammen aus der Tagesaktualität, können jedoch auch durch eigene Ideen der Journalistinnen und Journalisten ergänzt werden. Die Themenvergabe wird im Rahmen einer Redaktionskonferenz durchgeführt. Dort kommt es zu einem regen Austausch an Ideen. Neue Ideen können auch erst in der Runde entwickelt werden.

Beitrag

Ein Beitrag ist Teil einer Nachrichtensendung. Dieser entsteht durch die Verteilung der Themen in der Redaktionskonferenz. Das Ziel ist einen inhaltlich und technisch fertigen Beitrag zu erstellen. Der Prozess wird mehrmals täglich innerhalb des Produktionsprozesses einer Sendung durchlaufen. Die einzelnen Durchführungen laufen jedoch parallel mit mehreren Journalistinnen und Journalisten ab. Da die Anfertigung eines Beitrags ein Unterprozess der Produktion einer Sendung ist, darf dieser nicht länger dauern als der Überprozess.

Wird der Beitrag durch die eigenen Mitarbeiterinnen und Mitarbeiter produziert, wird zu Beginn eine textuelle Skizze angefertigt. Je länger der Beitrag, desto tiefer sollte die Beschreibung sein. Diese wird ebenfalls von der Redaktion abgenommen. Im nächsten Schritt wird das Thema recherchiert. Die Recherche erfolgt nicht nur inhaltlich, sondern auch nach Ansprechpartnerinnen und Ansprechpartnern sowie Aufnahmegelegenheiten. Außerdem muss ein technisches Team zusammengestellt werden. Abhängig von den beabsichtigten Medienformen müssen entsprechendes Equipment sowie Expertinnen und Experten im Team vorhanden sein. Letztere führen die Aufnahmen durch und ingestieren das gewonnene Material in den Speicher der Rundfunkanstalt.

Dabei wird es von einer Reporterin bzw. einem Reporter begleitet, welche bzw. welcher die Übersicht über die zu drehenden Szenen hat und Interviews führt. Jedes Material, das in Vorbereitung und Dreh entsteht, wird gespeichert und steht allen Mitarbeiterinnen und Mitarbeitern des Beitrags zur Verfügung, um auch paralleles Arbeiten und damit Zeitersparnis zu ermöglichen. Bei besonders aktuellen Ereignissen ist statt eines Beitrages auch eine Liveschleife zur Reporterin bzw. zum Reporter denkbar.

In der Rundfunkanstalt wird das Rohmaterial gesichtet, geordnet und zu einem Beitrag zusammengeschnitten. Dies führt aus Zeitgründen nicht zwangsläufig die Journalistin bzw. der Journalist, die noch zuvor beim Dreh unterwegs war, durch. In der Postproduktion wird für eine Fernsehsendung bei Bedarf eine Stimme aus dem Off unter das Bildmaterial gelegt. Dafür muss ebenfalls die Journalistin bzw. der Journalist die Texte verfassen und diese entweder selbst oder von einer weiteren Stimme einsprechen lassen. Für eine Hörfunksendung wird sich nur auf das Audiomaterial konzentriert. Anschließend wird der Beitrag erneut inhaltlich sowie technisch abgenommen, damit er in die Nachrichtensendung eingebaut werden kann. Die Speicherung des fertigen Beitrags erfolgt im CMS, sodass er in der Sendung veröffentlicht werden kann. Die Zeit zwischen Fertigstellung des Beitrags und Beginn der Sendung kann mitunter sehr gering sein.

Produktion

Die Produktion läuft durch die verschiedenen Medienformen jeweils unterschiedlich ab. Für eine Fernsehsendung wird ein audiovisuelles Studio benötigt. Das technische Personal im Studio erhält das Grundgerüst der Sendung über das Redaktionssystem. Es arbeitet mit unterschiedlichen grafischen Benutzungsoberflächen, sodass jeder Berufsgruppe genau die Informationen angezeigt werden, die sie für ihre Arbeit benötigt. Vor dem Beginn der Studioproduktion muss die Planung der Sendung erneut von der Redaktion abgenommen werden.

Im Anschluss wird eine Probe durchgeführt, um für Moderation, Redaktion und technisches Personal einen Eindruck der Sendung zu erzeugen und Schwachstellen aufzudecken. Danach folgt die Aufzeichnung der eigentlichen Sendung. Vorher produzierte Beiträge – seien sie eigen- oder fremdproduziert – sowie Liveschleifen werden zwischen die Moderationsteile eingefügt. Die Aufzeichnung der Sendung wird im Speicher der Rundfunkanstalt abgelegt, sodass sie beispielsweise auch über Social Media oder die Mediathek ausgespielt werden kann.

In einem Hörfunkstudio muss weit weniger vorbereitet werden, da hier nur Audioaufnahmen gemacht werden. Die Radiosprecherinnen und Radiosprecher sprechen die Nachrichten live zur vollen Stunde nach einem vorbereiteten Sprechertext ein.

Die Produktion im Onlinejournalismus erfolgt zunächst textuell. Es wird eine Skizze des geplanten Beitrags angefertigt. Diese wird von der Redaktion abgenommen, bevor mit einer tieferen Recherche und dem anschließenden Verfassen des Beitrags begonnen wird. Jeder Fortschritt wird im CMS des Onlinemagazins gespeichert. Die Produktion erfolgt gleich in mehreren Formaten. Ein Artikel in Textform mit Bildern wird verfasst. Dieser wird wiederum von einer Stimme eingelesen und wird

ebenfalls auf der Webseite des Onlinemagazins angeboten. Nutzerinnen und Nutzer können den Artikel dann beispielsweise unterwegs anhören. Dafür können Interviewaussagen als Audioaufnahme verwendet, transkribiert oder Bilder und kurze Sequenzen aus dem Videomaterial entnommen werden. Außerdem kann das Thema auch in einer Podcast-Folge weitergeführt werden, die auf Streaming-Plattformen veröffentlicht wird. Eine Präsentation in einem längeren Videoformat zieht ebenfalls das Publikum an. Ein Video ist leichter zu konsumieren als ein Text, daher lassen sich Nutzerinnen und Nutzer die Inhalte auch gerne audiovisuell präsentieren. Der Textbeitrag wird mit Bildern angereichert und in ein Layout gesetzt. Die fertigen Produktionen werden vor der Veröffentlichung nochmals von der Redaktion abgenommen. Entspricht der Beitrag nicht den Erwartungen der Redaktion, muss die Journalistin bzw. der Journalist nacharbeiten.

Kurzüberblick

Ein Kurzüberblick ist eine Sonderform aus dem Fernsehjournalismus. Er soll kürzer als die primäre lange Nachrichtensendung sein und eine Zusammenfassung der Nachrichtenlage geben. Eine solche Sendung dauert ungefähr zwei Minuten. Der Prozess entspringt in der Produktion der Nachrichtensendung und endet ebenfalls mit einem ausspielfertigen Produkt. Daher muss auch eine gewisse Anzahl an Stunden vor der Veröffentlichung mit der Produktion des Kurzüberblicks begonnen werden. Der Prozess wird täglich durchlaufen und ist damit ein planbarer Prozess.

Um einen Kurzüberblick zu erstellen, müssen aus der bestehenden Themenauswahl für die lange Nachrichtensendung erneut die wichtigsten Informationen und Bilder ausgewählt werden. Im Schnitt werden gemeinsam mit den Beiträgen für ausgewählte Themen auch kürzere Beiträge geschnitten. Über einen zentralen Speicher wird auf das Beitragsmaterial, das schon für die Zusammenstellung der langen Nachrichtensendung verwendet wurde, zugegriffen. Der Kurzüberblick benötigt jedoch eine erneute Studioproduktion, allerdings in kleinerem Umfang als die Nachrichtensendung. Das Format kann im linearen Fernsehen, in der Mediathek und in eventuell nochmals verkürzter Form auf Videoplattformen ausgespielt werden.

Informationsbeitrag

Ein Informationsbeitrag wird außerhalb der vollen Stunde im Hörfunk gesendet. Er vertieft ein Thema aus den Nachrichten. Bei einem solchen Beitrag kann es sich um einen nüchternen Bericht oder sogar ein freundschaftliches Gespräch unter den Radiomoderatorinnen und Radiomoderatoren handeln, bei dem weitergehende Informationen zu einem Thema aus den Nachrichten vermittelt werden. Auch die persönlichen Meinungen der Radiomoderatorinnen und Radiomoderatoren können hier zum Ausdruck kommen.

Der Auslöser für diesen Produktionsprozess ist ein Auftrag der Redaktion. Das Ziel ist ein sendefertiger Beitrag, inhaltlich sowie technisch. Der Fertigstellungszeitpunkt wird vom geplanten Sendeplatz bestimmt. Das Auftreten des Prozesses folgt keinem festen Rhythmus, er wird bei Bedarf ausgeführt. Ein hoher Bedarf kann jedoch zur parallelen Ausführung durch mehrere Journalistinnen und Journalisten führen.

Der Arbeitsablauf beginnt mit einer Ideenskizze, die für den weiteren Verlauf von der Redaktion bestätigt werden muss. Sie wird durch Recherche und kreative Arbeit weiter ausgeführt. Stehen die Texte für den Beitrag fest, wird der Beitrag live zum Sendezeitpunkt eingesprochen. Bei der Studioproduktion wird der aufgenommene Beitrag im CMS gespeichert und wird wieder von dort abgerufen, damit er mehrmals ausgespielt werden kann. Eventuelle Musik und Geräusche werden im Vorfeld aus dem Archivmaterial gesucht und per Knopfdruck in den Beitrag eingefügt.

Feature

Ein Feature wiederum ist ein Format aus dem Onlinejournalismus. Es liefert weitere Hintergrundinformationen zu einem Thema, das bereits in einem Artikel behandelt wurde. Es ist kürzer als der ursprüngliche Beitrag und wird oft in Kombination mit diesem veröffentlicht. Der Produktionsprozess eines Features ist also eng mit dem des Artikels verbunden, da dieser innerhalb der Produktion des Artikels ausgelöst wird. Es handelt sich also um einen Unterprozess. Ziel des Prozesses ist einen informativen (Unter-)Beitrag zu verfassen, der fertig zur Veröffentlichung ist, und diesen zu veröffentlichen. Der Ablauf erfolgt mehrmals täglich und parallel, jedoch immer in Kombination mit der Erstellung eines Artikels. Dadurch handelt es sich ebenfalls um einen plan- und vorbereitbaren Prozess.

Direkt bei der Recherche zu einem Beitrag wird differenziert, welche Informationen im Beitrag und welche im Feature veröffentlicht werden. Oberflächliche für das Verständnis des Beitrags wichtige Informationen fließen in den Beitrag mit ein. Tiefere und weiterführende Informationen bleiben dem Feature vorbehalten. Durch die Planung des Beitrags und des zugehörigen Features wird also auch die Recherchetiefe festgelegt. Durch die Speicherung der Rechercheergebnisse im CMS können diese für alle Mitarbeiterinnen und Mitarbeiter sichtbar zwischen Beitrag und Feature verteilt werden.

Gemeinsam mit der Produktion des Artikels wird auch das Feature verfasst. Es wird ebenfalls von einer Stimme eingesprochen, sodass die Nutzerin bzw. der Nutzer sich auch diesen Teil des Beitrags anhören kann. Eventuell kann das Feature auch in Videoform veröffentlicht werden. Synergieeffekte von Fernsehen und Radio können auch hier von Vorteil sein. Das Feature wird wie der Beitrag in ein Layout gesetzt. Vor der Veröffentlichung erfolgt die Abnahme durch die Redaktion. Das Feature kann entweder gleichzeitig mit dem Beitrag oder mit zeitlichem Abstand veröffentlicht werden.

Ausspielkanäle

Die Ausspielung der journalistischen Erzeugnisse erfolgt zum geplanten Sendezeitpunkt auf den für die Medienform geeigneten Kanälen. Für audiovisuelle Beiträge sind es das lineare Fernsehen, die Mediathek, Videoplattformen sowie in verkürzter Form Soziale Medien. Hörfunkbeiträge können ebenfalls über den linearen Hörfunk, die Mediathek, das Webradio und Soziale Medien veröffentlicht werden. Im Onlinejournalismus ist der primäre Veröffentlichungsplatz die Webseite des Onlinemagazins, jedoch sind auch Videoplattformen und Soziale Medien denkbar.

Für die Ausspielung auf dieser Vielzahl von Kanälen muss auch der Beitrag in verschiedenen technischen Formaten sowie Längen im CMS vorliegen. Soziale Medien können beispielsweise nur Bilder und Videos mit einer bestimmten Größe, Länge oder Bildseitenverhältnis annehmen. Dies muss vor der Ausspielung berücksichtigt und angepasst werden. Eventuell können die Ausspielzeitpunkte auf verschiedenen Kanäle variieren. So kann eine Nachrichtensendung bereits in der Mediathek oder in Sozialen Medien veröffentlicht werden, bevor diese im linearen Fernsehen erscheint.

4.3 BREAKING NEWS

Im Breaking News-Fall soll eine aktuelle Nachricht möglichst schnell in ein sendefertiges Format gebracht werden. Dabei handelt es sich meist um (welt-)politische Ereignisse oder Katastrophenmeldungen, die das Interesse der gesamten Bevölkerung einnehmen.

Eckdaten

Der Prozess beginnt also mit dem Eintreffen einer neuen Meldung im Medienunternehmen. Die Meldung stammt von einem anderen Medienunternehmen, einer Presseagentur, der Regierung, der Polizei oder aus Sozialen Medien. Ziel ist ein kurzer, informativer und sendefertiger Beitrag. Dieser unterbricht das lineare Sendeprogramm bzw. verdrängt Artikel auf der Titelseite. Der Prozess wird bei Bedarf gestartet und soll in möglichst kurzer Zeit ablaufen. Hier bleibt keine Zeit eigene Dreharbeiten durchzuführen. Es wird fremd produziertes Material anderer Sender oder einer Presseagentur übernommen. Es ist möglich, dass der Breaking News-Prozess mehrfach hintereinander abläuft, wenn es sich um eine anhaltend wichtige Nachrichtenlage handelt.

Im Onlinejournalismus gibt es jedoch die Besonderheit, in diesem Falle einen Live Ticker zu verwenden. Dabei handelt es sich um eine Artikelseite auf der Webseite eines Onlinemagazins, auf der in zeitlichem Abstand Kurzmeldungen zu einer Nachrichtenlage veröffentlicht werden. Die Kurzmeldungen sind jeweils mit der Uhrzeit versehen, zu der sie hinzugefügt wurden, um die zeitliche Abfolge darzustellen. Da der Prozess durch eine einkommende Meldung ausgelöst wird, ist der Prozess zeitlich nicht planbar und es können auch keine konkreten Themen vorbereitet werden.

Arbeitsablauf

Nach dem Eintreffen einer neuen Meldung wird das Thema von der Redaktion an mehrere Journalistinnen und Journalisten übergeben. Das eintreffende Live-Material wird so schnell wie möglich gesendet. Ein vorüberlegtes Grundgerüst wie bei den terminierten Nachrichten gibt es nicht. Es wird spontan die passendste und einfachste Darstellungsform gewählt. In dieser kurzen Zeitspanne zwischen Themeneingang und Veröffentlichung ist keine Zeit für die eigene Recherche vorgesehen. Es kann auch der Fall sein, dass während das initiale Material ausgespielt wird, erst Entwürfe und Texte für die weiteren Darstellungsformate verfasst werden.

Wird im Fernsehen eine Moderation benötigt, müssen Moderationstexte verfasst und ein Studio vorbereitet werden. In Rundfunkanstalten gibt es für den Breaking-News-Fall meist ein kleines Ausweich-Studio, um nicht direkt ein vollständiges Studioteam anzufordern oder eine gerade stattfindende Sendung zu stören. Tritt der Breaking News-Fall beispielsweise in den frühen Morgenstunden auf, zu denen sich wenig technisches Personal im Studio befindet, kann dieses spontan über Bereitschaft an den Arbeitsplatz gerufen werden. Trotzdem sind die Abläufe im Studio so weit automatisiert, dass außer der Moderation wenig Personal benötigt wird (Selbstfahrer-Studio). Bilder für den Hintersetzer sowie Grafiken werden nicht selbst erstellt, sondern aus dem eintreffenden Material genommen. Es können auch Symbolbilder aus dem Archiv verwendet werden.

Der Breaking News-Fall im Radio verläuft ähnlich wie im Fernsehen. Ziel ist ebenfalls, möglichst schnell Informationen an das Publikum zu geben. Dadurch, dass hier nur die auditive Komponente bedient wird, ist die Berichterstattung flexibler. Für die Radiomeldung werden zunächst die Stimmen der Moderatorinnen und Moderatoren verwendet, um Fakten zu vermitteln. Dazu kommen eventuell O-Töne von Betroffenen, die entweder voraufgezeichnet oder live eingespielt werden. Hier wird ebenfalls kein Ablauf der Sendung festgelegt. Das Ende der Programmunterbrechung steht nicht vorher fest. Eine solche Situation bedarf eines effizienten Teams und eingespielter Arbeitsabläufe.

Für den Onlinejournalismus wird meist direkt ein textueller Beitrag aus dem gelieferten Material übernommen, editiert und veröffentlicht. Dieser wird im Nachgang von einer Stimme eingesprochen. Für eine redaktionelle Abnahme bleibt keine Zeit. Bei der besonderen Form des Live Tickers wird dieser Prozess, eine kurze Text- oder Audiomeldung auszuspielen, immer wieder durchlaufen, sobald sich eine neue Meldung einstellt.

Ausspielung

Die Ausspielung erfolgt möglichst schnell und auf möglichst vielen Kanälen, da die Nachrichten für einen Großteil der Bevölkerung relevant sind. Die produzierten Meldungen sind generell kurz. Trotzdem werden je nach Ausspielkanal verschiedene Längen und Formate benötigt. Im linearen Fernsehen kann ein Einschub ungefähr fünf Minuten gesendet werden. Im Hörfunk benötigt man wenige Sätze und auf Sozialen Medien wiederum ein Bild oder ein kurzes Video.

4.4 INFORMATIONSNACHRICHTEN

Der Zweck dieser Nachrichtenform ist die pure Übermittlung von Informationen. Der Beitrag bedarf keiner kreativen Ausarbeitung oder stilistischen Besonderheiten. Wetter-, Verkehrs- und Börsennachrichten folgen in regelmäßigen Veröffentlichungsabständen demselben Schema.

Eckdaten

Wetter- und Verkehrsnachrichten sind Teil der stündlichen Nachrichtensendung im Hörfunk und müssen für jede Ausstrahlung aktualisiert werden. Einige Radiosender senden sogar zu jeder halben Stunde Verkehrsnachrichten. Der Prozess wird also stündlich bis halbstündlich durchlaufen und muss zu jeder Ausstrahlung vollendet sein. Ziel des Prozesses ist ein kurzer verständlicher Beitrag in sendefertigem Format. Der Prozess ist also planbar und kann aufgrund seines standardisierten Aufbaus stark vorbereitet werden. Im Onlinejournalismus sind dagegen Wetter- und Börsennachrichten prominent. Diese folgen bei jeder Aktualisierung demselben visuellen Schema. Vorbereitete Grafiken müssen nur noch mit Texten und Symbolen befüllt werden.

Arbeitsablauf

Zu jeder Iteration des Prozesses erhält die Rundfunkanstalt rohe Daten von der Verkehrsüberwachung, der Börse und den Wetterdiensten. Diese werden ständig aktualisiert. Im Verkehr muss zunächst gefiltert werden, ob es sich um eine Meldung handelt, die im Rahmen der nächsten Nachrichten weitergegeben wird, oder um eine dringende Nachricht, die sofort gemeldet werden muss. Dies sind besondere Gefahrensituationen wie z.B. eine Falschfahrerin bzw. ein Falschfahrer. Für Meldungen, die bis zur nächsten Ausstrahlung der Nachrichten warten können, werden Notizen verfasst und Grafiken kreiert. Dafür existieren bereits vorgefertigte Bausteine im CMS, welche nur noch mit den aktuellen Informationen bestückt werden müssen. Radiomoderatorinnen und Radiomoderatoren lesen die Wetter- und Verkehrsinformationen und fassen diese während des Sprechens direkt in ganze Sätze. Dies kann für die ZuhörerIn bzw. den Zuhörer natürlicher wirken. Die Nachrichten werden zum Sendezeitpunkt im Radio live gesprochen. Für dringende Nachrichten wird das aktuelle Programm unterbrochen und die Nachricht sofort ausgesprochen.

Ausspielung

Wetter- und Verkehrsnachrichten sprechen eine breite Zielgruppe an, daher werden diese regelmäßig auf möglichst vielen Kanälen veröffentlicht. Neben der Ausspielung im Audioformat über das lineare Radio und in Textform werden die Nachrichten auch im Webradio, Onlinemagazin und der Mediathek veröffentlicht. Die Texte und die Audiobeiträge, die bei der Live-Produktion aufgenommen wurden, werden von einem zentralen Speicher übertragen, in ein sendefertiges Format transferiert und ausgespielt.

4.5 RÜCKMELDUNG VON NUTZERINNEN UND NUTZERN

Die Beantwortung der Rückmeldungen von Nutzerinnen und Nutzern ist kein produktiver Prozess. Es handelt eher sich um einen Service am Publikum. Die Möglichkeit, Feedback zu geben, fördert die Partizipation der Nutzerin bzw. des Nutzers und das Interesse an journalistischen Erzeugnissen. Die Nachrichten der Nutzerinnen und Nutzer können auf ganz unterschiedlichen Kanälen eintreffen. Dabei werden zwei Gruppen unterschieden: Formen, bei denen eine direkte Antwort für die Annahme der initialen Nachricht nötig ist, wie beispielsweise ein Telefonanruf und Formen, die keine direkte Antwort benötigen, da die Nachricht „hinterlassen“ wird, wie beispielsweise eine Nachricht über einen Messenger-Dienst.

Der Prozess der Rückmeldung von Nutzerinnen und Nutzer wird von der Ankunft einer Nachricht bzw. durch einen Anruf ausgelöst. Da eine große Zahl an Nachrichten von Nutzerinnen und Nutzern täglich bei Medienunternehmen eintreffen, läuft dieser Prozess sehr häufig und parallel ab. Ein Durchlauf sollte möglichst kurz andauern, damit die Nutzerin bzw. der Nutzer ein Gefühl von Anwesenheit der Journalistinnen und Journalisten erhält. Da eine Journalistin bzw. ein Journalist während ihrer oder seiner Arbeitszeit nicht ununterbrochen Zeit hat, Anfragen von einer Nutzerin oder einem Nutzer zu beantworten, gibt es in großen Medienhäusern eine eigens dafür geschaffene Abteilung sowie zeitlich Beschränkungen bei Telefonanrufen. Ziel des Prozesses ist, eine zufriedenstellende Antwort zu verfassen und an die Nutzerin oder den Nutzer abzuschicken. Für häufig gestellte Fragen können inhaltlich Vorbereitungen in Form von Textbausteinen getroffen werden. Der Prozess wird nur bei Bedarf durchlaufen, ist daher nicht zeitlich planbar. Man kann jedoch davon ausgehen, dass durch die hohe Auftretenshäufigkeit ständig mit dem Auslöser dieses Prozesses gerechnet kann.

Der Ablauf besteht aus dem Eintreffen der Nachricht, dem Beantworten der Nachricht und der Rückmeldung an die Nutzerin bzw. den Nutzer. Abhängig vom Inhalt der Nachricht kann es jedoch vorkommen, dass die Nachricht direkt gelöscht bzw. der Anruf abgebrochen wird. Dies passiert bei unhöflichen, unsachlichen und beleidigenden Anfragen. Eine oder mehrere sachlich formulierte Nachrichten zum gleichen Thema können jedoch innerhalb des Medienunternehmens eventuell einen Anstoß für Erneuerung und Verbesserung bilden.

5 ANFORDERUNGEN AN EINE SYSTEMLANDSCHAFT MIT KÜNSTLICHER INTELLIGENZ

Nach den Einführungen sowie der Definition der vorherrschenden Workflows werden nun die Anforderungen an ein systemisches Konstrukt mit einem KI-Modell im Newsworkflow dargelegt. Die Anforderungen sind in drei Kategorien unterteilt: Rechtliche und ethische Anforderungen, funktionale Anforderungen und organisatorische Anforderungen. Rechtliche Anforderungen werden von Gesetzen zur KI oder zur Berichterstattung bestimmt. Die ethischen Anforderungen sind Best Practices aus der Forschung zum verantwortungsvollen Umgang mit KI. Die funktionalen Anforderungen entspringen aus regelmäßigen Tätigkeiten im Newsworkflow und Case Studies. Die organisatorischen Anforderungen bestimmen die Prozesse sowie die Anforderungsprofile für die Mitarbeiterinnen und Mitarbeiter. Alle Anforderungen werden in einer Matrix in Anhang eins zusammengefasst.

5.1 RECHTLICHE UND ETHISCHE ANFORDERUNGEN

Das oberste Ziel und Gebot, das bei jeglicher Regulierung von KI im Medienumfeld zu beachten ist, ist das Recht auf freie Meinungsäußerung in Artikel fünf Absatz eins des Grundgesetzes der Bundesrepublik Deutschland (GG). Eingriffe in dieses Recht werden nur geduldet, wenn einer der vorangegangenen Artikel des GG verletzt wird [26, S. 445]. „Die Pressefreiheit und die Freiheit der Berichterstattung durch Rundfunk und Film werden gewährleistet. Eine Zensur findet nicht statt.“ [27, Art. 5 Abs. 1]. Diese Freiheit ist Grundlage für die freie Meinungsbildung innerhalb der Demokratie, für Selbstbestimmung und freie und fundierte Wahlentscheidungen. Dieser Prozess darf nicht durch Einflussnahme gestört werden [28, S. 9]. Deshalb ist die Ausübung der Berichterstattung in der Presse und im Rundfunk mit einer großen Verantwortung verbunden. Im Journalismus tätige Personen sind der individuellen Unterrichtung der Öffentlichkeit und dem Ansehen der Presse verpflichtet. Journalistische Erzeugnisse müssen für die allgemeine Öffentlichkeit der Bundesrepublik Deutschland zugänglich sein [26, S. 444, 29, Präambel].

Die Aufgabe des öffentlich-rechtlichen und privaten Rundfunk ist laut dem Medienstaatsvertrag der Länder die „freie[-] individuelle[-] und öffentliche[-] Meinungsbildung sowie [die] Medienvielfalt“ [30, Präambel]. Der Staatsvertrag regelt außerdem die Produktion, Diffusion und den Zugang zum Rundfunk und den Telemedien.

Einhaltung der Anforderungen

Die Brücke zwischen der technischen Natur der KI und der textuellen logischen Natur der juristischen Vorgaben zu bauen und damit eine Erfüllung der Anforderungen zu erreichen, ist nicht unbedingt einfach. Hacker et al. nennt zwei Vorgehensweisen, damit KI gewinnbringend und sicher in Medienunternehmen eingesetzt werden kann:

Zum einen sollen existierende Vorgaben durch Technologien eingehalten werden. Dafür müssen die Gesetzestexte analysiert und die richtigen datenwissenschaftlichen Verfahren gefunden werden, um diese umzusetzen. Eine Übersetzung von Gesetz in Software muss erfolgen. Zum anderen sollen an

geeigneter Stelle die bestehenden Gesetze neugestaltet werden, sodass die Übereinstimmung mit diesen für die Medienunternehmen einfacher und effizienter wird. Diese Maßnahme ist in angemessenem Umfang einzusetzen [31, S. 2].

5.1.1 DATENSCHUTZ

Bei der Verarbeitung von personenbezogenen Daten von Journalistinnen und Journalisten müssen die Kriterien von „Rechtmäßigkeit, Verarbeitung nach Treu und Glauben [und] Transparenz“ [32, Art. 5 Abs. 1] einhalten werden. Sollten diese Vorgaben nicht erfüllt werden, dürfen personenbezogene Daten nicht verwendet werden. Das zugrunde liegende Gesetz in Deutschland ist das Recht auf informationelle Selbstbestimmung im GG Artikel zwei Absatz eins in Verbindung mit Artikel eins des GG. Das Recht schützt die einzelne Person vor der unbegrenzten Verarbeitung ihrer personenbezogenen Daten [33].

Bei personenbezogenen Daten handelt es sich um Informationen, die eine natürliche Person eindeutig identifizieren. Dazu gehören auch Angaben von Eigenschaften wie Aussehen, Bewegungsbild, Verwandtschaft, Geschlecht und Kulturzugehörigkeit. Verarbeitung bedeutet die Erhebung, Erfassung, Organisation, Kategorisierung, Speicherung, Anpassung, Veränderung, Abfrage, Verwendung, Offenlegung, Abgleich, Verknüpfung, Löschung oder Vernichtung von personenbezogenen Daten. Dies kann mit oder ohne der Hilfe von Software durchgeführt werden [32, Art. 4].

Laut der Datenschutzgrundverordnung (DSGVO) der Europäischen Union (EU) ist die Verarbeitung von personenbezogenen Daten zunächst verboten. Der betroffenen Person ist es jedoch vorbehalten, in die Verarbeitung einzuwilligen [34, S. 24]. Die Person muss ihre Erlaubnis zu einem bestimmten Zweck der Datenverarbeitung geben. Eine Einwilligung wird vorausgesetzt, wenn beispielsweise die Datenverarbeitung für die Erfüllung eines Vertrages, den die betroffene Person in Auftrag gegeben hat, nötig ist [32, Art. 6 Abs. 1]. Die Einwilligung zur Datenverarbeitung kann nur freiwillig und für einen spezifischen Fall gegeben werden [32, Art. 4]. Der Beweis für eine gültige Einwilligung obliegt der Datenverarbeiterin bzw. dem Datenverarbeiter [34, S. 24].

Im Rahmen der DSGVO kann die betroffene Person ebenfalls eine Auskunft über die verarbeitenden Systeme verlangen, wenn die Daten zu einer automatisierten Entscheidungsfindung verwendet werden (z.B. bei der automatisierten Berechnung von Versicherungssummen) [32, Art. 13 Abs. 2]. Dabei müssen auch die Tragweite der Auswirkungen für die betroffenen Personen offengelegt werden [34, S. 25]. Außerdem müssen Datenverarbeiter auf Anfrage der betroffenen Person mitteilen, ob und für welche zeitliche Dauer ihre personenbezogenen Daten gespeichert werden [32, Art. 15 Abs. 1, 32, Art. 13 Abs. 2]. Dazu kommt, dass diese ihre Einwilligung auch jeder Zeit widerrufen kann. Jegliche Verarbeitung vor dem Widerruf bleibt davon jedoch unberührt [32, Art. 7 Abs. 3, 34, S. 24].

Der Datenschutz im Journalismus ist für die Privatperson zu bewahren. Insbesondere schutzbedürftige Personen in Zusammenhang mit einer Straftat müssen in der Berichterstattung

anonymisiert werden. Dies nennt sich redaktioneller Datenschutz. Handelt es sich aber um eine Person öffentlichen Interesses, hat das Informationsinteresse der Öffentlichkeit Priorität. Es darf also in der Presse über diese Person berichtet werden [29, Ziffer 8]. Personenbezogene Daten, die in Situationen erworben wurden, die dem Pressecodex widersprechen, sind zu sperren oder zu löschen [29, Ziffer 4.3].

Rundfunkanstalten kommt jedoch eine besondere Stellung bei der Verarbeitung von personenbezogenen Daten zu. Das sogenannte Medienprivileg erlaubt die Verarbeitung von personenbezogenen Daten zu journalistischen Zwecken. Eine Verarbeitung zu anderen Zwecken ist verboten, denn die Daten fallen dann unter das Datengeheimnis [30, § 12, § 23]. Nur wenn durch die Berichterstattung die Persönlichkeitsrechte einer Person verletzt werden, kann eine Person Auskunft über ihre personenbezogenen Daten erhalten [30, § 12]. Dasselbe gilt für Telemedien [30, § 23].

5.1.2 URHEBERRECHT²

„Die [Urheberinnen und] Urheber von Werken der Literatur, Wissenschaft und Kunst genießen für ihre Werke Schutz nach Maßgabe dieses Gesetzes.“ [35, § 1]

Recht am eigenen Bild

Das Recht am eigenen Bild wird im Kunsturhebergesetz definiert. Abbildungen in Bild oder Video dürfen nur mit Einwilligung der abgebildeten Person veröffentlicht werden. Wurde die Person allerdings für die Anfertigung der Aufnahmen bezahlt, gilt die Einwilligung als erteilt [36, § 22]. Allerdings dürfen zeitgeschichtliche Bildnisse sowie solche, die einem höheren Interesse der Kunst dienen, ohne Einwilligung der abgebildeten Person öffentlich zur Schau gestellt werden. Dazu kommen Werke, auf denen Menschen als Beiwerk oder als Teil eines Ganzen (z. B. eine Großaufnahme einer Demonstration) abgebildet sind. Dabei ist dann nicht der einzelne Mensch von Bedeutung, sondern das größere Motiv [36, § 23].

Urheberrecht an Trainingsdaten

In der letzten Aktualisierung des Urheberrechtsgesetzes im Jahr 2021 wurde unter anderem der Paragraph 44b hinzugefügt. Dieser besagt, dass Kopien von „rechtmäßig zugänglichen Werken“ [35, § 44b] für das Training von KI zulässig sind. Die Kopien müssen jedoch nach Erfüllung ihres Zwecks wieder gelöscht werden. Urheberrechtlich geschützte Werke dürfen also zum Training von KI-Modellen verwendet werden, wenn das Werk vorher schon ohne die Verletzung von Gesetzen zugänglich war. Die Urheberin oder der Urheber kann jedoch widersprechen, indem sie oder er sich die Nutzung zum Training selbst vorbehält [35, § 44b].

² Im Folgenden bezieht sich das Wort „Urheberrecht“ auf alle Geschlechter. Für besseren Wiedererkennungswert ist „Urheber“ im Maskulinum verwendet.

Urheberrecht am Output der KI

Wenn nun eine KI einen Text generiert, ist zunächst unklar, wer nun das Urheberrecht am Erzeugnis der KI hat. Dies ist wichtig, wenn beispielsweise aus dem Beitrag der KI für wissenschaftliche Zwecke zitiert wird. Schließlich stellen im Journalismus Informationen und geschriebener Text ein wichtiges Wirtschaftsgut dar [37, S. 36]. Voraussetzung für den urheberrechtlichen Schutz ist eine „persönliche geistige Schöpfung“ [35, § 2 Abs. 2]. Dies impliziert nach der Enquete-Kommission Künstliche Intelligenz des Deutschen Bundestages sowie Kunschke et al, dass die Schöpferin bzw. der Schöpfer eines Werkes immer ein Mensch ist [26, S. 72, 38, S. 49]. Da die Selbstständigkeit und das Erkenntnisvermögen eines KI-Modells nur durch den von einem Menschen erstellten Quellcode und Trainingsdaten zustande kommt, werden die Ergebnisse der KI auch urheberrechtlich einem Menschen zugesprochen [38, S. 49]. Dadurch, dass die Frage nach dem Urheberrecht zum Zeitpunkt der Veröffentlichung dieser Thesis noch nicht abschließend geklärt ist, müssen vor der Nutzung eines KI-Modells im Journalismus Nachforschungen dazu angestellt werden.

Kennzeichnungspflicht

Im deutschen Recht gab es bis 2022 keine Kennzeichnungspflicht von KI-generierten Inhalten in journalistischen Werken. In der Praxis kennzeichneten Redaktionen bereits ihre KI-generierten Beiträge auf freiwilliger Basis [28, S. 6]. Nach einer Umfrage der Universität Erlangen-Nürnberg im Jahr 2021 sprachen sich 79 Prozent der befragten 88 Journalistinnen und Journalisten, 23 Ausbilderinnen und Ausbilder sowie sechs Vertreterinnen und Vertreter von Journalismusverbänden für eine Kennzeichnungspflicht aus. Ein ähnlicher Prozentsatz aus derselben Gruppe plädierte ebenfalls für einheitliche Standards bei der Darstellung der Kennzeichnung [28, S. 7, 39].

Für den Einsatz von KI in der Generierung von Beiträgen ist es ab 2022 nach Paragraph 18 des Medienstaatsvertrages Pflicht, in Sozialen Medien kenntlich zu machen, wenn ein Beitrag automatisiert durch ein Computerprogramm erstellt wurde. Dem Beitrag ist ein gut sichtbarer Hinweis voranzustellen [30, § 18]. Eine Kennzeichnungspflicht wird auch bei solchen Beiträgen außerhalb von Sozialen Medien gefordert, um die Transparenz zu fördern [37, S. 34].

Exkurs: KI-basierte Uploadfilter

Bei der Veröffentlichung von User Generated Content auf Medienintermediären können KI-basierte Filter eingesetzt werden. Diese prüfen, ob es sich um Inhalte handelt, die urheberrechtlich geschützt sind [26, S. 407]. Dabei macht es keinen Unterschied, ob das urheberrechtlich geschützte Werk nur ausschnittsweise verwendet wird. Auch eine Manipulation wie eine Spiegelung oder das Hinzufügen von Rauschen kann erkannt werden. Die Häufigkeit, nicht-geschützten Content als geschützt zu klassifizieren und deshalb nicht zu veröffentlichen, ist laut der Industrie gering [26, S. 477].

Im juristischen Sinne sind Uploadfilter aber nicht geeignet, um urheberrechtlich geschützten Inhalt zu identifizieren. Sie nehmen keine Rücksicht auf den Kontext des Videos. Handelt es sich beispielsweise nach Paragraph drei des Urheberrechts um eine Bearbeitung des ursprünglichen

Inhalts, bei der die persönliche geistige Schöpfung der Bearbeiterin bzw. des Bearbeiters klar erkennbar ist, wird keine Urheberrechtsverletzung begangen. Auch Lizenzierungen von urheberrechtlich geschützten Inhalten werden mit diesen Filtern nicht erkannt. Ein Einsatz ohne menschliche Überwachung wird also von der Enquete-Kommission Künstliche Intelligenz des Deutschen Bundestages nicht empfohlen [26, S. 408].

5.1.3 INFORMATIONSSICHERHEIT

Auch aus Sicht der Informationssicherheit bergen LLMs ein Risiko gegen Angreiferinnen und Angreifer von außen, die entweder auf die Trainingsdaten zugreifen wollen oder sich die Funktionen des LLMs zu Nutze machen wollen.

Missbrauchsszenarien

Mögliche Missbrauchsszenarien sind die Einführung von Falschmeldungen in das Wissen des LLMs und damit in die journalistischen Erzeugnisse. Eine solche Manipulation wirkt sich stark auf die Glaubwürdigkeit des Medienunternehmens aus und muss unter allen Umständen verhindert werden. Über Eingaben, die in einer anderen Sprache oder Form gestaltet sind als die Trainingsdaten und somit stark von ihnen abweichen, können Anweisungen unbemerkt in das KI-Modell eingebracht werden [40, S. 18, 40, S. 10-11]. Es darf dadurch nicht zu einer sogenannten Datenvergiftung, einer Manipulation des Trainingsdatensatzes, kommen [41, Art. 15].

Ebenfalls ein schädliches Szenario ist eine Model Stealing Attacke. Dadurch wird die Funktionalität des Modells von Angreifern extrahiert. Es wird die Struktur des Modells kopiert. Dabei handelt es sich um den Diebstahl geistigen Eigentums. Mit dem erbeuteten Modell können ebenfalls weitere Angriffe vorbereitet werden [42, S. 5].

Ein weiteres Szenario ist die Rekonstruktion von Trainingsdaten durch gezielte Eingaben. Verwendet das Modell die hauseigenen Mediendaten, könnten dadurch für Außenstehende sensible und personenbezogene Daten offengelegt werden. Dies würde der DSGVO und dem Pressecodex widersprechen und könnte damit rechtliche Konsequenzen nach sich ziehen [40, S. 16].

Sicherheitskriterien

Zu den Kriterien für die Informationssicherheit gehören Sicherheit und Robustheit. Damit ist der Schutz vor unautorisierten Zugriffen auf die Trainingsdaten sowie die trainierten Gewichte des KI-Modells gemeint. Manipulation und Diebstahl von Daten muss ausgeschlossen sein. Selbst in einem Havarieszenario darf noch immer kein Fremdzugriff möglich sein [43]. Das KI-System muss also entsprechend widerstandsfähig und robust konzipiert werden, sodass kein Missbrauch möglich ist. Daher ist es ebenfalls wichtig, die Zeiten der Verfügbarkeit des KI-Modells möglichst unterbrechungsfrei zu halten. So kann gewährleistet werden, dass zu jeder Tageszeit die Unterstützung der KI in Anspruch genommen werden kann. Gibt es Ausfallzeiten, sollen diese durch Logging aufgezeichnet werden. Außerdem müssen Lösungen für Aufträge an die KI, die durch Ausfallzeiten pausiert oder abgebrochen werden, gefunden werden [44, S. 8].

Logging

Logging bezeichnet die automatische Aufzeichnung von Ereignissen während des Betriebs. Die Protokollierung sorgt dafür, dass nachgewiesen werden kann, dass das KI-Modell nur zu seinem ursprünglichen Zweck eingesetzt wurde. Für die Nachforschung von Ausfall- oder Überlastsituationen sollen ebenfalls Informationen aus dem Log gezogen werden. Es erleichtert die Messung der Güte und liefert eine Grundlage, auf der Audits und Tests entworfen werden. Mögliche Risiken und Schwachstellen werden ebenfalls durch Logging identifiziert [41, Art. 12].

Qualitative Anforderungen

Zur Erfüllung der rechtlich-ethischen Kriterien müssen ebenfalls qualitative Kriterien von dem KI-Modell erfüllt werden. Es wird eine gewisse Genauigkeit des Outputs vorausgesetzt. Es darf keine Art des Overfittings auftreten (siehe Abschnitt 2.1.3). Außerdem muss eine Vollständigkeit, auch Recall genannt, gegeben sein. Diese gewährleistet, dass das KI-Modell alle gegebenen Sachverhalte Variablen zuordnen kann. Bei einem Sprachmodell wäre der Recall beispielsweise die Zuordnung eines Wortes im Input zu einem Wort in seinem Vokabular [25, S. 33]. Diese Kriterien mindern die Wahrscheinlichkeit, dass ein Risiko durch ungenauen oder falschen Output auftritt. Daher müssen sie über den Lebenszyklus des KI-Modells auch überwacht und eventuell angepasst werden. Diese Kennzahlen müssen in der technischen Dokumentation angegeben werden [41, Art. 15]. Die jeweiligen Kennzahlen müssen jedoch an die Beschaffenheit und Anwendung des KI-Modells angepasst sein [42, S. 6].

Strukturelle Anforderungen

Neben diesen praktischen Anforderungen empfiehlt die European Broadcasting Union (EBU) die Informationssicherheit als Teil der Unternehmenskultur zu betrachten. Die ersten Schritte beinhalten das Engagement der Unternehmensführung zur Informationssicherheit. Sie soll als Teil der Unternehmensziele betrachtet und in die Prozesse miteinbezogen werden. Innerhalb des Unternehmens werden Rollen dafür definiert und finanzielle Ressourcen allokiert [45, S. 3].

Teil der Umstrukturierung zu mehr Informationssicherheit ist die Ernennung einer bzw. eines Chief Information Security Officers (CISO). Sie oder er ist verantwortlich für alle Belange der Informationssicherheit innerhalb des Medienunternehmens und berichtet der Leitung des Medienunternehmens. Aufgaben der bzw. des CISO sind die Durchführung von Audits und die Evaluation von Risiken, die die Informationssicherheit bedrohen. Die Person arbeitet je nach Größe des Medienunternehmens mit einem Team zusammen [45, S. 5].

Regelmäßige Verbesserungen

Die Informationssicherheit bei KI-Modellen soll durch regelmäßige Überprüfungen aufrechterhalten werden. Die Überprüfungen beinhalten die Umsetzung von Nutzungsfeedback, Fehlerberichten und Logs. Außerdem soll von Nutzerinnen und Nutzern aktiv verlangt werden, mögliche Schwachstellen der Abteilung für Informationssicherheit zu melden, damit diese schnellstmöglich ausgebessert werden können [44, S. 33]. Ebenso sollen in kurzen zeitlichen

Abständen Aktualisierungen oder Anpassung der Kriterien und Metriken durchgeführt werden, um mit dem Modell auf dem aktuellen Stand zu bleiben [44, S. 32].

Dazu soll das KI-Modell auch in einem festgelegten Rhythmus auf den Prüfstand gestellt werden. Dabei wird bewertet, wie zufriedenstellend das KI-Modell die Anforderungen des Medienunternehmens erfüllt. Diese Tests sollen ebenfalls nach größeren Veränderungen erfolgen [44, S. 31].

5.1.4 ERKLÄRBARKEIT

Erklärbarkeit³ in Zusammenhang mit KI bedeutet, die Möglichkeit in einfacher menschlicher Sprache zu erklären, wie KI arbeitet und zu Ergebnissen kommt [46, S. 2]. Das KI-Modell soll seine Vorgehensschritte in logischer Reihenfolge für einen Menschen verständlich präsentieren [46, S. 1]. Im Entwurf des Gesetzes über Künstliche Intelligenz von 2021 der EU ist verankert, dass Nutzerinnen und Nutzer eines KI-Systems dieses angemessen interpretieren können müssen. Dafür muss das KI-System Informationen über seine Daten und Abläufe im Inneren bereitstellen können [41, Art. 13].

Der Wunsch nach Erklärbarkeit stammt aus einer Unvollständigkeit des Wissens über KI-Modelle, die eine Grenze zwischen Optimierung und Evaluation schafft. Die heutigen KI-Modelle arbeiten stark autonom, sodass der Mensch ein KI-Modell als Black Box empfindet. Er hat also keine Kenntnis über die Abläufe im Inneren. Hinzukommt die hohe Rechengeschwindigkeit. Ein Mensch könnte nur noch mit großer Verzögerung nachvollziehen, welche Rechengänge innerhalb des Modells durchgeführt wurden. Da der Mensch also nur unvollständiges Wissen hat, macht dies eine Evaluation der Ergebnisse und Optimierung des Modells mühsam. Deshalb sollte ein KI-Modell in der Lage sein, sich selbst und seine Abläufe zu erklären [46, S. 3].

Wie verständlich die Erklärungen des KI-Modells sind, lässt sich anhand von drei Ansätzen nach Doshi-Valez und Been evaluieren. Zunächst wird eine solche Erklärung im Kontext der Aufgabe betrachtet, anhand der das KI-Modell sein Vorgehen offenlegt. Es werden Experimente mit Testpersonen durchgeführt, um zu ermitteln, ob diese die Erklärung zur Vorgehensweise des KI-Modells für die Aufgabe oder Teilaufgabe hilfreich fanden. Für diese Art der Experimente müssen die jeweiligen Anwenderinnen und Anwender des KI-Modells ihre Einschätzung angeben. Handelt es sich also um eine KI-Anwendung im medizinischen Bereich, muss auch medizinisches Personal Teil der Testpersonen sein [46, S. 4-5].

Der zweite Ansatz zielt auf das generelle menschliche Verständnis ab. Hier müssen keine Sachverständigen die Probandentests durchführen. Es wird untersucht, welche Arten von Erklärungen die meisten Menschen in kurzer Zeit verstehen. Außerdem wird die Qualität und Komplexität durch die Testpersonen geprüft. Diese Experimente laufen ohne konkreten Anwendungsfall ab [46, S. 5].

³ auf Englisch Interpretability

Für den dritten Ansatz werden keine Testpersonen benötigt. Die Erklärungen werden mit der formalen Definition bzw. mit den formalen Anforderungen an die Erklärbarkeit abgeglichen. Sie wird als Stellvertretung für eine menschliche Testperson benutzt. Dieser Ansatz wird verwendet, nachdem die Erklärungen durch Tests mit Probandinnen und Probanden durchgeführt wurden. Auch wenn geringe Verbesserungen am KI-Modell durchgeführt wurden, müssen nicht zwangsläufig aufwendige Tests mit Probandinnen und Probanden erneut durchgeführt werden. Hier reicht dann eine Überprüfung anhand der formalen Definition aus. Die Auswahl der Definition ist dabei schwierig [46, S. 5-6] und sollte daher den Anforderungen an das KI-Modell angeglichen werden [46, S. 9]. Je nach Anwendung haben KI-Modelle eventuell unterschiedliche Bedürfnisse an Erklärbarkeit. Die Tiefe und Qualität der Erklärungen kann also variieren.

5.1.5 FAIRNESS

Fairness im Kontext von KI bedeutet die Identifizierung, Analyse und Quantifizierung von Verzerrungen, auch Bias genannt. Bias liegt vor, wenn Personengruppen nach Eigenschaften wie Geschlecht, Herkunft, Aussehen oder sexueller Orientierung unzulässig diskriminiert werden. Voreingenommenheit hängt stark von den verwendeten Trainingsdaten ab. Je nachdem, wie häufig bestimmte Personengruppen in den Trainingsdaten auftreten, desto eher empfindet die KI diese Personengruppe als die (mengenmäßig) vorherrschende Gruppe. Dies hat dann Einfluss auf die potenzielle Fairness der KI [43].

Mediale Voreingenommenheit, die rechtliche Probleme bereiten kann, muss einerseits vorsätzlich und andererseits systematisch erfolgen [47, S. 392]. Fälle von unabsichtlicher oder unterbewusster Voreingenommenheit sind außerhalb des Rahmens dieses Kapitels. Voreingenommenheit ist bereits bei menschlichen Journalistinnen und Journalisten ein unterbewusster oder sogar absichtlicher Effekt. Um die Entstehung von Voreingenommenheit in KI-Modellen zu erkennen und zu beseitigen, muss zunächst geklärt werden, woher die Voreingenommenheit der Menschen stammt. Schließlich wird das KI-Modell mit Daten trainiert, die Menschen erzeugt haben. Außerdem repräsentieren Daten nur einen begrenzten Ausschnitt der Realität. Sie können durch ihre Selektivität und menschliche Beschriftung nicht völlig wertfrei sein [48].

Verschiedene Arten von Voreingenommenheit

Voreingenommenheit kann viele Gründe haben. Es beginnt mit den politischen und ideologischen Ansichten der Redaktionsmitglieder. Diese können unbewusst in der journalistischen Arbeit zum Ausdruck kommen. Es ist aber auch möglich, dass das Medium eine gewisse politische oder weltanschauliche Richtung vorgibt. Medien arbeiten zunehmend profitorientiert. Sie passen also ihre Nachrichten der Zielgruppe an, sodass Nachrichten außerhalb dieses Interessensgebiets vernachlässigt werden (siehe Abbildung neun oben) [47, S. 394].

Die Voreingenommenheit fängt in der Wertschöpfungskette schon bei der Auswahl der Themen an. Je nachdem, ob über ein bestimmtes Ereignis oder eine Veranstaltung berichtet wird, kann dies schon eine Meinung der medienschaffenden Person ausdrücken. Der sogenannte Event Selection

Bias kann dadurch untersucht werden, dass man eine möglichst objektive Quelle (z.B. Polizeiberichte) der Berichterstattung in den Nachrichten gegenübergestellt werden (siehe Abbildung neun Fact Selection) [47, S. 398, 49]

Nicht nur bei der Auswahl der Themen, sondern auch bei der Auswahl der Quellen kann eine Voreingenommenheit entstehen. Um eine Überfrachtung an Informationen zu vermeiden, dürfen nur die wichtigsten Informationsquellen genutzt werden. Gerade in der Berichterstattung über Veranstaltungen und Ereignisse, an denen viele Menschen beteiligt sind, können Journalistinnen und Journalisten nicht alle Sichtweisen in ihre Beiträge miteinfließen lassen und müssen daher auswählen, welche Quellen sie zulassen. Dies nennt sich Source Selection Bias (siehe Abbildung neun Fact Selection) [47, S. 400].

Generell wird durch Betonung oder Weglassen ebenfalls der Informationsgehalt des Beitrags gelenkt [47, S. 401]. Dabei spielt die Wortwahl beim Ausdruck von Voreingenommenheit eine große Rolle. Durch gefärbte Bezeichnungen für Personen(-gruppen), Orte oder Gegenstände kann ein neues Bild von einer Tatsache vermittelt werden. Es kann so schnell eine positive oder negative Stimmung vermittelt werden oder eine bestimmte Sichtweise in den Vordergrund gerückt werden (siehe Abbildung neun Writing Style) [47, S. 403].

Schließlich kann auch durch die örtliche und zeitliche Platzierung eines Beitrags die Aufmerksamkeit des Publikums gelenkt werden. Ort und Größe implizieren einen Wert, der dem Beitrag zugesprochen wird. Das Publikum behandelt den Beitrag als dem Wert entsprechend und konsumiert vorrangig die prominenten Beiträge (siehe Abbildung neun Presentation Style). In Onlinemagazinen und Mediatheken kann die Platzierung von Beiträgen dynamisch verändert werden, sodass sich immer an die Präferenzen der Nutzerinnen und Nutzer angepasst werden kann. In gedruckten Zeitungen oder dem linearen Fernsehen ist dies eher schwierig, daher muss das Layout bzw. der Sendeplan vor Veröffentlichung feststehen [47, S. 404].

Eine weitere Art der Beeinflussung ist die Auswahl und Anordnung von Bildern oder Videos in einem Beitrag. Durch Bilder, auch gerade Teaserbilder in Verbindung mit einer Schlagzeile, erhält die Konsumentin bzw. der Konsument einen ersten Eindruck vom Beitrag und seinem Thema. Dadurch kann schon vorgegeben sein, wie die Konsumentin bzw. der Konsument den Rest des Beitrags aufnimmt. Ebenfalls einen Einfluss haben Bildunterschriften. Diese beschreiben den Bildinhalt, sind aber nicht immer völlig objektiv und können dadurch auch den Bildeindruck beeinflussen [47, S. 406].

Außerdem einen Einfluss auf die Perzeption von Nachrichtenbeiträgen hat der Hintergrund der Nutzerinnen und Nutzer. Vorwissen, Haltung, sozialer Status und Aufenthaltsland spielen eine große Rolle beim Lesen und Verstehen von Beiträgen. Diese Eigenschaften bestimmen beispielsweise, ob eine Leserin bzw. ein Leser von einem Ereignis betroffen ist. Wird beispielsweise negativ über ein bestimmtes Land berichtet, kann dies in der Leserschaft ebenfalls unterschiedliche Gefühle hervorrufen je nach Herkunfts- oder Aufenthaltsland (siehe Abbildung neun Consumer Context) [47, S. 395].

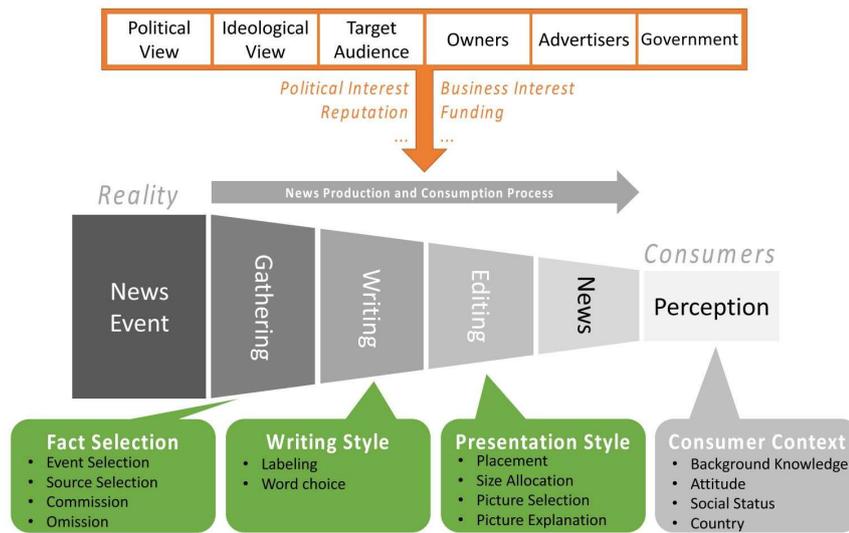


Abbildung 9: Evolution von Voreingenommenheit in den Medien [47, S. 394].

Entstehung von maschineller Voreingenommenheit

Voreingenommenheit tritt sowohl bei menschlichen Journalistinnen und Journalisten als auch bei deren maschineller Unterstützung auf. Ein solcher Effekt ist jedoch bei einer KI weitaus gravierender als bei einem einzelnen Menschen, weil bei häufigem und weitreichendem Einsatz diese eine Breitenwirkung entfaltet. Die Skaleneffekte können journalistische Erzeugnisse öfter und systematischer beeinflussen als ein Mensch [50, S. 167].

Die Voreingenommenheit stammt zunächst aus den Trainingsdaten. Diese wurden von Menschen erzeugt und bilden die Vergangenheit und die Gegenwart ab. Daher kann es passieren, dass Diskriminierungen aus Vergangenheit und Gegenwart auch die Entscheidungen der KI beeinflussen. Diese Entscheidungen können durch innere Abläufe der KI eventuell verschleiert oder verstärkt werden [50, S. 167].

Meist ist eine zahlenmäßig ungenügende Repräsentativität von bestimmten Personengruppen der Grund für die Diskriminierung. Innerhalb eines Datensatzes für die Erkennung von Gesichtern sind beispielsweise Trainingsbeispiele von People of Color seltener vorhanden als Beispiele von weißen Menschen. Dem KI-Modell sind damit die Eigenschaften von dunkler Hautfarbe nicht vollständig bekannt und das führt zu Fehlklassifikationen [50, S. 168].

Abgesehen von den Trainingsdaten können jedoch auch durch die Programmierung von KI-Modellen Diskriminierungseffekte auftreten. Durch die Wahl von Labels und Variablen können Personen und Personengruppen nicht ausreichend beschrieben werden und somit ihre Charakteristiken zur fehlerfreien Bearbeitung nicht extrahiert werden. Außerdem stellen sich Probleme ein, wenn ein KI-Modell nicht für den Zweck verwendet wird, für den es entworfen und trainiert wurde. Durch ein solch unvorhergesehenes Einsatzumfeld können die Ergebnisse der KI schnell verzerrt wirken [50, S. 168].

Effekte der Voreingenommenheit

Durch Voreingenommenheit im Journalismus wird nicht vollständig über einige Themen berichtet, während über andere Themen überdurchschnittlich viele Inhalte veröffentlicht wird. Dies kann dazu führen, dass die öffentliche Meinung polarisiert und sich dies auch auf Wahlergebnisse auswirken kann [47, S. 393]. Bei Wahlen in Österreich und in den USA wurden Zusammenhänge zwischen Wählerentscheidungen und dem Konsum von bestimmten Medien festgestellt. Politikerinnen und Politiker wurden aufgrund von unterschwellig beeinflussender Berichterstattung in ein positiveres Licht gerückt. Damit ist beispielsweise die Auswahl von Bildern und Videoaufnahmen der Politikerinnen und Politiker, auf denen sie freundlich, kompetent und durchsetzungsfähig erscheinen, gemeint. Je häufiger eine Politikerin bzw. ein Politiker positiv medial in Erscheinung tritt, desto wahrscheinlicher wird die Person gewählt [51, S. 1045, 52, S. 124].

Mit dem Einsatz von KI-basierten Methoden, Beiträge zu generieren, werden sich die Effekte auf die Meinung der Leserinnen und Leser wahrscheinlich verstärken. Wenn Beiträge in einer großen Dichte auf verschiedenen Ausspielkanälen distribuiert werden und somit möglichst viele Menschen auf dem Medium ihrer Wahl angesprochen werden, könnten die öffentliche Meinung und damit die Ergebnisse von Wahlen bis zu einem gewissen Grad beeinflusst werden [47, S. 409].

Rechtliche Aspekte zur Voreingenommenheit

Niemand darf in der Presse wegen „seines Geschlechts, einer Behinderung oder seiner Zugehörigkeit zu einer ethnischen, religiösen, sozialen oder nationalen Gruppe“ [29, Ziffer 12] diskriminiert werden. Außerdem verpflichtet sich die Presse keine „religiöse[n], weltanschauliche[n] oder sittliche[n] Überzeugungen“ [29, Ziffer 10] herabzuwürdigen. Diese Ziffern werden aus dem Diskriminierungsverbot im GG Artikel drei abgeleitet. In Verlängerung bedeutet dies auch, dass unter Einsatz einer KI keine Diskriminierung auftreten darf, da es sich sonst um einen Verstoß gegen das GG handelt [33].

Wird eine Person aufgrund eines persönlichen Merkmals diskriminiert, kann diese nach dem Allgemeinen Gleichbehandlungsgesetz (AGG) Schadensersatzansprüche stellen [53]. Die Merkmale sind nach Paragraf eins des AGG die Rasse, ethnische Herkunft, Geschlecht, Religion, Weltanschauung, Behinderung, Alter und sexuelle Identität [54, § 1]. Das Gesetz beschreibt also Schutzrechte für betroffene Personen. Die Beweislast für eine solche Diskriminierung liegt jedoch bei diesen selbst. Wichtig für die Argumentation ist, dass die betroffene Person dieses Merkmal erfüllt und deswegen ungerechtfertigt schlechter behandelt wurde als eine Person, die dieses Merkmal nicht erfüllt [53]. Eine Erleichterung bei der Beweisführung soll der Paragraf 22 des AGG schaffen. Dieser bestimmt, dass, wenn eine Betroffene bzw. ein Betroffener Indizien für eine Diskriminierung vorweisen kann, der Diskriminierende nachweisen muss, dass keine Ungleichbehandlung vorliegt [53, 54, § 22]. Diese gesetzlichen und ethischen Grundsätze werden für den Einsatz von KI-Systemen ebenfalls verwendet.

Beispiel: Bias bei der Aggregation, Selektion und Präsentation von Beiträgen

Die Zusammenstellung eines Newsfeeds für eine Nutzerin bzw. einen Nutzer erfolgt automatisiert mithilfe von Algorithmen. Diese sammeln Daten über das Nutzungsverhalten und können anhand dessen weitere, zu den Interessen passende Beiträge vorschlagen [1, S. 61]. Diese algorithmischen Systeme können damit einen Einfluss auf die Meinungsbildung von Menschen haben. Insbesondere bei öffentlich interessanten Themen, Fakten und Meinungen ist es möglich, dass Nutzerinnen und Nutzer von der Auswahl der Algorithmen beeinflusst werden [55].

Ein Phänomen dieser algorithmischen Vorschläge sind sogenannte Filterblasen und Echokammern. Eine Filterblase entsteht, wenn Nutzerinnen und Nutzer sich bevorzugt nur innerhalb einer Gruppe bewegen und kommunizieren, sodass nur wenige Meinungen von außen an die Nutzerin bzw. an den Nutzer herankommen. Eine Echokammer ist die Weiterentwicklung der Filterblase. Dort erhalten Nutzerinnen und Nutzer nur Vorschläge für Beiträge, die ihrer eigenen Meinung entsprechen ohne Einfluss von kontroversen Ansichten. Diese Beobachtungen sowie der Einfluss auf die Meinungsbildung sind aber noch nicht hinreichend wissenschaftlich erforscht [55].

Damit Nutzerinnen und Nutzer aber eine Vorstellung haben, wie die Auswahl der Beiträge, die ihnen angezeigt wird, zusammengestellt wird, ist es wichtig, die Mechanismen dahinter zu kennen. Nutzerinnen und Nutzer möchten wissen, warum sie gerade über bestimmte Themen informiert werden und über andere nicht. Das Vertrauen in die Medienunternehmen würde ebenfalls durch eine solche Offenlegung gestärkt [1, S. 61].

Der Mechanismus der algorithmischen Zusammenstellung von Beiträgen zu einem Newsfeed wird in Sozialen Medien verstärkt angewandt. Diese unterliegt nicht der journalistischen Sorgfaltspflicht. Daher sind die Kriterien für die Auswahl und Anordnung der angezeigten Beiträge nicht unbedingt an den Werten der Sorgfalt und Vielfalt orientiert. Stattdessen ist es Ziel der Sozialen Medien, die Nutzungszeit der Benutzerinnen und Benutzer zu erhöhen. Das wird dadurch erreicht, dass bestimmte Interessen und Meinungen bedient werden, um die Aufmerksamkeit der Nutzerinnen und Nutzer zu lenken und so einen höheren Gewinn zu erzielen [55].

Arbeit zum De-Biasing

Um gegen Voreingenommenheit vorzugehen, bedarf es menschlicher sowie maschineller Kontrolle. Dabei können verschiedene Maßstäbe angelegt werden. Es kann auf kurze oder lange Zeit beobachtet werden, wie oft oder wie intensiv ein Thema in den Nachrichten behandelt wurde. Ziel der Analyse ist daher, wiederkehrende Muster in der Themen- und Quellenwahl zu finden [47, S. 397]. Aufgrund dessen ist es wichtig eine Art Kontrolle einzuführen, um eine Antwort auf die Frage der Voreingenommenheit zu liefern [1, S. 62]. Kriterien für diesen Kontrollmechanismus wären Originalität, Diversität, Faktentreue, Anspruch an die Leserschaft und Objektivität [1, S. 58]. Auch bei der Auswahl der Quellen kann eine menschliche oder technische Unterstützung Voreingenommenheit abmildern. Diese Kontrolle kann durch einen Menschen oder durch eine weitere KI-gestützte Software durchgeführt werden. Dadurch wird transparent, welche Themen oder

Quellen ausgewählt werden [56, S. 17]. Mit bereits etablierten Methoden der Plagiatserkennung kann in bestehenden Artikeln analysiert werden, welche Quellen ursprünglich verwendet wurden. Anhand der gefundenen Quellen kann man dann bestimmen, ob die Auswahl objektiv abgelaufen ist. Im Journalismus – anders als im akademischen Bereich – ist es nämlich keine Pflicht, verwendete Quellen anzugeben [47, S. 400].

Weitere KI-gestützte Techniken wie Aspect-Level-Browsing und Sentiment Analysis tragen ebenfalls zur Minderung von Voreingenommenheit bei. Durch das Aspect-Level-Browsing werden verschiedene Sichtweisen auf ein Thema sichtbar gemacht. Wenn eine Journalistin bzw. ein Journalist im Archiv nach Material zu einem Recherchethema sucht, werden ihr oder ihm Beiträge anhand der Metadaten zu diesem Thema angezeigt. Bei dem Verfahren liegt jedoch der Unterschied darin, dass die Schlagwörter jedes gefundenen Beitrags nach ihrer Prominenz im Beitrag gewichtet werden. Dadurch kann die Liste an gefundenem Material in verschiedene Aspekte zu dem Überthema aufgegliedert werden, sodass die medienschaffende Person einen ausgewogenen Überblick auf das Thema erhält [47, S. 402].

Die Sentiment Analysis kann wiederum bei Voreingenommenheit durch Bezeichnung und Wortwahl helfen. Wird Text (eventuell auch aus einer Audiospur transkribiert) auf seinen Gefühlsausdruck hin überprüft, sollte sich das Gefühl für eine objektive Berichterstattung als neutral herausstellen. Ist dies nicht der Fall, liegt eventuell eine Färbung und damit Voreingenommenheit vor [47, S. 403]. Dadurch, dass KI aufgrund ihrer Trainingsdaten ebenfalls nicht frei von Voreingenommenheit ist, lernen immer wieder Mensch und Maschine voneinander, was Objektivität bedeutet. Dieser Zusammenhang kann als Kreislauf beschrieben werden. Die Menschen korrigieren den Bias der KI durch Feedback und die KI korrigiert die bewusste oder unterbewusste Bevorzugung von Themen durch die Menschen [1, S. 60].

5.1.6 TRANSPARENZ

Ist ein KI-Modell transparent, legt dieses seine Informationen offen und verliert den Charakter einer Black Box. Die Informationen sind zunächst die Trainingsdaten, deren Herkunft und sein Quellcode, um die Abläufe im Inneren ansatzweise nachvollziehen zu können. Ferner ist damit auch die Qualität der Trainingsdaten und die Absichten, mit denen diese gelabelt wurden, gemeint. Transparenz ermöglicht, dass die Maßnahmen zur Erfüllung von rechtlichen und ethischen Anforderungen im Zusammenhang mit KI geprüft und gegebenenfalls korrigiert werden kann. Transparenz steigert außerdem das Vertrauen in die KI [43]. Es ist also eine unbedingte Voraussetzung für die Qualität eines KI-Modells.

Qualitätsmerkmale der Datensätze

Die Eigenschaften für qualitativ gute Daten sind vielschichtig. Nach Wang und Strong wurden bereits 1996 vier Kategorien an Merkmalen aufgestellt, die man mit etwas Modifikation auch für Trainingsdaten einer KI anwenden kann [57, S. 20].

Die Kategorien sind:

System/Zugang	Inhalt
Zugänglichkeit (7)	Hohes Ansehen (12)
Zugangssicherheit (18)	Korrektheit (4)
Bearbeitbarkeit	Objektivität (8)
Effizienz beim Abruf	Glaubwürdigkeit (1)
Darstellung	Nutzung
Konsistenz (13)	Aktualität (9)
Verständlichkeit (6)	Mehrwert (2)
Übersichtlichkeit	Vollständigkeit (10)
Eindeutige Auslegbarkeit (17)	Angemessener Umfang (19)
Erklärbarkeit (5)	Relevanz (3)

Abbildung 10: Kategorien der Datenqualität mit Beispielen nach [57, S. 20, 58].

In der Arbeit von Wang und Strong wurden die Merkmale nach ihrer Priorität bei der Datenverarbeitung sortiert (siehe Zahlen in Klammern in Abbildung zehn). So hat die Glaubwürdigkeit von Daten die größte Wichtigkeit, während dem angemessenen Umfang von Daten eine eher geringere Rolle zukommt. Merkmale ohne Priorisierung wurden in der Arbeit von Wang und Strong nicht betrachtet, sind aber für heutige KI-Anwendungen Voraussetzung [58].

Im Entwurf des Gesetz über Künstliche Intelligenz von 2021 der EU sind die Kriterien für sämtliche verwendeten Datensätze Relevanz, Repräsentativität, Fehlerfreiheit und Vollständigkeit [41, Art. 10 Abs. 3]. Sie entsprechen also im weitesten Sinne den Kategorien Inhalt, Darstellung und Nutzung von Wang und Strong.

Prüfmetriken

Für die obengenannten Merkmale eignen sich drei Metriken, in die jeweils die Merkmale eingeteilt werden. Ist ein Merkmal messbar, so kann man seine Ausprägung in Zahlenwerten ablesen. Ein Beispiel ist die Effizienz. Hier kann genau gemessen werden, mit welcher Geschwindigkeit Daten eingelesen oder von einem Speicherort zum nächsten transportiert werden (siehe Abbildung elf links). Kann man ein Merkmal überprüfen, kann man eine binäre Entscheidung fällen, ob das Merkmal zutrifft oder nicht. Bei der Bearbeitbarkeit beispielsweise kann man genau feststellen, ob man die Daten nun bearbeiten kann oder nicht (siehe Abbildung elf Mitte). Ein bewertbares Merkmal kann nicht objektiv festgestellt werden, sondern bedarf einer menschlichen Einschätzung. Ob Daten glaubwürdig sind, kann nicht mit Zahlen gemessen werden. Das hängt von mehreren Faktoren ab, die durch einen Menschen betrachtet und subjektiv bewertet werden müssen (siehe Abbildung elf rechts) [59].



Abbildung 11: Kriterien für Datenqualität [59].

Um nun Datensätze anhand der Merkmale zu vergleichen, muss die Messung der Merkmale standardisiert werden. Eine Normung der Prüfmethode bildet außerdem die Grundlage für eine Zertifizierung. Eine Normung ist offen für Sachverständige aus Wirtschaft, Wissenschaft, Verbänden und Politik. Durch einheitliche Prüfmethode und Zertifizierung von Trainingsdatensätzen können außerdem Haftungsrisiken gemeistert werden [60].

Herkunft der Trainingsdaten

Da die Trainingsdaten die Grundlage eines KI-Systems sind, muss bei der Herkunft der Daten auf Unabhängigkeit der Lieferantin bzw. des Lieferanten geachtet werden. Denn je unabhängiger die Bezugsquelle, desto weniger beeinflusst sind die Daten. Generell ist die geografische Herkunft der Daten auch von Bedeutung. Handelt es sich beispielsweise um Daten, die aus einem anderen Teil der Welt stammen, ist es möglich, dass andere kulturelle und traditionelle Gepflogenheiten Einzug in die Trainingsdaten erhalten. In Europa beispielsweise wird ein Elefant als Zootier bezeichnet, in Afrika als Wildtier und auf dem indischen Kontinent eventuell als Nutztier. Sind die Datenlieferanten Teil eines großen Tech-Unternehmens, können sich dort Verzerrungen in die Daten einschleichen.

Außerdem spielt die Sprache der Trainingsdaten eine große Rolle. Da der Großteil der fortgeschrittenen KI-Modelle aus dem englisch-sprachigen Raum oder aus einer Forschungsgemeinschaft mit Englisch als Lingua Franca stammt, ist die Sprache der Trainingsdaten meist Englisch. Daher ist es wichtig, möglichst neutrale Daten geografisch nah aus unabhängiger Quelle zu beziehen. Die Einrichtung von sogenannten Open-Data-Portalen für Medienunternehmen kann dabei helfen, Daten nicht von kommerziellen Drittanbietern zu beziehen. Eine unabhängige Datenausgabe- und Verwaltungsstelle für Rundfunkanstalten in Deutschland oder Europa wäre geeignet [37, S. 37].

Die Enquete-Kommission der Deutschen Bundesregierung beschreibt die Schaffung von staatlichen Open-Data-Portalen als kritischen Schritt, um Daten, die unabhängig von kommerziellen Drittanbietern sind, für das Training in Rundfunkanstalten zur Verfügung zu stellen [26, S. 449].

Dazu kommen Daten, die das Medium selbst aus dem Nutzungsverhalten der Leserschaft erhebt. Durch Tracking im eigenen Webauftritt und auf Drittseiten erhalten Journalistinnen und Journalisten einen Eindruck von den Präferenzen ihres Publikums. Daraus lassen sich auch neue Diskussionen und Recherchethemen ableiten [26, S. 453].

Annotation der Datensätze

Bei der Annotation der Datensätze arbeiten Menschen alle Beispiele eines Datensatzes durch und versehen diese mit Labeln. Handelt es sich um einen Datensatz zur Objekterkennung, sind die Beispiele eine große Menge an Bildern und die Label die Beschreibung, was auf dem jeweiligen Bild zu sehen ist. Ziel der Annotation ist es, eine möglichst genaue und wahrheitsgetreue Lösung des Beispiels zu liefern.

Bei der Auswahl des geeigneten Datensatzes für das eigene KI-Modell muss also darauf geachtet werden, mit welchen Zielen und unter welchen Vorgaben der Datensatz gelabelt wurde. Wird ein Datensatz mit dem Ziel einer möglichst präzisen Beschreibung oder nur mit einer generellen Beschreibung versehen, kann dies beim Training und bei den Ergebnissen des KI-Modells einen Unterschied machen.

Dazu kommt, dass in manchen Fällen auch zwischen den einzelnen Personen, die die Beispiele annotieren, Uneinigkeit herrschen kann. Je nach Wissensstand und kulturellem Hintergrund der labelnden Personen können verschiedene Lösungen für dasselbe Beispiel entstehen. Außerdem kann auch die Annotatorin bzw. der Annotator bereits voreingenommen sein und ihre oder seine Sichtweise in die Trainingsdaten einfließen lassen. Das nennt sich Annotator Bias [48]. Die Beschriftung der Daten wird häufig in Niedriglohnländern ausgelagert. In solchen Datenfarmen herrschen schlechte Arbeitsbedingungen. Arbeiterinnen und Arbeiter machen beispielsweise bei der Erkennung von jugendschutzkritischen Inhalten traumatische Erfahrungen [5].

Weitere Aspekte von Transparenz

Dokumentation erhöht die Transparenz eines KI-Systems. Laut dem Gesetz über Künstliche Intelligenz der EU muss über KI-Systeme eines gewissen Risikobereiches eine technische Dokumentation erstellt werden. Diese muss während des Betriebs des KI-Systems vorhanden und auf dem neusten Stand gehalten werden. Die Dokumentation muss unter anderem die Verwendung sämtlicher Datensätze zu Training, Validierung und Testen sowie die Einrichtung von Schnittstellen angeben [41, Art. 64, 41, Art. 11].

Die Transparenz wird ebenfalls durch die Offenlegung des Quellcodes des KI-Modells gesteigert. Hier ist außerdem wichtig, dass die verwendeten Mechanismen für den Menschen verständlich dargelegt werden. Es muss außerdem darauf geachtet werden, ob der Quellcode unter Open Source fällt oder patentiert ist. Dadurch darf er nicht veröffentlicht werden.

Während der Verwendung des KI-Modells in der Rundfunkanstalt oder im Onlinemedium muss ebenfalls geklärt werden, ob es beim produktiven Einsatz dazulernt und neu generierte Beiträge wiederum als Trainingsmaterial genutzt werden oder ob das KI-Modell in seinem

Auslieferungszustand verbleibt. Ist ersteres der Fall, ist zu überlegen, ob das KI-Modell nach einer gewissen Zeit erneut überprüft werden muss. Bei einer solchen Prüfung würde festgestellt, ob der Datensatz bzw. der Output der KI noch den obengenannten Kriterien entspricht. Dabei muss geklärt werden, wer und auf welche Weise die Prüfung durchführt. Es ist denkbar, dass eine automatisierte Teststellung verschiedene Testfälle definiert und testet.

5.1.7 REGULIERUNG

Mögliche Ansätze der Regulierung sind die Sicherung der journalistischen Qualität oder der reglementierte Zugang zu Inhalten und Plattformen [26, S. 444]. Letzteres gilt insbesondere für jugendliche Nutzerinnen und Nutzer, die Inhalte von KI-gestützten Plattformen konsumieren [26, S. 405]. Dies mindert ebenfalls die Unsicherheit im Umgang mit KI und zeigt die Grenzen derselben auf [1, S. 62]. Nach einer Umfrage der Universität Düsseldorf und des Center of Advanced Internet Studies sprechen sich auch 68 Prozent der Personen mit deutscher Staatsangehörigkeit ab 18 Jahren für eine generelle Regulierung der KI im Medienbereich aus [61, S. 6]. Im Journalismus sowie in der deutschen Regierung spricht man sich für zusätzliche branchenweite Richtlinien, ähnlich dem Pressecodex, oder sogar eine Erweiterung desselben aus [26, S. 449, 28, S. 8, 62].

Im geänderten Medienstaatsvertrag von 2022 wird unter Paragraph 93 eine Pflicht zur Bereitstellung von Informationen, beispielsweise über die Zusammenstellung des Newsfeeds auf den Startseiten der Sozialen Medien, beschrieben. Diese Pflicht gilt für alle Medienintermediäre. Ein Intermediär ist im Sinne des Medienstaatsvertrags ein Telemedium, das „auch journalistisch-redaktionelle Angebote dritter Personen aggregiert, selektiert und allgemein zugänglich präsentiert, ohne diese zu einem Gesamtangebot zusammenzufassen“ [30, § 2]. Beispiele sind Suchmaschinen, Netzwerkplattformen, Multimediaplattformen und Instant-Messaging-Dienste. Sie stellen Verbindungen zwischen Menschen und vielfältigen Inhalten her [26, S. 418]. Das Geschäftsmodell solcher Intermediäre beruht auf der Anzeige von Werbung basierend auf den gesammelten persönlichen Verhaltensdaten der Nutzerinnen und Nutzer. Obwohl die Angebote von Intermediären zunächst kostenlos sind, „zahlen“ Nutzerinnen und Nutzer mit dem Konsum von Werbung und mit ihren Nutzungsdaten [26, S. 418].

Laut dem Medienstaatsvertrag müssen Medienintermediäre offenlegen, wie und nach welchen Kriterien die Beiträge angesammelt, ausgewählt und präsentiert werden [30, § 93]. Im folgenden Paragraphen wird außerdem ein Verbot der Diskriminierung von Inhalten durch die Anbieterinnen und Anbieter von Medienintermediären ausgesprochen. „Journalistisch-redaktionell gestaltete Inhalte“ dürfen nicht außerhalb der offengelegten Kriterien, die angeben, wie Beiträge zusammengestellt werden, behandelt werden. Diese dürfen ebenfalls nicht systematisch behindert werden [30, § 94]. Diese Vorgaben gelten für Medienintermediäre, die innerhalb von sechs Monaten mehr als eine Million Nutzerinnen und Nutzer erreichen, keine Verkaufsplattform sind und nicht ausschließlich familiären und privaten Zwecken dienen [30, § 91]. Beispielsweise müssen Anbieterinnen und Anbieter von Online-Plattformen in verständlicher Sprache offenlegen, auf Basis welcher Parameter ihre Empfehlungssysteme Beiträge auswählen. Außerdem müssen Nutzerinnen

und Nutzer darüber informiert werden, wie sie diese Parameter ändern können [63, Art. 27 Abs. 1]. Im Jahr 2024 tritt außerdem das Gesetz über digitale Dienste der EU in Kraft. Dieses soll für eine größere Sicherheit im digitalen Raum sorgen. Online-Plattformen sind für die Inhalte verantwortlich, die dort veröffentlicht werden und müssen Nutzerinnen und Nutzer über die Richtlinien der Plattform informieren. Außerdem sind die Anbieterinnen und Anbieter dieser Plattformen stärker in Bezug auf Jugendschutz in der Pflicht. Es müssen Kontroll- und Beschwerdestellen für Behörden sowie für Nutzerinnen und Nutzer eingerichtet werden [55]. Plattformen, die ihren Nutzerinnen und Nutzern Inhalte in Videoformat bereitstellen, sind nach dem Telemediengesetz verpflichtet, Beschwerden über rechtswidrige Inhalte zu erfassen. Die Erfassung muss elektronisch und für die Nutzerin bzw. den Nutzer leicht zugänglich sein. Die Plattform muss die Beschwerden zur Kenntnis nehmen und bearbeiten [64, § 10a].

Haftung für journalistischen Inhalt

Journalistische Erzeugnisse unterliegen den publizistischen Grundlagen, genannt Pressecodex. Dabei handelt es sich um ethische Grundsätze, denen jede medienschaffende Person verpflichtet ist. Die obersten Gebote sind die „Achtung vor der Wahrheit, die Wahrung der Menschenwürde und die wahrhaftige Unterrichtung der Öffentlichkeit“. Ausgehend davon soll die Vertrauenswürdigkeit der Medien gewahrt werden [29, Ziffer 1]. Daher müssen vor der Veröffentlichung von Informationen diese auf ihren Wahrheitsgehalt geprüft werden. Ihre Aussage darf, egal mit welchen Mitteln, nicht verändert werden. Insbesondere bei Bildern darf der Sinn nicht durch Bearbeitung oder Beschriftung verloren gehen. Nicht fundierte Berichte, Gerüchte und Vermutungen dürfen ohne Beweis nicht weiterverbreitet werden oder müssen als solche gekennzeichnet werden [29, Ziffer 2].

Telemedien, also der Rundfunk, unterstehen ebenfalls dem Pressecodex, sofern sie „regelmäßig Nachrichten oder politische Informationen enthalten“ [30, § 19]. Die Redaktion sorgt für die Einhaltung von journalistischen Grundsätzen in ihrem Hoheitsbereich. Dort trägt die Presse ebenfalls die Verantwortung von User Generated Content, der auf ihrer Plattform veröffentlicht wird. Daher müssen die von Nutzerinnen und Nutzern bereitgestellten Inhalte auch als solche erkennbar gemacht werden [29, Ziffer 2.7]. Trotzdem können Fehler, die unwissentlich oder nicht auf mangelnde Sorgfalt zurückzuführen sind, entstehen. Journalistinnen und Journalisten sind in diesem Fall nicht haftbar [28, S. 6].

KI als juristische Person

Bei der Frage nach der Haftung von Schäden, die durch eine KI verursacht werden, kommt kontinuierlich die Diskussion auf, ob KI als eine eigene juristische Person zu betrachten ist. Der Begriff „elektronische Person“ wird ebenfalls in dieser Diskussion genannt [34, S. 27]. In der Regelung zur zivilrechtlichen Haftung beim Einsatz von Künstlicher Intelligenz des Europäischen Parlaments aus dem Jahr 2020 wird jedoch festgestellt, dass „es nicht erforderlich ist, KI-Systemen Rechtspersönlichkeit zu verleihen“ [65, Abs. 7]. Die Begründung lautet, dass KI-Systeme als direktes Produkt eines menschlichen Schaffens gesehen werden. In Verlängerung wird also auch ein Schaden durch die KI durch einen Menschen verursacht. Daher sind diejenigen Personen haftbar, die das KI-

System betreiben oder auf es einwirken. Durch die Komplexität von KI-Systemen kann jedoch eine fehlerhafte Ausgabe nicht unbedingt direkt einer fehlerhaften menschlichen Eingabe zugeordnet werden, wodurch die Frage nach der Haftung juristisch nicht eindeutig beantwortet werden kann [65, Abs. 7].

Wird nun eine KI in den Prozess des Nachrichtenschaffens aufgenommen, ergeben sich Unklarheiten, wie die obengenannten Anforderungen an guten Journalismus auch von einer Software erfüllt werden können. Diese trägt nämlich nicht selbst die Verantwortung für ihre Erzeugnisse. Die Verantwortung für die Ergebnisse der KI liegt hier also bei ihren Anwenderinnen und Anwendern, den Journalistinnen und Journalisten. Weder die Software noch die Herstellerin bzw. der Hersteller können hier haftbar gemacht werden [21, S. 6]. Dadurch, dass KI nicht völlig zuverlässig ist, bleibt eine gewisse Unsicherheit gegenüber der Sorgfaltspflicht. Gerade bei der Verbreitung von Falschmeldungen wäre das Vertrauen in die Presse erschüttert [1, S. 57]. Daher ist es wichtig, dass nicht alle Entscheidungen der KI überlassen werden. Die finalen Entscheidungen muss der Mensch treffen [1, S. 63].

5.2 FUNKTIONALE ANFORDERUNGEN

Ziel der Einführung von KI in den Newsworkflow ist eine Effizienzsteigerung. Repetitive Aufgaben, die im Arbeitsalltag einer medienschaffenden Person viel Zeit in Anspruch nehmen, können automatisiert werden. Die Aufgaben werden ebenfalls in einer viel kürzeren Zeit abgearbeitet im Vergleich zu einem Menschen [25, S. 32].

Die Aufgaben, die von einer KI übernommen werden können, sind in drei Bereiche gegliedert: Assistierende, generative und distributive Aufgaben [28]. Unter assistierenden Aufgaben versteht man in Abbildung zwölf den Kreisabschnitt von Organize bis Discover im Uhrzeigersinn. Die KI arbeitet der menschlichen Journalistin bzw. dem menschlichen Journalisten zu, gibt Vorschläge ab (z.B. Curate) und bearbeitet das bestehende Material (z. B. Verify). Die Journalistin bzw. der Journalist nutzt dann die Vorarbeit und die Vorschläge der KI, um eigenes Material zu erzeugen. Die generativen Aufgaben erschaffen neues Material nach Auftrag des Menschen (Create). Die Erstellung erfolgt auf Basis der assistierenden Aufgaben, jedoch übernimmt die KI hier auch den nächsten Schritt der Inhaltserstellung. Dies wird von einem Menschen überwacht und beurteilt. Die distributiven Aufgaben nehmen in Abbildung zwölf den Bereich von Distribute bis Personalize im Uhrzeigersinn ein. Die KI kann aufgrund von Nutzungszahlen und -feedback herausfinden, welche Beiträge an welcher Stelle zu welcher Zeit die meiste Kundschaft ansprechen und platziert die Beiträge entsprechend.

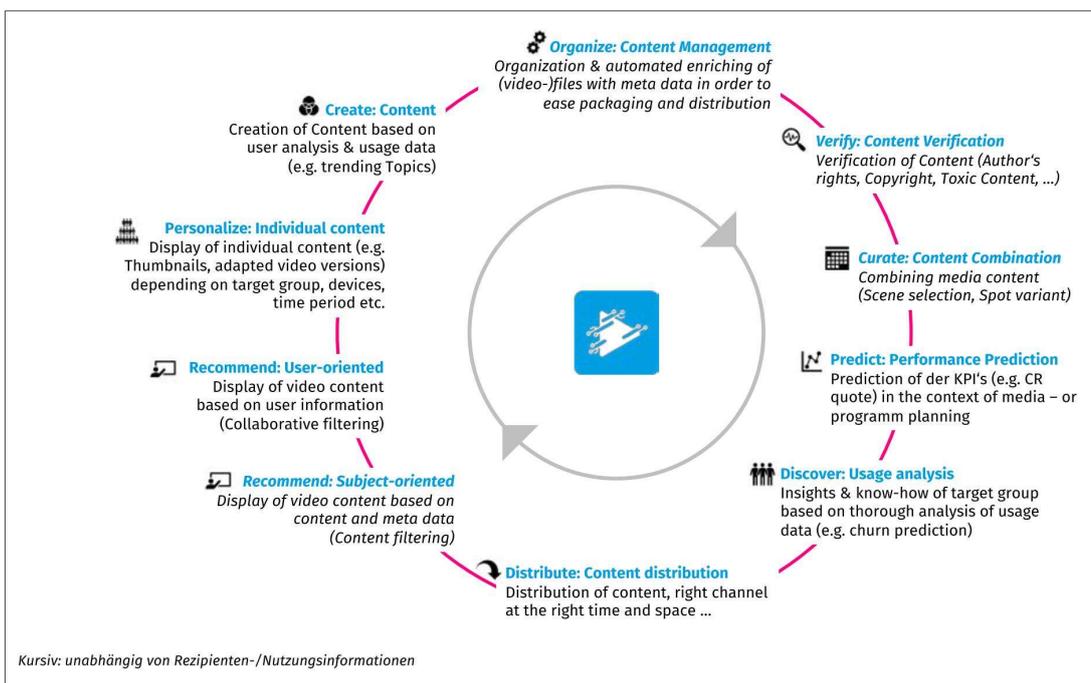


Abbildung 12: Kreislauf der Content-Wirtschaft [25, S. 33].

5.2.1 ASSISTIERENDE AUFGABEN

Assistierende Aufgaben unterstützen die Journalistin bzw. den Journalisten, indem sie eine Essenz editieren oder in eine andere Form umwandeln. Sie unterscheiden sich von den generativen Aufgaben in dem Punkt, als dass sie eine bestehende Essenz als Arbeitsgrundlage benötigen.

Themenvorschläge

Die Grundlagen für einen journalistischen Beitrag sind die Wahl eines Themas und eine gründliche Recherche. Journalistinnen und Journalisten treffen ihre Themenwahl nach Aktualität, Interessen der Leserinnen und Leser und zuletzt auch nach persönlichem Bauchgefühl. Um auf Themen aufmerksam zu werden, durchforsten sie aktuelle Weltereignisse, den lokalen Veranstaltungskalender oder auch Soziale Medien. Dabei kann die Themenwahl subjektiv und der Kreis der Auswahlmöglichkeiten unvollständig sein. Durch KI, die das Netz nach Trends durchsucht, sollen Themen gefunden werden, die der Mensch übersehen hat [56, S. 17]. Ebenfalls sollen aus der Moderation von Kommentaren von Nutzerinnen und Nutzern neue Ansatzpunkte für ein Thema gewonnen werden. Damit erhält das Publikum ebenfalls das Gefühl, dass die Redaktion die Partizipation der Nutzerinnen und Nutzer begrüßt und wertschätzt [1, S. 20].

Recherche

Durch die Suche nach Trends soll die KI auch weitere Informationen sowie Quellen zur Recherche liefern. Bei großen unübersichtlichen Themen ist es hilfreich, sich Daten von der KI vorsortieren zu lassen. Gerade bei heterogenen Datensätzen, die aus mehreren Medientypen bestehen, ist es einer KI möglich, mehr Informationen in kürzerer Zeit zu erfassen als ein Mensch. Diese ist ebenfalls in der Lage, die Daten zusammenzufassen und wiederum Schlussfolgerungen daraus zu ziehen. Trotzdem kann der Mensch nicht die alleinige Arbeit der KI überlassen. Er muss die Arbeit der KI beaufsichtigen und auf Plausibilität prüfen [37, S. 17, 56, S. 8-10, 66].

Faktenprüfung

Faktenprüfung bedeutet den Beweis, dass eine Aussage der Wahrheit entspricht, zu erbringen und damit die Glaubwürdigkeit des eigenen Mediums zu steigern. Die Faktenprüfung ist Teil der journalistischen Sorgfaltspflicht. KI-basierte Systeme sollen hierbei helfen. Sie schätzen die Glaubwürdigkeit von Quellen auf Sozialen Medien ein, gleichen die vorliegenden Informationen mit anderen Quellen ab und kennzeichnen falsche Fakten. Diese Ergebnisse werden der menschlichen Journalistin oder dem menschlichen Journalisten vorgelegt. Ihr oder ihm ist es möglich, direkt bei der Recherche die unseriösen Quellen zu umgehen und falsche Fakten zu berichtigen [67, S. 46].

Falschnachrichten werden desinformierende Inhalte genannt. Sie beinhalten manipulierende, irreführende oder (nachweislich) falsche Texte oder Bilder. Durch ihre reißerischen Inhalte erzielen diese Beiträge eine hohe Aufmerksamkeit beim Lesepublikum und werden daher von Algorithmen oft der konsumierenden Zielgruppe angezeigt [55]. Die Verbreitung von DeepFake-Videos wächst stark, deshalb müssen etablierte Medien - der öffentlich-rechtliche Rundfunk eingeschlossen – sich gerade in der glaubwürdigen Berichterstattung engagieren, um nicht das Vertrauen der konsumierenden Zielgruppe zu verlieren [68]. Daher muss gerade beim Einsatz von KI die fehlerfreie Erkennung von Falschmeldungen gewährleistet sein [1, S. 57]. Dabei soll eine technische Lösung mit KI helfen, Ursprung, Authentizität und Aussagekraft über mehrere Quellen hinweg zu überprüfen [26, S. 452].

Content Management

Eine weitere Anwendung von KI in der Wertschöpfungskette ist die Verarbeitung von Archivmaterial. Archive in Rundfunkanstalten und Onlinemagazinen sind mit einer großen Menge von Beiträgen gefüllt. Dabei kann es sich um Legacy-Material oder um den Beitrag, der gestern erst publiziert wurde, handeln. Um ein Archiv als Recherche- und Materialquelle zu nutzen, muss gewährleistet werden, dass man das Archiv schnell durchsuchen kann. Schließlich ist im Nachrichtenschaffen die Ressource Zeit nicht in großen Mengen vorhanden. Damit man ein Archiv schnell durchsuchen kann, müssen alle Beiträge – egal welcher Medienform – mit Schlagwörtern, also beschreibenden Metadaten, angereichert werden. So kann man, ähnlich wie in einer Suchmaschine, den gesuchten Begriff eingeben und man erhält eine Liste von Suchergebnissen. Wenn Menschen Beiträge mit Metadaten versehen, dauert das eine gewisse Zeit und wäre bei der großen Menge an Archivmaterial auf lange Sicht zu langsam. Daher soll die Aufgabe von einer KI übernommen werden. Die Bilder, Videos, Tonaufnahmen und Texte werden verarbeitet und ihnen Merkmale entzogen. Dies soll bei Bildern und Videos über eine Objekt-, Gesichts- und Bauchbindenerkennung erfolgen. Dabei sind zwei Schritte notwendig: Die Erkennung eines Motivs und die Klassifikation. Im ersten Schritt wird im Bild oder Video das zu erkennende Motiv mit einer Bounding Box umrahmt. Das Motiv kann ein Gesicht, ein Objekt oder Text sein. Damit ist dieser Bildbereich vom Rest des Bildes abgetrennt. Im zweiten Schritt sollen die Motive innerhalb der Bounding Boxes klassifiziert, also benannt werden. Für die Gesichts- und Objekterkennung wird anhand von vortrainierten Informationen entschieden, um wen oder was es sich handelt. Außerdem kann gespeichert werden, welches Objekt oder Gesicht an welcher zeitlichen Stelle im Video auftritt. Bei der Texterkennung werden die einzelnen Buchstaben klassifiziert. Dies bedarf aufgrund der Begrenztheit des Alphabets keinem extensiven Training [69, S. 23]. Bei der Auswertung von Audiospuren ist der erste Schritt die Transkription durch eine KI. Innerhalb des Texts werden nun Schlagworte extrahiert. Dies geschieht auch mit Material, das schon in Textform vorliegt. Anhand der Schlagwörter soll außerdem das generelle Thema und Genre des Beitrags festgestellt werden [9, S. 37] sowie bereits eine Zusammenfassung verfasst werden [70].

Die Klassifikation von Motiven wird dadurch erleichtert, dass Wörter zueinander in Beziehung gesetzt werden. Diese Beziehung ist rein inhaltlich [69, S. 24]. Dadurch kann Material deutlich einfacher im Archiv gefunden und letztendlich auch genutzt werden. Die zeitintensive Vorsortierung muss die Journalistin bzw. der Journalist also nicht mehr von Hand erledigen [66].

Eine weitere Anwendung für KI im Content Management ist die Auswahl von Drittcontent, der von externen Produktionsfirmen eingekauft wird. Hier gibt ein KI-Modell Vorschläge ab, welche Art Beitrag beim Publikum den größten Anklang findet, aufgrund von Nutzungszahlen vergangener Beiträge. Außerdem kann geprüft werden, ob der zu kaufende Content dem ähnlich ist, was sich bereits im Archiv des Medienunternehmens befindet [9, S. 34].

Assetverarbeitung

Im nächsten Schritt in der Wertschöpfungskette extrahiert die KI nicht nur Informationen aus einem Beitrag, sondern nutzt diese, um die Beiträge zu verändern und mit anderem Material zusammenzustellen. Das KI-Modell arbeitet also mit vorhandenem Material und produziert auf Basis dessen ein neues Produkt. Mögliche Anwendungen sind hier die Transkription, sprachliche Übersetzung oder der automatische Schnitt von Videomaterial. Dabei handelt es sich um repetitive Tätigkeiten, die sonst von Journalistinnen und Journalisten mit hohem Zeitaufwand durchgeführt werden müssten.

Die Transkription kommt beim crossmedialen Arbeiten zum Einsatz, beispielsweise wenn die Journalistin bzw. der Journalist aus einem Fernseh- oder Radiobeitrag einen Textbeitrag generieren möchte. Auch bei der Erstellung von Untertiteln ist Speech-to-Text hier von Nutzen. Die umgekehrte Richtung, Text-to-Speech, ist für die Arbeit mit mehreren Medienformen interessant. So können Artikel in Textform auch zum Hören verfügbar gemacht werden. Eine weitere Anwendung ist die Übersetzung von Texten. Liegt eine Quelle in einer anderen, dem Medienschaffenden unbekanntem Sprache vor, kann dies durch den Einsatz von KI verändert werden.

Auch im Bereich der Bild- und Videobearbeitung soll KI die Prozesse beschleunigen. So soll Material schnell anhand eines Algorithmus zusammengestellt oder gefiltert werden. Die Anwendungen reichen von der Detektion von besonders emotionalen Momenten [71, S. 30], über den automatischen Schnitt [56, S. 10-13] bis hin zur Erkennung und Entfernung von jugendschutzkritischen Inhalten [72, S. 14-15].

5.2.2 GENERATIVE AUFGABEN

Im Vergleich zur Assetverarbeitung, die ein bestehendes Asset verändert, wird bei der Assetproduktion ein völlig neues Asset erzeugt. Durch Natural Language Processing sollen ganze Texte nach wenigen Vorgaben generiert werden. Zunächst werden Rubriken befüllt, die keine starke Variation in ihren Formulierungen benötigen, wie Sport-, Wetter- oder Verkehrsnachrichten. Die KI befüllt vorgeschriebene Textbausteine mit den aktuellen Daten und veröffentlicht sie. So werden beispielsweise die Beiträge zum Feinstaubradar der Stuttgarter Zeitung komplett von einer KI geschrieben [37, S. 19].

KI ist bereits in der Lage komplexere Texte zu generieren und diese dem Stil einer Autorin bzw. eines Autors anzupassen [37, S. 17, 70]. Die Journalistin bzw. der Journalist gibt dem KI-Modell wenige Vorgaben und Textbeispiele und dieses arbeitet selbstständig an einem Beitrag, meist schneller als es ein Mensch könnte. Für den Radiojournalismus werden die Stimmen der Sprecherinnen und Sprecher von der KI trainiert, sodass diese nun Beiträge nicht mehr selbst vorlesen müssen, sondern die KI die Audiospur generiert [68]. Im Fernsehjournalismus werden entsprechend schon Prototypen von KI-gesteuerten Nachrichtensprecherinnen und Nachrichtensprechern eingesetzt [56, S. 10-13].

Dadurch, dass die automatisierte Berichterstattung weniger personelle Ressourcen benötigt, ist es möglich auch über Themen, die erfahrungsgemäß weniger rezipierende Personen haben, zu berichten. Dazu gehört z.B. die Berichterstattung über Sportarten wie Curling oder Bogenschießen [62].

Zusätzlich soll die KI auch Vorschläge abgeben, wie ein Beitrag ansprechender zu gestalten ist, sodass dieser von einem größeren Publikum konsumiert wird. Die Empfehlungen können eine stärkere Bebilderung, eine andere Schlagzeile oder eine Umstrukturierung sein [67, S. 47].

5.2.3 DISTRIBUTIVE AUFGABEN

Distributive Aufgaben beschreiben alle Tätigkeiten, die direkt mit dem Publikum zu tun haben. Dazu gehören die Ausspielung sowie die Beantwortung von Anfragen aus dem Publikum.

Publikation

KI soll auch bei der Veröffentlichung von Beiträgen eine wichtige Rolle spielen. Sie sammelt Daten, welche Beiträge das Publikum bevorzugt konsumiert und in welcher Form diese am gewinnbringendsten präsentiert werden [67, S. 47]. Um das zu erreichen, müssen Daten über die rezipierenden Personen extrahiert werden. Die Interessen werden aus bereits gesehenen Beiträgen geschlossen. Die Medienvorlieben einer Nutzerin bzw. eines Nutzer hängen jedoch auch zu einem gewissen Maß von ihrem Geschlecht, Alter, Beruf, Einkommen oder Familienstand ab. Auf Grundlage dieser Daten wird der konsumierenden Person dann gezielt ein passender Beitrag empfohlen, bei dem die Wahrscheinlichkeit hoch ist, dass sie ihn öffnen wird [25, S. 31]. Außerdem ist die Art der Darbietung für die Person im Publikum entscheidend. Es ist abhängig von der Schlagzeile, dem Teaserbild, der Größe oder der Position auf der Webseite, ob ein Beitrag gesehen wird oder nicht [67, S. 47]. Die Auswertung wird von Vorhersagemodellen durchgeführt [37, S. 20]. Bei der Onlinepräsentation (z.B. in Mediatheken oder bei Onlinemagazinen) kommen sogenannte dynamische Paywalls zum Einsatz. Diese leiten aus den gesammelten Daten eine Auswahl von Beiträgen, die für die Nutzerin bzw. den Nutzer interessant sein können, ab. Außerdem bestimmen sie, wie lange, mit welcher Schlagzeile und in welcher Medienform diese der Konsumentin bzw. dem Konsumenten dargeboten werden, um ihre oder seine Aufmerksamkeit möglichst lange zu erhalten [67, S. 47]. Es kann sogar in den Metadaten der Essenz gespeichert werden, für welche Zielgruppen die Beiträge geeignet sind und welche inhaltlich verwandten Beiträge ebenfalls empfohlen werden können [9, S. 34].

Für jeden Beitrag wird ein Return of Investment prädiziert. Dieser beschreibt, wie lohnenswert die Veröffentlichung des Beitrags für die Rundfunkanstalt oder Zeitung ist. Er wird anhand von bereits veröffentlichten Beiträgen, die inhaltlich oder von der Medienform her ähnlich sind, geschätzt. Damit soll auch entschieden werden, ob ein Beitrag kostenpflichtig angeboten wird [73].

Funktionsweise

Für die Personalisierung von Inhalten werden Algorithmen verwendet, die auf bestimmte Metriken ausgelegt werden. Daher müssen vor der Implementierung solcher Algorithmen Überlegungen über die Ziele und die daraus resultierenden Metriken angestellt werden. Außerdem muss geklärt werden, welche Arten von Daten verfügbar sind und mit welchen Mitteln sie gesammelt und verarbeitet werden [74, S. 79].

Es gibt mehrere Arten von Algorithmen, die für die Personalisierung eingesetzt werden können.

- **Popularity Filtering:** Basierend darauf, was die Mehrheit der Nutzerinnen und Nutzer für relevant befindet, werden den einzelnen Nutzerinnen und Nutzern diese Inhalte angezeigt.
- **Collaborative Filtering:** Dieser Algorithmus teilt die Nutzerschaft in Gruppen und prädiziert die Interessen der einzelnen Nutzerin bzw. Nutzer anhand der Ähnlichkeit zu einer Gruppe.
- **Content-Based Recommendations:** Anhand der Inhalte, die die Nutzerin bzw. der Nutzer bereits konsumiert hat, werden ähnliche Beiträge vorgeschlagen.
- **Contextual Recommendations:** Diese Analyse zieht Informationen aus kontextuellen Nutzungsdaten (Ort und Zeit, an dem der Inhalt konsumiert wird) und schlägt dazu passende Inhalte vor.
- **Public Service Algorithms:** Diese schlagen Inhalte vor, die von öffentlichem Interesse sind [74, S. 79-80].

Erhält ein Medium eine neue Nutzerin bzw. einen neuen Nutzer, sind noch keine Daten über sie oder ihn bekannt. In diesem Fall wird der Collaborative Filtering-Ansatz verwendet. Ähnliches passiert, wenn ein Beitrag neu in einem Medium veröffentlicht wird. Es sind noch keine Nutzungsdaten zu diesem Beitrag angelegt worden. Hier werden Content-Based Recommendations genutzt [74, S. 79]. Daten fallen in unterschiedlichen Formen an: Daten zur Beschreibung der Inhalte, Nutzungsdaten und Daten über den Nutzungskontext: Daten können aktiv von Nutzerinnen und Nutzern abgegeben werden, indem diese Feedback geben, nach Stichwörtern suchen oder sich zwischen Inhalten entscheiden sollen. Dabei handelt es sich aber nur um einen Teil der Daten. Implizit gegebene Daten machen den anderen Teil der Daten aus. Dabei handelt es sich um die Speicherung von Parametern wie der Besuchshistorie, der konsumierten Beiträge und dem Zeitraum, den eine Nutzerin bzw. ein Nutzer mit einem Beitrag verbringt [74, S. 80-81]. Für den Entwurf eines Personalisierungsalgorithmus muss ebenfalls feststehen, ob die Personalisierung an jede Nutzerin bzw. jeden Nutzer individuell angepasst oder für alle Nutzerinnen und Nutzer gleich präsentiert werden soll [74, S. 81].

Schließlich müssen ebenfalls die gewünschten Messmethoden und Metriken für die Optimierung der Algorithmen ausgewählt werden. Dabei müssen Kriterien wie Transparenz und Erklärbarkeit in Betracht gezogen werden. Außerdem muss definiert werden, wie akkurat die Personalisierung ausfallen soll [74, S. 87]. Mögliche Metriken sind Reichweite, Engagement, Loyalität und Auswirkung. Davon lassen sich wiederum nicht alle mit Messmethoden in Zahlen darstellen (siehe Abbildung 13).

Metrik	Definition	Messmethode
Reichweite	Anzahl Personen, die das Medium nutzen	Anzahl der gesamten Nutzer*innen des Mediums, Anzahl der angemeldeten Nutzer*innen zu einem bestimmten Zeitpunkt
Engagement	Zeitspanne, die eine Nutzer*in mit einem Beitrag verbringt	Anzahl der konsumierten Beiträge einer Nutzer*in, Zeit, die mit dem Medium verbracht wird, Scrolltiefe, Anzahl angezeigte Seiten pro Besuch
Loyalität	Frequenz der wiederkehrenden Besuche des Mediums	Anzahl der Besuche des Mediums, Anzahl Links, die ein Nutzer generiert, um den Beitrag auf Sozialen Medien zu teilen
Auswirkung	Auswirkung des Beitrags auf das individuelle oder gesellschaftliche Leben	Keine einheitliche Metrik

Abbildung 13: Metriken und Messmethoden der Personalisierung nach [74, S. 52] und [74, S. 54].

Bei der Auswahl der Metriken und deren Messmethoden muss für jedes Medienunternehmen ein eigener Weg gefunden werden. Je nach Ziel und Beschaffenheit des Mediums setzt jeder Newsroom andere Prioritäten. Unumstößlich ist jedoch, dass eine Kombination aus mehreren Metriken und Messmethoden das beste Ergebnis liefert [74, S. 52].

Kommunikation mit Leserinnen und Lesern

Als zweiter Schritt in der Publikation folgt die Rückmeldung des Publikums. Dieses kann seine Meinung über Brief, Telefon, Email oder Kommentare zum Ausdruck bringen. Kommentare werden auf der Webseite des Mediums veröffentlicht, damit andere Nutzerinnen und Nutzer auf diese eingehen und weiter diskutieren können. Dabei handelt es sich immer wieder um unangebrachte, beleidigende und diskriminierende Aussagen. Diese müssen vor der Veröffentlichung von der Redaktion gelöscht werden. Hier soll eine KI die Kommentare vorsortieren und Empfehlungen für die Löschung aussprechen. Das KI-Modell durchsucht die Kommentare nach bestimmten Wörtern [62].

Auch bei der Beantwortung der anderen Kommunikationsformen soll die KI die Arbeit erleichtern. Diese können Brief, E-Mail, Telefon, Chat, Kommentar auf der Webseite des Mediums sowie Direct Message, Videochat oder Kommentar über Soziale Medien sein [21, S. 160, 21, S. 158]. Da die Leserin bzw. der Leser eine schnelle Antwort erwartet, kann sie entweder durch einen Menschen oder durch eine Maschine erfolgen. Je nachdem, wie hoch die verfügbare Kapazität bei den menschlichen Mitarbeiterinnen und Mitarbeitern ist, kommt es immer häufiger vor, dass Anfragen von Leserinnen und Lesern von Maschinen beantwortet werden. Dabei ist es wichtig, dass zunächst die Maschine fehlerfrei und für die Leserin bzw. den Leser auf annähernd menschliche Weise antwortet, als auch dass sämtliche Kommunikation mit der Leserin bzw. dem Leser im MAM des Mediums gespeichert

wird. Die Einbindung ins CMS ist von Vorteil, damit die Leserückmeldung auch wiederum als Material für zukünftige Recherchen verwendet werden kann.

5.2.4 BENUTZBARKEIT

Neben den konkreten Use Cases für den Einsatz von KI sind jedoch auch Anforderungen an die Benutzbarkeit des KI-Modells von Wichtigkeit. Die Entwicklung einer KI-gestützten Anwendung muss in Zusammenarbeit mit den primären Nutzerinnen und Nutzern erfolgen [25, S. 30]. Dabei handelt es sich um einen iterativen Prozess. Anwenderinnen und Anwender sollen bei jeder Auslieferungsstufe ihr Feedback zur Anwendung geben und die Ziele für die nächste Iteration mitbestimmen. Dabei ist jederzeit zu klären, inwieweit die Anwendung den Arbeitsalltag der Anwenderinnen und Anwender voranbringt und inwiefern sie den Wert des geschaffenen Produkts steigert. Daraus ergeben sich die Vorgaben für die technische Lösung (siehe Abbildung 14) [25, S. 33].

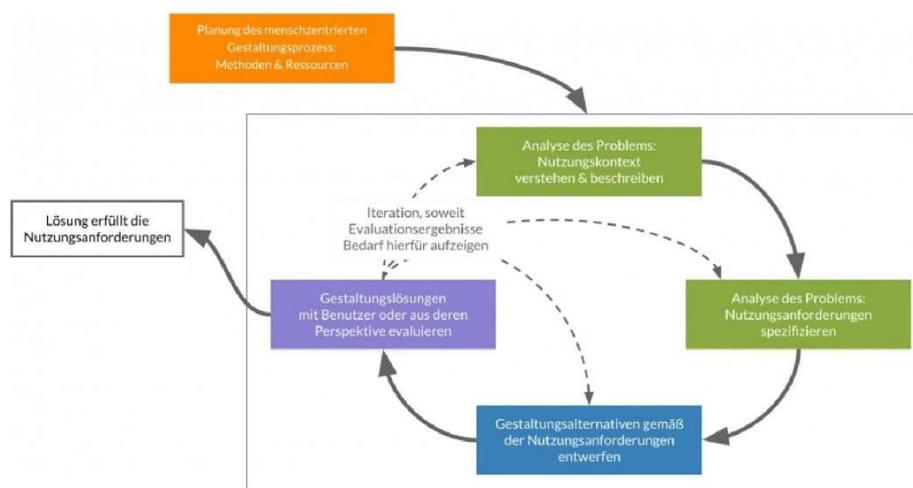


Abbildung 14: Menschzentrierte iterative Gestaltung eines Produkts [75].

*Austausch zwischen CMS, MAM, KI und Nutzer*in*

Für den reibungslosen Austausch von Daten zwischen den Akteuren des Workflows muss eine Schnittstelle zwischen dem zentralen Speicher der Rundfunkanstalt, den Nutzerinnen und Nutzern und dem KI-Modell existieren. Insbesondere, wenn Assets durch die KI bearbeitet oder zusammengefasst werden sollen, muss sie selbstverständlich Zugang erhalten. Es muss also die Unterscheidung zwischen automatischen Prozessen wie der Extraktion von Metadaten aus Essenzen und manuellen Prozessen wie die sprachliche Korrektur eines menschenverfassten Beitrags erfolgen. Automatische Prozesse laufen ohne die Aktion eines Menschen im Hintergrund ab. KI und CMS kommunizieren in maschinenlesbaren Formaten miteinander. Eine wichtige Voraussetzung ist außerdem, dass die KI das hauseigene Metadatenformat des Medienunternehmens versteht, sonst ist keine Kompatibilität zwischen den Systemen gewährleistet. Eventuell muss ein Mapping stattfinden.

Manuelle Prozesse werden von einem Menschen ausgelöst und bei der KI in Auftrag gegeben. Dabei ist es wichtig, dass für die menschlichen Akteure eine grafische Benutzungsoberfläche gestaltet wird. Sie kann direkt ins MAM integriert werden. Bei der Interaktion zwischen Mensch und KI muss die

KI jedoch im menschenlesbaren Format antworten. Bei jeder Art der Kommunikation mit KI muss außerdem darauf geachtet werden, dass die Rechen- und Übermittlungszeiten so gering wie möglich gehalten werden. Dies hängt jedoch davon ab, in welchem Servicemodell das KI-Modell angeboten wird. Befindet sich das Modell on premise, sind die Übertragungszeiten schätzungsweise gering. Befindet es sich jedoch off premise oder in der Cloud, muss auf die Begrenzung der Latenzzeiten geachtet werden.

Rollenbasierte Zugänge

Für die rollenbasierten Zugänge müssen mehrere Gruppen von Benutzerinnen und Benutzern gebildet werden. Diese sind zunächst die Frontend Personas. Sie sind diejenigen, die mit dem KI-Modell produktiv über eine grafische Benutzungsoberfläche arbeiten. Sie sind nicht tiefer mit der Funktionsweise des KI-Modells vertraut, kennen jedoch dessen Anwendungsbereiche und wissen, es zu nutzen.

Sogenannte Plattform Personas sind die Personen, die mit dem KI-Modell und seinen Trainingsdaten arbeiten, aber kein Graphical User Interface (GUI) benötigen. Dabei handelt es sich um Data-Scientistinnen und Data-Scientisten. Sie überwachen die Arbeit der KI und sorgen für die Erhaltung der Qualität.

Die dritte Gruppe benötigt weder GUI noch Zugang zu den Daten des KI-Modells. Die IT (Informationstechnik)-Administratorinnen und -administratoren arbeiten an der Einbindung des KI-Modells in das bestehende System des Medienunternehmens. Sie stellen Verbindungen zwischen den Systemen über Programmierschnittstellen her und ermöglichen den verlustfreien Informationsaustausch. Dafür muss der Output des KI-Modells in entweder maschinen- oder menschenlesbares Format gebracht werden. Das maschinenlesbare Format muss darüber hinaus noch in das hauseigene Metadatenformat des Medienunternehmens überführt werden [9, S. 35-36].

Infrastruktur

Eine weitere Anforderung, die die Benutzbarkeit erhöht, sind eine hohe Verarbeitungsgeschwindigkeit des KI-Modells. Diese setzt sich aus der Übertragungsgeschwindigkeit von Informationen von und zum KI-Modell und der Rechengeschwindigkeit zusammen. Für Aufgaben wie die Bearbeitung von Beiträgen muss das KI-Modell den ursprünglichen Beitrag aus dem Speicher der Rundfunkanstalt erhalten, daher müssen auch dort hohe Übertragungsgeschwindigkeiten möglich sein. Ist dies nicht gewährleistet, kann es zu einer Überlastsituation kommen.

Ebenfalls ein Kriterium für die Benutzbarkeit ist die Verfügbarkeit. Mögliche Ausfallzeiten des KI-Modells müssen so gering wie möglich gehalten werden, um die effiziente Arbeit der Journalistinnen und Journalisten aufrecht zu erhalten. Dadurch, dass die KI nur unterstützend tätig ist, gibt es keine Tätigkeit, die nicht auch ohne KI auszuführen wäre. Die Arbeit ohne KI würde nur mehr Zeit in Anspruch nehmen. Eine Ausfallzeit löst also keinen Havariefall aus, trotzdem wird die Arbeit dadurch umständlicher.

5.3 ORGANISATORISCHE ANFORDERUNGEN

Die Einführung von KI in den Newsworkflow umfasst neben tiefgreifenden Veränderungen im technischen Konzept auch Veränderungen in der Verteilung der Arbeit und ihrem Ablauf. Der Beginn einer jeden Veränderung bedarf einer übergeordneten Strategie. Sie gibt die Richtung für Ziele und dadurch auch Aufgaben und Ressourcen innerhalb des Veränderungsprojekts vor. Anhand dieser Strategie können nun Abläufe innerhalb des Medienunternehmens optimal ausgerichtet, erweitert, verknüpft oder abgeschafft werden, um eine möglichst schnelle und effiziente Weise des Nachrichtenschaffens zu etablieren. Diese Prozesse werden von und mit Ressourcen bearbeitet, die Prozessschritte werden damit einzelnen Ressourcen zugeordnet. Mit der Einführung von KI erhält sie auch den Status einer Ressource, da sie annähernd menschliche Arbeiten verrichtet. Menschliche Ressourcen arbeiten innerhalb von Strukturen, die Verantwortlichkeiten und Entscheidungsbefugnisse definieren. Bekannt ist die sogenannte Linienstruktur, jedoch werden diese Gebilde immer mehr aufgebrochen und flexiblere Strukturen eingeführt [76].

Die Einheiten innerhalb der Struktur bearbeiten die unterschiedlichsten Aufgaben. Daher gibt es in den Einheiten die Annahme von Werten und Normen, die zur Erreichung ihrer Ziele beitragen. Diese werden von bestehenden Mitarbeiterinnen und Mitarbeitern gepflegt und an neue Mitglieder weitergegeben. Das Verhalten wird von der Leitung des Medienunternehmens vorgelebt und von den Mitarbeiterinnen und Mitarbeitern widergespiegelt [76]. Wird eine Modernisierung der vorhandenen Kultur angestrebt, bedeutet dies eine fortschreitende Evolution hin zu neuen Werten. Dabei darf der Wandel keine kurzfristige Richtungsänderung sein, sondern eine reflektierte Anpassung an neue Technologien und Arbeitsweisen [76].

5.3.1 AUFBAUORGANISATION

Die Aufbauorganisation beschreibt die Identifikation und Bündelung von Aufgaben sowie die Verteilung auf Mitarbeiterinnen und Mitarbeiter innerhalb von Strukturen. Für die Bearbeitung der Aufgaben werden außerdem Befugnisse und Verantwortlichkeiten an die jeweiligen Personen übertragen [76].

Das Newsroom-Prinzip

Ein Newsroom ist ein zentraler Ort, an dem alle Informationen zu einer Sendung gesammelt, sortiert und aufbereitet werden. Der Newsroom ist in den meisten Medienhäusern ein großer Raum, in dem Journalist*innen, Redakteur*innen, Grafiker*innen und Mediengestalter*innen zusammensitzen. Diese kurzen Kommunikationswege sind für das Prinzip des Newsrooms Voraussetzung, denn Newsrooms sind aus dem crossmedialen Arbeiten entstanden. Ausgehend von diesem Raum werden mehrere Informationskanäle bespielt. Synergien zwischen den einzelnen Medienformen werden genutzt [77, S. 140-141]. Im Newsroom sind die Rollen der Informationsgewinnung und der Informationsaufbereitung voneinander getrennt. Dabei steht die Anwendbarkeit des Materials auf mehrere Ausspielkanäle im Vordergrund [77, S. 141]. Die Aufsicht und Verantwortung über die Inhalte hat die Chefredaktion (siehe Abbildung 15) [77, S. 140-141].

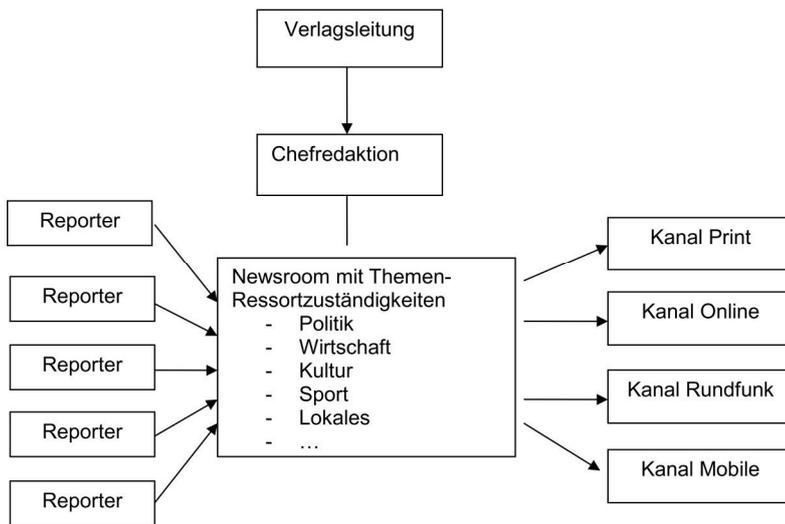


Abbildung 15: Newsroom-Organisation [77, S. 141].

Erweiterung um KI

Wird nun der Arbeitsaufbau der Redakteure im Newsroom um eine KI erweitert, werden Aufgaben, die vorher ein Mensch bearbeitet hat, der KI übertragen. Das erwünschte Ziel ist, durch parallelisierte Arbeit von Mensch und KI Zeit und Kosten zu sparen. Die KI soll die menschliche Arbeit ergänzen. Übernahme die KI die Routineaufgaben der Journalistinnen und Journalisten, bliebe den Menschen mehr Zeit für kreative Arbeiten. Dies würde zu einem Entlastungseffekt führen [28, S. 5].

Die Arbeit mit einer KI soll zu einer geringeren Time to Delivery führen. Dadurch sollen mehr Informationen verarbeitet und mehr Beiträge über verschiedene Kanäle ausgespielt werden. Der erhöhte Informations- und Arbeitsfluss kann einen höheren Koordinationsaufwand erfordern [28, S. 5].

Schließlich muss festgehalten werden, dass es im Prozess des Nachrichtenschaffens noch immer viele Aufgaben gibt, die eine KI nicht bewältigen kann und die zwangsläufig eine menschliche Journalistin bzw. ein menschlicher Journalist ausüben muss. Dazu zählen viele Bereiche der Informationsbeschaffung sowie das Verstehen von menschlichen Gefühlen. Aufgaben wie das Erfassen von Zusammenhängen, das Einordnen und Kommentieren von Informationen können in absehbarer Zeit ebenfalls nicht vollständig von einer KI übernommen werden [28, S. 11].

5.3.2 ABLAUFORGANISATION

Die Ablauforganisation beschreibt das räumliche und zeitliche Zusammenwirken der an der Aufgabenerledigung beteiligten Ressourcen. Die Aufgaben werden in zeitliche Arbeitsabläufe organisiert. Ziel eines solchen Ablaufs ist die Leistungserstellung. Im Falle einer Rundfunkanstalt wäre die Leistung die Ausspielung einer Sendung. Ein Prozess wird über einen Input angestoßen. Dann werden parallel oder nacheinander Tätigkeiten begonnen, durchgeführt und deren Teilergebnisse einem weiteren Prozessschritt zugeführt. Von einem Prozessschritt zum nächsten kann eine Abnahme des Produkts erfolgen. Das hat zur Folge, dass ein Prozessschritt eventuell

wiederholt werden muss, falls das Produkt nicht die ausreichende Qualität aufweist. Wird beispielsweise das Konzept für einen Beitrag von der Redaktion als nicht ausgereift genug empfunden, muss die Journalistin bzw. der Journalist dieses überarbeiten. Das Gesamtergebnis des Prozesses muss dem Publikum einen Mehrwert verschaffen [76].

Prozesse werden dadurch beschrieben, dass ihr Ablauf schematisch dargestellt wird sowie ihre Merkmale festgehalten werden. Diese sind die Prozessbeteiligten, verwendete Ressourcen, angewendete interne Regelungen und die angesiedelte Abteilung. Außerdem werden die Prozesse nach ihrer Auftretenshäufigkeit und Wichtigkeit priorisiert. Dabei können redundante und wenig verwendete Prozesse ebenfalls identifiziert werden [76]. Zur weiteren Analyse können Gesichtspunkte wie der Ressourcenverbrauch und die Kosten betrachtet werden. Als Konsequenz werden Prozesse verschlankt oder ganz abgeschafft. Eventuell kann ein Prozess parallelisiert und damit beschleunigt werden. Schnittstellen zwischen Systemen und Personen sorgen ebenfalls für Verzögerungen. Dabei treten auch Schwachstellen der Prozesse zu Tage. Schließlich werden die Prozesse nach der Qualität ihres Outputs bewertet. Auch Rückmeldungen des Publikums können Aufschluss über den Stand der internen Prozesse geben [76]. Den Prozessen übergeordnet, ist das Prozessmanagement. Dieses entspringt aus der generellen Strategie des Medienunternehmens. Es sorgt dafür, dass die Ziele des Medienunternehmens erreicht werden, und überwacht die Prozesse anhand von Kennzahlen. Prozessmanagement führt also dazu, dass Prozesse langfristig neugestaltet, gesteuert und optimiert werden, während die Prozessanalyse auf rein operativer Ebene kleine Veränderungen an Prozessen vornimmt [76].

Arbeit in Medienunternehmen

In Medienunternehmen sind der Zeitpunkt der Fertigstellung von Aufgaben und ihre Abfolge stark vom Ausspielungszeitpunkt eines Beitrags abhängig. Es wird vom Zeitpunkt der Veröffentlichung rückwärts gerechnet, um den Startzeitpunkt, an dem mit der Planung und Produktion des Beitrags begonnen wird, zu ermitteln. Bei einem Nachrichtenbeitrag im Rundfunk wird in der Regel sechs bis zwölf Stunden vorher mit der Arbeit begonnen. Bei sehr aktuellen Meldungen (Breaking News) ist die Bearbeitungszeit entsprechend kürzer [77, S. 151-153]. Anhand dieser Rechnung kann man also die erforderlichen Produktionsschritte festmachen, die aufgrund des Zeitmangels stark auf Effizienz und Zeitersparnis ausgelegt werden müssen [77, S. 146].

Journalistinnen und Journalisten sind in ihrer täglichen Arbeitszeit nicht ausschließlich mit dem Erstellen von Inhalten beschäftigt, sondern auch mit Recherche, Selbstorganisation und Kontakt zu Nutzerinnen und Nutzern. Letztendlich bearbeiten Journalistinnen und Journalisten ungefähr 60 Prozent der Zeit ihre Haupttätigkeiten und 40 Prozent ihre Nebentätigkeiten [77, S. 154]. Eine KI kann sowohl die Haupt- als auch die Nebentätigkeit unterstützen und Zeit einsparen.

5.3.3 PERSONALORGANISATION

Personalorganisation beschreibt die Tätigkeiten, die dafür sorgen, dass Mitarbeiterinnen und Mitarbeiter für bestimmte Arbeitsprozesse und in ausreichender Zahl und mit passenden Qualifikationen zur Verfügung stehen. Außerdem umfasst der Begriff die Sicherung von Arbeitsbedingungen sowie die Gegenleistung für die erbrachte Arbeit (Zahlung von Gehalt und Sozialleistungen) [77, S. 27]. Dazu kommen die funktionelle Zusammensetzung von Arbeitsteams und die Verteilung von Aufgabenpaketen [77, S. 28].

Ein Ansatz der Personalplanung ist die prozessorientierte Planung. Dabei wird bestimmt, welche Arbeitsschritte und damit welche Ressourcen ein Prozess benötigt. Anhand dieser Überlegung wird berechnet, wie viele unterschiedlich qualifizierte Mitarbeiterinnen und Mitarbeiter ein Prozess benötigt [77, S. 67]. Die Produktion einer täglichen viertelstündigen Nachrichtensendung bedarf modellhaft folgender Werte:

Prozessschritt	Ressourcenbedarf in Stunden			
	Reporter*innen	Redaktionsleitung	Redaktion	Technik
Vorbereitung, Planung		3	4	3
Recherche, Informationsbeschaffung	30		15	3
Produktion	3		30	15
Endkontrolle		1		1
Leitungsaufgaben		7		
Summe	33	11	49	22
Benötigte Anzahl Mitarbeiter*innen ⁴	3-4	2 ⁵	4-5	2-3

Abbildung 16: Modellhafter Ressourcenbedarf für eine tägliche viertelstündige Nachrichtensendung nach [77, S. 68].

Um nun eine Nachrichtensendung innerhalb eines festgelegten zeitlichen Rahmens fertigzustellen und auszuspielen, muss vom Zeitpunkt der Ausspielung rückwärtsgerechnet werden und die benötigten Arbeitsstunden auf Teammitglieder aufgeteilt werden. Daraus wird die benötigte Anzahl berechnet (siehe Abbildung 16).

Neuer Personalbedarf

Mit der Einführung von KI als Unterstützung im Newsroom verändern sich die Aufgabenzuweisungen und damit auch die Rollen innerhalb des Produktionsprozesses. Für die Einführung, Tests und Inbetriebnahmen werden in jedem Fall KI-Spezialistinnen und -Spezialisten sowie Data-Scientistinnen und Data-Scientisten benötigt. Diese finden weitere Beschäftigung in der

⁴ Es wird von einer Arbeitszeit von zwölf Stunden vor dem Sendezeitpunkt ausgegangen.

⁵ mit Kapazität für die Redaktion weiterer Sendungen.

Verwendung im Produktiveinsatz im Medienunternehmen. Schließlich muss in regelmäßigen zeitlichen Abständen der fehlerfreie Einsatz der KI kontrolliert werden [1, S. 44, 28, S. 11].

Personalentwicklung

Bei der Einführung von KI ist jedoch auch darauf zu achten, dass auch die bestehenden Mitarbeiterinnen und Mitarbeiter bei der Veränderung mitgenommen werden. Gezielte Maßnahmen der Personalentwicklung, die auch über längere Zeit hinweg die Leistungs- und Beschäftigungsfähigkeit der Mitarbeiterinnen und Mitarbeiter erhalten und fördern, sollten auch hier Teil des Einführungsprozesses sein. Dadurch erkennen die Journalistinnen und Journalisten den Mehrwert, der durch die neue Technologie Einzug in das Medienunternehmen hält [1, S. 45, 77, S. 103]. Die Journalistinnen und Journalisten erhalten also Zusatzwissen über die KI, mit der sie arbeiten werden.

6 VORSCHLAG EINER STRATEGIE UNTER EINBINDUNG VON CHATGPT

Die Thesis behandelt die Einbindung eines LLMs in den Newsworkflow. Die Strategie für die Einbindung erfolgt am Beispiel des Chatbots ChatGPT. Dieser basiert auf dem Modell GPT-3.5 von OpenAI. Nach der Definition der Anforderungen soll nun eine Herangehensweise dargelegt werden, mit der ChatGPT in einem Medienunternehmen im Newsworkflow eingesetzt werden kann. Die Herangehensweise besteht aus der Formulierung einer Strategie sowie der Sammlung mehrerer Konzepte für die Einhaltung der vorher genannten Anforderungen.

Für den Zeitraum der Einführung muss stark in die Kommunikation mit den Mitarbeiterinnen und Mitarbeitern investiert werden. Das Kapitel gibt Empfehlungen dazu. Außerdem werden Werkzeuge zum Projektmanagement sowie eine mögliche Roadmap präsentiert. Die Zusammenarbeit verläuft gemeinschaftlich mit dem Geschäftspartner OpenAI. Da Medienunternehmen vom Umfang des Angebotes des Konzerns abhängig sind, müssen sie ihre Prozesse und Infrastruktur daran anpassen. In Gegenzug muss OpenAI die Anforderungen aus Kapitel fünf erfüllen können. Im Folgenden werden Maßnahmen zur Erfüllung der rechtlichen und ethischen Anforderungen präsentiert, die zusätzlich zu den Funktionalitäten von ChatGPT vor der Einführung umgesetzt werden müssen. Schließlich werden die Arbeitsabläufe an den Stellen, an denen ChatGPT eingefügt wurde, beschrieben. Die Workflows unterscheiden sich bis auf die Integration der KI nicht tiefgreifend von denen in Kapitel vier.

Der neue Akteur in den Workflows zieht jedoch personelle Veränderung nach sich, weshalb sich bestehende Rollen im Medienunternehmen verändern. Zusätzlich werden neue Rollen kreiert. Diese sind entweder temporär, also nur für die Einführungsphase, oder längerfristig geplant. Das Ziel der Einführung von ChatGPT in den Newsworkflow ist nicht die Freisetzung von Personal. Zu den neuen Rollen kommen Nutzungsprofile, die mehrere Berufsgruppen einschließen und, aus Gründen der Sicherheit, jeweils nur einen beschränkten Zugang zu ChatGPT bieten.

Trotzdem ist es wichtig, dass das vorhandene Personal neue Kenntnisse über den hinzugekommenen Akteur erwirbt, um ein Verständnis für die neue Art der Arbeit zu entwickeln. Der Weg, diese neuen Kenntnisse zu erhalten, wird im Schulungskonzept beschrieben. Bei jeder tiefgreifenden Veränderung muss im Vorfeld eine Abschätzung der möglichen Risiken geschehen, um mit der gebührenden Verantwortung an die neue Materie heranzugehen und Untiefen zu umschiffen. Außerdem liefert eine Risikoabschätzung konkrete Handlungsempfehlungen, wenn denn das geschätzte Risiko tatsächlich eingetreten ist. Diese können während des und nach dem Einführungsprozess als Hilfen gegen das Scheitern des Projekts genutzt werden. Als Konsequenz von Abschnitt 5.3 muss ebenfalls eine wirtschaftliche Analyse der Veränderung erfolgen. Schließlich muss festgestellt werden, ob sich die Veränderung finanziell lohnt. Da Medienhäuser immer öfter auch einem finanziellen Druck unterstehen, muss dieser Aspekt ebenfalls beleuchtet werden.

6.1 AUSRICHTUNG DES PROJEKTS

Die Planung, Implementierung und Einführung von ChatGPT in den Prozess des Nachrichtenschaffens erfordert eine agile Herangehensweise. Aufgrund der Neuartigkeit der Unternehmung kann zu Beginn des Projekts noch nicht spezifiziert werden, was die Anforderungen an das Produkt sein werden. Daher muss der Einführungsprozess iterativ erfolgen.

Vision

Das Ziel der Einbindung von ChatGPT in den Newsworkflow ist, den Prozess des Nachrichtenschaffens schneller durchlaufen zu können und damit in der gleichen Zeit mehr Themen in eine veröffentlichungsfertige Form zu bringen und auszuspielen (A41)⁶. Durch die höhere Effizienz erhofft sich das Medienunternehmen höhere Zahlen an Nutzerinnen und Nutzern und damit eine längere Nutzungsdauer, außerdem eine höhere Reichweite und Attraktivität für möglichst viele Personengruppen. Dadurch sollen, wenn möglich, mehr Abonnements abgeschlossen bzw. für kostenpflichtige Inhalte gezahlt werden. Der Fortschritt wird an der Anzahl Use Cases gemessen, die nun mit KI abgebildet werden können. Der Erfolg wird jedoch an Anzahl der Beiträge innerhalb einer bestimmten Zeit und an der Zahl der Nutzerinnen und Nutzer, die dadurch erreicht werden, gemessen.

Strategie (A39)

Die Beschleunigung des Prozesses wird dadurch erreicht, dass eine KI die wiederkehrenden monotonen Aufgaben im Alltag der Mitarbeiterinnen und Mitarbeiter übernimmt und diese gegebenenfalls schneller ausführt als der Mensch. Mensch und Maschine können also parallel arbeiten. Der Mensch kümmert sich in der Zwischenzeit um kreative Aufgaben, die die KI nicht ausführen kann (A42). Die Einbindung von ChatGPT wird nur mit der Akzeptanz der Anwenderinnen und Anwender erfolgreich, daher müssen diese eng in den Prozess miteingebunden werden. Über frühzeitige Information, Schulung und Feedback soll dies erreicht werden. Durch die agile Arbeitsweise soll ChatGPT individuell an die Wünsche der Anwenderinnen und Anwender angepasst werden. Bei der Einführung von ChatGPT handelt es sich um eine umfassende Veränderung, daher werden Expertinnen und Experten mit verschiedenen Kenntnissen teilweise temporär, teilweise dauerhaft herangeholt. Priorität hat die journalistische Qualität. Alle Aktionen innerhalb des Projekts dienen zu deren Unterstützung.

Umsetzung

Zu Beginn der Umsetzungsphase wird die Vision sowie der Zeitplan, in dem die Strategie umgesetzt wird, im gesamten Medienunternehmen kommuniziert, sodass jede Mitarbeiterin und jeder Mitarbeiter über die kommenden Veränderungsmaßnahmen informiert ist. Wichtig ist, den Mitarbeiterinnen und Mitarbeiter früh die Fähigkeiten zur Nutzung von ChatGPT zu lehren, da sonst eine Abneigung gegen Unbekanntes entstehen könnte. Durch Schulungen, Demonstrationen und

⁶ Siehe Nummerierung in der Anforderungsmatrix in Anhang eins.

Möglichkeiten zum selbstständigen Testen soll das Vertrauen der Mitarbeiterinnen und Mitarbeiter gewonnen werden. Um die Mitarbeit der Anwenderinnen und Anwender zu aktivieren, wird regelmäßiges Feedback eingeholt und in den Product Backlog aufgenommen. Durch einfache und sichtbare Erfolge wird die Akzeptanz der Mitarbeiterinnen und Mitarbeiter ebenfalls gesteigert. Diese können die Implementierung einfacher Use Cases sein wie beispielsweise die Übersetzung von Texten durch ChatGPT. Danach werden erst komplexere Ziele in Angriff genommen. Sie liefern dann eine echte Wertschöpfung. Ein Beispiel ist die Extraktion von Metadaten aus Archivmaterial und die Erstellung eines mehrdimensionalen Vektorraums zur Verortung des Materials. Hier ist es wichtig, in Qualität zu investieren. Schließlich wird die neue Arbeitsweise in die Kultur des Medienunternehmens aufgenommen und auch in der Öffentlichkeit kommuniziert. Eine Erläuterung der Projektmanagementtechnik und Ideen zum Product Backlog befinden sich auf dem Datenträger in Kapitel vier und fünf.

Erfolgskontrolle

Um die Erfolge, die die Einführung von ChatGPT in den Newsworkflow mit sich bringt, auch sichtbar zu machen, müssen diese gemessen werden. Zunächst wird die Zeitdauer, die vorher benötigt wurde, um ein Thema sendefertig zu machen, derjenigen gegenübergestellt, die es nun mit der Unterstützung von ChatGPT dauert. Außerdem kann die Anzahl der Beiträge, die zu Nischenthemen erstellt werden, aufgenommen werden. Die gestiegenen Nutzungszahlen und die erhöhte durchschnittliche Verweildauer einer Nutzerin oder eines Nutzers auf dem Medium sind ebenfalls Anhaltspunkte für eine erfolgreiche Implementierung von ChatGPT. Über Soll-Ist-Vergleiche sowie Marktstudien ist es möglich, sich vom Fortschritt des Projektes zu überzeugen.

*Einbindung von Mitarbeiter*innen*

Bei der Mitnahme der Mitarbeiterinnen und Mitarbeiter soll das Konzept der sicheren Basis angewendet werden. Es besagt, dass Menschen sich nur gegen Veränderung wehren, wenn ihnen eine sichere Basis fehlt. Da Menschen von Natur aus neugierig und lernwillig sind, steht ihnen nur das Gefühl der Ungewissheit und Unsicherheit im Wege [74, S. 120]. Die sichere Basis ist hier eine frühzeitige Einbindung der Mitarbeiterinnen und Mitarbeiter sowie breit gestreute Informationen zum Thema KI. Den Anwenderinnen und Anwendern muss begrifflich gemacht werden, wie ChatGPT arbeitet und wie ein Output zustande kommt [1, S. 91]. Es ist nicht gefordert, dass die Anwenderinnen und Anwender ein tiefes Verständnis der Datenwissenschaft erhalten. Ein oberflächliches Wissen jedoch sorgt für ein sicheres Gefühl bei der Nutzung [74, S. 119]. Können die Mitarbeiterinnen und Mitarbeiter den Veränderungsprozess durch Feedback mitgestalten, ist die Schwelle ebenfalls niedriger, da sie das Gefühl haben, selbst die Veränderung herbei zu führen [74, S. 120].

6.2 EIGNUNG VON CHATGPT

OpenAI erlaubt die gewerbliche Nutzung seiner Modelle. Es ist jedoch keine Instanz on premise möglich. Die Kosten für die Nutzung der KI-Modelle werden in Token abgerechnet. Token sind logische Einheiten, in die der Input zerlegt und bearbeitet wird. Der Output besteht ebenfalls aus einer Anzahl Token [78].

Erfüllung der funktionellen Anforderungen

ChatGPT erfüllt als LLM sprachliche Use Cases. Das KI-Modell liefert Vorschläge zu Themen (A17), fasst Texte zusammen (A18) und extrahiert Schlagwörter aus einem Text (A20). Es ist in der Lage, Sprache zu übersetzen (A23) und diese sprachlich zu überarbeiten (A24). Außerdem kann es kurze Texte nach Vorgaben selbstständig generieren (A27). Durch die Chatfunktionalität ist es möglich, das KI-Modell als Chatbot in der Kommunikation mit Nutzerinnen und Nutzern einzusetzen (A31). Als letzten Use Case deckt es ebenfalls die Gefühlsanalyse (A32) ab. Damit kann eventuelle Voreingenommenheit identifiziert werden.

In der vorangestellten Projektarbeit der Autorin wurden die LLMs ChatGPT und Luminous von Aleph Alpha gegenübergestellt. Es wurde der Output auf festgelegte Fragen bewertet. Dabei wurden die Fragen in Kategorien eingeordnet und die Kategorien wiederum als kritisch für journalistische Anwendungen klassifiziert. Diese sind Voreingenommenheit und Diskriminierung, Ideenentwicklung, Falsche Fakten, Psychologisches Argumentieren, Leseverstehen, Semantik, Gefühlsanalyse (Erkennung von Hatespeech), Syntax, Rechtschreibung, Grammatik und Textproduktion. Das Ergebnis ist, dass ChatGPT in diesem stichprobenartigen Vergleich besser abschneidet als Luminous. Der Anteil der richtigen bzw. zufriedenstellenden Antworten an allen gegebenen Antworten von ChatGPT liegt bei über 75 Prozent. Luminous erreicht dagegen nur einen Anteil von ungefähr 40 Prozent. Dazu kommt, dass ChatGPT ausführliche Antworten mit zusätzlichen Informationen liefert [79, S. 13].

Erfüllung von rechtlichen und ethischen Anforderungen

OpenAI warnt, dass ChatGPT zuweilen falsche, diskriminierende und schädliche Antworten gibt. Durch textuelle Antworten, die der Ausdrucksweise eines Menschen sehr nahekommen, kann es falsche Informationen überzeugend darbieten. Dazu kommt, dass es nur begrenztes Wissen über Ereignisse nach 2021 hat, da es zu diesem Zeitpunkt trainiert wurde [13]. Da das KI-Modell keine Verbindung zum Internet hat, bleibt es ohne menschliche Einwirkung zunächst auf diesem Stand. In den Nutzungsbedingungen wird daher dringend darauf hingewiesen, dass jeder Output von ChatGPT von einem Menschen überprüft werden soll [80].

OpenAI weist ausdrücklich darauf hin, dass die Nutzerin bzw. der Nutzer, wenn sie oder er ChatGPT personenbezogene Daten verarbeiten lässt, die Einwilligung der betroffenen Personen benötigt. Die Beweislast dafür liegt bei der Nutzerin bzw. beim Nutzer [80]. Dagegen überträgt OpenAI der Nutzerin bzw. dem Nutzer die Rechte an sämtlichem Input und dem daraus generierten Output des KI-Modells. Der Output darf gewerblich genutzt und veröffentlicht werden (A11) [80]. In- und

Output, der über eine API übertragen wird, nutzt OpenAI nicht, um sein Modell weiterzuentwickeln. Content, der beispielsweise über die Teststellung im Browser übergeben wird, wird jedoch zur Verbesserung von ChatGPT genutzt. ChatGPT stellt allerdings die Vorgabe, dass mit seiner Funktionalität keinerlei Rechte einer natürlichen Person verletzt oder unterschlagen werden dürfen [80]. Die Verantwortung für die Erfüllung der rechtlichen und ethischen Anforderungen liegt also beim Medienunternehmen.

6.3 SOLL-WORKFLOWS

Die Ist-Workflows aus Kapitel vier erhalten unter Beachtung der Anforderungen aus Kapitel fünf einen neuen Akteur, die KI, im Besonderen das LLM ChatGPT. Es wird in der visuellen Darstellung der Workflows in Anhang zwei mit fett gedruckten Prozessschritten dargestellt. Die generellen Arbeitsabläufe ändern sich durch den Einsatz von KI nicht (A40). Die Start- und Endpunkte der Prozesse sind immer noch dieselben. ChatGPT wird auf zwei unterschiedliche Weisen in den Workflow integriert: Es wird zwischen Aufgaben unterschieden, die manuell von einem Menschen angestoßen werden und Aufgaben, die automatisiert innerhalb eines Workflows ablaufen. In beiden Fällen wird der Output der KI von einem Menschen kontrolliert. Die vollständigen Flussdiagramme der Workflows befinden sich in Anhang zwei. Eine Analyse der neuen Workflows befindet sich in Anhang drei.

6.3.1 INGEST

Beim Ingest von neuem Material ist das Ziel, dass in möglichst kurzer Zeit möglichst präzise Informationen zu einer Essenz erzeugt werden und diese mit derselben verbunden werden. Der Essenz wird eine eindeutige Identifikationsnummer zugeordnet und die technischen Parameter in den Metadaten abgespeichert. Im Folgenden wird nun die Extraktion von beschreibenden Metadaten nun nicht mehr von einem Menschen übernommen, sondern von der KI. Diese erkennt Schlagwörter und wichtige Informationen aus einem Text. Liegen Informationen in einer anderen Sprache vor, werden diese ebenfalls in die Sprache des Medienunternehmens übersetzt. Der Text in der Originalsprache und die Übersetzung liegen aufeinander referenziert ab. Stichprobenartig wird das Ergebnis von der Archivarin bzw. dem Archivar kontrolliert, um keine falschen Daten im Archiv zu speichern. Sind die Metadaten gesammelt, werden sie der Essenz zugeordnet, sodass sie ab sofort gemeinsam im Archiv gefunden werden können.

6.3.2 NACHRICHTENSENDUNG

Eine Nachrichtensendung fasst assistierende, generative und distributive Aufgaben zusammen, um an Ende ein ausspielfertiges Produkt zu erhalten. Die Use Cases für KI reichen von der Recherche über die Generation von Texten bis hin zur Verschlagwortung bei der Speicherung im Archiv.

Ideengeber

Fragt man ChatGPT nach einem Thema oder einem möglichen Ansatzpunkt für einen Beitrag, antwortet das KI-Modell mit plausiblen Vorschlägen. Diese bilden einen Ausgangspunkt für eine Redaktionskonferenz. Die Journalistin bzw. der Journalist lässt sich durch die Ideen der KI inspirieren und bringt Vorschläge in die Redaktionskonferenz ein. ChatGPT hat hier eine unterstützende Funktion. Die gesammelten Ideen inklusive Stichpunkte werden direkt in einen Arbeitsordner zur Erstellung eines Beitrags überführt und von dort aus weiterbearbeitet. Der Vorgang wird manuell von einer Anwenderin bzw. einem Anwender über die Benutzungsoberfläche ausgelöst werden, wenn die Person beispielsweise nach konkreten Ansatzpunkten zu einem gegebenen Thema sucht.

Unterstützung bei der Recherche

Durch den starken Detailgrad der Verschlagwortung und die Vernetzung der Assets mit einem Knowledge Graph wird die Archivsuche effizienter, da nun beispielsweise Synonyme bei der Stickwortsuche ignoriert werden können. Hier wird die Technik der Embeddings angewendet. Das bedeutet, dass jedes Asset in einem multidimensionalen Vektorraum verortet wird und mit einem Vektor beschrieben wird. Jedes Element des Vektors deckt eine messbare Eigenschaft des Assets ab. Die Definition und Bewertung dieser Eigenschaften obliegt dem KI-Modell. Außerdem werden zu einem Asset weitere vorgeschlagen, die zum Thema passen, ähnlich wie für die Nutzer in der Mediathek oder in Sozialen Medien.

Unterstützung bei der Textproduktion

ChatGPT ist als LLM in der Lage, Texte zusammenzufassen und damit einen Überblick über das Thema zu geben, ohne dass die Journalistin bzw. der Journalist viel Zeit für das Lesen aufwenden muss. Dies ist bei eintreffendem Text-Material von Bedeutung sowie bei der Erstellung eines Kurzüberblicks über eine bestehende Sendung. ChatGPT liefert eine Zusammenfassung der transkribierten Sendung mit den wichtigsten Meldungen des Tages. Trotzdem ist auch hier die Aufsicht eines Menschen entscheidend.

Verfasst die Journalistin bzw. der Journalist einen Beitrag in Textform übernimmt die KI die Kontrolle auf sprachliche Fehler und bringt Vorschläge zum Stil ein. Der Anknüpfungspunkt befindet sich also hier in der Vorbereitung und Produktion des Beitrags. Wichtig ist hier jedoch, die menschliche Kontrolle am Ende des Arbeitsschrittes der KI. Eine weitere Funktion von ChatGPT ist die Gefühlsanalyse (Sentiment Analysis). Diese kommt ebenfalls nach Verfassen eines Beitragstextes zum Einsatz. Sie prüft auf etwaige Subjektivität innerhalb des Textes. Erkennt die Analyse ein Gefühl, das über die Beschreibung „neutral“ oder „objektiv“ hinausgeht, ist die Wahrscheinlichkeit hoch, dass der Text nicht die ausreichende Objektivität besitzt und dieser bearbeitet werden muss.

Der Prozess wird manuell ausgelöst, wenn eine Journalistin bzw. ein Journalist beispielsweise einen gerade verfassten Text prüfen möchte. Er wird ebenfalls bei Archivmaterial automatisiert durchgeführt. Assets, die eine Sentiment Analysis erhalten, die nicht neutral ist, werden als solche markiert, sodass die Journalistin bzw. der Journalist diese Tatsache berücksichtigt. Die Journalistin bzw. der Journalist übergibt der KI die Referenz auf einen Text oder Textausschnitt im Speicher. Die Automation greift auf das Asset im Speicher zu und leitet es an ChatGPT weiter. Das Modell gibt dann die Antwort zurück. Bei bereits angefragten Texten wird das Ergebnis der sprachlichen Analyse in den Metadaten vermerkt und bei Ansicht des Textes in einer GUI am Rand an der entsprechenden Stelle im Text dargestellt. So müssen weitere Bearbeiterinnen und Bearbeiter des Textes keine neue Anfrage starten.

6.3.3 BREAKING NEWS

Breaking News sind eine Situation, bei der es besonders auf die Geschwindigkeit der Beitragsproduktion ankommt. Daher kann hier die KI Aufgaben zusätzlich oder sogar parallel zu den menschlichen Journalistinnen und Journalisten übernehmen. In einer solchen Situation kommt fremd produziertes Material eines anderen Medienunternehmens oder einer Presseagentur an und wird vom Ingest-Workflow bearbeitet.

Bei der Generation von Moderationstexten spart ChatGPT den Journalistinnen und Journalisten Zeit. Das Modell erzeugt eigenständig nach Vorgaben mögliche Texte zur Vermittlung von Nachrichten. Die Schlagwörter zu einer Meldung nimmt ChatGPT aus den vorher extrahierten Schlagwörtern, des fremd produzierten Materials. Die Moderation ist für das Fernsehen, den Hörfunk sowie den Onlinejournalismus geeignet. Beim selbstständigen Verfassen von Texten ist es jedoch besonders wichtig, diese nochmals von einem Menschen überprüfen zu lassen. Die Texte werden gespeichert und beispielsweise automatisch an den Teleprompter weitergegeben. In der umgekehrten Form überprüft auch die KI die Texte, die ein Mensch geschrieben hat, auf sprachliche Fehler. Auch hier ist in kleinerem Umfang eine Gefühlsanalyse von Vorteil. Mögliche bevorzugende oder diskriminierende Aussagen werden hier schnell erkannt.

6.3.4 INFORMATIONSNACHRICHTEN

Diese Art der Nachrichten benötigt keine stilistisch hochwertige Sprache. Im Vordergrund steht die Vermittlung von Informationen in kurzer Zeit. Dadurch, dass diese Nachrichtenform regelmäßig nach demselben Schema ausgespielt wird, wird das Verfassen von Nachrichtentexten hier von einer KI übernommen. ChatGPT benötigt dafür nur wenige Vorgaben und generiert einen Text, der auch von einem Menschen geschrieben sein könnte. Ähnlich wie im Breaking News-Fall nimmt sich ChatGPT hier die Informationen automatisch aus dem Speicher und verarbeitet sie zu Text. Das KI-Modell erkennt ebenfalls die Art der Daten und fügt diese in Grafiken und Texte ein. Bei der selbstständigen Generation von Sprache muss jedoch eine Kennzeichnung, dass der Beitrag automatisch erzeugt wurde, bei der Ausspielung erfolgen, welche von der Automation erstellt wird. Bei einer solchen generativen Aufgabe ist jedoch die nachträgliche Kontrolle durch einen Menschen nötig, damit die Texte auch sprachlich richtig und logisch sind.

6.3.5 RÜCKMELDUNG UND ANFRAGEN VON NUTZERINNEN UND NUTZERN

Bei der Kommunikation mit Nutzerinnen und Nutzern nimmt ChatGPT den Journalistinnen und Journalisten Arbeit ab. Trifft eine Anfrage einer Nutzerin oder eines Nutzers in Textform ein, entscheidet ChatGPT, ob es sich um eine einfache Anfrage oder um eine weitreichendere Anfrage handelt. Erfordert eine Anfrage nur eine Antwort ist das ein Fall für das KI-Modell. Zieht die Anfrage jedoch weitere innere Abläufe im Medienunternehmen nach sich, muss diese ein Mensch beantworten. Sie kann beispielsweise eine schwerwiegende Kritik an einem Beitrag sein. Es handelt sich also zunächst um einen automatischen Prozess. Wird jedoch die Entscheidung eines Menschen nötig, wird der Prozess angehalten und von Menschen wieder manuell weitergeführt. Bevor jedoch

eine Nachricht, die eine Nutzerin oder ein Nutzer verfasst hat, einen Menschen im Medienunternehmen erreicht, wird diese mit der Gefühlsanalyse überprüft, um grob beleidigende und diskriminierende Nachrichten zu kennzeichnen. Je nach Vorgabe des Medienunternehmens werden diese Nachrichten entweder ignoriert oder versuchsweise beantwortet.

Handelt es sich bei der Nachricht der Nutzerinnen und Nutzer um eine Beschwerde zu Inhalten oder der Präsentation derselben, muss diese an eine menschliche Beschwerdestelle weitergeleitet und dort geprüft werden. Bei einzelnen Bemerkungen muss geklärt werden, ob es sich um einen Verstoß gegen Richtlinien oder Gesetze handelt. Häufen sich die Beschwerden zum selben Thema, werden die Redaktion sowie die Leitung des Medienunternehmens miteinbezogen. Kommentare von Nutzerinnen und Nutzern, die man auf der Webseite von Onlinemagazinen hinterlassen kann, werden auf dieselbe Weise behandelt. Hier ist es besonders wichtig, die unpassenden Beiträge herauszufiltern, denn die Kommentare werden in einem öffentlichen Forum angezeigt, in dem andere Nutzerinnen und Nutzer miteinander interagieren. Unhöfliche Kommentare werden hier gelöscht. Diese Funktion der KI wird stichprobenartig von einem Menschen kontrolliert.

6.4 ROLLENKONZEPT

Die Einführung von KI in den Newsworkflow sorgt für die Veränderung, Abschaffung und Neugestaltung von Rollen. Einige Rollen verlangen weniger Personal, einige werden zusammengelegt oder aufgesplittet. Sie haben entweder nur für den Einführungsprozess Gültigkeit oder verbleiben dauerhaft in Medienunternehmen (A43). Im Folgenden sind Rollen beschrieben, bei denen eine Umstrukturierung erfolgt [74, S. 113]. Das vollständige Rollenkonzept befindet sich in Anhang vier.

*Journalist*in*

Journalistinnen und Journalisten arbeiten nach der Einbindung von ChatGPT mit den automatischen und manuellen Prozessen, die durch ChatGPT geschaffen wurden. Die automatischen Prozesse präsentieren den Journalistinnen und Journalisten zusätzliche Informationen oder bereits vorgenommene Bearbeitungsschritte. Die manuellen Prozesse werden bei Bedarf und auf Wunsch der Journalistin bzw. des Journalisten ausgelöst.

*Archivar*in*

Für die Mitarbeiterinnen und Mitarbeiter im Archiv ändert sich durch den Einsatz von KI der Arbeitsalltag. Die KI übernimmt während des Ingest-Workflows die Annotation der Essenzen mit beschreibenden Metadaten, welche vorher die Archivarin bzw. der Archivar noch manuell durchführen musste. Diese Art der Annotation spart Zeit und gibt ihr oder ihm Raum für systematisierende Arbeiten. Durch die Einführung von KI in seinen Arbeitsalltag, steht die Archivarin bzw. der Archivar in engem Austausch mit der Data-Scientistin bzw. dem Data-Scientisten. Gemeinsam entwickeln sie die optimale technische Lösung für die Verschlagwortung von Essenzen. Dabei untersuchen sie Metadatenmodelle auf ihre Tauglichkeit und führen Tests für die Präzision der Verschlagwortung durch. Da auch bei der Digitalisierung des Legacy-Materials so gesehen neues digitales Material ins Archiv eingepflegt wird, muss dieses ebenfalls mit Metadaten versehen werden.

*Data-Scientist*in*

Die Data-Scientistin bzw. der Data-Scientist beschreibt eine Rolle, die mit der Einführung von KI neu im Medienunternehmen sein wird. Die Aufgabe wird die technische Einführung von ChatGPT im Medienunternehmen sein. Dazu gehört Qualitätsmanagement und die Weitergabe von Unzulänglichkeiten an OpenAI. Die Data-Scientistin bzw. der Data-Scientist überprüft in regelmäßigen Abständen, ob ChatGPT sich noch auf dem Stand befindet, der im Medienunternehmen für aktuell befunden wird. Außerdem werden die Anregungen der Anwenderinnen und Anwender im Medienunternehmen berücksichtigt. Die Kompetenzen der Data-Scientistin bzw. des Data-Scientist liegen also im Bereich des Aufbaus und Trainings von KI-Modellen. Die Rolle ist Teil des Projektteams.

KI-Assistenz

Die KI-Assistenz ist eine Unterstützungsrolle für die Data-Scientistin bzw. den Data-Scientisten. Die Person übernimmt beispielsweise die Vorbereitung der Teststellung für die regelmäßigen Tests. Auch hier sind Kenntnisse von KI-Modellen, im Besonderen von ChatGPT notwendig. Die Rolle ist ebenfalls Teil des Projektteams.

*Compliance Manager*in*

Die Compliance Managerin bzw. der Compliance Manager ist ebenfalls eine neue Rolle, die durch die Einführung von ChatGPT geschaffen wird. Die Person sorgt in Zusammenarbeit mit der Data-Scientistin bzw. dem Data-Scientisten dafür, dass die ethisch-rechtlichen Anforderungen wie Datenschutz, Unvoreingenommenheit und Transparenz eingehalten werden. Dafür informiert sich die Person in der Rolle über die aktuellen Standards und führt regelmäßige Prüfungen durch. Die Rolle ist die menschliche Vertreterin bzw. der menschliche Vertreter von ChatGPT in der Redaktion und führt deren Vorgaben aus. Dafür ist ein datenwissenschaftlicher Hintergrund mit Kenntnissen zum Journalismus sowie den rechtlichen-ethischen Anforderungen an ein KI-System nötig.

Es ist ebenfalls Aufgabe der Compliance Managerin oder des Compliance Managers, die Brücke zwischen der Informationssicherheit des Medienunternehmens und des KI-Projektteams zu bilden. Die Rolle arbeitet eng mit dem CISO zusammen. Diese entwickelt gemeinsam mit OpenAI eine Policy für Sicherheitsnotfälle [44, S. 35]. Die Rolle ist auch Teil des Projektteams.

*IT-Administrator*in – KI / Systemintegrator*in*

Die IT- Administratorin bzw. der IT-Administrator, die oder der ausschließlich für die Integration von ChatGPT in die Systemlandschaft des Medienunternehmens zuständig ist, ist eine neue Rolle im Medienunternehmen. Die Ressourcen für diese Rolle stammen entweder aus der IT-Abteilung des Medienunternehmens oder werden neu beschafft. Externe Ressourcen sind in diesem Kontext auch möglich. Ihre Aufgabe ist die Schaffung von Verbindungen zwischen CMS, MAM, KI und Nutzer*in sowie die Automation von Workflows. Die Rolle arbeitet eng mit den Data-Scientistinnen und Data-Scientisten zusammen. Die Anforderungen für ihre Arbeit stammen von den Anwender*innen, den Journalist*innen, Planer*innen und Mediengestalter*innen. Sie arbeiten in einem iterativen Prozess die anwendungsfreundlichste Systemintegration heraus. Wichtige Voraussetzungen sind die genaue Kenntnis des Systemaufbaus des Medienunternehmens sowie Kenntnisse im IT-Bereich. Die Rolle ist Teil des Projektteams.

*UI (User Interface) Designer*in*

Die UI Designerin bzw. der UI Designer ist für den Entwurf der grafischen Benutzungsoberfläche zuständig. Die Rolle steht in Austausch mit den zukünftigen Anwenderinnen und Anwendern von ChatGPT und erfasst die Anforderungen an die GUI, um dem User-Centred-Approach gerecht zu werden. Über die GUI werden manuell Aufträge an das KI-Modell abgeschickt und die Ergebnisse betrachtet. Die UI Designerin bzw. der UI Designer legt ebenfalls gemeinsam mit der IT-Administratorin bzw. dem IT-Administrator fest, wie die Benutzungsoberfläche in das

Redaktionssystem und das CMS integriert wird. Hier sind also gestalterische Fähigkeiten und praktisches Denken erforderlich. Kenntnisse im IT-Bereich sind ebenfalls hilfreich. Die Rolle ist ebenfalls Teil des Projektteams.

*Change Manager*in / Product Owner*

Die Change Managerin bzw. der Change Manager ist der Projektleiter während der Einführung von ChatGPT. Bei ihm laufen alle Fäden des Veränderungsprojekts zusammen. Die Rolle überwacht die Qualität der Integration und setzt Ziele auf den Product Backlog. Sie berichtet an die Leitung des Medienunternehmens.

Projektteam

Das Projektteam besteht aus Data-Scientist*in, KI-Assistenz, Compliance Manager*in, IT-Administrator*in und UI Designer*in. Die Rollen können in einer oder mehreren Personen vorhanden sein, je nach Ressourcenbedarf. Das Team arbeitet in zeitlich begrenzten Sprints, die es selbst aus dem Product Backlog zusammenstellt.

Scrum Master

Der Scrum Master ist der Coach des interdisziplinären Teams. Die Rolle kümmert sich um die effektive Nutzung von Scrum, moderiert die Teambesprechungen und sorgt für die Einhaltung des Zeitplans. Sie stellt sicher, dass das Team gewinnbringend arbeiten kann.

6.5 COMPLIANCEKONZEPT

Um die Eigenschaften von ChatGPT mit den definierten rechtlichen und ethischen Anforderungen aus Kapitel fünf zu verbinden, bedarf es eines Konzepts zur Einhaltung dieser Vorgaben. Es werden Empfehlungen für Regeln gegeben, damit keine Vorgabe verletzt wird und keine rechtlichen Probleme daraus entstehen.⁷

Datenschutz

Der Datenschutz liegt zunächst in der Verantwortung der bzw. des Datenschutzbeauftragten. Mit der Einführung von ChatGPT ist ebenfalls die Compliance Managerin bzw. der Compliance Manager für den Datenschutz innerhalb des KI-Modells in der Verantwortung (A6 - A9). Die Verarbeitung von personenbezogenen Daten ist für journalistische Zwecke nach den Medienprivileg im Medienstaatsvertrag erlaubt [30, § 23, 30, § 12]. Daher fällt das Training eines KI-Modells, das den Newsworkflow unterstützt, unter diesen Zweck. Ob das Medienunternehmen nun die Einwilligung der betroffenen Personen besitzt, ist im journalistischen Kontext durch das Medienprivileg nicht von Bedeutung, daher können aus Sicht des Datenschutzes personenbezogene Daten zum Training verwendet werden. Die personenbezogenen Daten dürfen jedoch nicht außerhalb dieses Rahmens verwendet werden.

Urheberrecht

Das Medienunternehmen darf kein Material zum Training verwenden, an dem sich der Urheber die Nutzung zum Training von KI-Systemen vorbehalten hat. Dieses Material wird identifiziert und im MAM gekennzeichnet. Sofern Material rechtmäßig zugänglich ist, darf es jedoch zum Training verwendet werden. Das Medienunternehmen besitzt außerdem die Rechte am Output von ChatGPT und darf diesen gewerblich nutzen. Dies wird ebenfalls im MAM vermerkt (A10).

Informationssicherheit

Die Verantwortung für die Informationssicherheit liegt beim CISO und bei der Compliance Managerin bzw. beim Compliance Manager für ChatGPT. Die Informationssicherheit von KI-Modellen ist ein neues Themengebiet, das Best Practices aus der IT-Sicherheit und Eigenheiten von ChatGPT verbindet.

Ein Einstiegspunkt für die Informationssicherheit von KI-Modellen ist die Limitierung des Inputs. Ein textueller Input wird in Token zerlegt. Durch eine Obergrenze an Token pro Eingabe können große Eingaben verhindert werden. Durch bestimmte Inputs können Manipulationen ausgelöst werden, gegen die das Modell geschützt werden muss. Die Erstellung des richtigen und gewinnbringenden Inputs nennt sich Prompt Engineering. Zunächst bedeutet dies, dass ChatGPT eine Aufgabe so verständlich gestellt wird, dass es die gewünschte Antwort gibt. Im schadhaften Fall heißt es jedoch, dass Inputs auf jene Weise gestaltet werden, dass sie Gewichte verfälschen oder

⁷ Hierbei handelt es sich um Empfehlungen basierend auf der Recherche der Autorin. Diese hat keinerlei Ausbildung im juristischen Bereich, daher sind die Inhalte dieses Kapitels vor der Umsetzung zunächst mit geschulten Personen abzustimmen.

Trainingsdaten offenlegen. Ein solches Prompt Engineering bedarf einer Vielzahl an Eingaben, weshalb diese meist automatisiert oder sogar mit einem weiteren KI-Modell durchgeführt wird.

Zusätzlich muss die Authentizität der Eingaben geprüft werden. Durch mengenmäßige Limitierung von Inputs pro Zeiteinheit werden Manipulationen behindert. Eine Detektion maschinengeschriebener Texte kann ebenfalls Angriffe ausschließen [40, S. 13-15]. Jedoch wird auch durch den Einsatz von kryptografischen Verfahren die Authentizität der Eingaben – manuell oder automatisch – geprüft. Damit kann dann zweifelsfrei festgestellt werden, ob der Input aus einer vertrauenswürdigen Quelle stammt.

Für mehr Sicherheit sorgt auch die Limitierung des Outputs. Dies kann ebenfalls durch die Begrenzung der Outputtoken geschehen. Neben einer solchen quantitativen Beschränkung kann jedoch auch durch inhaltliche Beschränkungen Missbrauch verhindert werden. Jede Nutzerin bzw. jeder Nutzer kann durch rollenbasierte Zugänge nur diejenigen Aufgaben an das KI-Modell übergeben, die für die Arbeit der Person relevant sind [40, S. 16].

Neben technischen Vorkehrungen wird auch unter den Anwenderinnen und Anwendern eine Aufmerksamkeit für Informationssicherheit erzeugt. In Schulungen werden mögliche Angriffsszenarien über Social Engineering sowie deren Gegenmaßnahmen besprochen. Innerhalb solcher Schulungen werden auch die Grenzen der Funktionalität von ChatGPT besprochen, um ein Gefühl für mögliche Angriffe zu vermitteln [81]. Das Verhalten im Notfall wird ebenfalls gelehrt und geprobt.

Für zusätzlichen Schutz sorgen regelmäßige Gefahreneinschätzungen. Es werden vierteljährliche Überprüfungen empfohlen. Die Durchführbarkeit und Wahrscheinlichkeit von Missbrauchsszenarien wird beurteilt und daraus zukünftige Maßnahmen abgeleitet. Diese Audits werden außerdem bei einer hardware- oder softwareseitigen Veränderung durchgeführt. Dabei ist zwischen Sicherheit und Benutzbarkeit der Systeme abzuwägen. Eventuell bleibt ein akzeptables Restrisiko [44, S. 28]. Neben den KI-spezifischen Sicherheitsmaßnahmen muss das gesamte System ebenfalls mit den branchenüblichen Best Practices für IT-basierte Systeme gesichert werden.

Erklärbarkeit

Es arbeiten Personen innerhalb und durch die Nutzerrückmeldung auch außerhalb des Medienunternehmens mit ChatGPT. Daher muss für beide Personengruppen eine Dokumentation zur Beschaffenheit und Funktionsweise des KI-Modells vorgehalten und zugänglich gemacht werden. Die Erklärungen zu ChatGPT werden mithilfe von Probandinnen und Probanden evaluiert und verbessert. Das Verfahren von Doshi-Valez und Been, wie in Abschnitt 5.1.4 beschrieben, wird im Rahmen eines Testkonzepts angewandt. Die Testpersonen sind also zum einen Mitarbeiterinnen und Mitarbeiter des Medienunternehmens mit technischem Hintergrund sowie Personen ohne diesen (A16).

Fairness

Da die Voreingenommenheit eines KI-Modells oft mit der Unausgewogenheit seiner Trainingsdaten zusammenhängt, wird vor Beginn des Trainings zunächst die Fairness von ChatGPT bewertet. Daraus werden Schlüsse gezogen, ob OpenAI die Fairness seines KI-Modells global verbessern muss. Die Tests sind an den Umfang und die Art der Anwendung anzugleichen. Es empfiehlt sich, eine normierte Art der Prüfung zu verwenden oder zu entwerfen, um vergleichbare Ergebnisse zu erzielen. Da sich durch starken Gebrauch die Gewichte von ChatGPT ändern können, muss regelmäßig überprüft werden, ob die Vorgaben zur Voreingenommenheit noch eingehalten werden. Dabei können die Quellen- und Themenauswahl von geplanten und veröffentlichten Beiträgen statistisch erfasst werden (A1 und A13). Außerdem wird der Output Menschen vorgelegt, um daraus ebenfalls eine Bewertung ziehen zu können und Fehler zu erkennen, die nicht durch eine Statistik gefunden werden können [40, S. 18].

Transparenz

Ein KI-System wird der Nutzerin bzw. dem Nutzer mit einer Gebrauchsanweisung präsentiert. Diese wird digital und einfach zugänglich bereitgestellt. Sie ist in einfacher und verständlicher Sprache verfasst und muss vollständig, korrekt und eindeutig Informationen vermitteln. Die Veröffentlichung erfolgt beispielsweise in der Mediathek, aufgrund der dortigen Personalisierungsalgorithmen.

Die Informationen umfassen die Kontaktdaten des Anbieters sowie Details zum KI-System. Es werden Merkmale, Fähigkeiten und Leistungsgrenzen beschrieben. Zu den Details gehören auch Zweck, Sicherheitsbestimmungen sowie alle bekannten Möglichkeiten zur Fehlanwendung des KI-Modells. Es wird über die Fehlanwendungen aufgeklärt, da diese zu Sicherheitsrisiken und Grundrechtsverletzungen führen können.

Außerdem wird über Herkunft, Qualität, Diversität der Trainingsdaten sowie Methode und Arbeitsbedingungen der Annotation aufgeklärt. Teil der Dokumentation ist auch die Funktionsweise des Quellcodes und das Bilden von Kategorien und Metriken innerhalb von ChatGPT. Vor Betrieb des KI-Modells im Medienunternehmen müssen diese Kriterien geklärt und bewertet werden, ob sie den Qualitätsanforderungen des Medienunternehmens entsprechen (A14). Es könnte dem Medienunternehmen auf die Füße fallen, dass eventuell Trainingsdaten unter menschenunwürdigen Bedingungen annotiert wurden. Sprechen keine Aspekte gegen das KI-Modell, werden die Informationen in verständlicher Sprache der Nutzerin bzw. dem Nutzer zur Verfügung gestellt. Die Dokumentation enthält zusätzlich die erwartete Lebensdauer und erforderlichen Wartungsmaßnahmen, die in seinem Lebenszyklus auf den Betreiber zukommen [41, Art. 13].

Menschliche Aufsicht

Eine weitere Vorgabe des Gesetzes über Künstliche Intelligenz der EU ist die Versicherung der menschlichen Aufsicht. Ein KI-Modell darf also nicht völlig selbstständig betrieben werden. Es bedarf eines Human in the loop. Dafür muss eine Schnittstelle zwischen Mensch und Maschine geschaffen werden. Die menschliche Aufsicht soll Risiken minimieren oder verhindern [41, Art. 14].

Die menschliche Aufsicht wird zum einen über die regelmäßigen Audits der Data-Scientistin bzw. des Data-Scientisten realisiert. Zum anderen durch stichprobenartige Aufsicht der produktiven Anwenderinnen und Anwender. Gibt beispielsweise eine Journalistin bzw. ein Journalist die Transkription eines Interviews in Auftrag, prüft diese das Transkript vor der Weiterbearbeitung kurz auf Richtigkeit und Vollständigkeit. Auch die Archivarin bzw. der Archivar übersieht als Stichprobe die erstellten Metadaten im Ingest-Workflow. Außerdem werden Beiträge vor der Veröffentlichung durch die Redaktion abgenommen.

6.6 ZUGANGSKONZEPT

Zur Erhöhung der Informationssicherheit haben nur wenige Nutzerinnen und Nutzer den vollen Zugriff auf alle Funktionen von ChatGPT. Daher wird ein rollenbasierter Zugang mit Authentifizierung eingeführt. Jede Anwenderin bzw. jeder Anwender muss sich vor der Nutzung des KI-Modells authentifizieren und erhält danach Zugang zu einer GUI, die auf die jeweilige Tätigkeit abgestimmt ist. Durch diese Sicherheitsmaßnahmen wird ausgeschlossen, dass unbefugte Personen ChatGPT nutzen, Informationen auslesen oder Änderungen am Modell vornehmen (A36). In Anhang vier sind die Rollen aus Abschnitt 6.4 mit den Zugangsprofilen verbunden. Es folgen die Zugangsprofile für die jeweiligen Aufgabengebiete:

A: Text

Nutzerinnen und Nutzer, die dem Profil Text zugeordnet werden, arbeiten über die GUI mit ChatGPT. Sie können also manuell Prompts eingeben und erhalten einen Output. Außerdem können sie die Ergebnisse von automatischen Anfragen an das KI-Modell einsehen. Ein Recherche-Zugang zum Archiv ist ebenfalls im Profil Text enthalten. Die Rollen, die mit diesem Profil arbeiten, sind Journalist*innen, Redakteur*innen, Planer*innen und Co-Planer*innen.

B: Archiv I

Dieses Profil hat Zugang zur Verschlagwortung von Texten beim Ingest. Die Rolle arbeitet ebenfalls mit einer GUI. Sie überprüft stichprobenartig die Ergebnisse von ChatGPT und kann, wenn nötig, diese händisch korrigieren. Archivarinnen und Archivare haben Zugang zu diesem Profil.

C: Test

Das Test-Profil wird verwendet, um Voreingenommenheit und generelle Qualität des Outputs zu messen. Data-Scientist*innen und KI-Assistenz führen die Tests durch und präsentieren die Ergebnisse dem Product Owner und der Compliance Manager*in.

D: Admin

Das Admin-Profil hat keinen Zugang zu den Funktionalitäten von ChatGPT. Die Systemintegratorin bzw. der Systemintegrator nutzt dieses Profil, um das KI-Modell in die Systemlandschaft des Medienunternehmens einzubinden. Es werden Schnittstellen zwischen MAM, CMS und ChatGPT definiert. Dabei müssen die Ein- und Ausgangsdatenströme sowie das Format, in dem Informationen übergeben werden, beachtet werden.

E: GUI

Das GUI-Profil wird von der UI-Designerin bzw. vom UI-Designer verwendet. Mit dem Profil kann das Aussehen der Benutzungsoberfläche beschrieben und angepasst werden. Dabei handelt es sich um Entscheidungen der Benutzbarkeit.

6.7 SCHULUNGSKONZEPT

Für die Schulung der Mitarbeiterinnen und Mitarbeiter wird ein modulares Schulungskonzept verwendet. Die Belegschaft wird in mehrere Zielgruppen unterteilt. Diesen wird Wissen über die neue Technologie in kleinen Abschnitten dargeboten. Jede Zielgruppe erhält eine unterschiedliche Zusammenstellung an Modulen. Es können jedoch darüber hinaus auch weitere Module belegt werden. Die Module können jeweils in zwei Abstufungen gelernt werden: Grundlegendes oder tiefergehendes Verständnis. Die Darstellungsformen der Module reichen von Präsenzveranstaltungen über Onlineveranstaltungen bis zu Material zum selbstständigen Lernen, je nach der Vorliebe der Lernerin bzw. des Lerner (A43) [78]. Die Schulungsmatrix, die den Zielgruppen die Module zuordnet, befindet sich in Anhang fünf.

Zielgruppen

Die Zielgruppen werden nach den Berufsgruppen gebildet, da diese über ähnliche Kenntnisstände verfügen.

Zielgruppe	Berufsgruppe(n)
Nicht-Technisch	Journalist*in, Redaktion, Außenreporter*in, CvD (*)
Produktiv	Producer*in, Planer*in, Co-Planer*in
Gestalterisch	Grafiker*in, Mediengestalter*in, Regisseur*in
Archiv	Archivar*in

Abbildung 17: Zielgruppen des Schulungskonzepts [Eigene Darstellung].

Module

Die Module behandeln einzelne Aspekte des neuen Akteurs in den täglichen Prozessen. Hier wird besonderen Wert auf die Vermittlung der Funktionsweise und der Limitierungen von ChatGPT gelegt.

- Generelle Einführung in KI

Dieses Modul ist für alle Mitarbeiterinnen und Mitarbeiter als Einstieg empfohlen. Es soll die Hemmschwelle vor der Arbeit mit KI reduzieren.

- Funktionsweise der Algorithmen

Hier soll eine kurze Einführung in die verwendeten Algorithmen von ChatGPT gegeben werden. Das Modul soll ebenfalls einen Überblick über die Größe und Rechenkapazität vermitteln.

- Chancen und Limitierungen von ChatGPT

Für eine effiziente Arbeit mit ChatGPT ist es wichtig, die Fähigkeiten aber auch die Limitierungen desselben zu kennen. Ist nicht bekannt, was ChatGPT nicht kann, sind die Arbeitsergebnisse enttäuschend und die Frustration steigt.

- Einführung in textuelle Use Cases

Dieses Modul ist besonders für die nicht-technische Zielgruppe interessant. Es beleuchtet die möglichen Use Cases im Bereich der Textgeneration, -annotation und -bearbeitung. Dabei wird auch darauf eingegangen, wie die Prozesse in den jeweiligen Use Cases auszulösen sind und wie Verbesserungen vorgenommen werden können.

- Einführung in Archivworkflow

Die Arbeit im Archiv ist Aufgabe der Archivarin bzw. des Archivars. Daher zielt das Modul besonders auf ihre oder seine Bedürfnisse ab. Die Durchführung und Funktionsweise der automatischen Annotation von Essenzen wird hier beleuchtet.

- Rechtliche und ethische Informationen

Jede Zielgruppe, die dafür verantwortlich ist, dass Ergebnisse von ChatGPT veröffentlicht werden, muss sich über die rechtlichen und ethischen Implikationen im Zusammenhang mit KI im Klaren sein. Das Modul behandelt Aspekte von Datenschutz über Fairness bis hin zu Transparenz. Es gibt Empfehlungen, wie die Compliance mit den Anforderungen am besten zu erreichen ist (siehe Abschnitt 6.5).

- Verhalten bei Überlastung oder Ausfall der KI-Instanz

Zielgruppen, die direkt mit ChatGPT arbeiten und manuell Aufgaben in Auftrag geben, müssen mit einer eventuellen Überlast- oder Ausfallsituation rechnen. Diese hält den Arbeitsfluss auf. Das Modul gibt Informationen, wer in diesem Fall zu benachrichtigen ist und wie die Arbeit weitergeführt wird.

- Sensibilisierung für Informationssicherheit

Neben den infrastrukturellen Maßnahmen gegen einen Sicherheitsnotfall ist auch die menschliche Sensibilisierung ein wichtiger Schritt. Alle Zielgruppen lernen mögliche Szenarien und vorbereitende Gegenmaßnahmen kennen.

- Verhalten im Sicherheitsnotfall

Das richtige Verhalten im Sicherheitsnotfall ist für alle Zielgruppen relevant, um diesen schnellstmöglich zu beenden. Die Teilnehmerinnen und Teilnehmer lernen das Verhalten im Ernstfall sowie mögliche Recovery-Maßnahmen kennen.

6.8 FEEDBACKKONZEPT

Maßgeblich für die erfolgreiche Einführung und den produktiven Einsatz ist der Input der Anwenderinnen und Anwender, um die Funktionen und die Beschaffenheit von ChatGPT weiter zu verbessern. Verbesserungswünsche und zusätzliche Anforderungen können von allen produktiv arbeitenden Rollen an die bzw. den Product Owner gegeben. Sie oder er führt sie dem Projektteam zu.

Senkung der Hemmschwelle

Um möglichst viel Feedback von den Anwenderinnen und Anwendern zu erhalten, muss der Feedback-Prozess möglichst aufwandsfrei ablaufen. Da die Journalistinnen und Journalisten in ihrem Arbeitsalltag bereits unter Zeitdruck stehen, darf das Feedback nur wenig Zeit in Anspruch nehmen. Daher eignet sich die Formatierung in Tickets in einem Ticketmanagementsystem. Dieses gliedert sich ebenfalls in die hauseigene Systemlandschaft ein. Anwenderinnen und Anwender können über Texteingaben ihre Meinung zu ihrer Nutzungserfahrung abgeben. Außerdem werden regelmäßige Termine mit ausgewählten Anwenderinnen, Anwendern und dem Projektteam durchgeführt, damit das Feedback auch persönlich überbracht werden kann.

Kategorien

Das Feedback kann in mehrere Kategorien vorsortiert werden. Diese sind funktional, qualitativ, prozessual, rechtlich-ethische Anliegen sowie eine Kategorie für sonstige Vorkommnisse. Funktionale Anliegen sind beispielsweise zu lange Bearbeitungszeiten. Stimmt die Qualität des Outputs von ChatGPT nicht, wird ein Ticket in der Kategorie qualitativ angelegt. Ein prozessualer Fehler liegt beispielsweise vor, wenn Daten nicht korrekt von einer Stelle zur nächsten übergeben werden. Kommt es zu rechtlich-ethischen Verletzungen, werden sie in dieser Kategorie gemeldet. Die Kategorie Sonstiges nimmt alle übrigen Tickets auf.

Bearbeitung des Feedbacks

Die einkommenden Anliegen der Anwenderinnen und Anwender werden zunächst gelesen und in eine eher technische Sprache überführt, da die Absenderinnen und Absender der Tickets einen sprachlichen oder gestalterischen Hintergrund haben. Eventuell müssen Rückfragen zu Details gestellt werden. Es ist außerdem möglich, Feedback in Zusammenhang mit Log-Ereignissen zu geben. Diese werden ebenfalls zum besseren Verständnis des Sachverhalts ausgelesen. Treten mehrere Nachrichten zu einem Aspekt auf, werden diese gebündelt und quantitativ erfasst. Dann wird die voreingestellte Kategorie überprüft und die Tickets werden priorisiert. Aus den Tickets werden nun Anforderungen und wiederum daraus User Stories formuliert, die in den Product Backlog geschoben werden. Einträge im Product Backlog werden wiederum priorisiert und vom Projektteam abgearbeitet [44, S. 33].

Vulnerability Management

Für die Informationssicherheit ist ein solcher Feedback-Mechanismus ebenfalls wichtig zur Aufdeckung von Fehlerquellen. Das sogenannte Vulnerability Management behandelt die Meldung von Schwachstellen, die Ausgangspunkte für Havariefälle und Missbrauch des KI-Modells sowie des gesamten Veröffentlichungsbetriebs bilden. Es wird im gesamten Medienunternehmen kommuniziert, mögliche Schwachstellen zu melden. Der CISO bewertet das Risiko und arbeitet gemeinsam mit der Data-Scientist*in, der Compliance Manager*in und der Systemintegrator*in an Gegenmaßnahmen.

6.9 AUDITKONZEPT

Um die Sicherheit der eigenen Instanz von ChatGPT zu kontrollieren, sodass keine der vorangegangenen Anforderungen verletzt wird und dies auf das Medienunternehmen abfärbt, soll das KI-Modell regelmäßig getestet werden. Es wird zunächst die Benutzbarkeit und Sicherheit getestet. Die Benutzbarkeit wird von den Anwenderinnen und Anwendern beurteilt. Die Sicherheit wird wiederum mit objektiven Tests festgestellt, die sich als Best Practice in der Branche herausgestellt haben. Automatische Fehlermeldungen werden außerdem überprüft. Durch die hohe Veränderlichkeit der KI sollen die Tests innerhalb von kurzen Abständen erfolgen [44, S. 10]. Dies ist Aufgabe der Data-Scientistin bzw. des Data-Scientisten und der Compliance Managerin bzw. des Compliance Managers. In Reaktion auf die Testergebnisse werden Maßnahmen zu Verbesserung der gefundenen Schwachstellen durchgeführt. Der Testprozess wird immer in Zusammenarbeit mit einem Menschen durchgeführt, da dieser die Testaufgaben anpasst und die Ergebnisse auswertet.

Neben Sicherheit und Benutzbarkeit wird ebenfalls die Qualität des Outputs von ChatGPT getestet. Durch das Feedback der Anwenderinnen und Anwender erhält die Data-Scientistin bzw. der Data-Scientist einen Einblick in die Güte von ChatGPT. Mit regelmäßigen Tests wird gewährleistet, dass sich ChatGPT noch innerhalb der Anforderungen bewegt und keine verfälschten Ergebnisse liefert.

Dokumentationskonzept (A4 und A15)

Der Zugang zur Dokumentation muss jeder berechtigten Person jederzeit möglich sein, daher liegt die Dokumentation in einem geteilten Ordner. Die Dokumentation muss auch von außerhalb des Medienunternehmens und in einem Havariefall erreichbar sein [44, S. 23]. Die Dokumentation wird während der Einführung von ChatGPT in den Newsworkflow angelegt und muss für den produktiven Einsatz fertiggestellt werden. Bei jedem Test oder weiteren Veränderung an ChatGPT muss diese Maßnahme vermerkt und die Dokumentation angepasst werden. Die früheren Versionen werden ebenfalls gespeichert [41, Art. 11]. Ein Vorschlag für eine Gliederung befindet sich auf den Datenträger in Kapitel sechs.

6.10 HAVARIEKONZEPT

Eine Havarie ist ein vorübergehender Zustand, in dem Systeme aus verschiedenen Gründen nicht benutzt werden können. Die Benutzbarkeit hängt zum einen vom Medienunternehmen selbst und zum anderen von OpenAI als Hersteller von ChatGPT ab. Alle Vereinbarungen zwischen OpenAI und dem Medienunternehmen sowie Ansprechpartnerinnen und Ansprechpartnern müssen in einem Service Level Agreement (SLA) festgehalten werden. Darin wird ebenfalls spezifiziert, welche Aktionen im Notfall in Gang gesetzt werden. Der Hersteller verpflichtet sich, innerhalb einer festgesetzten Zeit nach einem Zwischenfall das KI-Modell wieder arbeitsfähig zu machen sowie in einer festgesetzten Zeit unterstützend erreichbar zu sein.

Überlastsituation

Kommt es zu einer übermäßigen Auslastung der Schnittstelle zu ChatGPT, muss mit Überlastsituationen gerechnet werden. Die Aufträge an die KI können eventuell nicht in der gewohnten Geschwindigkeit ausgeführt werden. Es wird eine Meldung an alle Benutzerinnen und Benutzer herausgegeben. Werden mehr Aufträge eingestellt als die KI bearbeiten kann, bildet sich eine Warteschlange. Innerhalb dieser Warteschlange ist ein manuelles Abbrechen möglich, wenn sich die Journalistin bzw. der Journalist entschließt, die Aufgabe der KI doch selbst zu übernehmen, beispielsweise bei Übersetzungen. Ist die Überlastsituation beendet, wird die verbleibende Warteschlange abgearbeitet. Die Überlastsituation wird in Logfiles aufgezeichnet, um dafür zu sorgen, dass eine solche Situation in Zukunft nicht mehr auftritt.

Ausfallzeiten

Ist die Verfügbarkeit von ChatGPT begrenzt, kann es zu unvorhergesehenen Ausfallzeiten kommen. Das KI-Modell ist dann auf unbestimmte Zeit nicht erreichbar. Die Verfügbarkeit ist im SLA mit OpenAI festgelegt. Tritt solch ein Fall auf, werden der Ausfallzeitpunkt sowie alle laufenden Aufträge an ChatGPT automatisch dokumentiert. Außerdem werden die Anwenderinnen und Anwender informiert, dass ChatGPT vorübergehend nicht erreichbar ist. Sobald das KI-Modell wieder Aufträge entgegennehmen kann, werden die Anwenderinnen und Anwender informiert und die abgebrochenen Aufträge neu gestartet. Die Nutzerin bzw. der Nutzer kann auch Aufträge aus der Warteschlange entfernen, wenn in diesem Moment die manuelle Bearbeitung des Auftrags schneller ist.

Backup

Eine Instanz der Logfiles sowie der Konfiguration der Automation und der Schnittstellen werden in regelmäßigen Abständen in einem externen Speicher außerhalb des Medienunternehmens festgehalten, damit im Schadensfall die gesicherte Instanz wiederhergestellt werden kann, abgespeichert. Eine solche Versionierung erfolgt auch außerhalb des Rhythmus, vor und nachdem eine größere Veränderung am System durchgeführt wurde. Es wird ebenfalls eine Sicherungskopie der Dokumentation zum KI-Modell angefertigt.

Sicherheitsnotfall

Das Verhalten in einem Sicherheitsnotfall muss ebenfalls im SLA mit OpenAI definiert werden. Gemeinsam wird eine Policy erarbeitet, die in solchen Fällen greift [44, S. 35]. Mögliche Missbrauchsszenarien sind in Abschnitt 5.1.3 beschrieben. Die Aktionen passen sich an die Art des Angriffs an. Welche Maßnahmen im spezifischen Fall angewendet werden, ist daher nicht berechenbar. Es ist jedoch möglich, sich auf mehrere der wahrscheinlichsten Szenarien vorzubereiten [44, S. 35].

Grundlegend für alle Situationen ist jedoch, dass eine klare schnelle Kommunikation an die Mitarbeiter erfolgen muss. Es darf kein Chaos durch fehlende oder falsche Informationen entstehen. Jede Aktion muss mit der bzw. dem CISO abgestimmt werden. Diese Person hat die Leitung während eines Notfalls. Es darf kein eigenmächtiges Handeln von anderen Personen aufkommen.

Die Priorität des Medienunternehmens ist nach einem solchen Notfall in kurzer Zeit wieder auf sichere Weise sende- bzw. veröffentlichungsbereit zu sein. Mögliche Gegenmaßnahmen werden also mit größter Eile durchgeführt. Nach Beendigung der Notfallsituation muss der Schaden begutachtet werden. Eventuell werden Backups eingespielt, Geräte ausgetauscht oder andere Recovery-Maßnahmen durchgeführt. Gleichzeitig beginnen Nachforschungen nach den Auslösern und Verantwortlichen.

6.11 RISIKOANALYSE

Sich durch eine Risikoanalyse bereits zu Beginn mögliche Fallstricke bewusst zu machen, liefert für den späteren Projektverlauf einen Vorteil, da mögliche Handlungsoptionen schon dokumentiert sind. Ein Risiko ist ein eventuell eintretendes Ereignis und hat eine Abweichung vom Soll-Zustand zur Folge. Es wird nach Eintrittswahrscheinlichkeit und Schadenshöhe bewertet [76]. Ein dokumentiertes Risikomanagement ist nach dem Gesetz zur Künstlichen Intelligenz der EU außerdem Pflicht [41, Art. 9].

Aufbau

Die Risiken, die eine erfolgreiche Einbindung von ChatGPT in den Newsworkflow erschweren können, werden in drei Kategorien geteilt: Strategische, finanzielle und operationelle Risiken. Erstere sorgen dafür, dass das Medienunternehmen seinem Auftrag als Informationsvermittler nicht ausreichend nachkommen kann. Die Anerkennung und Reputation in der Gesellschaft schwindet. Diese Art des Risikos führt dazu, dass die Existenz des Medienunternehmens langfristig gefährdet ist. Finanzielle Risiken betreffen die Haushaltssituation des Medienunternehmens. Wird dieses aus öffentlichen Mitteln finanziert, ist eine Gefahr der Insolvenz nicht absehbar, trotzdem muss sich auch eine öffentlich-rechtliche Rundfunkanstalt für ihre Ausgaben rechtfertigen. Bei einem privaten Medienunternehmen ist die Gefahr jedoch real. Die dritte Kategorie sind operationelle Risiken. Diese betreffen die unmittelbare Handlungsfähigkeit des Medienunternehmens. Sie werden durch das Versagen von inneren Abläufen, Menschen und Systemen ausgelöst [76]. Schließlich werden die Risiken nach Schadenshöhe und Eintrittswahrscheinlichkeit bemessen und mit einem Ampelsystem bewertet (siehe Risikomatrix in Anhang sechs).

Hohe Risiken

- Gesetze oder Pressecodex werden nicht eingehalten (Strategisch)

Die Verletzung von Gesetzen rund um den Journalismus und des Pressecodex zieht empfindliche Strafen nach sich. Mögliche Gründe für den Verstoß durch KI sind eine fehlerhafte Programmierung und mindere Trainingsdatenqualität. Dadurch, dass Menschen den Output von ChatGPT nur stichprobenartig kontrollieren können, kann es passieren, dass fehlerhafter Output weiterverwendet wird. Trotzdem sind nicht alle Verstöße das Werk der KI. Mangelhafter Umgang mit Daten beispielsweise ist den menschlichen Akteuren zuzuschreiben.

- Veröffentlichung von falschen Fakten (Strategisch)

Die Veröffentlichung von falschen Fakten ist ein hohes Risiko, da nach der Veröffentlichung direkt die Öffentlichkeit vom Fehler des Medienunternehmens erfährt. Dieser entsteht durch eine mindere Trainingsdatenqualität und fehlende menschliche Endkontrolle vor der Ausspielung. Je nach Ausmaß des Fehlers erlebt das Medienunternehmen einen gewissen Reputationsverlust. Sobald ein Beitrag veröffentlicht ist, ist es fast unmöglich ihn wieder zurückzunehmen. Daher bleiben nur wenige Handlungsoptionen. Diese sind die Veröffentlichung einer Berichtigung und Entschuldigung.

- Veröffentlichung von Bias und Diskriminierung (Strategisch)

Ein ebenfalls öffentlichkeitswirksames Risiko ist eine Veröffentlichung eines offensichtlich voreingenommenen Beitrags. Die Gründe sind dieselben wie bei der Veröffentlichung falscher Fakten. Auch hier wird eine Berichtigung und Entschuldigung nötig. Im konkreten Fall kann dies sogar rechtliche Probleme nach sich ziehen.

- Sicherheitsnotfall (Strategisch)

Ein Sicherheitsnotfall ist ein seltenes, jedoch gravierendes Ereignis mit großer Öffentlichkeitswirkung. Je nach Art des Angriffs kommt es entweder zu einem Sendeausfall, der Preisgabe von Informationen oder der Nicht-Nutzbarkeit von technischen Systemen. Um einem solchen Zwischenfall entgegenzuwirken, sind eine technische Absicherung sowie die Sensibilisierung der Mitarbeiterinnen und Mitarbeiter nötig.

Mittlere Risiken

- Anfrage von einer Nutzerin oder einem Nutzer wird falsch beantwortet (Strategisch)

Dieses Risiko bezieht sich auf den operativen Einsatz des KI-Modells. Antwortet ChatGPT auf die Anfrage einer Nutzerin oder eines Nutzers falsch, unhöflich oder ist die Antwort nicht für die Frage relevant, kann es zu einem Reputationsverlust gegenüber der einzelnen Nutzerin bzw. dem einzelnen Nutzer kommen. Dieser Fehler kann durch mangelhaftes Design, zu geringe Trainingsdatenqualität oder fehlende menschliche Endkontrolle entstehen.

- Zu hohe Kosten (Finanziell)

Das allgemeine finanzielle Risiko sind zu hohe Ausgaben im Projekt. Sie können durch höhere Ausgaben für Tests, Integration, zusätzliches Personal und Hardware entstehen. Werden weitere finanzielle Mittel nicht allokiert, leidet die Qualität des Projektergebnisses oder die Laufzeit verlängert sich.

- OpenAI erhöht die Preise (Finanziell)

Ein vom Medienunternehmen nicht zu beeinflussendes Risiko ist die Anpassung der Preise durch OpenAI. Erhöhen sich die Preise, werden der Einführungsprozess und der operative Gebrauch von ChatGPT teurer. Es muss abgewogen werden, ob sich der Einsatz der KI noch lohnt und ob höhere finanzielle Mittel für die KI aufgewendet werden sollen.

- Einführungsprozess dauert zu lange (Operationell)

Die Gründe für einen verlängerten Einführungsprozess sind vielfältig. Die Einführung von ChatGPT kann eine große Menge an Planung und Tests benötigen. Die Arbeit an der systemischen Integration oder der grafischen Benutzungsoberfläche können ebenfalls viel Zeit in Anspruch nehmen. Um die zusätzliche Zeit auszugleichen, müssen also entweder weitere finanzielle Mittel aufgewendet oder auf Qualität verzichtet werden.

- Qualität des Outputs zu gering (Operationell)

Durch zu geringe Datenqualität kann der Output von ChatGPT für die geforderten Use Cases nicht mehr verwendet werden. Dadurch kann effektiv keine Aufgabe von der KI übernommen werden. Das Risiko kann durch Prompt Engineering vermindert werden.

- Neue Prozesse sind zu umständlich (Operationell)

Sind die neugestalteten Prozesse tatsächlich umständlich und liefern mehr Arbeitsaufwand als Ergebnisse, müssen diese verschlankt und angepasst werden. Wichtig ist dabei, das Feedback der Mitarbeiterinnen und Mitarbeiter erneut einzuholen und die Prozesse entsprechend neu zu entwerfen.

- Neue Prozesse werden als zu umständlich empfunden (Operationell)

Die Anpassung von Arbeitsprozessen stößt bei einer routinierten Belegschaft schnell auf Widerstand. Wird die Veränderung als nicht sinnvoll empfunden, ist ChatGPT ebenfalls nicht vollständig nutzbar. Wichtig ist hier, die Mitarbeiterinnen und Mitarbeiter zu überzeugen, dass der Wandel seine Berechtigung hat. Außerdem kann durch zusätzliche Informationen und ein gutes Vorbild der Leitungsebene eine Einsicht erreicht werden.

- Niedrige Akzeptanz der Mitarbeiterinnen und Mitarbeiter (Operationell)

Der Veränderungsprozess steht und fällt mit dem Willen der Belegschaft zur Teilhabe am Wandel. Ohne Mitarbeiterinnen und Mitarbeiter, die ChatGPT nutzen, ist die Implementierung desselben nicht nötig. Durch Ausbau der Schulungen und Schaffung von Anreizen kann die Mitarbeiterschaft von der neuen Technologie überzeugt werden.

Niedrige Risiken

- Fehlende Kennzeichnung von KI-generierten Inhalten (Strategisch)

Eine Kennzeichnung von KI-generierten Inhalten bei der Veröffentlichung wird im Medienstaatsvertrag gefordert. Den Nachweis zu erbringen, dass ein Text nicht von einem Menschen verfasst wurde, ist jedoch schwierig, da ChatGPT mittlerweile Texte verfasst, die auch die eines Menschen sein könnten.

- Mehr Aufgaben an ChatGPT als geplant (Finanziell)

Findet das KI-Modell ChatGPT großen Anklang im Medienunternehmen und wird häufig genutzt, entstehen zusätzliche Kosten, da pro verwendetem Token abgerechnet wird. Da es sich hier um eine positive Entwicklung handelt, sollten hier die finanziellen Mittel erhöht werden, damit die Begeisterung für KI beibehalten oder sogar gesteigert wird.

- Übermittlungs- oder Rechengeschwindigkeit zu niedrig (Operationell)

Wenn die Arbeitsgeschwindigkeit mit ChatGPT zu gering ist, kann nicht die gewünschte Anzahl Aufträge in einer gewissen Zeit an das KI-Modell übergeben werden. Der Prozessablauf ist dadurch verzögert. Im besten Fall wird der Beitrag noch immer rechtzeitig veröffentlicht, da Aufgaben der KI auch von einem Menschen durchgeführt werden können.

- Zu geringe Verfügbarkeit (Operationell)

Dasselbe Problem tritt bei Ausfallzeiten von ChatGPT auf. Aufträge werden nicht abgearbeitet und Prozesse verzögert. In einem solchen Fall ist die spezifische Stelle des Ausfalls zu identifizieren und herauszufinden, in welchem Verantwortungsbereich die Schwachstelle liegt.

6.12 WIRTSCHAFTLICHE ANALYSE

Die Analyse behandelt zwei Bereiche. Zunächst werden Metriken für den Fortschritt im Projekt vorgestellt. Diese beschreiben die Arbeit in der Einführung von ChatGPT. Da ein mögliches finanzielles Risiko für das Scheitern des Projekts das übermäßige Andauern der Einführungsphase ist, muss ein solches Risiko mit den Kennzahlen erkannt und verhindert werden.

Der zweite Teil der Analyse behandelt die Kosten für den Einsatz von ChatGPT. Ebenfalls ein Risiko ist, dass der produktive Einsatz von ChatGPT hohe finanzielle Ressourcen verbraucht. Mit Hochrechnungen kann zumindest eine grobe Summe abgeschätzt und eingeplant werden. Außerdem können hier die Kosten im Vergleich zu einer Produktion ohne ChatGPT gesehen werden.

Kennzahlen für den Projektfortschritt

Die Erfassung des Projektfortschritts bei agilen Projekten wird auf andere Weise dokumentiert als bei Projekten nach dem Wasserfall-Modell. Im Wasserfall-Modell kann man anhand des Projektzeit- und des Meilensteinplans einen Soll-Ist-Vergleich ziehen. Bei agilen Projekten steht der Projektplan nicht von vornherein fest. In jedem Sprint werden neue Ziele gesteckt. Funktionalitäten, die zu Anfang noch im Product Backlog standen, können im Laufe des Projekts hinfällig werden. Daher müssen neue Metriken angelegt werden, um den Projektfortschritt festzustellen.

- **Sprint Burndown**

Der Sprint Burndown zeigt, wie viele Aufgaben in diesem Sprint von den geplanten Aufgaben bereits erledigt wurden und welche noch zu erledigen sind. Dabei handelt es sich um ein Koordinatensystem, in dem die Zeit auf der waagrechten Achse und der Arbeitsfortschritt auf der senkrechten Achse aufgetragen sind [82].

- **Story Completion**

Die Story Completion zeigt auf, wie viele und welche User Stories während des gesamten Projekts schon umgesetzt wurden. Dabei kann man ebenfalls ablesen, in welchem Sprint welche Story vollendet wurde. Diese Größe wird anteilig von der Gesamtzahl der umzusetzenden User Stories in Prozent angegeben [82]. Jede User Story wird wiederum in Punkte zerlegt, um sich innerhalb der Story Zwischenziele zu stecken und umfangreichere Stories mit mehr Punkten zu belohnen.

- **Release Burnup**

Der Release Burnup stellt dar, wie viele Story Points aus dem Product Backlog während des gesamten Projekts schon erledigt sind. Diese werden anteilig an den gesamten Story Points im Product Backlog erfasst [82].

- **Durchsatz**

Der Durchsatz gibt an, wie viele Story Points pro Sprint erledigt werden. Dadurch wird die kurzfristige Arbeitsgeschwindigkeit im Sprint dargestellt [82].

Beispielhafte Rechnung für textuellen Artikel

Möchte der Journalist einen Artikel von einer Seite Text⁸ mit dem Chatbot ChatGPT auf Sprachrichtigkeit und Stil prüfen lassen, müssen folgende Überlegungen angestellt werden. Laut Angaben von OpenAI entspricht eine Seite Text ungefähr 800 Token. Es müssen jeweils die Bearbeitung des Textes und die Generation einer Antwort berechnet werden. Ein genaueres Ergebnis kann mit dem Tokenizer TikToken von OpenAI berechnet werden [81].

$$\text{Kosten} = 800 \text{ Token} * (\$0,0015 * \frac{1}{1000 \text{ Token}} + \$0,002 * \frac{1}{1000 \text{ Token}}) = \$0,0028$$

Die Bearbeitung von 500 Wörtern Text kostet ungefähr 0,28 US-amerikanische Cents. Das entspricht 0,26 europäischen Cents⁹ [83].

Beispielhafte Rechnung für einen Chat

Wird der Chatbot ChatGPT im Rahmen einer Anfrage einer Nutzerin oder eines Nutzers nach der Postadresse des Medienunternehmens gefragt, werden für den Prompt ungefähr 37 Token benötigt. Die Antwort wiederum benötigt ungefähr 60 Token [78].

$$\text{Kosten} = 37 \text{ Token} * \$0,0015 * \frac{1}{1000 \text{ Token}} + 60 \text{ Token} * \$0,002 * \frac{1}{1000 \text{ Token}} = \$0,0001755$$

Ein solcher Chataustausch kostet also in Summe 0,01755 US-amerikanische Cents. Das entspricht ungefähr 0,016 europäischen Cents [83].

Einsparungen durch die Einführung eines KI-Modells

Um den wirtschaftlichen Erfolg des Projekts zu messen, müssen die obengenannten Kosten dem Nutzen, der durch ChatGPT eingebracht wird, gegenübergestellt werden. Öffentlich-rechtliche und damit beitragsfinanzierte Rundfunkanstalten sind weniger auf wirtschaftlichen Gewinn getrimmt. Trotzdem müssen sie begründen, wofür sie ihre finanziellen Mittel ausgeben. Private Medienunternehmen sind auf Gewinn durch die Vermarktung ihrer Inhalte angewiesen. Daher müssen folgende Fragen gestellt werden: Erhöht sich die Anzahl der veröffentlichten Beiträge durch ChatGPT? Wenn ja, wie stark? Werden mehr Kanäle bespielt? Hat sich die Anzahl der Nutzerinnen und Nutzer erhöht? Dazu kommen Fragen, die darauf abzielen, ob sich die Einführung wirtschaftlich gelohnt hat. Es werden also Vergleiche zwischen der aktuellen Situation mit ChatGPT und der fiktiven Situation ohne ChatGPT gezogen. Hätten mehr Mitarbeiterinnen und Mitarbeiter eingestellt werden müssen, um dieselbe Anzahl veröffentlichter Beiträge zu erreichen? Wie viel kostet der Einsatz von KI im Vergleich zur Einstellung neuer Mitarbeiterinnen und Mitarbeiter? Wann amortisiert sich die Einführung des KI-Modells?

⁸ hier 500 Wörter

⁹ Die Umrechnungskurse stammen aus dem Zeitraum der Veröffentlichung der Thesis.

7 TECHNISCHE LÖSUNGSBESCHREIBUNG

Zur Landschaft, die rund um die Einführung von ChatGPT kreiert wird, gehört ebenfalls eine technische Lösung, um die Konzepte aus Kapitel sechs auch umzusetzen. Dieses beginnt mit der Definition der Informationsflüsse. Sie geben an, wo wann welche Art von Material bewegt, bearbeitet, erstellt oder gelöscht wird. Anhand dieser Vorgaben wird die Hardware, die für die Umsetzung der Flüsse nötig ist, präsentiert. Geräte und deren Verbindungen werden in einem Blockschaltbild dargestellt.

Es wird zunächst eine API-Verbindung zu ChatGPT aufgebaut. Um einen gewinnbringenden Einsatz von ChatGPT zu ermöglichen, müssen ebenfalls Schnittstellen zwischen ChatGPT, CMS und MAM definiert werden. Zwischen diesen Schnittstellen werden Daten und Material ausgetauscht. Für den erfolgreichen Austausch über die Schnittstellen muss ein einheitliches Metadatenformat gewählt werden. Alle Komponenten müssen daher dieselben Klassen und Unterklassen sprechen und dieselben Bewertungskriterien anlegen. Die Informationsflüsse zwischen den Komponenten werden für jeden Anwendungsfall beschrieben. Ebenfalls in diesem Abschnitt enthalten sind die technischen Lösungen für Automation, Schnittstellen und Austauschformat dargelegt. Dadurch können einem Chatbot, der sonst auf Kommunikation in natürlicher Sprache spezialisiert ist, auch strukturierte Daten entnommen werden.

Es muss eine Lösung für die Automatisierung der Flüsse entwickelt werden. Beispielsweise der Informationsfluss beim Ingest-Workflow sollte, abgesehen von der initialen Einspeisung der Essenz in den Speicher des Medienunternehmens, keine Aktion eines Menschen enthalten. Die Übergabe der Aufgaben an ChatGPT müssen also automatisiert werden.

Ein ebenfalls wichtiger Bereich in der technischen Lösung ist die Planung der grafischen Benutzungsoberflächen. Es müssen mehrere Oberflächen für verschiedene Nutzungsprofile erstellt werden. Eine weitere Maßnahme, um die Sicherheit im Umgang mit dem KI-Modell zu erhöhen, ist die Authentifikation von Eingaben. Das KI-Modell muss also durch kryptografische Verfahren prüfen, dass die Eingabe entweder von einem befugten Menschen oder von einer beabsichtigten Aktion der Automation stammt. Mit dieser Direktive kann verhindert werden, dass Missbrauch mit ChatGPT betrieben wird.

Das sogenannte Prompt Engineering bezeichnet die Anpassung von Eingaben in ChatGPT, sodass die Aufgabe für ChatGPT möglichst präzise beschrieben wird und im Sinne des Auftraggebers ausgeführt wird.

7.1 INFRASTRUKTUR

Für die Einführung von ChatGPT in den Newsworkflow sind neben den prozessualen und sicherheitstechnischen Maßnahmen auch Eingriffe in die Systemlandschaft des Medienunternehmens nötig. Um ChatGPT anzusteuern, bedarf es entweder der Implementierung einer neuen Automation oder der Einbindung der bereits im Medienunternehmen bestehenden Automation.

Das System besteht aus einem Dreieck aus ChatGPT, Nutzer und MAM. Diese sind über die Automation sowie direkt verbunden. Die Schnittstellen zwischen den Systemkomponenten sind jeweils mit APIs gestaltet. An das MAM angegliedert ist das CMS. Alle Informationen, die vom CMS in das Dreieck gegeben werden, werden zunächst vom MAM verwaltet. Die Nutzerin bzw. der Nutzer schickt Anfragen an das KI-Modell über eine GUI an die Automation und diese leitet die Informationen an ChatGPT weiter (siehe Abbildung 18).

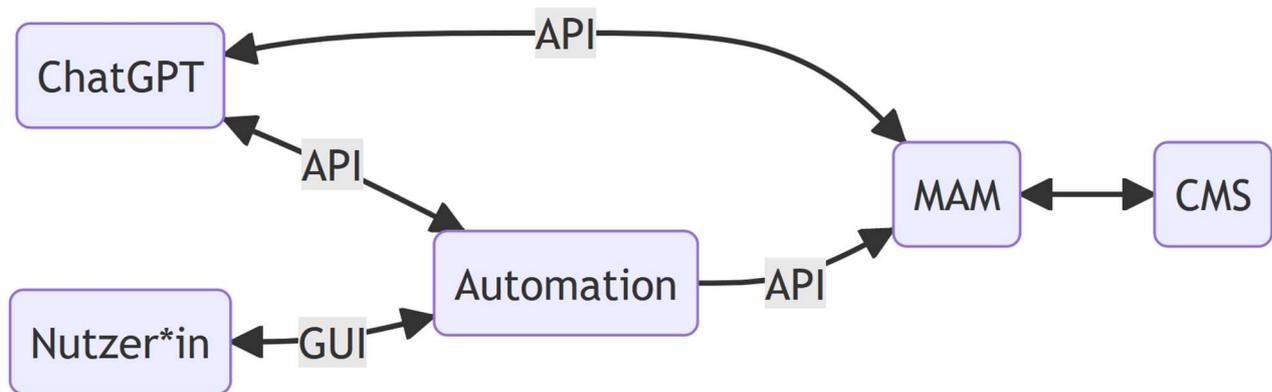


Abbildung 18: Systemaufbau mit ChatGPT nach [84, S. 51].

Automationssoftware

Die Automation fungiert hier als Vermittlung zwischen Nutzerin bzw. Nutzer, ChatGPT und MAM. Der Vorschlag für eine neue Automationssoftware für diesen Anwendungsfall ist Node-RED. Dabei handelt es sich um ein Tool, das Verbindungen zwischen Hardware, APIs und Onlinediensten schafft. Der browser-basierte Editor bietet eine grafische Benutzeroberfläche, um Flüsse aufzubauen, miteinander zu verbinden und zu testen [85].

MAM

Bei einem MAM-System handelt es sich um eine Medienverwaltungssoftware. Es ist also eine Datenbank, die Informationen zu den Essenzen enthält. Die primären Aufgaben sind Suche nach Schlagwörtern, Auswahl, Bestellung, Archivierung aller erzeugten Versionen, Erzeugung von technischen und beschreibenden Metadaten und die Integration von Schnitt- und Grafiksystemen. Das MAM-System stellt somit die Schnittstelle zwischen Nutzer und Speicher dar [86, S. 3].

CMS

Ein CMS ist ein System aus Software und Hardware, das für die Produktion, Publikation und Archivierung von Content entwickelt wurde. Ein CMS arbeitet mit referenzierten Essenzen im Speicher. Es ist über Schnittstellen mit der Studioautomation, dem Redaktionssystem und dem Schnittsystem verbunden. Nach dem Ingest wird der Content Beiträgen im Redaktionssystem zugeordnet und Metadaten werden erzeugt. Das CMS arbeitet sozusagen als Schnittstelle zwischen IT und Broadcast [86, S. 3].

7.2 INFORMATIONENFLÜSSE

KI ist maßgeblich von den Daten sowie dem Textmaterial abhängig, das ihr zugeführt wird. Daher muss definiert werden, welche Informationen wann ChatGPT erreichen und wie die verarbeiteten Informationen wiederum im Medienunternehmen verbreitet werden. Die Flüsse unterscheiden sich je nach Anwendung des KI-Modells. Die vollständigen Ablaufdiagramme befinden sich in Anhang sieben.

Manuelle Anfrage an die KI

Eine manuelle Anfrage an ChatGPT zeichnet sich dadurch aus, dass ein Mensch den Auftrag auslöst und auswählt, welche Essenz das Modell bearbeiten soll. Dabei kann es auch sein, dass dem KI-Modell eine gerade verfasste Essenz übergeben wird. Der Fluss wird durch die Eingabe eines Menschen in die GUI von ChatGPT ausgelöst. Dabei greift der Mensch entweder auf vorgefertigte Prompts für wiederkehrende Aufgaben zurück oder gibt einen individuellen Prompt ein. Ist für die gestellte Aufgabe eine Essenz nötig, wird diese ebenfalls in der Anfrage referenziert. Die Referenz ist dabei die eindeutige Identifikation (ID), die jeder Essenz im Speicher zugeordnet wird. Der Mensch schickt den Auftrag ab. Der Auftrag wiederum wird mit einer eindeutigen ID versehen. Zu seinen Informationen wird das Datum und die Uhrzeit des Auftrags ergänzt sowie Inhalt und Referenz des Auftrages dokumentiert. Schließlich wird der Auftrag an die KI übergeben und bearbeitet. Für die Bearbeitung greift die KI über die Referenz auf die Essenz im Speicher zu. Ist der Auftrag beendet, wird die Antwort mit derselben ID gespeichert. Eine neu entstandene Essenz wird in der Antwort referenziert. Außerdem wird der Zeitpunkt der Antwort geloggt. Sind diese Informationen gespeichert, wird die Antwort an den Menschen zurückgeleitet. Ist der Mensch mit der Antwort nicht zufrieden, kann er eine neue Anfrage stellen oder den Text händisch bearbeiten. Eine Möglichkeit, Feedback zu der Antwort zu geben, ist ebenfalls vorhanden.

Automatische Anfrage an die KI

Die automatische Anfrage an die KI ähnelt der manuellen Anfrage mit dem Unterschied, dass in diesem Fluss kein Mensch beteiligt ist. Eine automatische Anfrage ist beispielsweise Teil des Ingest-Workflows. Eine Automation gibt ähnlich wie der Mensch einen vorgefertigten Prompt und eine Referenz auf eine Essenz an ChatGPT. Es werden ebenfalls die Informationen zum Auftrag gespeichert und die KI bearbeitet den Auftrag. Die Antwort erhält dieselbe ID. Datum, Uhrzeit, Inhalt und Referenz werden dokumentiert. Die Antwort wird an das System gegeben, das der Ausgangspunkt der Anfrage war. Im Ingest-Workflow stellt zwar die Automation die Anfrage, aber die Antworten, hier beschreibende Metadaten, werden gemeinsam mit der Essenz im Speicher abgelegt. Ein Mensch hat nun stichprobenartig die Möglichkeit, die Antworten auf Plausibilität zu prüfen.

Ingest

Der Ingest-Workflow im Besonderen hat zum Ziel, möglichst viele präzise Metadaten anhand eines Metadaten-Standards einer Essenz zuzufügen. Der gesamte Workflow soll möglichst automatisiert ablaufen. Der Mensch soll nur über wenige Schnittstellen verfügen. Der Workflow beginnt mit der Übertragung einer neuen Essenz in den Speicher des Medienunternehmens. Der Ursprung der Essenz kann eine Studioproduktion, eine Außenaufzeichnung, der Schnitt oder die Zulieferung von externem Material sein. Alles Material, das ingestiert wird, wird zunächst in einem Zwischenspeicher aufgefangen. Ein sogenannter Watchfolder erkennt die Anwesenheit von neuem Material und der Ingest-Workflow wird in Gang gesetzt.

Jeder neuen Essenz wird durch das CMS eine eindeutige ID zugeordnet, damit diese im Speicher wiedergefunden und referenziert wird. Dann wird das Material auf seine technischen Eigenschaften und Qualität geprüft. Die Ergebnisse werden in den zugehörigen Metadaten vermerkt. Dies ist besonders bei externen Produktionen zu beachten. Liegt das Material nicht im korrekten Format vor, muss es transkodiert und erneut gespeichert werden. Danach stellt die Automation eine Anfrage an das KI-Modell. Diese läuft ähnlich einer automatischen Anfrage an ChatGPT ab. Die Antwort von ChatGPT sind Metadaten. Diese werden in das CMS überführt und mit der Essenz verbunden. Zuletzt werden Rechte und Nutzungsvorbehalte, die aus dem gelieferten Material ersichtlich sind, ebenfalls in den Metadaten gespeichert. Abschließend werden Essenz und Metadaten im Speicher des Medienunternehmens abgelegt.

Assetbearbeitung

Wird durch KI ein bestehendes Asset bearbeitet, verläuft der Informationsfluss ähnlich zur generellen Anfrage, jedoch mit kleinen Unterschieden. Die Bearbeitung schließt auch die Überführung in einen anderen Medientyp mit ein. Wenn beispielsweise eine Tonaufnahme zu Text transkribiert wird, bleibt der Inhalt derselbe, es ändert sich nur die Darstellungsform. Durch die Bearbeitung einer Essenz wird eine neue Version derselben erzeugt und wird daher wie eine neue Essenz im Speicher behandelt. Deshalb wird für diese Essenz der Ingest-Workflow durchgeführt. Dabei können schon Informationen von der ursprünglichen Essenz übernommen werden. Außerdem wird eine Versionierung in Bezug auf die Ausgangsessenzen vorgenommen. Die bearbeitete Essenz wird an die Anwenderin bzw. den Anwender zurückgeleitet und sie oder er bewertet die Antwort der KI. Durch Feedback in der GUI wird diese Art der Bearbeitung als gewinnbringend eingestuft. Ist die Anwenderin bzw. der Anwender nicht zufrieden, wird ein Folgeauftrag mit detaillierten Informationen zum Fehler an ChatGPT gestellt.

Assetgeneration

Die Assetgeneration läuft ähnlich ab wie die Assetbearbeitung. Es wird auf dieselbe Weise eine Anfrage an ChatGPT gestellt. Der Unterschied ist, dass in diesem Fall auf kein weiteres Asset referenziert wird. Das Asset wird ausschließlich aus dem Prompt oder einer Datendatei erzeugt. Keine Mediendatei dient hier als Vorbild. Ein Mensch gibt also den Prompt in die grafische

Benutzungsoberfläche von ChatGPT ein. Damit wird der Auftrag abgesendet. Der Mensch erhält eine Bestätigung über den Auftragseingang. Gleichzeitig wird der Auftrag mit ID, Zeitpunkt und Inhalt geloggt und an die KI gegeben. Diese verarbeitet die Aufgabenstellung und generiert eine Antwort. Die Antwort besteht hier aus einer neuen Essenz, die, bevor sie an den Menschen weitergeleitet wird, zunächst den Ingest-Workflow durchläuft. Dadurch, dass es sich hier um eine KI-generierte Essenz handelt, wird ein kurzer Disclaimer der Essenz angehängt. Diese Tatsache wird ebenfalls in den Metadaten vermerkt (A5). Ist der Ingest-Workflow beendet, wird die Antwort an den Menschen zurückgeleitet und dieser kontrolliert das Werk der KI. Er gibt Feedback über die Qualität der Essenz. Ist die Qualität nicht ausreichend, stellt er einen weiteren Auftrag mit verändertem Prompt an die KI. Bei der Generation jeder Essenz muss als Fehlerschutz eine nachgelagerte Sentiment Analysis erfolgen, um möglicherweise unhöfliche, voreingenommene oder gefährliche Texte zu identifizieren.

Rückmeldung von Nutzerinnen und Nutzern

Bei der Bearbeitung von Anfragen von Nutzerinnen und Nutzern wird aufgrund der Menge der Anfragen mit Unterstützung des KI-Modells gearbeitet. Eintreffende Nachrichten, die nicht sofort bearbeitet werden müssen, wie beispielsweise Emails oder Instant Messages, werden über Schnittstellen zu den jeweiligen Kommunikationsmitteln in den Speicher des Medienunternehmens übertragen. Dann werden die Anfragen zunächst von einer Sentiment Analysis bearbeitet, um unhöfliche und beleidigende Nachrichten im Vorhinein auszuschließen. Diese wird ebenfalls über eine automatisierte Anfrage durchgeführt. ChatGPT entscheidet ebenfalls, ob es sich bei der Anfrage um eine „einfache“ Anfrage handelt und diese von der KI beantwortet werden kann oder ob die Anfrage komplexer ist und die Bearbeitung eines Menschen bedarf. Entsprechend werden die Anfragen durch ChatGPT beantwortet und eine Nachricht zurück an die Nutzerin bzw. den Nutzer gesendet. Auf der anderen Seite werden die Nutzeranfragen – unabhängig von ihrem ursprünglichen Kommunikationsmittel – in einer grafischen Benutzungsoberfläche für die Journalistin bzw. den Journalisten aufbereitet. Dieser verfasst die Antwort und schickt sie über die GUI zur Nutzerin bzw. zum Nutzer. Jede Nachricht wird im Speicher des Medienunternehmens mit Absender, Inhalt, Zeitpunkt und gegebener Antwort abgelegt. Sie wird wie eine neue Essenz behandelt und durchläuft eine verkürzte Version des Ingest-Workflows. Jegliche Informationen, die zu dieser Nachricht der Nutzerin oder des Nutzers im Medienunternehmen bekannt sind, werden als Metadaten gespeichert. Dadurch können auch die Anfragen von Nutzerinnen und Nutzern für Ideen und Recherche verwendet werden. Es kann jedoch auch der Fall auftreten, dass sich Anfragen zu einem bestimmten Thema häufen. Durch die Speicherung und Behandlung mit den Ingest-Funktionalitäten liegen die Nachrichten strukturiert vor, sodass schnell festgestellt werden kann, wie viele Nachrichten in welcher Gefühlslage zu welchem Thema existieren.

7.3 INTERFACES

Der Abschnitt „Interfaces“ behandelt die Implementierung von ChatGPT. Es werden Techniken beschrieben, wie strukturierte Informationen zwischen den Systemen und der Nutzerin bzw. dem Nutzer ausgetauscht werden. Die Lösung umfasst einen Vorschlag zur Automatisierung, der Übertragung von Metadaten, der Interaktion mit der Anwenderin bzw. dem Anwender und die Verschlüsselung der Daten.

7.3.1 AUTOMATION

Workflow Management wird als die Automatisierung von Workflows verstanden. Inhalte und Aufgaben werden von einer Person zur anderen weitergegeben, sodass sie mit den geplanten Abläufen übereinstimmen. Die Automation nutzt eine Software, um die Flüsse zu managen. Eine wichtige Eigenschaft des Workflow Managements ist, dass eine Software Flexibilität erlaubt, sodass Medienunternehmen sich laufend anpassen und verbessern können, um ihre Geschäftsziele zu erreichen und Technologiewechsel mitzumachen [87]. Die Automationssoftware dient in diesem Fall als Vermittlung zwischen Mensch, ChatGPT und MAM. Sie leitet Informationen weiter, führt die Log-Einträge und gibt Aufgaben in Auftrag. Jede Nutzung des KI-Modells läuft über die Automation (A33 und A34).

Funktionsweise von Node-RED

Node-RED ist ein browser-basiertes Tool, um Hardware, APIs und Online-Dienste mithilfe einer grafischen Benutzungsoberfläche zu verbinden. Damit können komplexe Workflows geschaffen, getestet und ausgeführt werden. Umfangreiche Libraries und Vorlagen ermöglichen einen schnellen Einstieg [85].

Messages

Die Workflows basieren darauf, dass sogenannte Messages zwischen den Nodes ausgetauscht werden. Die Messages sind JavaScript-Objekte und können jedes Key-Value-Paar aufnehmen. Messages haben eine Payload. Darin werden alle Informationen gespeichert, die transportiert werden (siehe Abbildung 19) [88].

```

msp.payload = {
  "FirstName": "Fred",
  "Surname": "Smith",
  "Age": 28,
  "Phone":
  {
    "Type": "home",
    "Number": "071233456789"
  }
}

```

Abbildung 19: Beispiel für eine Message in JavaScript Object Notation (JSON) nach [88].

Nodes

Die Definition von Workflows wird mithilfe von verschiedenen Node-Typen durchgeführt:

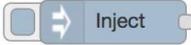
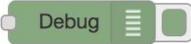
Node-Typ	Funktionalität	Beispielhafter Einsatz im Workflow	Darstellung im Editor
Inject-Node	Manueller Auslöser eines Prozesses, eine Message wird versandt	Beginn einer Anfrage an die KI	
Debug-Node	Darstellung von Messages	Fehlersuche bei der Einrichtung	
Function-Node	Ausführung von JavaScript-Code auf Messages	Bearbeitung der Messages	
Switch-Node	Informationen können je nach Parameter in unterschiedliche Richtungen weitergegeben werden	Einordnung der Rückmeldung von Nutzer*innen nach der Sentiment Analysis	
Change-Node	Verändert Informationen innerhalb einer Message	Status der Message ändert sich nach Durchlauf eines Workflows	
Template-Node	Text wird in eine vorbereitete Message gefüllt	Ergebnisse von ChatGPT werden an das MAM weitergegeben	

Abbildung 20: Node-Typen in Node-RED nach [89].

Prompt Chaining

Obwohl LLMs sehr vielseitig sind, können komplexe Aufgaben aus der praktischen Anwendung oft nicht mit nur einem Prompt erledigt werden. In manchen Fällen muss erst eine Aufgabe vollständig bearbeitet werden, bevor die nächste dem LLM übergeben werden kann [90, S. 1]. Daher werden die komplexen Aufgaben in mehrere Unteraufgaben geteilt und ihre Prompts verkettet. Der Output einer Anfrage bildet die Eingabe für die nächste [90, S. 2]. Mit der Kettenstruktur der Node-RED-Workflows ist eine solche Technik ebenfalls möglich. Die Daten werden dann mit Messages übergeben.

7.3.2 SCHNITTSTELLEN

Die Schnittstellen zwischen den Systemen (siehe Abschnitt 7.1) sind APIs. Für die Verbindungen zwischen MAM, Automation und ChatGPT wird eine sogenannte REST (Representational State Transfer)-API verwendet. Diese API beschreibt eine Sammlung von Architekturbeschränkungen im Vergleich zur herkömmlichen API. Die API wird mittels einer Programmiersprache aufgerufen. Die Informationen, die über die API übertragen werden, werden über das Hypertext Transfer Protokoll (HTTP) versandt [91, S. 5]. Die REST-Architektur schränkt die Auswahl der Methoden, mit denen die Informationen übermittelt werden, ein. Die verbleibenden Methoden sind GET, POST, PUT und DELETE. Damit verfolgt diese Schnittstelle den CRUD-Ansatz [91, S. 32]. Dieser beschreibt eine Kommunikation mit ausschließlich vier Operationen: Create, Read, Update und Delete. Diese Operationen stimmen mit den vier HTTP-Methoden überein [91, S. 34].

Eine REST-API greift mit einer Methode auf eine Ressource zu. Die Ressource muss über einen eindeutigen Uniform Resource Identifier (URI) erreichbar sein. Jede Anfrage wird von neuem aufgebaut und mit Informationen bestückt, denn das Protokoll arbeitet zustandslos. Es werden keine Informationen aus einer Anfrage für die nachfolgende in Cookies gespeichert [91, S. 33]. Eine häufige Darstellung der übermittelten Informationen ist die JavaScript Object Notation (JSON). In einer JSON-Datei werden Daten als Schlüssel-Wert-Paare gespeichert. Dadurch können individuelle Datenstrukturen aufgebaut werden. Als Datentypen für den Wert stehen die gängigen Typen wie Null, Boolean, Zahlen, Strings und Arrays zur Verfügung. Schlüssel können beliebig vergeben werden [91, S. 37]. Durch die Verwendung eines Modells benötigt die API auch nur einen URI, nämlich `.../v1/chat/completions`. Für Anfragen an ChatGPT wird die Methode POST verwendet [92].

Beispiel: Anfrage an ChatGPT

In diesem Beispiel soll ChatGPT den Prompt "Say this is a text!" erhalten. Ein Request mit mehreren Parametern wird in Node-RED durch die Abfolge in Abbildung 21 abgebildet.



Abbildung 21: Komplexerer Request mit Parametern in Node-RED [93].

Die Automation erhält über den Inject-Node den Prompt der Nutzerin bzw. des Nutzers. Der Function-Node `set_payload_and_headers` setzt die Payload und die Headerparameter (siehe grüne Markierung in Abbildung 22). Es können beliebige Parameter in der Payload und im Header gewählt werden. Der Prompt wird im realen Fall von der Nutzerin bzw. dem Nutzer übergeben. In diesem Beispiel ist er jedoch hardcoded (siehe blaue Markierung in Abbildung 22).

```
test_prompt = 'Say this is a test!'
def set_payload_and_headers(prompt)
    msg.headers = {};
    msg.headers['Content-Type'] = 'application/json';
    msg.headers['Authorization'] = f"Bearer {OPENAI_API_KEY}";
    msg.payload = 'model': 'gpt-3.5-turbo',
        'messages': [
            {
                'role': 'user',
                'content': prompt
            }
        ],
        'response' = openai.ChatCompletion.create(
            model="gpt-3.5-turbo-0613",
            messages=messages
        ),
    ;
    return msg

set_payload_and_headers(test_prompt)
```

Abbildung 22: Inhalt des Function-Nodes in Python nach [93] und [94].

Der folgende Knoten wurde aus einem HTTP-Node erstellt und schickt die Message, die im vorherigen Knoten erstellt wurde an den API-Endpunkt `https://api.openai.com/v1/chat/completions`. Der Debug-Node stellt die Message zur Kontrolle im browser-basierten Editor von Node-RED dar.

Beispiel: Antwort von ChatGPT ausgeben

Die Antwort von ChatGPT wird wiederum an einen API-Endpunkt von Node-RED gegeben. Der Fluss wird durch den Inject-Node angestoßen. Dieser reagiert auf einkommende HTTP-Responses. Der Change-Node verändert die Einstellungen für einkommende Responses so, dass die Response in die Payload einer Message übertragen wird, sodass diese im Workflow weitergegeben und bearbeitet werden kann. Der HTTP-Node liefert schließlich die Response, die über die

vorhergehenden Einstellungen bearbeitet wird. Ein Debug-Node stellt die Message zur Kontrolle dar (siehe Abbildung 23).

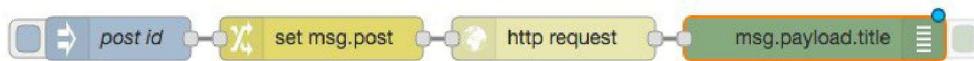


Abbildung 23: JSON-Response wird erhalten in Node-RED [95].

Funktionen für wiederkehrende Aufgaben

ChatGPT ist dafür bekannt, ein Gespräch mit einem Menschen zu simulieren. Daher antwortet das Modell auf gestellte Fragen in ganzen Sätzen und mit zusätzlichen Erklärungen zum Sachverhalt. Für Use Cases, in denen die Antwort von ChatGPT direkt in den Metadaten einer Essenz abgespeichert werden soll, ist dies jedoch nicht geeignet. Damit ChatGPT also strukturierte Antworten im JSON-Format ausgibt, werden einem Eingabeprompt direkt eine oder mehrere Funktionen hinzugefügt. Diese Funktionen definieren Eingabe- und Ausgabevariablen. Die Eingabevariablen werden mit den Informationen aus den textuellen Prompt gefüllt. Die Ausgabevariablen werden mit den Informationen aus der Antwort von ChatGPT versehen. Werden einem Prompt mehrere Funktionen mitgegeben, entscheidet ChatGPT, welche Funktion die geeignete für diesem Prompt ist, und diese wird ausgeführt. Innerhalb der Ausführung können auch Drittsysteme (z.B. das MAM) eingebunden werden. Die Definitionen der Funktionen zählen zu den Eingabetoken dazu [96]. Die Ausgabevariablen werden wiederum an ChatGPT zurückgegeben und das Modell formuliert daraus eine Antwort in natürlicher Sprache.

Der Ablauf einer Konversation mit Funktionen beginnt damit, dass ChatGPT eine Anfrage übergeben wird. Diese Anfrage enthält einen Prompt, eine oder mehrere Funktionen und deren Ein- und Ausgabeparameter. ChatGPT verarbeitet den Prompt und verbindet diesen mit einer Funktion. Daher ist es wichtig, den Funktionen einen aussagekräftigen Namen zu geben. Passt keine Funktion zum Prompt, wird eine unstrukturierte Antwort gegeben. Dann wird die Funktion mit ihren Parametern aufgerufen und bearbeitet. Die strukturierte Antwort wird an ChatGPT zurückgegeben, in Sprache umgeformt und an die laufende Konversation mit ChatGPT angehängt. Es können die Ausgabeparameter der Funktion sowie eine textuelle Antwort, die die Ausgabeparameter zusammenfasst, ausgegeben werden [96]. Soll die Antwort an einen Menschen zurückgegeben werden, eignet sich eine Antwort in natürlicher Sprache. Soll die Antwort wiederum direkt in die Metadaten einer Essenz eingepflegt werden, wird die strukturierte Antwort verwendet. Abbildung 24 visualisiert den Ablauf für eine Anfrage in einem Bestellsystem. Der Ablauf lässt sich jedoch auch auf die Workflows in einem Medienunternehmen übertragen. Ideen für Funktionen befinden sich in Anhang acht.

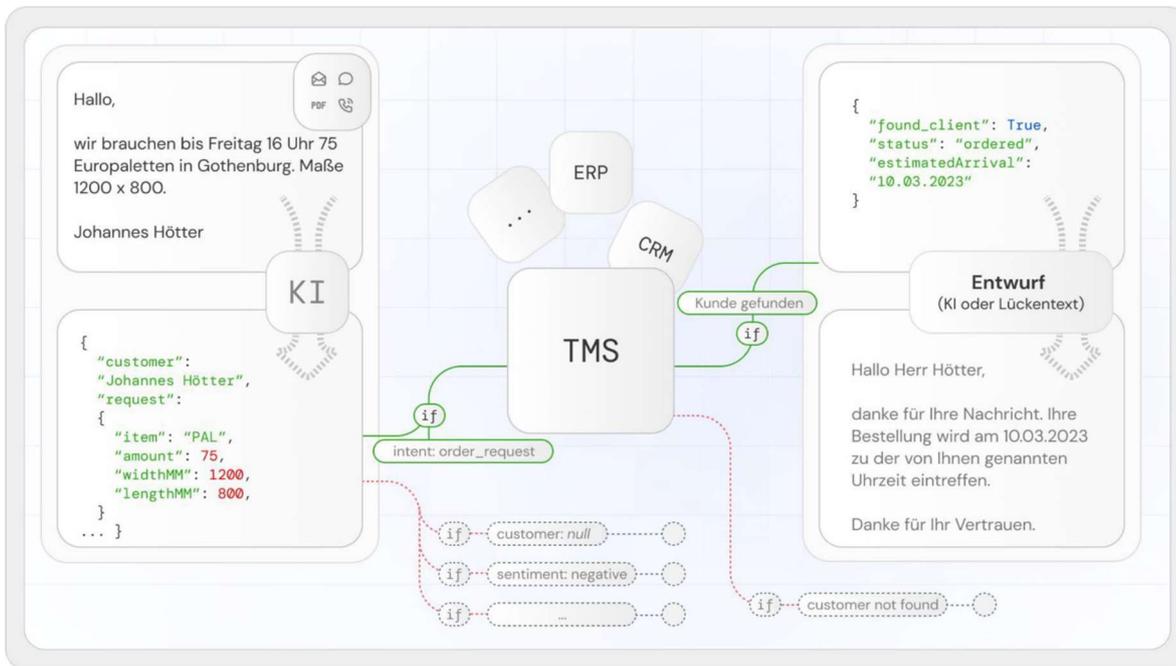


Abbildung 24: Visualisierung des Function Callings [97].

Da ChatGPT Anfragen und Antworten in Form von JavaScript-Objekten verarbeitet [94], muss für die Verbindung zum MAM eventuell ein Mapping zur Extensible Markup Language (XML) durchgeführt werden. Dafür stellt Node-RED einen eigenen XML-Node zur Verfügung (A35).

Benutzbarkeit der Schnittstellen (A38)

Zusätzlich zur lückenlosen Informationsvermittlung muss die API Anfragen und Antworten auch in einer gewissen Geschwindigkeit übergeben und empfangen können. OpenAI gibt sogenannte Rate Limits vor, die vom gesamten Medienunternehmen inklusive aller Nutzer eingehalten werden müssen. Ihr Zweck ist die Gleichberechtigung aller Nutzerinnen und Nutzer. In besonderen Fällen kann das Rate Limit erhöht werden. Die Rate Limits werden in Token pro Minute oder Anfragen pro Minute angegeben. Das Limit für Chatanwendungen ist beispielsweise 3500 Anfragen pro Minute [98].

7.3.3 METADATEN

Metadaten sind Daten über Daten. Sie beschreiben die Essenzen im technischen sowie im inhaltlichen Sinne. Durch die Beschreibungen ist es möglich, die Essenzen im Archiv durch eine Stichwortsuche zu finden. Die Daten müssen jedoch in einer Art Struktur ausgebaut sein, die für jede Essenz gleich ist. Ein Metadatenmodell besteht aus mehreren Elementen wie beispielsweise Name, Aufnahmedatum und Länge, denen jeweils ein Wert zugeordnet wird. Anhand dieser Struktur können Abfragen effizienter bearbeitet werden [99, S. 4].

Das gebräuchlichste Metadatenmodell, das Metadata Dictionary der Society of Motion Pictures and Television Engineers (SMPTE), besteht aus zwei Teilen. Der eine ist ein eindeutiges, zweiteiliges 16 bit-Label, dem Universal Material Identifier (UMID). Es ist aus Zahlen aufgebaut und daher

sprachunabhängig [99, S. 4]. Der zweite Teil der Referenzierung ist eine Bibliothek aus Symbolen, die die Essenz beschreiben. Jedem Symbol wird ein Wert zugeordnet. Die Darstellung erfolgt in XML [99, S. 4].

UMID

Der UMID wird zur Identifikation des Materials verwendet. MAM und CMS beziehen sich auf dasselbe Label, das für jede Essenz auf der Welt einmalig und eindeutig ist. Außerdem enthält der UMID Informationen über die Entstehung des Materials [100, S. 4]. Der UMID wird entweder bei der Entstehung der Essenz in ihren Metadaten abgespeichert oder beim Eintritt in den Speicher des Medienunternehmens erzeugt. Dabei ist es wichtig, dass der UMID standort- und systemunabhängig ist [100, S. 6]. Der UMID wird innerhalb des Speichers den Metadaten zugeordnet. Durch den „Indirection Mechanism“, also das MAM, referenzieren die Metadaten die Essenz (siehe Abbildung 25).

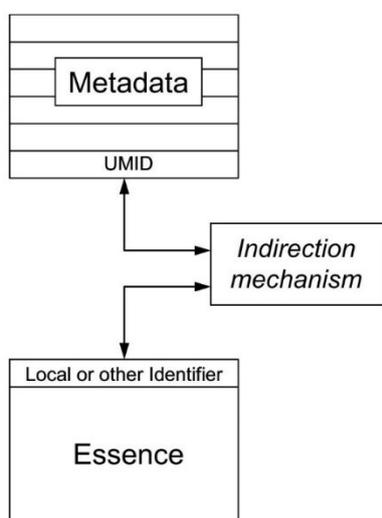


Abbildung 25: Verbindung von UMID zu Essenz [100, S. 7].

Klassenstruktur

Das SMPTE Metadata Dictionary ist in sieben Standard-Klassen unterteilt (Klassen eins bis sieben). Diese bilden hauptsächlich die technischen Metadaten zu einer Essenz ab. Dazu kommen drei Klassen, die von den Benutzerinnen und Benutzern nach Bedarf erweitert werden können (Klassen 13 bis 15) (siehe Abbildung 26). Der Katalog umfasst mittlerweile ungefähr 127 Klassen [23, S. 103].

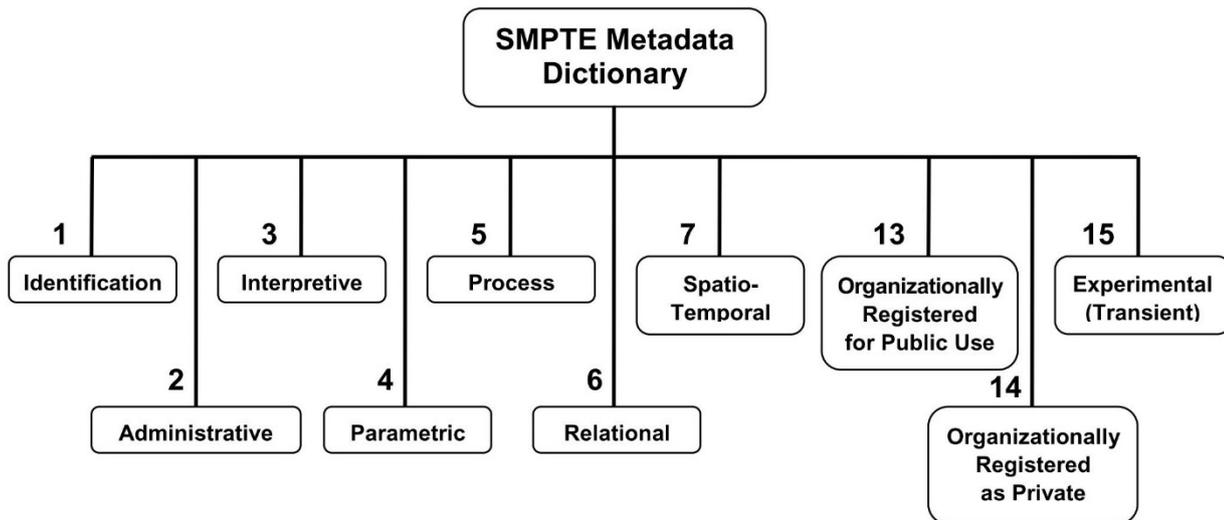


Abbildung 26: SMPTE Metadata Dictionary Klassenstruktur [99, S. 7].

Klasse 13 besteht aus individuellen Metadatenelementen, die von dritten Personen registriert wurden und daher für sie reserviert sind. Die Elemente werden ebenfalls von diesen Personen verwaltet und werden dadurch unabhängig von den Klassen eins bis sieben behandelt. Angemeldete Subklassen und Elemente in der Klasse 13 werden veröffentlicht und von der SMPTE anerkannt [99, S. 10]. In Klasse 14 können ebenfalls dritte Personen neue Subklassen und Metadatenelemente einführen, die unteren Metadatenelemente werden jedoch nicht veröffentlicht. Trotzdem müssen sie noch bei der SMPTE eingereicht und bestätigt werden [99, S. 10]. Um Einträge in beiden Klassen hinzuzufügen, müssen diese die Vorgaben des Metadata Dictionary erfüllen und von einer spezifischen Organisation eingereicht werden [99, S. 18].

Verbindung zwischen ChatGPT und dem Metadatenmodell

Um einen verlustfreien Austausch zwischen ChatGPT und dem Metadatenmodell zu gewährleisten, muss die Ausgabe von ChatGPT an das Metadatenmodell angepasst werden bzw. auch umgekehrt. Die Anpassung der Ausgabe von ChatGPT erfolgt über übergebene Funktionen beim Function Calling. Eine Anpassung des Metadatenmodells kann über zusätzliche Subklassen und Elemente in den Klassen 13 und 14 erreicht werden. Eine solche Anpassung in beide Richtungen ist für den Use Case Extraktion von Schlagwörtern bzw. beschreibenden Metadaten besonders wichtig. Eine beispielhafte Antwort von ChatGPT wird im JSON-Format ausgegeben. Ein möglicher Prompt für diesen Use Case ist „Nenne alle Schlagwörter aus dem Text“ sowie ein zu bearbeitender Text.

```

"choices": {
  "finish_reason": "stop",
  "index": 0,
  "message": {
    "content": "Künstliche Intelligenz, Medienunternehmen,
    ChatGPT, Journalismus, Systemlandschaft, Einführung",
    "role": "assistant"
  }
},
"created": 1677664795,
"id": "chatcmpl-7QyqpwdfhqwajicIEznoc6Q47XAYW",
"model": "gpt-3.5-turbo-0613",
"object": "chat.completion",
"usage": {
  "completion_tokens": 17,
  "prompt_tokens": 57,
  "total_tokens": 74
}

```

Abbildung 27: Antwort von ChatGPT in JSON [94].

Aus dieser Antwort ist hauptsächlich der Tag „Content“ (grüne Markierung in Abbildung 27) interessant. Er wird über ein Mapping in ein Beschreibungselement des Metadatenmodells überführt. Ein solches Beschreibungselement kann entweder aus einer der Standardklassen gezogen oder als ein eigenes in den Klassen 13 und 14 definiert werden. Ein weiteres wichtiges Element, das zum Metadatenmodell hinzugefügt werden muss, ist eine Abfrage, ob die Essenz von einem Menschen oder einer KI erzeugt wurde. Im Use Case Textgeneration muss also diesem Element ein wahrer Wert zugeordnet werden. Es ist jedoch auch möglich, jedes andere Metadatenmodell, außer das SMPTE Metadata Dictionary, für zusätzliche beschreibende Metadaten anzupassen.

7.3.4 GUI

Im Folgenden werden Vorschläge für mögliche GUIs für die jeweiligen Nutzungsgruppen (siehe Abschnitt 6.6), die keinen technischen Hintergrund haben, beschrieben. Die Personen müssen ebenfalls einen Zugang zu ChatGPT erhalten, auch wenn keine Programmierkenntnisse vorhanden sind. Neben jeder produktivarbeitenden GUI wird außerdem der aktuelle Rundown aus dem Redaktionssystem angezeigt (A37). Die Anforderungen an die jeweiligen GUIs befinden sich in Anhang neun. Zeichnungen befinden sich auf dem Datenträger in Kapitel sieben.

A: Text

Diese grafische Benutzeroberfläche ist für das Verfassen von Beiträgen mit gleichzeitiger Archivsuche geeignet, sodass Informationen oder Inhalte aus Archivbeiträgen direkt aus den Suchergebnissen entnommen werden können. Neben der Schreibfläche befindet sich ein Chat-Interface zu ChatGPT. Damit können ChatGPT mittels Prompts Fragen oder Aufgaben gestellt werden. Ein Chatverlauf zeigt die bereits bearbeiteten Prompts. Die Eingaben an ChatGPT können entweder über Drag and Drop aus dem Text getätigt werden oder eine weitere Essenz mit ihrer ID dem Prompt angehängt werden. Gleichzeitig generiert ChatGPT automatisch Anmerkungen, die neben dem geschriebenen Text auftauchen. Diese können über Funktionsschaltflächen bearbeitet, aktualisiert oder gelöscht werden.

Eine weitere Funktionalität in der GUI ist die Feedbackfunktion. Gibt ChatGPT eine fehlerhafte oder anstößige Antwort aus, kann dies über ein Feedback-Interface gemeldet werden. Die betreffende Antwort muss dafür nur in ein Texteingabefeld gezogen werden. Eine Beschreibung des Fehlers ist hier erwünscht.

B: Archiv I

Die grafische Benutzeroberfläche im Archiv enthält weniger Eingabeflächen, da die hauptsächliche Aufgabe hier nur die Kontrolle von automatisch erzeugten Metadaten für Texte ist. Es wird also der betreffende Text abgebildet, sowie die kategorisierten Metadaten, die ChatGPT bereits aus dem Text gezogen hat. Über Bearbeitungsschaltflächen kann die Archivarin bzw. der Archivar die Metadaten anpassen. Die Feedbackfunktion hilft bei der Meldung grob falscher Metadaten. Fehlerhafte Einträge werden über die Drag and Drop-Funktion in die Texteingabe des Feedbackfeldes gezogen und können mit erläuternden Worten versehen werden.

7.3.5 KRYPTOGRAFISCHE VERFAHREN

Um Missbrauch mit ChatGPT zu vermeiden und einen generellen Schutz gegen Angreifer zu bieten, muss vor der Verarbeitung von Eingaben festgestellt werden, ob der Input von der Automation oder einem autorisierten Menschen innerhalb des Medienunternehmens kommt. Damit kann auch das Zugangskonzept (siehe Abschnitt 6.6) umgesetzt werden, denn jede Zugangsgruppe hat ihren eigenen Schlüssel, mit dem sie sich bei der Automation wiederum authentifizieren muss und dann auf die jeweiligen Funktionen zugreifen kann. Zunächst muss sich jede Nutzerin bzw. jeder Nutzer mit ihrer oder seiner Rolle in der jeweiligen GUI authentifizieren. Es ist jedoch möglich, für Zugangsgruppen, die mehrere GUIs benötigen, einen gemeinsamen Zugang für diese Gruppe anzulegen. Die Automation wiederum authentifiziert jede Message gegenüber ChatGPT, damit auch hier kein Missbrauch entstehen kann.

Verschlüsselung in Node-RED

Die Verschlüsselung in Node-RED ist über den hmac-Node möglich (siehe Abbildung 28). Dieser unterstützt den Einsatz von Hash-Funktionen. Sie sind nicht zur Verschlüsselung von Nachrichten gedacht, sondern als Beweis, dass die Nachricht auf dem Übertragungsweg nicht verändert oder

ersetzt wurde. Kryptografische Hashfunktionen sind Einwegfunktionen. Die Art von Funktionen ist leicht zu berechnen, aber nur mit unverhältnismäßig großem Rechenaufwand invertierbar [101, S. 2]. Der Hashfunktion $h(x)$ wird also der zu bestätigende Text M übergeben. Die Funktion wandelt den Text in einen kürzeren Hash $h(M)$ um. Dieser Hash hat keinen augenscheinlichen Bezug zum Text M . Wenn jedoch der Absender den Hash berechnet und dem Text beifügt und der Empfänger diesen mit dem empfangenen Text ebenfalls berechnet und vergleicht, kann eine Manipulation des Textes ausgeschlossen werden, solange der Angreifer nicht auch eine Hashfunktion berechnet und dem Text anhängt [101, S. 3]. Eine Maßnahme, um dies zu umgehen, ist der hmac. Hier wird dem Text M noch ein geheimer symmetrischer Schlüssel k angehängt, bevor dieser durch die Hashfunktion bearbeitet wird:

$$h_k(M) = h(M||k)^{10}$$

Der hmac-Node unterstützt die Hashfunktionen MD5 sowie SHA1 bis SHA512. Mit den Rechenressourcen der heutigen Zeit wird für die höchste Sicherheitsstufe SHA512 empfohlen, da beispielsweise MD5 bereits mit überschaubarem Rechenaufwand umgangen wurde [101, S. 13].



Abbildung 28: hmac-Node in Node-RED [102].

¹⁰ $M||k$ bedeutet, dass M und k konkateniert werden [101, S. 18].

7.4 PROMPT ENGINEERING

Da es sich bei den Modellen von OpenAI um Sprachmodelle handelt, müssen jegliche Aufgabenstellungen, auch Prompt genannt, daher in natürlicher Sprache formuliert werden. Angepasste Prompts geben vor, in welchem Bereich des mehrdimensionalen Vektorraums die Suche nach geeigneten Inhalten durchgeführt werden soll [103, S. 5]. Wird der Prompt „Schreibe eine witzige Geschichte“ eingegeben, werden mögliche Wörter für eine Geschichte im grünen Bereich des Vektorraums gesucht. Wörter für eine traurige Geschichte werden dem roten Bereich entnommen (siehe Abbildung 29).

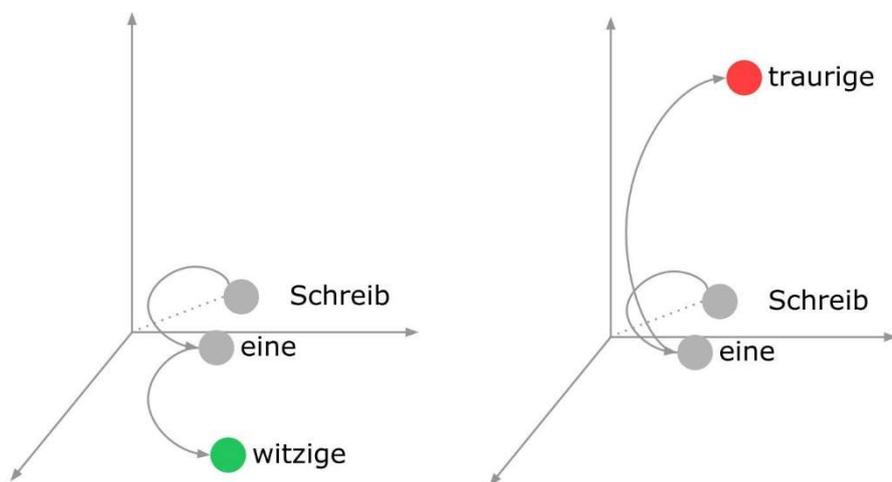


Abbildung 29: Beispiel für Prompt Engineering [103, S. 8, 103, S. 7].

Die Prompts wiederum werden als Key-Value-Paar in einer JSON-Message zu ChatGPT transportiert [94]. Die Prompts können entweder von einem Menschen formuliert worden sein oder einen vorgefertigten Prompt für wiederkehrende Aufgaben darstellen. Mit kontinuierlichem Einsatz von ChatGPT werden die effizientesten Prompts ermittelt [104, S. 4].

Die Idee für die Verbesserung von Prompts ist, die Instruktionen so verständlich wie möglich einzugeben. Mehr Details, Beispiele oder eine schrittweise Beschreibung der Aufgabe können das Ergebnis verbessern. In Anlehnung an die schrittweise Bearbeitung der Aufgabe kann eine Aufgabe auch direkt in mehrere Unteraufgaben und damit mehrere Prompts geteilt werden. Referenztexte helfen ebenfalls bei der Bearbeitung der Aufgabe. Ein möglicher Prompt kann sein: „Überarbeite diesen Artikel sprachlich, sodass der Stil ähnlich wie bei einem vorangegangenen Artikel wird“. Ein spezifischer vorangegangener Artikel wird dann bei der Aufgabenstellung referenziert. Ähnlich wie bei der menschlichen Problemlösung kann es ChatGPT auch helfen, den Prompts Formulierungen wie „Lass dir Zeit mit der Antwort“ oder „Lege deine Lösung schrittweise dar“ hinzuzufügen. Dadurch wird ChatGPT angehalten, gründlich die Lösung zu „überdenken“ und gibt einen Lösungsweg aus, der für den Menschen verständlich ist [105]. Beispiele für Prompts befinden sich in Anhang zehn.

Eine weitere Möglichkeit bietet die iterative Eingabe. Dabei wird die Aufgabe in mehrere Anforderungen geteilt. Zunächst wird ein Prompt eingegeben, der die unterste Anforderung bedient.

Das KI-Modell gibt eine Antwort für diesen Prompt aus. Ein weiterer Prompt fügt die zweite Anforderung hinzu und ChatGPT beantwortet ihn ebenfalls. Durch diese iterative Vorgehensweise entwickelt sich schließlich ein zufriedenstellendes Produkt. Bei Anwendungen, bei denen die Aufgabenstellung gleichbleibt, sich aber nur die Parameter ändern (Archivsuche), ist es über die Automation auch möglich, nur Stichwörter einzugeben und nicht den gesamten Prompt.

*Stufen der Nutzer*innenrückmeldung*

Die Beantwortung von Nachrichten von Nutzerinnen und Nutzern soll ebenfalls über den Chatbot erfolgen. Dieser operiert in fünf Stufen der Autonomie:

Das unterste Level beschreibt eine schlichte Sentiment Analysis. Unhöfliche, negative Nachrichten werden von ChatGPT erkannt und von der Automation markiert. Hier beantworten immer noch Menschen die Nachrichten, denn sie können jedoch leichter zwischen konstruktiver Kritik und Hass unterscheiden. Im nächsten Level wird nicht nur die Stimmung der Nachricht erkannt, sondern auch das Thema. Dadurch können die Nachrichten bereits vorsortiert und den richtigen menschlichen Mitarbeiterinnen und Mitarbeitern zugeordnet werden. Außerdem können Statistiken geführt werden, wie oft ein Thema von Nutzerinnen und Nutzern gelobt oder kritisiert wird. Im dritten Schritt werden die Informationen aus der Nachricht in strukturierte Daten überführt, wobei die Variablen von Menschen definiert werden. Die vierte Stufe umfasst als nächsten Schritt eine Beantwortung der Nachricht. Die Antwort beruht auf den vorher aus der Nachricht extrahierten Informationen, Domänenwissen und einer Antwortvorlage, die zuvor von einem Menschen verfasst wurde. ChatGPT setzt also lediglich neue Informationen in einen Lückentext ein. Im höchsten Level erfolgt die Beantwortung der Nachrichten ohne Hilfe eines Menschen. ChatGPT generiert die Antwort selbst, sodass auch ein gewisses Maß an Individualisierung für jede Nutzerin bzw. Nutzer gegeben ist [106]. Auf jeder Stufe werden die ausgehenden Nachrichten zumindest stichprobenartig von einem Menschen kontrolliert.

Domänenwissen

Damit ChatGPT auch Fragen beantworten kann, die das Medienunternehmen im Speziellen betreffen (z.B. die Postadresse), muss dem KI-Modell Domänenwissen gegeben werden. Nur mit seinem trainierten Weltwissen kann es solche Fragen nicht beantworten. Dieses Wissen wird ChatGPT durch Prompt Engineering übergeben. Man kann der API sogar eine Datei mit den Informationen übergeben, sodass diese nicht Teil des textuellen Prompts sein müssen [107].

7.5 ZWISCHENFAZIT

Nach der Beschreibung der nicht-technischen Konzepte sowie der technischen Lösung folgt eine Zusammenfassung des vorgestellten Systems. Es werden zunächst Vision und Strategie für ein solches Projekt beschrieben. Sie bilden die Grundlage, auf der alle folgenden Überlegungen und Tätigkeiten aufbauen. ChatGPT soll als Mittel der Effizienzsteigerung im Newsworkflow dienen und den Menschen repetitive Arbeiten abnehmen. ChatGPT eignet sich grundlegend für die gestellten Use Cases. Mit Techniken wie Prompt Engineering kann das Ergebnis verbessert werden. Für die Erfüllung der rechtlich-ethischen und organisatorischen Anforderungen sorgt eine Vielzahl an Unterkonzepten. Diese Konzepte werden durch die technischen Lösungen in Kapitel sieben umgesetzt. Außerdem werden in Abschnitt 6.3 die Workflows mit KI-Unterstützung beschrieben, die sich aus jenen in Kapitel vier entwickelt haben. Die technische Lösung nennt einen neuen Akteur in der Systemlandschaft des Medienunternehmens: Die workflowbasierte Automation Node-RED. Mithilfe dieser Software kann der Informationsaustausch zwischen Nutzer, ChatGPT und CMS sicher gestaltet werden. Das Kapitel enthält ein Blockschaltbild sowie Darstellungen der Informationsflüsse für verschiedene Szenarien. Außerdem werden die Schnittstellen zwischen Maschinen und zwischen Mensch und Maschine näher beleuchtet. Dadurch wird ein Austausch möglich, der den Anwenderinnen und Anwendern in ihrem Arbeitsalltag ein zusätzliches Werkzeug an die Hand gibt, das eine potenzielle Effizienzsteigerung bedeutet.

Trotzdem zeigen dieses und das vorherige Kapitel, dass ChatGPT nur etwas mehr als die Hälfte der funktionalen Use Cases abdeckt (siehe Anforderungsmatrix in Anhang eins). ChatGPT hat keinen Zugang zu den Medien Bild und Ton, da es nur auf die Verarbeitung von Texten ausgelegt ist. Zusätzlich ist es auch nicht möglich, ChatGPT durch eigenes Finetuning an den Stil des Medienunternehmens anzupassen. Für die Anwendung im Archiv wird ein Knowledge Graph mit Embeddings empfohlen. Auch diese Funktionalität ist für ChatGPT nicht gegeben. Bei ChatGPT handelt es sich eher um ein universelles Sprachmodell mit Weltwissen, das besonders auf Chatanwendungen spezialisiert ist. Entsprechend ist es auf keinen der Use Cases besonders spezialisiert.

8 WEITERFÜHRENDE ÜBERLEGUNGEN

Dieses Kapitel stellt über ChatGPT hinaus Lösungsansätze dar, um die Anforderungen zu erfüllen, die ChatGPT nicht abdecken kann. Dabei handelt es sich in erster Linie um Bild- und Tonverarbeitung sowie Embeddings und Finetuning. Zur Erfüllung dieser Use Cases oder Verbesserung der Qualität bei bereits umgesetzten Use Cases existieren weitere Modelle von OpenAI. Dabei handelt es sich um das Bildverarbeitungsmodell DALL-E und das Audiomodell Whisper. Dazu kommt die Gruppe der InstructGPT-Modelle.

8.1 WEITERES ANGEBOT VON OPENAI

OpenAPI besitzt zusätzlich zur GPT-Familie weitere Modelle:

DALL-E und Whisper

OpenAI bietet die Modelle Whisper und DALL-E zur Bild- und Tonverarbeitung. Whisper ist in der Lage, gesprochene und aufgezeichnete Sprache in Text zu transkribieren. Auch eine Übersetzungsfunktionalität wird angeboten. DALL-E ist ein KI-Modell, das realistische Bilder aus wenigen Texteingaben erstellt. Die Bilder können in verschiedenen Auflösungen ausgegeben werden. Dazu kommt die Möglichkeit, bereits existierende Bilder zu bearbeiten und mehrere Versionen zu erstellen [81]. Beide Modelle befinden sich zum Zeitpunkt der Veröffentlichung der Thesis noch in einer Betaphase [81, 108].

InstructGPT

Die Gruppe der InstructGPT-Modelle ist darauf spezialisiert, einzelne Aufgaben besonders gut zu lösen. Jedes Modell ist dabei für eine andere Aufgabe optimiert. Innerhalb der Gruppe gibt es vier Modelle: davinci, curie, babbage und ada. Es ist die Möglichkeit zum individuellen Anpassen der Modelle gegeben. Eine dedizierte Instanz, die individuell nachtrainiert werden kann, kann ebenfalls über eine API erreicht werden. Die ausgewählten Modelle sollen dann mit den hauseigenen (Medien)-Daten trainiert werden. Die vier Modelle bieten außerdem eine Funktionalität, Texte inhaltlich miteinander in Beziehung zu setzen. Sogenannte Embeddings messen die Ähnlichkeit zwischen Textausschnitten. Mögliche Use Cases sind Suche, Clustern, Empfehlungen, Diversitätsmessung und Klassifizierung [81].

8.2 ERGÄNZUNG ZU DEN WORKFLOWS

Durch den Einsatz zusätzlicher Modelle verändern sich auch die Workflows an einigen Stellen. Neue Funktionalitäten können noch mehr Aufgaben an die KI übergeben. Menschen führen weniger wiederkehrende Arbeiten durch und können sich dadurch auf die kreativen Aufgaben konzentrieren.

Ingest

Beim Ingest werden ebenfalls ausschließlich Metadaten aus Texten gezogen und in den beschreibenden Metadaten abgespeichert. Zusätzlich werden jedoch Audioaufnahmen sowie die Audiospuren von Videos verschlagwortet. Diese werden zu Text transkribiert, sodass auch hier beschreibende Metadaten aus dem Text gezogen werden. Die Transkription wird ebenfalls mit Referenz auf die Originalessenz abgespeichert. Die Anzahl der Essenzen, die mit KI verschlagwortet werden, erhöht sich also durch den zusätzlichen Einsatz eines KI-Modells, das Audioaufnahmen verarbeiten kann.

Archivsuche

Um bei der zukünftigen Archivsuche ein umfassenderes Sucherlebnis zu bieten, werden alle Essenzen im Archiv von einem auf Embeddings trainierten KI-Modell in einem mehrdimensionalen Vektorraum verortet. Jede Dimension des Vektorraums entspricht einer Eigenschaft. Diese Eigenschaft wird wiederum mit einer Skala bewertet. Daraus ergibt sich eine Zahl innerhalb des Vektors, der einem Wort zugeordnet wird. Dabei sind dann inhaltlich ähnliche Wörter in einem geringeren Abstand zueinander aufgetragen (siehe Abschnitt 2.1.2). Mit den Beiträgen wird genauso verfahren. Sie werden anhand ihrer inhaltlichen Übereinstimmung in diesem Raum positioniert. Die Koordinaten im Vektorraum werden in den Metadaten des Inhalts gespeichert. Die Auswahl der Eigenschaften trifft das KI-Modell größtenteils selbst und führt auch die Bewertung der Eigenschaften eigenständig durch. Um diesen Vektorraum aufzubauen, benötigt das KI-Modell also Zugriff auf das Archiv des Medienunternehmens. Anhand dieses Vektorraums werden bei der Archivsuche nicht nur Beiträge angezeigt, die der Stichwortsuche entsprechen. Es werden Beiträge aus dem räumlichen Umfeld des Stichwortes angezeigt, um Synonyme miteinzuschließen und Homonyme gezielt auszuschließen. Dadurch gestaltet sich die Archivsuche effizienter.

Unterstützung in Grafik

Für Symbolbilder oder Material für den Hintersetzer werden Bilder von einem KI-Modell generiert, um eine Alternative zur Archivsuche zu bieten. Dadurch werden individuelle Bilder nach textuellen Eingaben erzeugt, die es in dieser Form noch nicht gibt. Existieren bestehende Bilder im Archiv in einem ungeeigneten Bild-Seiten-Verhältnis oder ungeeigneter Auflösung, schafft ein KI-Modell hier ebenfalls Abhilfe. Das neu entstandene Bild wird gemeinsam mit dem Prompt, der es erzeugt hat, abgespeichert.

Personalisierte Ausspielung

Die Analyse von Nutzungsdaten ist für Journalistinnen und Journalisten von Vorteil, da es den Anklang ihrer Erzeugnisse im Publikum widerspiegelt. Daraus werden Erfahrungen für kommende Beiträge gezogen. Es können jedoch nur Aufnahmen der Vergangenheit durch die Darstellung der Nutzungszahlen aufgezeigt werden. Da die Stärken der meisten Journalistinnen und Journalisten jedoch im Bereich der Sprache und nicht in dem der Statistik liegt, wird der Rückkanal selten genutzt [74, S. 35].

Die Daten werden automatisch von eingeloggten Nutzerinnen und Nutzern erhoben, sofern diese der Erhebung zustimmen. Die Daten werden ebenfalls im Speicher des Medienunternehmens angesammelt. Die Menge der Daten aus der Nutzung des Mediums wird für die personalisierte Ausspielung von Inhalten von Nutzen genutzt. Die drei Stufen der Individualisierung sind Empfehlungen, Personalisierung der Präsentation und Personalisierung der Geschichte [74, S. 56].

Dem Nutzungsverhalten werden Daten, die auf die Eigenschaften und Interessen der Nutzerin bzw. des Nutzers schließen lassen, entzogen (siehe Abschnitt 5.2.3). Sie werden gespeichert und anhand dieser Daten neue Inhalte empfohlen. Möglich ist auch, dass die Beiträge mit einem anderen Titelbild oder Schlagzeile präsentiert werden, je nach Präferenz der Nutzerin bzw. des Nutzers. Bei mehrteiligen Serien kann sich auch die angebotene Reihenfolge der Episoden ändern. Sobald die Nutzerin bzw. der Nutzer einen neuen Beitrag konsumiert, werden die Daten aktualisiert und neue Vorschläge generiert. Dabei handelt es sich um einen iterativen Prozess. Bei der personalisierten Ausspielung werden der Nutzerin bzw. dem Nutzer dann auf der Startseite Beiträge vorgeschlagen, die inhaltlich zu den bereits konsumierten passen, jedoch noch nicht von der Nutzerin bzw. dem Nutzer gesehen wurden.

8.3 ERGÄNZUNGEN ZU DEN KONZEPTEN

Dieses Kapitel definiert keine eigenständigen Konzepte. Es lehnt sich an die Überlegungen in Kapitel sechs an. Durch den zusätzlichen Einsatz der InstructGPT-Modelle sowie Whisper und DALL-E, ergeben sich jedoch weitere Rollen, Profile, Schulungsbedarfe, Sicherheitsaspekte und Kosten, die die Konzepte vervollständigen.

8.3.1 ROLLENKONZEPT

Im Rollenkonzept kommt es zu Änderungen im Rollenverständnis zweier Berufsgruppen:

*Grafiker*in*

Die Erstellung von Grafiken ist in Rundfunkanstalten typischerweise in die 2D- und 3D-Grafik aufgeteilt. 2D-Grafiken werden über den Hintersetzer im Studio ausgespielt. Die Bildrecherche im hauseigenen Archiv gehört ebenfalls zu den Aufgaben. 3D-Grafiken sind individuelle Erklärgrafiken sowie Bauchbinden-Vorlagen. Durch die Unterstützung durch das bildgenerierende KI-Modell DALL-E sowie die verbesserte Archivsuche kann bei der Erstellung der 2D-Grafiken Zeit gespart werden.

*Mediengestalter*in*

Im Bereich des Schnitts kommt ebenfalls ein bildverarbeitendes KI-Modell zur Bildgeneration zum Einsatz. Mediengestalterinnen und Mediengestalter profitieren durch Arbeitserleichterungen aufgrund der KI, denn diese ersetzt die Suche nach einem geeigneten Symbolbild. Das KI-Modell generiert ein neues Bild anhand von wenigen textuellen Vorgaben.

8.3.2 ZUGANGSKONZEPT

Durch die Vielzahl an verschiedenen Modellen werden mehr Zugangsprofile geschaffen, die wiederum den Rollen zugeordnet sind. In Anhang vier sind die Rollen aus Abschnitt 8.3.1 mit den Zugangsprofilen verbunden.

B: Archiv II

Die InstructGPT-Modelle, optimiert für Embeddings, sorgen für die Einordnung von Beiträgen in den mehrdimensionalen Vektorraum, denn die Beiträge im Archiv sind ebenfalls untereinander vernetzt, um bei der Recherche gezielt semantisch ähnliche Beiträge zu finden. Das Profil ist produktiv, da über die GUI ebenfalls manuelle Annotationen zu Essenzen eingegeben werden können. Archivarinnen und Archivare arbeiten mit diesem Profil.

F: Training

Das Trainingsprofil erlaubt den Zugang zu den Trainingsdaten und Manipulation der Gewichte durch Nachtraining. Das Profil ist nicht produktiv, die Nutzerinnen und Nutzer erzeugen aktiv keine Beiträge in diesem Nutzungsprofil. Data-Scientist*innen und die KI-Assistenz benutzen dieses Profil.

G: Audio

Das Profil B zielt auf alle Personen ab, die in ihrem Arbeitsalltag Audio in Text umwandeln. Das Modell Whisper von OpenAI erlaubt die Transkription von gesprochener Sprache zu Text. Die Rollen, die Teil dieses Profils sind, sind Journalist*in und Archivar*in.

H: Grafik

Das Grafikprofil ist für Personen geeignet, die Bilder suchen, auswählen oder generieren wollen. Es enthält ebenfalls einen Zugang zum Archiv für die Bildrecherche. Hierfür wird das Modell DALL-E von OpenAI verwendet. Die Rollen innerhalb dieses Profils sind Grafiker*in, Mediengestalter*in und Journalist*in. Der Zugang zu den Grafik-Funktionalitäten erfolgt über eine grafische Benutzungsoberfläche.

I: Distributiv

Auf Basis der Nutzungsgewohnheiten wird für jede Nutzerin bzw. für jeden Nutzer ein Interessenprofil angelegt. Ausgehend davon, werden Empfehlungen für Nutzerinnen und Nutzer ausgesprochen, die im Vektorraum räumlich nah zu den Interessen der Nutzerin bzw. des Nutzers sind. Das Nutzungsprofil ist ein administratives Profil, das den sonst automatischen Workflow der Personalisierung betreut. Der Zugang sieht also keine GUI vor. Rollen, die Zugang zu diesem Profil haben, sind Data-Scientist*in und Systemintegrator*in.

8.3.3 SCHULUNGSKONZEPT

Dem Schulungskonzept werden drei neue Bausteine hinzugefügt, nämlich die Einführung in auditive Use Cases, die Einführung in Archivworkflows sowie Merkmale von Datenqualität. Die ersten beiden Module erreichen die gestalterische Zielgruppe. Hier werden den Teilnehmerinnen und Teilnehmern die Eigenschaften der Use Cases im Bereich Audio und Video nähergebracht. Die Anwendung und die Möglichkeiten stehen hier im Vordergrund. Bei der Einführung in Archivworkflows wird ergänzt, dass zusätzlich eine Transkription bei Audioaufnahmen und Videos stattfindet. Dadurch wird eine größere Menge an Essenzen präziser verschlagwortet. Außerdem wird für jede der verschlagworteten Essenzen ein Ort im mehrdimensionalen Vektorraum bestimmt, der in den Metadaten gespeichert wird und für eine inhaltliche Verbindung zwischen Essenzen sorgt. Im Modul Datenqualität wird die Wichtigkeit der Datenqualität für KI-Modelle behandelt. Dies ist besonders für die Archivarin bzw. den Archivar von Interesse, da die Person die Schaffung von neuem Trainingsmaterial überwacht. Es wird auf Metriken zur Messung der Datenqualität und Effekte wie Fehler und Bias eingegangen.

8.3.4 KONZEPT DER INFORMATION AN NUTZER*INNEN

Durch die personalisierte Ausspielung von Inhalten tritt das Publikum nun direkt mit der KI in Kontakt. Das Medienunternehmen ist damit verpflichtet, seinen Nutzerinnen und Nutzern Strukturen zur Verfügung zu stellen, die einen sicheren Umgang mit KI gewährleisten.

Erklärungen in einfacher Sprache (A4)

Nach dem Gesetz über digitale Dienste und dem Medienstaatsvertrag sind Anbieter von Plattformen, die Empfehlungssysteme verwenden, verpflichtet, in klarer und verständlicher Sprache eine Erklärung zur Funktionsweise abzugeben. Darin werden die wichtigsten Parameter und alle Möglichkeiten, diese zu beeinflussen, beschrieben [63, Art. 27 Abs. 1]. Eine Auswahl wird in Abschnitt 5.2.3 aufgezeigt. Der Nutzerin bzw. dem Nutzer muss deutlich werden, nach welchen Kriterien die Inhalte auf der Startseite angeordnet wurden [30, § 93]. Die Erklärungen werden ebenfalls durch strukturierte Tests mit Probandinnen und Probanden evaluiert (siehe Abschnitt 5.1.4). Außerdem ist wichtig, dass Nutzerinnen und Nutzer der Erhebung von Nutzungsdaten und einer folgenden Personalisierung vollständig widersprechen können.

Kontroll- und Beschwerdestelle (A3)

Zusätzlich zu den Erklärungen muss laut Telemediengesetz und Medienstaatsvertrag eine Beschwerdestelle für rechtswidrige Inhalte eingerichtet werden (siehe Abschnitt 5.1.7). Diese wird der Nutzerkommunikation im Medienunternehmen angegliedert. Wichtig ist hier, die prompte Bearbeitung jeglicher Beschwerden, da sich Inhalte im Internet mit hoher Geschwindigkeit verbreiten. Handelt es sich um ein öffentlich-rechtliches Medium, kann ebenfalls an die gemeinsame Schlichtungsstelle der Landesmedienanstalten verwiesen werden [30, § 99]. Ein privates Medienunternehmen hält eine eigene Beschwerdestelle vor.

8.3.5 REGLEMENTIERTER ZUGANG ZU INHALTEN

Um die Einhaltung von Jugendschutz und publizistischen Grundsätzen zu gewährleisten, ist es von Zeit zu Zeit nötig, den Zugang zu ausgewählten Inhalten bei der personalisierten Ausspielung zu beschränken. Dies ist ein Grundsatz, der in das Design der Workflows mit aufgenommen werden muss. Handelt es sich um User-Generated Content trägt das Medienunternehmen nach bestem Wissen und Gewissen die Verantwortung, wenn die Inhalte auf der eigenen Plattform veröffentlicht werden. Auch hier kann eine Sentiment Analysis die Inhalte filtern.

Inhalte, die ein Übermaß an Gewalt, Brutalität und Leid zeigen oder beschreiben, fallen unter den Jugendschutz (A2). Sie werden durch eine Kombination aus Sentiment Analysis und menschlicher Kontrolle herausgefiltert. Dabei werden Texte sowie Transkripte zur Verarbeitung herangezogen. Danach werden diese redaktionell geprüft und eventuell einer erwachsenen Zielgruppe nach vorheriger Altersprüfung zur Verfügung gestellt.

8.3.6 SICHERHEITSKONZEPT

Als vorbereitende Maßnahme werden die KI-Modelle, bei denen eine Individualisierung möglich ist, im Training gegen Manipulationen, wie in Abschnitt 5.1.3 aufgezeigt, gehärtet. Durch Training und Test mit unsauberen Daten kann die KI solche Angriffe ignorieren. Beispielsweise kann auch auf Word Embeddings trainiert werden. Damit sind verschiedene Schreibweisen eines Wortes gemeint. Wenn mehrere dieser Schreibweisen auftreten, erkennt das KI-Modell, dass sie das gleiche Wort meinen und legt nur eine Instanz des Worts im Vektorraum an [40, S. 16-17].

8.3.7 WIRTSCHAFTLICHE ANALYSE

Die Kosten für die Nutzung von KI-Modellen von OpenAI werden in Dollar pro 1000 Tokens berechnet. Der textuelle Input wird durch einen Algorithmus in Token umgerechnet. Die Rate variiert je nach Güte und Größe des KI-Modells. Außerdem unterscheiden sich die Raten auch beim Training und späteren Gebrauch des individualisierten KI-Modells.

Beispielhafte Rechnung für einen audiovisuellen Beitrag

Beim Text-to-Speech-Modell Whisper werden die Preise in Dollar pro transkribierter Minute berechnet. Für die Rechnung wird von einem dreißigsekündigen Nachrichtenbeitrag in Videoformat ausgegangen. Bei einem Schnittverhältnis von eins zu zehn geht man also von fünf Minuten Material aus, das im Schneiderraum ankommt. Würde das Material mit Whisper behandelt werden, um den gesprochenen Text aus dem Video zu extrahieren und dieses somit einfacher zu schneiden, wird folgende Rechnung angestellt:

$$\text{Kosten} = 5 \text{ min} * \$0,006 \frac{1}{\text{min}} = \$0,03$$

Die Transkription von fünf Minuten Material durch Whisper kostet also drei US-amerikanische Cents. Das entspricht zum Zeitpunkt der Erstellung der Thesis 2,8 europäischen Cents¹¹ [83].

Beispielhafte Rechnung für einen textuellen Artikel

Möchte der Journalist einen Artikel von einer Seite Text¹² mit dem Modell ada auf Sprachrichtigkeit und Stil prüfen lassen, müssen folgende Überlegungen angestellt werden. Laut Angaben von OpenAI entspricht eine Seite Text ungefähr 800 Token. Ein genaueres Ergebnis kann mit dem Tokenizer TikToken von OpenAI berechnet werden [81].

$$\text{Kosten} = 800 \text{ Token} * \$0,0016 \frac{1}{1000 \text{ Token}} = \$0,00128$$

Die Bearbeitung von 500 Wörtern Text kostet ungefähr 0,128 US-amerikanische Cents. Das entspricht 0,12 europäischen Cents [83].

¹¹ Die Währungskurse stammen aus der Zeit der Veröffentlichung der Thesis.

¹² hier 500 Wörter

Beispielhafte Rechnung für Training

Wird eine Seite Text als Trainingsmaterial zur Individualisierung des KI-Modells verwendet, erzeugt das ebenfalls Kosten. Bei der Verwendung des Modells *adav2*, das auf Embeddings spezialisiert ist, belaufen sich diese auf:

$$\text{Kosten} = 800 \text{ Token} * \$0,0004 \frac{1}{1000 \text{ Token}} = \$0,00032$$

Das Training von einer Seite Text kostet ungefähr 0,032 US-amerikanische Cents. Der Betrag in europäischen Cents lautet 0,03 [83].

8.4 ERGÄNZUNG ZUR TECHNISCHEN LÖSUNGSBESCHREIBUNG

Die Erweiterung der Landschaft um weitere KI-Modelle von OpenAI bringt auch weitere technische Überlegungen mit sich. Es wird somit für jeden Use Case das geeignete Modell eingesetzt. Die Informationsflüsse werden für weitere Use Cases wie z.B. die personalisierte Ausspielung angepasst. Außerdem ergibt sich durch die Option der Individualisierung eine neue Chance auf Antworten, die in den Stil des Medienunternehmens passen.

8.4.1 INFORMATIONENFLÜSSE

Die personalisierte Ausspielung ist ein neu geschaffener Informationsfluss. Er ist erst durch Modelle möglich, die Embeddings unterstützen. Zwar ändert sich beispielsweise auch der Ingestworkflow durch die neu eingeführten KI-Modelle, trotzdem war es auch schon davor möglich Essenzen zu verschlagworten, allerdings nur in geringerem Umfang. Daher hier nun die Beschreibung der personalisierten Ausspielung.

Personalisierte Ausspielung

Die personalisierte Ausspielung ist ein iterativer Fluss. Er läuft kontinuierlich und auch parallel für mehrere Nutzer ab, jedoch ist er nur für Ausspielplattformen geeignet, die ihre Präsentation ihren Nutzern individuell anpassen. Beginnend mit dem Eintritt einer Nutzerin bzw. eines Nutzers auf die Videoplattform oder das Onlinemagazin eines Medienunternehmens werden Nutzungsdaten über diese Person gesammelt. Für Nutzerinnen und Nutzer, über die noch keine Daten bekannt sind, wird zunächst das Collaborative Filtering angewandt. Werden mit zunehmender Nutzungszeit mehr Daten bekannt, werden andere Methoden verwendet. Die Daten werden extrahiert und im Speicher des Medienunternehmens angelegt. Diese sind beispielsweise die Nutzungsdauer, Scrolltiefe, Genres der konsumierten Beiträge und die Anzahl der Besuche. Aus den Daten werden Schlüsse über Interessen, Alter, Geschlecht und soziale Stellung gezogen. Diese wiederum werden mit ausspielbereiten Inhalten des Medienunternehmens verbunden.

Die Inhalte und die Nutzungsprofile sind in Embeddings in einem mehrdimensionalen Vektorraum organisiert. Jedes Objekt in diesem Raum kann durch einen Vektor beschrieben werden. Durch die räumliche Nähe von Inhalten und Nutzungsprofilen wird berechnet, welche Beiträge einer Nutzerin bzw. einem Nutzer präsentiert und mit welcher Wahrscheinlichkeit eine Nutzerin bzw. ein Nutzer einen ihr bzw. ihm vorgeschlagenen Beitrag konsumiert. Beiträge mit einer höheren Wahrscheinlichkeit werden prominenter auf der Startseite der jeweiligen Plattform platziert. Dazu kommt, dass Beiträge, die die Nutzerin bzw. der Nutzer schon gesehen hat, ausgeblendet werden. Wenn nun die Nutzerin bzw. der Nutzer den vorgeschlagenen Beitrag konsumiert, wird die Distanz zwischen beiden im Vektorraum verkürzt, indem das Nutzungsprofil angepasst wird. Gefällt der Inhalt nicht, wird die Distanz ebenfalls verändert. Das Nutzungsprofil entfernt sich vom Beitrag. Hier muss die Gefahr des Overfittings stets im Auge behalten werden. Je länger die Nutzerin bzw. der Nutzer auf der Plattform aktiv ist, desto mehr Daten werden zum Nutzungsverhalten gespeichert. Dadurch wird ihr oder sein Profil genauer und die Empfehlungen präziser.

Archiv, InstructGPT-Modell, Speicher der Nutzungsdaten und Automation stehen also im ständigen Austausch miteinander. Die Automation nimmt neu eingetroffene Nutzungsdaten aus dem Speicher und gibt sie an das KI-Modell. Dieses passt die Nutzungsprofile an und gibt Vorschläge für Inhalte aus. Die Inhalte werden dann durch die Automation über MAM und CMS der Nutzerin bzw. dem Nutzer präsentiert, sobald diese oder dieser die Webseite betritt. Die Nutzerin bzw. der Nutzer wählt einen Inhalt aus und generiert dadurch neue Nutzungsdaten, die wiederum den Kreislauf befeuern.

8.4.2SCHNITTSTELLEN

Im Vergleich zu ChatGPT besteht die Lösung für eine Systemlandschaft mit InstructGPT, Whisper und DALL-E aus mehreren KI-Modellen und damit auch aus mehreren Schnittstellen zu den jeweiligen Modellen. Jede Anwendung hat ihren eigenen Endpunkt.

Für die verschiedenen Aufgabenstellungen werden verschiedene Modelle von OpenAI verwendet. Für Archivanwendungen wird ein anderes Modell genutzt als für Transkriptionen und wieder ein anderes für Sprachanwendungen. Alle Modelle sind über eine API erreichbar. Die Art der Anfrage richtet sich nach dem jeweiligen Use Case (siehe Abbildung 30) [81].

Anwendungsfall	Endpunkt	Kompatibles Modell	Individualisierbar?
Chatanwendungen / Rückmeldung an Nutzer*innen (A31)	/v1/chat/completions	GPT-3.5	Nein
Sprachvervollständigung (Completion) (A17, A18, A23, A28)	/v1/completions	davinci, curie, ada	Ja
Sprachverbesserung (Edit) (A24)	/v1/edits	davinci	Ja
Transkription / Übersetzung von Audio (A22)	/v1/audio/transcriptions	whisper-1	Nein
Finetuning	/v1/fine-tunes	davinci, curie, ada, babbage	Ja
Archivanwendungen/ Personalisierung (A20, A21, A30)	/v1/embeddings	ada	Ja
Sentiment Analysis (A32)	/v1/moderations	text-moderation	Nein
Bildgeneration	/v1/images/generations	DALL-E	Nein

Abbildung 30: Kompatibilität der Modelle mit den jeweiligen Endpunkten nach [81].

Eine Anfrage sieht jedoch ähnlich aus wie für ChatGPT (siehe Abschnitt 7.3.2). Um einen Auftrag abzuschicken, wird ebenfalls POST verwendet [92]. Je nach Use Case wird ein anderer Endpunkt benötigt. Der Anfrage werden dann ebenfalls wie beim Function Calling vorbereitete Variablen in

Funktionen übergeben, um strukturierte Daten zurück zu erhalten. Beispiele für Variablen befinden sich in Anhang acht.

Auch hier kommen Rate Limits zum Einsatz. Die Grenze für Embeddings liegt bei 350000 Token pro Minute. Das entspricht etwa 438 Seiten Text pro Minute¹³. DALL-E generiert maximal 50 Bilder pro Minute. Whisper kann nur 50 Anfragen pro Minute verarbeiten [98].

Erweiterung des Metadatenmodells

Um mit den InstructGPT-Modellen weitere Use Cases abzubilden, bedarf es auch einer Weiterentwicklung des Metadatenmodells. Diese wird am Beispiel von Embeddings gezeigt. Ein Embedding ist die Einordnung eines Wortes, Satzes oder Textes in einen mehrdimensionalen Vektorraum, sodass sich inhaltlich ähnliche Wörter in räumlicher Nachbarschaft befinden. Damit lassen sich eine effizientere Archivsuche und eine personalisierte Ausspielung von Inhalten realisieren. Die Embedding-Modelle erhalten als Prompt den einzuordnenden Satz oder Text und geben einen Ort im Vektorraum in Form von Zahlen zurück. Anhand dieses Ortes können nah beieinanderliegende Inhalte identifiziert werden.

```
"data": {[
  "embedding": [
    -0.006929283495992422,
    -0.005336422007530928,
    ...
    -0.024047505110502243
  ],
  "index": 0,
  "object": "embedding"
}]
"model": "text-embedding-ada-002",
"object": "list",
"usage": {
  "prompt_tokens": 5,
  "total_tokens": 5
}}
```

Abbildung 31: Beispiel einer Antwort auf ein Embedding [94].

¹³ Bei einer Textformatierung ähnlich wie in dieser Thesis.

Entsprechend benötigt ein solcher Embeddingvektor ein Metadatenenelement, unter dem dieser abgespeichert werden kann. Er wird in den Klassen 13 oder 14 angelegt.

8.4.3 WEITERE GUIS

Da durch den Einsatz von InstructGPT weitere Use Cases entstanden sind, werden auch weitere Zugangsprofile (siehe Abschnitt 8.4.2) und weitere GUIs nötig. Es werden Benutzungsoberflächen verbessert oder neu entworfen. Neben jeder produktiv arbeitenden GUI wird außerdem der aktuelle Rundown aus dem Redaktionssystem angezeigt.

B: Archiv II

Eine Weiterentwicklung erhält die grafische Benutzungsoberfläche der Archivarinnen und Archivare. Da durch Whisper nun auch Audiospuren transkribiert werden können, werden zusätzlich Audioaufnahmen und Videos automatisch verschlagwortet. Statt Texten ist in der GUI eine Audioaufnahme oder ein Video eingebettet. Trotzdem werden die extrahierten Metadaten in Textformat der Nutzerin bzw. dem Nutzer angezeigt und können bei Bedarf bearbeitet werden. Die Feedbackfunktion meldet grob fehlerhafte Metadaten.

G: Audio

Die grafische Benutzungsoberfläche zur Transkription von Audioaufnahmen ähnelt der Archiv-GUI. Jedoch mit dem Unterschied, dass statt der Metadaten nun das Transkript bearbeitet werden kann. Die Arbeit mit der GUI läuft wie folgt ab: Über eine Archivsuche mit Texteingabe und Suchergebnissen sucht die Anwenderin bzw. der Anwender eine Audioaufnahme aus. Diese wird über Drag and Drop in den Player der GUI gezogen. Mit der Schaltfläche „Transcript“ wird die Essenz dem KI-Modell Whisper zugeführt und in Text umgewandelt. Der Text wird ebenfalls in der GUI angezeigt. Hier kann die Benutzerin bzw. der Benutzer ihn noch bearbeiten. Fehler im Transkript werden über eine Feedbackfunktion gemeldet.

H: Grafik

Die Grafik-GUI ähnelt wiederum der Audio-GUI. Hier wird jedoch nicht über eine Texteingabe im Archiv gesucht, sondern auf den Prompt hin werden Bilder generiert, die noch nicht existieren. Bei jeder Eingabe werden direkt mehrere Alternativen generiert. Die Nutzerin bzw. der Nutzer wählt das geeignete Bild zu den kommenden Nachrichten aus und verschiebt dieses in den Speicher, auf den der Hintersetzer im Studio zugreift oder ordnet das Bild einem textuellen Beitrag als Symbolbild zu. In der GUI besteht die Möglichkeit, das Bild auf einer großen Fläche zu betrachten und auf mögliche Fehler zu prüfen. Es ist außerdem möglich, das Bild über Funktionsschaltflächen zu bearbeiten, wenn das Bild im Großen und Ganzen gefällt, jedoch noch kleine Details ausgebessert werden sollen. Eine Feedbackeingabe nimmt fehlerhafte Bilder auf. Hier ist es nun möglich, Bilder per Drag and Drop in das Eingabefeld zu ziehen und einige erklärende Worte hinzuzufügen.

8.4.4 DATENAUFBEREITUNG UND TRAINING

Für das Nachtraining der KI-Modelle werden Daten des Medienunternehmens verwendet. Diese liegen in Form von Texten, Bildern und Videos vor. Für das Training von Sprachmodellen werden ausschließlich Texte verwendet. Jedoch können auch die Transkripte von Bildern und Videos als Texte gesehen werden. Durch das Nachtraining prägt sich das KI-Modell den Stil und die Haltung des Medienunternehmens ein.

Dafür durchlaufen die Daten und auch das KI-Modell den Data Science Life Cycle. In Abbildung 32 wird der Life Cycle für die Entwicklung eines neuen KI-Modells aufgezeigt. Trotz dessen, dass in diesem Fall das KI-Modell schon existiert und nur noch ein Fine Tuning vorgenommen wird, werden alle Schritte durchlaufen. Ein KI-Modell ist laufend diesem Kreislauf unterworfen. Das bedeutet, dass selbst nach einem erfolgreichen Durchlauf immer wieder aufs Neue beurteilt werden muss, ob das KI-Modell den aktuellen Anforderungen entspricht und ob nicht erneut nachgebessert werden muss [109, 110].

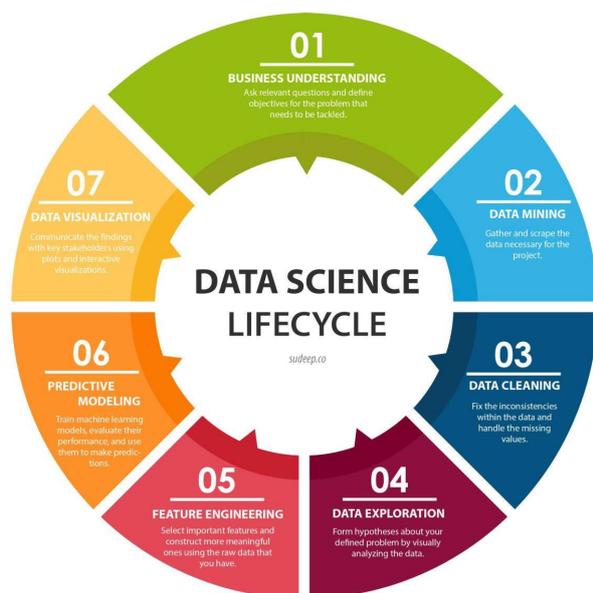


Abbildung 32: Data Science Life Cycle [111].

Im Unterschied zum Prompt Engineering verändert das Finetuning die Gewichte des KI-Modells. Es erfordert einen größeren Eingriff in das KI-Modell. Dieser nimmt wiederum mehr Zeit in Anspruch [104, S. 4].

Anforderungen definieren

Der Beginn jedes Projekts ist die Festlegung des Ziels und dessen Anforderungen. Damit ist die Grundlage für alle zukünftigen Schritte gelegt. Zunächst werden zusammenfassende Anforderungen definiert (siehe Kapitel fünf). Diese werden im weiteren Verlauf heruntergebrochen und in kleinere Einheiten definiert.

Datensammlung

Datensätze bilden den Ausgangspunkt für jedes KI-Modell. Medienunternehmen verfügen durch ihr langes Bestehen über ein umfangreiches Archiv. Dazu kommt die Erhebung von Nutzungsdaten auf

der Webseite und in der Mediathek. Bei dieser Art von Daten muss vorab geklärt werden, ob diese überhaupt verarbeitet werden dürfen (siehe Abschnitt 5.1.1).

Datenfluss und Zugang zu den Daten

Eine weitere Eigenschaft, die Trainingsdaten erfüllen müssen, ist die Zugänglichkeit. Damit ist gemeint, dass die Daten beispielsweise in einem Dateiformat vorliegen müssen, das verarbeitet werden kann, oder in einem Speicher mit ausreichend großer Übertragungsgeschwindigkeit liegt. Darüber hinaus werden Datenflüsse definiert, durch welche auf die Datensätze zugegriffen wird.

Datenbereinigung

Die Datenbereinigung stellt eine unumgängliche Vorarbeit dar. Durch fehlende Werte, Dubletten oder widersprüchliche Eingaben kann die Aussage eines gesamten Datensatzes verändert werden und somit das generelle Ziel des KI-Modells verfehlt werden (siehe Abbildung 33).

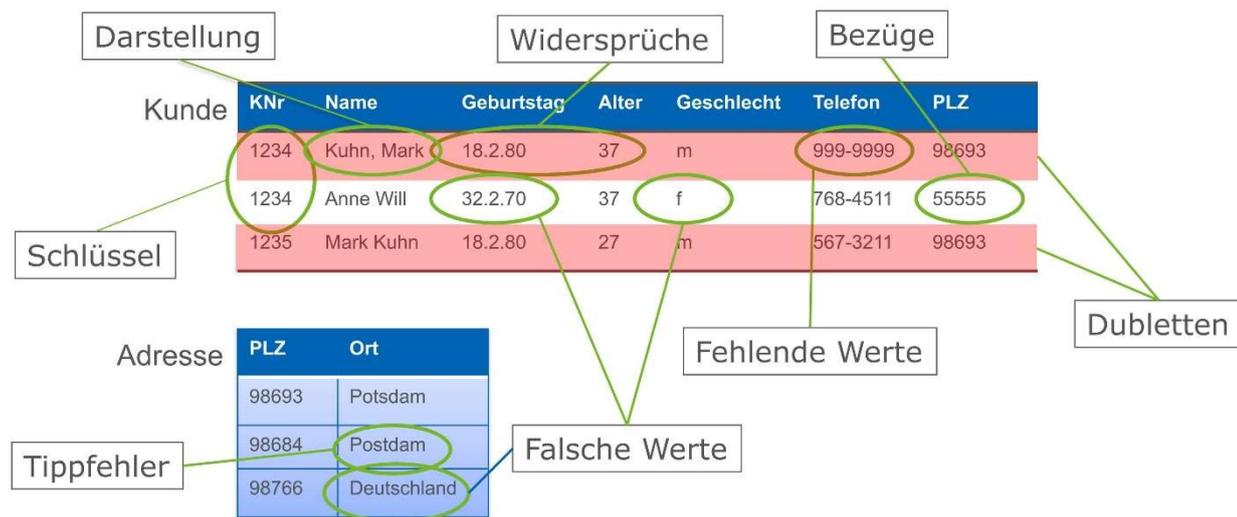


Abbildung 33: Typische Datenfehler [112].

Dabei müssen folgende Schritte durchlaufen werden, um eine gewinnbringende Bereinigung zu erzielen: Untersuchung, Diagnose, Behandlung und Dokumentation. Zunächst wird geprüft, welches Problem im Datensatz vorliegt. Danach wird das Problem klassifiziert. Es ist möglich, dass es sich um fehlende, falsche oder um besonders unhandliche, aber richtige Werte handelt. Sind diese Informationen beschafft, werden die Schwachstellen beseitigt. Dafür gibt es mehrere Methoden. Diese sind Löschen der Datenzeile oder Zurechnung von Mittelwerten. Ist die Bereinigung abgeschlossen, wird sie dokumentiert [113].

Eigene Trainingsdaten vorbereiten

Für die Anwendung der KI werden bereits im Vorfeld Prompts sowie Textbausteine, die bei den Nachrichten häufig zum Einsatz kommen, vorformuliert. Sie werden ebenfalls als Trainingsmaterial verwendet. Außerdem werden Namen (z.B. der hauseigenen Moderatoren) in der richtigen Schreibweise gelernt, sodass diese später keine Probleme liefern [28, S. 4].

Explorative Datenanalyse

Der nächste Schritt ist die statistische Analyse der Daten, um die Variablen kennenzulernen und ein Verständnis für die Daten zu erreichen. Damit sind zunächst die bestehenden Variablen des KI-Modells gemeint sowie die Variablen des neuen Trainingsmaterials. Es können hier bereits Korrelationen identifiziert werden, die für das spätere Training genutzt werden können. Die Anwendung der statistischen Methoden ist von der Beschaffenheit des Datensatzes abhängig. Eine Visualisierung der Daten hilft beim Verständnis.

Finetuning: Training und Validierung

Die neuen Trainingsdaten werden dazu verwendet, die Gewichte des NN zu verfeinern und an den konkreten Anwendungsfall anzupassen. Jedoch wird nur ein Teil der Daten zum Training verwendet. Der andere Teil dient als Validierungsmaterial, um die Qualität des Outputs zu beurteilen. Man kann ein KI-Modell nicht mit denselben Daten trainieren und testen, da es sonst nur für diese spezifischen Testdaten die richtigen Antworten liefern würde und nicht universell einsetzbar ist.

Finaler Einsatz der KI

Für den Anfang werden erst einzelne Use Cases umgesetzt, damit sie von Journalistinnen und Journalisten eingesetzt und bewertet werden können. Durch das Feedback werden ebenfalls die Gewichte im NN angepasst. Der Übergang zwischen der Testphase und dem produktiven Einsatz verläuft so fließend wie möglich, da auch im produktiven Einsatz immer wieder am KI-Modell nachgebessert werden muss, um es innerhalb der qualitativen und rechtlichen Schranken zu betreiben.

Kommunikation und Visualisierung der Ergebnisse

Bereits während der Trainingsphase wird laufend der Stand im Projektteam und im Medienunternehmen veröffentlicht. Ein kontinuierlicher Fortschritt fördert Akzeptanz und Interesse an dem Veränderungsprojekt. Um abstrakte Werte greifbarer zu machen, werden sie in geeigneten Formen für die jeweilige Zielgruppe visualisiert.

Dokumentation und Tests

Um die Aktualität der KI-Modelle zu erhalten, werden die neu erzeugten journalistischen Erzeugnisse zum Nachtraining in regelmäßigen Abständen verwendet. Wird eine Veränderung an der Instanz der KI-Modelle vorgenommen, muss auch die Dokumentation fortlaufend angepasst werden. Es wird für jeden Trainingszyklus eine neue Version der Dokumentation angelegt. Das Nachtraining findet immer unter Aufsicht eines Menschen statt.

Nach jeder großen Veränderung von ChatGPT muss das Modell erneut getestet werden. Die Tests erfolgen zunächst durch die Data-Scientistin bzw. den Data-Scientisten und schließlich auch durch eine Gruppe von Anwenderinnen und Anwendern. Durch das Feedback wird Ausmaß und Güte der Veränderung bestimmt.

8.5 ZWISCHENFAZIT

Für die Einführung der zusätzlichen Modelle von OpenAI in den Newsworkflow werden zusätzliche Konzepte und Lösungen gebraucht. Da nun mit mehreren Modellen statt nur einem einzigen gearbeitet wird, werden jedem Modell nur wenige Use Cases zugeordnet. ChatGPT ist auf Chats trainiert, also übernimmt es sämtliche Interaktionen mit Menschen. Es gibt also kein universelles Modell mehr, sondern mehrere spezialisierte. Die Konzepte und technischen Lösungen in diesem Kapitel stehen also nicht für sich alleine, sondern sind als Ergänzungen und Verbesserungen von jenen in Kapitel sechs und sieben zu verstehen.

Neue Erkenntnisse gibt es im Bereich der Workflows. Diese werden um neue Schnittstellen zur KI ergänzt. Durch Embeddings kann außerdem ein völlig neuer Ausspielungsworkflow gestaltet werden. Die größten Neuerungen durch die InstructGPT-Modelle sind die unterschiedliche Art der Schnittstellen zu den Modellen und das Finetuning.

Dadurch, dass nun mehrere Modelle über APIs mit der Automation verbunden sind, muss nicht mehr der Umweg über das Function Calling gegangen werden, um strukturierte Informationen vom KI-Modell zu erhalten. Da ein Modell nur thematisch ähnliche Use Cases verarbeitet, reicht ein Set an Variablen für die API aus. Das Modell muss sich also nicht die passende Funktion selbst herausuchen, sondern bearbeitet ausschließlich die mitgelieferte. Die weitere Funktionalität der InstructGPT-Modelle ist das Finetuning. Damit kann den KI-Modellen der Stil des Medienunternehmens durch Nachtraining mit den eigenen Daten aufgeprägt werden. Dies führt zu authentischeren Ergebnissen. Dafür werden die hauseigenen Daten des Medienunternehmens für das Training aufbereitet.

9 DISKUSSION

Nachdem in den vorherigen Kapiteln geklärt wurde, ob und unter welchen Voraussetzungen ein KI-Modell den Newsworkflow bereichert, soll folgende Diskussion das Für und Wider des Einsatzes von KI generell abwägen. Dabei geht der Umfang der Diskussion über den Anwendungsbereich Journalismus hinaus und befasst sich mit der generellen Verantwortung bei der Nutzung von KI.¹⁴

KI hat schon lange Einzug in den Arbeitsalltag der Menschen. Mit Use Cases auch außerhalb des Journalismus sehen Unternehmen KI als Mittel, um eine höhere Genauigkeit, Geschwindigkeit und damit Effizienz in ihrer Produktion zu erzielen. KI nicht in den Produktionsablauf aufzunehmen, könnte in naher Zukunft sogar als Nachteil gesehen werden [114]. Einige Beispiele sind die Verbesserung von medizinischen Diagnosen und die individuelle Patientenbehandlung. KI unterstützt bei der Analyse von Patientendaten und medizinischen Aufnahmen. Eine weitere Anwendung sind Sprachassistenten und die Unterstützung durch Smart Homes. Diese Art der Automatisierung im Alltag sorgt für eine Zeitersparnis und Erhöhung des Komforts. Als letztes Beispiel soll der Einsatz im Straßenverkehr genannt werden. KI wird in selbstfahrenden Autos eingesetzt und soll dort bei der Verringerung von Unfällen und Verletzungen helfen. Außerdem analysiert und optimiert KI Verkehrsströme, um den Durchfluss zu verbessern. KI-Modelle im Allgemeinen und LLMs im Speziellen haben die Eigenschaft, Menschen schwierige Sachverhalte abzunehmen und ihnen zu erklären [115].

Trotz dessen ist der Zugang zu KI-Modellen auf der Welt ungleich verteilt. Das liegt an der Monopolstellung einzelner Unternehmen. Ursachen dafür sind eine übermäßige Verfügbarkeit von Rechenleistung, Daten, Wissen und finanziellen Mitteln [116]. Dagegen behaupten sich die Open Source-Modelle gegen die geschlossenen Modelle. Bei Tests erzielen sie ähnlich gute Ergebnisse wie die kommerziellen Modelle mit dem Unterschied, dass ihr Quellcode und Trainingsdatensätze öffentlich einsehbar sind [117].

Die hauptsächliche Entwicklung von KI-Modellen findet in den USA statt. 73 Prozent der KI-Modelle seit 2017 stammen aus den USA, direkt gefolgt von 15 Prozent aus China [118, S. 6]. Entsprechend schwer haben es europäische und damit auch deutsche KI-Modelle auf dem internationalen Markt. Die Nutzung von amerikanischen Modellen für europäische Rundfunkanstalten ist jedoch auch nicht besonders effektiv, denn diese haben aus europäischer Sicht einige Defizite. In den Trainingsdaten der amerikanischen Modelle sind nämlich europäische und deutsche Inhalte unterrepräsentiert, was zu einer unausgewogenen Darstellung von Tatsachen führen kann. Ein KI-Modell würde also über zu wenig diversifizierte Daten verfügen, um verifizierbare Aussagen zu treffen, die auf der gesamten Welt gültig sind. Die Modelle wurden zudem nur auf Englisch trainiert. Andere europäische Sprachen werden nachträglich dazugelernt, weshalb sich die englische Semantik bei der Zuordnung von Wort und Bedeutung durchgesetzt hat.

Um also europäische KI-Modelle zu entwickeln und zu trainieren, müssen folgende Voraussetzungen erfüllt werden. Die Verfügbarkeit von kompetentem Personal ist die erste Hürde. Es muss eine

¹⁴ Der Stand der Diskussion bezieht sich auf den Zeitpunkt der Erstellung der Thesis.

ausreichende Anzahl an geschulten Expertinnen und Experten mit Erfahrungen im Bereich von KI geben. Weiterhin muss die Recheninfrastruktur verfügbar sein. Diese sollte eine effiziente Datenaufbereitung, Training und Evaluation ermöglichen. Schließlich benötigt das KI-Modell Datensätze zum Training. Sie müssen einen gewissen Qualitätsstandard und Umfang aufweisen [118, S. 60].

Die Entwicklung und der Einsatz von KI-Modellen in Deutschland sind aufgrund der besonderen Situation in Europa anspruchsvoller als beispielsweise in den USA. Die Medienunternehmen und die Medienwirtschaft sind hier stärker fragmentiert. Sie decken auch aufgrund der Mehrsprachigkeit in Europa jeweils nur kleine Bereiche des Publikums ab, was einen Nachteil im internationalen Wettbewerb bedeutet [25, S. 31]. Große europäische Medienhäuser, die eine eigene Forschungs- und Entwicklungsabteilung haben (z.B. British Broadcasting Corporation (BBC)), sind hier im Vorteil. Kleinere regionale Medienhäuser begeben sich jedoch in Abhängigkeit von KI-Herstellern [7].

Die europäische DSGVO bildet eine der strengsten Datenschutzregelungen der Welt. Entsprechend werden bei der Datenverarbeitung durch KI Bedenken formuliert. Im Gesundheitswesen verarbeitet die KI beispielsweise sensible Patientendaten. Diese sind unerlässlich für die Aufgabe, die dem KI-Modell gestellt wird, wirft jedoch Bedenken bei der Behandlung dieser Daten auf. Gleichzeitig müssen auch die ethischen Implikationen beim Einsatz von KI in der Medizin bedacht werden. Eine KI kann hier auf keinen Fall im Alleingang über die Behandlung eines Menschen oder seine Überlebenschancen bestimmen. Jeder Vorschlag der KI durchläuft eine Kontrolle durch einen Menschen. Die Folgen im Falle von unsachgemäßer Verarbeitung von Daten wären definitiv weitreichender als im Journalismus.

Auch in anderen Bereichen kann KI nicht ohne Vorsichtsmaßnahmen verwendet werden. Sicherheitsbedenken gibt es bei der Verwendung von Smart Home-Geräten und Sprachassistenten. Diese befinden sich in der privaten Wohnung und sind jederzeit bereit gesprochene Befehle der Bewohnerinnen und Bewohner aufzunehmen und umzusetzen. Daher werden sämtliche gesprochenen Worte in der Wohnung erfasst und auf Anweisungen geprüft. Dabei handelt es sich in den meisten Fällen um Gespräche, die gar nicht an den Sprachassistenten gerichtet waren. Dadurch erhalten die Sprachassistenten einen tiefen Einblick in das Leben ihrer Besitzerinnen und Besitzer.

Um den verschiedenen Gesetzen und Verordnungen in allen Teilen der Welt gerecht zu werden, ist eine Regulierung von KI notwendig. Die Diversität der Anwendungsfälle sowie die hohe Entwicklungsgeschwindigkeit fördern die Unberechenbarkeit. Eventuell wird es nötig sein, individualisierte Instanzen eines KI-Modells für jedes Land bzw. jeden Wirtschaftsraum auszurollen. Sam Altman hält dies zumindest nicht für ausgeschlossen [114].

Wie bei allen generellen schöpferischen Tätigkeiten spielt die Urheberbeschaft eine wichtige Rolle beim Design eines KI-Modells und dies gleich in zweierlei Hinsicht. Das Urheberrecht wird auf den Output des KI-Modells sowie auf die Trainingsdaten angewandt. Die rechtswissenschaftliche Forschung ist im Moment auf dem Stand, dass das Urheberrecht an generierten Werken bei der Anwenderin bzw. beim Anwender liegt. Die Schöpferin bzw. der Schöpfer eines Werkes kann sich die Nutzung seines

Werkes als Trainingsmaterial vorbehalten. Es ist jedoch nicht klar, wie die Nutzung von urheberrechtlich geschützten Werken zum Training in Zukunft vergütet werden soll [119]. Darüber hinaus stellt sich die Frage, ob und wie KI in einem akademischen Kontext genutzt werden kann. Da beispielsweise ChatGPT gerne Fakten halluziniert oder Quellen erfindet, wäre ein Einsatz desselben auf diesem Gebiet nicht unbedingt zielführend. Das KI-Modell präsentiert seine Aussagen sehr überzeugend, trotzdem müssen diese manuell auf ihren Wahrheitsgehalt geprüft werden [117]. Die rechtswissenschaftliche Forschung auf diesem Gebiet läuft ununterbrochen. Jedoch ist die Entwicklung neuer Anwendungsfälle und Lösungen im Bereich der KI schneller.

Die Güte einer KI wird zu einem großen Teil von der Qualität ihrer Trainingsdaten bestimmt. Dabei kommt es auf die Diversität, die Genauigkeit und Objektivität der Annotationen an. Trotzdem werden teilweise Daten zum Training verwendet, die diese Kriterien nicht erfüllen, deren Ursprung zusätzlich nicht geklärt oder deren Nutzung rechtlich gar nicht abgedeckt ist [119]. Ebenfalls in der Kritik stehen die Arbeitsbedingungen, unter denen Daten annotiert werden. Da ChatGPT auf Lernprozessen nach menschlichem Feedback beruht, werden hier also viele Menschen benötigt, die die Antworten des KI-Modells bewerten. Dabei handelt es sich zum Teil um kenianische Arbeiterinnen und Arbeiter, die für niedrige Löhne lange Arbeitszeiten bewältigen müssen. Wird beispielsweise ein Datensatz zur Erkennung von jugendschutzkritischen Inhalten benötigt, kann die Annotation zu traumatischen Erfahrungen bei den Annotatorinnen und Annotatoren führen [5].

Durch die Tatsache, dass Trainingsdaten nie völlig unvoreingenommen sind, ist auch kein KI-System jemals völlig objektiv. Daher sind alle generierten Texte nur als Kommentar zu verstehen und nie als Sachtext. Umso wichtiger ist es, dass die Nutzerin bzw. der Nutzer die Kontrolle über die Werte des KI-Modells hat [114]. Letzteres wird besonders dadurch erschwert, dass sich KI-Modelle von Natur aus als Black Boxes gestalten. Man sieht die Eingabe, man sieht die Ausgabe. Die Abläufe im Inneren sind jedoch so komplex, dass ein Mensch diese nicht nachvollziehen kann. Er kann also bei einem Output nicht beurteilen, warum ein KI-Modell eine Entscheidung getroffen hat. Dadurch kann der Mensch bei falschen Antworten auch den Ursprung des Fehlers nicht so einfach zurückverfolgen [3].

Dieses Phänomen tritt aufgrund von mangelnder Erklärbarkeit auf. Erklärbarkeit bedeutet im ML-Kontext, dass ein System seine inneren Vorgänge beschreiben kann, sodass ein Mensch diese nachvollziehen kann [46, S. 2]. Zur Förderung der Erklärbarkeit kann ein KI-Modell gefragt werden, ob es bei der Generation einer Antwort seine Argumentation mitteilt. Durch die Beschreibung der Argumentation kann ein Mensch leichter herausfinden, ob diese profund ist [46, S. 1]. Dadurch können auch andere Güte-Kriterien wie Fairness, Verlässlichkeit, Robustheit, Kausalität und Vertrauen weiter vorangetrieben werden [46, S. 3]. Denn der Wunsch nach Erklärbarkeit stammt aus einer unvollständigen Lösungsformalisierung. Die inneren Abläufe eines KI-Systems werden also nicht in eine eindeutige und unmissverständliche Form gebracht. Deshalb besteht eine Kluft zwischen Evaluierung und Optimierung, da ein außenstehender Mensch zunächst wenig Verständnis für die Rechenoperationen hat, um diese effizient zu verbessern [46, S. 3].

Es müssen alle Ausgaben eines KI-Modells hinterfragt werden. Ein LLM ist kein Autor, kein Taschenrechner, keine eigenständige Persönlichkeit, sondern lediglich ein Werkzeug, das zur Unterstützung für Menschen eingesetzt wird. Daher trägt es keine Verantwortung für seinen Output [117]. Die Verantwortung und auch die Haftung bei Fehlern liegen bei der Betreiberin bzw. dem Betreiber durch eine Aufsichtspflicht. Andererseits liegt auch eine gewisse Verantwortung bei der Anwenderin bzw. dem Anwender. Im Journalismus müssen also die Journalistinnen und Journalisten und Redakteurinnen und Redakteure den Output der KI kontrollieren und auf Plausibilität prüfen. Dabei handelt es sich nicht um ein technisches Problem, sondern eine Frage des kritischen Mindsets der Anwenderinnen und Anwender.

In der Aufzählung der Einsatzmöglichkeiten für KI gehört der Journalismus zu jenen mit durchschnittlichen Anforderungen. Ein Anwendungsfall, bei dem der Output eines KI-Modells direkt mit dem Wohlergehen von Menschen zu tun hat, wäre dagegen der Einsatz in selbstfahrenden Fahrzeugen. Aufgrund der Geschwindigkeit der Entscheidungen ist es einem Menschen – je nach Grad der Automatisierung – nur noch im Notfall möglich einzugreifen. Da die KI juristisch nicht als Person gewertet wird, stellt sich die Frage nach der Haftung bei Unfällen. Erst, wenn diese abschließend geklärt sind, kann der Einsatz von selbstfahrenden Fahrzeugen auf den Straßen erlaubt werden.

Zunächst sind die Chancen, die sich durch KI-Modelle ergeben, nicht zu übersehen. Aufgaben, die ein Mensch gar nicht oder nur in einem sehr langen Zeitraum bewältigen kann, werden von der KI übernommen. Mit der Entwicklung der LLMs ist die Kommunikation mit dieser um ein Vielfaches einfacher geworden. Durch die Entwicklung der Sprachmodelle müssen Menschen, die ein KI-Modell verwenden wollen, kaum oder gar nicht mehr programmieren können. Dadurch ergeben sich soziale Auswirkungen auf die Menschen und ihre Arbeitsweisen. Aufgaben werden in kürzerer Zeit und mit weniger Personen realisiert. Die Aussage von Sam Altman, dem Gründer von OpenAI, ist jedoch, dass Menschen nicht ersetzt werden, da es noch immer Aufgaben gibt, die die KI nicht lösen kann und die von einem Menschen bearbeitet werden müssen [114].

Ein Aspekt, den KI im Nachrichtenjournalismus nicht übernehmen kann, ist die menschliche Perspektive in der Berichterstattung. Die Beschreibung von menschlichen Erfahrungen eines Ereignisses bleibt nach wie vor einzigartig. KI kann beispielsweise automatisiert Berichte aus Sportergebnissen erstellen, aber keine Stadionatmosphäre beschreiben [120]. Die menschliche Arbeit sorgt außerdem für ein breites Vertrauen in die Güte und Richtigkeit der Informationen. Da Medienunternehmen als Filter für die Masse an Informationen arbeiten und mit der Veröffentlichung diesen ein Qualitätssiegel verleihen, besteht ein hohes Maß an Vertrauen in die etablierten Medien. Eine weitere Arbeit, die KI nicht leisten kann, ist die Korrespondententätigkeit. Große Medienhäuser unterhalten ein weltweites Netzwerk, um möglichst unmittelbar an Informationen auf der ganzen Welt zu kommen [120].

Generell hat der Journalismus eine gesellschaftliche Aufgabe innerhalb der Demokratie und muss daher von hoher Qualität sein. Daher dürfen KI-Modelle kein verzerrtes Wahrheitsbild abgeben. Es

bleibt die Pflicht der Medienunternehmen, solche Ergebnisse kritisch zu betrachten. Durch personalisierte Ausspielung von Inhalten in den Mediatheken und auf Sozialen Medien erhält jede Nutzerin und jeder Nutzer ein anderes Bild der Nachrichten [7].

Die Aufgabe öffentlich-rechtlicher Rundfunkanstalten ist die Unterrichtung der Bevölkerung nach den Standards des Pressecodex. Ihre Ziele sind die Vielfältigkeit der Inhalte sowie der Anspruch immer aktuelle Inhalte zu präsentieren. Dabei orientieren sie sich an Neuartigkeit und Serendipität, da im Weltgeschehen nicht vorausgesagt werden kann, welche Nachricht am nächsten Tag veröffentlicht wird. Gleichzeitig sind öffentlich-rechtliche Medien dazu angehalten, transparent mit der Auswahl der Themen und ihren Informationsquellen umzugehen. Die Rundfunkanstalten sind unabhängig vom Staat, um keinen Verdacht der Beeinflussung aufkommen zu lassen. Deshalb werden sie nicht über Steuergelder finanziert, sondern durch einen gemeinschaftlichen Beitrag [74, S. 85-86].

Daher kann der Eindruck entstehen, dass öffentlich-rechtliche Medien ihren Platz als gemeinsamen Ort für Informationen und Diskurs verlieren, wenn das Nachrichtenangebot nicht für alle Nutzerinnen und Nutzer dasselbe ist. Es besteht jedoch kein Zweifel, dass Personalisierung dabei hilft, Nutzerinnen und Nutzer für das Medium zu interessieren und eine gewisse Loyalität aufzubauen [74, S. 56]. Daher müssen öffentlich-rechtliche Medien im Besonderen bei der Wahl der Personalisierungsalgorithmen aufmerksam sein. Sie tragen die Verantwortung dafür, welche Inhalte gezeigt und welche ausgeblendet werden. Dies wird zum einen dadurch bestimmt, wie Daten für Algorithmen aufbereitet werden, aber auch wie Vorhersagen und Schlüsse aus Daten gezogen werden. Dazu kommen die Kriterien und Kategorien, mit denen das Nutzerverhalten gemessen wird. Den Rundfunkanstalten werden also viele Freiheiten bei der Gestaltung der Algorithmen gegeben. Einen Algorithmus zu finden, der aber völlig objektiv Vorschläge zu Inhalte gibt, existiert nicht. Daher wird eine personalisierte Ausspielung immer die Meinung der Nutzerinnen und Nutzer beeinflussen und in Verlängerung auch die Meinung der Öffentlichkeit [74, S. 85]. Daher stellt sich die Frage, welchen Einfluss Personalisierungsalgorithmen auf die Gesellschaft und in Verlängerung auf politische Meinungen und Wahlergebnisse haben sollten.

Als Antwort auf diesen Umfang an Verantwortung wird immer mehr gefordert, bereits in der Ausbildung Journalistinnen und Journalisten an die Datenwissenschaft heranzuführen. Durch zusätzliche Module im Studium oder einen Codex – ähnlich dem Pressecodex – zum Umgang mit KI, könnte mehr Transparenz geschaffen werden, so der deutsche Presserat [7]. Mit der Nutzung von KI-Modellen in jeder Branche, nicht nur im Journalismus, stellen sich immer wieder die gleichen Herausforderungen und damit auch Probleme an die Nutzerinnen und Nutzer. Es müssen Abwägungen getroffen werden, wie und ob diese Aufgaben die eigene Branche betreffen und wie sie gelöst werden können.

10 FAZIT

Bezogen auf den Mediensektor bietet der Einsatz von KI im Newsworkflow viele verschiedene Use Cases. Von Textkorrekturen über Bildgeneration bis zur personalisierten Ausspielung ist die Menge der möglichen Einsatzszenarien groß. Die Verwendung von KI wird jedoch von Gesetzen und ethischen Grundsätzen reglementiert, die die sichere Verwendung von KI-Systemen fördern. Die technische Entwicklung eilt jedoch der Gesetzgebung voraus.¹⁵

Die einzuhaltenden Gesetze stammen zu einem Teil aus dem Medienumfeld wie beispielsweise die publizistischen Grundsätze. Der andere Teil der Vorgaben sind Gesetze, die direkt auf die KI abzielen. Dazu kommen allgemeingültige Gesetze wie Artikel fünf des GG zur Meinungs- und Pressefreiheit sowie das AGG zur Gleichstellung aller Menschen.

Alle Anforderungen, funktionell, rechtlich-ethisch und organisatorisch, an eine Systemlandschaft mit KI werden in einer Anforderungsmatrix zusammengefasst und sind Grundlage für die folgenden Lösungskapitel. KI soll gewisse geforderte Aufgaben übernehmen und sich dabei innerhalb der Schranken des Gesetzes sowie des guten Journalismus bewegen.

Die Einführung von ChatGPT soll sich ebenfalls an organisatorische Anforderungen halten. Dabei gilt es, die bewährten Strukturen des Medienunternehmens nicht grundlegend zu verändern. Die Aufgabenverteilung in Form des Newsroom-Prinzips soll erhalten bleiben. Die Ablauforganisation, also die zeitlichen Arbeitsabläufe, verändern sich dadurch, dass die KI Aufgaben schneller und parallel bearbeiten kann und somit für eine Zeitersparnis sorgt. Als Konsequenz daraus werden sich die Fähigkeitsprofile der Mitarbeiterinnen und Mitarbeiter ändern. Über Maßnahmen zur Personalentwicklung werden daher Kenntnisse in der Datenwissenschaft an bestehende Mitarbeiterinnen und Mitarbeiter vermittelt sowie neues Personal mit den gewünschten Fähigkeiten akquiriert. Trotz der Arbeitserleichterung durch das KI-Modell ist und bleibt eine menschliche Arbeit und Aufsicht im Workflow sehr wichtig.

ChatGPT ist unter Voraussetzungen, die in dieser Thesis erläutert wurden, dazu geeignet, im Newsworkflow eingesetzt zu werden. Es ist theoretisch technisch möglich, dieses KI-Modell in den Newsworkflow einzubinden. Der Einführungsprozess sowie der spätere produktive Einsatz sind iterative Prozesse. Die einzelnen Prozessschritte ergeben sich aus den Erfahrungen während des Projekts und dem Feedback der Anwenderinnen und Anwender. Die Voraussetzungen, unter denen ChatGPT sicher und gewinnbringend in den Newsworkflow integriert wird, sind zunächst nicht-technischer Natur. Sie umfassen Aspekte von neuen Rollenbildern über Maßnahmen, um die rechtlich-ethischen Anforderungen einzuhalten, bis hin zu Feedback und Havarieszenarien.

Die technische Lösung trägt die Konzepte. Durch eine zusätzliche Automation wird ein Zwischenglied zwischen ChatGPT und die Nutzerin bzw. den Nutzer eingefügt. Dieser sorgt für die Weitergabe von Informationen, Sicherheit und den Anstoß von automatischen Workflows. Als

¹⁵ Die Autorin weist nochmals ausdrücklich darauf hin, dass sich der Stand des Fazits auf den Zeitpunkt der Erstellung der Thesis bezieht.

Vermittlung zwischen ChatGPT, Nutzer*in und MAM wird die Automationssoftware Node-RED empfohlen, da diese workflow-basiert arbeitet und mit einem grafischen Editor bedient wird. Um die Qualität der Antworten von ChatGPT zu verbessern, wird Prompt Engineering als Verfahren vorgeschlagen.

Mit dieser Lösung wird ungefähr die Hälfte der Use Cases in der Anforderungsmatrix umgesetzt. Um weitere Use Cases zu realisieren, eignen sich weitere Modelle von OpenAI für den Einsatz im Newsworkflow. Als Ergänzung zu ChatGPT decken die Modelle aus der Gruppe InstructGPT, DALL-E und Whisper mehr Use Cases ab. Sie umfassen damit auch die Verarbeitung von Bild und Ton. Für die Implementierung dieser weiteren Modelle werden Ergänzungen zu den bestehenden Konzepten und technischen Lösungen benötigt. Diese beschreiben jetzt auch das Finetuning, eine Methode, um dem KI-Modell hausgemachte Daten als Trainingsmaterial zu geben. Dadurch kann das KI-Modell den Stil und die Haltung des Medienunternehmens erlernen und beispielsweise in generierten Texten ausgeben. Trotzdem deckt auch die Kombination aus den Modellen nicht alle Use Cases ab. Es fehlt, um nur ein Beispiel zu nennen, die Verschlagwortung von Bildern und Videos anhand der visuellen Komponente.

Schließlich ist KI eine Technologie, die in den letzten Jahren eine rasante Entwicklung durchgemacht hat. Sie hat die Fähigkeit, eine hohe Effizienzsteigerung in Prozessen zu ermöglichen. Dennoch bedarf sie noch immer einer strengen Regulierung und Aufsicht durch Menschen.

Zukunftsausblick

Die Rolle des Menschen gegenüber der KI und auch gegenüber der Arbeitswelt generell bewegt sich von der ausführenden Rolle hin zur beauftragenden und überwachenden Rolle [121]. Im Journalismus und anderen Anwendungsfeldern wird die menschliche Arbeit nicht automatisch weniger wert sein, nur weil eine KI diese eventuell schneller ausführen kann. Als Analogie kann der Vergleich zwischen industriell gefertigten und handgefertigten Produkten verwendet werden. Die industrielle Produktion ist effizienter und schneller, trotzdem wird das Handwerk höher geschätzt und auch als höherwertiger gesehen [122]. Diese Ausprägung gilt es zu stärken und rechtlich für jede Anwenderin und jeden Anwender deutlich zu machen, damit das eigene selbstverantwortliche Handeln eines Menschen nicht auf reine Aufgabenerfüllung reduziert wird und damit missbräuchliche Aspekte einer KI die Oberhand bekämen. KI kann den Menschen nicht ersetzen, nur unterstützen.

Für den letzten Abschnitt soll nun auch der Hauptakteur ChatGPT zu Wort kommen. Der Prompt zur Erstellung des Textes lautet „Schreibe ein Statement zum Zukunftsausblick von KI in zehn Sätzen“:

„Die Zukunftsaussichten für KI sind äußerst vielversprechend. Wir werden eine zunehmende Integration von KI in verschiedenen Lebensbereichen erleben, von der Gesundheitsversorgung bis hin zur Mobilität. KI wird eine Schlüsselrolle bei der Bewältigung komplexer Herausforderungen wie dem Klimawandel und der Lösung gesellschaftlicher Probleme spielen. Neue Technologien wie

Machine Learning und Neuronale Netze werden zu immer leistungsfähigeren KI-Systemen führen, die in der Lage sind, menschenähnliche Denk- und Lernfähigkeiten zu entwickeln. Mit der Weiterentwicklung von KI wird auch die Frage der Ethik und Verantwortung im Umgang mit diesen Technologien an Bedeutung gewinnen. Die KI-Entwicklung wird zu einer verstärkten Automatisierung führen, was wiederum zu Veränderungen in der Arbeitswelt führen kann. Dennoch werden neue Arbeitsmöglichkeiten im Bereich der KI-Entwicklung und -Wartung entstehen. Die Zusammenarbeit zwischen Mensch und Maschine wird zunehmen, und KI-Systeme werden als Werkzeuge dienen, um menschliche Fähigkeiten zu erweitern. KI wird auch weiterhin eine wichtige Rolle bei der Datenanalyse und dem Erkennen von Mustern spielen, um fundierte Entscheidungen zu treffen und neue Erkenntnisse zu gewinnen. Insgesamt wird KI unser tägliches Leben revolutionieren und zu einer effizienteren, intelligenten und vernetzten Gesellschaft beitragen.“
[123]

Die Zukunftsprognose stimmt mit der der Autorin überein. Die ersten revolutionären Entwicklungen auf dem Gebiet der KI sind mit den Sprachmodellen in den letzten Jahren geschehen. Jetzt liegt es an Unternehmen und Regierungen als Early Adopters, diese neue Technologie für ihre Anwendungsfälle gewinnbringend einzusetzen. Als Schritt in diese Richtung schaffen die Vorschläge und Konzepte in der Thesis einen Rahmen für die erfolgreiche Zusammenarbeit von Mensch und Maschine.

11 VERZEICHNISSE

11.1 ABKÜRZUNGSVERZEICHNIS

AGG	Allgemeines Gleichbehandlungsgesetz
AI	Artificial Intelligence
API	Application Programming Interface
CISO	Chief Information Security Officer
CMS	Content Management System
DSGVO	Datenschutzgrundverordnung
EU	Europäische Union
GG	Grundgesetz der Bundesrepublik Deutschland
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
ID	Identifikation
IT	Informationstechnik / Information Technology
JSON	JavaScript Object Notation
KI	Künstliche Intelligenz
LLM	Large Language Model
MAM	Media Asset Management
ML	Machine Learning / Maschinelles Lernen
MU	Medienunternehmen
NN	Neuronales Netz / Neural Network
REST	Representational State Transfer
RNN	Rekurrentes Neuronales Netz / Recurrent Neural Network
SLA	Service Level Agreement
SMPTE	Society of Motion Picture and Television Engineers
UI	User Interface
UMID	Universal Material Identifier
URI	Uniform Resource Identifier
XML	Extensible Markup Language

11.2 ABBILDUNGSVERZEICHNIS

Abbildung 1: Neuronales Netz [4].	5
Abbildung 2: Rekurrentes Neuronales Netz [4].	6
Abbildung 3: Aufbau eines Transformers [4].	8
Abbildung 4: Anordnung und Beziehung von Wörter im Vektorraum [9, S. 35].	8
Abbildung 5: Darstellung der Entscheidungsgrenze bei Overfitting [10].	9
Abbildung 6: Zeitleiste der LLMs [5].	11
Abbildung 7: Vorgehensweise beim Training von ChatGPT [18].	13
Abbildung 8: Content Life Cycle in einem Medienunternehmen [22, S. 123].	21
Abbildung 9: Evolution von Voreingenommenheit in den Medien [47, S. 394].	50
Abbildung 10: Kategorien der Datenqualität mit Beispielen nach [57, S. 20, 58].	54
Abbildung 11: Kriterien für Datenqualität [59].	55
Abbildung 12: Kreislauf der Content-Wirtschaft [25, S. 33].	60
Abbildung 13: Metriken und Messmethoden der Personalisierung nach [74, S. 52] und [74, S. 54].	66
Abbildung 14: Menschzentrierte iterative Gestaltung eines Produkts [75].	67
Abbildung 15: Newsroom-Organisation [77, S. 141].	70
Abbildung 16: Modellhafter Ressourcenbedarf für eine tägliche viertelstündige Nachrichtensendung nach [77, S. 68].	72
Abbildung 17: Zielgruppen des Schulungskonzepts [Eigene Darstellung].	91
Abbildung 18: Systemaufbau mit ChatGPT nach [84, S. 51].	105
Abbildung 19: Beispiel für eine Message in JavaScript Object Notation (JSON) nach [88].	110
Abbildung 20: Node-Typen in Node-RED nach [89].	110
Abbildung 21: Komplexerer Request mit Parametern in Node-RED [93].	111
Abbildung 22: Inhalt des Function-Nodes in Python nach [93] und [94].	112
Abbildung 23: JSON-Response wird erhalten in Node-RED [95].	113
Abbildung 24: Visualisierung des Function Callings [97].	114
Abbildung 25: Verbindung von UMID zu Essenz [100, S. 7].	115
Abbildung 26: SMPTE Metadata Dictionary Klassenstruktur [99, S. 7].	116
Abbildung 27: Antwort von ChatGPT in JSON [94].	117
Abbildung 28: hmac-Node in Node-RED [102].	119
Abbildung 29: Beispiel für Prompt Engineering [103, S. 8, 103, S. 7].	120
Abbildung 30: Kompatibilität der Modelle mit den jeweiligen Endpunkten nach [81].	132
Abbildung 31: Beispiel einer Antwort auf ein Embedding [94].	133
Abbildung 32: Data Science Life Cycle [111].	135
Abbildung 33: Typische Datenfehler [112].	136

11.3 LITERATUR

- [1] C. Beckett, „New powers, new responsibilities: A global survey of journalism and artificial intelligence“, Nov. 2019. Verfügbar unter: https://www.ansa.it/documents/1583854311829_Ricerca_Journalism_AI.pdf. Zugriff am: 6. April 2023.
- [2] S. Rothe, „1.1 Einführung (Begriffe KI, Machine Learning, Deep Learning)“. [Online]. Verfügbar unter: <https://www.youtube.com/watch?v=8xQW2Q5UXog>
- [3] S. Rothe, „1.2 Neuronale Netze (Was ist ein Neuronales Netz? Wie funktioniert es?)“. [Online]. Verfügbar unter: <https://www.youtube.com/watch?v=oc3lPa9Ld8o>
- [4] M. Bischoff, „Wie man einem Computer das Sprechen beibringt“, *Spektrum.de*, 9. März 2023, 2023. [Online]. Verfügbar unter: <https://www.spektrum.de/news/wie-funktionieren-sprachmodelle-wie-chatgpt/2115924>. Zugriff am: 28. März 2023.
- [5] S. Rothe, „2.2 Überblick über Große Sprachmodelle“. [Online]. Verfügbar unter: https://www.youtube.com/watch?v=xBrx_U3UI_M
- [6] M. Ruby, „How ChatGPT Works: The Model Behind The Bot - Towards Data Science“, *Towards Data Science*, 30. Jan. 2023, 2023. [Online]. Verfügbar unter: <https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286>. Zugriff am: 27. März 2023.
- [7] *ChatGPT and friends - wie KI Medien und Gesellschaft verändert*, This is Media NOW - der Podcast der MEDIENTAGE MÜNCHEN. Medien.Bayern GmbH, Jg. 101. Verfügbar unter: <https://thisismedianow.podigee.io/103-folge-101-chatgpt-and-friends-wie-ki-medien-und-gesellschaft-verandert>
- [8] B. Probiez, J. Kozak und A. Hrabia, „Clustering of scientific articles using natural language processing“, *Procedia Computer Science*, Jg. 207, S. 3443–3452, 2022, doi: 10.1016/j.procs.2022.09.403.
- [9] A. Rossner *et al.*, „ASSET MATCH: DATENGESTÜTZTE ENTSCHEIDUNGSHILFEN FÜR EMPFEHLUNGEN IN MEDIENUNTERNEHMEN“, *FKTG*, Nr. 11, S. 33–38, 2022.
- [10] A. Dawari, „Overfitting: Causes and Remedies“, *Towards AI*, 3. Aug. 2022, 2022. [Online]. Verfügbar unter: <https://towardsai.net/p/1/overfitting-causes-and-remedies>. Zugriff am: 6. Juni 2023.
- [11] T. Taulli, *Grundlagen der Künstlichen Intelligenz: Eine nichttechnische Einführung*, 1. Aufl. Berlin, Heidelberg: Springer Berlin Heidelberg, 2022.
- [12] D. Ruby, „ChatGPT Statistics for 2023: Comprehensive Facts and Data“, *Demand Sage*, 8. Feb. 2023, 2023. [Online]. Verfügbar unter: <https://www.demandsage.com/chatgpt-statistics/>. Zugriff am: 15. März 2023.
- [13] *ChatGPT General FAQ | OpenAI Help Center*. [Online]. Verfügbar unter: <https://help.openai.com/en/articles/6783457-chatgpt-general-faq> (Zugriff am: 17. März 2023).
- [14] Nerdy Nav, *91 Important ChatGPT Statistics & Facts For March 2023 (Gpt-4, ChatGPT Plugins Update) - Nerdy Nav*. [Online]. Verfügbar unter: <https://nerdynav.com/chatgpt-statistics/#usercount> (Zugriff am: 27. März 2023).
- [15] S. Frieder *et al.*, *Mathematical Capabilities of ChatGPT*. arXiv, 2023. Verfügbar unter: <https://arxiv.org/pdf/2301.13867.pdf>.
- [16] *ChatGPT plugins*. [Online]. Verfügbar unter: <https://openai.com/blog/chatgpt-plugins> (Zugriff am: 8. Mai 2023).
- [17] M. Ramponi, „How ChatGPT actually works“, *News, Tutorials, AI Research*, 23. Dez. 2022, 2022. [Online]. Verfügbar unter: <https://www.assemblyai.com/blog/how-chatgpt-actually-works/>. Zugriff am: 27. März 2023.
- [18] *Introducing ChatGPT*. [Online]. Verfügbar unter: <https://openai.com/blog/chatgpt> (Zugriff am: 27. März 2023).
- [19] G. Hooffacker und K. Meier, *La Roches Einführung in den praktischen Journalismus*. Wiesbaden: Springer Fachmedien Wiesbaden, 2017.

- [20] A. Buchholz und G. Schult, *Fernseh-Journalismus*. Wiesbaden: Springer Fachmedien Wiesbaden, 2016.
- [21] G. Hooffacker, *Online-Journalismus*. Wiesbaden: Springer Fachmedien Wiesbaden, 2020.
- [22] S. Spörrer, *Content Management Systeme*. Wiesbaden: Springer Fachmedien Wiesbaden, 2009.
- [23] M. Christmann, „Vorlesung Studioteknik: Kapitel IX. Broadcast IT“, Hochschule RheinMain, 2020.
- [24] M. Kaiser, *Journalistische Praxis: Automatisierte Kommunikation Im Journalismus und in der Public Relation*. Wiesbaden: Vieweg, 2019.
- [25] A. Rossner, C. Lellinger, V. Denkel, J. Köhler, P. Schramm und S. Pagel, „AI4MEDIADATA: VON DATA THINKING ZU DATEN-GESTÜTZTEN ENTSCHEIDUNGSHILFEN“, *FKTG*, Nr. 10, S. 30–36, 2022.
- [26] Deutscher Bundestag, „Unterrichtung der Enquete-Kommission Künstliche Intelligenz: Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale“, 28. Okt. 2020. [Online]. Verfügbar unter: <https://dserver.bundestag.de/btd/19/237/1923700.pdf>. Zugriff am: 20. April 2023.
- [27] „Grundgesetz für die Bundesrepublik Deutschland: GG“ in *Gesetze im Internet*. [Online]. Verfügbar unter: <https://www.gesetze-im-internet.de/gg/BJNR000010949.html>
- [28] „KÜNSTLICHE INTELLIGENZ IN REDAKTIONEN: IST JOURNALISMUS BERECHENBAR?“, *tbd - der Debattenmonitor*, Nr. 3, 2021. Verfügbar unter: https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Themen/Intermediaere/tbd-Debattenmonitor/Ausgabe3_Debattenradar_04_0821.pdf.
- [29] *Publizistische Grundsätze (Pressekodex)*, 1973. [Online]. Verfügbar unter: <https://www.presserat.de/pressekodex.html>
- [30] *Medienstaatsvertrag: MStV*, 2022. [Online]. Verfügbar unter: <https://www.die-medienanstalten.de/service/rechtsgrundlagen/gesetze>
- [31] P. Hacker, F. Naumann, T. Friedrich, S. Grundmann, A. Lehmann und H. Zech, „AI Compliance – Challenges of Bridging Data Science and Law“, *J. Data and Information Quality*, Jg. 14, Nr. 3, S. 1–4, März 2022, doi: 10.1145/3531532.
- [32] P. Office, „Datenschutz-Grundverordnung: DSGVO“ in *Amtsblatt der Europäischen Union*, 2016. [Online]. Verfügbar unter: <https://dsgvo-gesetz.de/>
- [33] F. Prof. Dr. Dr. Rostalski, „1.2.2 Die juristische Perspektive“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023>
- [34] C. Armbrüster, „Digitalisierung und Nachhaltigkeit – rechtliche Herausforderungen für den Versicherungssektor, insbesondere beim Einsatz von Künstlicher Intelligenz“, *ZVersWiss*, Jg. 111, Nr. 1, S. 19–31, 2022, doi: 10.1007/s12297-022-00518-3.
- [35] „Gesetz über Urheberrecht und verwandte Schutzrechte: UrhG“ in *Gesetze im Internet*, 1965. [Online]. Verfügbar unter: <https://www.gesetze-im-internet.de/urhg/BJNR012730965.html#BJNR012730965BJNG000101377>
- [36] *Gesetz betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie: KunstUrhG*, 1907. [Online]. Verfügbar unter: <https://www.gesetze-im-internet.de/kunsturhg/BJNR000070907.html>
- [37] K. Prof. Dr. Goldhammer, K. Dieterich und T. Prien, „Künstliche Intelligenz, Medien und Öffentlichkeit: Wissenschaftlicher Bericht“, Berlin/Bern, Apr. 2019. Verfügbar unter: https://www.bakom.admin.ch/dam/bakom/de/dokumente/bakom/elektronische_medien/Zahlen%20und%20Fakten/Studien/studien-kuenstliche-intelligenz-medien-oeffentlichkeit.pdf.download.pdf/K%C3%BCnstliche%20Intelligenz,%20Medien%20und%20%C3%96ffentlichkeit.pdf. Zugriff am: 5. April 2023.
- [38] D. Kunschke, M. F. Spitz und J. Pohle, „Regulierungsrahmen für Künstliche Intelligenz“ in *FinTech: Digitalisierung, Künstliche Intelligenz und aufsichtsrechtliche Regulierung von Finanzdienstleistungen*, D. Kunschke, M. F. Spitz und J. Pohle, Hg., Berlin: Erich Schmidt Verlag GmbH & Co. KG, 2022, S. 41–88, doi: 10.37307/b.978-3-503-20689-6.02.

- [39] M. Kaiser, *Studie zu ethischen Aspekten im Roboterjournalismus*. [Online]. Verfügbar unter: <https://medien-kaiser.de/2020/02/22/studie-zu-ethischen-aspekten-im-roboterjournalismus/> (Zugriff am: 21. April 2023).
- [40] Bundesamt für Sicherheit in der Informationstechnik, „Große KI-Sprachmodelle - Chancen und Risiken für Industrie und Behörden“, 3. Mai 2023. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Grosse_KI_Sprachmodelle.pdf?__blob=publicationFile&v=2.
- [41] *Gesetz über Künstliche Intelligenz*, 2021. [Online]. Verfügbar unter: <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52021PC0206&from=FR>
- [42] Bundesamt für Sicherheit in der Informationstechnik, „Sicherer, robuster und nachvollziehbarer Einsatz von KI - Probleme, Maßnahmen und Handlungsbedarfe“. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Herausforderungen_und_Massnahmen_KI.pdf?__blob=publicationFile&v=6. Zugriff am: 6. Juni 2023.
- [43] *Kitqar - Glossar zur Datenqualität*. [Online]. Verfügbar unter: <https://www.kitqar.de/de/veroeffentlichungen/dq-glossar> (Zugriff am: 1. März 2023).
- [44] Bundesamt für Sicherheit in der Informationstechnik, „AI Cloud Service Compliance Criteria Catalogue (AIC4)“, 2021. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/CloudComputing/AIC4/AI-Cloud-Service-Compliance-Criteria-Catalogue_AIC4.pdf?__blob=publicationFile&v=4.
- [45] EBU, Hg., „R 144: Cybersecurity Governance for Media Companies“, Genf, Sep. 2016. [Online]. Verfügbar unter: <https://tech.ebu.ch/docs/r/r144.pdf>. Zugriff am: 22. Mai 2023.
- [46] F. Doshi-Velez und B. Kim, „Towards A Rigorous Science of Interpretable Machine Learning“, 28. Feb. 2017. [Online]. Verfügbar unter: <https://arxiv.org/pdf/1702.08608>.
- [47] F. Hamborg, K. Donnay und B. Gipp, „Automated identification of media bias in news articles: an interdisciplinary literature review“, *Int J Digit Libr*, Jg. 20, Nr. 4, S. 391–415, 2019, doi: 10.1007/s00799-018-0261-y.
- [48] J. PD Dr. Heesen, „1.2.7 Wie werden Daten beschriftet?“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023>
- [49] J. Gruenewald, J. Pizarro und S. M. Chermak, „Race, gender, and the newsworthiness of homicide incidents“, *Journal of Criminal Justice*, Jg. 37, Nr. 3, S. 262–272, 2009, doi: 10.1016/j.jcrimjus.2009.04.006.
- [50] Datenethikkommission der Bundesregierung, „Gutachten der Datenethikkommission“, Okt. 2019. [Online]. Verfügbar unter: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6. Zugriff am: 19. April 2023.
- [51] J. N. Druckman und M. Parkin, „The Impact of Media Bias: How Editorial Slant Affects Voters“, *The Journal of Politics*, Jg. 67, Nr. 4, S. 1030–1049, 2005, doi: 10.1111/j.1468-2508.2005.00349.x.
- [52] J.-M. Eberl, M. Wagner und H. G. Boomgaarden, „Are Perceptions of Candidate Traits Shaped by the Media? The Effects of Three Types of Media Bias“, *The International Journal of Press/Politics*, Jg. 22, Nr. 1, S. 111–132, 2017, doi: 10.1177/1940161216674651.
- [53] F. Prof. Dr. Rostalski, „2.2.3 Mögliche Rechtsfolgen minderer Datenqualität“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023>.
- [54] *Allgemeines Gleichbehandlungsgesetz: AGG*, 2006. [Online]. Verfügbar unter: <https://www.gesetze-im-internet.de/agg/BJNR189710006.html>
- [55] B. Stegmann, *TAB - Themen und Projekte - Projekteübersicht - Algorithmen in digitalen Medien und ihr Einfluss auf die Meinungsbildung*. [Online]. Verfügbar unter: https://www.tab-beim-bundestag.de/projekte_algorithmen-in-digitalen-medien-und-ihr-einfluss-auf-die-meinungsbildung.php (Zugriff am: 5. April 2023).
- [56] C. Prof. Dr. Bieber, J. PD Dr. Heesen, A. Prof. Dr. Lauber-Rönsberg, C. Prof. Dr. Neuberger, C. Prof. Dipl.-Journ. Elmer und T. Hühnert, *Künstliche Intelligenz und Journalismus*:

- Potenziale und Herausforderungen für Medienschaffende*, 2023. Verfügbar unter: https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG3_WP_KI_und_Journalismus.pdf.
- [57] R. Y. Wang und D. M. Strong, „Beyond Accuracy: What Data Quality Means to Data Consumers“, *Journal of Management Information Systems*, Jg. 12, Nr. 4, S. 5–23, 1996. [Online]. Verfügbar unter: <https://www.jstor.org/stable/40398176?origin=JSTOR-pdf>
- [58] F. Prof. Naumann, „1.2.1 Dimensionen der Datenqualität“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023>
- [59] F. Prof. Naumann, „1.2.4 Datenqualität messen“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023>
- [60] S. Dr. Hallersleben, „1.2.3 Was ist eigentlich Normung?“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023>
- [61] Center for Advanced Internet Studies, Heinrich Heine Universität Düsseldorf, Hg., „Künstliche Intelligenz im Journalismus: Wie nimmt die Bevölkerung den Einfluss von Künstlicher Intelligenz auf die Zukunft des Journalismus wahr?“, *Meinungsmonitor Künstliche Intelligenz* 4, Juli 2022. Verfügbar unter: <https://www.cais-research.de/wp-content/uploads/Factsheet-4-Journalismus.pdf>. Zugriff am: 11. April 2023.
- [62] B. Gruber, *Fakten, Falschmeldungen, Zahlen*. [Online]. Verfügbar unter: <https://www.goethe.de/prj/k40/de/lan/aij.html> (Zugriff am: 5. April 2023).
- [63] *Gesetz über digitale Dienste*. [Online]. Verfügbar unter: <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32022R2065&from=en>
- [64] Bundesministerium der Justiz, *Telemediengesetz: TMG*, 2007. [Online]. Verfügbar unter: <https://www.gesetze-im-internet.de/tmg/BJNR017910007.html>
- [65] Europäisches Parlament, *Regelung der zivilrechtlichen Haftung beim Einsatz künstlicher Intelligenz - Dienstag, 20. Oktober 2020*. [Online]. Verfügbar unter: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_DE.html (Zugriff am: 18. April 2023).
- [66] S. Rothe, „1.0 KI-nema: KI und Film“. [Online]. Verfügbar unter: <https://www.youtube.com/watch?v=-sLLnJTzfxU>
- [67] D. B. Wilczek und P. D. M. Haim, „Wie kann Künstliche Intelligenz die Effizienz von Medienorganisationen steigern?: Eine Systematisierung entlang der Nachrichtenwertkette mit besonderer Berücksichtigung lokaler und regionaler Medien“, *Medienwirtschaft*, Jg. 19, Nr. 4, S. 44–50, 2022.
- [68] *re:publica 2021 - WDR präsentiert KI-Prototypen und diskutiert über die Zukunft synthetischer Medien - Presselounge - WDR*. [Online]. Verfügbar unter: https://presse.wdr.de/plounge/wdr/programm/2021/05/20210520_synthetische_medien_republica.html#goToHead (Zugriff am: 5. April 2023).
- [69] M. Kaiser, „Visual Data Mining – Möglichkeiten und Grenzen einer KI-gestützten Verschlagwortung von Videoarchiven – betrachtet am Beispiel von DeepVA“. Bachelorthesis, Informationstechnologie und Elektrotechnik, Hochschule RheinMain, Wiesbaden/Rüsselsheim, 2022.
- [70] G. Voigt-Müller, *ChatGPT für Medien-Workflows - film-tv-video.de*. [Online]. Verfügbar unter: <https://www.film-tv-video.de/equipment/2023/01/30/chatgpt-fuer-medien-workflows/> (Zugriff am: 1. März 2023).
- [71] A. Rossner *et al.*, „HIGHLIGHT SEARCH: IDENTIFIZIERUNG EMOTIONALER HIGHLIGHTS IN REALITY-TV-FORMATEN“, *FKTG*, Nr. 12, S. 30–35, 2022.
- [72] M. Grotticelli, „Streamlines Multiplatform AI-Enabled MAM Production“, *FKTG*, Nr. 1, S. 14–15, 2023.
- [73] kress, *Anwendungen, die Medienprofis kennen sollten: So setzt die FAZ Künstliche Intelligenz ein*. [Online]. Verfügbar unter: <https://kress.de/news/detail/beitrag/149293-anwendungen-die-medienprofis-kennen-sollten-so-setzt-die-faz-kuenstliche-intelligenz-ein.html> (Zugriff am: 5. April 2023).

- [74] A. Jääskeläinen und M. Olij, *News Report 2019: The Next Newsroom*. [Online]. Verfügbar unter: https://www.amic.media/media/files/file_352_2172.pdf (Zugriff am: 16. Mai 2023).
- [75] *Menschzentrierte Gestaltung nach DIN EN ISO 9241-210 | German UPA*. [Online]. Verfügbar unter: <https://germanupa.de/wissen/berufsbild-usability-ux-professional/grundlegend/menschzentrierte-gestaltung> (Zugriff am: 6. Juni 2023).
- [76] Bundesministerium des Innern, für Bau und Heimat und Bundesverwaltungsamt, „Organisationshandbuch“. [Online]. Verfügbar unter: <https://www.verwaltung-innovativ.de/OHB/DE/node.html>. Zugriff am: 24. April 2023.
- [77] S. Hillebrecht, *Personalwirtschaft der Medienunternehmen*, 1. Aufl. Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 2018.
- [78] *OpenAI API Tokenizer*. [Online]. Verfügbar unter: <https://platform.openai.com/tokenizer> (Zugriff am: 18. Juni 2023).
- [79] M. Kaiser, „Gegenüberstellung von ChatGPT und Luminous“. Dokumentation zur Projektarbeit, Informationstechnologie und Elektrotechnik, Hochschule RheinMain, Wiesbaden/Rüsselsheim, 2023.
- [80] *Terms of use*. [Online]. Verfügbar unter: <https://openai.com/policies/terms-of-use> (Zugriff am: 11. Mai 2023).
- [81] *OpenAI API Overview*. [Online]. Verfügbar unter: <https://platform.openai.com/docs/introduction/overview> (Zugriff am: 11. Mai 2023).
- [82] S. Simli, *How To Effectively Track Progress In Agile Projects?* [Online]. Verfügbar unter: <https://www.nimblework.com/blog/agile-project-metrics/> (Zugriff am: 30. Mai 2023).
- [83] *Pricing*. [Online]. Verfügbar unter: <https://openai.com/pricing> (Zugriff am: 26. Mai 2023).
- [84] P. Dorn, „Herausforderungen und Lösungsansätze für eine Digitalisierungsstrategie in der Stadtentwicklung mithilfe von Broadcast-IT-Komponenten“. Masterthesis, Ingenieurwissenschaften, Hochschule RheinMain, Wiesbaden/Rüsselsheim, 2022.
- [85] *Node-RED*. [Online]. Verfügbar unter: <https://nodered.org/> (Zugriff am: 12. Juni 2023).
- [86] M. Kaiser, „MEDIA ASSET MANAGEMENT, ESSENZ MANAGEMENT UND CONTENT MANAGEMENT SYSTEME IM BROADCAST-UND PRODUKTIONSUMFELD“. Hausarbeit, Hochschule RheinMain, Wiesbaden/Rüsselsheim, 2021.
- [87] D. Austerberry, „Workflow management“, *TV Tech*, 1. Mai 2008, 2008. [Online]. Verfügbar unter: <https://www.tvtechnology.com/opinions/workflow-management>. Zugriff am: 14. Juni 2023.
- [88] *Working with messages : Node-RED*. [Online]. Verfügbar unter: <https://nodered.org/docs/user-guide/messages> (Zugriff am: 12. Juni 2023).
- [89] *The Core Nodes : Node-RED*. [Online]. Verfügbar unter: <https://nodered.org/docs/user-guide/nodes> (Zugriff am: 12. Juni 2023).
- [90] T. Wu et al., *PromptChainer: Chaining Large Language Model Prompts through Visual Programming*, 2022. [Online]. Verfügbar unter: <https://arxiv.org/pdf/2203.06566>
- [91] P. Bichiou, „Webbasierte Anwendungen“, 24. Apr. 2020.
- [92] Postman, *OpenAI | OpenAI*. [Online]. Verfügbar unter: <https://www.postman.com/devrel/workspace/openai/documentation/13183464-90abb798-cb85-43cb-ba3a-ae7941e968da> (Zugriff am: 20. Juni 2023).
- [93] *Set a request header : Node-RED*. [Online]. Verfügbar unter: <https://cookbook.nodered.org/http/set-request-header> (Zugriff am: 20. Juni 2023).
- [94] *Open AI API Reference*. [Online]. Verfügbar unter: <https://platform.openai.com/docs/api-reference> (Zugriff am: 15. Mai 2023).
- [95] *Get a parsed JSON Response : Node-RED*. [Online]. Verfügbar unter: <https://cookbook.nodered.org/http/parse-json-response> (Zugriff am: 21. Juni 2023).
- [96] *Open AI API Function calling*. [Online]. Verfügbar unter: <https://platform.openai.com/docs/guides/gpt/function-calling> (Zugriff am: 20. Juni 2023).
- [97] J. Hötter und C. Warmuth, „2.2 LLM als Puzzleteil“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/6yxDDL1kb3cjeALSfwIVRc>

- [98] *OpenAI Platform Rate Limits*. [Online]. Verfügbar unter: <https://platform.openai.com/docs/guides/rate-limits/overview> (Zugriff am: 21. Juni 2023).
- [99] *Metadata Element Dictionary Structure*, ST 335:2012, SMPTE, 2012. [Online]. Verfügbar unter: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7291959>
- [100] *Application of Unique Material Identifiers in Production and Broadcast Environments*, RP 0205-2009, SMPTE, 2009. [Online]. Verfügbar unter: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7290981>
- [101] P. D. J. Winter, „Vorlesung Secure Networking: Kapitel 4 Authentifikation und Digitale Signatur“, Hochschule RheinMain, 2022. Zugriff am: 27. Juni 2023.
- [102] *node-red-contrib-crypto-js*. [Online]. Verfügbar unter: <https://flows.nodered.org/node/node-red-contrib-crypto-js> (Zugriff am: 27. Juni 2023).
- [103] J. Hötter und C. Warmuth, „1.10 Prompt Engineering und Vector Datenbanken“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/51ldVWk41rYbGQgMr6vAQ3>
- [104] J. Hötter und C. Warmuth, „1.5 Wege nach Rom - Grundlagenausblick“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/5iSrPROVhLEg8ISTyaGpMy>
- [105] *OpenAI API: GPT Best Practices*. [Online]. Verfügbar unter: <https://platform.openai.com/docs/guides/gpt-best-practices/six-strategies-for-getting-better-results> (Zugriff am: 19. Juni 2023).
- [106] J. Hötter und C. Warmuth, „2.8 Postfächer“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/4DYha911kTdHW4BHjPbI6w>
- [107] J. Hötter und C. Warmuth, „2.6.2 Conversational AI (2)“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/6UUPLVv7hdnrmPw6bUaGD3>
- [108] *Introducing ChatGPT and Whisper APIs*. [Online]. Verfügbar unter: <https://openai.com/blog/introducing-chatgpt-and-whisper-apis> (Zugriff am: 24. Mai 2023).
- [109] M. Rogati, „The AI Hierarchy of Needs“, 12. Juni 2017, 2017. [Online]. Verfügbar unter: <https://hackernoon.com/the-ai-hierarchy-of-needs-18f11fcc007>. Zugriff am: 6. April 2023.
- [110] M. Elhayany und H. Steinbeck, *Overview of the Data Science life-cycle | Data Science Bootcamp* |. [Online]. Verfügbar unter: <https://open.hpi.de/courses/datascience2023/items/7pTarBSX8YpQirSBgDBNOa> (Zugriff am: 9. Juni 2023).
- [111] S. Agarwal, *Understanding the Data Science Lifecycle*. [Online]. Verfügbar unter: <https://www.sudeep.co/data-science/2018/02/09/Understanding-the-Data-Science-Lifecycle.html> (Zugriff am: 15. Juni 2023).
- [112] F. Prof. Naumann, „2.2.8 Datenreinigung“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kidaten2023/items/2RkoZqGVrNcaIJE4rkYfWO>
- [113] M. Elhayany und H. Steinbeck, *Techniques for Handling Missing Values*. [Online]. Verfügbar unter: <https://open.hpi.de/courses/datascience2023/items/slHCBGdD42oOoCkrgfHXv> (Zugriff am: 15. Juni 2023).
- [114] S. Altman, „A conversation with OpenAI's CEO, Sam Altman“, 25. Mai 2023. [Online]. Verfügbar unter: <https://www.youtube.com/watch?v=uaQZIK9gvNo>
- [115] S. Rothe, „2.1 ChatGPT“. [Online]. Verfügbar unter: https://www.youtube.com/watch?v=G2MnIOiR_Ds
- [116] J. Hötter und C. Warmuth, „3.10 Monopolstellung und die Rolle von Europa“. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/2Ql8xLIi3x6BsE5k6vXVAP>
- [117] S. Rothe, „2.3 KI als Co-Autor?“. [Online]. Verfügbar unter: <https://www.youtube.com/watch?v=shN4Ynsuw1A>
- [118] Akademie für Künstliche Intelligenz AKI gGmbH im KI Bundesverband, „Große KI-Modelle für Deutschland“, 2023. [Online]. Verfügbar unter: <https://leam.ai/feasibility-study-leam-2023/>. Zugriff am: 25. April 2023.
- [119] DeepVA, „Open Source auf dem Vormarsch – wie wir uns mit DeepVA diesem Wandel stellen!“, *DeepVA*, 13. Juni 2023, 2023. [Online]. Verfügbar unter:

- <https://deepva.ai/de/open-source-auf-dem-vormarsch-wie-wir-uns-mit-deepva-diesem-wandel-stellen/>. Zugriff am: 20. Juni 2023.
- [120] S. Anderl, *KI im Journalismus*. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/6CLRpNM2nEYMj5cLTguOsR>. Zugriff am: 7. Juli 2023.
- [121] S. Wieczorek, *Zukunft von Unternehmen mit generativer KI*. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/42J9K3zNS9n5fIc8FzgOeK>. Zugriff am: 8. Juli 2023.
- [122] J. Hötter und C. Warmuth, „4.9 Die Rolle menschlicher Arbeit“, 8. Juli 2023. [Online]. Verfügbar unter: <https://open.hpi.de/courses/kizukunft2023/items/19EDkAfuB6SM8ObuvegrcL>
- [123] ChatGPT, *Statement zum Zukunftsausblick*. [Online]. Verfügbar unter: <https://chat.openai.com/share/od2345d8-5ab8-40ff-b517-ac46115413c4> (Zugriff am: 10. Juli 2023).
- [124] D. Jandel, „Modular Learning im Onboarding: In 5 Schritten zum gelungenen Schulungskonzept – ein Praxisbeispiel“, *ICON Blog*, 10. Nov. 2020, 2020. [Online]. Verfügbar unter: <https://blog.iconnewmedia.de/modular-learning/>. Zugriff am: 24. Mai 2023.
- [125] G. Kamradt, *Exploring ChatGPT Function Calling*. [Online]. Verfügbar unter: https://github.com/gkamradt/langchain-tutorials/blob/main/data_generation/Exploring%20ChatGPT%20Function%20Calling.ipynb (Zugriff am: 22. Juni 2023).
- [126] GitHub, *GitHub - openai/openai-python: The OpenAI Python library provides convenient access to the OpenAI API from applications written in the Python language*. [Online]. Verfügbar unter: <https://github.com/openai/openai-python> (Zugriff am: 26. Juni 2023).

11.4 INHALTSVERZEICHNIS DES DATENTRÄGERS

	Dateiname	Dateiformat	Inhalt
1	Masterthesis		Digitale Version der Masterthesis
1.1	Masterthesis MaikeKaiser	.docx	
1.2	Masterthesis MaikeKaiser	.pdf	
2	Informationen zum Mermaid Editor	.txt	Link zur Dokumentation und zum browser-basierten Live-Editor
3	Kapitel vier Ist-Workflows		Flussdiagramme der Workflows vor Einführung von ChatGPT
3.1	Crossmedial		
3.1.1	Tägliche Nachrichten	.txt, .png	
3.1.2	Kurzüberblick	.txt, .png	
3.1.3	Breaking News	.txt, .png	
3.1.4	Informationsnachrichten	.txt, .png	
3.2	Fernsehen		
3.2.1	Tägliche Nachrichten	.txt, .png	
3.2.2	Kurzüberblick	.txt, .png	
3.2.3	Breaking News	.txt, .png	
3.2.4	Informationsnachrichten	.txt, .png	
3.3	Hörfunk		
3.3.1	Tägliche Nachrichten	.txt, .png	
3.3.2	Kurzüberblick	.txt, .png	
3.3.3	Breaking News	.txt, .png	
3.3.4	Beitrag zu täglicher Nachrichtensendung	.txt, .png	
3.4	Onlinejournalismus		
3.4.1	Tägliche Nachrichten	.txt, .png	
3.4.2	Informationsnachrichten (Wetter, Verkehr)	.txt, .png	
3.4.3	Breaking News	.txt, .png	
3.4.4	Informations- /Unterhaltungsbeitrag	.txt, .png	
3.5	Sonstige Workflows		
3.5.1	Ingest	.txt, .png	
3.5.2	Nutzerrückmeldung	.txt, .png	

4	Abschnitt 6.1 Erläuterungen zu den Projektmanagementtechniken	.pdf	Erläuterungen zu Scrum, Change Management und Erfolgsmetriken
5	Abschnitt 6.1 Ideen zum Product Backlog	.xlsm	Ideen für Arbeitspakete während der Einführung
6	Abschnitt 6.9 Beispielhaftes Inhaltsverzeichnis einer technischen Dokumentation	.pdf	Inhaltsverzeichnis anhand des Gesetzes über Künstliche Intelligenz
7	Abschnitt 7.3.4 GUI		Vorschläge für verschiedene GUIs als Zeichnungen
7.1	A Text	.vsdx	
7.2	B Archiv I	.vsdx	
7.3	G Archiv II	.vsdx	
7.4	H Audio	.vsdx	
7.5	I Grafik	.vsdx	
8	Abschnitt 12.2 Soll-Workflows		Flussdiagramme der Workflows nach der Einführung von ChatGPT
8.1	Crossmedial		
8.1.1	Tägliche Nachrichten	.txt, .png	
8.1.2	Kurzüberblick	.txt, .png	
8.1.3	Breaking News	.txt, .png	
8.1.4	Informationsnachrichten	.txt, .png	
8.2	Fernsehen		
8.2.1	Tägliche Nachrichten	.txt, .png	
8.2.2	Kurzüberblick	.txt, .png	
8.2.3	Breaking News	.txt, .png	
8.2.4	Informationsnachrichten	.txt, .png	
8.3	Hörfunk		
8.3.1	Tägliche Nachrichten	.txt, .png	
8.3.2	Kurzüberblick	.txt, .png	
8.3.3	Breaking News	.txt, .png	
8.3.4	Beitrag zu täglicher Nachrichtensendung	.txt, .png	
8.4	Onlinejournalismus		
8.4.1	Tägliche Nachrichten	.txt, .png	
8.4.2	Informationsnachrichten (Wetter, Verkehr)	.txt, .png	
8.4.3	Breaking News	.txt, .png	

8.4.4	Informations- /Unterhaltungsbeitrag	.txt, .png	
8.5	Sonstige Workflows		
8.5.1	Ingest	.txt, .png	
8.5.2	Nutzerrückmeldung	.txt, .png	
9	Abschnitt 12.3 Analyse der Prozesse mit KI-Unterstützung	.xslm	Tabellarische Auflistung der Prozesseigenschaften
10	Abschnitt 12.7 Informationsflüsse		Sequenzdiagramme zur den Informationsflüssen nach der Einführung von ChatGPT
10.1	Manuelle Anfrage	.txt, .png	
10.2	Automatische Anfrage	.txt, .png	
10.3	Ingest	.txt, .png	
10.4	Assetbearbeitung	.txt, .png	
10.5	Assetgeneration	.txt, .png	
10.6	Nutzerrückmeldung	.txt, .png	
10.7	Personalisierte Ausspielung	.txt, .png	
11	Abschnitt 12.8.2 Programmierbeispiele		Beispiele für Funktionen für verschiedene Use Cases
11.1	Function Calling Template	.py	
11.2	Ideengeber	.py	
11.3	Zusammenfassung	.py	
11.4	Stilistische Korrektur	.py	
11.5	Übersetzung	.py	
11.6	Textgeneration aus Prompt	.py	
11.7	Textgeneration aus Datenbank	.py	
11.8	Schlagwörter	.py	
11.9	Sentiment Analysis	.py	
11.10	Nutzeranfragen	.py	

12 ANHANG

ANHANG 1: ANFORDERUNGSMATRIX

A	Abschnitt	Art der Anforderung	Beschreibung	Gesetz	Definition of Compliance	Erfüllung durch
		Rechtlich-ethisch				
1	5.1	Freie Meinungsäußerung	Oberstes Gebot des Journalismus	Art. 5 Abs. 1 GG	Keine Meinungen ausblenden, keine Meinungen bevorzugen	6.5 Compliancekonzept
2	5.1	Jugendschutz gewährleisten		Jugendschutzgesetz	Keine Jugendschutzkritischen Inhalte an Jugendliche ausspielen, Alterskontrollen	8.3.5 Reglementierter Zugang zu Inhalten
3	5.1	Einrichtung von Kontroll- und Beschwerdestellen		Art. 27 Abs. 1 Gesetz über digitale Dienste Medienstaatsvertrag § 93	Erreichbare, effiziente Stellen, die Ergebnisse weiterleiten und in Aktionen umsetzen	8.3.4 Konzept der Information an Nutzer*innen
4	5.1	Funktionsweise von Empfehlungssystemen offenlegen		Art. 27 Abs. 1 Gesetz über digitale Dienste Medienstaatsvertrag § 93	Einfache Sprache, verständlich, zugänglich	6.9 Auditkonzept
5	5.1.7	Kennzeichnung von KI-generierten Beiträgen		Medienstaatsvertrag § 18	An Beiträge vorangestellt, klar, verständlich, sichtbar	7.2 Informationsflüsse
6	5.1.1	Personenbezogene Daten in den Trainingsdaten (außer Person des öffentlichen Interesses) nur nach Einwilligung verarbeiten	Personenbezogene Daten anonymisieren, sperren oder löschen	DSGVO Art. 6 Pressecodex Ziffer 4-3	Einwilligung einholen oder Daten löschen	6.5 Compliancekonzept

7	5.1.1	Workflows im Falle von Widerruf der Erlaubnis zur Datenverarbeitung		DSGVO Art. 7	Modularer Aufbau der Trainingsdaten, zusätzliche Information zu welcher Person welche Daten personenbezogen sind, Möglichkeit des Vergessens von Trainingsdaten?	6.5 Compliancekonzept
8	5.1.1	Workflows im Falle von Auskunft, ob die eigenen personenbezogenen Daten verarbeitet werden		DSGVO Art. 15	zusätzliche Information zu welcher Person welche Daten personenbezogen sind,	6.5 Compliancekonzept
9	5.1.1	Auskunft mit welchen Systemen die eigenen personenbezogenen Daten verwendet werden		DSGVO Art. 13	Einfache Sprache, verständlich, zugänglich	6.5 Compliancekonzept
10	5.1.2	Urheberrecht an Trainingsdaten		?	Urheberrecht an den Trainingsdaten liegt beim jeweiligen Urheber, KI-Hersteller und MU in Verlängerung dürfen diese nutzen	6.5 Compliancekonzept
11	5.1.2	Urheberrecht am Output des KI-Modells		?	Urheberrecht muss beim MU liegen, um die Beiträge veröffentlichen und gewerblich nutzen zu dürfen	-
12	5.1.3	Informationssicherheit	KI-Modell wird gegen Missbrauch und Angriffe von außen und innen gehärtet		Die Wahrscheinlichkeit für einen Missbrauch, Angriff oder Diebstahl ist möglichst gering, bei zumutbaren Maßnahmen	-
13	5.1.5	Unvoreingenommenheit gegenüber jeglicher Personengruppe		AGG § 1 Pressecodex Ziffer 12	Alle Personengruppen müssen ausreichend im Trainingsmaterial vertreten	6.5 Compliancekonzept

					sein, keiner Personengruppe dürfen übermäßig viele positive oder negative Eigenschaften zugeordnet werden, Korrektur von „alten“ Daten z.B. weibliche Sekretärinnen	
14	5.1.5	Trainingsdatensatz muss hohe Qualität haben	Damit richtige Ergebnisse und keine Voreingenommenheit		Verlangen von Qualitätsstandards für Trainingsdaten	6.5 Compliancekonzept
15	5.1.6	Technische Dokumentation muss vorgehalten sein	Technische Dokumentation zum KI-Modell und dessen Trainingsdaten	Art. 11, Art. 64 Gesetz über Künstliche Intelligenz	Verlangen einer Dokumentation und regelmäßige Aktualisierung	6.9 Auditkonzept
16	5.1.4	KI-System muss sich selbst erklären können		Art. 13 Gesetz über Künstliche Intelligenz	Einfache Sprache, verständlich, zugänglich, hochwertig	6.5 Compliancekonzept
		Funktional				
17	5.2.1	Hilfreiche Themenvorschläge liefern	KI schlägt aus Web, Sozialen Medien oder eigenen Daten Themen für neue Beiträge vor	-	Relevante, dankbare Themen für Beiträge	6.2.1 Erfüllung der funktionellen Anforderungen
18	5.2.1	Datensätze von der KI vorsortieren und zusammenfassen lassen → Hilfreiche, genaue Zusammenfassungen	Einkommendes Material in großen Mengen	-	Hilfreiche, genaue Zusammenfassungen, Wichtigste Aspekte aufgegriffen	6.2 Eignung von ChatGPT
19	5.2.1	Fakten werden korrekt überprüft, Deepfakes erkennen		-	Korrekte Fakten werden also solche erkannt, falsche Fakten werden ebenfalls als solche erkannt	-

20	5.2.1	Präzise Metadatenextraktion aus Essenzen	Extraktion von beschreibenden Metadaten bei neuen oder bestehenden Essenzen im Archiv	-	Aussagekräftige Metadaten, vorgegebene Anzahl/Tiefe der Informationen	6.2 Eignung von ChatGPT
21	5.2.1	Verortung aller Assets in einem mehrdimensionalen Vektorraum	Assets werden geclustert, um inhaltlich ähnliche Vorschläge für die Archivsuche oder die personalisierten Empfehlungen zu liefern	-	Präzise und korrekte Verortung, Nutzung sinnvoller Parameter	-
22	5.2.1	Speech-to-Text	Gesprochenes Interviewmaterial wird zu einem Text transkribiert	-	Präzise Transkription, Wörter richtig erkannt, kaum/keine Nacharbeit nötig	8.1 Weiteres Angebot von OpenAI
23	5.2.1	Übersetzungen	Text wird von einer Sprache in eine andere überführt	-	Genauere Übersetzung, trotzdem muss der Sprachfluss in der neuen Sprache gewährleistet sein	6.2 Eignung von ChatGPT
24	5.2.1	Sprachliche Korrekturen	Rechtschreib- und Grammatikkorrektur sowie Vorschläge zum Schreibstil		Korrekturen nach der neuen Rechtschreibung, Vorschläge für einen guten journalistischen Stil	6.2 Eignung von ChatGPT
25	5.2.1	Unterstützung bei Schnitt	Szenenerkennung, Setzen von In- und Outpoints	-	Korrekte Auswahl	-
26	5.2.1	Kuration von Videos, Bildern, Textpassagen...	Auswahl von relevanten Aussagen und Szenen	-	Korrekte Auswahl	-
27	5.2.2	Moderationstexte generieren	Generation von Text nach Vorgaben	-	Sprachlich richtige und für die Moderation angemessene Texte	6.2 Eignung von ChatGPT
28	5.2.2	Informationsnachrichten (Wetter, Verkehr, Börse) generieren	Generation des Nachrichtentextes nach Vorgaben	-	Sprachlich richtige und für die Nachrichten angemessene Texte	8.1 Weiteres Angebot von OpenAI

29	5.2.2	Text-to-Speech	Generation von Sprachaufnahmen auf Basis der Stimme eines Menschen aus einem Text	-	Realistisch klingende Aufnahme in Sendequalität und -format, Wörter sind korrekt ausgesprochen	-
30	5.2.3	Personalisierte Ausspielung	Individuelle Aggregation von Beiträgen für registrierte Nutzer*innen auf Basis ihrer in der Vergangenheit gewählten Beiträge	-	Anzeige von für die Nutzer*in interessanten Beiträgen, höhere Nutzungszeit	8.1 Weiteres Angebot von OpenAI
31	5.2.3	Chatbot zur Beantwortung von Nutzer*innenanfragen	Beantwortung einfache Anfragen von Nutzer*innen ohne menschliches Zutun via Chat	-	Richtige, höfliche Antworten, Weiterleitung von komplexeren Anfragen an Menschen	6.2 Eignung von ChatGPT
32	5.2.3	Sentiment Analysis zur Filterung von Hatespeech in Anfragen und Kommentaren	Erkennung von Hatespeech, Markierung solcher Aussagen	-	Korrekte Klassifizierung von Hatespeech	6.2 Eignung von ChatGPT
33	5.2.4	Schnittstelle zwischen KI und CMS, Einspeisung von Inhalten ohne menschliches Zutun	Informationen sowie Arbeitsanweisungen sollen direkt zwischen KI und CMS ausgetauscht werden, KI greift direkt auf den Speicher des Medienunternehmens zu	-	Arbeit der KI soll nicht durch menschliche Eingaben verzögert werden	7.3.1 Automation
34	5.2.4	Automatische Aufgaben an KI sowie manuelle Aufgaben von Menschen getriggert	Unterscheidung zwischen manuell ausgelösten (Rechtschreibkorrektur) und automatischen Aufgaben (Extraktion von Metadaten bei der Speicherung eines Essenz)	-	Arbeit mit der KI soll in hoher Geschwindigkeit ablaufen	7.3.1 Automation
35	5.2.4	KI darf nur Antworten im XML-Format ausgeben	Der Chatbot ChatGPT gibt beispielsweise Informationen	-	Die KI soll Informationen in maschinenlesbarem Format	7.3.2 Schnittstellen

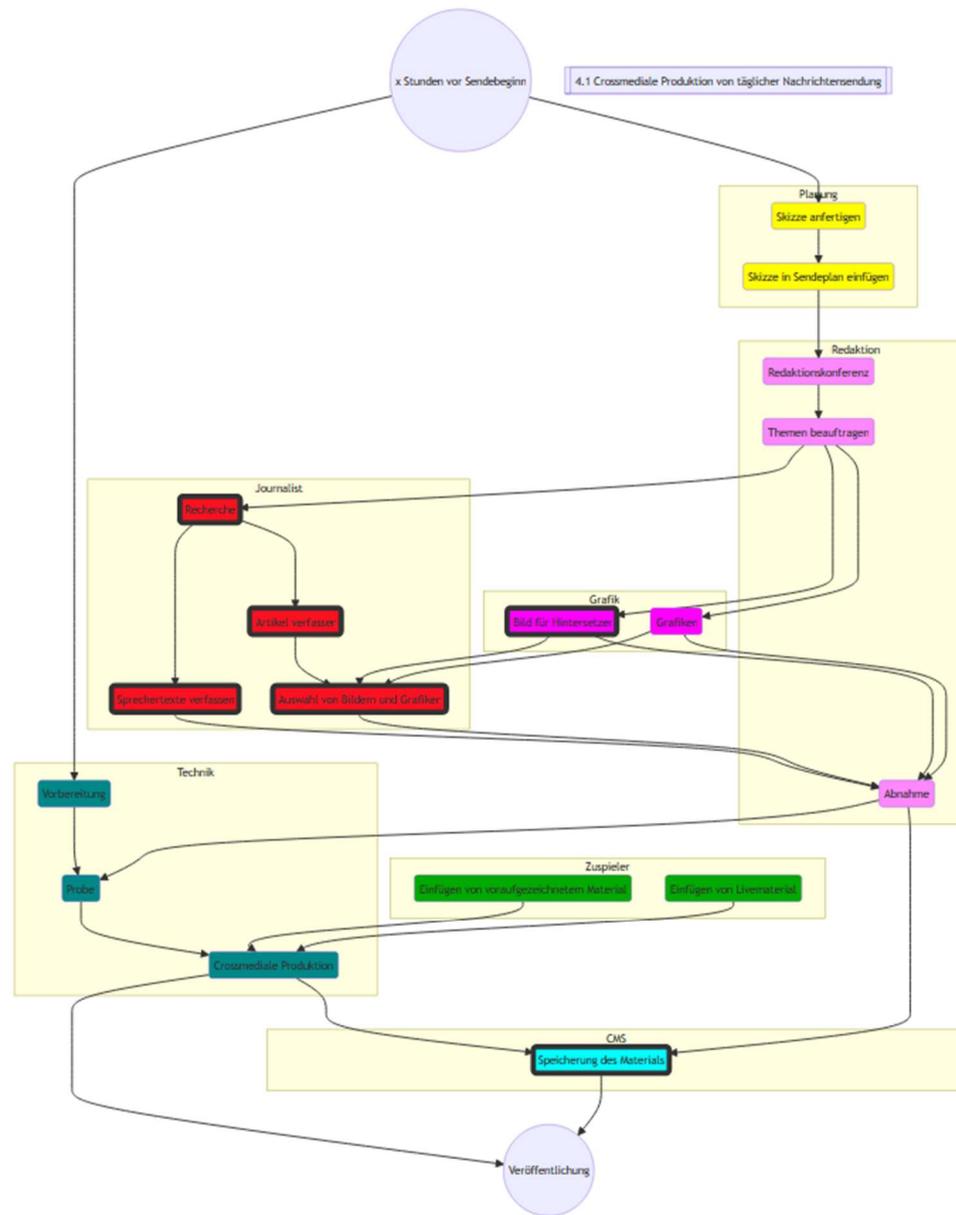
		(keine zusätzlichen Formulierungen)	in menschenlesbaren langen Texten aus		(XML) im hauseigenen Metadatenstandard ausgeben	
36	5.2.4	Bedienung von mehreren Profilen (Frontend, Admin, Plattform...)	Rollenbasierte Zugänge zum KI-Modell für erhöhte Informationssicherheit	-	Jede Person erhält nur so viele Zugriffsrechte, wie für ihre Arbeit unbedingt nötig ist	6.6 Zugangskonzept
37	5.2.4	Intuitive GUI	Für manuelle ausgelöste Aufgaben muss ein Frontend des KI-Modells existieren	-	Programmierunkundige Personen müssen ebenfalls mit dem KI-Modell arbeiten können	7.3.4 GUI
38	5.2.4	Schnelle Rechen- und Übertragungszeiten	Zu lange Rechen- und Übertragungszeiten verzögern den Arbeitsablauf und verringern die Akzeptanz für die Arbeit mit KI	-	Rechen- und Übertragungszeiten von Informationen und Assets vom CMS zur KI und zurück dürfen einen gewissen Rahmen nicht überschreiten	7.3.2 Schnittstellen
39	5.3	Vorhandene Strategie	Leitbild, an das alle folgenden Projekte, Tätigkeiten angelehnt werden	-	Nachvollziehbare, weitsichtige, machbare Strategie wird im Unternehmen kommuniziert	6.1 Ausrichtung des Projekts
40	5.3.1	Newsroom Prinzip bleibt erhalten	Der Newsroom ist zentraler Austauschort für Journalist*innen, Reporter*innen und Redaktion	-	KI wird als unterstützende Funktion den Mitarbeiter*innen zur Verfügung gestellt, der generelle Workflow wird nicht verändert	6.3 Soll-Workflows
41	5.3.1	Geringere Time to air/delivery	Die Arbeit der Journalist*innen wird effizienter genutzt. Es kann eine Sendung mit geringerem zeitlichen Vorlauf vorbereitet werden	-	Die Zeit vom Eintreffen einer neuen Meldung bis zur Veröffentlichung eines Beitrags wird kürzer	6.1 Ausrichtung des Projekts

42	5.3.2	Übernahme von repetitiven oder Nebentätigkeiten durch die KI	Eine Journalist*in ist ca. 40% ihrer Zeit mit Arbeiten außerhalb der Erstellung von Inhalten beschäftigt	-	Die Journalist*in kann sich in größerem Maße um kreative Tätigkeiten kümmern	6.1 Ausrichtung des Projekts
43	5.3.3	Das MU hat einen erweiterten Bedarf an Personal mit bestimmten Fähigkeiten	Es werden Berufsgruppen mit Hintergrund in Data Science benötigt	-	Mitarbeiter*innen werden geschult oder neu eingestellt	6.4 Rollenkonzept, 6.7 Schulungskonzept

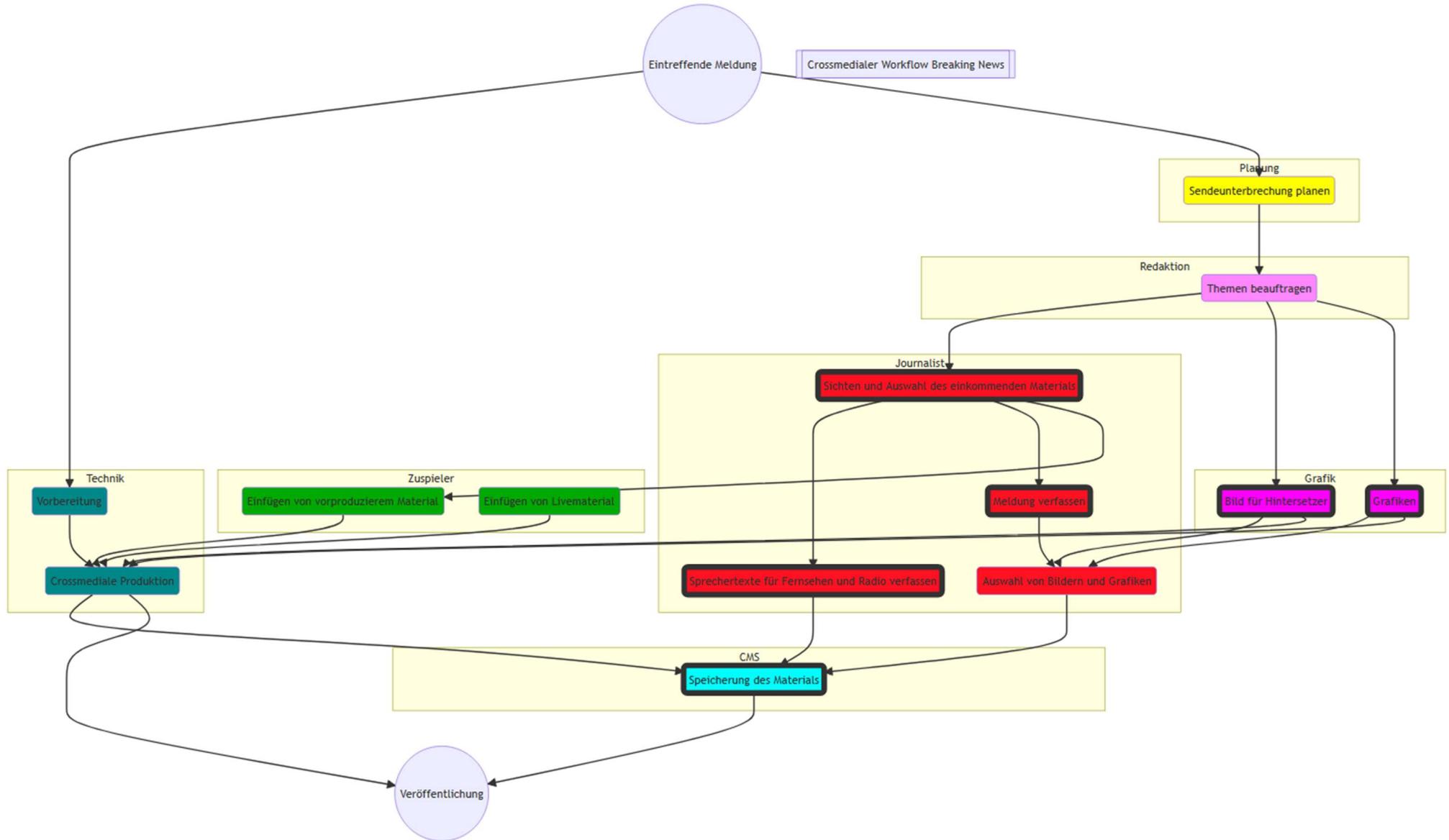
ANHANG 2: SOLL-WORKFLOWS

Workflow 1: Abschnitt 6.3 Crossmediale Produktion von täglichen Nachrichten mit KI-Unterstützung [Eigene Darstellung]	XXVII
Workflow 2: Abschnitt 6.3 Crossmediale Produktion im Breaking News-Fall mit KI-Unterstützung [Eigene Darstellung]	XXVIII
Workflow 3: Abschnitt 6.3 Crossmediale Produktion eines Kurzüberblicks bzw. Features mit KI-Unterstützung [Eigene Darstellung].....	XXIX
Workflow 4: Abschnitt 6.3 Crossmediale Produktion von Informationsnachrichten mit KI-Unterstützung [Eigene Darstellung].....	XXX
Workflow 5: Abschnitt 6.3 Ingest mit KI-Unterstützung [Eigene Darstellung].....	XXXI
Workflow 6: Abschnitt 6.3 Nutzerrückmeldung mit KI-Unterstützung [Eigene Darstellung].....	XXXII
Workflow 7: Abschnitt 8.2.4 Personalisierte Ausspielung mit KI-Unterstützung [Eigene Darstellung].....	XXXIII

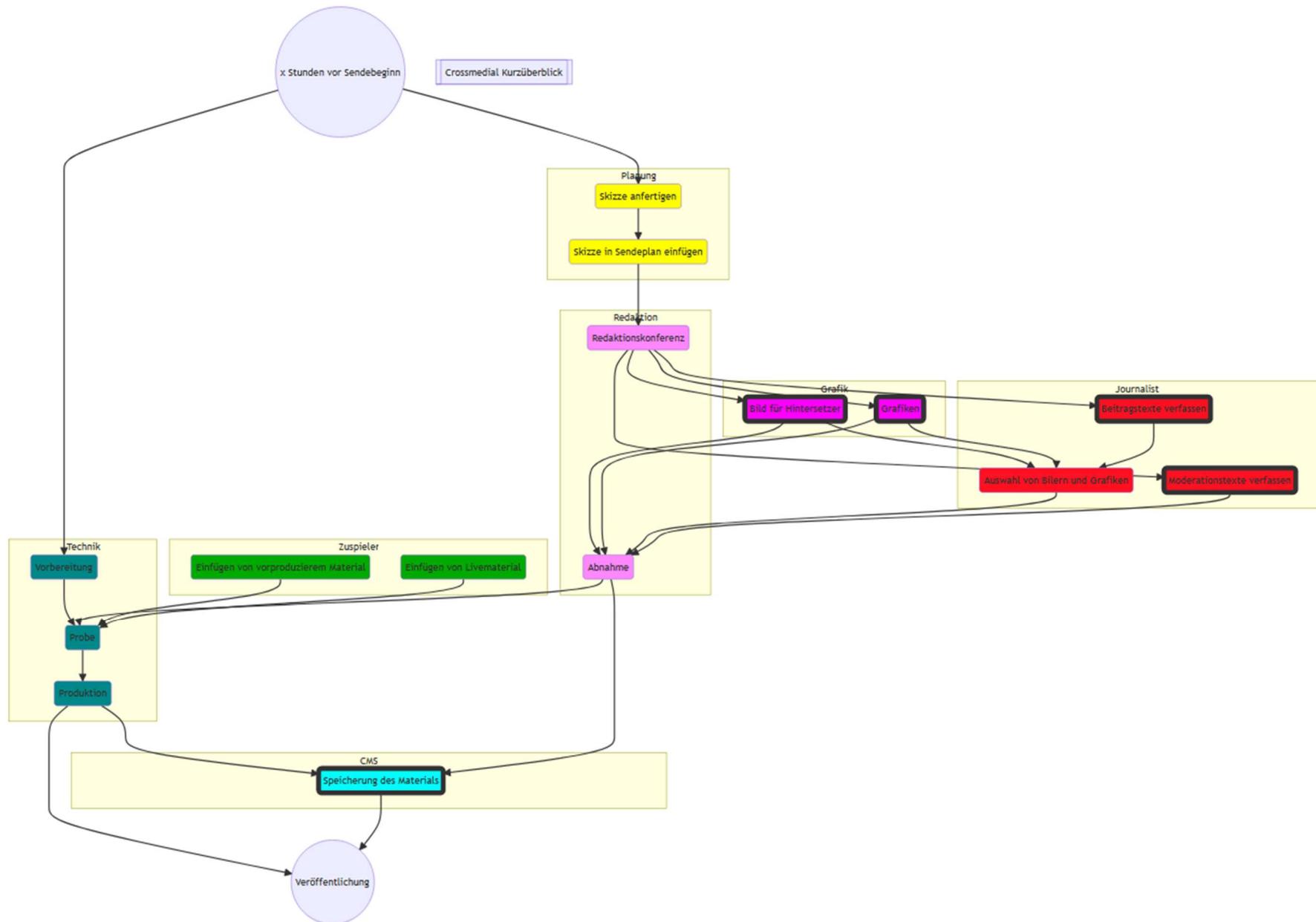
Alle Workflows inklusive jener aufgeschlüsselt nach Medientyp (Fernsehen, Hörfunk, Onlinejournalismus) befinden sich in digitaler Form auf dem Datenträger in Kapitel acht.



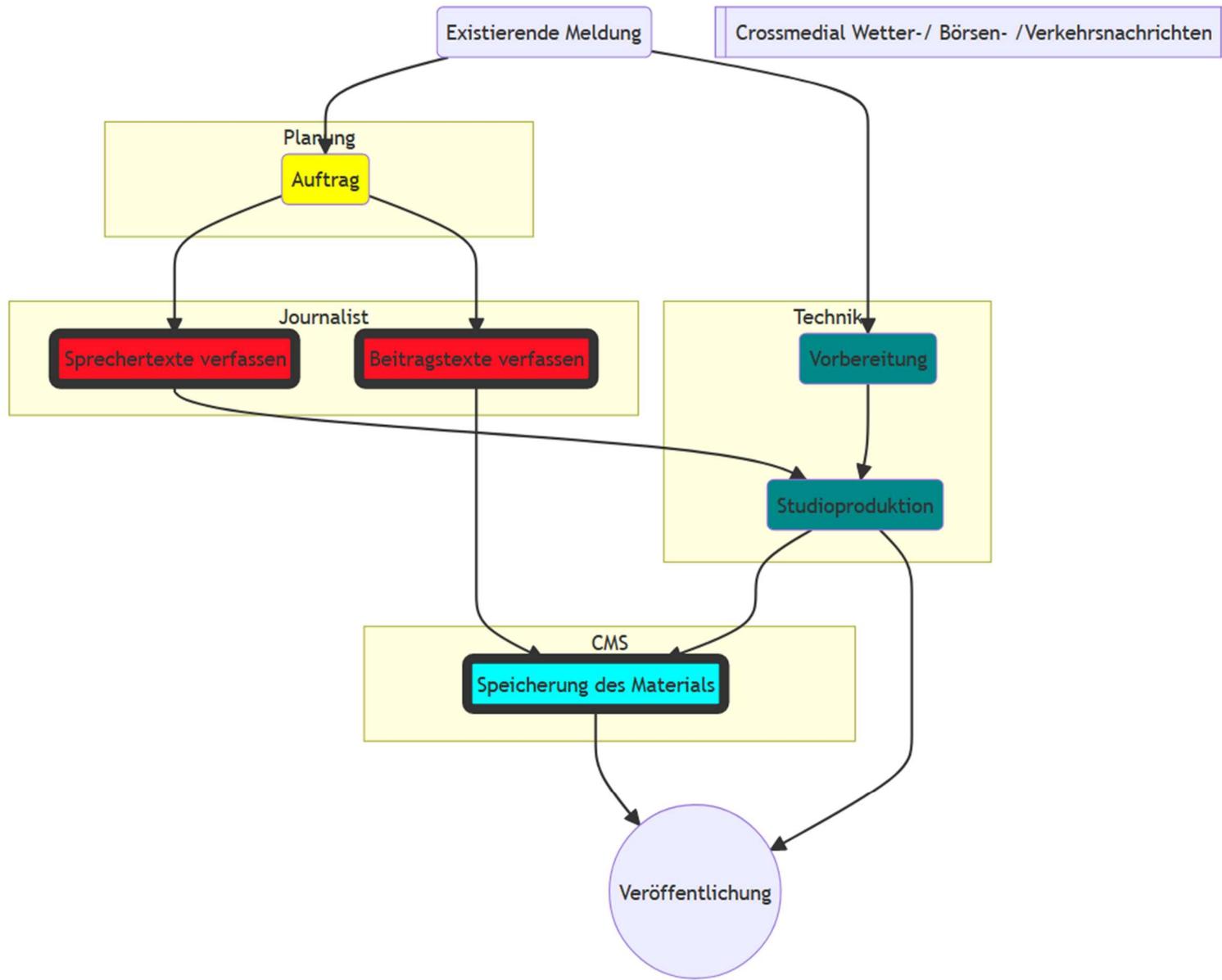
Workflow 1: Abschnitt 6.3 Crossmediale Produktion von täglichen Nachrichten mit KI-Unterstützung [Eigene Darstellung]



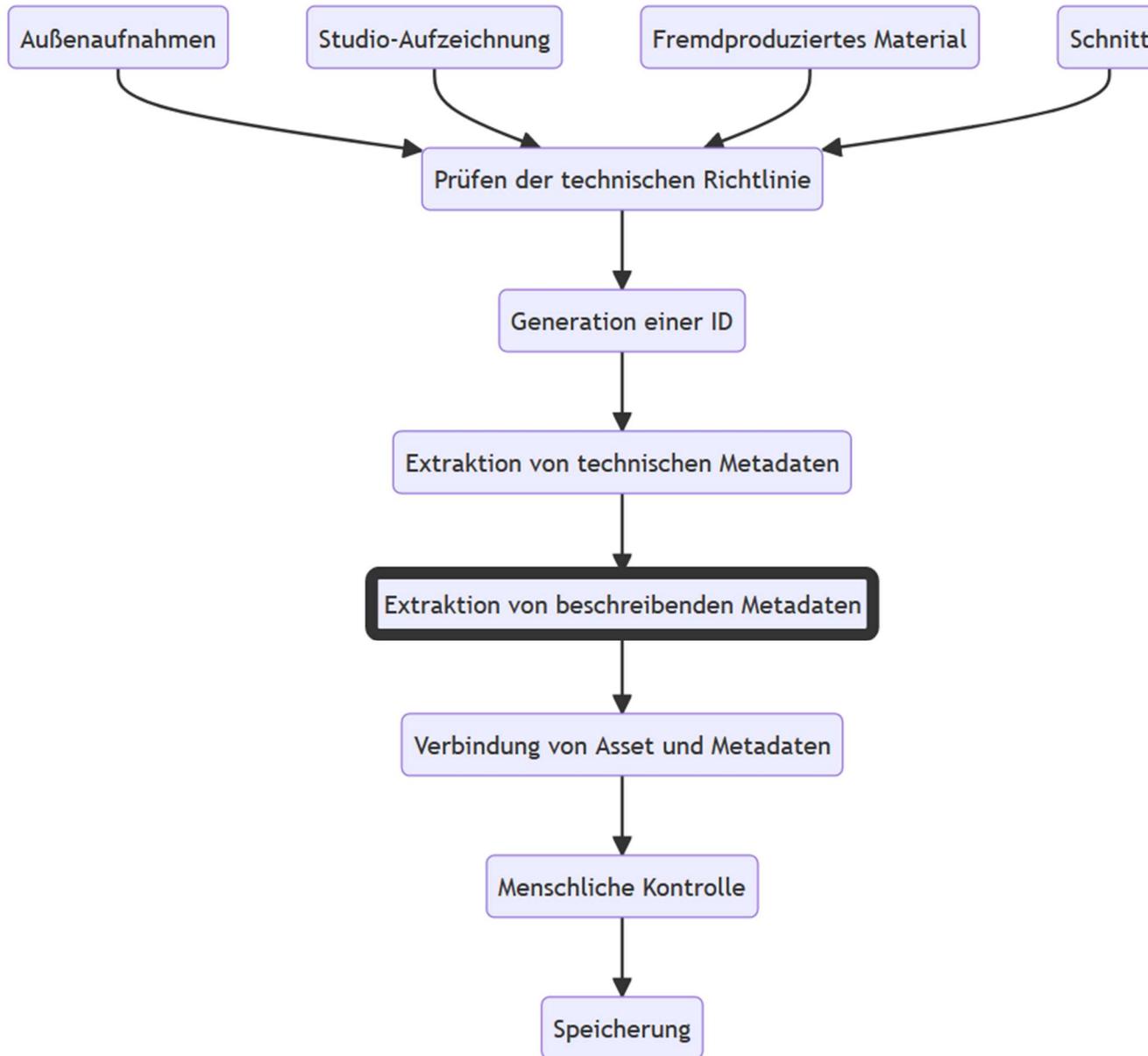
Workflow 2: Abschnitt 6.3 Crossmediale Produktion im Breaking News-Fall mit KI-Unterstützung [Eigene Darstellung]



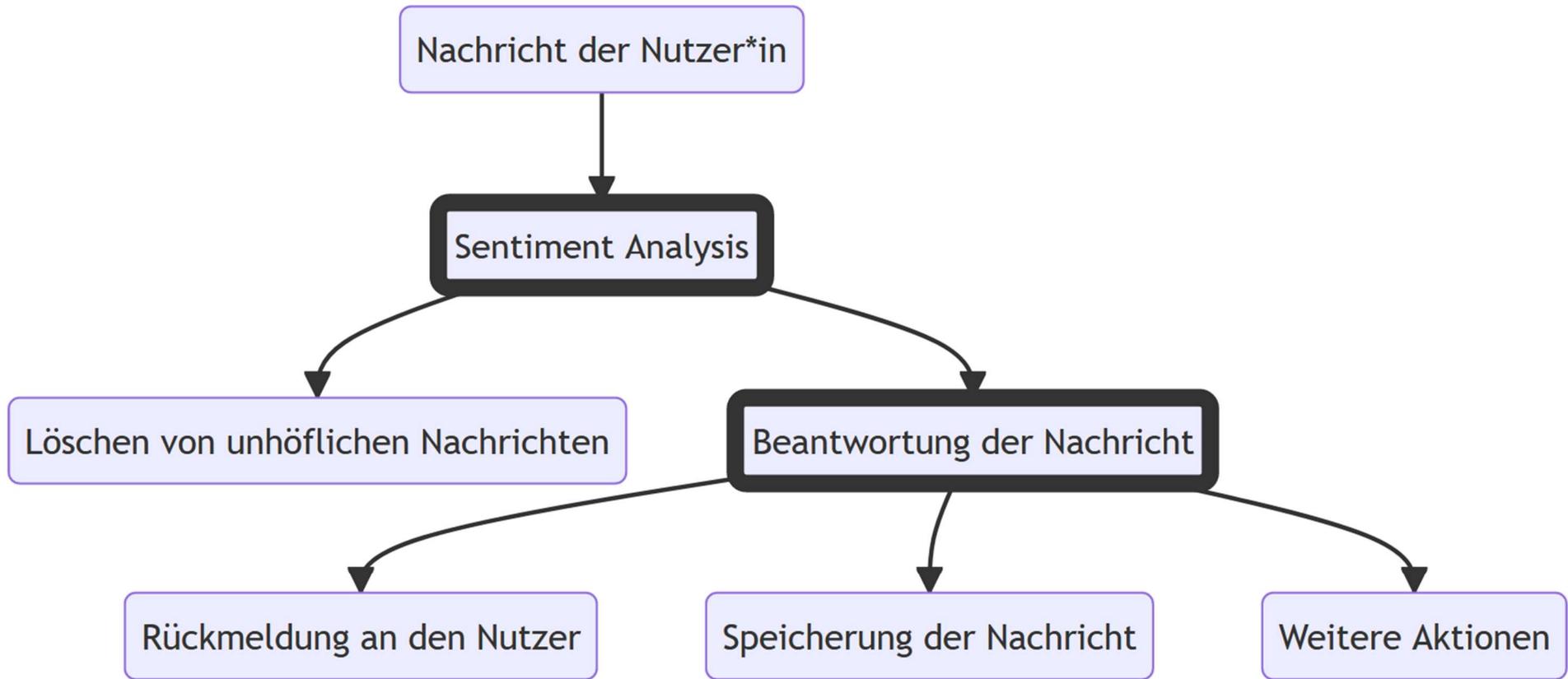
Workflow 3: Abschnitt 6.3 Crossmediale Produktion eines Kurzüberblicks bzw. Features mit KI-Unterstützung [Eigene Darstellung]



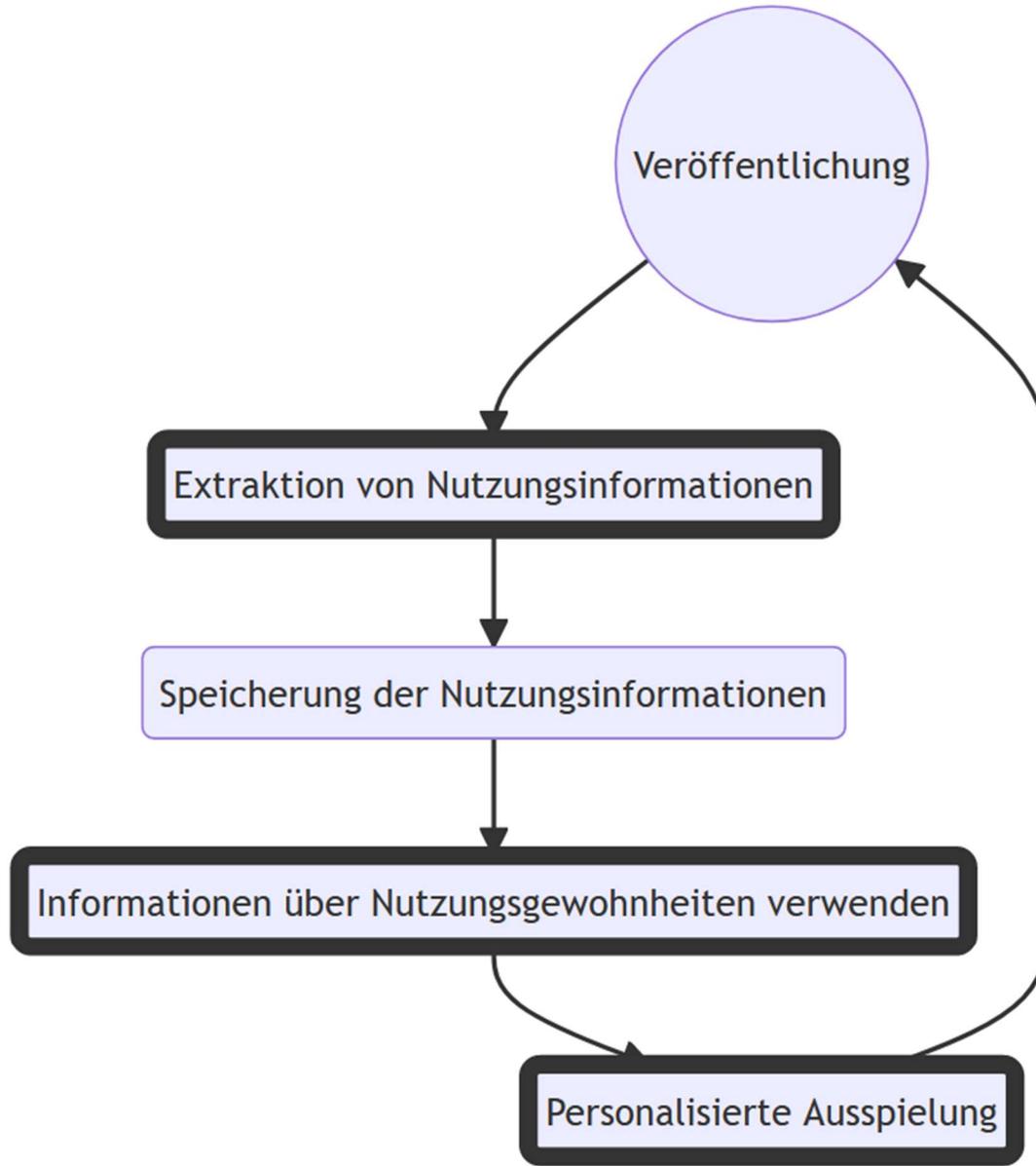
Workflow 4: Abschnitt 6.3 Crossmediale Produktion von Informationsnachrichten mit KI-Unterstützung [Eigene Darstellung]



Workflow 5: Abschnitt 6.3 Ingest mit KI-Unterstützung [Eigene Darstellung]



Workflow 6: Abschnitt 6.3 Nutzerrückmeldung mit KI-Unterstützung [Eigene Darstellung]



Workflow 7: Abschnitt 8.2.4 Personalisierte Auspielung mit KI-Unterstützung [Eigene Darstellung]

ANHANG 3: ANALYSE DER CROSSMEDIALEN WORKFLOWS MIT KI-UNTERSTÜTZUNG

Prozess	Start	Prozessbeteiligten	Ziel	Wie hoch ist die Bearbeitungszeit?	Wie oft wird der Prozess durchlaufen?	Planbarer Prozess?
Crossmedialer Workflow Tägliche Nachrichtensendung	x Stunden vor Sendebeginn	Journalist*in, Mediengestalter*in, Grafiker*in, Technik, Planung, CMS, KI, Redaktion, Sendeabwicklung	Ausspielung/Veröffentlichung	6-12 Stunden	Täglich	Ja
Crossmedialer Workflow Breaking News	Neue Meldung	Journalist*in, Grafiker*in, Technik, Planung, CMS, KI, Redaktion, Sendeabwicklung	Ausspielung/Veröffentlichung	1 Stunde	Bei Bedarf	Nein
Crossmedialer Workflow Kurzüberblick	x Stunden vor Sendebeginn	Journalist*in, Grafiker*in, Technik, Planung, CMS, KI, Redaktion, Sendeabwicklung	Ausspielung/Veröffentlichung	4 Stunden	Täglich	Ja
Crossmedialer Workflow Informationsnachrichten	Neue Meldung	Journalist*in, Grafiker*in, Technik, Planung, CMS, KI, Redaktion, Sendeabwicklung	Ausspielung/Veröffentlichung	0,5 - 1 Stunde	Halbstündlich	Ja
Ingest	Bei jeder Art einkommendem Material	Archivar*in, CMS, KI	Essenz ist mit Metadaten versehen und kann im Archiv wieder gefunden und verwendet werden	Wenige Minuten	Bei Bedarf	Nein
Nutzer*innenrückmeldung	Neue Nachricht	Journalist*in, CMS, KI, Nutzer*in	Rückmeldung an die Nutzer*in	Bis zu 6 Stunden	Bei Bedarf	Nein

Personalisierte Auspielung	Veröffentlichung eines journalistischen Erzeugnisses	KI, CMS	Empfehlung von zu den Interessen der Nutzer*innen passenden Inhalten	Wenige Minuten	Kontinuierlich	Ja
-------------------------------	---	---------	--	----------------	----------------	----

Die Prozessanalyse in digitaler Form befindet sich auf den Datenträger in Kapitel neun.

ANHANG 4: ROLLENÜBERSICHT

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
Producer*in	Zulieferungen organisieren, Leitungen bestellen, technische Infos im Rundown eintragen.	Producer*in	Gleich	Zulieferungen organisieren, Leitungen bestellen, technische Infos im Rundown eintragen.	-	-
Grafiker*in	Bereitet die Grafiken für die Hintersetzer vor und legt diese auf einem Austauschserver ab, Eventuelle Anpassungen sind nach Rücksprache mit dem Regisseur basierend auf den Kamerapositionen notwendig, berät über die Umsetzbarkeit und macht die Bildrecherche.	Grafiker*in	Weniger	Bereitet die Grafiken für die Hintersetzer vor und legt diese auf einem Austauschserver ab, Eventuelle Anpassungen sind nach Rücksprache mit dem Regisseur basierend auf den Kamerapositionen notwendig, berät über die Umsetzbarkeit und macht die Bildrecherche.	-	-
Erklärgrafiker*in	Entwickelt und realisiert spezielle Erklärgrafiken.	Erklärgrafiker*in	Gleich	Entwickelt und realisiert spezielle Erklärgrafiken.	-	-
Planer*in	Konzipieren der Sendung, Erstellen des Sendeplans und Bestellen der Beiträge.	Planer*in	Gleich	Konzipieren der Sendung, Erstellen des Sendeplans und Bestellen der Beiträge.	Ideenquelle für Beiträge und thematische Ansatzpunkte	A: Text
Co-Planer*in	Entwirft gemeinsam mit dem Grafiker Erklärgrafiken und entwirft die	Co-Planer*in	Gleich	Entwirft gemeinsam mit dem Grafiker Erklärgrafiken und lädt Studiogäste ein und	Zusammenfassung für Kurzüberblicke, Ideenquelle	A: Text

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
	Kurzüberblicke, weiteren Beiträge und Features.			entwirft die Kurzüberblicke, weiteren Beiträge und Features.		
Journalist*in	Ist für die inhaltliche Produktion der Beiträge verantwortlich.	Journalist*in mit Schulung in Grundlagen der KI	Mehr	Ist für die inhaltliche Produktion der Beiträge und Moderationstexte verantwortlich, bringt den Kolleg*innen sowie dem Publikum den Einsatz mit KI näher, gestaltet in der Einführungsphase gemeinsam mit den Data-Scientist*innen Workshops.	Ideenquelle, Korrekturvorschläge, Stilvorschläge	A:Text
Außenreporter*in	Hat die inhaltliche Verantwortung über Außenreportagen, spricht Aufseher, führt Interviews, sorgt für die Einhaltung des Drehplans und den pünktlichen Ingest des produzierten Materials, arbeitet bereits bei der Aufnahme für mehrere Medienformen.	Außenreporter*in	Gleich	Hat die inhaltliche Verantwortung über Außenreportagen, sorgt für die Einhaltung des Drehplans und den pünktlichen Ingest des produzierten Materials, arbeitet bereits bei der Aufnahme für mehrere Medienformen.	-	-

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
Bildingenieur*in – Außen	Führt bei der Außenreportage die Kameraarbeit nach Vorgabe der Außenreporter*in aus.	Bildingenieur*in – Außen	Gleich	Führt bei der Außenreportage die Kameraarbeit nach Vorgabe der Außenreporter*in aus.	-	-
Toningenieur*in – Außen	Führt bei der Außenreportage die Tonarbeit nach Vorgabe der Außenreporter*in aus.	Toningenieur*in – Außen	Gleich	Führt bei der Außenreportage die Tonarbeit nach Vorgabe der Außenreporter*in aus.	-	-
Mediengestalter*in	Schneidet das einkommende Material zu einem fertigen Beitrag zusammen, achtet auf die Einhaltung der Länge, bearbeitet Audio- und Videomaterial nach.	Mediengestalter*in	Weniger	Schneidet das einkommende Material zu einem fertigen Beitrag zusammen, achtet auf die Einhaltung der Länge, bearbeitet Audio- und Videomaterial nach.	-	-
Nachrichtenredakteur*in	Befüllt in Absprache mit dem CvD den Nachrichtenblock der Sendung, schreibt die Nachrichtenmeldungen und Anmoderationen zu den Nachrichtenfilmen im Nachrichtenblock.	Nachrichtenredakteur*in	Gleich	Befüllt in Absprache mit dem CvD den Nachrichtenblock der Sendung, Kontrolle von generierten Nachrichtentexten	Generation von Moderationstexten	A: Text
CvD (*)	Übernimmt die Sendeplanung vom Planer und ist für die Weiterentwicklung des	CvD (*)	Gleich	Übernimmt die Sendeplanung vom Planer und ist für die Weiterentwicklung des	-	-

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
	Sendeplans verantwortlich, Abnahme der Beiträge (für den aktuellen und eventuell auch Folgetage), Grafik- und Regiebesprechungen durchführen, Änderung in Open Media während der Sendung pflegen, Absprache mit Studiogästen und Reporter*innen, ist verantwortlich für den Inhalt der Sendung (Presserechtlich).			Sendeplans verantwortlich, Abnahme der Beiträge (für den aktuellen und eventuell auch Folgetage), Grafik- und Regiebesprechungen durchführen, Änderung in Open Media während der Sendung pflegen, Absprache mit Studiogästen und Reporter*innen, ist verantwortlich für den Inhalt der Sendung (Presserechtlich) gemeinsam mit der Compliance Manager*in.		
Regisseur*in	Verantwortlich für die gestalterische Umsetzung der Sendung, ist bei allen redaktionellen Sitzungen dabei.	Regisseur*in	Gleich	Verantwortlich für die gestalterische Umsetzung der Sendung, ist bei allen redaktionellen Sitzungen dabei.	-	-
Studio Operator*in	Sendung technisch vorbereiten (Kameraeinstellung, Licht). Eventuell Nachregelung bei Licht und Kamera während der	Studio Operator*in	Gleich	Sendung technisch vorbereiten (Kameraeinstellung, Licht). Eventuell Nachregelung bei Licht und Kamera während der	-	-

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
	Sendung, ruft bei nicht-automatisierten Sendungen die Kamerafahrten ab.			Sendung, ruft bei nicht-automatisierten Sendungen die Kamerafahrten ab.		
Toningenieur*in	Sprachmikrofonsignale aufbereiten (Pegeln, EQ, Komprimieren), Töne mischen.	Toningenieur*in	Weniger	Sprachmikrofonsignale aufbereiten (Pegeln, EQ, Komprimieren), Töne mischen.	-	-
Archivar*in	Verwaltung von sämtlichem Material im MU, manuelle Annotation von Essenzen, Unterstützung bei der Recherche.	Archivmitarbeiter*in	Weniger	Verwaltung von sämtlichem Material im MU, steht im Austausch mit der Data-Scientist*in, um Systematisierung des Archivs so gewinnbringend wie möglich zu gestalten, Unterstützung bei der Recherche, Überwachung der automatischen Annotation.	KI übernimmt die Annotation der Essenzen	B: Archiv
Rollen für den Einführungsprozess (und darüber hinaus)						
-	-	Data-Scientist*innen	Mehr	Ist verantwortlich für die Einhaltung der Qualitäts- und Aktualitätsstandards der KI, führt regelmäßige Tests durch, steht im	-	D: Test, F: Training

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
				Austausch mit den Journalist*innen, geht auf deren Anregungen ein.		
-	-	KI-Assistenz	Mehr	Unterstützt die Data-Scientist*in, ist für die Aufbereitung von Daten zuständig.	-	D: Test, F: Training
-	-	Compliance Manager*in	Mehr	Ist verantwortlich für die Einhaltung der ethischen und rechtlichen Anforderungen der KI, führt regelmäßige Tests durch, Bericht an den CISO.	-	C: Test
-	-	IT-Administrator*in – KI / Systemintegrator*in	Mehr	Ist für die Verbindung von KI, CMS und MAM verantwortlich, schafft Schnittstellen zwischen den Systemen, gleicht die Metadaten aneinander an, sorgt für die Automation von Workflows.	-	D: Admin
-	-	UI Designer*in	Mehr	Entwirft nach der Vorgabe der Anwender*innen die Benutzungsoberfläche	-	E: GUI

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
				für das KI-Modell, Einbindung ins Redaktionssystem und CMS, führt zu Beginn ein Requirements Engineering durch.		
-	-	Change Manager*in	Mehr	Projektleiter während der Einführung von ChatGPT, hat den Überblick und beauftragt Arbeitspakete, berichtet an die Leitung des MU.	-	-
-	-	Projektteam	Mehr	Besteht aus Data-Scientist*in, KI-Assistenz, Compliance Manager*in, IT-Administrator*in und UI Designer*in, arbeitet in zeitlich begrenzten Sprints, die es selbst aus dem Product Backlog zusammenstellt.	-	-
-	-	Scrum Master	Mehr	Coach des interdisziplinären Teams, effektive Nutzung von Scrum, moderiert die Teambesprechungen und sorgt für die Einhaltung des Zeitplans, stellt	-	-

Rolle alt	Beschreibung	Rolle neu	Veränderte Anzahl	Beschreibung	KI Unterstützung in der neuen Rolle	Nutzungsprofil für Authentifizierung
				sicher, dass das Team gewinnbringend arbeiten kann.		

Die orange Einfärbung zeigt den Einsatz von KI in dieser Rolle an.

ANHANG 5: SCHULUNGSMATRIX

	Nicht-technisch: Journalist*innen, Redaktion, Außenreporter*innen, CvD (*)	Produktiv: Producer*innen, Planer*innen, Co- Planer*innen	Gestalterisch: Grafiker*innen, Mediengestalter*innen, Regisseur*innen	Archiv: Archivar*innen
Generelle Einführung in KI				
Funktionsweise der Algorithmen				
Chancen und Limitierungen von KI				
Einführung in textuelle Use Cases				
Einführung in Archivworkflow				
Einführung in auditive Use Cases				

Datenqualität				
Rechtliche und ethische Informationen (Datenschutz, Fairness, Transparenz...)				
Verhalten bei Überlastung oder Ausfall				
Sensibilisierung für Informationssicherheit				
Verhalten im Sicherheitsnotfall				

Eigene Darstellung nach [124]

	Tiefes Verständnis nötig
	Nur oberflächliche Kenntnisse nötig
	Keine Kenntnisse zwingend nötig

ANHANG 6: RISIKOMATRIX

	Beschreibung	Gründe	Konsequenzen	Bewertung	Handlungsoptionen
Strategische Risiken	Gesetze oder Pressecodex werden nicht eingehalten	Fehlerhafte Programmierung, Mindere Trainingsdatenqualität	Rechtliche Probleme, Reputationsverlust		Kontakt zu OpenAI, Überprüfen des eigenen Materials, Nachtraining
	Veröffentlichung von falschen Fakten	Fehlende menschliche Kontrolle, Fehlerhafte Programmierung, Mindere Trainingsdatenqualität	Reputationsverlust, Verlust von Nutzer*innen		Berichtigung veröffentlichen
	Veröffentlichung von Bias und Diskriminierung	Fehlende menschliche Kontrolle, Fehlerhafte Programmierung, Mindere Trainingsdatenqualität	Reputationsverlust, Verlust von Nutzer*innen		Berichtigung veröffentlichen
	Fehlende Kennzeichnung von KI-generierten Inhalten	Fehlende menschliche Kontrolle, Fehlerhafte Programmierung	Reputationsverlust, Verlust von Nutzer*innen		Berichtigung veröffentlichen
	Nutzer*innenanfrage wird falsch beantwortet	Fehlende menschliche Kontrolle, Fehlerhafte Programmierung, Mindere Trainingsdatenqualität	Reputationsverlust, Verlust von Nutzer*innen		Berichtigung nachsenden
	Sicherheitszwischenfall	Fehlende Sicherheitsmaßnahmen, geringe Sensibilisierung der Mitarbeiter*innen	Sendeausfall, Preisgabe von Daten, Nicht-Nutzbarkeit von technischen Systemen		Jeweiliges Missbrauchsszenario erkennen und Gegenmaßnahmen einleiten
Finanzielle Risiken	Zu hohe Kosten	Viele Tests nötig, zusätzliches Personal, Hardware...	Projekt dauert länger, Projektergebnis wird qualitativ schlechter		Weitere finanzielle Mittel allokalieren, weitere Aktionen effizienter gestalten
	OpenAI erhöht die Preise	-	Projekt und operativer Betrieb wird teurer		Weitere finanzielle Mittel allokalieren, Prozesse in

					Zusammenarbeit mit den Mitarbeiter*innen verschlanken
	Mehr Aufgaben an ChatGPT abgegeben als geplant	Fehlplanung, ChatGPT findet hohen Anklang und wird von vielen Nutzer*innen verwendet	Hohe Auslastung von ChatGPT, hoher Rechenaufwand, hohe Kosten		Weitere finanzielle Mittel allokalieren, System hochskalieren
Operationelle Risiken	Einführungsprozess dauert zu lange	Viele Tests und Nachbesserungen	Hohe Kosten		Weitere finanzielle Mittel allokalieren, Prozesse verschlanken, Akzeptanz der Mitarbeiter*innen hoch halten
	Qualität des Outputs zu gering, Output ist für die Use Cases nicht nutzbar	Prompt Engineering nicht effektiv, Datenqualität ist zu niedrig	ChatGPT nicht vollständig nutzbar, Aufgaben müssen von Menschen übernommen werden		Weitere finanzielle Mittel allokalieren, Kontakt zu OpenAI, weitere Expert*innen heranholen, Datensäuberung und Nachtraining
	Übermittlungs- oder Rechengeschwindigkeit zu niedrig	Schnittstelle ist nicht für Datenraten ausgelegt, Ratenlimit von OpenAI, hohe Auslastung	ChatGPT nicht vollständig nutzbar, Aufgaben müssen von Menschen übernommen werden		Kontakt zu OpenAI, Übertragungswege verstärken
	Zu geringe Verfügbarkeit	Unzuverlässige Schnittstelle, unzuverlässige Hardware, zu hohe Auslastung	ChatGPT nicht vollständig nutzbar, Aufgaben müssen von Menschen übernommen werden		Kontakt zu OpenAI, Übertragungswege verstärken
	Neue Prozesse sind zu umständlich	Prozesse haben zu viele Akteure und Elemente	ChatGPT nicht vollständig nutzbar		Prozesse in Zusammenarbeit mit den Mitarbeiter*innen verschlanken
	Neue Prozess werden als zu umständlich empfunden	Nicht ausreichende Schulung, niedrige Akzeptanz der Mitarbeiter*innen	ChatGPT nicht vollständig nutzbar, Mitarbeiter*innen bearbeiten die Aufgaben bevorzugt manuell		Weitere Schulungsmodule entwerfen und veröffentlichen, Wandel in der Leitungsebene

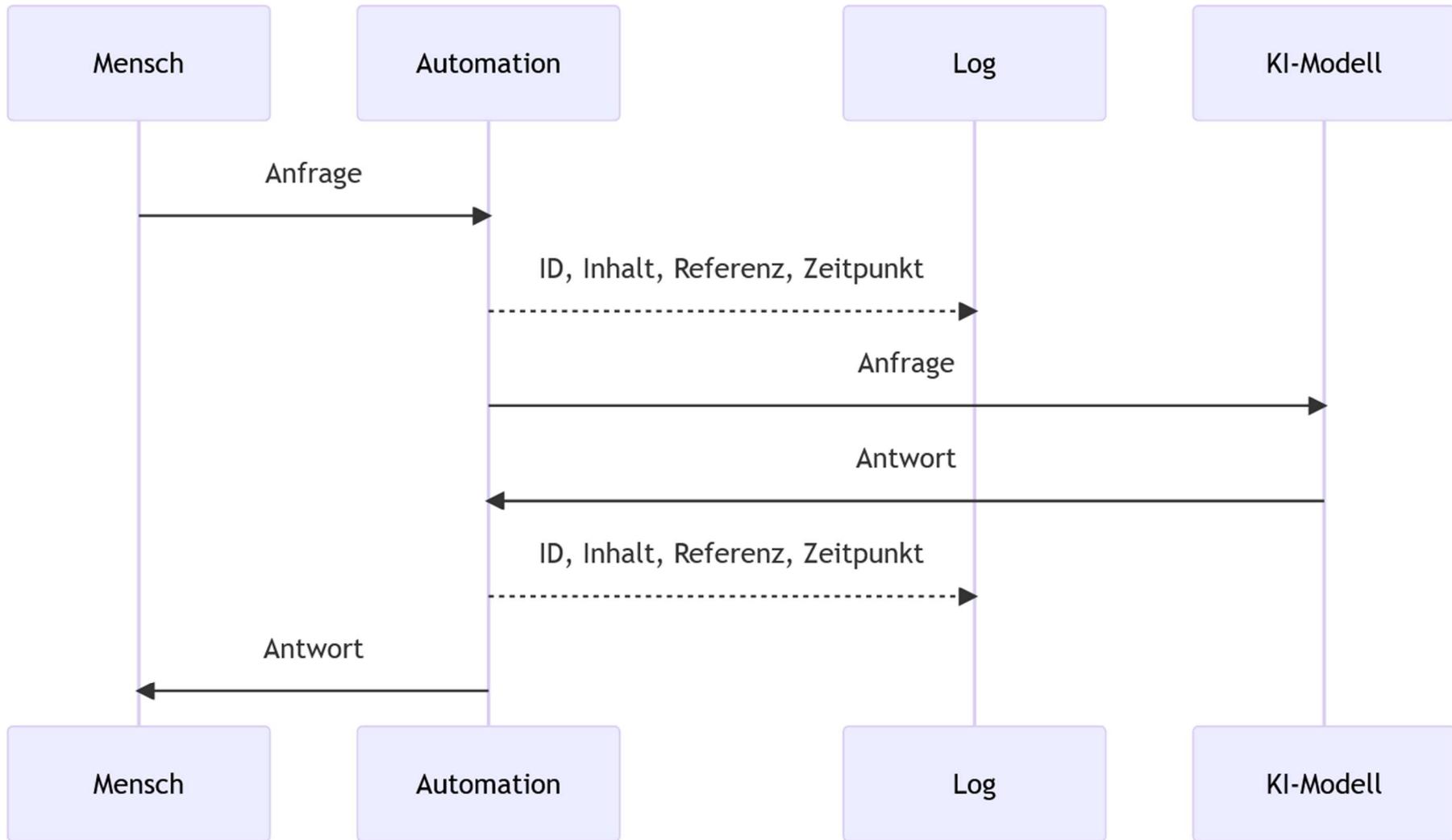
					vorleben, Peer-to-Peer-Schulungen
	Niedrige Akzeptanz der Mitarbeiter*innen	Unzureichende Informationskampagne und Schulung	ChatGPT nicht vollständig nutzbar		Weitere Schulungsmodule entwerfen und veröffentlichen, Wandel in der Leitungsebene vorleben, Peer-to-Peer-Schulungen

	Risiko birgt einen hohen Schaden, sofort handeln!
	Risiko birgt einen mittleren Schaden
	Risiko birgt einen niedrigen Schaden

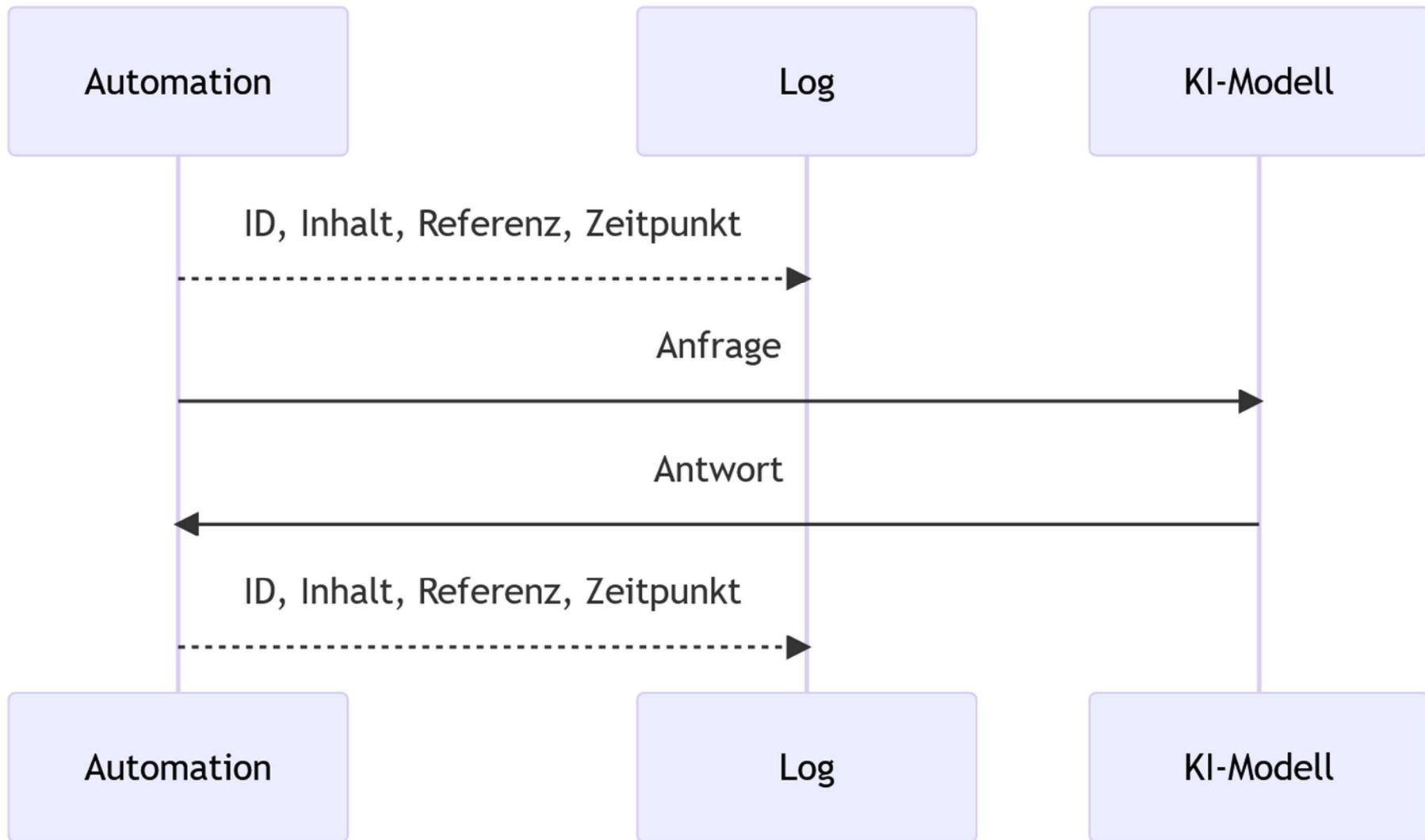
ANHANG 7: INFORMATIONSFÜSSE

Informationsfluss 1: Abschnitt 7.2 Manuelle Anfrage [Eigene Darstellung].....	L
Informationsfluss 2: Abschnitt 7.2 Automatisierte Anfrage [Eigene Darstellung]	LI
Informationsfluss 3: Abschnitt 7.2 Ingest [Eigene Darstellung]	LII
Informationsfluss 4: Abschnitt 7.2 Assetbearbeitung [Eigene Darstellung].....	LIII
Informationsfluss 5: Abschnitt 7.2 Assetgeneration [Eigene Darstellung]	LIV
Informationsfluss 6: Abschnitt 7.2 Personalisierte Ausspielung [Eigene Darstellung]	LV
Informationsfluss 7: Abschnitt 8.5.1 Nutzer*innenanfrage [Eigene Darstellung]	LVI

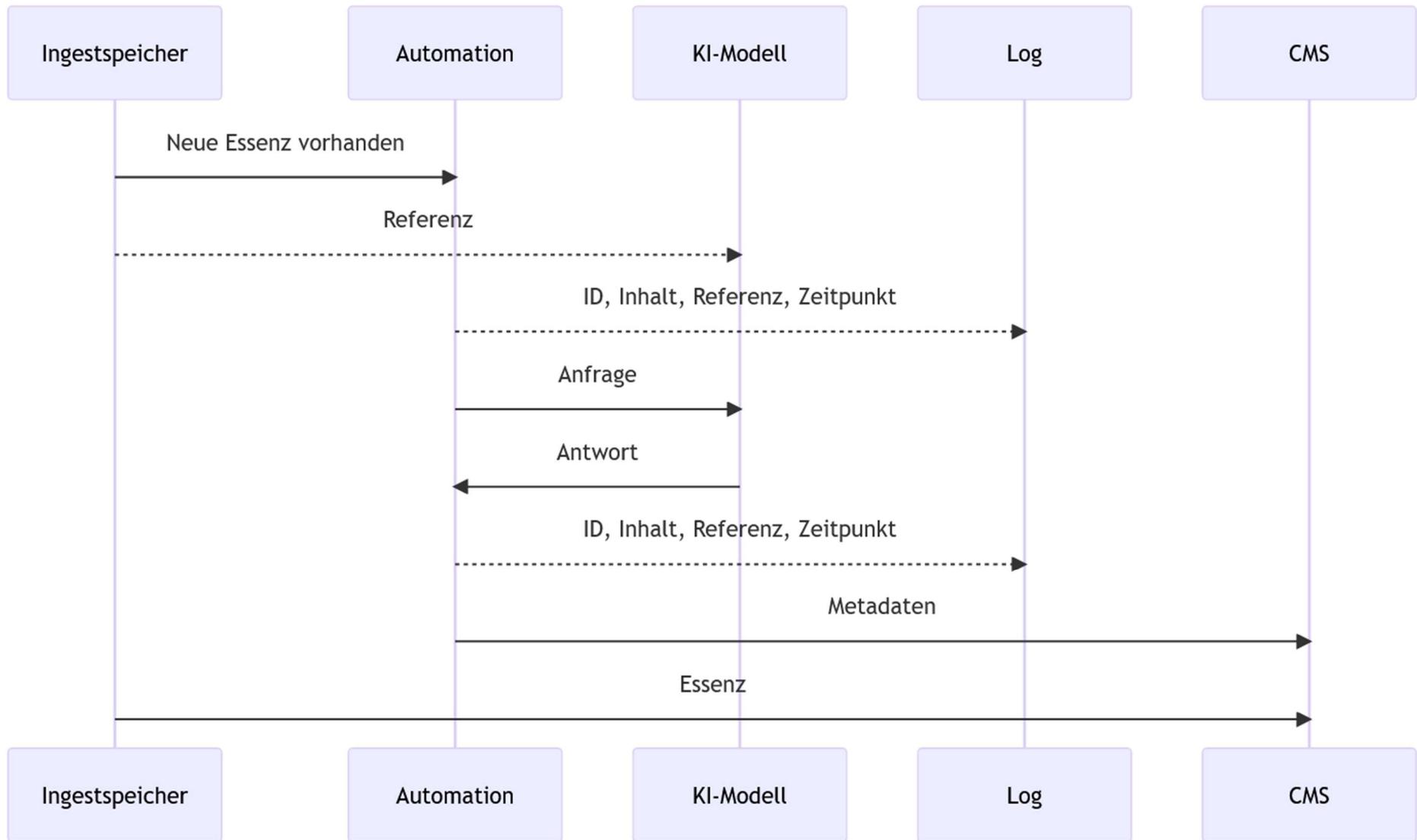
Die Zeichnungen befinden sich in digitaler Form auf dem Datenträger in Kapitel zehn.



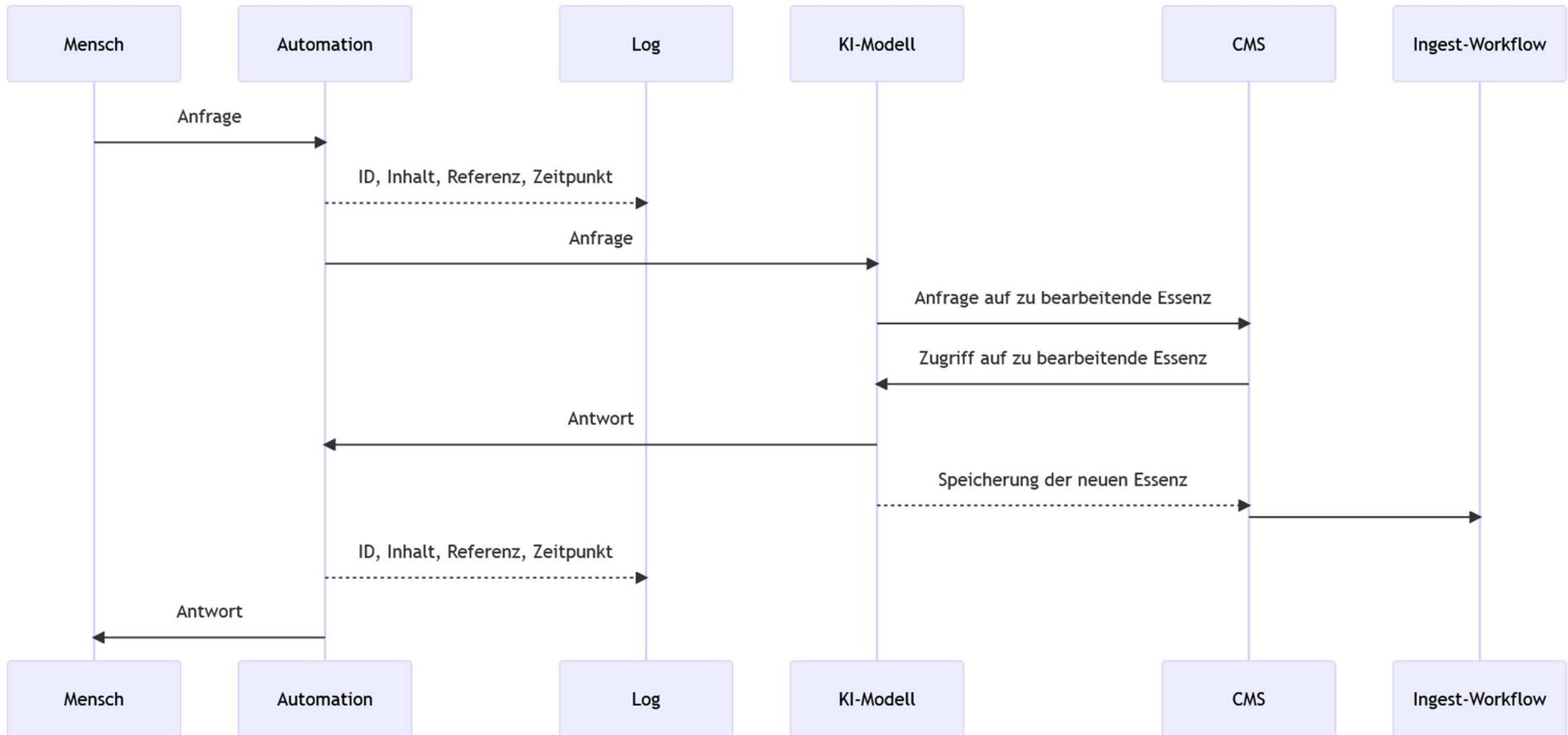
Informationsfluss 1: Abschnitt 7.2 Manuelle Anfrage [Eigene Darstellung]



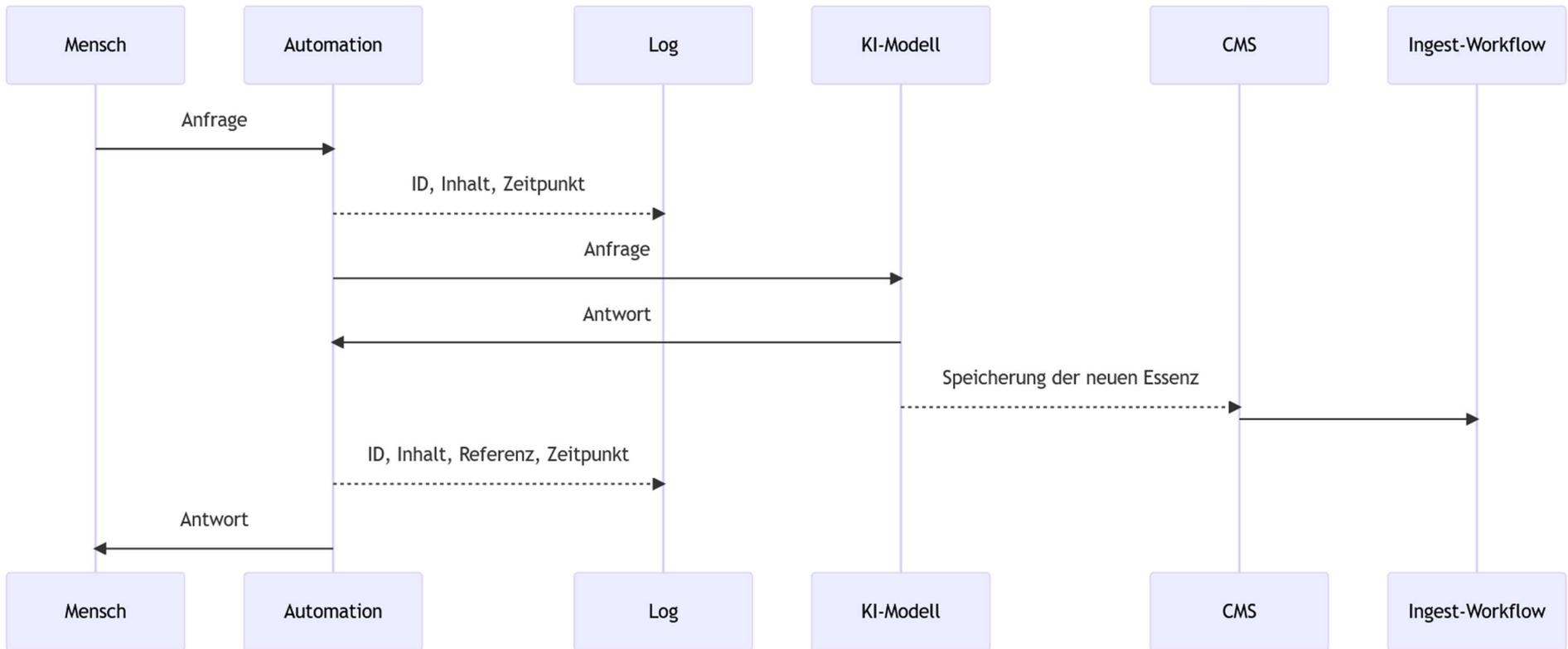
Informationsfluss 2: Abschnitt 7.2 Automatisierte Anfrage [Eigene Darstellung]



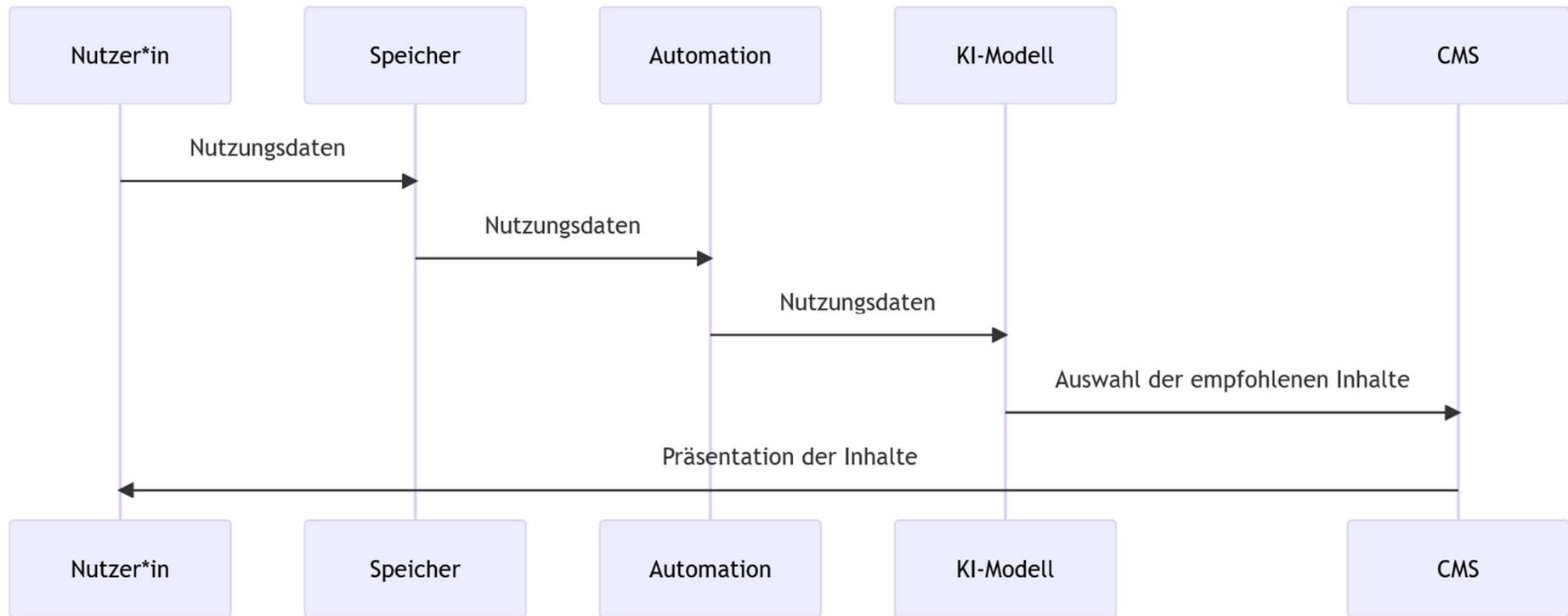
Informationsfluss 3: Abschnitt 7.2 Ingest [Eigene Darstellung]



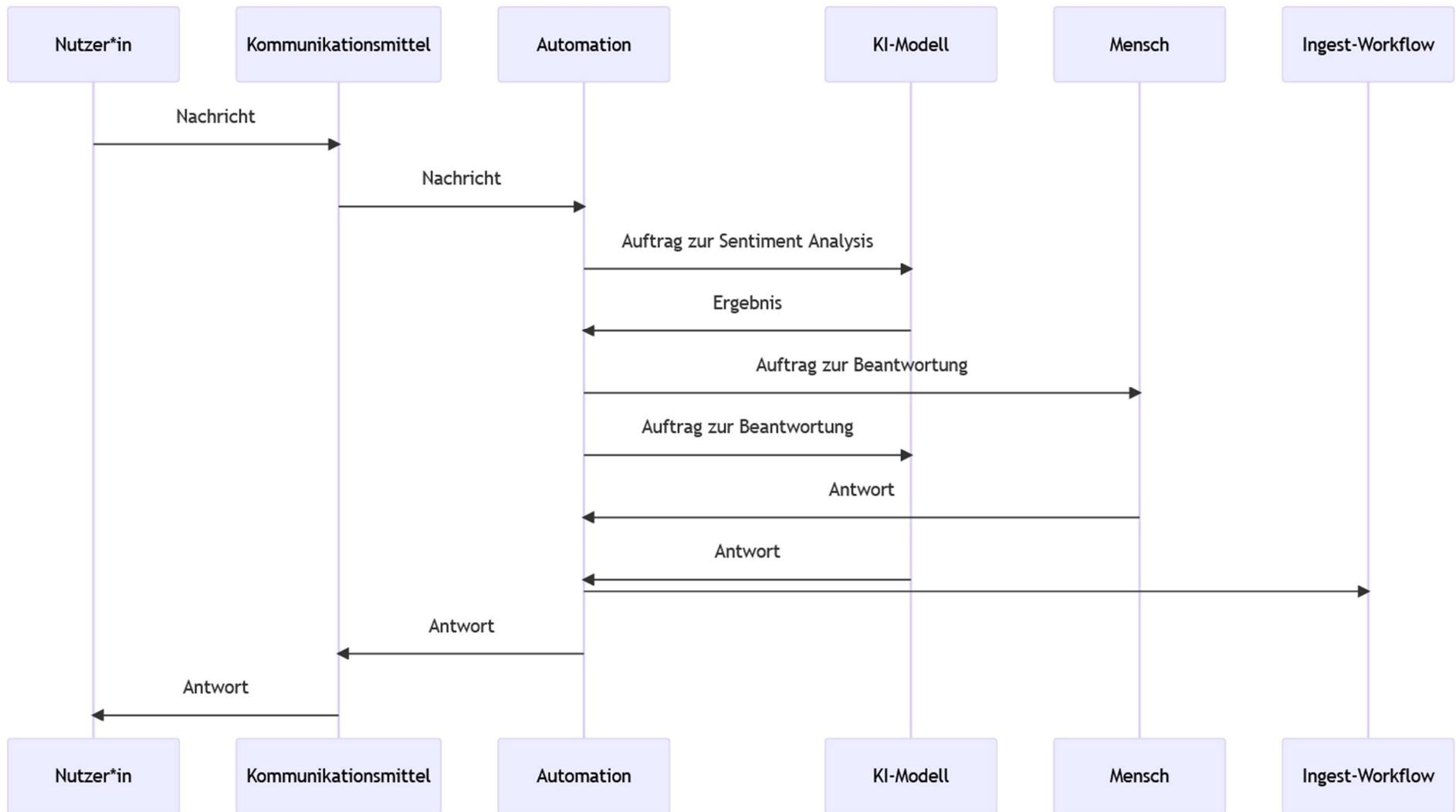
Informationsfluss 4: Abschnitt 7.2 Assetbearbeitung [Eigene Darstellung]



Informationsfluss 5: Abschnitt 7.2 Assetgeneration [Eigene Darstellung]



Informationsfluss 6: Abschnitt 7.2 Personalisierte Ausspielung [Eigene Darstellung]



Informationsfluss 7: Abschnitt 8.5.1 Nutzer*innenanfrage [Eigene Darstellung]

ANHANG 8: API-PROGRAMMIERVORSCHLÄGE

Dieser Anhang bietet Codeausschnitte, die für die API-Verbindung zwischen Node-RED und ChatGPT benötigt werden. Es wird der Fall abgedeckt, in dem ChatGPT als KI-Modell zum Einsatz kommt sowie der Fall, in dem InstructGPT verwendet wird. Mit ChatGPT bzw. InstructGPT werden die einzelnen Use Cases realisiert. Dafür werden zunächst Variablen, die in die Funktion eingesetzt werden, definiert. Informationen wie Prompt und Ergebnis werden in Node-RED Messages zum und vom jeweiligen KI-Modell übergeben. ChatGPT sucht sich seine Funktion je nach Use Case, der im Prompt beschrieben ist, aus. Bei InstructGPT muss das gewünschte Modell bzw. der Use Case von der Nutzerin bzw. vom Nutzer bereits bei der Anfrage ausgewählt werden.

VARIABLEN

Damit die jeweiligen Use Cases abgebildet werden können, werden beispielhaft für jeden Use Case Ein- und Ausgabevariablen definiert. Die endgültige Auswahl an Variablen entwickelt sich aus den vorherrschenden Strukturen im Medienunternehmen sowie aus dem praktischen Einsatz.

Ideengeber

Eingabevariablen

Variable	Typ	Beschreibung
General_topic	String	Überthema, für das Ideen ausgegeben werden sollen
Number_of_ideas	String	Anzahl der Ideen, die ausgegeben werden sollen
Ignore_idea	String	Ideen, die bereits bekannt sind und ignoriert werden sollen
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Ideas	List	Liste der Ideen
Ideas.topic	String	Thema der Idee
Ideas.description	String	Beschreibung der Idee
Ideas.target	String	Zielgruppe für einen möglichen Artikel

Zusammenfassungen

Eingabevariablen

Variable	Typ	Beschreibung
Id	String	Identifikation der Ausgangsessenz
Desired_length	Float	Geforderte Länge der Zusammenfassung
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Summary	String	Zusammenfassung
Length	Float	Länge der Zusammenfassung

Stilistische Korrektur

Eingabevariablen

Variable	Typ	Beschreibung
Id	String	Identifikation der Ausgangsessenz
Applied_correction	String	Angewandte Veränderung auf den Text z.B. Text wird in Umgangssprache gesetzt
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Old_id	String	Identifikation der Ausgangsessenz
New_id	String	Identifikation der neu geschaffenen Essenz
Title	String	Titel der Essenz
Body	String	Inhalt der Essenz
Length	Float	Länge der Essenz
Correction	String	Angewandte Veränderung

Übersetzung

Eingabevariablen

Variable	Typ	Beschreibung
Id	String	Identifikation der Ausgangsessenzenz
Desired_language	String	Zielsprache
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Old_id	String	Identifikation der Ausgangsessenzenz
New_id	String	Identifikation der neu geschaffenen Essenzenz
Title	String	Titel der Essenzenz
Body	String	Inhalt der Essenzenz
Length	Float	Länge der Essenzenz
Language	String	Sprache der neuen Essenzenz

Texte generieren aus Prompt

Eingabevariablen

Variable	Typ	Beschreibung
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Id	String	Identifikation der neu geschaffenen Essenzenz
Title	String	Titel der Essenzenz
Body	String	Inhalt der Essenzenz
Length	Float	Länge der Essenzenz
Prompt	String	Prompt, aus dem die Essenzenz entstanden ist

Texte generieren aus Datenbankeinträgen

Eingabevariablen

Variable	Typ	Beschreibung
Database	String	Datenbank, aus der die Informationen stammen
Topic	String	Kategorie der Daten
Date	Int	Aktuelles Datum
Location	String	Sendegebiet
Data	List	Daten aus der Datenbank
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Id	String	Identifikation der neu geschaffenen Essenz
Title	String	Titel der Essenz
Body	String	Inhalt der Essenz
Length	Float	Länge der Essenz

Schlagwörter extrahieren

Variable	Typ	Beschreibung
Id	String	Identifikation der Ausgangsessenzen
Number_of _buzz_words	Int	Anzahl der geforderten Schlagwörter
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Buzz_words	List	Liste der Schlagwörter
Buzz_words.body	String	Schlagwort
Buzz_words.length	Float	Länge des Schlagworts
Buzz_words.category	String	Überkategorie des Schlagworts

Sentiment Analysis / Moderations

Eingabevariablen

Variable	Typ	Beschreibung
Id	String	Identifikation der Ausgangssensenz
Messages.content	String	Prompt des Nutzers

Ausgabevariablen

Variable	Typ	Beschreibung
Positive	Boolean	Abfrage, ob der Inhalt ein positives Gefühl überbringt
Neutral	Boolean	Abfrage, ob der Inhalt ein neutrales Gefühl überbringt
Negative	Boolean	Abfrage, ob der Inhalt ein negatives Gefühl überbringt
Hate	Boolean	Abfrage, ob der Inhalt Hass aufgrund von Ethnie, Geschlecht, Hautfarbe, Religion, Nationalität, sexueller Orientierung, Behinderung oder gesellschaftliche Schicht enthält
Hate/Threatening	Boolean	Siehe oben inklusive Gewalt oder ernsthafter Bedrohung
Self-Harm	Boolean	Abfrage, ob der Inhalt Darstellungen von selbstverletzendem Verhalten enthält
Sexual	Boolean	Abfrage, ob dieser Inhalt von sexuellen Inhalten handelt
Sexual/Minors	Boolean	Siehe oben inklusive der Beteiligung von Minderjährigen
Violence	Boolean	Abfrage, ob dieser Inhalt von Gewalt handelt
Violence/Graphic	Boolean	Abfrage, ob der Inhalt Gewalt übermäßig grafisch darstellt

CODE MIT CHATGPT

Die API-Anfragen an ChatGPT laufen über einen Endpunkt, da dieses Modell nur mit diesem kompatibel ist. Die Anfrage an die API wird immer mit der Methode POST an den URI `https://api.openai.com/v1/chat/completions` durchgeführt [92]. Um strukturierte Informationen zu erhalten, werden ChatGPT bei jeder Anfrage mehrere Funktionen mitgegeben. Das Grundgerüst wird unten dargestellt. Die blau markierten Teile des Codes ändern sich für die jeweiligen Use Cases.

```
import openai
import json
import os
openai.organization = "YOUR_ORG_ID"
openai.api_key = os.getenv("OPENAI_API_KEY")

"""Dies ist eine Beispielfunktion, die hier fest implementiert ist. Im
produktiven Einsatz wäre dies der Zugang zum MAM. """
def get_current_weather(location, unit="fahrenheit"):
    """Get the current weather in a given location"""
    weather_info = {
        "location": location,
        "temperature": "72",
        "unit": unit,
        "forecast": ["sunny", "windy"],
    }
    return json.dumps(weather_info)

def run_conversation():
    """ Schritt 1: Konversation und Funktionen an GPT senden. In diesem
    Beispiel wird nur eine Funktion mitgeschickt. """
    messages = [{"role": "user", "content": "What's the weather like in
    Boston?"}]
    functions = {
        "name": "get_current_weather",
        "description": "Get the current weather in a given location",
        "parameters": {
            "type": "object",
            "properties": {
                "location": {
```

```

        "type": "string",
        "description": "The city and state, e.g. San
        Francisco, CA",
    },
    "unit": {
        "type": "string",
        "enum": ["celsius", "fahrenheit"]
    },
},
"required": ["location"],
},
}
response = openai.ChatCompletion.create(
    model="gpt-3.5-turbo-0613",
    messages=messages,
    functions=functions,
    function_call="auto",
)
response_message = response["choices"][0]["message"]

# Schritt 2: Prüfen, ob eine Funktion auf den Prompt passt
if response_message.get("function_call"):
    # Schritt 3: Funktion aufrufen
    available_functions = {
        "get_current_weather": get_current_weather,
    }
    function_name = response_message["function_call"]["name"]
    function_to_call = available_functions[function_name]
    function_args = json.loads(
        response_message["function_call"]["arguments"]
    )
    function_response = function_to_call(
        location=function_args.get("location"),
        unit=function_args.get("unit"),
    )

```

```

# Schritt 4: Strukturierte Daten zu GPT senden
messages.append(response_message)
messages.append({
    "role": "function",
    "name": function_name,
    "content": function_response,
})
second_response = openai.ChatCompletion.create(
    model="gpt-3.5-turbo-0613",
    messages=messages,
)
return second_response

# Schritt 5: Konversation ausgeben
print(run_conversation())

```

Function Calling nach [96] und [125]

Programmierbeispiele mit den eingefügten Variablen befinden sich auf den Datenträger in Kapitel elf.

CODE MIT INSTRUCTGPT

Die Modelle von InstructGPT geben von sich aus ihre Daten in strukturierter Form aus, nicht in textueller. Daher muss hier kein Function Calling angewandt werden. Es müssen auch keine vorbereiteten Variablen übergeben werden. Die Struktur der Variablen kann jedoch trotzdem zur Übergabe der Informationen durch Node-RED Messages genutzt werden. Nachfolgend sind hier die Modelle mit ihren Fähigkeiten beschrieben.

```

import os
import openai
openai.organization = "YOUR_ORG_ID"
openai.api_key = os.getenv("OPENAI_API_KEY")

```

Beginn einer API Abfrage [94]

Sprachkomplettierung / Completion

Aufgabe	Endpunkt	Methode
Create Completion	https://api.openai.com/v1/completions	POST

[92]

```
completion = openai.Completion.create(  
    model="text-davinci-003",  
    prompt="Hello world"  
)
```

Anfrage an die Completions-API - Codeausschnitt in Python [126]

Die Antwort wird in folgendem Format gegeben:

```
"choices": [{  
    "finish_reason": "length",  
    "index": 0,  
    "logprobs": null,  
    "text": "\n\n\"Hello World!\"  
}],  
"created": 1683130927,  
"id": "cml-7C9Wxi9Du4j1lQjdjhxB1O22M61LD",  
"model": "text-davinci-003",  
"object": "text_completion",  
"usage": {  
    "completion_tokens": 16,  
    "prompt_tokens": 10,  
    "total_tokens": 26  
}}
```

Antwort auf eine Anfrage an die Completions-API – Codeausschnitt in Python [126]

Mit dem folgenden Befehl wird explizit die textuelle Antwort „Hello World“ abgefragt und ausgegeben.

```
print(completion.choices[0].text)
```

Ausgabebefehl zur Anfrage an die Completions-API – Codeausschnitt in Python [126]

Bildgeneration und -veränderung / Images

Aufgabe	Endpunkt	Methode
Create Image	https://api.openai.com/v1/images/generations	POST
Create image variation	https://api.openai.com/v1/images/variatiions	POST
Create image edit	https://api.openai.com/v1/images/edits	POST

[92]

```

image_resp = openai.Image.create(
    prompt="two dogs playing chess, oil painting", n=4, size="512x512"
)
image_url = response["data"][0]["url"]

```

Bildgeneration - Codeausschnitt in Python [126]

Hier wird das Ergebnisbild als URL ausgegeben und in der Variable `image_url` gespeichert.

Archivanwendungen und Personalisierung / Embeddings

Aufgabe	Endpunkt	Methode
Create Embeddings	https://api.openai.com/v1/embeddings	POST

[92]

```

text_string = "sample text"
model_id = "text-similarity-davinci-001"
response = openai.Embedding.create(
    input=text_string,
    model=model_id
)
embedding = ['data'][0]['embedding']

```

Embeddings – Codeausschnitt in Python [126]

Die Antwort wird in der Variable `embeddings` gespeichert. Die Antwort besteht aus dem Vektor, der die genaue Position des Prompts im mehrdimensionalen Vektorraum angibt. Weitere Methoden, die Klassifikation, Clustering oder semantische Textsuche mithilfe von Embeddings würden den Rahmen dieser Arbeit sprengen, daher ist in diesem Fall auf die Quelle [126] und ihre verlinkten Seiten verwiesen.

Transkription / Audio

Aufgabe	Endpunkt	Methode
Create Transcription	https://api.openai.com/v1/audio/transcriptions	POST

[92]

```

f = open("path/to/file.mp3", "rb")
transcript = openai.Audio.transcribe("whisper-1", f)

```

Audio Transkription – Codeausschnitt in Python [126]

Die Variable `transcript` speichert den transkribierten Text und kann der Audioaufnahme als Metadaten angehängt werden.

Sentiment Analysis / Moderations

Aufgabe	Endpunkt	Methode
Create Moderation	https://api.openai.com/v1/moderations	POST

[92]

```
moderation_resp = openai.Moderation.create(  
    input = "Here is some perfectly innocuous text that follows all OpenAI  
    content policies."  
)  
output = response["results"][0]
```

Moderations - Codeausschnitt in Python [126]

Die Antwort wird in folgender Form gegeben:

```
"id": "modr-XXXXX",  
"model": "text-moderation-001",  
"results": [{  
    "categories": {  
        "hate": false,  
        "hate/threatening": false,  
        "self-harm": false,  
        "sexual": false,  
        "sexual/minors": false,  
        "violence": false,  
        "violence/graphic": false  
    },  
    "category_scores": {  
        "hate": 0.18805529177188873,  
        "hate/threatening": 0.0001250059431185946,  
        "self-harm": 0.0003706029092427343,  
        "sexual": 0.0008735615410842001,  
        "sexual/minors": 0.0007470346172340214,  
        "violence": 0.0041268812492489815,  
        "violence/graphic": 0.00023186142789199948  
    },  
    "flagged": false  
}]
```

Detaillierte Antwort auf Anfrage an Moderations-API – Codeausschnitt in Python [126]

Zunächst werden für die Kategorien bool'sche Werte abgegeben. Die genauen Ergebnisse im Datenformat float in sogenannten Scores angegeben. Außerdem wird der bool'sche Wert `flagged` ausgegeben. Dieser gibt an, ob diese Essenz eine jugendschutzkritische Eigenschaft hat. Die Antwort wird wie folgt abgefragt:

```
output = response["results"][0] #Binäre Werte
output = response["results"][1] #Scores
output = response["results"][2] #flagged
```

Strukturierte Antwort auf Anfrage an Moderations-API – Codeausschnitt in Python [126]

Nachtraining / Finetuning

Aufgabe	Endpunkt	Methode
Create fine-tune	https://api.openai.com/v1/fine-tunes	POST
List fine-tunes	https://api.openai.com/v1/fine-tunes	GET
Retrieve fine-tune	https://api.openai.com/v1/fine-tunes/:fine tune id	GET
Cancel fine-tune	https://api.openai.com/v1/fine-tunes/:fine tune id/cancel	POST
List fine-tune events	https://api.openai.com/v1/fine-tunes/:fine tune id/events	GET
Delete a fine-tuned model	https://api.openai.com/v1/models/:model	DELETE
Use a fine-tuned model	https://api.openai.com/v1/completions	POST

[92]

Eine detaillierte Erklärung zum Finetuning sowie Codeausschnitte für die jeweiligen Situationen würden den Rahmen dieser Arbeit sprengen, daher ist an diesem Punkt auf die Quelle [126] und ihre verlinkten Seiten verwiesen.

ANHANG 9: ANFORDERUNGEN AN DIE GRAFISCHEN BENUTZUNGSOBERFLÄCHEN

	Zugangsprofil	Eingabe	Ausgabe	Technischer Hintergrund	Feedbackbutton	Überwachung	Zugang zum MAM
A	Text	Text	Text	Nein	Ja	Nein	Ja
B	Archiv I	Text	Text, Daten	Nein	Ja	Ja	Ja
G	Archiv II	Text, Audio, Video	Text, Daten	Nein	Ja	Ja	Ja
H	Audio	Audio	Text	Nein	Ja	Nein	Ja
I	Grafik	Text	Bild	Nein	Ja	Nein	Ja

Beispielhafte Zeichnungen befinden sich auf den Datenträger in Kapitel sieben.

Zugangsprofil: Mit welchem Zugangsprofil erreicht man die GUI?

Eingabe: Welche Medienform wird eingegeben?

Ausgabe: Welche Medienform wird ausgegeben?

Technischer Hintergrund: Hat die Person, die mit der GUI arbeitet, einen technischen Hintergrund?

Feedbackbutton: Benötigt die Anwenderin bzw. der Anwender eine Möglichkeit, Feedback zu hinterlassen?

Überwachung: Wird über die GUI ein automatischer Prozess überwacht?

Zugang zum MAM: Kann über diese GUI das MAM erreicht werden?

ANHANG 10: BEISPIELE FÜR PROMPT ENGINEERING

Anwendung	Beispielprompt	Templates möglich?	Verbesserter Prompt
Themenvorschläge	„Gib mir mögliche Ansatzpunkte für einen Artikel über das Essen an Schulen.“	Ja	Wenn du Journalist wärst und einen Artikel über das Mensaessen an deutschen Schulen schreiben würdest, welche Themen könnte der Artikel behandeln?
Zusammenfassung	„Fasse den Text zusammen.“	Ja	„Extrahiere die wichtigsten Aussagen aus dem Text und formuliere damit einen neuen, aber kürzeren Text.“
Übersetzungen	„Übersetze diesen Text ins Deutsche.“	Ja	„Übersetze den Text in formelles Deutsch, das in Deutschland gesprochen wird. Nimm die vorherige Übersetzung als Beispiel.“
Chatbot	„Nenne mir bitte die Adresse des Medienunternehmens.“	Nein	Stell dir vor, du möchtest einen Leserbrief an das Medienunternehmen schicken, an welche Postadresse würdest du den schicken?
Textgeneration	„Fasse diese Verkehrsdaten in ganze Sätze.“	Ja	„Fasse diese Verkehrsdaten in ganze Sätze. Orientiere dich stilistisch an den gestrigen Nachrichten.“
Metadatenextraktion / Extraktion von Schlagwörtern	„Nenne die Schlagworte aus diesem Text.“	Ja	„Nenne die Schlagworte aus diesem Text. Ich möchte mindestens 10 Schlagworte.“
Sprachliche Verbesserungen	„Schreibe diesen Text um in formelle Sprache.“	Ja	„Schreibe diesen Text um in formelle Sprache. Lass dir Zeit mit der Antwort und wähle deine Formulierungen gründlich.“
Sentiment Analysis	Welches Gefühl drückt dieser Satz aus?	Ja	„Drückt dieser Satz Aggression aus?“ „Gibt dieser Satz Hinweise auf eine Straftat?“