

Inhalt



Selbstlernkurs

In der ARD-Normdatenbank gibt es jetzt einen interaktive Selbstlernkurs für die Personenansetzung



Schwerpunkt KI in der Praxis

Für viele Menschen - auch in unserer Branche - ist KI immer noch ein Buch mit vielen Siegeln. Es erschließt sich nicht immer sofort, wozu "man das braucht"... Daher ist der Schwerpunkt dieses Heftes darauf ausgerichtet, praktische Beispiele zu finden, wie KI in der Praxis eingesetzt werden kann und an welchen Stellen sie wirklich nutzbringend ist



Wie kann das Zusammenspiel von Mediendokumentation und KI funktionieren, wo kann Automatisierung verbessert werden, wo sind Verschlagwortung und inhaltliche Erschließung zu erleichtern?



Und natürlich muß überall mitbedacht werden, welche Auswirkungen der Einsatz von KI auf die Gesellschaft haben kann und wohin dies in der Zukunft führen wird



Hybrid nach Freiburg

Berichte über die Freiburger Frühjahrstagung des vfm vom 22.-24. April 2024



Rückblick

"Arbeiterleichterung" durch Beschallung in Unternehmen und Betrieben ab dem Ende der 1920er Jahre



Das Letzte

Im Jahr 2074: Rückblick auf die Zeiten der Entdeckung Künstlicher Intelligenz in der Mediendokumentation



Automatisierte Erschließung

Automatisierung der Formalerschließung von Manuskripten und weiterem Schriftgut im BR mit der Software-Anwendung MES



Ticket in die Zukunft

Um den Service zu optimieren wurde für den IDA-Desk beim SWR ein Ticketsystem konzipiert

Inhalt

Editorial

2 Innovation & Dialog

Thiemo Kremser

Aktuell

3 Aktuelle Meldungen

Schwerpunkt KI und Praxis

4 Bis zum Hals im digitalen Öl?

Öffentlich-rechtliche Rundfunkarchive als Lieferanten für KI-Trainingsdaten

Brecht Declercq

9 Künstliche Intelligenz trifft auf Medien-dokumentation

Erfahrungen und Perspektiven aus dem TIB AV-Portal – Plattform für wissenschaftliche Videos

Dr. Sven Strobel

14 Der digitale Archivar

Lessons learned aus einem KI-Erschließungsprojekt

Michael Koltan

19 „Zeitfunk“ – Sendungen des Rundfunks der DDR

Exploration automatisierter Inhaltserschließung von historischen Tondokumenten im Deutschen Rundfunkarchiv als Proof of Concept

Claudia Hunold und Alexander Wolff

26 Automatisierte Playlisterstellung in der Musikdokumentation

KI-basierte Analyse und Verschlagwortung von Songtexten

Sabrina Ladenburger

31 AI Next Level

Was Künstliche Intelligenz alles verändern wird

Prof. Dr. Hannah Bast

Rezensionen

32 Künstliche Intelligenz und ihre Auswirkung auf die Gesellschaft

Rezension von Stefan Holländer

FaktenFokus

33 FaktenFokus für Faktencheck und Verifikation

Thiemo Kremser

Marianne-Englert-Preis

34 newcomer-forum im vfm

Neues aus den Hochschulen

Heiko Linnemann

35 Gewusst wie

Ein interaktiver Selbstlernkurs für die Personenansetzung in der ARD-Normdatenbank

Jasmin Casarano

39 „Automate the boring stuff...“

MES-Software-Anwendung für die automatisierte Formalerschließung von Manuskripten im BR

Stefan Merl

43 Ticket in die Zukunft

Konzeption eines Ticketsystems als Optimierung für den IDA-Desk von SWR und SR

Lena Cara Wernhöfer

Seminare

51 Wir hinken wahnsinnig hinterher

Der vfm verstärkt seinen Weiterbildungsauftrag im Bereich KI mit Seminarankündigung

Herbert Staub

Tagungen

55 Hybrid nach Freiburg

Impressionen und Tagungsbericht von der vfm-Frühjahrstagung vom 22.-24. April 2024

Eva Krause, Thiemo Kremser und Ute Mader

57 Doku ex Machina

Nutzung von künstlicher Intelligenz und Automatisierung in der Mediendokumentation – Tagungsbericht

Eva Krause, Florian Reichinger, Jonas Wagner

71 Begrüßung

auf der vfm-Frühjahrstagung in Freiburg

Vanessa Sautter

Rückblick

72 „Zweifellos schafft also der Rundfunk frohe Menschen ...“

Musik aus dem Radio in den Fabrikhallen

Jörg-Uwe Fischer

Das Letzte

75 Nostalgische Reise

in den Newsroom und ins Medienarchiv von 2024

Eva Krause und Ute Mader

Innovation & Dialog

Thiemo Kremser



Thiemo Kremser
Hessischer Rundfunk
Frankfurt a. M.
Thiemo.Kremser@info7.de

Liebe Mitglieder, liebe Kolleginnen und Kollegen,

die diesjährige Frühjahrstagung hat die Dynamik und Innovationskraft unseres Berufsfeldes erneut eindrucksvoll unterstrichen. Die Vorträge und Diskussionen präsentierten spannende Anwendungsfälle neuronaler Netze und Einsatzmöglichkeiten von generativer KI und großen Sprachmodellen (LLMs) in Medienarchiven.

Neben dem Hauptfokus auf Künstliche Intelligenz in Medienarchiven sowie neuen Aufgabenfeldern und (alten) Herausforderungen im Metadatenmanagement spannte die Tagung einen weiten Bogen über aktuelle und zukunftsweisende Themen- und Aufgabengebiete. Neben großartigen Vorträgen, einer herausragenden Podiumsdiskussion und sehr guten Gesprächen waren für mich persönlich vor allem die innovativen und explorativen Lösungsansätze zur automatisierten Inhaltserschließung ein Highlight der Konferenz. Diese wertvollen Einblicke, Diskussionen und persönlichen Gespräche werden sicherlich auf dem ein oder anderen Weg zur Weiterentwicklung unseres Arbeitsfeldes beitragen.

Die aktuelle Ausgabe der info7 knüpft direkt an die erfolgreiche Frühjahrstagung 2024 an und bietet eine Fülle von Anregungen für die Entwicklung und Erforschung von KI-Anwendungen in der Praxis. Die Essenz der Tagung – sozusagen hier noch einmal zum Nachlesen. Neben der Tagungsdokumentation möchte ich Ihnen besonders den Aufsatz des FIAT/IFTA-Präsidenten Brecht Declerq über öffentlich-rechtliche Rundfunkanstalten als Anbieter von KI-Trainingsmaterial empfehlen. Eine Frage, über die Sie in Ihrem Unternehmen vielleicht auch schon nachgedacht haben? Vielleicht haben Sie nach der Lektüre des Aufsatzes Anmerkungen, Kritik oder eine andere Meinung, die Sie uns mitteilen möchten? Schreiben Sie uns! Last but not least möchte ich an dieser Stelle besonders auch auf die spannenden Beiträge der Preisträgerinnen und Preisträger

des Marianne-Englert-Preises 2024 hinweisen, die auch ohne den Einsatz von KI wichtige Aspekte unseres Berufsfeldes innovativ beleuchten.

Nach all den lobenden Worten über die diesjährige Tagung möchte ich aber auch auf einen kritischen Punkt hinweisen, der mir persönlich während der Tagung aufgefallen ist, nämlich die ungleiche Beteiligung an den Diskussionen und Fragerunden. Trotz der hybriden Ausrichtung der Veranstaltung, die theoretisch eine sehr breite Beteiligung ermöglichen sollte, gab es leider nur wenige aktive Teilnehmer in den Fragerunden. Dieses Ungleichgewicht wirft Fragen über die Beteiligung und den Austausch innerhalb unserer Gemeinschaft auf. Meiner Meinung nach ist es wichtig, dass wir als Fachleute eine inklusive Atmosphäre schaffen, in der sich jeder ermutigt fühlt, Fragen zu stellen und Beiträge zu leisten.

Die geringe Beteiligung könnte auf Unsicherheiten bezüglich der eigenen Kompetenz oder auf technische Barrieren bei der Online-Beteiligung zurückzuführen sein. Es liegt jedoch in unserer Verantwortung im vfm, solche Barrieren zu erkennen und zu überwinden. Es ist daher wichtig, dass wir uns Gedanken darüber machen, wie wir die Diskussionen und Fragen zu den Vorträgen auf der Tagung besser strukturieren und insbesondere die Online-Teilnehmer effektiver einbinden können, um eine gleichmäßigere Beteiligung zu gewährleisten. Bitte schreiben Sie mir gerne eine Nachricht, wenn Sie Ideen haben, wie wir Fragerunden und Diskussionen effizienter und integrativer gestalten können.

Denn angesichts der rasanten technologischen Entwicklungen in unserem Bereich ist der Austausch von Wissen und Ideen wichtiger denn je. Ich ermutige daher alle, sich aktiv an diesem Diskurs zu beteiligen, sei es durch Vorträge, Fragen, info7-Artikel oder insbesondere durch die Teilnahme an den Diskussionen und Fragerunden im Rahmen der künftigen Frühjahrstagungen.

Lassen Sie uns nicht nur interdisziplinär, sondern auch noch stärker kooperativ zusammenarbeiten, damit unsere Gemeinschaft so vielfältig, engagiert und informiert ist wie die innovative Themenvielfalt, die Sie in den Ausgaben der info7 finden. Ihre Meinungen und Ihr Fachwissen sind unverzichtbar, um unser gemeinsames Wissen zu erweitern und die Herausforderungen unserer Zeit zu meistern.

Herzliche Grüße,
Euer Thiemo Kremser

Aktuelle Meldungen

Faktencheck beim SPIEGEL mit KI-Unterstützung

Spiegel-KI-Chef Ole Reißmann auf dem EPC: „Wir nutzen Künstliche Intelligenz dazu, unsere eigenen Vorurteile zu entdecken“.

Medienhäuser sind heutzutage nahezu gezwungen, KI strategisch in ihre Geschäftsprozesse zu integrieren. Dies gilt im Grundsatz für alle Abläufe, sei es in den Redaktionen, den Publikations oder der Vermarktung sowie bei der Akquise neuer Zielgruppen.

Dieses Thema wurde unter anderem auf dem in Wien stattfindenden „European Publishing Congress“ behandelt.

Um den Gefahren, die KI birgt, etwas entgegenzustellen, hat der SPIEGEL ein eigenes Tool entwickelt. Auf dem Kongress hat SPIEGEL-Geschäftsführer Ole Reißmann darüber berichtet.

Texte werden beim Nachrichtenmagazin nicht von KI verfasst – aber im Rahmen der traditionell ausgeprägten Faktenprüfung soll das eigenentwickelte Tool dazu beitragen, die Nachteile von generativer KI auszugleichen. So werden im SPIEGEL die Defizite populärer KI-Anwendungen kritisch gesehen – etwa die Unzuverlässigkeit von Antworten auf Chat-GPT-Anfragen, die oft auch zu Falschangaben, zum „Halluzinieren“ oder zu KI-Lügen führen.

Die anfänglich große Euphorie über neue technische Möglichkeiten und Arbeitserleichterungen für Journalistinnen und Journalisten ist mittlerweile behutsamer Skepsis und einem vorwärts gerichteten Pragmatismus beim Finden tragfähiger KI-Anwendungen gewichen.

„Die Maschine atmet einmal alles ein, danach weiß sie nichts mehr Neues“, bemerkt Reißmann. Um dem entgegenzutreten, erlaubt der SPIEGEL den KI-Tools, mit denen das Haus aktuell arbeitet, Zugriff etwa auf die eigenen faktenbasierten Archive. KI schreibt beim SPIEGEL keine Texte. „Wir stehen für Originalität, nicht für Durchschnittlichkeit“, so Ole Reißmann, „für Exzellenz, nicht für Flüchtigkeitsfehler.“

Beim Fact Checking des SPIEGEL lässt Reißmann KI dagegen einsetzen: um die Arbeit der hauseigenen Dokumentationsabteilung in Teilen zu entlasten. Mit dem eigenentwickelten Tool können einzelne Artikel auf Satzebene geprüft werden, sodass jede Fakteninformation verifiziert werden kann. Erst wenn KI keine ausreichenden, belastbaren Belege finden kann, greife die eigentliche Dokumentationsabteilung ein.

Das Tool kann im Hintergrund Fehler finden und gibt Hinweise, denen die Dokumentation dann nachgehen kann – wenn sie nicht ohnehin schon fündig geworden ist. „Das Tool ist sozusagen ein Sicherungsnetz“, so Reißmann. „Als nächstes schauen wir, ob wir es auch der Redaktion zur Verfügung stellen wollen, zum Beispiel am Wochenende und in Randzeiten.“

Von einem weiteren KI-Tool beim SPIEGEL geht der Blick auch rückwirkend auf bereits publizierte SPIEGEL-Texte. So wird etwa untersucht, wie viele Frauen und Männer in früheren Texten vorkommen – in einer Art „Gender-Check“, um die Redaktionen zu sensibilisieren. „Wir nutzen KI dazu, unsere eigenen Vorurteile zu entdecken“, betont Ole Reißmann.

Mehr Informationen: [LINK
https://kress.de/news/beitrag/147600-spiegel-ki-chef-ole-reissmann-auf-dem-epc-quot-wir-nutzen-kuenstliche-intelligenz-dazu-unsere-eigenen-vorurteile-zu-entdecken-quot.html](https://kress.de/news/beitrag/147600-spiegel-ki-chef-ole-reissmann-auf-dem-epc-quot-wir-nutzen-kuenstliche-intelligenz-dazu-unsere-eigenen-vorurteile-zu-entdecken-quot.html)

Live-Faktencheck für Politik-Talkshows: h_da-Studierende entwickeln KI-Prototyp

Am 27. 06. findet das erste TV-Duell zwischen US-Präsident Joe Biden und Herausforderer Donald Trump statt. Künftig könnten solche Wortgefechte einem Echtzeit-Faktencheck unterzogen werden. Studierende der Hochschule Darmstadt (h_da) haben den Prototyp eines KI-Tools entwickelt, das in TV-Duellen oder Talkshows wie die von Markus Lanz (ZDF) oder Hart aber fair (ARD) die Aussagen von Politikern und weiteren Gästen unmittelbar überprüfen kann.

Zusammen mit dem Journalisten und Technologie-Experten Rafael Bujotzek haben die Onlinekommunikation-Studierenden einen Chatbot konzipiert, der mittels Spracherkennung den Live-stream der Sendung transkribiert und die Aussagen direkt auf Widersprüche und Fake News analysiert. Grobe Fehler und Widersprüche sollen so schnell erkannt und den Zuschauenden auf dem Bildschirm angezeigt werden können. Redakteurinnen und Redakteure können die erkannten Widersprüche oder Fake News wiederum schneller tiefergehend überprüfen und zusätzlich qualitätssichern.

„Das KI-Tool könnte dazu beitragen, die Transparenz, Genauigkeit und letztendlich die Qualität von Live-Diskussionen zu erhöhen“, sagt Rafael Bujotzek, Lehrbeauftragter am Mediacampus der h_da und früher selbst Reporter für das „heute-journal“ im ZDF. „Auf Smart-TVs oder in Mediatheken könnte diese Technologie Fake News und redaktionellen Fehlern vorbeugen.“ An Testsendungen haben die Studierenden ihren demonstrationsfähigen Prototyp erfolgreich erprobt und zuvor an Chat-GPT angebunden. Weitere Trainingsquellen sind denkbar.

Weitere Informationen: [LINK
https://nachrichten.idw-online.de/2024/06/20/live-faktencheck-fuer-politik-talkshows-h-da-studierende-entwickeln-ki-prototyp](https://nachrichten.idw-online.de/2024/06/20/live-faktencheck-fuer-politik-talkshows-h-da-studierende-entwickeln-ki-prototyp)

Bis zum Hals im digitalen Öl?

Öffentlich-rechtliche Rundfunkarchive als Lieferanten für KI-Trainingsdaten¹

Brecht Declercq



Brecht Declercq
President, FIAT/IFTA and
Head of Archives, Radiotele-
visione Svizzera Italiana
www.rsi.ch
brecht.declercq@rsi.ch

Die sehr umfangreichen Datenbestände der Öffentlich-Rechtlichen Rundfunkanstalten werden zunehmend gerne genutzt, um Anwendungen der Künstlichen Intelligenz (KI) zu trainieren. Dieser Trend hat an Bedeutung gewonnen, seit große Technologieunternehmen Zugang zu diesen reichen Datenbeständen suchen. In diesem Artikel werden daher strategische, rechtliche und ethische Aspekte dieser Entwicklung diskutiert.

Während einige Archive die sich daraus ergebenden finanziellen Vorteile begrüßen, bestehen andererseits Bedenken hinsichtlich des Schutzes geistigen Eigentums, der Verantwortung gegenüber dem Publikum und potenzieller Auswirkungen auf demokratische Prozesse durch KI-Algorithmen. Die Entscheidung, Archive für KI-Training zu öffnen, bleibt eine komplexe Herausforderung, die eine ausgewogene Berücksichtigung der genannten Faktoren erfordert.

Freiburg im Breisgau, Germany, late in April. The German Association of Media Documentalists (VfM) kicked off its annual gathering. If you are now thinking of a dust coat défilé, you're mistaken, yet it's agreed that this conference, delving into the depths of the German-speaking media landscape, remains largely under the radar for most. And yet, the focal point this year?

Artificial intelligence, naturally. One phenomenon has been brewing for weeks, whispered about in international press: a burgeoning awareness of the va-

lue of media archives for AI training. One archivist aptly remarked, 'We're neck-deep in the digital oil.'

Should public broadcast archives consent when approached by AI algorithm developers and exploiters, seeking access to their collections for algorithm training purposes? I'll explore this issue from strategic, legal, and ethical vantage points, but let's first set the scene.

The substantial value inherent in public broadcast archives has long been acknowledged. Over 15 years ago, some public broadcaster's archives endeavored to recoup the hefty, albeit one-time costs of digitizing archival audio and moving image carriers through sales platforms for archival footage and stock shots, a market valued at hundreds of millions of euros worldwide. Such endeavors failed miserably.

It demanded a shrewd sales strategy in a cut-throat market, and Anderson's long tail exacted its toll: the revenue from selling a few exceedingly popular snippets failed to cover the costs of preserving thousands of others that were never even asked for. It was a reaffirmation of an ancient law of the audiovisual archive economy: the price an archive must pay to be able to provide that unique fragment someone is longingly searching for, is having to preserve thousands of others nobody is ever interested in. The only problem being you never know beforehand which fragment it will be.

Nutzen = Wert

The capitalist logic above prompted many commercial media companies to reduce their archive departments to the bare bones. Publicly funded audiovisual archives, on the other hand, pivoted. If 'Return on Investment' couldn't cover the costs, then it ought to be 'Return on Society'. Tobias Golodnoff, a Dane who previously revolutionized the digitization market with a groundbreaking approach to public tendering, [posited the maxim 'use = value'](#), and it resonated as revolutionary.

¹ This article reflects the opinion of the author and not his employer's or FIAT/IFTA's. Thanks are due to the experts who provided comments on the manuscript: Sarah-Haye Aziz (Radiotelevisione Svizzera Italiana), Virginia Bazàn-Gil (Radio-televisión Española), Louise Broch (DR), Dale Grayson (Northbound Television), Maartje Hülsenbeck (Netherlands Institute for Sound and Vision), Thiemo Kremser (Hessischer Rundfunk) and Johan Oomen (Netherlands Institute for Sound and Vision).

To augment societal value, the utilization of archival collections had to be maximized. Publicly funded archives worldwide embraced his rationale and continue to do so until today. In countries that have the financial means for it, projects, platforms and infrastructures were devised allowing others to make use of the data: [media professionals](#), the [creative industry](#), the [general audience](#), [scientists and students](#), but also [teachers, pupils, tourists, DJs](#), even [Alzheimer's patients](#). And nothing indicates that the public's hunger for content from the treasure trove of public broadcasters' archives has waned. However, a different tune has emerged in recent months.

Wenn die Kapitalrendite die Kosten nicht decken kann, dann müsste es eine Rückgabe an die Gesellschaft geben

On December 27, 2023, it was revealed that The New York Times [filed a complaint against OpenAI](#), the company behind, among other products, [ChatGPT](#) and [DALL-E](#). OpenAI allegedly employed the Times' archive to train algorithms without their permission. The New York Times perceived it as an infringement on its intellectual property rights and seeks compensation totaling 4.5 billion USD. There isn't much scholarly literature on the value of newspaper archives, but heritage librarians, petitioning their Ministries of Culture for funding digitization projects for old newspapers, surely had that figure jotted down for their next grant application.

Die Jagd nach Trainingsdaten

These events shifted the perspective on the matter for many. OpenAI's management recently [declined to confirm](#) whether it utilized YouTube videos to train its algorithms, as it lacked the consent of owner Google—a strong competitor of theirs in the AI market. According to a [recent article](#) (once again) by The New York Times, Meta employees were even inclined to disregard copyright in internal discussions last year, in their pursuit of training data. Their need was (and is) simply too pressing.

Notwithstanding the foregoing, OpenAI has now opted to invest heavily in gaining access to media archives. In past months it struck deals with [Axel Springer](#), [Le Monde](#), [AP](#), and [Prisa Media](#), the [Financial Times](#) and many more thereby gaining access to not only Europe's largest commercial newspaper archives. Specific figures are rarely disclosed, but



insiders allude to a lump sum, supplemented with an annual fee, totaling tens of millions of euros in the next years.

A broader perspective seems to corroborate AI developers' willingness to pay. Market research firm Business Research Insights asserts that the [global market for AI training datasets](#) was already valued at 2.5 billion USD in 2023, with an expected growth to 30 billion USD by 2032. A third of this is represented by video collections, a quarter by audio collections. Critics highlight downward price pressure due to the inevitability of freely available high-quality data being [scraped](#) online. Yet, [impending legislative restrictions](#), especially in the EU, might bolster prices, while surely not all valuable data sources are readily accessible online.

Algorithmentwickler verbrauchen Daten schneller, als sie generiert werden. Diese Knappheit zwingt auch die Verwalter öffentlicher Datensätze, Stellung zu beziehen

Meanwhile, the data hunger of major players remains insatiable. According to Epoch Research Institute the freely available high-quality data is [nearly depleted](#). Algorithm developers are consuming data faster than it is being generated. This scarcity also forces the managers of public datasets to take a stand. The Dutch Royal Library does not want its data to be used, at least not by commercially inspired parties, but Wikipedia [chooses not to prevent it](#).

The audiovisual data of public broadcasters, however, is different. First of all due to its richness in information. A digitized radio archive of a million hours (and there are a few tens of them across Europe), contains an estimated number of words equivalent to 10 million newspaper pages. For comparison, The New York Times has produced an estimated 2.5 million pages since its founding in 1851. The television archives of Western European public broadcasters



are at least as vast as the aforementioned radio archives, and they evidently add to the audio the information in the form of moving images.

Additionally, the more reliable the data, the better [the results it yields](#). The most valuable data is that which is meticulously crafted, filmed, recorded by professionals. If alongside such audiovisual material, human descriptions of that material are also available, all the better. It will enable immediate validation of the results produced by the algorithms.

Jackpot for public broadcaster archives? We're not quite there yet. The market is evolving at breakneck speed, and nearly every day new information surfaces, altering the landscape. Nevertheless, it seems that the conventional relationship between broadcast archives and algorithm providers is undergoing a reversal.

Since the early 2010s, archives were the customers, paying large technology companies for the use of descriptive AI, typically per hour of audio or video transcribed, for example, using speech-to-text. Today, a consensus seems to have emerged that for those archives, no payment is due anymore for using algorithms to describe collections.

A few years ago already, at least one public broadcaster in Europe struck a deal with an AI provider, allowing the algorithm to be trained on archival material in exchange for transcripts. At least one big tech company now offers something more to those granting access to their archival material. While it may not be called a quid pro quo, it certainly resembles one: the big tech company digitizes certain types of videotapes from [public broadcaster archives for free](#), albeit with all related costs (selection, transportation, download and quality checks on files, project management, etc.) borne by the broadcasters themselves.

The European Broadcasting Union (EBU) lends legitimacy to these deals by offering a model contract. The terms and conditions stipulate that the company will not use the data generated in any way other than as specified, but whether AI algorithms are trained on it is often hard to verify. Many specialists harbor no illusions. Moreover, it remains questionable who would be willing to take on a legal battle against a big tech giant, even if the usage for AI training could be determined, in court. Since 2019, many European public broadcasters have been approached; while many have declined, at least two said yes.

Überlegen Sie es sich zweimal, bevor Sie sich verpflichten

The potential for significant financial gain associated with public broadcaster archives as datasets for AI algorithm training is evident. However, is it prudent for public broadcasters to hastily engage with AI developers and pursue maximum financial gain? Among specialists, this topic has been under discussion since 2018 when Virginia Bazàn-Gil, now Head of Archives at the Spanish broadcaster RTVE, delivered a groundbreaking presentation at the FIAT/IFTA World Conference in Venice. Bazàn-Gil made it clear that the sudden interest from AI companies was no coincidence and urged her colleagues, now besieged from all sides, [to think twice before committing](#). With the AI training dataset market clearly accelerating, it is time to further elaborate on Bazàn-Gil's wake-up call. Broadcasters considering to strike a deal should first address several strategic, legal, and ethical questions.

Strategically, the fundamental question is whether an archive access for AI training agreement would positively or negatively contribute to the strategic objectives of public broadcasters. On the benefits side, there is, of course, a sum of money, the magnitude of which is difficult to gauge given the rapidly evolving market. Some broadcasters might be inclined to offer preferential rates for research purposes.

However, one should be aware that algorithms developed by nonprofit entities (as for example OpenAI Inc. was, from its founding in 2015 until 2018) can easily find their way into more profit-driven contexts (e.g. via spin-offs). In addition to the financial aspect, such a deal could also generate goodwill among political circles and public opinion

for the involved broadcaster, particularly if domestic companies request access to archive data within the context of the development of a „sovereign“ LLM or other kind of algorithm, as is already being developed in some countries.

Sender, die über den Abschluss eines Vertrags nachdenken, sollten sich zunächst mit mehreren strategischen, rechtlichen und ethischen Fragen auseinandersetzen

Considering potential negative strategic consequences, reputational damage could occur. This might happen if the collaboration is perceived negatively by stakeholders, for instance, due to concerns that technological development is being pushed in the wrong direction, because the collaborating companies have a negative reputation, because the broadcaster is seen as overly profit-oriented, or because the deal is perceived as not transparent. Furthermore, some experts warn that providing archive content for AI algorithm training could lead to [long-term online invisibility](#) of journalistic work, potentially displacing traditional journalism as an information source for a great deal of media consumers.

From a legal perspective, several considerations must be taken into account. Firstly, there is the legal framework within which public broadcasters (and their archives) operate. In some countries (including for example [Switzerland](#)), the law obliges public broadcasters to grant access to their archives to anyone, profit-driven or not, as long as copyright is respected. Sometimes, such obligations are limited to the audio and video content itself, and sometimes metadata is included.

Additionally, it must be examined whether the respective public broadcaster is legally permitted to generate additional income. For example, in the EU, strict rules exist regarding this matter to prevent market distortion. One of the guiding principles is that [one-time revenues can only be used to cover one-time expenses](#). Parenthetically, this is why in several European countries, part of the [digitization of archives was financed](#) with revenues from the sale of broadcasting antenna infrastructure.

A third type of limitation could arise from existing contracts with third parties. For instance, some broadcasters have agreed with their speech-to-text providers that the transcriptions obtained through their algorithms cannot be passed on to

third parties. And a fourth type of limitation pertains to copyright and neighboring rights. As mentioned earlier, there is still much [uncertainty worldwide](#) about whether using archive material for algorithm training constitutes copyright infringement. If it does, this would severely limit the possibilities. Public broadcaster archives typically contain a substantial amount of material for which they are the sole rights holder, but filtering out unshareable data can incur significant costs, and the overall data volume would inevitably decrease.

Finally, there are restrictions imposed by [personal data protection legislation](#). Here as well, a possible obligation to filter out certain amounts of data could prove to be a costly endeavor.

If a public broadcaster decides to enter into a contract with an AI company, the following points deserve particular attention: covering costs for the technical setup of data transmission or filtering, indemnifying the organisation against third-party claims, and provisions regarding the potential resale of the algorithm itself or the datasets supplied or obtained by the archive.

Zweifellos eine moralische Entscheidung

Next to strategical and legal, also ethical considerations may also come into play. While much has been published on the ethical aspects of AI use, research has thus far focused mainly on the ethical aspects of descriptive and generative AI, rather than on the decision to deliver datasets to third parties to train their algorithm.

Nevertheless, the latter remains indisputably a moral choice. The expectation is that the influence of AI algorithms on society will significantly increase in the future. There is a risk that some generative AI algorithms will accelerate the spread of disinformation, pollute the information ecosystem and thus threaten democracies.

Should this encourage or discourage public broadcasters from contributing to their development? Public broadcaster values are implicitly or explicitly present in their archive material. Bias in this material is thus inevitable. Public broadcasters have provided a certain view of the society they have been reporting on throughout their history. The evolution of that view, as well as the evolution of society as a whole, are omnipresent in public broadcasters' archives.

Es besteht das Risiko, dass einige generative KI-Algorithmen die Verbreitung von Desinformation beschleunigen, das Informationsökosystem verschmutzen und damit Demokratien bedrohen

They contain vocabulary and opinions from the past that may no longer align with today's ethical standards. On top of that, it is very unlikely that our tomorrow's vocabulary and opinions will align with our ethical standards today. However, if we can choose between an algorithm influenced by content created by an institution that has been subject to democratic control, or the contrary, will we really choose the last option?

It seems advisable for the involved broadcaster to point out this implicit bias to the entity it intends to collaborate with, without assuming legal responsibility for any damage claims, as purifying the audio, video, and/or metadata of ideas, opinions, or language that no longer meet current ethical standards is simply an impossible task for an archive.

For the rest, it seems entirely legitimate to demand ethically responsible use of the dataset from every user and re-user, although it is not easy to precisely articulate what should and should not be permissible. Documents such as the EU's "[Ethics guidelines for trustworthy AI](#)" seem to be primarily directed towards those who create algorithms, rather than to those who provide the training data.

But they might be inspirational for public broadcasters who would like to contractually impose ethical conditions for a collaboration. It is my personal conviction that content created with public funds should not be privatized. Leaving aside copyright, exclusivity regarding the datasets provided or obtained seems incompatible with the public mandate of a public broadcaster.

Eine Haltung festlegen

Once all arguments have been considered, it is time to make a decision, ideally rooted in a corporate policy in this regard. Many broadcasters already have guidelines in place for the use of generative AI in production. However, few include provisions regarding data for training AI algorithms. Such guidelines inevitably also touch upon the overall strategy for opening up their archives.

While the dominant discourse throughout the last decade in the audiovisual archiving world was centered on openness, even the staunchest advocates

of such openness are now conceding to restrictions and respecting the choices of content creators who do not agree to make their material available for any form of reuse. For them, the principle of openness [no longer applies universally](#), to every potential audience. Before anything is made available, it seems to be legitimate again for the intentions to be checked.

When determining their stance and course of action, many public broadcasters will take the actions of major players into account. INA (which manages the archives of French public broadcasters), RTVE in Spain, and RAI in Italy have not yet announced their plans as far as is known. The BBC seems to be [hedging its bets](#). On March 21, it announced that it wants to train its own LLM based on its own (textual) archive material, but it is also negotiating with several technology companies to sell access to its archives for algorithm training. The one, apparently, doesn't have to exclude the other.

Schließlich scheint das Archivieren – bislang wirtschaftlich kaum beachtet – plötzlich auch in finanzieller Hinsicht sehr wertvoll zu werden

After careful consideration of strategic, legal, and ethical aspects, the question of the role of public broadcaster archives as AI training datasets remains complex and intriguing. After all, archiving, an activity that was never really considered very valuable in financial terms, suddenly appears to be so.

Although the lure of financial benefits may seem attractive, especially for those who are increasingly the victim of heavy budget cuts, significant risks and limitations must not be overlooked. One may wonder whether it is ethically acceptable to work with companies that are not known for their adherence to public values. It is a decision that every broadcaster and every archive can – and I would even say must – make for itself.

The evolution of this dynamic between archives and AI developers, symbolized by many ongoing negotiations, illustrates the delicate balance between intellectual property rights, ethical responsibilities, and funding models of public broadcasters. It is clear that archives are not only neck-deep in digital oil but also in a sea of complexity and challenges. Those who are navigating these new territories must not forget that, even amidst digital revolutions, the preservation of public values and ethical integrity remains an invaluable asset. ■

Künstliche Intelligenz trifft auf Mediendokumentation

Erfahrungen und Perspektiven aus dem TIB AV-Portal – Plattform für wissenschaftliche Videos

Dr. Sven Strobel

In den letzten 50 Jahren hat sich die Mediendokumentation von der Indexierung auf Karteikarten über die digitale Erfassung in Datenbanken hin zu hochkomplexen KI-Systemen erheblich weiterentwickelt. Das TIB AV-Portal, eine Open-Access-Plattform für wissenschaftliche Videos, nutzt fortschrittliche Technologien wie automatische Sprach-, Text- und Bilderkennung, um die Suche und Auffindbarkeit von Videoinhalten zu verbessern. Besonders hervorzuheben ist das KI-basierte Spracherkennungssystem Whisper von OpenAI. Whisper transkribiert gesprochene Sprache in maschinenlesbaren Text und bietet Übersetzungen ins Englische, wodurch die Durchsuchbarkeit und Zugänglichkeit der Videos erheblich verbessert wird. Der Einsatz solcher Technologien revolutioniert die Mediendokumentation, indem er eine effiziente und präzise Erschließung großer Mengen audiovisueller Inhalte ermöglicht.

Wandel in der Mediendokumentation

Die Mediendokumentation hat in den letzten 50 Jahren einen signifikanten Wandel durchlaufen. In den 1970er Jahren erfolgte die Indexierung von Medien ausschließlich manuell mithilfe von Karteikarten und gedruckten Katalogen. Mit der Digitalisierung und Computerisierung in den 1980er und 1990er Jahren wurden Medien zunehmend in digitalen Datenbanken indiziert, wenn auch weiterhin manuell. Diese Datenbanken waren zu jener Zeit noch lokal begrenzt, d. h., sie befanden sich auf den Festplatten einzelner Computer oder innerhalb lokaler Netzwerke.

Mitte der 1990er Jahre, mit dem Aufkommen des Internets und der ersten Online-Datenbanken, änderte sich dies grundlegend. In den 2000er Jahren erfolgte die Medienauswertung zunehmend in Media-Asset-Managementsystemen (MAM). Diese Systeme ermöglichen es, Medien und Metadaten an einem Ort zusammenzuführen. AV-Medien konnten nach Bild-

inhalten und O-Tönen ausgewertet werden, die im integrierten Player sofort abgespielt werden konnten. Mit dem MAM war es auch möglich, Sequenzen für den Schnitt zu bestimmen und digital zu transferieren. Dokumentare, Redakteure und Cutter konnten nun von überall auf den Content zugreifen und ihn weiter bearbeiten.

In den 2010er Jahren nahm die Automatisierung weiter zu, und erste Deep-Learning-Modelle wurden in die Analyse-Pipelines digitaler Dienste integriert. Obwohl die manuelle Indexierung weiterhin eine Rolle spielte, boten automatische Analysen wie Schnitt-, Sprach-, Text- und Bilderkennung mittlerweile eine erhebliche Unterstützung. In diesem Zeitraum entstand auch das TIB AV-Portal, bei dem die Technische Informationsbibliothek (TIB) von Anfang an ganz auf automatische Indexierung setzte.

Mit der Veröffentlichung der ersten Large-Language-Modelle der großen Player, insbesondere durch ChatGPT von OpenAI, erleben wir in den 2020er Jahren erneut einen erheblichen Wandel. Institutionen im Medienbereich integrieren verstärkt KI-Modelle in ihre Dienste, um diese zu optimieren. In allen Bereichen der Videoanalyse – von der Optical Character Recognition und Automated Speech Recognition über Shot Boundary Detection und Visual Concept Detection bis hin zum Named-Entity Linking – kommen jetzt Deep-Learning-Modelle zum Einsatz. Diese zeichnen sich durch ihre hohe Genauigkeit und Effizienz aus, was die Analyse- und Indexierungsprozesse erheblich verbessert. Wir stehen möglicherweise an der Schwelle zur vollautomatischen Indexierung, und es ist nicht auszuschließen, dass die manuelle Indexierung in weiten Teilen der Dokumentation bald der Vergangenheit angehören wird.



Dr. Sven Strobel
Product Owner (av.tib.eu)
TIB – Leibniz Information
Centre for Science and
Technology, Hannover
Research & Development
Sven.Strobel@tib.eu

Video in der Wissenschaft

Die Digitalisierung und das Internet haben die Nutzung von Videos in der Wissenschaft grundlegend verändert. Heute kann nicht nur jeder Videos konsumieren, sondern dank Smartphones und Open-Source-Software auch kostengünstig eigene produzieren.

Wissenschaftliche Videos sind längst nicht mehr nur ergänzendes Material zu Textpublikationen, sondern haben sich als primäre wissenschaftliche Quellen etabliert. Diese Entwicklung hat weitreichende Auswirkungen auf die Verbreitung, Vermittlung und Zugänglichkeit von Wissen.

Ein herausragendes Beispiel für den Einsatz wissenschaftlicher Videos findet sich in der Lehre. Universitäten und Bildungseinrichtungen nutzen Videos, um komplexe Konzepte anschaulich darzustellen und das Verständnis der Studierenden zu erleichtern. Online-Kurse und MOOCs (Massive Open Online Courses) bieten hochwertige Bildung weltweit und oft kostenlos an.

In der Forschung dienen Videos als wertvolle Dokumentationen von Experimenten und Untersuchungen. Zeitrafferaufnahmen von Zellkulturen etwa veranschaulichen das Zellwachstum und -verhalten über längere Zeiträume. Solche visuellen Daten helfen Forschern, Muster und Anomalien zu identifizieren, die in Textform schwer zu beschreiben wären. In der Umweltforschung dokumentieren Drohnenaufnahmen Veränderungen in Ökosystemen und unterstützen so die Analyse. Grundsätzlich bieten Videos eine dynamische Darstellung zeitlicher Entwicklungen und Bewegungen, die Texte oder Einzelbilder nicht leisten können.

Auch in der Wissenschaftskommunikation spielen Videos eine zentrale Rolle, beispielsweise in Form von Video-Abstracts. Diese kurzen, visuell ansprechenden Videos fassen die Kernaussagen wissenschaftlicher Studien zusammen. Sie bieten eine dynamische und leicht verständliche Alternative zu traditionellen Textabstracts, indem sie komplexe Inhalte durch Grafiken, Animationen und gesprochene Erklärungen veranschaulichen. So können Forscher ihre Ergebnisse effizient verbreiten und Laien sowie Fachkollegen einen schnellen Überblick über die jeweilige Studie ermöglichen.

Ein weiterer Vorteil von wissenschaftlichen Videos ist ihre Rolle im Bereich der offenen Bildungsressourcen (Open Educational Resources). Die freie Verfügbarkeit von Lehr- und Lernmaterialien in

Videoform demokratisiert Wissen und macht es einer breiteren Öffentlichkeit zugänglich. Studierende können durch diese Videos Lerninhalte flexibel und im eigenen Tempo vertiefen.

Die Nutzung von Videos in der Wissenschaft bringt zahlreiche Vorteile mit sich: Sie erleichtert die Vermittlung komplexer Inhalte, fördert das Verständnis durch visuelle Darstellungen und macht wissenschaftliche Erkenntnisse einem breiteren Publikum zugänglich. Zudem ermöglichen Videos eine bessere Dokumentation und Nachvollziehbarkeit von Forschungsergebnissen, was die Transparenz und Reproduzierbarkeit wissenschaftlicher Arbeiten erhöht. Durch die fortschreitende Digitalisierung und die zunehmende Verfügbarkeit von Videotechnologien wird diese Entwicklung in Zukunft weiter an Bedeutung gewinnen.

Das TIB AV-Portal für wissenschaftliche Videos

Das Videoportal der Technischen Informationsbibliothek (TIB AV-Portal, av.tib.eu) ist eine Open-Access-Plattform für wissenschaftliche Videos, die sich auf Naturwissenschaften und Technik konzentriert, aber inzwischen auch alle anderen Fachrichtungen abdeckt. Zu den Zielgruppen gehören Studierende, Lehrende und Forschende.

Zwischen 2011 und 2014 wurde das Portal gemeinsam von der TIB und dem Hasso-Plattner-Institut entwickelt und am 29. April 2014 live geschaltet. Seitdem hat es sich als produktiver Dienst etabliert, der zunehmend an Beliebtheit gewinnt: Die Nutzerzahl stieg von 19.000 im Jahr 2014 auf 660.000 im Jahr 2023.

Das AV-Portal bietet eine Vielzahl von Dienstleistungen an, darunter das Hosting und die Langzeitarchivierung von Videos, die Klärung von Rechten sowie Lizenzberatung. Jedes Video erhält einen Digital Object Identifier (DOI), der eine dauerhafte Zitierfähigkeit gewährleistet. Sprache, Text und Bildinhalte der Videos werden automatisch indexiert. Das Portal wird von einem Community-Manager redaktionell begleitet und kontinuierlich von einem Scrum-Team weiterentwickelt (vgl. Strobel 2024). Die Plattform ist offen zugänglich, werbefrei und DSGVO-konform, was sie im Vergleich zu kommerziellen Anbietern zu einer besonders vertrauenswürdigen und benutzerfreundlichen Quelle für wissenschaftliche Videos macht.

Der Inhalt des TIB AV-Portals umfasst rund 46.000 Videos, von denen etwa 42.000 unter einer

Sonografie Hals 5 von 6 - Glandula parotidea

Parenchym beurteilen (Links)
MHH Medizinische Hochschule Hannover

SERIE TRANSKRIPT ANNOTATIONEN

Transkript: Deutsch (automatisch erzeugt) ▾

schilddrüse × Suchen

8 von 23 Abschnitten angezeigt

03:02
gesundem **Schilddrüsengewebe**. Veränderungen können, ähnlich der **Schilddrüse**, diffuse oder fokal sein. Lymphknoten können physiologischerweise sowohl intra- als auch extra glandulär in der Parotis vorkommen, häufig präaurikulär oder im apikalen Drüsenanteil,

04:21
accessorische Drüsen möglich. Insgesamt haben wir gesundes Parenchym beobachtet, ohne Inhomogenitäten, wie fokale Läsionen oder Konkreme und ohne Duktectasien. Ein kleiner Exkurs zum anatomischen Verständnis. Der Musculus masseter liegt Rostrei des

5 Zitieren Teilen

Medizinische Hochschule Hannover (MHH)

Abbildung 1: Suche im gesprochenen Text des Videos

Creative-Commons-Lizenz stehen. Das Spektrum der Videos umfasst Vorlesungen und Konferenzen, Animationen und Simulationen, Experimente und Tutorials, Interviews und Dokumentationen, Mikroskopie und Makroaufnahmen sowie Video Abstracts. Zusätzlich stellt das Portal den Bestand des ehemaligen Instituts für den Wissenschaftlichen Film (IWF) bereit, der einen hohen kulturellen und wissenschaftshistorischen Wert besitzt (vgl. Technische Informationsbibliothek 2024). Herausgeber der Videos sind Konferenzen, Hochschulen, Bundesämter, Online-Journals, wissenschaftliche Gemeinschaften, Forschungsinstitute und herausragende Persönlichkeiten der Open-Science-Bewegung.

Mit diesen umfassenden Dienstleistungen und Inhalten stellt das TIB AV-Portal eine wichtige Ressource für die wissenschaftliche Gemeinschaft dar.

Videoanalyse und Suche

Im AV-Portal erfolgt keine manuelle Indexierung. Stattdessen stellen die Videoanbieter einige grundlegende Metadaten wie Titel und Autor über das Upload-Formular oder XML-Dateien bereit. Darüber hinaus wird eine Reihe automatischer Videoanalysen durchgeführt:

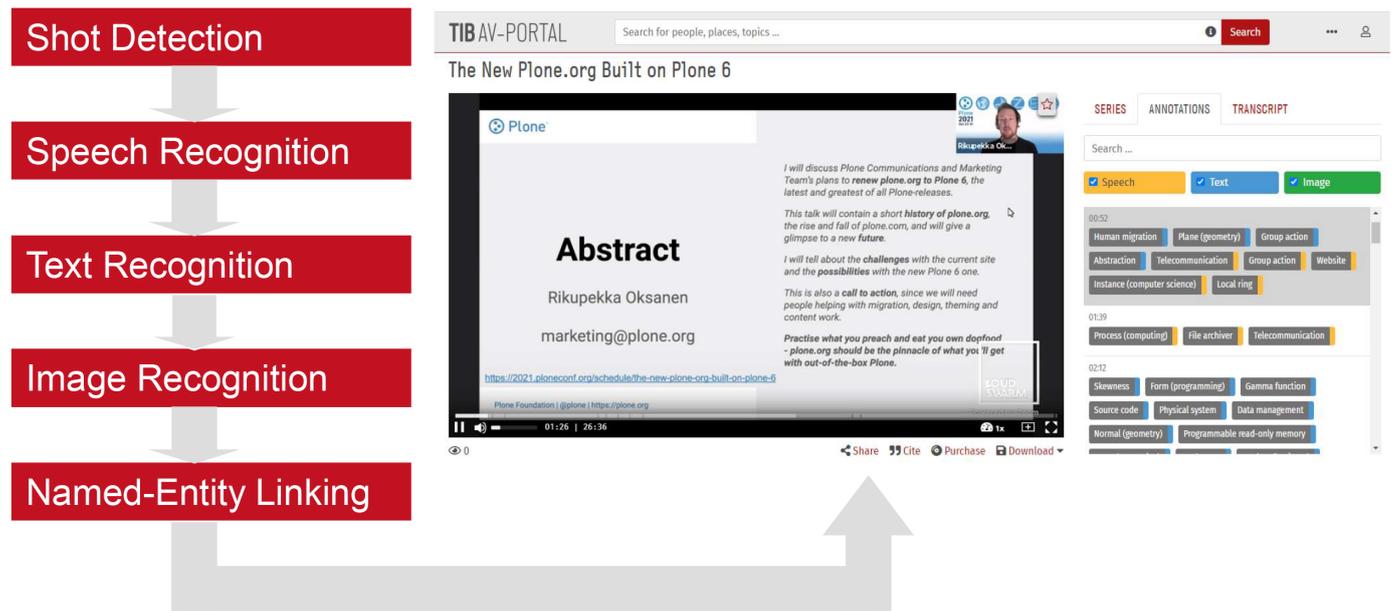
1. Die Shot Boundary Detection (PySceneDetect) teilt das Video in einzelne Segmente auf.
2. Die Automatische Spracherkennung (Whisper) wandelt gesprochene Sprache in maschinenlesbare Transkripte um.

3. Die Texterkennung (tesseract) extrahiert Text aus eingblendeten Folien und Bauchbinden.
4. Die Bildererkennung (eine Eigenentwicklung der TIB-Forschungsgruppe Visual Analytics) annotiert visuelle Konzepte wie „Satellit“ oder „chemisches Experiment“.
5. Das Named-Entity-Linking (eine Eigenentwicklung des Hasso-Plattner-Instituts) annotiert Entitäten wie Personen oder Sachbegriffe und verlinkt diese mit der Gemeinsamen Normdatei (GND, vgl. Deutsche Nationalbibliothek 2024).

Durch die Integration der GND, die Synonyme und englische Übersetzungen zu den Hauptbezeichnern enthält, können bei Suchanfragen sämtliche Begriffe mitabgefragt werden, was die Anzahl relevanter Treffer potenziell erhöht.

Whisper – open source

Seit Juli 2023 nutzen wir für die automatische Spracherkennung im AV-Portal Whisper von OpenAI. Das Modell transkribiert 98 Sprachen und übersetzt diese ins Englische (vgl. OpenAI 2024b). Es basiert auf tiefen neuronalen Netzen, die mit großen Mengen an Audiodaten (680.000 Stunden) und deren Transkriptionen trainiert wurden (vgl. Radford u. a. 2023). Diese Daten stammen aus verschiedenen Quellen und umfassen verschiedene Sprachen, Dialekte und Akzente. Dies ermöglicht es dem Modell, eine hohe Genauigkeit bei der Spracherkennung zu erzielen.



Video-Analyse

Im AV-Portal sind Englisch (67% des Bestands) und Deutsch (22% des Bestands) die häufigsten Sprachen (Stand: Juni 2024). Das large-v3-Modell von Whisper erreicht für Englisch eine Wortfehler-rate (Word Error Rate, WER) von 9,3 bei Material aus dem Common Voice 15 Dataset und 4,1 bei Material aus Fleurs (vgl. OpenAI 2024a). Die WER ist eine gängige Metrik zur Bewertung der Leistung von Spracherkennungssystemen. Eine WER von 9,3 bedeutet, dass Whisper durchschnittlich 9,3 Fehler pro 100 Wörter macht. Diese Fehler umfassen falsche Ersetzungen, fehlende Wörter oder zusätzliche Wörter. Eine WER von 9,3 ist ein guter Wert und deutet auf eine hohe Genauigkeit hin. Eine WER von 4,1 gilt als exzellent. Für Deutsch liegen die WERs des large-v3-Modells bei 5,7 (Common Voice 15) bzw. 4,9 (Fleurs) (vgl. OpenAI 2024a).

Whisper wurde als Open Source unter der MIT-Lizenz veröffentlicht (vgl. Open Source Initiative 2024), was uns die Integration des Codes im AV-Portal ermöglicht. Wir nutzen Whisper nicht nur zur Videountertitelung, sondern auch für die zielgenaue Suche und Auffindbarkeit der Videoinhalte. Das gesamte Whisper-Transkript wird indiziert und kann durchsucht werden. Über die Timestamps der Transkripte erfolgt die Zuordnung zu den einzelnen Segmenten des Videos. Nutzende können somit in den Transkripten suchen und gezielt zu den Stellen im Video springen, die den Suchbegriff enthalten (vgl. Abbildung 1). Da wissenschaftliche

Videos größtenteils über gesprochene Sprache ihre Inhalte transportieren, bietet die Transkription eine ausgezeichnete Möglichkeit, inhaltlich in den Videos zu suchen.

Halluzinationen – ein Problem von Sprachmodellen

Darüber hinaus nutzen wir Whisper, um unsere nicht-englischsprachigen Videos ins Englische zu übersetzen. Die Übersetzungen können neben der Originalsprache als Untertitel angezeigt und durchsucht werden. Dies unterstützt das mehrsprachige Verstehen und Suchen im Portal. Mit Englischkenntnissen können somit alle Videos im AV-Portal verstanden und durchsucht werden. Zudem bieten wir die Transkripte zum Download an, sodass sie in anderen Kontexten weiterverwendet werden können.

Bei der Implementierung waren wir vom Scrum-Team mit diversen Problemen konfrontiert. Ein häufiges Problem von Sprachmodellen sind Halluzinationen, also generierte Wörter oder Phrasen, die im Audioinput nicht vorhanden sind. Besonders in stillen Passagen und bei leichten Hintergrundgeräuschen traten Transkriptionen wie „Thanks for Watching“, „Bye-bye“ oder „Okay“ auf, die nicht geäußert wurden. Dieses Problem haben wir weitgehend reduziert, indem wir Faster Whisper –eine optimierte und leistungsfähigere Reimplementierung des Whisper-Modells –verwendet haben (vgl. SYSTRAN 2024). Faster Whisper setzt einen automatischen

Stillefilter ein, der alle stillen Passagen im Video überspringt, die zu Halluzinationen führen könnten. Ein zusätzlicher Vorteil von Faster Whisper ist, dass es viermal schneller ist als Whisper und deutlich weniger Speicher benötigt.

Ein weiteres Problem war der initial hohe Speicherverbrauch von Whisper, insbesondere bei langen Videos. Bei Videos über einer Stunde überschritten wir regelmäßig das 8 GB Speicherlimit unserer Kubernetes-PODs. Die Lösung bestand darin, lange Videos in zwei Teile aufzuteilen, die einzeln analysiert wurden. Dadurch entstanden zwei Transkripte, die am Ende zusammengefügt wurden. Diese Vorgehensweise hielt die Speicherlast konstant bei 7 GB.

Schließlich waren 15% der von Whisper erstellten Transkripte nicht für die Videountertitelung geeignet, da sie zu lange Cues (mehr als 150 Zeichen) enthielten. Dies führte dazu, dass das Video teilweise überdeckt wurde und die Untertitel nicht mehr richtig gelesen werden konnten. Um dieses Problem zu lösen, erstellten die Entwickler VTTs, bei denen die Cues auf maximal 150 Zeichen verkürzt wurden, ohne dabei die Sinnzusammenhänge zu vernachlässigen.

Fazit & Ausblick

In den letzten 50 Jahren hat die Mediendokumentation einen tiefgreifenden Wandel durchlaufen, geprägt durch technologische Fortschritte und die Integration künstlicher Intelligenz. Die Indexierung von Videos, die früher ausschließlich von Menschen durchgeführt wurde, erfolgt nun zunehmend automatisiert. KI-Technologien wie maschinelles Lernen und künstliche neuronale Netze setzen neue Maßstäbe und werden in Zukunft eine immer größere Rolle spielen. Eine unserer Produktvisionen ist es, stets modernste Techniken im TIB AV-Portal einzusetzen. Daher planen wir, in den kommenden Jahren verstärkt KI in unser Portal zu integrieren.

Bereits jetzt haben wir wichtige Projekte identifiziert: Derzeit stellen wir die visuelle Konzepterkennung auf Open-CLIP um, ein leistungsfähiges Modell für Bilderkennung und maschinelles Lernen (vgl. mlfoundations 2024). Parallel dazu evaluieren wir neuere KI-Modelle im Bereich Text Detection und Text Recognition. Zusätzlich planen wir eine Aktualisierung des Named-Entity Linking in Zusammenarbeit mit der TIB-Forschungsgruppe

Visual Analytics. Diese Gruppe experimentiert mit Wikidata und dem Wikifier (vgl. Wikifier 2024). Der Wikifier analysiert Texte, identifiziert wichtige Entitäten wie Personen, Orte oder Begriffe und verlinkt diese mit den entsprechenden Identifiern aus Wikidata, um zusätzliche Informationen bereitzustellen.

Wir vom Scrum-Team AV-Portal sind gespannt darauf, welche technologischen Innovationen die nächsten Jahre bringen werden. ■■■

Literatur

- Deutsche Nationalbibliothek (2024). Gemeinsame Normdatei (GND). Zugriff am 17. Juni 2024. url: https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html.
- mlfoundations (2024). GitHub: Open Clip. Zugriff am 17. Juni 2024. url: https://github.com/mlfoundations/open_clip.
- Open Source Initiative (2024). The MIT License (MIT). Zugriff am 17. Juni 2024. url: <https://opensource.org/licenses/mit>.
- OpenAI (2024a). GitHub: Whisper. Zugriff am 17. Juni 2024. url: <https://github.com/openai/whisper>.
- OpenAI (2024b). Platform: Speech to text. Zugriff am 17. Juni 2024. url: <https://platform.openai.com/docs/guides/speech-to-text/supported-languages>.
- Radford, Alec u.a. (2023). „Robust speech recognition via large-scale weak supervision“. In: International Conference on Machine Learning. PMLR, S. 28492–28518.
- Strobel, Sven (2024). TIB Blog: 10 Jahre wissenschaftliche Videoplattform: Die Geschichte des TIB AV-Portals von den Anfängen bis zur Gegenwart aus der Perspektive der Entwicklung. Zugriff am 17. Juni 2024. url: <https://blog.tib.eu/2024/04/08/10-jahre-wissenschaftliche-videoplattform-die-geschichte-des-av-portals-von-den-anfaengen-bis-zur-gegenwart-aus-der-perspektive-der-entwicklung>.
- SYSTRAN (2024). GitHub: Faster Whisper. Zugriff am 17. Juni 2024. url: <https://github.com/SYSTRAN/faster-whisper>.
- Technische Informationsbibliothek (2024). Medienbestand der ehemaligen IWF Wissen und Medien gGmbH. Zugriff am 17. Juni 2024. url: <https://www.tib.eu/de/research-entdecken/sondersammlungen/iwf-medienbestand>.
- Wikifier (2024). Semantic Annotation Service for 100 Languages. Zugriff am 17. Juni 2024. url: <https://wikifier.org>.

Der digitale Archivar

Lessons learned aus einem KI-Erschließungsprojekt

Michael Koltan



Michael Koltan
Archiv Soziale Bewegungen e.V.
Adlerstraße 12, 79098 Freiburg
Michael.Koltan@
archivsozialebewegungen.de

In diesem Vortrag geht es letztlich darum, die neuen Möglichkeiten, die *large language models* für die Erschließung digitalisierter Textmaterialien bieten, etwas genauer zu beleuchten. Dies wird anhand bestimmter Probleme diskutiert werden, die bei einem abgebrochenen (oder zumindest eingebrochenem) Projekt des Archivs Soziale Bewegungen in Freiburg aufgetreten sind.

Bevor ich aber auf dieses Projekt eingehe, werde ich zunächst einigermaßen grundsätzlich über KI reden. Dabei wird es um die Unterscheidung zwischen statistischer und symbolischer KI gehen. Zur statistischen KI gehören beispielsweise die aktuell so gehypten *large language models* wie ChatGPT, aber nicht ausschließlich. Ich werde ein bisschen ausführlicher auf statistische KI eingehen, als es eigentlich für diesen Vortrag notwendig wäre, um dem aktuellen Hype etwas entgegenzusetzen. Ich komme dann auf die andere Art der künstlichen Intelligenz, symbolische KI, zu sprechen, die einen von statistischer KI sehr verschiedenen Ansatz verfolgt. Diese Unterscheidung benötige ich, um dann unser Projekt, den Digitalen Archivar vorzustellen. Dieser hatte sich zur Aufgabe gesetzt, statistische mit symbolischer KI zu verbinden.

Ich gehe dann auf eines der Probleme ein, an dem der Digitale Archivar gescheitert ist. An einem konkreten Beispiel werde ich dann zeigen, wie dieses Problem durch *large language models* wie ChatGPT zwar nicht vollständig gelöst, aber doch zumindest deutlich entschärft werden kann.

Statistische KI beruht auf einem speziellen Programmierparadigma. Bei einem „normalen“ Programm findet erst einmal ein Gespräch mit einer Domänenexpertin statt, damit die Programmiererin versteht, mit welchen Gegenständen sie es zu tun hat und welche Regeln für diese Gegenstände existieren. Diese werden in Datenstrukturen und

Algorithmen umgesetzt, die dann weitgehend manuell ausprogrammiert werden. Jetzt haben wir ein Computerprogramm, das auf den Eingabedaten Operationen ausführt, die dann hoffentlich zum intendierten Ergebnis führen.

Das, was wir so gemeinhin künstliche Intelligenz nennen, darunter, aber nicht ausschließlich neuronale Netze, beruht auf einem anderen Paradigma: Wir haben zunächst einmal eine mehr oder minder große Datensammlung. Auf dieser Datensammlung legen wir eine Zielvariable und deren mögliche Werte an. Wenn wir es mit strukturierten Daten zu tun haben, handelt es sich bei dieser Zielvariable um eine der Eigenschaften des Datensatzes selbst, bei unstrukturierten Daten wird das gewünschte Ergebnis als Label hinzugefügt.

Statistische KI

Als nächstes teilen wir die Daten in Trainings- und Testdaten auf. Und schließlich erzeugt ein KI-Algorithmus ein Modell, das mit Hilfe statistischer Verfahren versucht, die Ausgangsdaten möglichst gut auf die gewünschte Zielvariable abzubilden. Das statistische Modell wird dann mit den Testdaten geprüft und wenn die Erfolgsquote suboptimal ist, wird so lange am Algorithmus herumgeschraubt, bis es einigermaßen passt. Am Ende steht ein Modell, das aus neuen Eingangsdaten mit Hilfe statistischer Korrelationen die Zielwerte prognostiziert.

In den weitaus meisten Fällen in der Praxis kommen dabei überhaupt nicht die gefeierten neuronalen Netze zum Einsatz. Dort, wo die Daten einigermaßen strukturiert vorliegen, sind andere statistische Verfahren in der Regel überlegen. Der Brot-und-Butter-KI-Algorithmus für strukturierte Daten ist nach wie vor die lineare Regression, ein statistisches Verfahren, das Anfang des 19. Jahrhunderts entwickelt wurde. Besonders neu und originell ist das also nicht.

Neuronale Netze sind nur spezielle Varianten dieses Vorgehens. Sie eignen sich vor allem dann, wenn die vorliegenden Daten unstrukturiert sind.

Paradebeispiel dafür ist Bilderkennung: Die Pixel einer digitalen Abbildung lassen sich nur sehr schwer in strukturierte Daten überführen, weshalb klassische statistische Verfahren wenig nützen. Doch auch neuronale Netze sind älter, als man vermutlich denkt. Tatsächlich wurde die Theorie dahinter 1943 entwickelt, noch bevor es überhaupt Computer im modernen Sinn gab. Und die erste Maschine, die das Verfahren in die Praxis umsetzte, das Perceptron Mark I, wurde 1957 gebaut.

Nicht nur sind die neuronalen Netze keine besonders neue Entwicklung, sondern sie sind letztendlich auch nur statistische Modelle, in die sehr, sehr viele unstrukturierte Daten eingeflossen sind. Auch wenn sie von einem sehr vereinfachten Modell einer Nervenzelle inspiriert sind, haben sie mit der tatsächlichen Funktion eines Gehirns nichts zu tun.

Wahrscheinlichkeiten sind keine Wahrheiten

Warum betone ich immer so, dass es sich um statistische Verfahren handelt? Weil derartige statistische Modelle nie Wahrheiten abbilden, sondern nur Wahrscheinlichkeiten kennen. Wenn ich beispielsweise ein neuronales Netz mit Hunde- und Katzenbildern darauf trainiere, Hunde und Katzen zu unterscheiden, dann wird es das, bei einer ausreichend großen Zahl von Trainingsdaten, irgendwann ziemlich zuverlässig hinbekommen. Wenn ich diesem Modell dann aber das Bild eines Aktenordners gebe, wird es mir einfach sagen, das ist ein Hund oder eine Katze, je nachdem, was das statistische Modell für wahrscheinlicher hält. Das Ergebnis ist natürlich einigermassen willkürlich, weil ein Aktenordner weder eine Katze noch ein Hund ist. Doch das Modell selbst hat keinen Begriff von Katzen oder Hunden. Alles was man mit seiner Hilfe machen kann, ist Pixelstrukturen daraufhin absuchen, ob sie eher in Kategorie 1 oder Kategorie 2 fallen. Und man wird auf jeden Fall eine Antwort erhalten, egal, ob das überhaupt einen Sinn ergibt oder, im Fall eines Aktenordners, nicht.

Es gibt aber eine konkurrierende Schule künstlicher Intelligenz, die nicht auf statistischen Methoden beruht. Sie wird symbolische KI genannt. Diese basiert, im Gegensatz zu den bislang thematisierten statistischen Methoden, auf tatsächlichem Wissen, das in maschinell verarbeitbarer Form vorliegt. Die Idee dazu ist ebenfalls alt, noch älter als die statistischen Methoden. Sie wurde ursprünglich vom

Die Väter der symbolischen KI



Abbildung 1: René Descartes
(1596 - 1650)

Abbildung 2: Gottfried Wilhelm
Leibniz (1646 - 1716)

philosophischen Vater des neuzeitlichen Denkens, René Descartes, in der ersten Hälfte des 17. Jahrhundert aufgebracht. Ein halbes Jahrhundert später versuchte sich dann Gottfried Wilhelm Leibniz an einer *characteristica universalis*, zu Deutsch: Universalsprache. Die dem ganzen zu Grunde liegende Idee war es, alle Gegenstände und ihre Beziehungen untereinander mit eindeutigen mathematischen Symbolen zu fixieren. Damit sollte es dann möglich sein, über die reale Welt mit quasi mathematischer Präzision Aussagen zu treffen.

Sybolische KI

Was zunächst wie eine völlig abstruse Idee klingt, wird heute durchaus ernsthaft umgesetzt, wenn auch nicht mit den übertriebenen Hoffnungen, die ein Leibniz damit verband. Zur Implementierung einer solchen *characteristica universalis* verwendet man heute Graph-Datenbanken. Das hat nichts mit Graphiken zu tun, ein Graph ist vielmehr eine mathematische Struktur, die man sich am besten als ein Netz vorstellt, wo verschiedene Fäden, die sogenannten Kanten, sich an bestimmten Punkten, den sogenannten Knoten treffen. Knoten können dabei für irgendwelche Entitäten stehen: Gegenstände, Personen, Ideen etc. Die Kanten, die die Knoten verbinden, repräsentieren die Beziehungen zwischen den Knoten. Wenn also ein Knoten eine konkrete Person repräsentiert, sagen wir Max Müller, und ein anderer Knoten die abstrakte Vorstellung „Mensch“, dann werden sie durch die Kante „ist ein Exemplar von“ verbunden. Die mathematische Graphentheorie



Abbildung 3: Beispielsatz aus dem Flugblatt „Frauen schlagen Alarm“ (Abbildung 4)

beschäftigt sich im wesentlichen mit den Eigenschaften derartiger Netzwerke aus Knoten und Kanten. Mit solchen Graphen können ganz verschiedene Dinge abgebildet werden, zum Beispiel ein Nahverkehrsnetz oder ein Familienstammbaum. Im Fall einer *characteristica universalis* repräsentieren die Knoten alle möglichen konkreten oder abstrakten Gegenstände, die Kanten ihre Relation zueinander.

Wir nutzen derartige Graphen viel öfter, als wir vielleicht denken. Gerade als Archivar:innen und Dokumentar:innen greifen beispielsweise auf die Gemeinsame Normdatei (GND) zurück, die nichts anderes ist als eine derartige Graph-Datenbank. Und wahrscheinlich nutzen wir eine andere noch viel häufiger, nämlich WikiData. WikiData ist eine Datenbank, die ebenso wie die Wikipedia von der Wikimedia-Stiftung betrieben wird und die im Hintergrund der Wikipedia läuft. Und sobald ein Wikipedia-Artikel aufgerufen wird, werden Informationen aus WikiData geholt. Jede Seite der Wikipedia hat eine eindeutige WikiData-Id, die es dann erlaubt, weitergehende Informationen zu dem entsprechenden Gegenstand abzurufen, die dann meistens die *side bars* des Artikels füllen.

KI-Projekt digitaler Archivar

Soviel zur Theorie. Nun zur Praxis des KI-Projektes „Digitaler Archivar“. Dabei ging es um die Erschließung von gedruckten Materialien mit Hilfe von KI. Der eigentliche Witz des Digitalen Archivars sollte es sein, statistische mit symbolischer KI zu verbinden. Der geplante Workflow sah ungefähr folgendermaßen aus: Wir beginnen mit einem gescannten Dokument. Zunächst sollte mittels OCR, also optischer Texterkennung, der Text eines Dokuments ermittelt werden (das ist übrigens auch schon eine Form von KI; heutiger Texterkennungssoftware liegen trainierte Modelle zu Grunde). Im nächsten Schritt wollten wir dann mit Hilfe eines Verfahrens, das sich *entity recognition* nennt, im Text semantisch bestimmte Begriffe herauspräparieren: Zu-

nächst einmal Personen, Körperschaften und Orte, in einer späteren Iteration dann auch Sachterme. Und im dritten Schritt wollten wir dann von diesen statistisch ermittelten Termini übergehen zu symbolischer KI und die Termini eindeutig mit einem Wissensgraphen verbinden. Zunächst dachten wir dabei an die Gemeinsame Normdatei, die sich in der Praxis dann aber als zu schlecht erwies (zu viele Knoten, zu wenige Kanten), weshalb wir dann auf WikiData umgeschwenkt sind.

Warum das Ganze? Um den Schritt von der Wahrscheinlichkeit hin zu gesichertem Wissen zu gehen. Symbolische KI repräsentiert wirkliches Wissen. Archivarinnen und Archivare wissen, wie hilfreich eine Erschließung mit einem Normvokabular ist. Und ein Wissensgraph ist ein Normvokabular auf Steroiden. Damit erweitern sich die Möglichkeiten ungemein, in den Beständen nach bestimmten Dokumenten zu suchen.

Woran scheiterte der digitale Archivar? Der Hauptgrund, auf den ich hier gar nicht eingehen will, war die Finanzierung. Wir haben während der Pandemie in unserer Freizeit einen Prototypen gebaut, weil wir die Zeit dafür hatten. Um das zu einer gewissen Praxisreife zu entwickeln, hätten wir Geld gebraucht, das wir letztlich nicht auftreiben konnten.

Davon abgesehen will ich aber auf einen anderen, technischen Aspekt zu sprechen kommen: Die Qualität der mittels OCR extrahierten Texte. Wenn wir unseren Prototypen mit nativ digitalen Texten, also beispielsweise einem email-Bestand, gefüttert haben, dann hat er erfreulich gut funktioniert. Wenn wir aber erst einmal Texte per OCR extrahieren mussten, wurde es ganz schnell schwierig. OCR funktioniert bei guten, sauber gedruckten Vorlagen inzwischen praktisch fehlerfrei. Aber bei archivarischem Material, wo wir es beispielsweise mit unregelmässiger Schreibmaschinenschrift auf dünnem Durchschlagpapier oder – in unserem Fall – mit ausgebleichten Flugblättern auf stark vergilbtem Papier zu tun haben, sind die Ergebnisse immer noch sehr fehleranfällig.

Large language models

Und damit komme ich zum letzten Teil, wie uns hier nämlich *large language models*, also beispielsweise ChatGPT, weiterhelfen können. Wie gesagt, auch neuronale Netze sind letztendlich nur sehr komplexe statistische Modelle. Und *language models* sind, wie der Name schon sagt, statistische Modelle für natürliche Sprache. In ihnen ist festgehalten, was sprachlich in einem gewissen Kontext wahrscheinlich ist und was nicht. Und dieses Wissen können wir nutzen, um Fehler in der Texterkennung auszubügeln. Dazu habe ich ein Beispiel mitgebracht.

Den folgenden Ausführungen liegt ein historisch wichtiges Flugblatt der Kaiserstühler Bäuerinnen während der Proteste gegen das Atomkraftwerk Wyhl zu Grunde (Abbildung 4). Die Qualität ist durchaus zeittypisch. Nehmen wir einen Beispielabsatz heraus (Abbildung 3):

Die OCR, in diesem Fall Tesseract, macht daraus folgendes:

Was beim Erörterungstermin noch geleugnet wurde, ist durch das Beispiel Merckolsheim jetzt bewiesen: Das KKW zieht weitere lebensgefährliche Industrie nach sich. Um das Bleiwerk zu verhindern, haben die 21 badisch-Ä, elsässischen Bürgerinitiativen an 20. September zur Bauplatzbesetzung aufgerufen. Und fast 5000 sind bisher gekommen, um sich an der Besetzung zu beteiligen. Warum? = | | ee ee

Die Texterkennung ist eigentlich ziemlich gut. Natürlich kann sie die gelochte Stelle nicht korrekt lesen und am Anfang und am Ende von Zeilen treten einige sinnlose Artefakte in Form seltsamer Zeichen auf.

Wir können den Text aber verbessern, indem wir ihn an ChatGPT übergeben mit der Aufforderung, Tippfehler zu entfernen. Das Ergebnis sieht folgendermaßen aus:

Was beim Erörterungstermin noch geleugnet wurde, ist durch das Beispiel Merckolsheim jetzt bewiesen: Das KKW zieht weitere lebensgefährliche Industrie nach sich. Um das Bleiwerk zu verhindern, haben die 21 badisch-elsässischen Bürgerinitiativen am 20. September zur Bauplatzbesetzung aufgerufen. Und fast 5000 sind bisher gekommen, um sich an der Besetzung zu beteiligen. Warum?

Immer noch falsch ist der Ortsname „Merckolsheim“ statt „Marckolsheim“. Aber es gibt einiges an Verbesserung: Überflüssige Zeichen wurden entfernt, die fehlenden Zeichen durch die Lochung korrekt er-

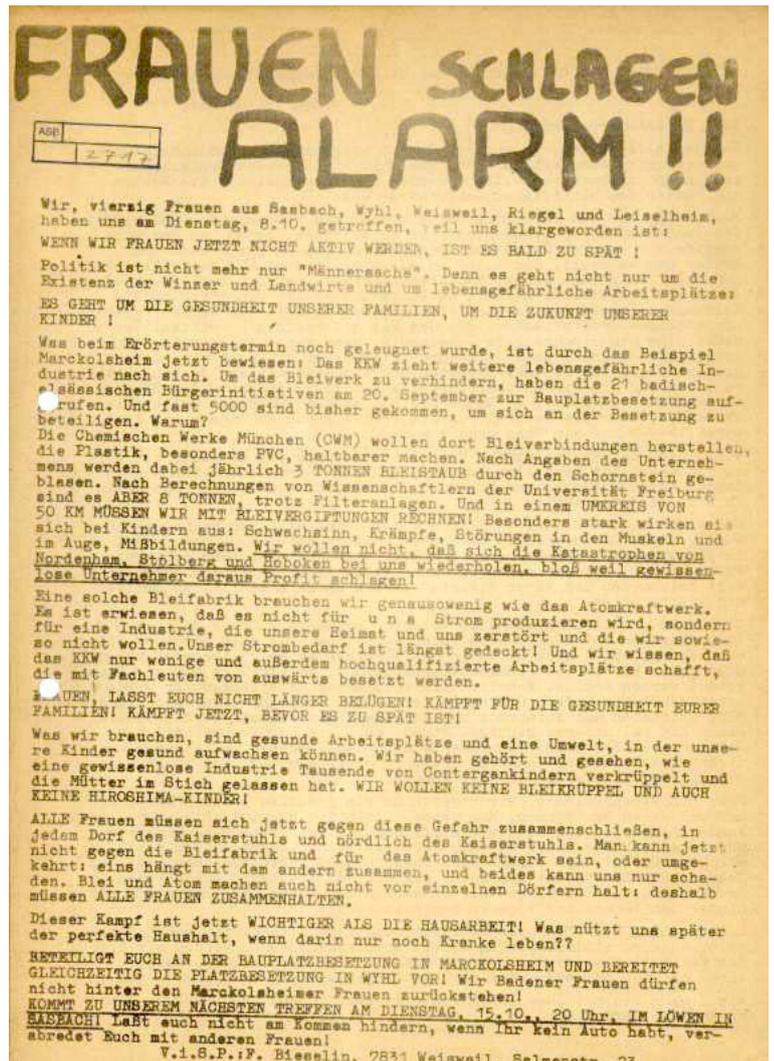


Abbildung 4: Flugblatt vom Oktober 1974, Archiv Soziale Bewegungen e.V., Bestand 12.1.5 Marckolsheim, Signatur 2717

setzt und die Trennungen bei Umbrüchen zusammengefügt. Damit ist ein deutlich saubererer Text entstanden.

Das wirkt sich auf die nächsten Bearbeitungsschritte aus. Führen wir – jetzt auf dem gesamten Text des Flugblattes – eine Entitätserkennung aus, verbessern sich die Resultate. Dabei kam die *natural language processing*-Software *spacy* zum Einsatz. Der Screen shot in Abbildung 5 zeigt die Differenz zwischen unkorrigierter und korrigierter Version.

Die Städte Hoboken und Nordenham werden nun korrekt erkannt, Duisburg wird fälschlich identifiziert, das sollte eigentlich die Gemeinde Stolberg sein, in der in den 60er Jahren schwere Bleivergiftungen festgestellt wurden. Das fehlerhaft erkannte „Merckolsheim“ findet sich unkorrigiert auch hier, außerdem „KReisweil“ statt „Weisweil“. Durch die Textkorrektur verschwinden aber zwei fehlerhafte



Abbildung 5: Ortserkennung mit spacy im Vergleich korrigierter/unkorrigierter Text

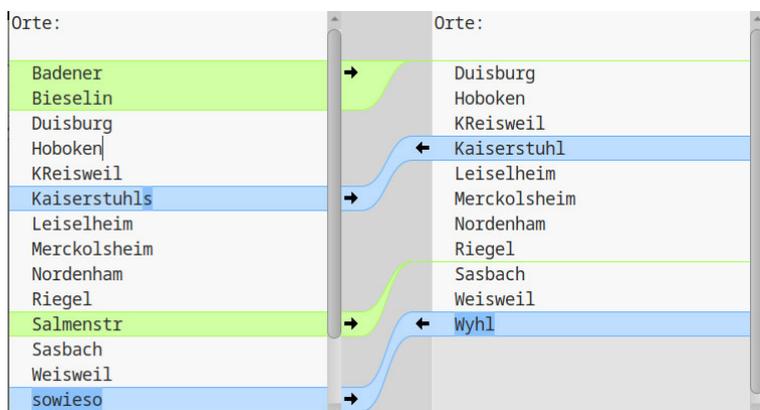


Abbildung 6: Ortserkennung im Vergleich spacy/ChatGPT korrigiert

Identifizierungen, dafür kommt eine andere („sowieso“) hinzu. Drei weitere fehlerhafte Identifizierungen („Badener“, „Bieselin“ und „Salmenstr“) bleiben.

Wir haben also, allein durch diesen zusätzlichen Verarbeitungsschritt, in dem wir den Ausgangstext durch ChatGPT haben verbessern lassen, eine signifikante Verbesserung der Resultate erhalten. Allerdings haben wir uns auch den Fehler mit Duisburg eingehandelt.

Wir können aber noch einen Schritt weiter gehen. Statt spacy zu benutzen, können wir ChatGPT direkt nach den Orten im Text fragen. Wenn wir diese Liste, die wir dann erhalten, mit der vorigen vergleichen, die bereits auf der verbesserten Version beruhte, erhalten wir das in Abbildung 6 dargestellte Ergebnis.

„Duisburg“ statt „Stolberg“ ist immer noch falsch, ebenso „KReisweil“ und „Merckolsheim“. Aber hier gibt es keine fehlerhaften Identifizierungen und die im Kontext sehr wichtige Ortschaft „Wyhl“, die *spacy* überhaupt nicht als Ort erkannt hatte, wird von ChatGPT korrekt als Ort identifiziert. Wir sind also noch weit von Perfektion entfernt, aber grundsätzlich sind wir unserem Ziel, benannte Entitäten zu finden, ein gutes Stück näher gekommen. Und das ist die Voraussetzung dafür, die bislang rein statistisch ermittelten Ergebnisse auf die symbolische Ebene zu heben und mit einer Wissensdatenbank zu verknüpfen.

Fazit

Was ergibt sich daraus als Fazit? Auch wenn *large language models* keineswegs perfekte Werkzeuge sind, können sie im *workflow* für eine automatisierte Erschließung von Textdokumenten ein wichtiger Baustein sein. Sie ermöglichen es, fehlerhafte Texterkennung aufgrund statistischer Wahrscheinlichkeiten unterm Strich deutlich zu verbessern. Dabei sollte man sich aber durchaus der potentiellen Fehlerhaftigkeit der statistischen Textverbesserung bewusst sein, wie das Beispiel „Stolberg“ – „Duisburg“ zeigt.

Und auch in Sachen kategorisierter Entitäten erweisen sich LLMs als den Standardverfahren überlegen. Zumindest in unserem Beispiel hatten wir signifikant weniger fehlerhafte Identifizierungen und mehr korrekte Identifizierungen als es das zum Vergleich herangezogene Modell von *spacy* geschafft hat. Auf dem Weg zur automatischen Erschließung digitalisierter Textdokumente lösen *large language models* zwar nicht alle Probleme, aber sie bringen uns einen guten Schritt voran. ■

„Zeitfunk“ – Sendungen des DDR-Rundfunks

Exploration automatisierter Inhaltserschließung von historischen Tondokumenten im Deutschen Rundfunkarchiv als Proof of Concept

Claudia Hunold und Alexander Wolff

Mit dem Projekt „Zeitfunk“ soll eine automatisiert unterstützte Inhaltserschließung für die Bestände des Deutschen Rundfunkarchivs entwickelt und implementiert werden.

Das Deutsche Rundfunkarchiv (DRA) ist ein rein historisches Archiv. In seinen Beständen befinden sich die Überlieferungen des Rundfunks der DDR und des Rundfunks vor 1945. Der Einsatz von KI und Automatisierungstools soll die Erschließung der historischen Aufnahmen erleichtern und beschleunigen, stellt die Beteiligten aber auch vor besondere Herausforderungen.

Im Jahr 2020 haben wir als DRA dem Verwaltungsrat und somit unseren Stiftern das „Leistungsversprechen Digitalisierung und Erschließung 2034“ gegeben. Hinsichtlich der Erschließung bedeutet das, dass wir bis 2034 für alle digitalisierten Bestände die jeweils angemessenen Metadaten zur Verfügung stellen, um unsere Bestände recherchierbar und zugänglich zu machen. Dieses Versprechen kann ohne einen hohen Grad an Automatisierung nicht erfüllt werden. Im Jahr 2022 wurde daher das Team „Automatisierung in Retrieval & Dokumentation“ gegründet, in dem die Automatisierungskompetenzen im DRA zusammengefasst werden. In diesem Team kommen sowohl hauptamtliche Automatisierungsspezialisten als auch Mitarbeiter an der Schnittstelle zwischen Dokumentation und Automatisierung zusammen.

Das DRA hat bereits zuvor, seit Gründung dieses Teams aber verstärkt, in verschiedenen Projekten Erfahrungen mit Automatisierung sammeln können. Beispielsweise wurden projektspezifische Python-Skripte geschrieben. Zwei Erkenntnisse aus diesen vergangenen Projekten waren dabei zentral: Erstens gibt es wiederkehrende Projektaufgaben, die man standardisieren könnte. Zweitens besteht oft ein gewisser Nachbearbeitungsbedarf, insbesondere bei automatisiert, zum Beispiel mit Hilfe von KI, erzeugten Metadaten. Wie effizient Automatisierung ins-

besondere durch automatische Erzeugung von Metadaten tatsächlich ist und wieviel Arbeitsaufwand gegenüber der händischen Erzeugung gespart werden kann, hängt nicht nur von der Qualität der erzeugten Daten ab, sondern auch von ihrer Darstellung und der Möglichkeit, diese einfach überprüfen

und ggf. korrigieren zu können. In unseren vergangenen Projekten erfolgte dies überwiegend über eine Excel-Datei, in die die vorgeschlagenen Daten geschrieben, dort korrigiert und schließlich per Python-Skript von der Excel-Liste in die Datenbank übertragen wurden. In der Praxis erwies sich dies aber als zu kompliziert und unhandlich, so dass hier viel an potenziellem Effizienzgewinn liegen blieb. Dies wollten wir für kommende Projekte beheben.

Der Zeitfunk

Vor diesem Hintergrund wurde Mitte 2023 das Projekt „Zeitfunk“ gestartet. Benannt ist das Projekt nach dem Bestand, der hier erschlossen werden soll. Zeitfunk war eine Redaktion des Deutschlandsenders des Rundfunks der DDR, die unter anderem zeitgenössische Reportagen und Features produzierte. Die Aufnahmen, beispielsweise O-Töne, Interviews und Berichte, stammen aus den 1950er und 1960er Jahren und sollten als Produktionsmaterial vorgehalten werden. Über den Inhalt der insgesamt 232 analogen Bänder liegen nur sehr spärliche Informationen vor. Die handschriftlich vermerkten Zählungen auf den Bandkartons geben keinen Hinweis auf eine Chronologie. Die Beschriftung der Bandkartons ist nicht informativ und teilweise schlecht lesbar. (Abbildung 1)



Claudia Hunold und Alexander Wolff
Deutsches Rundfunkarchiv,
Stiftung von ARD und Deutschlandradio
Potsdam
alexander.wolff@dra.de
Claudia.Hunold@dra.de



Abb. 1: Zeitfunk-Bänder im Magazin

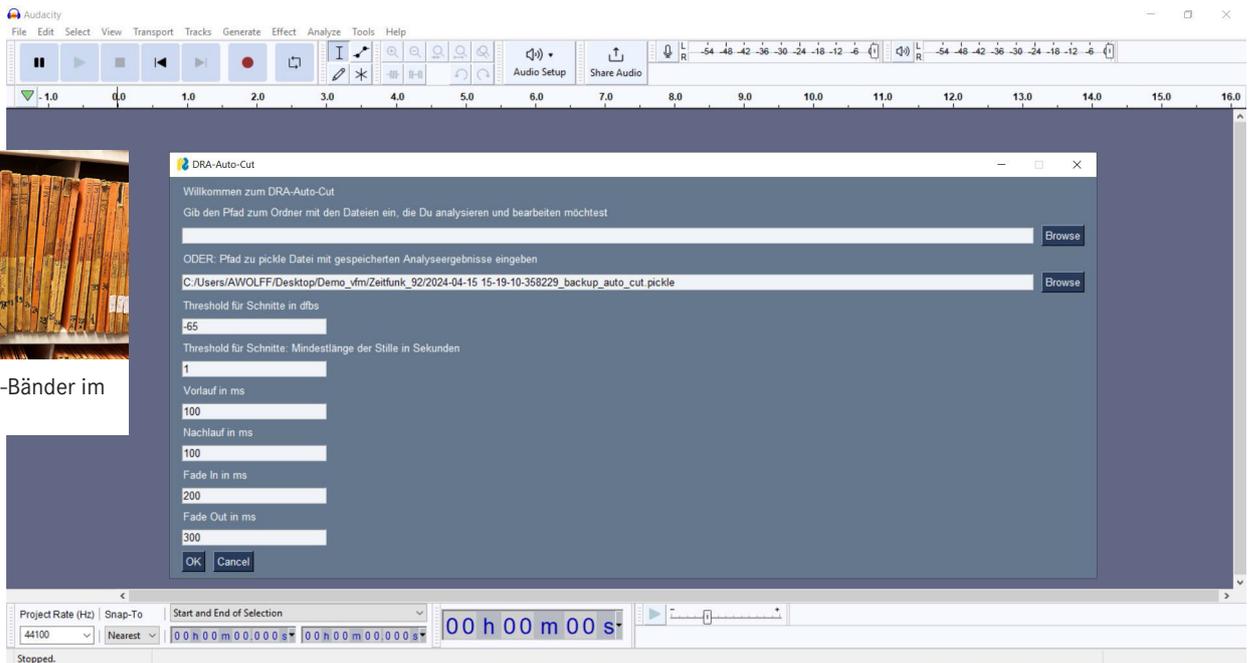


Abb. 2: Das DRA-Auto-Cut, im Hintergrund Audacity

Umso interessanter sind diese bereits vollständig digitalisierten Bänder, über deren Inhalt wir zu Beginn des Projekts sehr wenig wussten, für eine automatisierte bzw. durch automatische Verfahren unterstützte Erschließung. Der Auftrag an das Team Automatisierung lautete hier schlicht, den Bestand Zeitfunk soweit automatisiert wie möglich zu erschließen.

Nach erster Exploration stellte sich schnell heraus, dass eine Komplettautomatisierung wie schon bei vorigen Projekten nicht möglich ist, und zwar bei keinem der benötigten Arbeitsschritte. Nichtsdesto trotz zeigten die explorierten automatischen Verfahren bei allen Arbeitsschritten vielversprechende Ergebnisse; sie waren nur eben nicht so gut, dass sie einfach ohne weitere Bearbeitung übernommen werden konnten. An dieser Stelle hätten wir in der Vergangenheit ein Projekt vielleicht sogar abgebrochen, da Automatisierung keine ausreichend guten Metadaten versprach. Wir haben uns aber unter dem Einfluss unserer Erfahrungen aus anderen Projekten zu einem anderen Ansatz entschieden: Von der komplett automatisierten zur automatisiert unterstützten Erschließung.

DRA-Auto-Cut

Bei vielen der Zeitfunk-Bänder handelt es sich um so genannte „Sammelbänder“, das heißt mehrere „Positionen“, also inhaltlich abgeschlossene Einheiten, können sich auf einem Band befinden. Diese Bänder wur-

den als Ganzes digitalisiert, was bedeutet, dass die Audiofiles vor einer Erschließung in ihre einzelnen dokumentarischen Einheiten zerlegt werden müssen.

Auch dies wollten wir weitgehend automatisieren. Wir suchten also messbare Audiomerkmale, um die dokumentarischen Einheiten eines Audiofiles erkennen und dieses somit automatisch schneiden zu können. Hier kam uns das zwischen die einzelnen Positionen auf dem analogen Ausgangsmaterial eingeschnittene „Weißband“ zugute. In unseren Analysen zeigte sich schnell, dass das Grundrauschen der historischen Bänder dort, wo Weißband eingeschnitten wurde, deutlich leiser ist als beim Rest des Bandes.

In einem ersten Experiment wurden automatisiert Schnitte an den Stellen gesetzt, an denen mindestens über eine Sekunde der Pegel des Audios unter -65 dfbs fiel. Mit Hilfe der Python-Library `pydub` war hier schnell ein entsprechendes Python-Skript programmiert. Damit konnten wir bei ca. 70 bis 80 Prozent der Fälle korrekte Schnitte setzen. Auch wenn aus der Sicht eines Automatisierungsspezialisten diese Quote durchaus zufriedenstellend wäre, genügt sie nicht den Ansprüchen der Dokumentation. Wir beschlossen also, einen anderen Ansatz zu wählen und uns zu fragen „Wie bringen wir Algorithmus und Dokumentar*in so zusammen, dass sie effektiv zusammenarbeiten können?“

So entstand der DRA-Auto-Cut, dessen Aufbau auch eine Blaupause für Tools liefert, die die Er-

schließung durch Automatisierung unterstützen. Ziel war es, die durch den Algorithmus erzeugten Metadaten, hier die potenziellen Schnittstellen für das jeweilige Gesamt-Audiofile, möglichst einfach überprüf- und korrigierbar zu machen. Dazu war es wichtig, die Hüllkurve anzuzeigen und Schnittmarken automatisch zu setzen, aber auch verschieben zu können.

Wir bedienen uns hier des beliebten Open-Source Audio-Editors Audacity, mit dem wir ein entsprechendes, in Python geschriebenes Tool umsetzen konnten, indem wir auf die Audacity-Standard-Schnittstelle zugreifen.

Das Tool leitet die Dokumentarin mit mehreren Fenstern durch den Prozess. Beim Start hat sie zunächst die Möglichkeit, einen Ordner zu wählen, in dem sich die Ausgangsfiles, die analysiert und automatisiert unterteilt werden sollen, befinden. Es werden während des Prozesses auch Backup-Dateien geschrieben, die hier ebenfalls eingesehen werden können, um nicht von Null, sondern von einem bestimmten Arbeitsstand aus zu starten. Zudem lassen sich die Grenzwerte hier variieren, die für einen automatisch vorgeschlagenen Schnitt angesetzt werden. Parameter für Vor- und Nachläufe bei den Schnitten sowie für Fades können ebenfalls gesetzt werden, damit das Tool einen audioteknisch möglichst sauberen Schnitt setzt. (Abbildung 2)

In der Folge werden die zum Schneiden vorgesehenen Files analysiert und die Schnittmarken definiert. Da der Prozess etwas dauert, hält ein Statusfenster die Bearbeiterin über den aktuellen Stand auf dem Laufenden.

Nach Abschluss des Analyseprozesses werden die jeweiligen Audiofiles nacheinander in Audacity geladen, wo der Dokumentarin alle Funktionen von Audacity zur eigenen Analyse des Files zur Verfügung stehen. In einer zweiten Spur werden die Schnittmarken angezeigt und können einfach per Drag & Drop versetzt werden. (Abbildung 3)

Ist die Kontrolle und Korrektur der vorgeschlagenen Schnitte nun abgeschlossen, wird dies in der Python UI mit „Ok“ bestätigt. Im Hintergrund werden die Schnittmarken aus Audacity nun wieder ausgelesen und die Dateien per Python-Skript genau an diesen Stellen geschnitten und im Zielordner abgelegt. Dies passiert innerhalb weniger Sekunden. Anschließend wird die nächste Datei mit Schnittmarken in Audacity angezeigt. (Abbildung 4)

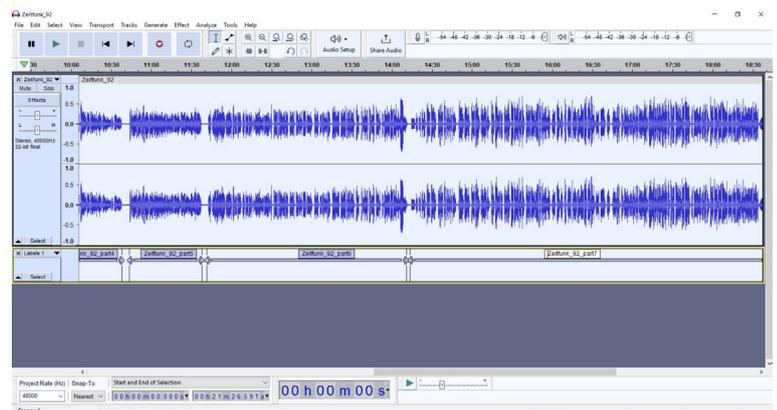


Abb. 3: Das zu schneidende File wurde in Audacity geladen. In der unteren Spur sind die vorgeschlagenen Schnittmarken sichtbar.

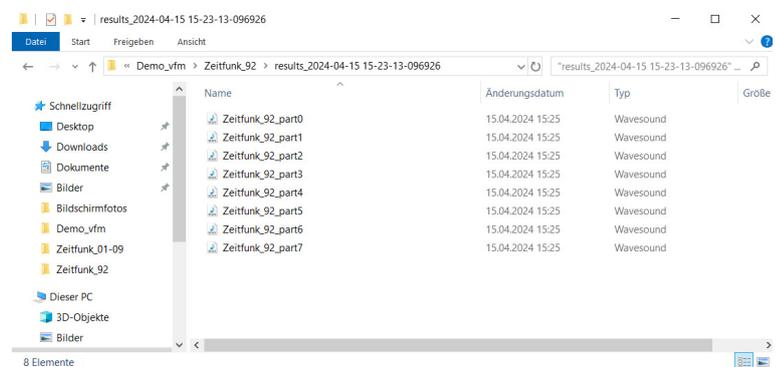


Abb. 4: Die durch den Schnitt entstandenen Files wurden im Zielordner abgelegt

Der DRA-Auto-Cut ist bereits in einer ersten Testversion verfügbar und wird zeitnah zum automatisch unterstützten Schnitt der Zeitfunk-Audiofiles verwendet werden. Das Tool lässt sich zukünftig auch auf andere Bestände mit ähnlichen Herausforderungen anwenden und wird hinsichtlich Funktionalität und Usability weiter optimiert werden.

DRA-Auto-Dok

Kern des Zeitfunk-Projekts ist ein zweites Tool namens DRA-Auto-Dok. In dieses Tool fließen die vielfältigen Erfahrungen des DRA hinsichtlich automatisierter Metadatenbearbeitung und -generierung hinein, um den/die Dokumentar*in zukünftig möglichst effizient bei der täglichen Arbeit zu unterstützen.

So hat das DRA zum einen vielfältige Erfahrungen mit der Schnittstelle der ARD-Hörfunkdatenbank (HFDB) gemacht und hier schon viele Daten automatisiert erzeugt und bearbeitet. Zum anderen haben wir in der Vergangenheit bereits mit Musikererkennung, Transkriptionstools und Metadatenextraktion per Large Language Model (LLM) gearbeitet. Beim

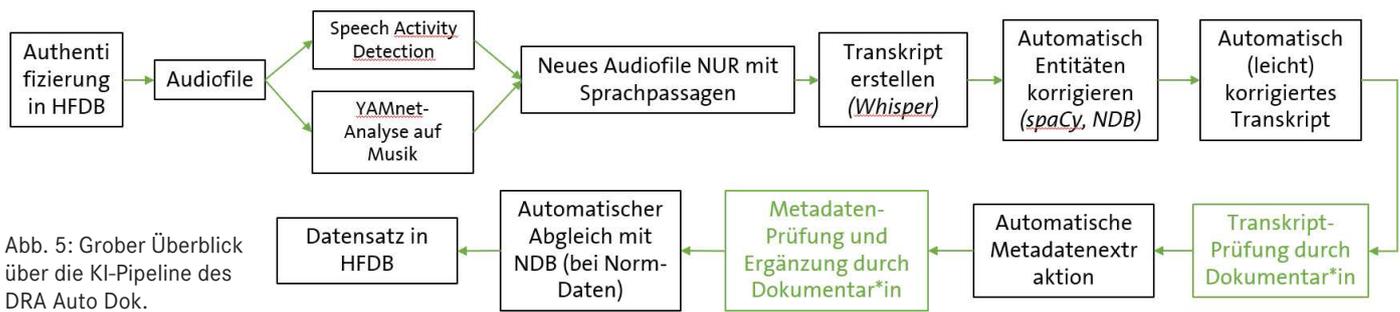


Abb. 5: Grober Überblick über die KI-Pipeline des DRA Auto Dok.



Abb. 6: HFDB-Authentifizierungsfenster mit Möglichkeit, sich in HFDB-Produktiv- oder Konsolidierungsinstanz anzumelden.

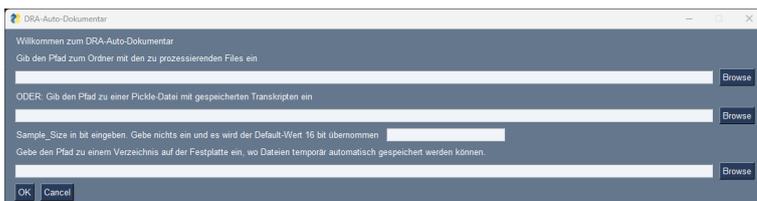


Abb. 7: Fenster, in dem u.a. die zu analysierenden Dateien definiert werden.

DRA-Auto-Dok kommt es nunmehr darauf an, all diese Tools in eine Pipeline zu gießen und die vorgeschlagenen Metadaten bestmöglich überprüf- und korrigierbar zu machen. Wir legen hier in der Entwicklung zunächst Wert auf die Usability und die Reihenfolge und das Zusammenspiel der einzelnen Tools (sprich die „Pipeline“), um, wenn das Tool so weit fertig ist, die Schritte der Metadatengenerierung bzw. -extraktion zu verfeinern. (Abbildung 5)

Die Pipeline

Beim Start des DRA-Auto-Doks wird zunächst ein Authentifizierungsfenster angezeigt, mit dem sich der/die User*in in der HFDB authentifizieren kann. Genauer gesagt authentifiziert sich das Tool hier mit den Userdaten der HFDB an der HFDB-Schnittstelle. (Abbildung 6)

Nach erfolgreicher Authentifizierung wird ein weiteres Fenster angezeigt, in dem der/die Dokumentar*in die Möglichkeit hat, entweder einen Ordnerpfad mit zu analysierenden Files anzugeben oder, wie auch beim Auto-Cut-Tool eine pickle¹-Datei mit

einem gespeicherten Arbeitsstand auszuwählen. Es muss zudem ein temporärer Ordner für die Zwischenspeicherung von Arbeitsdateien, die während des Prozesses temporär entstehen, angegeben werden. (Abbildung 7)

Sind die zu analysierenden Audiofiles definiert, werden diese zunächst mittels einer Speech Activity Detection analysiert. So wird festgestellt, an welchen Stellen des Audios tatsächlich Sprache zu hören ist (dazu gehört auch Gesang). Wir nutzen hier „silero VAD“², wobei es auch verschiedene andere Voice Activity Detections („VAD“) für Python gibt. Da hier Gesang und somit auch Musik mit Gesang nicht herausgefiltert wird, nutzen wir zudem YAMNet³ (von uns bereits vorgestellt auf der vfm-Tagung 2022), um alle Stellen mit Musik zu finden. Wir schneiden nun vollautomatisiert ein File, das wirklich ausschließlich die reinen Sprachpassagen enthält. Hintergrund ist der nächste Schritt der Pipeline, die Transkription mit „Whisper“⁴, einem Tool der Firma OpenAI, hierzulande vor allem bekannt durch Chat-GPT. Whisper ist open-source und derzeit das wohl beste Tool zur Transkriptionserstellung, hat jedoch Probleme, wenn es nicht mit reiner Sprache, sondern mit Geräuschen oder Musik gefüttert wird. Dies führt zu Halluzinationen⁵ in den Transkripten. Diese konnten wir in ersten Tests mit dem Pre-Processing mittels VAD und YAMNet weitgehend verhindern.

Da die meisten Algorithmen „von der Stange“ vor allem auf zeitgenössische Daten trainiert sind, wie auch Chat-GPT und Whisper, werden vor allem Namen der DDR-Geschichte oft falsch transkribiert, da diese im Trainingsmaterial, das vor allem auf dem zeitgenössischen Internet beruht, nicht oder nur wenig präsent sind. Dieses Problem dürfte bei

² <https://github.com/snakers4/silero-vad>

³ <https://www.tensorflow.org/hub/tutorials/yamnet>

⁴ <https://openai.com/index/whisper>

⁵ Halluzination im KI-Kontext: KI gibt Content aus, der mit der eigentlichen Abfrage nichts zu tun hat.

¹ Pickle = Python-spezifisches Fileformat zur Datensicherung – <https://docs.python.org/3/library/pickle.html>

der Erschließung zeitgenössischer Audioquellen weniger stark ausgeprägt sein.

Bevor wir für das DRA Metadaten generieren, die auf den Whisper-Transkripten basieren, schieben wir daher noch einen Zwischenschritt ein, indem wir mit dem Tool spaCy⁶ Personennamen aus den Transkripten extrahieren. Es wird dann automatisiert geprüft, ob ein plausibler Eintrag in der ARD Normdatenbank vorliegt, die zum Beispiel auch DDR-Persönlichkeiten enthält. Findet sich kein gleichlautender Eintrag, wird mittels einer Levenshtein⁷-Distanz ermittelt, ob ein ähnlich geschriebener Name in der Normdatenbank vorliegt. Im Transkript wird anschließend der von Whisper transkribierte Name durch den in der Normdatenbank gefundenen, ähnlichen Namen ersetzt. Dieses Verfahren hat seine Tücken und muss im weiteren Testprozess noch weiter evaluiert werden. So wurden hier auch richtig geschriebene Namen durchaus verfälscht, da es keine gleich lautenden, aber zufällig ähnliche, hier aber nicht relevante Namen in der Normdatenbank gibt. Es gibt aber Hinweise darauf, dass wir die Quote der korrekt geschriebenen Personennamen im Transkript im Durchschnitt etwas verbessern können. In Zukunft gibt es zur Optimierung der Personentranskription aber auch andere Möglichkeiten, die zu evaluieren sind, zum Beispiel das Nachtrainieren des Whisper-Modells mit DDR-Material, was jedoch aufwendig ist.

Es ist anzumerken, dass wir derzeit mit dem Whisper-Modell „large V2“ arbeiten und bereits eine neuere Version „large V3“ verfügbar ist, die wir zeitnah testen werden. Somit ist es gut möglich, dass einzelne oben beschriebene Probleme mit der neuen Version obsolet sind oder weniger stark auftreten. Unsere Pipeline ist so modular aufgebaut, dass sich einzelne verbaute Tools problemlos austauschen oder entfernen lassen.

Die eigentliche Metadatenextraktion erfolgt per Large Language Modell. Man kann sich dies so vorstellen, als würde man Chat-GPT (hinter dem auch ein LLM steht) Fragen zu einem vorgegebenen Text stellen, in diesem Fall zu unserem Transkript, das wir mit Whisper angefertigt haben. Wenn man dem LLM einen Quelltext mitgibt und es ausschließlich zu diesem Text befragt, so bekommt man auch recht zuverlässige Antworten, die dem Inhalt des Textes entsprechen. Zu Halluzinationen oder einfach gesagt falschen Antworten kommt es bei LLM-Abfragen

```
Die Person spricht selbst
Die Person spricht selbst
Die Person spricht selbst
Die Person spricht selbst
Die Person Strauss spricht nicht selbst in dem Audio.
Die Person spricht selbst
Die Person spricht selbst
```

Abb. 8: Nicht immer hält sich das LLM komplett an die Formatierungsvorgaben, dies lässt sich aber meist leicht abfangen, indem typische Ausgabevarianten entsprechend berücksichtigt werden.

normalerweise eben immer dann, wenn man der Maschine eigenes Weltwissen unterstellt und sie wie Wikipedia benutzt. Dies genauer zu erklären, würde an dieser Stelle zu weit führen. Um es kurz zu halten: Metadatenextraktion und das Schreiben von Zusammenfassungen aus vorgegebenen Texten funktioniert per LLM sehr gut.

Wir stellen dem LLM nun also gezielt Fragen zu dem gegebenen Transkriptionstext, z.B. „Welche Personen spielen in diesem Transkript eines Radiobeitrags eine wichtige Rolle? Nenne in Stichpunkten (*Vorname, Nachname*): *Transkript*“

Dabei hat sich in vergangenen Projekten gezeigt, dass man durch Ausprobieren mit verschiedenen Prompts (Abfragen an das LLM) die Ergebnisse sehr gut immer weiter optimieren kann. Der erste Prompt ist hier selten der Beste. Zudem ist es ratsam, dem LLM einfache Fragen zu stellen. Komplexere Sachverhalte sollte man in mehrere einzelne Abfragen aufteilen. So fragen wir stets erst „Kommen in diesem Transkript relevante Personen vor? Antworte mit Ja oder Nein: *Transkript*“ Und erst wenn das LLM hier bejaht, stellen wir die Frage, welche relevanten Personen dies nun sind. Man könnte beide Fragen auch in eine Abfrage zusammenfassen, erhält hier aber schlechtere Ergebnisse. Die Frage nach dem „Ja oder Nein“ ist zudem wichtig, denn LLMs wollen immer eine Antwort geben, auch wenn es keine gibt. Fragten wir das LLM nur nach relevanten Personen, ohne vorher zu fragen, ob es überhaupt welche gibt, wird es im Zweifel welche erfinden und wir sind wieder beim Halluzinations-Problem. Dies taucht bei der Metadaten-Extraktion per LLM aber tatsächlich nur dann auf, wenn ein Metadatum im Text nicht vorhanden ist, und lässt sich wie gezeigt durch die Vorab-Abfrage nach der Existenz des Metadatums leicht umgehen.

⁶ <https://spacy.io>

⁷ <https://de.wikipedia.org/wiki/Levenshtein-Distanz>

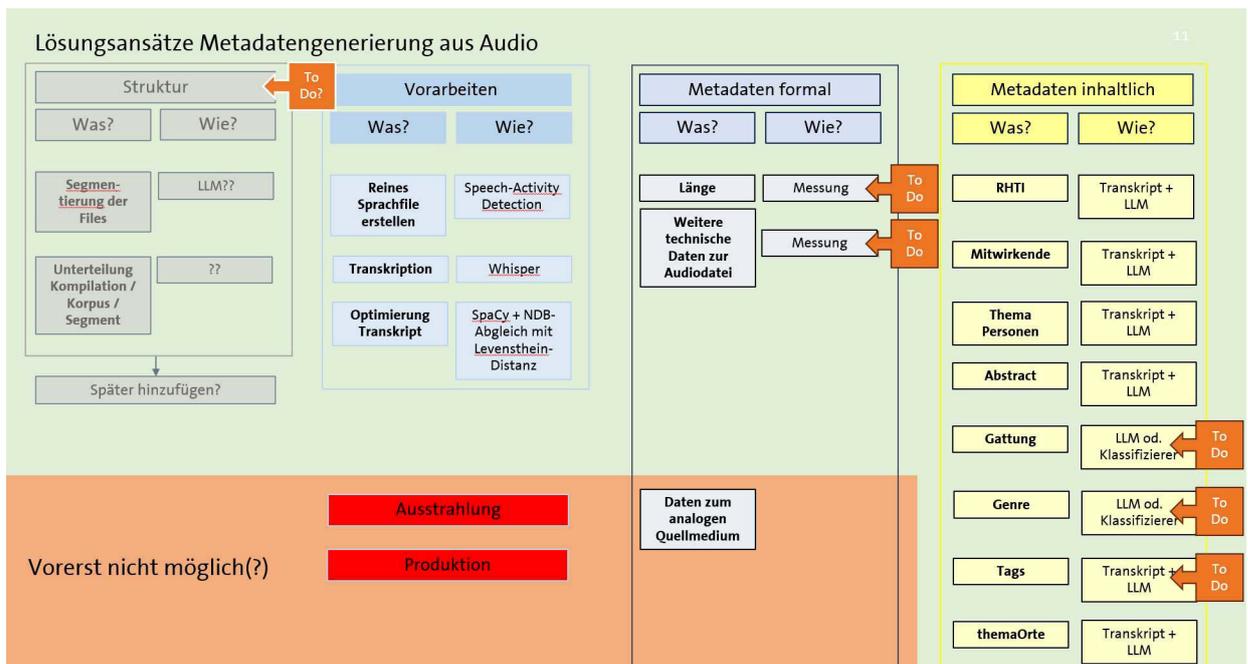


Abb. 9: Für fast alle zu einem Tondokument gehörenden Metadaten haben wir Lösungsansätze in der Hinterhand, um Daten vorzuschlagen zu können.

Die Formatierung der LLM-Antworten kann trotz Vorgaben im Prompt stets leicht abweichen, was sich aber durch ein paar Skriptzeilen, die typische Versionen der Rückgabe abfangen, leicht berücksichtigen lässt. (Abbildung 8)

Eine mögliche Alternative wäre ein Tool wie LangChain⁸, das die Benutzung von LLMs mit strukturellen Vorgaben bei den Antworten verbindet. In anderen Projekten, in denen wir dies in der Vergangenheit getestet haben, funktioniert die Strukturierung der Antworten zwar sehr gut, aber es gehen dabei teilweise Informationen verloren, die bei „freier“ Beantwortung durch das LLM enthalten sind. Beispielsweise werden bei einer Abfrage nach vorkommenden Personen nicht alle vorkommenden Personen genannt, sondern nur ein Teil. Aber auch hier werden wir in Zukunft weitere Tests fahren und der strukturelle, objektbasierte Aufbau unseres Tools ermöglicht es uns hier schnell umzustiegen, sollte sich dies lohnen.

Was das Large Language Modell an sich angeht, so setzen wir zurzeit auf Google Gemini in der Version „gemini-1.0-pro-002“⁹. Dies hat derzeit strukturelle Gründe, da wir in letzter Zeit sehr gute Erfahrungen mit Google Cloud Services gemacht haben, über die sich auch die Gemini Modelle einfach an-

sprechen lassen und wir Chat-GPT noch nicht auf datensicherem Wege verfügbar haben. Mittlerweile gibt es auch hier eine neue Version 3. Gemini-Modelle sind multimodal, das heißt sie lassen sich nicht nur mit Text, sondern auch direkt mit Audios, Bildern etc. füttern, zu denen man wiederum direkt Fragen stellen kann. Dies wird uns bei der Extraktion bestimmter Metadaten wiederum neue Möglichkeiten eröffnen. Denkbar ist, in Zukunft auch GPT-Modelle einzusetzen bzw. Gemini gegen GPT auszutauschen, sofern dies die Ergebnisse noch weiter verbessern sollte.

Derzeit werden noch nicht alle theoretisch möglichen Metadaten, die für einen HFDB-Eintrag sinnvoll sind, extrahiert, wir werden aber in den nächsten Wochen die weiteren Daten ergänzen. Bei wichtigen Metadaten wie Titel, Abstract, Sprechenden, Orten und genannten Personen erzielen wir bereits ohne allzu viel Prompt- und LLM-Optimierung recht gute Ergebnisse. (Abbildung 9 und 10)

Wie gut am Ende Mensch und Maschine zusammenarbeiten können, hängt beim DRA-Auto-Dok maßgeblich vom User Interface ab. In diesem Bereich stimmen sich Team Automatisierung und Team Metadatengenerierung im DRA eng ab. Es gilt hier wie auch beim DRA-Auto-Cut die Metadaten möglichst so anzuzeigen, dass diese leicht überprüfbar und korrigierbar sind. Zu den meisten Metadaten gibt es daher eine zusätzliche Kontextabfrage an das LLM, z.B. „In welchem Kontext kommt Person XY in diesem Text

⁸ https://python.langchain.com/docs/get_started/introduction

⁹ <https://gemini.google.com>

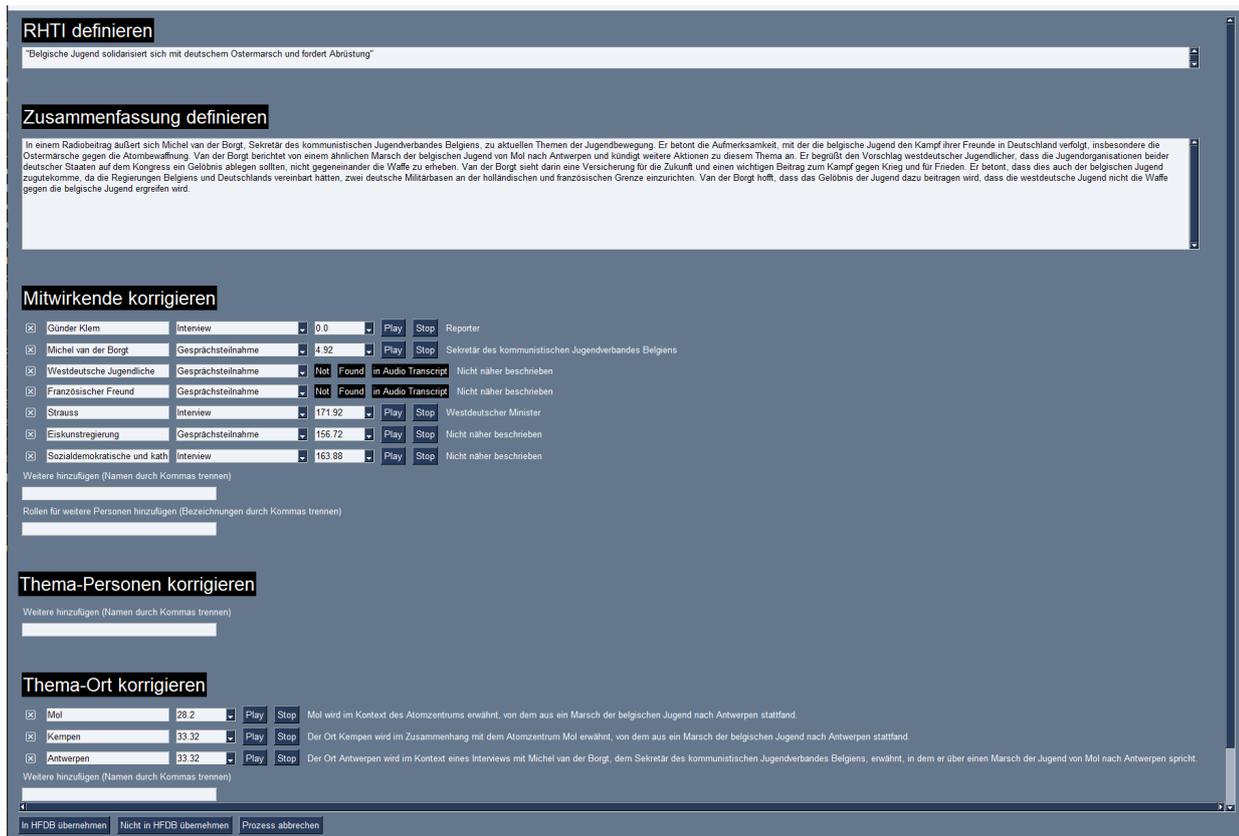


Abb. 10: DRA-Auto-Dok Hauptfenster, in dem die vorgeschlagenen Metadaten angezeigt werden. Wie zu sehen ist, sind nicht alle davon schon wirklich hilfreich, das Optimierungspotential ist aber auch noch lange nicht ausgeschöpft.

vor?: *Transkript*⁴. Der hierbei zurückgegebene Text wird zusätzlich zum eigentlichen Metadatum angezeigt. Auf Wunsch der Dokumentar*innen wurde zudem ein Audioplayer eingebaut, der jeweils die Stellen des Audios direkt abspielt, an denen ein Metadatum im Transkript vorkommt. Das macht die vorgeschlagenen Metadaten besonders einfach überprüfbar.

Sämtliche Metadaten können durch das User Interface ergänzt und abgeändert oder auch entfernt werden, wenn sich herausstellt, dass sie doch nicht benötigt werden. Neben der einfachen Überprüfbarkeit ist die Handhabung und Editierbarkeit der Metadaten ein wichtiges Element für ein gelungenes Tool, auf das wir bei der Entwicklung viel Wert legen.

Mit dem Tool PySimpleGUI¹⁰ gelingt es uns bereits sehr gut, flexible und nutzerfreundliche GUIs zu bauen, die sich auch gut erweitern und editieren lassen.

Am Ende steht die Normierung der Daten mit Einträgen der ARD-Normdatenbank. Jeder Eintrag, egal ob in der Normdatenbank vorhandene oder neue, dort noch nicht erfasste Person, muss am Ende mit einem Normdatenbank-Eintrag verknüpft sein. Derzeit haben wir hier einen automatischen Abgleich eingebaut, der aber nicht alle Namen umfasst. Auch hier wollen

wir noch weiter optimieren, um auch den Normdatenbankabgleich bestmöglich und automatisiert unterstützt, anstatt vollautomatisiert handhabbar zu machen, was in der nächsten Zeit ein großes Arbeitspaket sein wird. Normierte, mit der Normdatenbank verknüpfte Metadaten sind auch bei anderen Metadatenfeldern in der Hörfunkdatenbank Pflicht und ein großer Wunsch der Dokumentar*innen für das Tool.

Ausblick

Wir glauben, dass uns mit dem Ansatz, Dokumentationsprozesse im DRA nicht komplett zu automatisieren, sondern Tools zu bauen, die die Dokumentar*innen bestmöglich in ihrer Arbeit unterstützen, ein wichtiger strategischer Meilenstein gelungen ist. Wir merken aber auch, dass ein Tool wie der DRA-Auto-Dok sehr aufwendig zu bauen und für uns eine echte Energieleistung ist. Die ersten Ergebnisse und Reaktionen auf das Tool bestätigen uns aber in unserem Weg. Wir planen, den DRA-Auto-Dok im dritten oder vierten Quartal 2024 so weit fertig gestellt zu haben, dass wir damit den Bestand „Zeitfunk“ als Proof of Concept erschließen können. Die Erfahrungen aus diesem POC werden dann wiederum in die Weiterentwicklung des Tools fließen. ■

¹⁰ <https://www.pysimplegui.com>

Automatisierte Playlisterstellung in der Musikdokumentation

KI-basierte Analyse und Verschlagwortung von Songtexten

Sabrina Ladenburger



Sabrina Ladenburger
Bayerischer Rundfunk
Sabrina.Ladenburger@br.de

Im ersten Abschnitt des folgenden Artikels wird das PARAS-Projekt vorgestellt, ein Dienstleistungsprojekt der Dokumentation Musik, das exemplarisch dafür steht, wie durch KI- und Automatisierungstools neue Produkte aus den Archiven angeboten werden können. In der Durchführung dieses Projektes folgte eine ausführliche Auseinandersetzung

damit, wie Large Language Models – als Beispiel für KI- und Automatisierungstools – im Arbeitsalltag der Musikrecherche zur Songkategorisierung eingesetzt werden können. Hierzu liefert der zweite Teil des Artikels Beispiele, Handlungsempfehlungen und Lessons learned.

PARAS – Projektvorstellung

Projekthalt

Bei PARAS handelt es sich um das „Playlist Aufbereitungs & Recherche Archiv Service“-Projekt der Dokumentation Musik im BR. Dem Dienstleistungsgedanken folgend sollen in diesem Projekt die (Musik-)Redaktionen mit Musiksammlungen zur Programmvorbereitung bespielt werden.

Als geeignete Methode der Materialportionierung stellten sich Playlists heraus, mittels derer unseren Redaktionen Musik zu ausgewählten relevanten Themen zur Verfügung gestellt werden kann. Die Playlists sind jedoch nicht als Aneinanderreihung von Musikstücken zu verstehen, also keine reinen Listen, sondern sollen vielmehr mit Zusatzdaten und -material angereichert werden, um unsere redaktionellen KollegInnen schnell mit den richtigen Infos zu versorgen. Gleichzeitig liefern die Musiksammlungen einen Überblick zu thematisch relevantem Archivmaterial und fungieren als erster Anhaltspunkt für weiterführende Recherchen.

Die Präsentation dieses Angebots soll demnächst auf einer eigenen Website/Landingpage er-

folgen, da hier, im Vergleich zu unseren Archivoberflächen, die Informationen optisch und redaktionell leichter aufbereitet werden können. Zusätzlich wird es Schnittstellen zu unseren gängigen Datenbanken geben, sodass Titel vorgehört oder direkt ins Produktionssystem bestellt werden können. Dafür ist sichergestellt, dass alle Musikstücke in der korrekten Version, voll erfasst und digitalisiert vorliegen.

Zielsetzung

PARAS zielt darauf ab, sowohl auf Kunden- als auch auf Archivseite mehrere Probleme auf unterschiedlichen Arbeitsebenen zu lösen. Während einige Recherchebedarfe extrem wellenspezifisch sind, werden andere Probleme von allen Wellen geteilt. An erster Stelle ist hier auf die Datenmenge hinzuweisen, die auf unsere immer größer werdenden Archivbestände zurückzuführen ist. Eine schnelle Recherche wird durch diese Datenflut erschwert. Die Archive können hier als Leuchtturm im Datenmeer ein „Finden“ statt „Suchen“ ermöglichen. Verstärkt werden die genannten Rechercheprobleme dadurch, dass es stellenweise zu wenige, beziehungsweise nicht die richtigen Daten gibt. Wurden treffende Deskriptoren beispielsweise noch nicht vergeben oder zu inflationär verwendet, sind die Trefferzahlen gegen 0 gehend oder viel zu hoch, sodass auf ein genaueres Sichten der Ergebnisse ob der Menge an Treffern verzichtet wird. Dieser Zustand, gepaart mit den bereits vorhandenen Zeitproblemen im Produktionsprozess, erschwerte oft ein ausführliches Recherchieren. Durch die Bündelung von Musiken und Informationen in Playlists lässt sich eine überschaubare Materialsammlung anbieten, die gegen geringen Zeitaufwand einen ersten Überblick zum Thema liefern und Material zur Verfügung stellen kann.

Nicht nur auf Redaktionsseite können mit PARAS Probleme gelöst werden, auch einige archivseitige Wünsche lassen sich mittels dieses Projektes vorbringen. So ermöglichen die Playlists erstens, proaktiver in die Programmgestaltung einzugrei-

fen, zum Beispiel über eine gezielte Platzierung von Archivschätzen. Zweitens führen die Einbindung in den Produktionsprozess und die damit einhergehenden Gespräche mit den Redaktionen zu einer natürlichen Kooperationssteigerung. Drittens ermöglicht dieser stärker kooperative Produktionsprozess, uns deutlicher als Partner der Redaktionen zu positionieren – trotz Dienstleistungscharakter des Projekts. Hierüber erhöhen wir unsere Sichtbarkeit im Haus, auch als diejenigen, die die Deutungshoheit über die Daten besitzen. Zuletzt ermöglichen uns die Rückmeldungen zu den Playlists, eine detailliertere und passgenauere Bedarfsanalyse durchzuführen, sowohl mit Blick auf die einzelnen Wellen als auch zu Abstimmung der Recherchen für folgende Großereignisse und Themenschwerpunkte. Die zukünftigen KI- und Automatisierungstools werden es uns währenddessen erlauben, unseren Fokus stärker auf die Servicequalität zu legen.

Playlist-Design

Auf der Suche nach Playlistdesigns, mit denen sich möglichst viele relevante Themenkomplexe abdecken lassen, haben sich drei unterschiedliche Design-typen hervorgetan, anhand derer sich auch die unterschiedlichen relevanten Thementypen erläutern lassen.

1. Ereignisplaylists: Diese können im Kontext anstehender Jubiläen oder aktueller Ereignisse erstellt werden, gerade wenn diese Ereignisse eine größere Relevanz für den Rundfunk haben, zum Beispiel Olympische Spiele 1972 in München. Diese Playlists zeichnen sich durch ihren einmaligen Erstellungs- und minimalen Aktualisierungsaufwand aus, da das Ereignis in einem meist relativ kurzen Zeitraum erfolgte. Durch feststehende Jahrestage und Jubiläen weisen sie zudem hohen Wiederverwendungswert auf. Zur Berichterstattung über solche Ereignisthemen gehören meist Specials, weswegen es lohnenswert ist, das Thema aus mehreren Blickwinkeln zu betrachten. Einige der Unterkategorien, die für die Playlist zu Olympia '72 genauer beleuchtet wurden, lauteten zum Beispiel: (1) zeitgenössische Titel, die im Rahmen der Spiele liefen; (2) Songs, die sich seitdem mit Olympia oder Sportereignissen beschäftigt haben; oder (3) Werke, die Attentate thematisieren. Gerade aufgrund dieser Mannigfaltigkeit an Themen eignen sich die Ereignisplaylists besonders um die unterschiedlichen Wellen mit ihren eigenen musikalischen Anforderungen zu bespielen.

2. Dauerbrenner-Playlists: Diese behandeln Themenkomplexe, die in Abhängigkeit vom Tagesgeschehen jederzeit relevant werden können, zum Beispiel der Klimawandel, Naturkatastrophen oder Migration. Diese Playlists weisen einen deutlich höheren Aktualisierungsaufwand auf, dafür können sie häufiger punktuell verwendet werden, wenn man beispielsweise nach der Berichterstattung über erneute Waldbrände einen der Songs spielt, die als Reaktion auf die Brände in Kalifornien 2021 komponiert wurden – und das waren aufgrund der geographischen Lage L.A.S in der Populärmusik einige. Diese Playlists dienen somit als Materialsammlung zur schnellen Selbstbedienung, wenn passende Musik zu bestimmten Themen benötigt werden. Und musikwissenschaftlich lassen sich dabei schöne Analysen durch die Jahrzehnte und Musikstile hindurch vornehmen, da die meisten größeren Themenkomplexe schon länger Bestandteil der kulturellen Auseinandersetzung sind. Aufgrund dieser jahrzehntelangen Beschäftigung mit den Themen in der Popkultur lassen sich auch hier wieder alle Wellen mit ihren eigenen Fokussen bespielen.

3. Einmaliges: Einmalige Materialsammlungen werden im Zusammenhang mit Ereignissen der Musikszene meist in zwei Fällen benötigt: größere Berichterstattungen und Specials zu Jubiläen, z.B. 50 Jahre Rolling Stones oder ausführliche Berichterstattungen im Todesfall eines Künstlers. Für längere Berichterstattungen sollten die notwendigen Recherchen den redaktionellen Experten überlassen werden. Gerade im Falle einer Todesnachricht wird jedoch meist sehr schnell passendes Material in Form von 2-3 bekannten, thematisch passenden, oder ruhigen Songs des jeweiligen Künstlers oder seiner Band benötigt, mit dem die Recherchedauer für weiteres, wellenspezifisches Material überbrückt werden kann. Über die Erstellung kleinerer Dossiers im Medienbroker lässt sich sicherstellen, dass diese Songs schnell auffindbar, bereits digitalisiert und damit sofort einsatzbereit sind.

KI-Möglichkeiten in der Musikrecherche – LLMs als Recherche- & Analysetool

Die Erstellung dieser Playlists ist derzeit noch mit einem hohen händischen und intellektuellen Arbeitsaufwand verbunden. Dieser lässt sich durch den gezielten Einsatz von KI- und Automatisierungstools

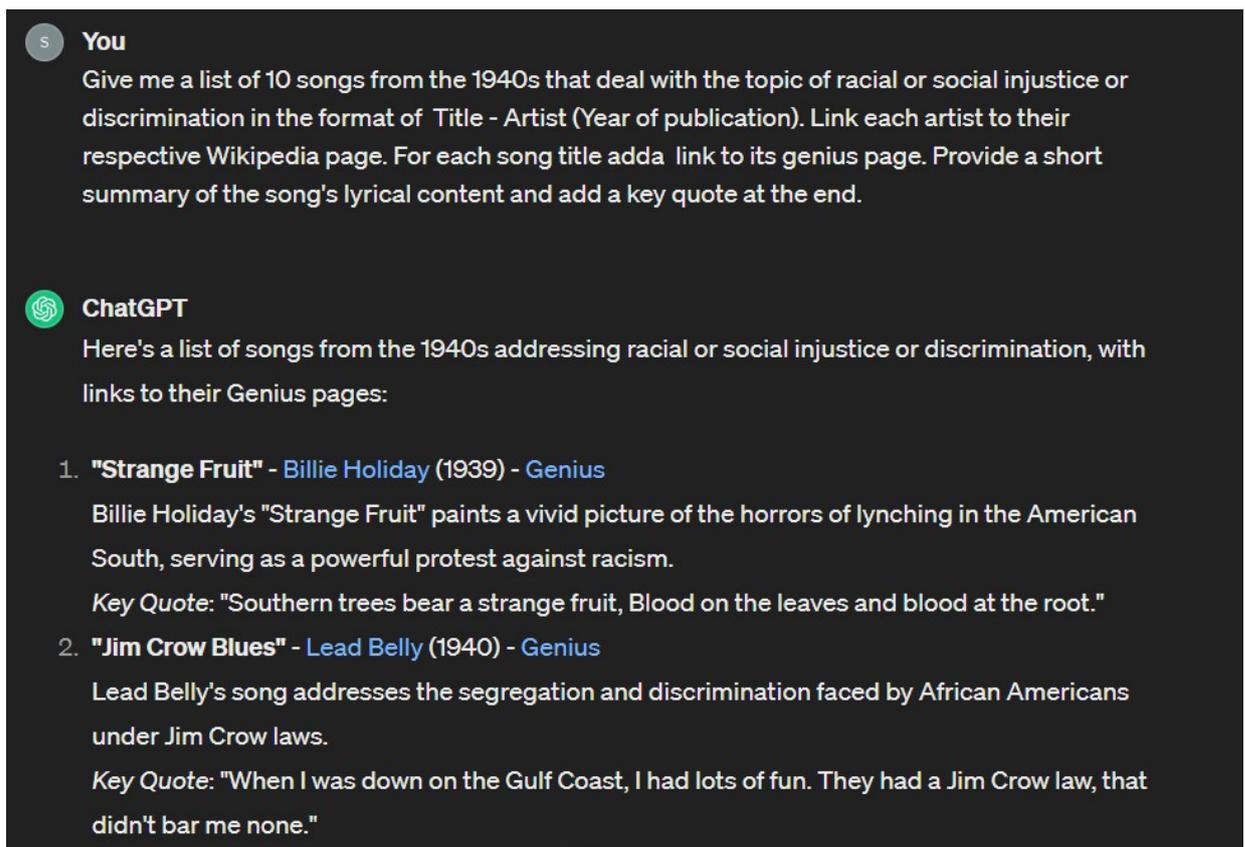


Abb. 1.: Screenshot Prompt Songsausgabe Thema Rassismus 1940er

reduzieren, wie nun am Beispiel von Large Language Models (LLMs) gezeigt wird. Die Auswahl und der genaue Einsatz der KI-Tools müssen sich hierfür jedoch nach der Aufgabe richten. Die Wahl von LLMs als Songanalysetool erfolgte hier, da diese aufgrund der semantischen Fähigkeiten besser zur Analyse von lyrischen Texten geeignet sind als statistische Textmining-Methoden, wie zum Beispiel topic modeling. In Abhängigkeit von den gerade analysierten Arten von Playlists müssen differenzierte Fragen und Prompts an die LLMs gestellt werden, die sich ebenfalls in drei Kategorien gliedern lassen.

Ausgabe von Unterthemen/Kategorien/ etc. zu Themen

Wie vorangehend erläutert, lohnt sich themenabhängig häufig eine Gliederung der Playlists in Unterkategorien, um den Themenkomplex aus diversen Blickwinkeln zu beleuchten und Anhaltspunkte für weitere Rechercheperspektiven zu geben. Durch die semantische Analyse ihres Trainingsmaterials sind LLMs in der Lage, solche möglichen Perspektiven auszugeben und erste Beispiele zu den Unterkategorien zu nennen. Betrachten wir hierzu das Beispiel Olympia '72 in München. ChatGPT wurde hierzu aufgefordert, als Journalist zu agieren und

einen Beitrag über die Olympischen Spiele zu konzipieren, wofür es in einem ersten Schritt relevante Aspekte nennen sollte. Diese lauteten: „Political Tensions and Cold War Context“, „Security and Terrorism“, „Legacy and Impact“, „Memorialization and Remembrance“, „Athlete Experiences and Perspectives“, „International Relations and Diplomacy“ sowie „Cultural Exchange and Diplomacy“. Von der Doppelnennung der Diplomatie abgesehen, kann diese Liste in einem zweiten Schritt gewinnbringend genutzt werden, um zu diesen Unterthemen jeweils nach passender Musik zu suchen. Der Mehrwert der Verwendung von ChatGPT ergibt sich hierbei vor allem aus der Erweiterung der eigenen Perspektive, sodass zum Beispiel Kategorien in die Playlisterstellung einfließen können, an die man selbst nicht gedacht hätte, die aber in den Trainingsdaten des Language Models aufgetaucht waren.

Recherche im Internet mit LLMs

Eine weitere Einsatzmöglichkeit der LLMs in der Playlisterstellung ergibt sich aus der, vor allem den Dauerbrennerthemen inhärenten, Eigenschaft, über mehrere Dekaden hinweg analysiert werden zu können. So lassen sich zum Thema Klimawandel

beispielsweise Folksongs recherchieren, die bereits in den 1950ern entstanden und ebenso Naturkatastrophen thematisierten wie Titel, die in den letzten 3 Jahren als Reaktion auf die Waldbrände in Kalifornien komponiert wurden. Um diese Verbindungen zu finden und Vergleiche anstellen zu können, ist jedoch ein erheblicher Rechercheaufwand nötig. Dieser kann nun durch die Recherchefunktion gängiger LLMS erheblich reduziert werden, indem die Recherche im Hintergrund erfolgt und in die Antwort auf den Prompt einfließt (geprüft wurden zum Zeitpunkt des Artikelverfassens OpenAI's ChatGPT 4.0, Microsofts Copilot sowie Perplexity). Zur Portionierung der Anfragen eignet sich eine Paketgröße von 10 Songs pro Abfrage. Hier ein Beispielprompt zur Erstellung einer Playlist, die das Thema Rassismus / Diskriminierung behandelt: „Gib mir eine Liste mit 10 Songs aus den 1940ern, die die Themen soziale Ungerechtigkeiten, Rassismus oder Diskriminierung behandeln, im Format Titel – Interpret (Jahr). Verlinke die Interpreten mit ihren jeweiligen Wikidataseiten. Füge für jeden Song einen Link zur Genius-Seite ein. Gib außerdem eine kurze Zusammenfassung des Songinhalts und ein prägnantes Zitat aus.“ Über die verlinkte Wikidata-Seite kann eine erste Verifikation des Artists erfolgen. Genius, eine Seite für Songtexte, erfüllte dieselbe Funktion dann für die Existenz der jeweiligen Titel. (Abbildung 1)

Die genannten LLMS liefern unterschiedliche Ergebnisse, wobei sich die Auswahl des Tools nach der gewünschten Musik richten kann. So weist ChatGPT durch das Trainingskorpus mehr Stärken in der Recherche westlicher (US-amerikanischer und europäischer) Musik auf. CoPilot hingegen liefert diversere Ergebnisse, während hier zum Beispiel mit Hinweis auf Unbekannte Urheber in der Volks-/Folk-Musik die Verlinkungen zu Artists und Lyrics ab und an nicht funktionieren. Als letzter Nutzen der LLMS kann in dieser Form der Playlisterstellung auf die Zusammenfassungen der Songs hingewiesen werden, die meist gute Ergebnisse liefern.

Analyse von Songtexten

Wie eingangs erwähnt, eignen sich LLMS zur Analyse lyrischer Texte vor allem aufgrund des höheren „semantischen Verständnisses“ im Vergleich zu statistischeren Methoden des Textmining. Zwei Methoden sollen hier kurz beschrieben werden. Erstens kann dem LLM der Songtext – eingebettet

in einem Prompt – gegeben werden, in dem die Auswahl einer oder mehrerer Deskriptoren gefordert wird, die entweder aus einer vorgefertigten Liste selektiert oder frei vergeben werden sollen. Diese Deskriptoren kann man dann entweder zur Verschlagwortung der Datensätze in der eigenen Datenbank oder zur Verifikation der zuvor ausgegebenen Songs verwenden. Zweitens kann die Inhaltsangabe aber auch in Form ganzer Sätze gefordert werden, die dann wiederum als Begründung der Songauswahl in die Playlists eingebaut werden können.

Handlungsempfehlungen und Learnings

Abschließend sollen hier noch die aktuellen Probleme der LLM-Verwendung zur Playlisterstellung betrachtet werden sowie die aus dem Projekt resultierenden Learnings.

Erstens stellt sich bei den meisten KIs weiterhin die Frage nach der Herkunft und Qualität der Datenquellen und Trainingsdaten sowie dem rechtlichen Zugang zu diesen. Damit sind in erster Linie die Trainingsdaten insofern gemeint, dass die auf Wahrscheinlichkeiten beruhenden Antworten der LLMS durch falsche Informationen im Trainingsmaterial verzerrt werden können. Es geht aber auch bei der Einbindung von Onlinerecherchen um die Frage, welche Quellen und Suchmaschinen verwendet werden. So liefert Microsofts CoPilot, der vor allem über Bing recherchiert, deutlich andere Ergebnisse als Perplexity. Zweitens und ebenfalls teilweise auf die Trainingsdaten zurückzuführen, ist der teilweise stark US-amerikanische Fokus der Recherche und Textausgabe, der auf das zu Beginn vornehmlich englischsprachigen Trainingsmaterial zurückzuführen ist. Eine Auslegung auf US-amerikanische Nutzer wird jedoch auch in den Richtlinien ChatGPTs deutlich, wenn beispielsweise Themen, die in den USA aktuell medial aufgeladen sind, als „gegen die Richtlinien verstoßend“ eingestuft werden. So wird zum Beispiel der Songtext „Brenda's got a baby“ von 2Pac, der Teenagerschwangerschaften im Ghetto thematisiert, derzeit nur bei 2/5 Versuchen analysiert. In 60% der Fälle wird auf einen Verstoß gegen die Richtlinien hingewiesen: (Abbildung 2)

Da normalerweise nicht jeder Prompt mehrmals eingegeben wird, ist hier davon auszugehen, dass die benötigten Informationen nicht erhalten werden würden.

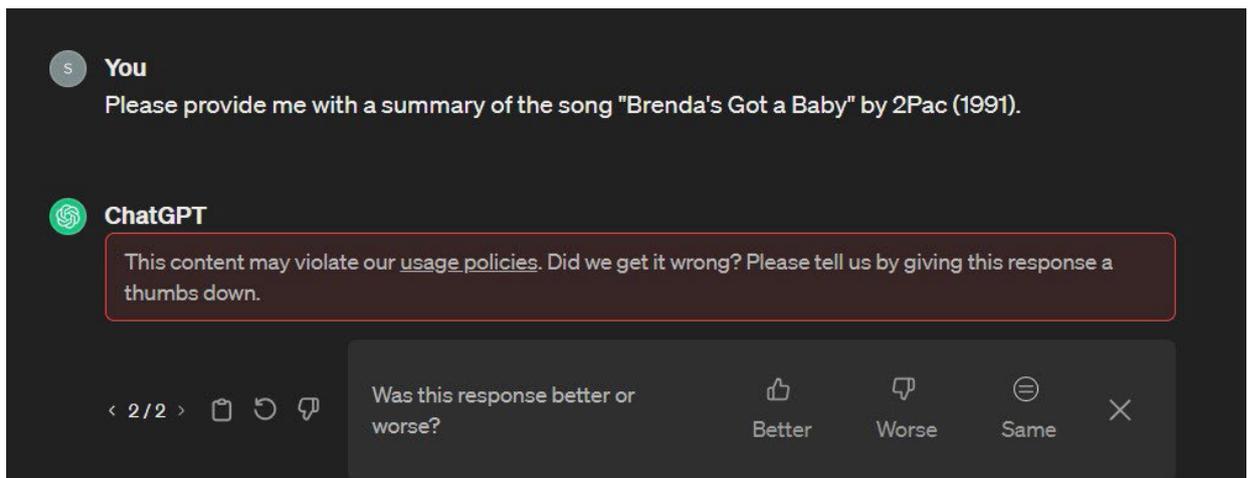


Abb. 2: Screenshot Prompt Zusammenfassung Songtext zu „Brenda’s got a baby“

Drittens ist als wohl bekanntestes Problem zu nennen, dass auch Songtexte nicht von Halluzinationen verschont bleiben: Stellenweise werden neue Titel oder Künstler erfunden oder aber Songtexte von Songs zitiert, die es gar nicht gibt. Nicht so essentiell, aber dennoch einschränkend, ist die Tokenbeschränkung je nach Bezahlmodell und Anbieter. Große Mengen an Songtexten sollten für die Eingabe daher gut portioniert werden.

Viertens bleibt die rechtliche Lage zur Speicherung, Verwendung und Ausgabe von Songtexten weiterhin ungeklärt. Früher konnte man ChatGPT noch um Songtexte bitten, heute kommt der Hinweis, dass diese dem Urheberrecht unterliegen und höchstens eine Zusammenfassung möglich ist. Und zuletzt bleibt der menschliche Kontrollaufwand, da alle Angaben aus den LLMs stets nochmal auf ihre Richtigkeit überprüft werden müssen, bevor diese an die Redaktionen weitergegeben werden können.

Die Erstellung der Playlists ist derzeit also auch unter Verwendung von ChatGPT und CO. noch mit einem hohen händischen und intellektuellen Arbeitsaufwand verbunden. Die Learnings aus den Anwendungstests der LLM-Nutzung können als wie folgt zusammengefasst werden: Als Recherchetool – vor allem zum Brainstorming – können die unterschiedlichen Modelle bereits gewinnbringend genutzt werden. Die menschliche Verifikation, der „human in the loop“, wird aber noch lange notwendig bleiben. Der KI-Einsatz kann dabei die intellektuelle Arbeit zwar nicht ersetzen, aber stellenweise erleichtern und definitiv anreichern.

Für die Zukunft der Medienarchive werden hierdurch interessante Perspektiven eröffnet, wie ich am Beispiel des PARAS-Projektes abschließend

nochmal zusammenfassen möchte. PARAS steht exemplarisch dafür, wie sich unser Arbeitsfeld in Zukunft entwickeln wird: Weg von Teilen der Erfassungstätigkeit, hin zur Dienstleistung. KI-Tools wie nun schon die LLMs, werden es uns ermöglichen, mit geringem Zeitaufwand neue Produkte zu entwickeln, die proaktiv an die Redaktionen herangezogen werden können. So können zum Beispiel Playlists zahlreiche Bedürfnisse der Redaktionen im zeitintensiven Produktionsprozess erfüllen. Getrieben durch die aktuellen Innovationen im KI-Bereich findet hier gerade ein Aushandlungsprozess statt, bei dem es unter anderem gilt, Arbeitsweisen, Rollen und Hierarchien neu auszuhandeln. Durch eine aktive Gestaltung dieses Prozesses – u.a. durch den bewussten Einsatz von KI – positionieren wir uns als Partner der Redaktionen und behalten die Deutungshoheit über die Daten. Gerade unser Wissen um den Bestand und die Infrastruktur prädestinieren Archive dafür, in Zukunft solche Dienstleistungen wie PARAS anzubieten. Dieser Prozess muss jedoch bewusst gesteuert werden. ■■■

AI Next Level

Was Künstliche Intelligenz alles verändern wird

Prof. Dr. Hannah Bast

Frau Prof. Bast hielt auf der Frühjahrstagung des vfm am 22. April 2024 den Eröffnungsvortrag über die Entwicklung künstlicher Intelligenz. Für info7 hat sie eine Zusammenfassung ihres Vortrags verfasst, die hier abgedruckt ist. Der Vortrag selbst steht als Mitschnitt zur Verfügung. Neu in info7 ist, dass wir auf bestimmte Abschnitte des Vortrags verlinkt haben. Aus der digitalen Version von info7 können diese direkt angesteuert werden. Für Leser:innen des analogen Heftes helfen die Minutenangaben, innerhalb des Mitschnitts zu navigieren.

Es ist keine Übertreibung, die Entwicklungen in der KI im letzten Jahrzehnt als revolutionär zu bezeichnen, mit langfristigen Folgen für die Wirtschaft, die Gesellschaft und die Wissenschaft, die mindestens so einschneidend sein werden wie die Erfindung des Internets oder der Dampfmaschine.

In diesem Vortrag wird anhand einiger sorgfältig ausgewählter Beispiele gezeigt, was KI schon alles beherrscht, das vor wenigen Jahren noch Science Fiction war. Dazu gehören insbesondere das praktisch fast

schon perfekte Beherrschen von menschlicher Sprache, gepaart mit einem schon beeindruckend tiefgehenden Bildverständnis, sowie die Fähigkeit, auch gesprochene Sprache nicht nur zu verstehen, sondern auch so zu generieren, dass sie von der eines Menschen praktische nicht mehr zu unterscheiden ist. Als ein weiteres, für die Professionen des Publikums besonders relevantes Beispiel wird die automatische Extraktion von Metadaten aus einem Zeitungstext vorgeführt. Ebenfalls ist KI bereits in der Lage, komplexe Datenbankanfragen zu erzeugen und selbstständig selber Programme zu schreiben.

Hinter all diesen, fast magisch anmutenden, Leistungen stehen sog. tiefe neuronale Netze, deren Funktionsweise im Vortrag von Grund auf und anschaulich erklärt wird. Damit diese Netze funktionieren, müssen sie hinreichend groß sein und aufwendig trainiert werden. Das war bis vor Kurzem noch das Privileg sehr weniger großer Firmen, allen voran OpenAI. Seit dem Überraschungserfolg von ChatGPT vor zwei Jahren ist die Entwicklung in diesem Ökosystem explodiert, und es gibt inzwischen zahlreichen Firmen und Initiativen, die solche Netze zur Verfügung stellen, teilweise ohne jegliche Gegenleistung oder sogar open source, das heißt inklusive dem zugehörigen Programmcode.

Der Vortrag schließt mit einem Ausblick in eine ungewisse Zukunft. In der kurzfristigen Perspektive ist die KI vor allem ein sehr mächtiger Assistent und kein Orakel. Allerdings schreitet die Entwicklung weiter weiter rasant voran und es ist noch kein Ende in Sicht. ■



Prof. Dr. Hannah Bast
Lehrstuhl für Algorithmen und Datenstrukturen gruppiert nach Funktion
Albert-Ludwigs-Universität Freiburg
bast@informatik.uni-freiburg.de

www.info7.de/2024-4-22_Eroeffnungsvortrag_Bast.mp4

Einleitung: Das Trolley-Problem

Wie funktioniert ein neuronales Netz?

▶ 06:00 Minute

Schematische Darstellung eines NNs

▶ 15:34 Minute

Es gibt (inzwischen) viel mehr als ChatGPT

▶ 22:28 Minute

Anwendungen

▶ 27:00 Minute

Frequently Asked Questions

▶ 44:28 Minute

Diskussion

▶ 47:05 Minute



Künstliche Intelligenz und ihre Auswirkung auf die Gesellschaft

Rezension von Stefan Holländer

Künstliche Intelligenz (KI) ist in aller Munde, seit ChatGPT Ende 2022 auf den Markt kam. Im gleichen Masse haben auch die Fachpublikationen zu diesem Thema einen neuen Höhepunkt erreicht. Das vorliegende Buch entstand aus einer Ringvorlesung zur KI an der Hochschule Augsburg, die interdisziplinär von verschiedenen Beiträgern gestaltet wurde, unter der wissenschaftlichen Leitung von Prof. Dr. László Kovács. Künstliche Intelligenz kommt in vielen Anwendungen in unserem Alltag vor, man denke nur an Online-Suchmaschinen, Soziale Netzwerke oder Streaming-Plattformen. Für Industrie und Wirtschaft bringt dies grosse Herausforderungen. Sie müssen grosse Investitionen tätigen und über grosse Rechenleistungen verfügen, um weiterhin marktfähig zu bleiben. KI scheint alle Bereiche des Alltags durchdringen zu wollen. Auch Bibliotheken müssen sich die Frage stellen, welche Künstliche Intelligenz sie brauchen und welche nicht?

Das Buch erklärt einige Grundprinzipien der KI, bietet Einblicke in verschiedene Anwendungen in den unterschiedlichsten Lebensbereichen und liefert interessante Anregungen. Die technischen Beiträge halten, was der erste Teil des Buchtitels verspricht: eine Erklärung der Funktionsweise von KI in Anwendungen. Gleich der erste Beitrag von Claudia Meitinger über eine kognitive Anwendung bei der Entscheidungsfindung am Beispiel des Schafkopfspiels zeigt einige Grundprinzipien künstlicher Intelligenz auf. Auch ein Rückblick auf die Entwicklungsgeschichte von KI durch Rudolf Seising ist informativ zu lesen. Das auf YouTube¹ verfügbare Video seines Vortrags zum gleichen Thema anlässlich der Ringvorlesung fällt jedoch wesentlich detailreicher aus. Einen interessanten Einblick in die Schnittstelle Mensch – Maschine und was uns die Zukunft noch bringen kann, gibt uns der kurz und prägnant gehaltene Beitrag von Mathias Vukelic und Nekataria Tagalidou. Einen ganz wichtigen und konzisen Beitrag zu KI und Security bei der Bilderkennung haben Hella Hollman und Philipp Schurk beigesteuert.



László Kovács (Hg.)
Künstliche Intelligenz und
menschliche Gesellschaft
231 S., De Gruyter Verlag,
ISBN: 9783111034492

Hier, wie bei allen Beiträgen muss man sich vor Augen halten, dass deren Konzeption und Ausführung auf dem Stand von 2022 beruhen und dass die Vorträge für die Buchpublikation bestenfalls noch einmal überarbeitet wurden. Bei der rasanten Entwicklung auf diesem Gebiet sind Anmerkungen zum jüngsten Stand der Erkenntnisse nur bedingt eingearbeitet.

Bibliotheksrelevante Themen, dies sei hier vorweggenommen, werden im Buch nicht oder nur sehr entfernt angesprochen. Trotzdem lohnt sich die Lektüre der Beiträge. Der Beitrag zu KI-gestützten Entscheidungsunterstützungssystemen im Gesundheitssystem von Susanne Gaube und Eva Lermer zeigt einleuchtend auf, wie sehr die Technologieakzeptanz von den Anwendern abhängig ist und wie wichtig es deshalb ist, die Determinanten zur Nutzungsbereitschaft für neue Technologien gut zu berücksichtigen, wenn diese Technologien erfolgreich eingesetzt werden sollen. Ein Thema, mit dem sich die Bibliotheken künftig wohl mehr beschäftigen müssen, ist die Ermittlung der Authentizität von Information. Der Beitrag von Marco Viviani und Elisabeth Mess zur Informationssuche beim Information Retrieval im Gesundheitsbereich für Verbraucher beschreibt einige Forschungseinrichtungen und ihre Forschungslösungen zur Authentizitätsüberprüfung von Informationen. Solchen Technologien sollten sich Bibliotheken auch nicht verschliessen, um gegen künftige Herausforderungen gewappnet zu sein.

Der abschliessende Beitrag vom Herausgeber László Kovács zu den Leistungsgrenzen der KI in der Zukunft entspricht dem Versprechen des Buchtitels in Bezug auf gesellschaftliche Auswirkungen der KI-Technologie nicht völlig. Einig ist man sich mit dem Verfasser, dass die KI-Technologie gemäss ihrem heutigen Entwicklungsstand eine unterstützende Funktion hat, aber den Menschen nicht vollständig ersetzen wird. Auch wenn man sich vor Augen hält, dass der abschliessende Beitrag konzeptionell auf dem Erkenntnisstand von 2022 beruht, bleibt er mit seinen Prognosen zu sehr im Ungefähren.

Andere Publikationen, die im Jahre 2022 erschienen sind, enthalten wesentlich informativere Beiträge, so etwa die Publikation von Thomas Hören und Stefan Pinelli mit dem Titel „Künstliche Intelligenz und Recht“². Abschliessend bleibt festzuhalten, dass die technischen Beiträge des Buches dem Buchtitel gerecht werden. ■

¹ <https://www.youtube.com/watch?v=aPVIoZGIODQ>

² Künstliche Intelligenz – Ethik und Recht / Thomas Hören, Stefan Pinelli u.a.. – München : C.H.Beck, 2022. – 393 S. ; Abb., Tab. – (Schriftenreihe Information und Recht, Bd. 87). – dt. und engl. – ISBN 978-3-406-77388-4

FaktenFokus

für Faktencheck und Verifikation

Thiemo Kremser

Willkommen beim „FaktenFokus“, der neuen Rubrik für Faktencheck und Verifikation in der info7! Auch in dieser Ausgabe stellen wir wieder drei ausgewählte Ressourcen vor, die vertiefende Einblicke und praktische Werkzeuge für die tägliche Überprüfung von Informationen bieten. Ob aus beruflichem oder privatem Interesse an zuverlässigem Faktencheck und digitaler Recherche, diese Links bieten wertvolle Hilfestellungen und interessante Einblicke. Viel Spaß beim Lesen, Anschauen und Ausprobieren!

Tool-Tipp: Fact Check Explorer für die Bilderverifikation

Google hat zum „Internationalen Tag des Faktenchecks“ vier neue Möglichkeiten vorgestellt, die dabei helfen, Informationen schneller zu bewerten und Zusammenhänge zu analysieren. Besonders interessant ist der „Fact Check Explorer“, mit dem Faktenchecks zu einem Thema gefunden werden können. Diese Funktion, die sich mittlerweile nicht mehr in der Beta-Phase befindet, ermöglicht es auch, ein Bild hochzuladen oder zu verlinken, um per Bilder-Rückwärtssuche zu überprüfen, ob es bereits in einem Faktencheck verwendet wurde und wann Google das Bild zum ersten Mal indexiert hat.

Hier geht es zum Blog-Artikel

 <https://blog.google/products/search/google-search-fact-checking-resources>

Hier zum Tool

 <https://toolbox.google.com/factcheck/explorer/search/list:recent;hl=de>

Impulsvortrag: Neues Open-Source-Recherche-Toolkit von Johanna Wild (Bellingcat)

Johanna Wild, OSINT- & Tech-Expertin, gibt in einem kurzen Vortrag (10 min) einen Einblick in ihre Erkenntnisse bei der Entwicklung eines neuen kollaborativen „Open Source Research Toolkits“. Das Online-Recherche-Toolkit von Bellingcat hat eine lange Tradition, doch die neueste Version ist etwas Besonderes: Sie wird in Zusammenarbeit mit der

„Bellingcat Volunteer-Community“ angeboten. Vorteil: Die „Volunteers“ kennen „ihre Werkzeuge“ in- und auswendig und können Auskunft über deren Vorteile und Grenzen geben.

Hier den Vortrag ansehen

 <https://www.youtube.com/watch?v=CqjMQwLoVK8&t=3867s>

Hier geht es zur neuen Bellingcat-Toolbox

 <https://bellingcat.gitbook.io/toolkit>

Prompt-Tipp: Effektives Faktenchecken mit „Chain-of-Verification“

Zum Abschluss der aktuellen Ausgabe stellen wir eine innovative Methode zur Faktenüberprüfung vor, die auf der Website von Mister SEO beschrieben wird: die „Chain-of-Verification“. Diese Methode eignet sich nicht nur zur Überprüfung von Fakten und Behauptungen, sondern auch zur Identifizierung und Korrektur von Fehlinformationen durch KI-Systeme. Die Methode verwendet einen schrittweisen Ansatz, um die Zuverlässigkeit der generierten Antworten zu erhöhen. Prompt:

Du bist ein Faktenchecker und Experte im Recherchieren und Verifizieren von Informationen. Nutze den 'Chain of Verification'-Ansatz, um Halluzinationen in Antworten zu reduzieren. Beginne mit einer der folgenden Aussage, die Halluzinationen enthalten kann. Stelle dann Verifikationsfragen, um Fakten zu überprüfen. Nutze Methoden wie Joint, 2-Step, Factored und Factor+Revise, um Antworten zu verifizieren. Erstelle abschließend eine verifizierte Antwort unter Berücksichtigung entdeckter Inkonsistenzen. Hier ist die Aussage: [AUSSAGE].

Mehr zum Konzept der „Chain-of-Verification“ im Hintergrundartikel von Mister SEO

 <https://mister-seo.com/chatgpt-faktencheck-chain-of-verification>

und bei DW-Journalist Patrick Große

 <https://ajjournalist.de/prompt-tipp-faktencheck-mit-chain-of-verification>

Hier geht es zum GPT: Beyond GPT Hallucinations – Fact Checker von Mister SEO

 <https://chatgpt.com/g/g-EiksB3luQ-beyond-gpt-hallucinations-fact-checker>

Wie immer freuen wir uns über eure Anregungen für weitere Inhalte, Tipps & Tricks und Tools oder auch Feedback zu den vorgestellten Ressourcen.

Viele Grüße und bis zur nächsten Ausgabe,
Thiemo. ■■■



Thiemo Kremser
Hessischer Rundfunk
Frankfurt a. M.
Thiemo.Kremser@info7.de

Marianne-Englert-Preis

newcomer-forum im vfm e.V. – Neues aus den Hochschulen

Heiko Linnemann



Heiko Linnemann
Documentation Greenpeace
heiko.linnemann@
greenpeace.org

An dieser Stelle berichten wir vom letzten Programmpunkt eines ereignisreichen ersten Tages auf der Frühjahrstagung 2024 in Freiburg. Wir hatten bis dahin beispielsweise gelernt, dass das Wissen über neuronale Netze, die zum Betrieb von Anwendungen der Künstlichen Intelligenz die Grundlage bildet, bereits mehr als 70 Jahre alt ist. Außerdem, dass erst jetzt die vielfältigen Anwendungen in der KI zum Leben erweckt werden können,

weil die Rechenleistungen der IT-Systeme inzwischen schnell genug dafür sind (vgl. Eröffnungsvortrag von Prof. Dr. Hannah Bast).

Wir haben Werkstattberichte zum Stand von KI-Entwicklungsprojekten in sowohl öffentlich-rechtlichen als auch privaten Rundfunk-Archiven gehört und den aktuellen Stand zu Problembereichen und Chancen in diesem Umfeld erörtert. In der Sponsoren-Session haben sich im Anschluss Unternehmen vorgestellt, die uns bei der Bewältigung dieser Problemstellungen behilflich sein können.

Danach ging es weiter mit dem newcomer forum, welches der vfm bei den Frühjahrstagungen seit 2012 kontinuierlich ausrichtet. Für diejenigen, die noch an keiner der Veranstaltungen teilgenommen haben, sei an dieser Stelle erwähnt, was wir uns da-

von versprechen. Es geht uns darum zu erfahren, welche Themen unseren „Branchennachwuchs“ an den Hochschulen, bzw. den Bildungseinrichtungen mit bibliothekarischer oder dokumentarisch-technischer Ausrichtung umtreiben. Wir wollen von den drei besten (Hochschul)-Projekten hören, die wir seit der Tagung im vergangenen Jahr kennengelernt haben und die Menschen vorstellen, die an diesen Projekten maßgeblich beteiligt waren.

Das waren in diesem Jahr: Jasmin Casarano, Lena Wernhöfer und Stefan Merl. Sie sind damit die Gewinner:innen des Marianne-Englert-Preises 2024, zu dem wir noch einmal herzlich beglückwünschen! Der Preis, den die jungen Kolleg:innen mit ihren herausragenden Arbeiten gewonnen haben, setzt sich aus folgenden Teilen zusammen:

- einem finanziellen Beitrag (jeweils 500€),
- der Teilnahme inklusive Vortrag vor Fachpublikum auf der Frühjahrstagung,
- der Möglichkeit, den Vortrag in der Vereinszeitschrift info7 zu veröffentlichen.

Auch im kommenden Jahr werden wir wieder drei Marianne-Englert-Preise vergeben.

Worum es in den 2024 prämierten Arbeiten im Einzelnen geht, können Sie in den folgenden Artikeln nachlesen. ■



Verleihung des Marianne-Englert-Preises



Gewusst wie

Ein interaktiver Selbstlernkurs für die Personenansetzung in der ARD-Normdatenbank

Jasmin Casarano

Die ARD-Normdatenbank fungiert als Repository für Entitäten und Vokabulare der Rundfunkanstalten der ARD. Für den Aufbau eines cross-medialen Mediendatensystems ist es wichtig, die Personendatensätze einheitlich anzulegen, da die Datensätze ARD-weit und von verschiedenen Systemen genutzt werden. Das Regelwerk für die Personenansetzung ist allerdings knapp 100 Seiten lang – das liest sich kaum jemand freiwillig durch. Im NDR gibt es dafür nun ein E-Learning-Angebot.

Die Datenqualität der ARD-Normdatenbank (NDB) spielt eine wichtige Rolle, denn die Datensätze werden auch von automatisierten Prozessen genutzt. Die angeschlossenen Systeme im NDR und der ARD (FESAD, HFDB etc.) greifen auf die Datensätze der NDB zu, wenn dort z.B. Personennamen für die Musikalische Darbietung, Komposition, Rede u.v.m. erfasst werden. Ist der Name nicht recherchierbar, so wird von den Dokumentar*innen entweder direkt aus dem angeschlossenen System oder in der NDB ein neuer Namens- oder Personendatensatz angelegt. Dafür gibt es bestimmte Regeln und Empfehlungen, die durch die AG Regelwerk Metadaten herausgegeben werden.

Problemstellung

Nicht alle im NDR-Archiv kennen das Regelwerk Metadaten sicher, denn es ist umfangreich und kann beim Einstieg überfordern. In der Vergangenheit gab es Workshops zum Regelwerk und der Personenansetzung, aber keine terminunabhängige Lösung oder andere Lernangebote. So kommt es immer wieder zu doppelten Ansetzungen von Personennamen und einem hohen Nachpflegeaufwand für das Datenpflege-team im NDR, weil das Zusammenspiel der Datenbanken nicht bekannt ist. Die Aufgabe des Datenpflege-teams ist es, die Datensätze zu korrigieren, anzureichern und / oder zu verschmelzen, sodass die unterschiedlichen Anset-

zungen als Synonyme oder nicht zugelassene Variante am selben Datensatz hängen. In der NDB werden dafür Aufgabenlisten angezeigt, die nach dem Verursacherprinzip gefüllt werden. Das heißt, Datensätze, die aus den angeschlossenen Systemen des NDR in der NDB landen, werden auch nur dem NDR-Datenpflege-team in der Aufgabenliste angezeigt. In den letzten Jahren hat sich in den Listen einiges angesammelt und das Datenpflege-team wurde 2023 mit Kolleginnen aufgestellt.

Zurzeit gibt es ca. 7100 Bearbeitungsfälle für die Entität Personen (siehe Abbildung 1). Mit einem Arbeitsaufwand zwischen einer Minute und einer Stunde (je nach Nachvollziehbarkeit und Rechercheaufwand) ist es neben den Hauptaufgaben schwierig, diese Datensätze abzarbeiten, während täglich neue dazu kommen. Weiterhin fallen auch im Arbeitsalltag ältere Datensätze auf, die korrigiert oder verschmolzen werden müssen. Um die Datenqualität schon bei der Ansetzung zu verbessern und die anfallende Korrekturarbeit des Datenpflege-teams zu verringern, wurde ein Lernangebot für die Dokumentar*innen entwickelt. Aus Kapazitätsgründen habe ich mich auf die Entität Personen beschränkt, weil diese tagtäglich am häufigsten angesetzt werden.

Ziele

Die Idee bestand darin, ein kompaktes und interaktives E-Learning-Angebot zu entwickeln, von den Mitarbeitenden zeitunabhängig und selbstständig durchgeführt werden kann. Das Ziel war es, den Mitarbeitenden das (grundlegende) Regelwerk für die Personenansetzung interaktiv zu vermitteln, sodass weniger Korrekturfälle und doppelte Ansetzungen aus den angeschlossenen Systemen heraus entstehen. Es sollten Inhalte vermittelt werden, die



Jasmin Casarano
Norddeutscher Rundfunk
Berufsbegleitender Master-
Studiengang Bibliotheks- und
Informationswissenschaft,
Technische Hochschule Köln
j.casarano@ndr.de

Aufgabenlisten	
Personen (7133)	↑
Namen ungeprüft (1)	
Namen zu bearbeiten (23)	
Personen ungeprüft (3192)	
Personen ungeprüft mit Notizen (2693)	
Personen doppelt (1224)	
Personen zu individualisieren (0)	
Künstlergruppen (1459)	↓
Sonst. Institutionen (1789)	↓

Abbildung 1 Aufgabenlisten für das NDR-Datenpflege-team (Stand 01.05.24)

das Verständnis über die **Relevanz der NDB** und das Zusammenspiel der verschiedenen Datenbanken erhöhen. Außerdem sollten **Kommunikationswege** für Anfragen an das Datenpflege-team festgelegt und dargestellt werden. Idealerweise sollte das Angebot gleichzeitig als Nachschlagewerk für schnelle Wissensauffrischungen dienen und die häufigsten Fehler in der Ansetzung aufgreifen. Durch die Vermittlung des Regelwerkes können die Dokumentar*innen auch in den angeschlossenen Systemen besser nach Namen recherchieren, wenn sie wissen, wie sie korrekt angelegt sein sollten. Zum Beispiel würde in der HFDB „Ingrid von Holt“ gefunden werden, wenn nach „Holt, Ingrid von“ gesucht wird, aber nicht bei einer Suche nach „von Holt, Ingrid“. Im zweiten Fall besteht das Risiko einer Doppelansetzung nach falschen Regeln. Dieses sollte durch das Lernangebot verringert werden.

Konzept

Nach längeren Überlegungen, auf welche Art und Weise die Inhalte am besten aufbereitet werden können, schien ein E-Learning-Kurs die beste Lösung zu sein. Da aus Kostengründen kein Web Based Training mit der ARD-ZDF-Medienakademie entwickelt werden konnte, fiel die Wahl auf eine eigenständige **Umsetzung in PowerPoint**. Falls zu einem späteren Zeitpunkt ein Auftrag für die Medienakade-

Einstieg
Relevanz der ARD-Normdatenbank
Grundlagen & Regeln
Ansetzung aus verschiedenen Systemen
Übung: Fehlersuche
Kontakt zum NDB-Datenpflege-team
Nützliche Links
Abschluss

Zurück Weiter

Abbildung 2 Kurs-Folie: Inhalt Selbstlernkurs Personenansetzung

mie erfolgen sollte, kann die PowerPoint-Präsentation dafür als Grundlage genutzt werden. Außerdem ist durch die Bekanntheit der Software eine **niedrige Technischschwelle** vorhanden. Bei der Erstellung des Konzeptes orientierte ich mich auch an schon bestehenden Kursen der Lernplattform CAMPUS, vor allem, was die Struktur betrifft. Gleichzeitig floss meine Erfahrung aus der Datenpflege und als Erfasserin mit ein, indem ich überlegte, welches Wissen notwendig ist. So ist eine Struktur entstanden, die sich in **fünf inhaltliche Kapitel** aufteilt und gerahmt ist von einem **Einstieg** und einem **Abschluss** (siehe Abbildung 2).

Diese Konstruktion kann einfach erweitert werden, z.B. wenn noch ein Abschnitt über die Ansetzung von Institutionen oder ein Kapitel zur Datenpflege hinzugefügt werden soll. Aktuell umfasst der fertige Kurs 65 Folien, die sich in die Abschnitte Einführung (5), Relevanz der ARD-Normdatenbank (3), Grundlagen & Regeln (31), Ansetzung auf verschiedenen Systemen (13), Übung: Fehlersuche (5), Kontakt zum NDB-Datenpflege-team (3), Nützliche Links (1) und Abschluss (1) aufteilen. Dazu sind zwei Tests in Microsoft Forms entstanden.

Didaktische Aufbereitung

Die Auswahl der Inhalte erfolgte auf Grundlage der Regelwerksdokumente und der Dokumentations- und Datenpflegepraxis im NDR. Die Kapitel basieren auf der Struktur vorheriger Workshops (die an typischen Ansetzungsfehlern ausgerichtet waren), der Regelwerke, sowie einem Schulungskonzept der

Didaktische Reduktion: Blumenstrauß-Methode

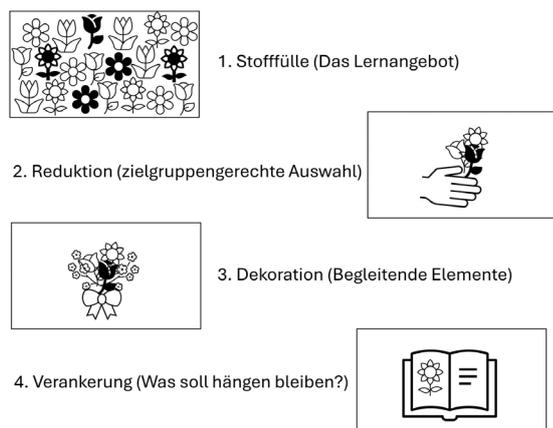


Abbildung 3 Blumenstrauß-Methode der didaktischen Reduktion. Abbildung in Anlehnung an Wüst (2022, S. 58)

AG REM. Für die Füllung der Kapitel bin ich nach dieser Reihenfolge vorgegangen: Sichtung der dazugehörigen Kapitel im Regelwerk, Extraktion der wichtigen Inhalte, Didaktische Reduktion, Aufbereitung als PowerPoint Folie. Als Literaturgrundlage diente das Mini-Handbuch Didaktische Reduktion von Yvo Wüst (2022). Durch dieses Handbuch habe ich gelernt, worauf ich bei der Reduktion des Regelwerkes achten sollte, und bin angelehnt an die **Blumenstrauß-Methode** vorgegangen (siehe Abbildung 3).

Inhalte

Der Kurs startet nach der Anleitung mit einer **Wissensabfrage** in Microsoft Forms. Die Ergebnisse sollen den Teilnehmenden Orientierung bieten, wo ihr Lernbedarf liegt. Nach dem Kapitel zur Relevanz bilden die Grundlagen & Regeln den größten Teil des Kurses. Hier werden wichtige Kenntnisse zur Normdatenbank vermittelt und das Regelwerk zur Personenansetzung in verkürzter Form dargestellt. Der Kurs ist **individuell durchführbar**, da die Zielgruppe heterogen ist. Einige Kolleg*innen kennen sich schon gut mit der Normdatenbank aus, andere weniger, manche legen Personen nur aus der HFDB heraus an, manche aus anderen Datenbanken. Um dem Lernbedarf jeder dieser Personen gerecht zu werden, ist anwählbar, welche Inhalte sie sich anschauen möchten. Das betrifft vor allem das Kapitel „Ansetzung aus verschiedenen Systemen“ (siehe Abbildung 4).

Die Person kann sich über die Buttons durch die Präsentation klicken, wie sie es für den eigenen Lernweg benötigt. Dadurch, dass der Kurs direkt in

Zusammenfassung:

Neuanlage einer Person

Name*

Vorname

Zusatz

Geschlecht

Geburtsdatum

Sterbedatum

Beruf/Funktion

Notizen

OK Abbrechen

Familienname der Person

Vorname der Person + Namenspräfix (je nach Regel)

Namenszusatz (z.B. sr. / jr., nicht Titel)

Geschlecht der Person

Geburtsdatum der Person

Sterbedatum der Person

Beruf der Person ohne Zusatzinformationen (wenn das Geschlecht bekannt ist, dann angepasst), mehrere Berufe mit Komma trennen

Weitere Informationen, die zur Identifizierung der Person relevant sind, z.B. Titel

! Für **Namensdatensätze** nur die Felder Name, Vorname, Zusatz und Geschlecht ausfüllen (maximal). Für **Personendatensätze** auch weitere Felder. Wann ein Name oder eine Person angelegt werden soll, findest du im Kapitel „Grundlagen und Regeln“ im Abschnitt „Normierungsstatus“.

Abbildung 4 Kurs-Folie: Zusammenfassung der Ansetzung aus der HFDB.

einer Bildschirmpräsentation startet (es ist eine .ppsx-Datei) und die Navigation über die Buttons erfolgt, ist eine **intuitive Bedienbarkeit** gegeben. Er kann jederzeit unterbrochen und später fortgesetzt werden, sodass er gut in den Alltag integrierbar ist. Die **Dauer der Durchführung** des Kurses insgesamt liegt bei etwa 50 Minuten.

Im Laufe des Kurses sind immer wieder **Interaktionen** (Klicke auf die Felder etc.) und **Denkaufgaben** eingebaut, um die Aktivität und Konzentration der Teilnehmenden zu unterstützen. Außerdem wurden **Screencasts** aufgenommen und Videos eingebettet. Wegen administrativer Einschränkungen war es nicht möglich Makros in PowerPoint zu aktivieren (die Lückentexte oder Drag and Drop Aufgaben ermöglicht hätten). Deshalb beschränkt sich die Interaktivität darauf, mit bestimmten Triggern (Klicks) Animationen auslösen zu lassen und Verlinkungen zwischen den Folien herzustellen. Für die „**Fehlersuche**“ wurden Datensätze aus der NDB ausgewählt, die besonders häufige Ansetzungsfehler enthalten (siehe Abbildung 5).

Die **Farbgebung** ist in Anlehnung an das Corporate Designs gewählt. Im Kurs werden daher hauptsächlich Blautöne, Weiß und Schwarz verwendet. Für einzelne Elemente und interaktive Felder werden auch andere, kräftige Farben genutzt. So kann der Fokus bei den Inhalten bleiben und die Aufmerksamkeit auf interaktive Elemente gelenkt werden. Außerdem habe ich darauf geachtet, dass wiederkehrende Elemente (Fragen, Abspiel-Symbole,

Fehler gefunden!
Die Titel „Prof. Dr.“ sind kein Namenszusatz, sondern gehören in das Feld „Akademischer Titel“ (oder bei der Ansetzung aus Fremdsystemen ins Feld „Notiz“).

Fehler gefunden!
Der Beruf „Bioinformatiker“ ist hier korrekt, jedoch sollte der Zusatz „an der Universität Düsseldorf“ im Steckbrief stehen (oder aus Fremdsystemen als Notiz erfasst werden).

Name	Vorname	Namenszusatz	Namensart	Bemerkung Name	Gültig von	Gültig bis	NameID
Marschall	Tobias	Prof. Dr.	Hauptname (Name)				451813
Akademischer Titel				Andere Titel			
Geschlecht	männlich		Instrument/Sti...				
Beruf/Funktion	Bioinformatiker an der Universität Düsseldorf						
Steckbrief							

Hier ist alles in Ordnung.

Fehler insgesamt: 2

Zurück

Weiter

Abbildung 5 Kurs-Folie: Fehlersuche.

Denkaufgaben, Hinweise) in der gleichen Weise gestaltet sind, damit diese Elemente direkt für die Lernenden erkennbar sind. Außerdem werden in den Übersichten die Felder grün gefärbt, auf die schon geklickt wurde.

Die Tests am Anfang und Ende umfassen sieben inhaltliche Fragen, die sich auf das Regelwerk und die NDB beziehen. Eine Frage erfasst, wie sicher sich die Person in der Ansetzung aus den verschiedenen Systemen fühlt. Im letzten Element wird abgefragt, ob die Person weiß, wie sie das Datenpflege-Team kontaktieren kann. Diese Test-Elemente sind bei beiden Tests identisch. Der Unterschied besteht darin, dass im Abschlusstest noch drei weitere Elemente hinzugefügt sind, die die Möglichkeit für **Selbstreflexion und Feedback** anbieten. Sie sind optional auszufüllen.

Evaluation und Ausblick

Während der Entwicklung wurden mehrere **Testläufe** sowohl mit den Kolleginnen aus dem Datenpflege-Team als auch mit Laien durchgeführt, die sonst wenig mit dem Thema zu tun haben. Die Rückmeldungen daraus wurden noch 2023 im Kurs umgesetzt. Es stellte sich aber auch heraus, dass es in der Webanwendung und Teams-App von PowerPoint Probleme mit der Darstellung der Folien und der Funktion der Verlinkungen gab. Für eine alltägliche Nutzung als Nachschlagewerk wäre aber sicher hilfreich, den Kurs direkt über Teams oder im Browser

nutzen zu können. Da für dieses Problem noch keine Lösung gefunden wurde, liegt der Kurs bisher auf einem internen Laufwerk. Seit Anfang 2024 steht der Kurs allen im NDR-Archiv zur Verfügung und wird den Kolleg*innen empfohlen, die Personen ansetzen oder recherchieren müssen, ihr Wissen überprüfen oder auffrischen möchten oder die neu im Archiv sind. Bei Regeländerungen oder neuen Empfehlungen seitens der AG REM können die Folien unkompliziert und schnell überarbeitet werden, ohne z.B. auf Dritte angewiesen zu sein. Insgesamt fiel das **Feedback** von meinen Kollegen und Kolleginnen positiv aus und das Angebot wird von vielen dankend angenommen. Kleinere Anpassungen werden weiterhin in den Kurs eingebaut. Aufgrund der Relevanz für die anderen Rundfunkanstalten der ARD wird aktuell geprüft, wie und in welcher Form der Kurs auch über den NDR hinaus bereitgestellt werden kann. Vielen Dank an dieser Stelle für die Interessensbekundungen auf der vfm-Frühjahrstagung! ■■■

Literatur

Wüest, Y. (2022). Mini-Handbuch Didaktische Reduktion: Mit E-Book inside und Online-Materialien. Beltz.

„Automate the boring stuff...“

MES-Software-Anwendung für die automatisierte Formalerschließung von Manuskripten im BR

Stefan Merl

In digitalen Arbeitsumgebungen liegen auch Textessenzen und Schriftgut immer mehr als *born digitals* vor, was die Automatisierung von zugehörigen Workflows ermöglicht und erleichtert. Die Nutzung dieses Potentials zur Effizienzsteigerung von Geschäftsprozessen und der damit einhergehenden Freisetzung von zeitlich-personellen Ressourcen ist daher ein Desiderat in der dokumentarisch-archivarischen Informationsverarbeitung.

Vorbetrachtungen und Rahmenbedingungen des Projekts „MES“

Als Unternehmensarchiv ist die Bewahrung der Programm- und Unternehmensgeschichte die Hauptaufgabe des Historischen Archivs des Bayerischen Rundfunks. Der Bestandsfokus liegt auf Textessenzen wie z.B. Drehbücher, Produktions- und Sendeunterlagen, Protokolle, Korrespondenzen, Hörerpost und Verwaltungsakten. Einen Bestandteil bilden dabei auch Manuskripte von Hörfunksendungen.

Aus der Programmdirektion Kultur, Programmbereich Bayern 2, übernimmt das Historische Archiv jährlich mehrere 100 Hörfunkmanuskripte zur Archivierung. (2023 bspw. ca. 1800 Stück). Diese liegen als *born digitals* entsprechend in digitaler Form vor.

Die manuell-intellektuelle Erfassung grundlegender Metadaten (z.B. Autor, Titel, Untertitel, Seitenanzahl) im Rahmen der Formalschließung ist äußerst ressourcenintensiv.

Das Projektziel ist daher die Automatisierung des Geschäftsprozesses der Formalerschließung, d.h. die Übernahme einzelner Prozessschritte durch die Entwicklung und Implementierung einer Software-Anwendung, um so den manuell-intellektuellen Arbeitsaufwand zu minimieren, den Prozess effizienter zu gestalten und so zeitlich-personelle Ressourcen freizusetzen.

Problem- und Lösungsbeschreibung

Ausgangspunkt ist die Analyse des Ist-Zustandes des manuell-intellektuellen Erschließungsprozesses und die dabei anfallenden Arbeitsschritte mit ihren jeweiligen Prozessergebnissen, benötigten Ressourcen sowie personellen und technischen Schnittstellen (Systeme, Anwender).

Der Formalerschließungsprozess wird angestoßen durch die Übernahme und den nachfolgenden Ingest neuer Hörfunkmanuskripte in das digitale Archivsystem des Historischen Archivs. Hierbei werden die Essenzen in den zugehörigen Essenzspeicher FEDORA ingestiert und ein korrespondierender Eintrag in der Datenbank FAUST erstellt.

Beim Ingest werden bereits einige technische und formale Metadaten automatisch generiert oder erfasst. Hierzu zählen u.a. formal-Inhaltliche Metadaten (Signatur, Provenienz, Sendedatum, Sendereihentitel) und technische Metadaten (Dateiname SIP, AIP-ID).

Die weitere Nacherschließung von für die Formalerschließung benötigten Metadaten, i.e. Autor, (Manuskript-)Titel und Umfang, erfolgt jedoch manuell-intellektuell: Im entsprechenden Datenbank-eintrag wird die Nutzungskopie des Manuskripts geöffnet, die jeweiligen Metadaten einzeln aus den Dateien ausgelesen und in die korrespondierenden Metadatenfelder eingetragen. Diese Schritte müssen für jedes Manuskript wiederholt werden.

Der gesamte Nacherschließungsprozess, insbesondere hochgerechnet auf mehrere 100 Manuskripte pro Jahr, ist somit enorm zeitaufwändig und bindet personelle Ressourcen. Bei einer Bearbeitungszeit von 1 Minute pro Manuskript entspricht dies allein für die ca. 1800 im Jahr 2023 übernommenen Manuskripte einer Arbeitszeit von ca. 30 Stunden.



Stefan Merl
Mediendokumentar im Content Management bei der RTL NEWS GmbH. Zuvor wissDok beim Bayerischen Rundfunk/ Hochschule Darmstadt
Stefan.merl@rtl.de



Abbildung 1 Deckblatt Bayern 2 Hörfunkmanuskript

Jedoch bieten sich zwei Schritte des Geschäftsprozesses für eine potenzielle Automatisierung an: Das Auslesen der einzelnen Metadaten aus den Manuskripten und die Übernahme der Metadaten in das Digitale Archivsystem. Durch eine automatisierte Stapelverarbeitung können dabei gleichzeitig die anfallenden Iterationen eliminiert werden.

Anforderungsbestimmung: Ausgangsmaterial, Systemumgebung und Schnittstellen

Die digitale Verfügbarkeit und Stukturiertheit der Manuskripte bzw. der in ihnen enthaltenen Metadaten ermöglicht ein maschinelles Auslesen bzw. eine automatisierte Weiterverarbeitung:

Die Manuskripte liegen als born digitals im PDF-Format vor. Die gesuchten Metadaten sind im Text (Autor, Titel) bzw. den Dateien (Umfang/Seitenzahl) enthalten. Die (Daten-)Werte finden sich in der Regel auf dem Deckblatt bzw. der ersten Seite des jeweiligen Manuskripts und sind in der Regel mit vorangestellten textlichen Markern versehen, die den (Daten-)Elementen der Metadaten entsprechen: Beispielsweise ist der Wert Autor*in als Element durch ein vorangestelltes „Autor:“ gekennzeichnet. (Vgl. Abb. 1)

Eine hier einsetzbare Möglichkeit der informatischen Verarbeitung von Text-Korpora bieten Reguläre Ausdrücke. In der theoretischen Informatik bezeichnen Reguläre Ausdrücke (engl. regular expression, Abk.: RegEx) eine Zeichenkette, die der Beschreibung von Mengen von Zeichenketten mit Hilfe

bestimmter syntaktischer Regeln dient. Mittels RegEx als Filterkriterien können Textkorpora nach bestimmten Zeichenketten durchsucht werden, indem der Text mit dem Muster des regulären Ausdrucks abgeglichen wird. Dieser Vorgang wird „Pattern Matching“ genannt. Zur automatisierten Anwendung können RegEx in viele Programmiersprachen implementiert werden.

Auf diese Weise kann der Text des Manuskriptdeckblatts nach den Text-Markern bzw. den (Daten-)Elementen durchsucht und die diesen nachfolgenden Zeichenketten, den (Daten-)Werten, ausgegeben werden. Der Wert des Daten-Elements „Umfang“ kann mittels einfacher automatischer Zählung der jeweiligen PDF-Seiten erfolgen und ausgelesen werden.

Für die Anschlussfähigkeit der Software-Anwendung sind im Weiteren die Funktionen und Spezifikationen des digitalen Archivsystems des Historischen Archivs zu beachten. Diese setzt sich aus der Datenbankanwendung FAUST 9 und dem damit verbundenen Essenzspeicher FEDORA Repository zusammen: Diese beinhalten Funktionalitäten zum Export von Sammlungen von digital vorliegenden Essenzen im Original oder als Nutzungskopie im PDF-Format, die zur weiteren Verarbeitung vorgehalten werden können. Metadaten können über eigenhändig definierbare maschinenlesbare Import- und Export-Formate (z.B. CSV) für beliebig viele Metadaten und Datensätze gesammelt, ausgegeben und eingelesen sowie anhand eines eindeutigen Indexes zu bereits vorhandenen Datenbankeinträgen hinzugefügt werden.

Neben technisch-systemische Anforderungen (Betriebssystem, Schnittstellen) sollten bei der Entwicklung einer Software-Anwendung auch Anforderungen der Anwender*innen im Sinne der Usability beachtet werden. Ziel ist es, MES so zu entwickeln, dass es für alle Mitarbeiter*innen des Historischen Archivs innerhalb der üblichen Systemumgebung und ohne weitere technische Voraussetzungen oder informatisches Hintergrundwissen ausführbar ist.

Systementwicklung und Implementierung von MES

Als Grundlage für die maschinell-automatisierte Verarbeitung der Manuskripte bzw. für die Anschlussfähigkeit der Software-Anwendung MES dient jeweils ein Export- und Importformat für die Daten-

bank-Anwendung FAUST. In diesem werden einzelnen Zeilen Datenbankeinträgen bzw. Manuskripten und zugehörigen Feldern dieser Zeilen entsprechenden Metadatenfeldern der Datenbankeinträge zugeordnet.

Das CSV-Exportformat beinhaltet die technischen Metadaten AIP-ID – zur eindeutigen Zuordnung der nachzuerschließenden Metadaten zu den bereits existierenden Datenbankeinträgen der Manuskripte – sowie Dateiname SIP – zur eindeutigen Zuordnung der PDF-Dateien (und der aus ihnen extrahierten Metadaten) zur jeweiligen AIP-ID (und damit zum zugehörigen Datenbankeintrag).

Das CSV-Importformat für die bei der maschinellen Erschließung gesammelten Metadaten beinhaltet neben den beiden Metadaten des Exportformats die Zuordnungen für die Felder Umfang (Text), Autor/in, Untertitel in Abstimmung mit dem Output der Software-Anwendung.

Zur Entwicklung der Anwendung selbst wurde die Programmiersprache Python in der Entwicklungsumgebung PyCharm verwendet. In das Script integriert sind die Python-Module/Bibliotheken PyPDF2 zur Verarbeitung von PDF-Dateien und Openpyxl zur Verarbeitung von CSV/Excel-Dateien.

Die Funktionalität des Scripts lässt sich wie folgt beschreiben: Das Dateiverzeichnis der Anwendung wird nach einer CSV-Datei (vordefiniertes FAUST-Exportformat) durchsucht und die darin enthaltenen Metadaten der zu erschließenden Manuskripte (AIP-ID, Dateiname SIP) ausgelesen und zum spätere-

ren Abgleich vorgehalten. Das Dateiverzeichnis der Anwendung wird nach PDF-Dateien (zu erschließende Manuskripte) durchsucht und für jede gefundene PDF-Datei werden folgende Schritte ausgeführt: Auslesen der Anzahl der Seiten der PDF-Datei, Auslesen des Texts der ersten Seite, Durchsuchen des Texts mittels definierter RegEx nach Markern für Metadaten (Autor/Titel) und Auslesen der Zeichenketten, die den Markern folgen. Schließlich erfolgt die Zusammenführung der Metadaten-Werte und Metadaten-Elemente als Schlüssel-Wert-Paare und die Zuordnung über einen Vergleich des Dateinamens des durchsuchten Manuskripts mit der in der Input-Datei enthaltenen Metadaten. Die gesammelten Metadaten der bearbeiteten Manuskripte werden schließlich in Form einer CSV-Datei ausgegeben. Eine Zeile entspricht hierbei einem Manuskript/PDF-Datei bzw. zugehörigem Datenbankeintrag, die jeweiligen Felder jeweils einem Metadatum. Die Output-CSV-Datei kann im Anschluss mittels dem erstellten entsprechenden Import-Format in die Archivdatenbank FAUST eingelesen und die Datenbankeinträge um die automatisch erschlossenen Metadaten angereichert werden.

Um die Usability für die Anwender*innen zu gewährleisten, wurde PyInstaller verwendet, welches das Python-Script und ihre Abhängigkeiten in einem Paket bündelt und in eine Standalone-Anwendung (Windows-EXE) umwandelt. Dies ermöglicht es, die Anwendung ohne Installation von Python bzw. Python-Interpreter auszuführen.

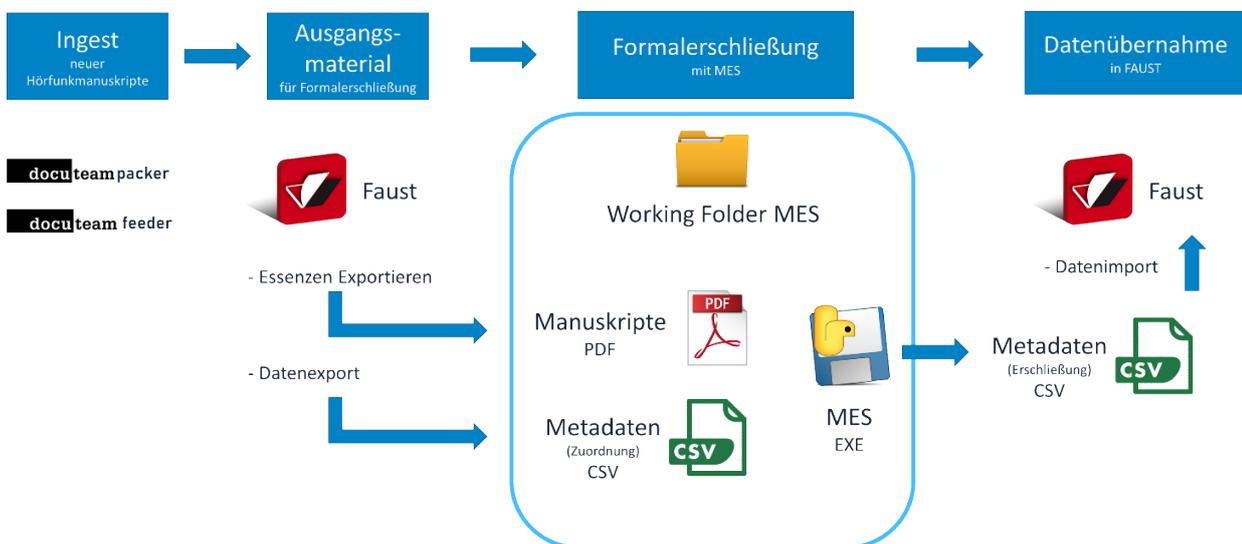


Abbildung 2 Workflow Darstellung: Formale Nacherschließung mit MES

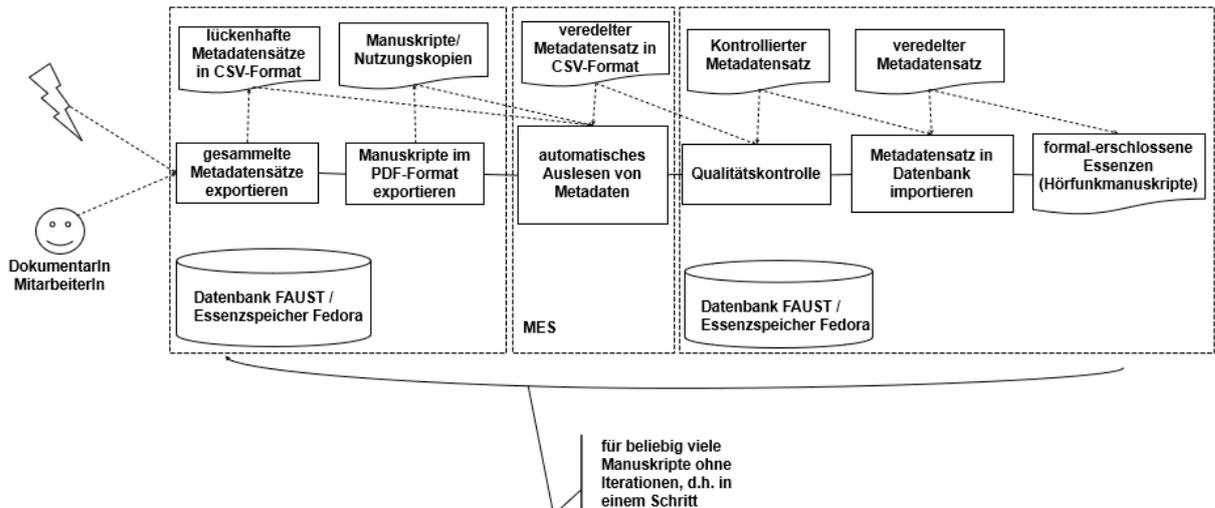


Abbildung 3 Prozessmodell: Automatisierter Formalerschließungsprozess mit MES

Des Weiteren wurde mittels TKinter ein GUI erstellt, das den Minimalanforderungen und Prinzipien der User Familiarity, Feedback und Guidance entspricht: Der Start der Anwendung öffnet ein Fenster mit Startbutton, es wird eine Startmeldung sowie eine Fortschrittsanzeige während der Ausführung sowie eine Benachrichtigung beim erfolgreichen Beenden des Scripts ausgegeben.

Die Implementierung erfolgte nach der Freigabe für das BR-Rechner-Netzwerk nach Prüfung durch die BR IT-Sicherheit (AG Software). Für die Inbetriebnahme wurde ein neues Geschäftsprozessmodell sowie eine entsprechende Workflowbeschreibung/Bedienungsanleitung (Visualisierung s. Abb.2) entworfen. Die Anwendung MES ist seit Januar 2024 im Historischen Archiv im Einsatz.

Mehrwert

In einem Testbeispiel mit 500 Manuskripten betrug die Gesamtbearbeitungszeit der Nacherschließung inkl. Wartezeiten und Qualitätssicherung ca. 15 Minuten. Eine manuell-intellektuelle Bearbeitung der gleichen Menge würde ca. 8 Stunden in Anspruch nehmen.

Best Practices und Lessons Learned

Die Automatisierung von Erschließungsprozessen ist im besonderen Maße auf die verlässliche Formalisierung und Standardisierung von Metadaten im Entstehungskontext angewiesen. Der Erfolg hängt dabei von der Kommunikation mit den jeweiligen Metadaten-Produzent*innen ab. Um die dauerhafte Produktivität von MES durch gleichbleibende Qualität der Ausgangsmaterialien zu gewährleisten, wurde daher in Absprache mit den Redaktionen ein einheitliches Deckblatt für Manuskripte eingeführt.

Als Beispiel kann MES zeigen,

-dass die Automatisierung von Geschäftsprozessen auch im Kleinen bereits zu erheblichen Mehrwerten führen kann und es sich daher lohnt, entsprechende Potentiale zu erkennen und zu nutzen,
- ... dass die Umsetzung von Automatisierungsprojekten in manchen Fällen bereits mit geringen Ressourcen machbar ist,
- ... dass Softwareentwicklung im kleinen Rahmen einen machbaren und sinnvollen Einstieg ins Programmieren bieten kann, und damit die Möglichkeit eröffnet, die Schnittstellenfunktion zwischen IT und Dokumentation/Archiv besser auszufüllen. ■

Ticket in die Zukunft

Konzeption eines Ticketsystems als Optimierung für den IDA-Desk von SWR und SR

Lena Cara Wernhöfer

Der Infodesk ist die Hauptannahmestelle für Anfragen, Rechercheaufträge und Services der HA Information, Dokumentation & Archive des SWR & des SR (IDA) und somit häufig erste Kontaktstelle für die Redaktionen. Täglich werden hier handlungsrelevante Informationen verarbeitet: Redaktionen kontaktieren den Infodesk entweder telefonisch oder über eine Funktionsmailadresse. Diese Anfragen werden von den Desk-Mitarbeitenden in Start- und Verstärkungsteam angenommen und je nach Zeitaufwand und -kapazität arbeitsteilig bearbeitet.

Der Workflow sowie die interne Kommunikation unter den Mitarbeitenden des Infodesk ist derzeit auf unterschiedliche Officeanwendungen gestützt. Durch die Konzeption eines Ticketsystems für den Infodesk in Hinblick auf interne und softwaregestützte Workflowoptimierung, soll eine Verbesserung der Arbeitsteilung, internen Kommunikation, statistischen Auswertungsmöglichkeit sowie eine geringere Fehleranfälligkeit erreicht und eine allgemeine Professionalisierung forciert werden.

Im Sommer 2024 startet die Hauptabteilung IDA mit dem ‚Ticket in die Zukunft‘ eine neue Reise in Richtung Professionalisierung und Arbeitswelt 4.0: Am IDA Desk, dem anfragenstärksten Produkt, wird ein Ticketsystem Einzug halten.

Das Projekt fußt auf einer Abschlussarbeit der wissDok-Weiterbildung an der Hochschule Darmstadt, deren Projektaufgabe die Konzeption, Evaluierung sowie Konfiguration und anschließende Einführung einer Softwarelösung zur Auftragsannahme, -verwaltung und -analyse von Serviceanfragen für den IDA Vorüberlegungen, wonach es um die konkreten Rahmenbedingungen des IDA Desks und die praktische Umsetzung der Projektphasen – Konzeption, Konfiguration und Produktivnahme – geht. Abschließend erfolgt ein Erfahrungsbericht, ergänzt durch einen Ausblick auf die kommenden Monate. (Abbildung 1)



Lena Cara Wernhöfer
IDA Information,
Dokumentation und Archive
des SWR und des SR
Südwestrundfunk
lena.wernhoefer@swr.de

TITEL	KUNDENBENUTZER-NAME	BESITZER	SERVICE	STATUS	SPERREN	ALTER	DEADLINE	EXPORTZIELE	TYP
Musik und Presse	Herr Christian Reck	Jan Seipel	Bewegtbild	offen	in Bearb.	6 d 20 h			1 - Recherche
Frohe Musik	Frau Monja Eigenschenk	Jan Seipel	Online	offen	in Bearb.	6 d 20 h		VPMS MNZ	1 - Recherche
Presse Autoindustrie	Herr Christian Reck	--	Bewegtbild	offen	frei	6 d 21 h			1 - Recherche
Musikanfrage	Herr Christian Reck	--	Presse	offen	frei	6 d 21 h			3 - Bereitstellung (Physisches Material)
Musik und Presse	Herr Christian Reck	Jan Seipel	Bewegtbild	offen	in Bearb.	6 d 22 h			1 - Recherche
Bewegtbild Bussen	Herr Christian Reck	--	Musik	offen	frei	6 d 22 h			1 - Recherche
musik background color check	Herr Jan Seipel	Jan Seipel	Online	offen	in Bearb.	15 d 21 h			1 - Recherche
test background color	Herr Jan Seipel	Jan Seipel	Presse	offen	frei	15 d 21 h			1 - Recherche
SWR-Aktuell-RP-Anfrage	Frau Michelle Habermehl	Lena Wigand-Steinmetz	Musik	offen	in Bearb.	27 d 20 h			1 - Recherche
Schäuble Bücher	Herr Thomas E. Schneider	Jan Seipel	Wort	offen	in Bearb.	106 d 18 h			3 - Bereitstellung (Physisches Material)

Abbildung 1: Screenshot des Dashboards als Landingpage des IDA-Ticketsystems ‚Anfragenpool‘

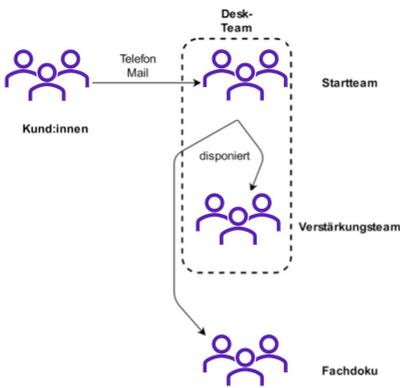


Abbildung 2: Darstellung der Arbeitsteilung am IDA Desk

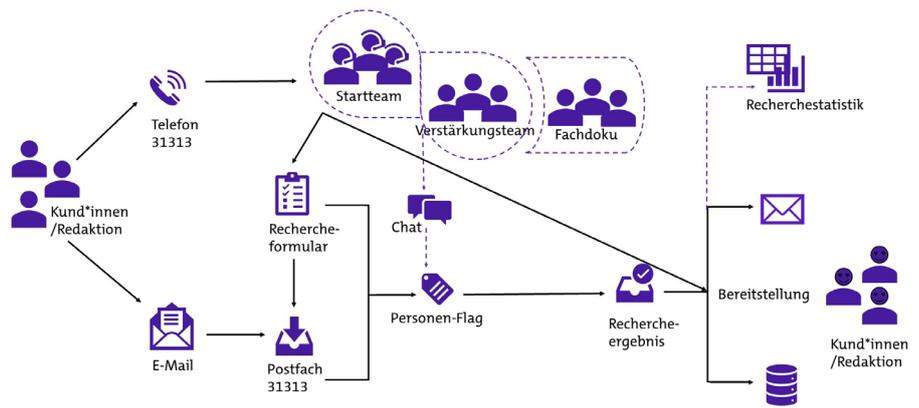


Abbildung 3: Darstellung des IST-Workflows am IDA Desk

Wissen in Aktion

„Information ist Wissen in Aktion“¹ – Das Zitat, das eigentlich der Definition des Begriffs Information als handlungsrelevant² dient, ist auf vielerlei Weise interpretier- und für die Arbeitspraxis nutzbar. So kann es beispielsweise für den Wandel des Berufsbilds der wissenschaftlichen Dokumentation im Allgemeinen stehen: Ein Berufsfeld, das sich durch den technischen und digitalen Wandel stetig in Bewegung und in komplexer Aushandlung und Neuausrichtung befindet. Im Kontext der vorliegenden Projektarbeit steht es vor allem aber als Minimaldefinition für den *IDA Desk der HA IDA des SWR und des SR*³, an dem das nachfolgend vorgestellte Ticketsystem etabliert wird. Das Projekt widmet sich der Konzeption eines Ticketsystems für den, für Redaktionen angebotenen, IDA Desk in Hinblick auf interne und softwaregestützte Workflowoptimierung, um eine Verbesserung von Arbeitsteilung, interner Kommunikation, statistischer Auswertungsmöglichkeit sowie eine geringere Fehleranfälligkeit zu erreichen und allgemeine Professionalisierung zu forcieren. An dieser Stelle schließt sich der Kreis von der informationswissenschaftlichen Theorie des Eingangszitats hin zur beruflichen Praxis der Information Specialists:

1 Kühlen & Semar, 2023,3.

2 Der hier zugrundeliegende Begriff der Information dient als Arbeitsbegriff gemäß eines Werkzeugs zum Erkenntnisgewinn der Geisteswissenschaften. Somit handelt es sich in der Informationswissenschaft zunächst um ein Konstrukt (vgl. Kühlen & Semar, 2023, 3f,22). Im Sinne der *actionable information*, zu Deutsch handlungsrelevanten Information ist er sowohl abhängig von Nutzungskontext als auch weiteren subjektiven Faktoren und entsteht „in sozialen Situationen durch Auswerten der vorhandenen Informationsobjekte“ (ebd. 22) Jene Faktoren bedingen, inwiefern die Handlungsrelevanz fremd oder frei bestimmt sei (vgl. ebd., 22).

3 *IDA Desk der HA IDA des SWR/SR* folgend synonym als IDA Desk bezeichnet.

Der IDA *Info-Desk* ist ein Ort, an dem sich *Wissen in Aktion* befindet.

Der IDA Desk

Der IDA Desk ist die Hauptannahmestelle für Anfragen, Rechercheaufträge und Services der Hauptabteilung Information, Dokumentation und Archive und somit häufig erste Kontaktstelle für die Redakteur*innen, auch benannt als Kund*innen. Täglich werden hier handlungsrelevante Informationen verarbeitet. Information, in ihren vielfachen Ausführungen, ist ein hohes Gut für Produktionsarchive, wie das des SWR/SR. Der Aufgabenbereich von Produktionsarchiven geht über die Tätigkeitsfelder klassischer Archive (wie Erhalten und Erschließen) hinaus. Während die Tätigkeit der Archivierung die „Überlieferungsbildung, Erschließung, Verwahrung, Erhaltung und Bereitstellung von Archivalien“⁴ umfasst, steht im Selbstverständnis von Medien- und Produktionsarchiven nicht ausschließlich dauerhafte Bewahrung im Vordergrund, sondern ebenso der Aspekt der Wiederverwendung von physischem und digitalen Material innerhalb des Senderprogramms⁵. Der IDA Desk dient folglich der Aufgabe, das Archivmaterial in Form von Information zweckgebunden und ortsunabhängig ‚in Verkehr zu bringen‘. Zu diesem Zweck bedarf es einer Informationsinfrastruktur. (Abbildung 2)

Innerhalb der Hauptabteilung ist seit 2020 eine agile Organisationsstruktur und Arbeitskultur etabliert. Das bedeutet, die HA ist in vier Geschäftsbereiche untergliedert, innerhalb derer die Mitarbeitenden in selbstständig zusammengesetzten Teams zu-

4 Schwarz 2023, 93.

5 Vgl. ebd.

sammenarbeiten. Die verschiedenen Dienstleistungsprodukte sind innerhalb des Geschäftsbereichs Service angesiedelt, hier befindet sich auch das *Team Desk*. Darüber hinaus dominiert seit der Zeit der Pandemie eine hybride Arbeitsweise, also unterschiedlich abgestufte Mischformen der Arbeit in Büro, im Home Office sowie in der Mobilien Arbeit.

Der IDA Desk ist die zentrale und standortübergreifende⁶ Anlaufstelle für Anfragen und Recherchen aller Art. Bei den Anfragen handelt es sich um die Bereitstellung von Archivgut, die Beratung zu Services sowie Rechercheaufträge aus den Bereichen Bewegtbild, Hörfunk, Musik, Wort, Geräusch sowie Presse und Online (also bspw. auch Social Media), das heißt Material und Information aus in- und externen Datenbanken sowie dem Internet. Das Team IDA Desk bearbeitet im Jahr etwa 20.000 Anfragen, was 1700 Anfragen im Monat oder 80 am Tag entspricht. Das Auftragsaufkommen wird in wechselnden Schichten von einem festen Team⁷ bewältigt.

Der Workflow

Innerhalb des Workflows begleiten die Mitarbeitenden des IDA Desk den Prozess von Informationsbedarfsermittlung über Informationsrecherche bis zur Informationsbeschaffung und -vermittlung, jeweils inklusive komplexer Unterprozesse (z.B. Qualitätssicherung, Quellenbewertung)⁸. Um den Prozess der Informationsverarbeitung angemessen, proaktiv und flexibel im Sinne der agilen Struktur der Hauptabteilung zu begleiten und damit die redaktionellen Kund*innen mit dem existierenden Wissen zu versorgen, muss mit zeit- und fachgemäßen Werkzeugen in Form von Software sowie Workflows gearbeitet werden. Hier befindet sich die HA in einem stetigen Prozess der Professionalisierung.

Dies bedeutet konkret: Redaktionen kontaktieren den IDA Desk entweder telefonisch oder über eine Funktionsmailadresse. Diese Anfragen werden von den Desk-Mitarbeitenden in Start- und Verstärkungsteam angenommen und je nach Zeitaufwand

und -kapazität arbeitsteilig bearbeitet. Die Aufträge werden je nach Dringlichkeit und Arbeitsaufwand vom aktiven Desk-Team, das aus Startteam (ST) oder Verstärkungsteam (VT) besteht, in Echtzeit oder der ausgelagerten Fachdokumentation (FD) im Nachgang bearbeitet. Der Workflow sowie die interne Kommunikation unter den Mitarbeitenden des IDA Desk ist derzeit auf unterschiedliche Officeanwendungen gestützt. (Abbildung 3)

Der gegenwärtige IST-Workflow sieht wie folgt aus: Das ST nimmt eine Telefonanfrage auf, bearbeitet diese entweder direkt oder verschriftlicht sie als Rechercheformular, welches dann an das zentrale Postfach gesendet wird. Dort treffen dann die E-Mail-Anfragen der Redaktionen und die Rechercheformulare in MS Outlook zusammen und die E-Mails werden je nach Medium geflaggt. Anschließend kann dann die Aufgabenverteilung über Personenflags und Chat-Absprachen in MS Teams erfolgen. Entscheidend ist hier, ob es sich um eine schnell zu bearbeitende Anfrage (ST), eine etwas umfangreichere Recherche (VT mit Schwerpunkt Presse, Bewegtbild, Musik o.Ä.) oder eine zeitaufwändige, eventuell mehrtägige Recherche (FD) handelt. Darüber hinaus gibt es Anfragen, die direkt vom ST an weitere Teams der HA IDA weitergeleitet werden (Materialbereitstellung/-einspielung, Programmaustausch, Notenarchiv, Musikbeschaffung, Buchbeschaffung u. v. m.). Ist eine Recherche abgeschlossen, bekommt der*die Kund*in eine E-Mail mit dem Rechercheergebnis, während je nach Medium parallel Essenzen oder Merklisten in den Datenbankstrukturen bereitgestellt (z.B. VPMS) werden. Abschließend wird dann von den Bearbeitenden händisch die Recherchestatistik in MS Excel gepflegt. Bei dieser Erläuterung des IST-Zustands werden drei Punkte augenfällig, die illustrieren, dass der momentane Workflow nicht ideal ist: Erstens sind Telefonanfragen nicht immer dokumentiert, da manche Kolleg*innen manuell Notizzettel schreiben. Zweitens ist die Verwendung von Personenflags fehleranfällig, zum Beispiel bei Namendopplungen oder weil sich Zuständigkeiten ändern, ohne dass Flags dabei angepasst werden. Drittens kann die Pflege der Statistik am Ende des Prozesses leicht vergessen werden. Der Hauptpunkt ist aber, dass der aktuelle Workflow auf verschiedene Programm Lösungen (also von Outlook über Teams bis zu Excel) gestützt ist. An dieser Stelle ergeben sich die Optimierungsbedarfe.

⁶ Seit der zentrale Desk 2020 als Remote-Lösung eingeführt wurde, arbeitet das Team standortübergreifend sowohl an den Hauptstandorten des SWR und SR (Mainz, Baden-Baden, Stuttgart und Saarbrücken) sowie aus verschiedenen Regionalstudios zusammen. Zuvor hatte jeder Standort, seit 2005, eine eigene Desk-Variante.

⁷ Insgesamt sind circa 75 IDA-Mitarbeitende und damit ein Großteil der Hauptabteilung unterschiedlich stark am IDA Desk beteiligt.

⁸ Vgl. Seidler-de Alwis, 2023, 477ff.



Abbildung 4: Übersicht der Optimierungsbedarfe am IDA Desk

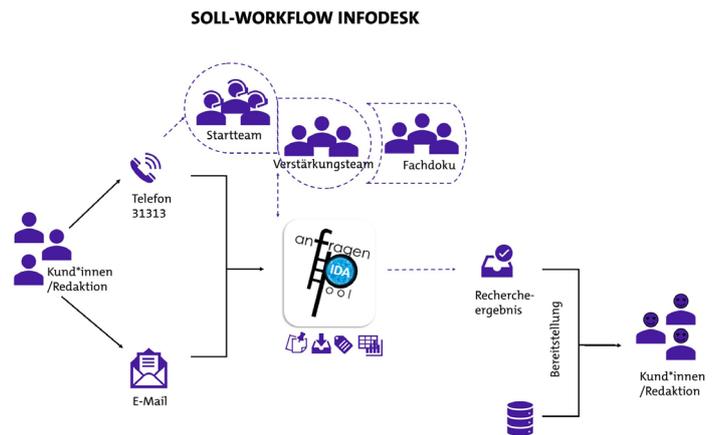


Abbildung 5: Darstellung des künftigen SOLL-Workflows am IDA Desk

Optimierungsbedarfe

Im Laufe der Vorüberlegungen war schnell klar, dass statt der Arbeit mit verschiedenen Programmen eine One-Hand-Lösung in Form eines Ticketsystems gewünscht ist. Mithilfe dieser Softwarelösung sollen, anhand aufgeführter Komponenten, nachfolgende Optimierungen erreicht werden:

1. Eine vereinfachte und eindeutige Verteilung von Zuständigkeiten durch ein Feld ‚Bearbeiter*in‘ mit API/Schnittstelle zu MS Identity und der ARD-B-Kennung.
2. Eine effektivere interne Kommunikation durch eine Kommentar-/Dialogfunktion am betreffenden Ticket und eindeutige Ticket-ID.
3. Ein einheitliches Template für alle Anfragen, was auch unter dem Aspekt Metadaten interessant werden kann.
4. Eine effektivere externe Kommunikation durch Nachvollziehbarkeit des Bearbeitungsstands anhand von Statusmeldung an Kund*innen.
5. Eine strukturierte Archivierung und Durchsuchbarkeit abgeschlossener Aufträge, unter anderem zur potenziellen Wiederverwertung von Rechercheergebnissen bei häufig oder wiederholt gestellten Anfragen.
6. Eine statistische Erhebung und Auswertung, die ohne Mehraufwand für die Mitarbeitenden ins System integriert ist, soll eine lückenlose Pflege gewährleisten.

Znuny – das Ticketsystem

Aber, was ist ein Ticketsystem überhaupt und wie wird es eingesetzt? Ein Ticketsystem ist eine Software, die als Werkzeug zur Verwaltung von Anfragen dient. Verschiedene Arten von Anfragen, wie Kunden- und Mitarbeiteranfragen oder -aufträge,

werden jeweils zu einem Ticket transferiert und können somit strukturiert und formatiert ver- und bearbeitet werden. Die Tickets werden um Informationen (Metadaten) ergänzt, mit einer ID und einem Status versehen, wodurch Zuständigkeiten klar verteilt werden können.⁹ Im Bereich des IT-Supports werden Ticketsysteme bereits seit Jahrzehnten verwendet, die jeweiligen Softwarelösungen sind gegenwärtig dementsprechend zunehmend hochindividuell.¹⁰

Znuny ist eine Open-Source Ticketing Software, die als vormals OTRS (Open Ticket Request System) bekannt ist. Die Software dient als Helpdesk- und Issue-Tracking-System dazu, sowohl Anfragen zu verwalten als auch mit Kund*innen darüber zu kommunizieren. Zudem verfügt Znuny über eine deutsche Benutzeroberfläche. Neben der Abdeckung vieler der zuvor genannten Anforderungen bietet es den Vorteil, bereits hausintern als IT-Ticketsystem in Verwendung zu sein, im Gegensatz zu einer Eigenentwicklung sind die Prüfungen verschiedener Gremien von Datenschutz bis zu Barrierearmut folglich bereits abgeschlossen.

Mit Einführung des Znuny-basierten Ticketsystems – welches IDA-spezifisch **Anfragenpool** getauft wurde – wird der IST-Workflow um einiges verschlankt und ist nicht mehr auf verschiedene Programme gestützt. E-Mailanfragen treffen automatisiert als Tickets im Anfragenpool ein, Telefonanfragen werden vom Startteam als Telefontickets aufgenommen und auch die Bereitstellung der Er-

⁹ Vgl. <https://otrs.com/de/otrsdiag/woher-kommt-der-be-griff-ticketsystem-und-passt-er-heute-ueberhaupt-noch>

¹⁰ Der Znuny-Vorläufer OTRS wurde 2001 erfunden Vgl.

https://doc.znuny.org/znuny_lts/about/index.html

<https://otrs.com/de/use-cases/solution-desk>

<https://otrs.com/de/otrsdiag/it-service-desk>

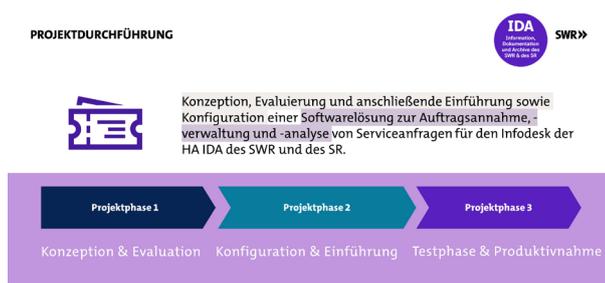


Abbildung 6: Projektverlauf, dargestellt anhand der drei Projektphasen

gebnisse läuft aus dem Anfragenpool heraus und kommt als E-Mail bei den Kund*innen an. Das bedeutet: das Postfach ist integriert, E-Mails werden aus dem Anfragenpool generiert und versendet. Gleichzeitig entstehen durch die Telefonticket-Maske keine manuellen Gesprächsnotizen mehr und die Rechercheformulare werden ersetzt. Die Personenzuordnung ist durch eine Schnittstelle zu MS Identity und die B-Kennung eindeutig, während Informationen wie Kontaktdaten automatisch mitgegeben werden. Zeitgleich werden Informationen, wie der Name der Redaktion, welche für die Statistik benötigt werden, mitgeliefert und die statistische Erhebung von Daten passiert automatisch, gleichzeitig und ohne zusätzlichen Aufwand für die Bearbeiter*innen.

Der Projektverlauf

Um den Projektverlauf nachvollziehbar aufzuschlüsseln, folgt ein Einblick in die bisherige Vorgehensweise sowie ein abschließender Ausblick. Das Projektteam ist zusammengesetzt aus dem Kernteam, bestehend aus drei Mitarbeitenden des Team Desk, sowie einem Soundingboard, bestehend aus fünf Mitarbeitenden. Letzteres wurde im Projektverlauf kontinuierlich in rekursiven Feedbackschleifen einbezogen. Das Projekt ist in drei große Projektphasen gegliedert: Projektphase eins umfasst die **Konzeption und Evaluation**, Projektphase zwei beinhaltet die **Konfiguration und Planung der Einführung** und Projektphase drei besteht aus **Testphase und Produktivnahme**.¹¹ (Abbildung 6)

In **Projektphase eins** wurde zunächst ein detaillierter Anforderungskatalog mit User Stories entworfen. Dieser diente einerseits dazu, dem Team

einen eingehenden Überblick zu den Anforderungen zu verschaffen und andererseits den Softwareexpert*innen vor den Gesprächen eine klare Vorstellung der Anwendungsfälle zu geben, um die Kommunikation zu erleichtern. Insgesamt wurden im Rahmen der Marksondierung fünf (inhouse) Softwarelösungen miteinander verglichen¹². Um eine annähernde Objektivität zu gewährleisten, wurden die Mindest- und Minimalanforderungen sowie Optimierungswünsche aus dem Anforderungskatalog in eine Evaluationsmatrix überführt. Da die Anforderungen an ein Ticketsystem von unterschiedlicher Priorität sowie unterschiedlich quantifizier- und messbar sind, wurde im Anschluss an die jeweiligen Expert*innengespräche mithilfe einer abgewandelten Form der *MoSCoW-Methode*¹³ gewichtet. Diese Bewertung geschah in Orientierung am aktuellen Workflow als ‚Nullpunkt‘ – das bedeutet, unter Bezug auf die Fragestellung: wäre diese Lösung besser oder schlechter als der aktuelle Workflow, oder würde sie vielleicht keine Änderung mitbringen? Das Ergebnis der Evaluation von Softwarelösungen ist das Ergebnis der Projektphase eins: Keines der betrachteten Programme deckt zu 100% sämtliche, in der Evaluationsmatrix festgelegte, Anforderungen und Wünsche des Projektteams ab. Mit Zguny hat sich jedoch ein Favorit herausgebildet, der die meisten Anforderungen erfüllt und damit zur abschließenden Softwareempfehlung für den Einsatz am IDA Desk wurde. Die Empfehlung wurde dem strategischen Entscheidungsgremium der Hauptabteilung IDA Gesamt vorgestellt, in dem der Vorschlag abgestimmt und angenommen wurde. (Abbildung 7-9)

Anschließend konnte mit einer eigenen Zguny-Testumgebung (als eigene Instanz) in **Projektphase zwei**, Konfiguration und Planung der Einführung, gestartet werden. Bei der Konfiguration der Grundlagen ging es zunächst um Grundsätzliches, wie die Namensfindung, die für die URLs für

¹² Parallel zur Klärung der Anforderung und Bewertung wurde sich der Entwicklung eines Prototyps für eine Eigenentwicklung gewidmet, wofür zunächst innerhalb eines Workshop-tags Mockups erstellt wurden. Der Prototyp einer Eigenentwicklung wurde basierend auf Python und Django programmiert, die Entwicklung wurde aber während der Evaluationsphase eingestellt.

¹³ MoSCoW steht für eine Einteilung in Must, Should, Could und Won't. Da im vorliegenden Fall selbstredend keine Anforderungen aufgenommen wurden, die in die Kategorie ‚Won't‘ fallen würden, wurde Won't durch Whould ersetzt. Orientiert an den vier Kategorien wurde dann zwischen 4-1 gewichtet.

¹¹ Das Projekt umfasste im Rahmen der wissDok Weiterbildung ursprünglich nur Phase 1. Hier konnten schnell so gute Fortschritte erzielt werden, dass es um Projektphase 2 (also Konfiguration & Einführung) erweitert wurde.



Abbildung 7:
Logo des
IDA Anfragespools

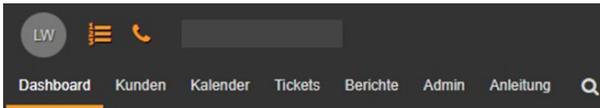


Abbildung 8: Logo im Header der Landingpage des Anfragespools

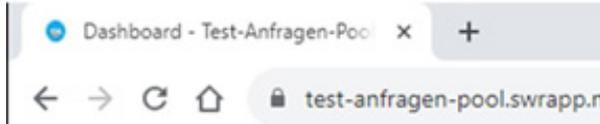


Abbildung 9: Fav-Icon der Anwendung

Produktiv- und Testsystem relevant ist. Darüber hinaus wurde die IDA-Umgebung des Systems, der sogenannte Anfragespool, mit Logo und Fav-Icon ausgestattet¹⁴. Zeitgleich wurde sich mit den Rollenbeschreibungen für verschiedene User auseinandergesetzt und die wichtigsten Seiten und Datenfelder definiert. Beruhend auf den Definitionen ist dann das fertige Tickettemplate, die Formularmaske für Telefontickets, entstanden. Insofern eine Redaktion den IDA Desk telefonisch kontaktiert, kann der* die Bearbeiter*in im ST direkt in einem Drop-down auswählen, um welche Form von Anfrage (Tickettyp) es sich handelt; Recherche, Bereitstellung von digitalem oder physischem Material, Service/Auskunft oder Störung des Anfragespools. Innerhalb des Felds Service werden dann die verschiedenen Medien ausgewählt, um die Zuständigkeiten zu verteilen (z.B. Bewegtbild, Wort, Musik, Presse, etc.)¹⁵. Das Ticket kann zusätzlich priorisiert und terminiert oder bereits während der Annahme direkt einem*r Bearbeiter*in zugeordnet werden. Nach Definition und Konfiguration wurde das Template daraufhin getestet, ob es alle Eventualitäten für Vorkommnisse am IDA Desk abbildet. Hierzu haben zwei Kolleginnen ihre Desk-Schichten mit dem Template begleitet. Um alle weiteren Funktionen vor der potenziellen Produktivnahme hinreichend zu testen und zu evaluieren, wurde eine dreistufige Testphase entworfen. Mit Abschluss der Grundlagenkonfiguration, in enger Zusammenarbeit mit der IMP¹⁶, und dem Entwurf der Testung wurde dann Projektphase zwei abgeschlossen. (Abbildung 10-12)

Mit Beginn des Jahres 2024 sind Kernteam und Soundingboard in die Testphase und somit **Projektphase drei** gestartet. Diese Testphase ist wie-

derum in drei Testzyklen unterteilt, mit deren Verlauf sich die Schwerpunkte zunehmend konkretisieren: Zunächst ging es um das Testen der einfachen Funktionen des Anfragespools, dann die Simulation einer Desk-Schicht und abschließend die Begleitung einer realen Desk-Schicht. Der jeweils nächste Testzyklus beginnt erst, sobald der vorherige abgeschlossen ist. An dieser Stelle ist wichtig zu erwähnen, dass die Konfiguration des Systems weiterhin parallel zur Testphase läuft, also dabei wie in einem Dev-Op-Prozess (Development Operations) **nicht** als **abgeschlossen** gilt. Das Feedback des Soundingboards wird dementsprechend parallel eingearbeitet, um eine für die Mitarbeitenden angenehme GUI (engl. graphical user interface, also zu deutsch grafische Benutzeroberfläche) zu erreichen und die User Experience zu berücksichtigen. Beispielsweise konnte infolgedessen das Dashboard des Anfragespools mithilfe eines Browser-Addons intuitiver gestaltet, durch Symbole und Farben ergänzt sowie insgesamt verschlankt werden.

Innerhalb des dritten Testzyklus hat ein sechsköpfiges Team, bestehend aus Kernteam und Soundingboard, zunächst eine zweistündige und zuletzt eine gesamte Desk-Schicht mit dem Anfragespool übernommen. Zudem wurde ein ‚Cheat-Sheet‘, also ein Merkblatt mit den wichtigsten Funktionen und Ersteinrichtungsschritten, für die Mitarbeitenden erstellt, welches sowohl im IDA-Wiki zu finden als auch auf der Landingpage des Anfragespools verlinkt ist.

Die konkreten nächsten Schritte bestehen aus der Testung des E-Mailversands, der Erstellung von Schulungsmaterial sowie der Planung der Schulungen. Die Produktivnahme wird dann in der ‚Sommerpause‘ Ende Juli erfolgen, ein Zeitraum, in dem es erfahrungsgemäß etwas ruhiger am Desk ist. Auf diese Art und Weise haben die Mitarbeitenden mehr Raum, sich an das neue Tool zu gewöhnen und eventuelle Ungereimtheiten würden weniger ins Gewicht fallen und daher den Arbeitsablauf nicht behindern.

Sofern sich der Anfragespool dann am IDA Desk erfolgreich etabliert hat, besteht die Möglichkeit, die Software auch für die Verwaltung anderer Anfragen und Dienstleistungen innerhalb der Hauptabteilung zu verwenden. Beispiel hierfür wäre die Verwendung in Teams, die ohnehin bereits intern mit dem Desk kommunizieren, wie der Programmaustausch, das Notenarchiv oder das Unternehmensarchiv – diesbezüglich gibt es bereits erste Gespräche.

¹⁴ Dies dient unter anderem dazu, damit es schon früh auf der Landingpage zu keinen Verwechslungen kommen konnte.

¹⁵ Damit es im Arbeitsalltag parallel zu einem Anruf schnell gehen kann, gibt es neben dem Dropdown auch weitere autofill Möglichkeiten.

¹⁶ Hauptabteilung IT, Medien- und Produktionstechnik.

Des Weiteren ist der ARD-weite IT-Service im Januar ebenso auf Znuny umgestiegen. Daher wäre es als Entwicklungsperspektive theoretisch technisch möglich, den Anfragenpool (bzw. ein Ticketsystemen wie dieses) für die Anfragen- und Auftragsverwaltung in den standort- und medienübergreifenden Recherchedesks ARD-weit einzusetzen.

Learnings & Best Practices

Im Folgenden werden die wichtigsten Erkenntnisse und bewährten Methoden vorgestellt, die im Projektverlauf identifiziert wurden und durch ihre Allgemeingültigkeit in Form von Learnings und Best Practices auch für zukünftige Projekte von Nutzen sein können. An diesem Punkt lassen sich drei Herausforderungen hervorheben, mit denen das Projektteam konfrontiert war: Zunächst die Abhängigkeiten von anderen Stellen innerhalb und außerhalb des SWR, wie der IMP oder Znuny. Insbesondere bei den technischen Komponenten, die nicht durch das Projektteam selbst realisiert werden konnten (wie die Einrichtung einer Firewall oder das Klonen der Testumgebung zu einem Produktivsystem), kam es zu Verzögerungen. Des Weiteren musste für die Expert*innengespräche (in Projektphase eins) und die Konfiguration in Kooperation mit der IMP (in Projektphase zwei), in kurzer Zeit ein gemeinsames Vokabular abgestimmt werden, um sicherzugehen, über dieselben Anforderungen zu sprechen. Denn – dies führt zum dritten Punkt – der Anfragenpool als Ticketsystem für den IDA-Desk unterscheidet sich von herkömmlichen Ticketsystemen in wichtigen Details, da es um die Verwaltung von Informations- und Recherche-Serviceleistungen geht und nicht um EDV-Support.

Aus der erfolgreichen Bewältigung dieser Herausforderungen resultierten wertvolle Erkenntnisse, die im Folgenden in Form von Learnings erläutert, aufzeigen, was zukünftig vermieden werden sollte. Punkt eins scheint zunächst offensichtlich: Missverständnisse in der Kommunikation sollten vermieden werden. Hierzu zählt es, den Flurfunk nicht zu unterschätzen und Anforderungen immer klar zu definieren, insbesondere in der Kommunikation mit abteilungsfremden Teams. Punkt zwei betont, dass man dem menschlichen Irrtum entgegenwirken sollte, das eigene Produkt aufgrund der bereits investierten Arbeit und Energie für das Beste zu halten („IKEA-Effekt“¹⁷). Bei der Entwicklung neuer

REENSHOT – ANFRAGENPOOL, DAS TOOL

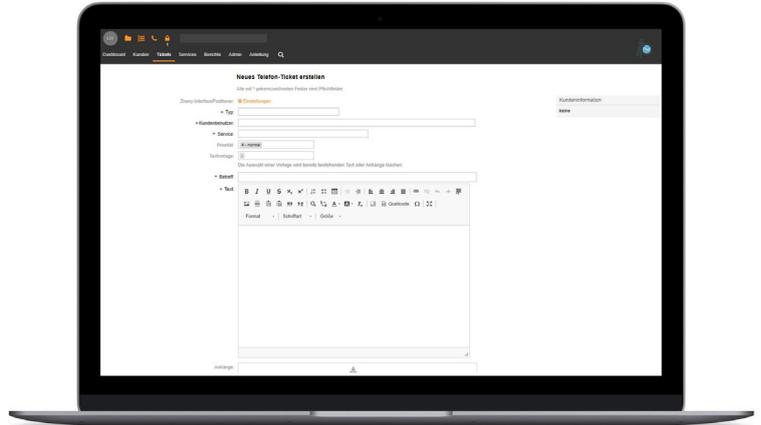


Abbildung 10: Telefonticket Formularmaske des Anfragenpools



Abbildung 11: Detailansicht des Telefontickets – Tickettyp

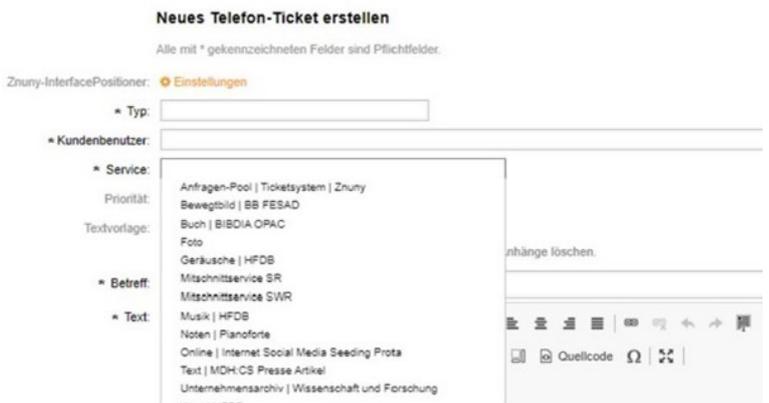


Abbildung 12: Detailansicht des Telefontickets – Service

Software, Workflows oder Prozesse ist es entscheidend, immer wieder objektiv gegenzuprüfen, ob und wie der Weg fortgeführt werden sollte (wie es im vorliegenden Projekt mit der Einstellung des

(<https://www.sueddeutsche.de/wissen/psychologie-ikea-effekt-selbstgebaut-moebel-1.5350871>) berufend auf die Studie von 2011 der Psychologen um Michael Norton und Dan Ariely, 'The IKEA effect: When labor leads to love' (<https://myscp.onlinelibrary.wiley.com/doi/abs/10.1016/j.jcps.2011.08.002>)

¹⁷ „Wer Mühe in etwas steckt, schätzt das Ergebnis.“

Prototyps einer Eigenentwicklung war). Punkt drei bezieht sich ebenso auf die Entwicklung neuer Workflows. Da Sonderfälle oft erst auffallen, wenn sie auftreten, ist es auch hier entscheidend, den Blick immer möglichst frisch zu halten. Punkt vier beschreibt ein häufig auftretendes Problem in der Projektplanung. Es sollte von Beginn an immer genügend Puffer eingeplant werden, da Unvorhergesehenes eintreten könnte, beispielsweise in der Personalplanung.

Des Weiteren können aus den Erfahrungen auch einige Best Practices abgeleitet werden, die als universelle Empfehlungen dienen können. Hier hat sich das Einbeziehen eines Soundingboards in den gesamten Prozess des Projektverlaufs als sehr konstruktiv erwiesen. Zum einen führen engmaschige Feedbackschleifen zur Vermeidung von eventuellen Fehleinschätzungen, zum anderen fördert es bei grundlegenden Umstrukturierungen oder Neuerungen in großen Teams die Akzeptanz, wenn ebendieses aktiv am Projekt und den Entscheidungen beteiligt ist. Darauf zählt auch die Transparenz gegenüber den Betroffenen ein, um Fragen zu klären oder potenziellen Bedenken zuvorzukommen und dafür durchgehend Zeit einzuplanen. Abschließend ist es gewinnbringend, das Projektteam möglichst interdisziplinär und divers aufzustellen (beispielsweise bezogen auf Altersstruktur, Geschlechter usw.), um Expertise und Erfahrung aus verschiedenen Schwerpunkten zu bündeln und von Anfang an viele Perspektiven und Eventualitäten zu berücksichtigen.

Soweit nun also zur Projektbeschreibung des Wegs der HA IDA zu einem Ticketsystem in naher Zukunft. Wie bereits angeschnitten, steht der Begriff Ticketsystem gegenwärtig in Diskussion, da er dem Leistungsumfang der heutigen, meist hochindividuellen Lösungen nicht mehr gerecht zu werden scheint. Für das Projektvorhaben bot Znumy eine passende Umgebung, um den auf die IDA und den Desk individuell zugeschnittene Lösung – den vorgestellten Anfragenpool – zu konfigurieren.

Um bei der eingehenden Reise-Metapher des ‚Tickets in die Zukunft‘ zu bleiben, lässt sich nun noch mit einem Augenzwinkern sagen:

Sicher warten nach der angenehmen und erfolgreichen Fahrt um die nächste Ecke noch die ein oder anderen Höhen und Tiefen, aber wir haben für alles gesorgt, also wird niemand auf der Strecke bleiben. Der Anfragenpool wartet, also: „Steigen Sie ein – lösen Sie ein Ticket!“ ■■■

Literatur:

Kuhlen, Rainer & Semar, Wolfgang (2023⁷): *Information – ein Konstrukt mit Folgen*. In: Grundlagen der Informationswissenschaft. Kuhlen, Rainer; Lewandowski, Dirk; Semar, Wolfgang; Womser-Hacker, Christa (Hg.). Berlin/Boston: Walter de Gruyter GmbH (3-26).

Schwarz, Karin (2023⁷): *Archive*. In: Grundlagen der Informationswissenschaft. Kuhlen, Rainer; Lewandowski, Dirk; Semar, Wolfgang; Womser-Hacker, Christa (Hg.). Berlin/Boston: Walter de Gruyter GmbH (93-102).

Seidler-de Alwis, Ragna (2023⁷): *Informationsrecherche*. In: Grundlagen der Informationswissenschaft. Kuhlen, Rainer; Lewandowski, Dirk; Semar, Wolfgang; Womser-Hacker, Christa (Hg.). Berlin/Boston: Walter de Gruyter GmbH (477-484).

Wir hinken wahnsinnig hinterher

Der vfm verstärkt seinen Weiterbildungsauftrag im Bereich KI

Herbert Staub

KI ist die Zukunft, kein Stein bleibt auf dem anderen. Um irgendwie beurteilen zu können, was da passiert, und um KI sinnvoll nutzen zu können, braucht man KI-Kompetenz. Und da hinkt Deutschland wahnsinnig hinterher. Der weitverbreitete digitale Analphabetismus fängt schon bei der Digitalisierung an. Das ist ein echtes Problem. Das Einzige, was da hilft, ist Bildung in der Breite – auch in Berufen, wo man denkt, das braucht man jetzt erstmal nicht.

So hat sich Prof. Hanna Bast, Albert-Ludwig-Universität Freiburg, in ihrem Eröffnungsvortrag an der diesjährigen Frühjahrstagung in Freiburg auf die Frage geäußert, ob KI-Kompetenz wichtig sei. Ihr Apell ist beim vfm auf offene Ohren gestossen. Der vfm hat seit 2022 sechs Seminare veran-

staltet, die sich ausschliesslich mit KI befassen. Weil diese Seminare Workshop-Charakter haben, sind jeweils nur 15 Teilnehmende zugelassen. D.h. der vfm hat bisher knapp 100 Leute im Bereich KI weitergebildet. Weil man da noch nicht von Bildung in der Breite sprechen kann, wird der vfm sein Konzept bei den KI-Fortbildungen ändern. Ab Oktober dieses Jahres wird er Crash-Kurse per Zoom anbieten, die knapp vier Stunden dauern. Die Kurse richten sich an KI-Laien und sollen Leute ansprechen, die noch nicht direkt mit KI zu tun haben. Für diese Kurse setzt sich der vfm folgende Lernziele:



Herbert Staub
Studienleiter vfm
herbert.staub@vfm-online.de

Ich bei der Arbeit...



Bilder der Vorstellungsrunde des KI-Seminars vom Juni 2024

<https://docs.google.com/presentation/d/1dqXJT8noxJu8hR8HkpNI7JS8T4ufaxcz6ebowrOWPIM/edit?usp=sharing>

Was ist echt?



„Deutsche bekommen immer länger Rente“ (Quelle)

tagesschau.de brachte einen Beitrag zum Thema Rente. Die Teilnehmenden illustrieren diesen Beitrag mit KI-Bildern. Welches ist das ursprüngliche Bild der Tagesschau?

- Verständnis der grundlegenden Konzepte und Technologien hinter KI
- Verständnis für das Anfertigen guter Prompts zur effizienten Bedienung grosser Sprachmodelle
- Kennenlernen spezifischer KI-Anwendungen und -Tools für die Wissensarbeit/Arbeit in Medienarchiven
- Sensibilisierung für ethische Fragen im Zusammenhang mit KI

Diese knapp vierstündigen Workshops bieten einen fundierten Einstieg in die Welt der Künstlichen Intelligenz, speziell zugeschnitten auf die Bedürfnisse von Mediendokumentarinnen und Mediendokumentaren. Ziel ist es, den Teilnehmenden nicht nur theoretisches Wissen zu vermitteln, sondern sie auch durch praktische Übungen direkt in die Anwendung von KI-Werkzeugen einzubinden. Der Kurs deckt grundlegende KI-Konzepte ab, ermöglicht einen praktischen Zugang und Austausch, erläutert spezifische Anwendungsfälle der KI in der Mediendokumentation und sensibilisiert für ethische Aspekte im Umgang mit dieser Technologie. Leiten wird diese KI-Seminare Thiemo Kremser, Hessischer Rundfunk.

Jan Eggers, der Leiter der bisherigen KI-Seminare, wird einen Schritt weiter gehen. Das letzte KI-Seminar vom Juni 2024 hat viele positive Rückmeldun-

gen generiert. Etliche zielten aber in die gleiche Richtung: „Gefallen hat mir der lebhaftige Vortrag mit den aufwendig gestalteten Folien, die vielen Praxisbeispiele und die Aufgaben in den Gruppenräumen. Ich hätte gerne noch viel mehr von Jan Eggers gelernt, es war wirklich sehr interessant, Danke!“ Oder: „Sollte fortgesetzt und vertieft werden. Wünsche mir mehr technischen Hintergrund (verständlich erklärt wie im Seminar) und mehr Zeit für Übungen.“ Mehr Zeit, mehr Übungen, mehr von Jan Eggers.

Jan Eggers wird dies aufnehmen und im Herbst ein erstes KI-Vertiefungsseminar leiten. In drei halben Tagen wird er die Teilnehmenden mehr selbst machen lassen – Text-Aufgaben lösen, einen Klassifikator bauen („Fanmail oder Werbemail“?), einen RAG-Assistenten („Fragen zu 1000 Seiten“), eine Materialrecherche, einen Faktencheck, eine Geolokation, um nur ein paar Beispiele zu nennen.

Man sieht schon diesen Beispielen an, dass KI eine Querschnittsfunktion hat. Wie Frau Professor Bast in ihrem Referat bemerkte, ist KI universell interessant. Nicht nur in bestimmten Fachgebieten. KI-Kompetenz wird überall gebraucht.

Der vfm nimmt diese Querschnittsfunktion von KI ernst. Alle seine Seminare, von Musikberatung bis Arbeiten mit Daten, vom Rechteseminar bis zum Faktencheck, schliessen KI ein. Denn – nochmals Professor Bast – „KI wird alles verändern“. ■

Seminare des vfm im Herbst 2024

MUSIKBERATUNG/MUSIKDOKUMENTATION 30.9. bis 2.10.2024

Präsenz bei der Deutsche Welle in Bonn

Nach 13 Jahren führt der vfm wieder ein Musikseminar durch. Weil der Austausch das Ziel dieses Seminars ist, wird es in Präsenz in Bonn durchgeführt. Kurze Input-Referate werden den Anstoss zur Diskussion liefern. Natürlich steht dabei KI im Zentrum. Wo wird KI bereits eingesetzt? Wohin bewegen sich Musikdokumentation und Musikberatung?

Es sind aber auch Themen vorgesehen wie Zukunft von Shared Services wie ZEMI, Roller der ZSK, Projekte zur Ähnlichkeitssuche, juristische Fragestellungen usw.

Im Rahmen dieses Seminars haben Sie die Möglichkeit, ein Konzert am Bonner Beethovenfest zu besuchen. Vorgesehen ist das Konzert mit dem jungen Pianisten Fabian Müller, der am 1. Oktober um 19:30 Uhr im Plenarsaal des Bundestages Beethovensonaten spielt.

KI-VERTIEFUNG 15. bis 17.10.2024

Mit Jan Eggers (Online)

In diesem Seminar geht es nicht mehr darum, wie KI funktioniert, sondern wie KI anzuwenden ist. Sie haben die Möglichkeit, mit generativer KI Aufgaben zu lösen:

- Mit ROMANE Textaufgaben bewältigen
- Einen Klassifikator bauen („Fanmail oder Werbemail?“)
- Mit einem RAG-Assistenten grosse Textmengen durchsuchen (Retrieval-Augmented Generation)
- KI für den Faktencheck einsetzen
- Eine Geolokation mit KI durchführen

Das Seminar bietet einen hohen Anteil von praktischen Demonstrationen – KI zum Anfassen und Ausprobieren, inkl. KI-Quiz! Es wird geleitet von Jan Eggers, Datenredakteur im HR und KI-Fachjournalist, bei uns bestens bekannt als Trainer in den Faktencheck-Seminaren.

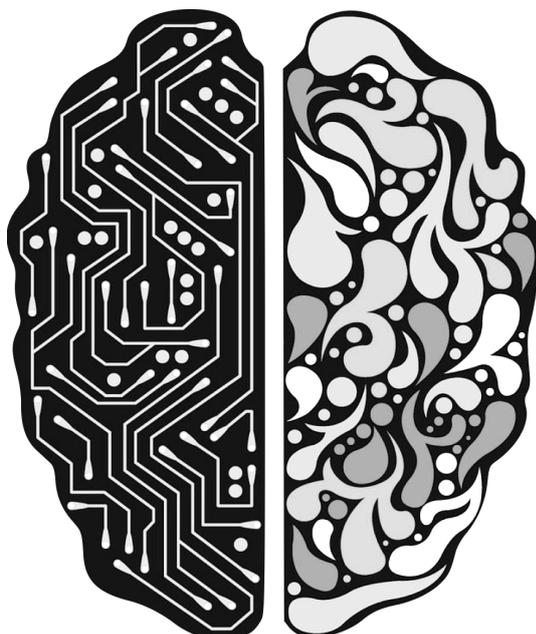
KI-CRASH-KURSE 28.10./11.11./25.11./2.12.2024

Thiemo Kremser (Online)

Dieses knapp vierstündige Seminar richtet sich explizit an KI-Laien in Mediendokumentationen. Ziel ist es, den Teilnehmenden nicht nur theoretisches Wissen zu vermitteln, sondern sie auch mit praktischen Übungen direkt an die Anwendung von KI-Werkzeugen heranzuführen.

Thema des Seminars:

- Grundlagen der KI
Definition und Schlüsseltechnologien der KI
- KI-Tools und Anwendungen
Wie schreibe ich einen guten Prompt?
Experimentieren mit KI-Tools
- Ethische Aspekte der KI
Wie könnte ich KI in der Dokumentation einsetzen?
Diskussion zu Ethik, Datenschutz und Transparenz
Verantwortung in der KI-Nutzung



FAKTENCHECK – GRUNDLAGENSEMINAR 21.10. bis 25.10.2024

Julia Maas und Thiemo Kremser (Zoom)

Weil die Nachfrage so gross ist, bietet der vfm das Seminar „Faktencheck in den Social Media“ erneut an.

Die beiden neuen Referent:innen Thiemo Kremser (hr) und Julia Maas (ZDF) legen in fünf aufeinanderfolgenden Halbtagen die Basis für die Recherche in den Social Media und zeigen Wege auf, wie Fakten gecheckt werden können. Sie werden auch der Frage nachgehen: „Hilft KI beim Aufdecken von KI?“

Dem Input der Referenten folgen jeweils praktische Übungen.

Zusätzliche Aufgaben, für deren Bearbeitung Sie sich zusätzlich 45 Minuten pro Seminartag reservieren sollten, helfen den Stoff zu vertiefen.

UNTERNEHMENSARCHIV – INFORMATIONEN, AUSTAUSCH UND VERNETZUNG 5. bis 7.11.2024

Präsenz bei der Deutsche Welle in Bonn

An der Frühjahrstagung 2023 in Nürnberg war einer der Schwerpunkte dem Thema Unternehmensarchiv gewidmet. 2024 möchte der vfm dieses Thema mit einem Seminar vertiefen. Weil auf Wunsch der Tagungsteilnehmenden Austausch und Vernetzung im Vordergrund stehen, wird der vfm dieses Seminar vom 05.-07. November 2024 erstmals wieder in Bonn durchführen.

Zu folgenden Inhalten wird es kurze Impulsreferate mit anschliessenden Gesprächsrunden geben:

- Was kommt ins Unternehmensarchiv? Auch Objekte?
- Wie gehen wir mit hybridem Schriftgut um?
- Wie kommen die entscheidenden Unterlagen ins Archiv?
- Hat ein Unternehmensarchiv einen Mehrwert für die Bevölkerung (public value) oder nur für die Wissenschaft?

Falls Sie weitere Aspekte diskutieren möchten, nimmt das Büro des vfm Ihre Anregungen gerne entgegen.

ARBEITEN MIT DATEN 18. bis 22.11.2024

Claus Hesseling und Uli Lang (Zoom)

Seit Zahlenreihen, Statistiken, Umfragen, Hochrechnungen vermehrt öffentlich zugänglich sind, hat der Datenjournalismus grossen Auftrieb erhalten.

Medien-Dokumentationen sind gefordert, den Datenjournalismus zu unterstützen: Scrapen, reinigen, aufbereiten, auswerten, visualisieren von Daten sollten zum Alltag in einer Dokumentation gehören.

In diesem Workshop werden die Grundlagen für die Arbeit mit Daten geschaffen und ein Einblick in die praktische Arbeit vermittelt.

Zusätzlich zur Seminarzeit am Morgen sollten die Teilnehmenden noch ca. ½ Stunden pro Tag für Hausarbeit reservieren. Excel-Grundkenntnisse werden vorausgesetzt.

Alle Seminare sind auf der Seite www.vfm-online.de Fortbildung mit genaueren Angaben ausgeschrieben.

Hybrid nach Freiburg

Impressionen von der vfm-Frühjahrstagung vom 22.-24. April 2024

Eva Krause, Thimeo Kremser und Ute Mader

„Wow, was für ein Innovationsfestival der Medienarchive in Freiburg“

Erstmals in der fast 65-jährigen Geschichte der Frühjahrstagungen der Medienarchivar:innen, fand die diesjährige Tagung des vfm – Verein für Medieninformation und Mediendokumentation – in der charmanten Stadt Freiburg als hybride Veranstaltung statt. Trotz des kalten Wetters, obwohl Freiburg ja eigentlich zu den wärmsten Orten Deutschlands gehört, fanden sich die Teilnehmenden vor Ort im Historischen Kaufhaus ein. Gleichzeitig schalteten sich viele Zuhörer:innen vor den Bildschirmen dazu, um den spannenden Vorträgen zu lauschen.

Die vfm-Vorsitzende **Vanessa Sautter Freudrich** begrüßte die Teilnehmenden vor Ort und online vor den Bildschirmen in diesem „stilvollem Ambiente“. Sie betonte, dass unser Berufsbild durch Technologien, KI und New work transformiert wird und wir diejenigen sein können, die sich an diesen Veränderungen aktiv beteiligen. Für diese Veränderungen (und auch die Vorträge vor Ort) ist sehr viel Mut von Nöten. So befand sich in jeder vfm-Tasche ein Tütchen mit Konfetti. Das könne jeder in die Luft schmeißen, um sich selbst einmal ausgiebig zu feiern. Nebeneffekt: Aus diesem Konfettiregen, kann eine wunderschöne vfm-Blumenwiese entstehen.

Die stellvertretende Leiterin des SWR-Studio Freiburg **Jana Lange** begrüßte ebenfalls die Teilnehmenden zu der diesjährigen Tagung. Anschaulich gab sie einen kurzen Einblick in die Rundfunkgeschichte in Freiburg. Dazu beschrieb sie die historische Entwicklung des Tagungsortes. Das Historische Kaufhaus ist noch heute ein Ort des Handels, der Begegnung und der Sprache. Und genau das hat sie allen Teilnehmenden gewünscht: spannende Gespräche und viel Austausch. Als Baderin gab sie noch einen kulinarischen Tipp: Auf dem Markt eine Wurst (gibt es auch als vegetarische Alternative) zu essen – gerne am Vormittag, wie es die Freiburger:innen tun würden.

Auch wenn technisch nicht alles rund lief, waren die Vorträge umso besser. So vermittelte

Prof. Dr. Hannah Bast von der Universität Freiburg in ihrem Eröffnungsvortrag Grundlagenwissen über neuronale Netze und Large Language Model (LLM) und nahm den Teilnehmenden den Schrecken vor der vermeintlichen black box KI.

Die Vorträge zeigten beeindruckende Fortschritte in generativen KI-Technologien für Medienarchive. Die Automatisierung von Anwendungsfällen, insbesondere mit dem Fokus auf der Inhaltserschließung sowie den Chancen und Herausforderungen der Medienarchive in Bezug auf den Umgang mit der Menge an Archivbeständen und der Qualität der KI-Ergebnisse, wurde von **Stefan Doğanay**, **Dr. Dirk Maroni**, **Gabriele Wenger-Glemser**, **Alexander Wolff**, **Claudia Hunold**, **Jakob Rosinski**, **Michael Koltan**, **Dr. Sven Srobel**, **Stefanie Kaiser**, **Helena Knuf** und **Sabrina Ladenburger** gut und differenziert veranschaulicht. Die drei Gewinner:innen des Marianne-Englert-Preises stellten ihre preisgekrönten Arbeiten vor, die sich alle mit aufgetretenen Problemen beschäftigten und eine Lösung anbieten. Daraus entsteht ein direkter Mehrwert für die tägliche Archivarbeit in den jeweiligen Häusern. Schon während der Tagung stieß der E-Learning-Kurs von **Jasmin Casarano** für die Personenansetzung in der ARD-Normdatenbank auf großes Interesse. Eine eigene MES-Software-Anwendung für die automatisierte Formalerschließung von Manuskripten im BR-Unternehmensarchiv program-



Redaktionsteam Vorort in Freiburg: Eva Krause, Thimeo Kremser und Ute Mader (Foto: Birgit Caroline Grill)



Veranstaltungs-Saal im Historischen Kaufhaus (Foto: Birgit Caroline Grill)

mierte **Stefan Merl** und zeigte damit, wie gut die „langweilige Arbeit“ automatisiert werden kann. Auch das Ticket-System von **Lena Cara Wernhöfer** traf auf großes Interesse und fand schon potenzielle Nutzer:innen in den Reihen der Zuhörer:innen, die auch ein Ticket in die Zukunft benötigen.

Neben dem faszinierenden Feld der KI-Technologien gibt es weitere spannende, zukunftsorientierte Bereiche: Social Listening, vorgestellt von **Veronika Förstl** und **Laura Perschon** vom BR, Metadaten-Management präsentiert von **Judith Markert** und **Peter Steltzer**, das Potenzial von Datenauswertungen im beeindruckenden Content-Management der RTL News GmbH durch **Stefanie Goldbecker**, das Management einer Crawler-Datenbank zu Inhalten politischer Akteure in der Schweiz, aufgezeigt von **Flurin Planta** sowie die Verbindung aus Archiv- und Rechtsmanagement mit interessanten Beispielen von erfolgreichen Produktentwicklungen im FAZ-Archiv von **Olivera Kipic**.

Der Höhepunkt der Tagung war zweifelsohne eine fesselnde KI-Diskussion mit **Sarah Ball**, **Peter Effenberg**, **Jakob Rosinski** und **Eva Wolfangel**, meisterhaft moderiert vom geschätzten ORF-Kollegen **Michael Vielhaber**. Die vielfältigen Perspektiven auf KI, die auf der Konferenz geteilt wurden, unterstreichen die Bedeutung des interdisziplinären Austauschs.

Durch die Tagung und die einzelnen Sessions führten verschiedene Moderator:innen, die informativ die einzelnen Präsentationen und Vortragenden anmoderierten und auch durch die ein oder andere technische Panne spontan mit amüsanten Anekdoten überbrückten. So haben die Teilnehmenden beispielsweise gelernt, welche kulinarischen Highlights (Große Rote und Käsekuchen) man in Freiburg genießen kann. Hier zu nennen: **Thomas**

Beckers (WDR), **Christoph Forster** (Sportcast, Köln), **Heiko Linnemann** (Greenpeace, Hamburg), **Sebastian Seng** (ZDF, Mainz), **Michael Vielhaber** (ORF Wien), **Lea Plattner** (SRF, Basel), **Jörg Wehling** (Deutschlandradio, Berlin), **Sabrina Bernhöfer** (ZDF, Berlin).

Eine gute Gelegenheit zum Erfahrungsaustausch und „Zämme kumme“ war der vfm-connect-Abend auf dem Schlossberg, der nicht nur mit seiner beeindruckenden Aussicht auf das Stadtpanorama bestehend war, sondern dazu durch das kulinarische Angebot. Auch die Newcomer wurden dort im Kreise der vfm-Mitglieder willkommen geheißen. Aber auch viele andere Gelegenheiten wurden genutzt um Kontakte zu knüpfen, Wissen auszutauschen und sich kennenzulernen. Zusätzlich zu dem klassischen Stadtrundgang in Freiburg, konnten die Teilnehmenden an verschiedenen spannenden Führungen teilnehmen: Archiv Soziale Bewegungen, Herder Verlag, Historia GmbH, Universitätsbibliothek Freiburg, Abteilung Militärarchiv des Bundesarchivs und NS-Dokumentationszentrum.

Optimistisch in die Zukunft blicken: „What a time to be alive!“

Auch verdienen die Sponsoren Erwähnung, die erneut auf der vfm-Tagung präsent waren. Vor Ort vertreten waren **Archipanion by 4eyes GmbH** aus Basel, **DataScan Computersysteme GmbH** aus Königstein, das **Munzinger-Archiv GmbH** aus Ravensburg, **FUJIFILM Recording Media GmbH** aus Kleve, **medialoopster – nachtblau GmbH** aus Hamburg, **startext GmbH** aus Bonn und **DeepVA** aus Freiburg. Sie stellten in spannenden und unterhaltsamen Präsentationen die aktuellen Möglichkeiten und Entwicklungen auf dem Markt dar. Also anstatt langweiliger Webblöcke viele spannende Neuheiten.

Der diesjährige Impuls to go wurde von **Dr. Ladislaus Ludescher** von der Universität Heidelberg den Teilnehmenden mitgegeben. Er beschäftigt sich mit der „medialen Vernachlässigung des globalen Südens“ und zeigte nicht nur seinen aktuellen Forschungsstand auf, sondern auch die möglichen Ursachen für die mangelnde Berichterstattung. Am Schluss appellierte er, dass mehr Aufmerksamkeit auf den globalen Süden gerichtet werden muss, denn „die größte Katastrophe ist es, nicht über eine Situation zu berichten“.

Im Anschluss der Tagung fanden sich die vfm-Mitglieder zur Mitgliederversammlung zusammen. Neu im Vorstand ist **Sonja Duschek** vom Hessischen Rundfunk. ■■■

Bericht der vfm-Frühjahrstagung vom 22. bis 24. April 2024 in Freiburg

Doku ex Machina

Nutzung von künstlicher Intelligenz und Automatisierung in der Mediendokumentation

Eva Krause, Florian Reichinger, Jonas Wagner

Einleitung

„Was ein wunderschöner Tagungsort“, sagten einige Teilnehmer:innen beim Betreten des Historischen Kaufhauses, das sich direkt gegenüber des Freiburger Münsters auf dem Münsterplatz befindet. Das rote Gebäude strahlte allerdings nicht nur von außen eine Faszination aus. Auch der Tagungsraum, der sogenannte Kaisersaal, verzauberte mit seinem Ambiente. Der große Raum, ausgestattet mit einer riesigen Leinwand für die Präsentationen und verschiedenen Kameras, sorgte für eine perfekte Tagungsatmosphäre. So konnten sich die Teilnehmer:innen vor Ort und die Onlineteilnehmer:innen auf spannende und interessante Vorträge freuen.

Eröffnung



Während die Onlineteilnehmer:innen sich zuschalteten und die Tagungsteilnehmer:innen mit ihren am Eingang erhaltenen vfm-Taschen einen Platz suchten, ging es auch schon los.

Die vfm-Vorsitzende **Vanessa Sautter** (SRF, Zürich) begrüßte die Teilnehmenden vor Ort und zu Hause vor den Bildschirmen mit einem „herzlichen guten Morgen“. Mit kräftigem Applaus würdigte sie das Engagement des Programmkomitees, der Helfer:innen und aller Vortragenden. Sie betonte, dass es großen Mut erfordere, sich auf die Bühne zu stellen und einen Vortrag zu halten. Sie habe in einer Studie gelesen, dass durchschnittlich die größte Angst des Menschen darin bestehe, vor anderen zu sprechen. Erst an zweiter Stelle wurde die Angst vor dem Tod genannt. Sautter hob hervor, dass die Vortragenden nicht nur den Mut haben, ihre Ideen hier zu präsentieren, sondern auch die Courage besaßen und besitzen, Veränderungen anzustoßen. „Wir leben in einer spannenden Zeit mit so vielen Veränderungen“, so Sautter „und in einer Zeit, in der unser Berufsbild merklich durch Technologien, KI oder New Work transformiert wird.“ Sie appellierte daran, dass wir eben nicht nur zuschauen sollen, sondern „für unseren Berufsstand etwas tun,

damit er so herauskommt, wie wir das gerne möchten und damit wir aufgestellt sind, auch für die Zukunft.“ Doch dafür bedarf es eben auch Mut und damit wir alle uns auch im Alltag selbst einmal feiern können, befand sich in jeder vfm-Tasche ein Tütchen mit Konfetti. Aus diesem Konfetti kann – sollte man einen grünen Daumen haben – eine wunderschöne vfm-Blumenwiese wachsen.



Auch **Jana Lange** (SWR Freiburg), Redaktionsleiterin und stellvertretende Studioleiterin des SWR Studios Freiburg, begrüßte alle Teilnehmer:innen vor Ort sowie

zu Hause und hieß alle in Freiburg herzlich willkommen.

Nach der Begrüßung bot Frau Lange den Teilnehmer:innen einen faszinierenden historischen Abriss über die Geschichte des Veranstaltungsortes und des SWR in Freiburg.

Besonders hervorgehoben wurde dabei, dass das Historische Kaufhaus immer „ein Haus des Handelns, der Begegnung und damit auch der Sprache“ sei, so Lange. Und genau das wünschte sie allen Teilnehmer:innen: „gute Gespräche und spannende Sessions“.

Als kulinarische Empfehlung – und es wird definitiv nicht die letzte auf dieser Tagung sein – nannte sie die Wurst auf dem Münstermarkt, die es auch als vegetarische Alternative geben würde. Vor allem sollte man sie morgens essen, wie es die Freiburger:innen selbst bevorzugen. Nämlich bevor der Touristensturm auf dem Markt eintrifft.

Eva Krause,
Jonas Wagner,
Florian
Reichinger
WDR
eva.krause@
wdr.de, florian.
reichinger@
wdr.de, jonas.
wagner@
wdr.de



Eröffnungsvortrag



Den fachlichen Auftakt machte Prof. Dr. Hannah Bast (Albert-Ludwigs-Universität Freiburg). Nach ersten technischen Schwierigkeiten, konnte sie ihren Vortrag „AI Next Level – Was KI alles verändern wird“ beginnen. Sie fragte die Teilnehmenden, ob sie das sogenannten „Trolley-Problem“ kennen. Dies ist eine ethische Fragestellung: Es geht darum, ob man einen herannahenden Zug so umleiten sollte, dass er statt fünf nur eine Person überfährt. Im Anschluss demonstrierte sie u.a. anhand dieses Problems live die verschiedenen Funktionen von ChatGPT, darunter die Bilderkennung, Spracherkennung und computergenerierte Sprache, um die Möglichkeiten moderner KI-Technologien aufzuzeigen.

Um zu verstehen, wie es im Maschinenraum der KI aussieht, erläuterte Frau Prof. Dr. Bast auf sehr verständliche und anschauliche Weise, was ein neuronales Netz ist und wie es funktioniert. Ein System, das ähnlich wie das menschliche Gehirn funktioniert und durch das Training mit einer großen Menge an Daten lernt.

In einer weiteren Live-Demonstration präsentierte Frau Prof. Dr. Bast, wie KI in Echtzeit auf externe Datenquellen zugreifen und relevante Informationen abrufen kann. In diesem Kontext wies sie darauf hin, dass es wichtig sei, freundlich zu der KI zu sein, um sie positiv zu beeinflussen bzw. zu trainieren.

Frau Prof. Dr. Bast ermutigte die Teilnehmer:innen, sich schrittweise mit der Technik vertraut zu machen, indem sie experimentieren und über die Jahre hinweg Erfahrungen sammeln, anstatt sich aufgrund der Komplexität von ihr fernzuhalten. Dies sei entscheidend, um die Vorteile ausschöpfen zu können.

Zusammenfassend bot der Eröffnungsvortrag von Frau Prof. Dr. Bast einen faszinierenden Einblick in die Welt der KI und deren Anwendungsmöglichkeiten in Medienarchiven, was in den folgenden Vorträgen und Diskussion immer wieder aufgegriffen wurde.

SESSION 1: KI im Medienarchiv



Die erste Session der Frühjahrstagung „KI im Medienarchiv“ wurde im Anschluss von Thomas Beckers (WDR, Köln) eröffnet. Er hatte ursprünglich thematisch passend geplant, seine Moderation von ChatGPT formulieren zu lassen, war jedoch mit dem Ergebnis unzufrieden und schrieb seine Eröffnungsrede deswegen dann letztendlich doch lieber

selbst. Inwiefern KI aber nichtsdestotrotz auch heute schon produktiv in medienarchivarischen Kontexten eingesetzt werden kann, sollte in den folgenden drei Vorträgen der Session geklärt werden.



Den Auftakt machte Gabriele Wenger-Glemser (BR, München) mit ihrem Vortrag „Anspruch und Wirklichkeit – eine Bestandsaufnahme von KI in den Medienarchiven“. Sie hielt zunächst fest, dass die Automatisierung von Prozessen mit Hilfe von KI in Medienarchiven aktuell häufig verstanden wird als eine Art Deus ex Machina zur Effizienzsteigerung und als ein Allheilmittel zur Lösung von Sparmaßnahmen, Stellenabbau und Metadaten-Problemen.

Entgegen dieser euphorischen Perspektive beschrieb Gabriele Wenger-Glemser KI nüchtern als eine Maschine, die durch viel Arbeit überhaupt erst einmal ans Laufen gebracht werden muss. Sie konstatierte, dass Anspruch und Wirklichkeit heute oft noch weit auseinander lägen. Die meisten KI-Projekte glichen zum gegenwärtigen Zeitpunkt noch einer Berg- und Talfahrt zwischen Optimismus, Erfolgserlebnissen, Enttäuschung und Ernüchterung.

Um schneller aus den Tälern herauszukommen und die Integration von KI in Zukunft noch produktiver vorantreiben zu können, skizzierte der Vortrag genauer, wie es um die AI-Readiness in den Medienarchiven der ARD aktuell bestellt ist. Die vier Leitfragen lauteten hierbei: Wie können wir KI schnell produktiv nutzen? Welche Daten haben wir bzw. benötigen wir? Wofür setzen wir KI ein? Wie passen wir uns neuer Technologie an?

Obwohl aktuell wohl noch einige Täler durchschritten werden müssen, schloss der Vortrag optimistisch und resümierte zum Abschluss wichtige Erfolgsfaktoren für das Gelingen bestehender und künftiger KI-Projekte, wie klare Zielvorstellungen, geeignete Technologien und skalierbare Produkte, hochwertige Trainingsdaten, Rechtssicherheit und ausreichend finanzielle Ressourcen. Ebenso wichtig außerdem klare ethische Prinzipien, eine Innovationskultur in interdisziplinären Produktteams, ein umfassendes Change Management, genügend Zeit, sowie der grundsätzliche Wille zur Veränderung. Die technologische Transformation muss also auch mit einer Veränderung der Arbeitsorganisation und der Arbeitskultur Hand in Hand gehen, damit der Einsatz von KI in unseren Medienarchiven langfristig Früchte tragen könne.



Im zweiten Vortrag der Session mit dem Titel „Einsatz (generativer) KI-Technologie bei RTL News“ gab **Stefan Doğanay** (RTL News GmbH, Köln) einen interessanten und konkreten Einblick zur aktuellen und zukünftigen Verwendung von KI in der dokumentarischen Praxis. Der Vortrag rekapitulierte zunächst den Einsatz von KI bei RTL News ab 2010, wobei sich die Entwicklungen nach der Veröffentlichung von ChatGPT 3.5 & Co ab 2023 noch einmal rasant beschleunigt haben. Veränderungen der dokumentarischen Arbeit werden bei RTL News in Zukunft insbesondere auch die Entwicklungen im Feld der Computer-Vision-Technologien mit sich bringen. Stefan Doğanay gab diesbezüglich einen kleinen Ausblick auf das Tool Videologger, an dem aktuell gearbeitet wird und das nach Fertigstellung für die dokumentarische Arbeit ein „riesen Gamechanger“ sein werde. Denn mit dem Videologger sollen in Zukunft mit Hilfe von multimodaler KI alle Metadaten zu allen möglichen Essenzen in einem Tool gepflegt werden können.

Wie KI bei RTL News jedoch schon heute in der dokumentarischen Praxis genutzt werden kann, zeigte der Vortrag anschließend am Beispiel der Anwendung „Roberts AI-Tool“, das sich gerade in der Testphase befindet. Der pragmatische Name geht auf den Programmierer des Tools Robert Zängerle zurück. Das Tool ermöglicht es, über Hotkeys und ein Kontextmenü LLMs niedrigschwellig nutzbar zu machen. So können mit dem Tool beispielsweise mit wenigen Klicks KI-generierte dokumentarische Titel, Keywords oder Zusammenfassungen zu Transkripten erstellt werden. Die jeweiligen Prompts können die Kolleg:innen dabei individuell konfigurieren und bearbeiten. „Roberts AI-Tool“ ermöglicht es also, einfach und unkompliziert Erfahrungen im Umgang mit LLMs zu sammeln und integriert die Verwendung von KI niedrigschwellig in den dokumentarischen Alltag. Neben der dokumentarischen Verwendung wird überlegt, das Tool in Zukunft auch zur Unterstützung der Recherche einzusetzen. Geplant ist hierbei genauer eine LLM-gestützte natürlichsprachige Suche im Redaktionsportal.

Zum Ende des Vortrags ging Stefan Doğanay noch auf die Veröffentlichung von Googles Gemini 1.5 Pro ein. Das KI-Modell wurde erst Anfang April veröffentlicht und kann auch aus sehr umfangreichem Videomaterial über einen einfachen dokumentarischen Prompt blitzschnell Titel, Entitäten und

Keywords extrahieren und in erstaunlicher Qualität Transkripte, Abstracts und Sequenzlisten erstellen. Trotz der enormen Geschwindigkeit der technologischen Innovation im Bereich der automatischen inhaltlichen Erschließung blickte Doğanay jedoch zum Abschluss seines Vortrags optimistisch in die Zukunft. Dokumentar:innen würden auch zukünftig noch benötigt. Das Aufgabenfeld könnte sich in Hinblick auf die Integration von KI allerdings erweitern und verändern, beispielsweise in Richtung Prompting und Qualitätskontrolle von automatisierten Prozessen. Es bleibt spannend, wie sich das Berufsfeld in den nächsten Jahren, Monaten, Wochen weiterentwickeln wird.



Die dritte und letzte Präsentation dieser Session hielt **Dr. Dirk Maroni** (WDR, Köln). Sein Vortrag trug den Titel „Potenziale von großen KI-Sprachmodellen in Medienarchiven“.

Zu Beginn arbeitete Dr. Dirk Maroni an Hand von eindrucksvollen Zahlen heraus, dass eine Besonderheit der Medienarchive des öffentlich-rechtlichen Rundfunks ihre enorme Größe ist. Allein der Video-Archivbestand des WDR beläuft sich auf circa 1,9 Millionen Files mit einer Datenmenge von 21 Petabytes (1 Petabyte = 1000 Terabytes). Der Zuwachs an Videomaterial belief sich beim WDR allein im Jahr 2023 auf stolze 1,65 Petabytes. Der Audio-Archivbestand des WDR liegt aktuell bei etwa 4,2 Millionen Files und der Text-Archivbestand in der gesamten ARD beläuft sich auf circa 43 Millionen Datensätze. Dr. Maroni konstatierte, dass diese enormen Datenmengen eine große Herausforderung für die Integration von KI-Sprachmodellen darstellen. Hinzu kommt, dass auch die Anzahl an Sprachmodellen stetig steigt. Die Kombination von riesigen Datenmengen auf der einen Seite und einem höchst dynamischen technologischen Feld auf der anderen Seite macht die Integration von KI in Medienarchiven so anspruchsvoll. Es war in diesem Sinne umso erfreulicher, dass der Vortrag im Folgenden auch aktuelle Fortschritte präsentieren konnte. Dr. Maroni stellte die Ergebnisse eines vom WDR durchgeführten Proof of Concepts vor. Es ging hierbei um die Potenziale von LLMs für die Konzeption einer semantischen Suche und damit um Fragen des Retrievals. Grundidee war, dass Journalist:innen über ein Suchfenster dialogisch im Bestand der ARD recherchieren können und gut belegte Antworten erhalten. Unter Rückgriff auf die Retrieval-Augmented-Generation-Technologie wurde ein Sprach-

modell mit einem Textkorpus von 1000 Lokalzeit-Transkripten optimiert, um anschließend wissensbasierter antworten zu können. Der Vortrag erläuterte die genauere Funktionsweise des eindrucksvollen Modells und stellte dazu auch noch weitere Optimierungsmöglichkeiten über die Anpassung der Algorithmen vor.

Als Fazit hielt Dr. Maroni fest, dass KI-Unterstützung die Nutzungsmöglichkeiten von Archiven auch auf Seiten des Information Retrievals bedeutend erweitern kann. Die Funktionalität muss dabei aber immer am spezifischen Use Case bewertet werden. Allgemein betonte der Vortrag zum Schluss noch einmal, dass bei allen Implementierungen von KI in unseren Medienarchiven sowohl die enorm großen Datenmengen als auch der dynamische Technologiewandel stets mitgedacht werden müssen. Die Herausforderungen sind also groß und es gibt viel zu tun.

SESSION 2: newcomer-forum im vfm e.V. – Neues aus den Hochschulen



Heiko Linnemann (Greenpeace, Hamburg) führte durch das Newcomer-Forum und die Vorträge der Preisträger:innen des Marianne Engler-Preises. Hierbei werden alljährlich die drei besten Projekte des Branchennachwuchses prämiert, die in Bildungseinrichtungen mit bibliothekarisch-dokumentarischer Ausrichtung umgesetzt wurden.

Hierbei werden alljährlich die drei besten Projekte des Branchennachwuchses prämiert, die in Bildungseinrichtungen mit bibliothekarisch-dokumentarischer Ausrichtung umgesetzt wurden.



Den Anfang machte **Jasmin Casarano** (NDR, Hamburg) Masterstudentin der Bibliotheks- und Informationswissenschaften an der TH Köln und Dokumentarin beim NDR,

mit ihrem Projekt „Gewusst wie: Ein interaktiver E-Learning-Kurs für die Personenansetzung in der ARD-Normdatenbank“. Die ARD-Normdatenbank dient als senderübergreifendes Repositorium für Entitäten und Vokabulare und muss daher einen hohen Qualitätsanspruch erfüllen. Dieser Standard der Datenqualität ist im Regelwerk Metadaten formuliert, das den Nutzen die Ansetzungsregeln vorgibt, aber nicht gerade „zum Selbstlernen einlädt“. Um durchwachsender Datenqualität und langen Korrekturlisten zu begegnen, erstellte Jasmin einen 50-minütigen E-Learning-Kurs, dessen Vorbereitung, Konzeption, Umsetzung und Auswertung die Newcomerin präsentierte. Für die Planung hospitierte sie in verschiedenen Teams, legte Zielgruppen fest und entwickelte auf Basis

von Powerpoint einen intuitiven und flexibel in den Arbeitsalltag integrierbaren Kurs, der die wichtigsten Grundlagen der Personenansetzung in der NormDB vermittelt und hierfür einen Wissenstest, interaktive Elemente, beispielhafte Problemfälle und Anwendungsfälle verschiedener Datenbankanbindungen wie FESAD oder HFDB bietet. Der Selbstlernkurs wurde hierbei bewusst auf die Ansetzung von Personen beschränkt, da dies den häufigsten Ansetzungsfall in der NormDB darstellt. Neben geringfügigen Problemen bei der Teams-App und der Webanwendung des Selbstlernkurses und fehlenden Nutzungszahlen wurde der Kurs im NDR gut angenommen und auch beim Tagungspublikum im Saal zeigte sich großer Zuspruch und reges Interesse an dem auf die ARD zugeschnittenen Lernangebot.



Der Titel des zweiten Vortrags „Automate the boring stuff... mit MES – Software-Anwendung für die automatisierte Formalerschließung von Manuskripten“ vom Absolventen des

Volontariats Information Specialist beim BR **Stefan Merl** (RTL, Köln) sprach dem Moderator der Session geradezu aus dem Herzen und nahm die automatisierte Formalerschließung von Manuskripten in den Blick. Im historischen Archiv des BR lagert ein großer Bestand an Manuskripten für Hörfunk und Fernsehen, die jedoch seit 2011 digital erstellt und als „Born Digitals“ archiviert werden. Im Jahr 2023 waren das beim BR 1.800 an der Zahl. Hier setzte das Projekt an und formulierte das Ziel, beliebig viele Manuskripte automatisiert auslesen und gemäß den Schnittstellen der ArchivDB weiterverarbeiten zu lassen. Anstatt die fehlenden Metadaten Autor:in, Titel und Seitenzahl wie bisher von „irgendeiner armen Seele“ intellektuell erfassen zu lassen, was bei der genannten Jahresmenge an Manuskripten etwa dreißig Stunden dauern würde und eine eher undankbare Aufgabe darstellt, sollte nun mittels der Software MES eine Softwareanwendung entwickelt werden, die über eine benutzerfreundliche Bedienoberfläche von den Mitarbeitenden des historischen Archivs bedient werden kann und die Titel und Autor:in der Manuskripte aufgrund der RegEx-Strukturiertheit des Titelblatts auslesen und in einem entsprechenden Import-/Export-Format für die Faust-Datenbank anschlussfähig ablegen kann. Stefan hatte vor dem Projekt keine Programmierkenntnisse und betont die tolle Gelegenheit, mit Tools wie dem „Metadata Extraction Script“ (MES) niedrigschwellig ins Thema Python-

Programmierung einzusteigen. Auch im Bereich Usability betonte er die niedrighschwellige Möglichkeit, mit PyInstaller eine Windows-Anwendung zu erstellen und diese mit Tkinter mit einem simplen Interface zu versehen. Der entwickelte Workflow für die formale Nacherschließung mit MES beinhaltet letztlich den Ingest der Hörfunkmanuskripte, den Export der Essenzen und Metadaten, die Formalerschließung mittels MES und die Datenübernahme in FAUST. Das „Vorbeischleusen an der IT“ (Lacher im Saal waren dem Referenten bei dieser charmanten Formulierung sicher) gestaltete sich als „auffallend einfach“, da die Anwendung dieses vergleichsweise kleinen Skripts von den Verantwortlichen als sicherheitstechnisch unproblematisch wahrgenommen wurde. Zentral für den Erfolg des Projekt war die Kommunikation mit den Menschen, die bisher diese nun automatisiert erhobenen Metadaten produzierten und entsprechende Bedenken hegten, um auch sie davon zu überzeugen, den „boring stuff“ doch den Maschinen zu überlassen.



Mit dem Vortrag „Ticket in die Zukunft – Konzeption eines Ticket-systems als Optimierung für den IDA-Infodesk“ machte **Lena Cara Wernhöfer** (SWR, Mainz) als dritte

Preisträgerin den Abschluss des Newcomer-Forums. Sie ist wie ihr Vorredner ebenfalls Absolventin des Volontariats Information Specialist und ging wie er ebenfalls ein Optimierungsproblem mit ihrem Projekt an. Ihre Spielweise waren hierbei die internen Arbeitsabläufe beim Recherche- und Infodesk der „IDA“ von SWR und SR, die verbessert und weiter in Richtung Arbeitswelt 4.0 professionalisiert werden sollten. Lena nahm uns mit auf die Projektreise durch die Vorbetrachtungen, Konzeption, Konfiguration und Produktivnahme des entwickelten Ticketsystem. Der Infodesk ist seit 2020 eingebettet in eine agile Organisations- und Arbeitskultur als ein Produkt des IDA-Servicebereich und bildet die zentrale und standortübergreifende Anlaufstelle für Rechercheaufträge per Telefon und Mail für die Medientypen Bewegtbild, Musik/Wort, Presse und Online. Von diesen Aufträgen fallen täglich 80 Stück an, die in Echtzeit oder bei niedriger Dringlichkeit im Nachgang bearbeitet werden. Die Projektaufgabe bestand in der Planung, Auswertung und Konfiguration sowie letztlich die Einführung einer Softwarelösung zur An-



nahme, Verwaltung und Analyse von solchen Serviceanfragen. Das Projektziel war hierbei die Optimierung von Workflow, Kommunikation und statistischer Auswertung/Durchsuchbarkeit. Der bisherige Workflow erlaubte es, dass Anfragen entweder direkt oder im Nachgang der Recherche verschriftlicht und ans Sammelpostfach gesendet wurden, woraufhin diese Formulare mit Personen-Flags versehen und mit einem Rechercheergebnis verknüpft wurden. Diese Möglichkeit zur Nacharbeitung und die Problematik, dass mehrere Programme (Outlook, Teams und Excel) für die Dokumentation verwendet wurden, führten u.a. zu hoher Fehleranfälligkeit oder zum Vergessen von Dokumentationsschritten. Die beschlossenen Veränderungen sollten mit einem integrierten System funktionieren, das u.a. transparente Zuständigkeiten und Kommunikation ermöglichen und eine einfache Möglichkeit zum Erstellen von Statistiken bieten soll. Mittels einem ausführlichen Anforderungskatalog mit User Storys wurde nach dem Versuch einer hauseigenen Lösung und einer Marktsondierung letztlich die Open Source Ticketing-Software „Znuny“ gewählt. Nach einer ausführlichen Testphase mit dem Infodesk ist im Falle einer Erfolgreichen Inbetriebnahme auch eine Ausweitung der Software auf andere Anwendungsfälle wie den Programmaustausch, das Noten- oder das Unternehmensarchiv denkbar. Lena hob die Sounding Board-Methode hervor, die es dem Projektteam ermöglicht hat, gut zusammenzuarbeiten und gleichzeitig Akzeptanz für das neue System zu schaffen. Belohnt wurde die Referentin mit begeistertem Applaus und Interesse an einem ARD-weiten Einsatz des Systems.

Die Session wurde mit der feierlichen Verleihung des **Marianne-Englert-Preises** beendet. Der Archivar Dr. Joseph Nietfeld sponserte zudem je eine Champagnerflasche. Herzlichen Glückwunsch an

alle Preisträger:innen!

SESSION 3: KI in der Mediendokumentation



Die erste Session des zweiten Tages widmete sich dem Thema „KI in der Mediendokumentation“ und wurde von **Sebastian Seng** (ZDF, Mainz) eröffnet. Die Session startete nach dem sehr gelungenen Zämme kumme im Schlossbergrestaurant Dattler am vorangegangenen Abend doch recht früh. Auch einer der drei Referenten wurde noch gesucht, tauchte dann aber letztendlich pünktlich zu seinem Vortrag auf. Der Moderator Sebastian Seng versprach dem Publikum jedoch drei spannende Präsentationen, welche die „müden, kleinen Augen“ ganz schnell in „wache, große Augen“ verwandeln würden.



Mit dem Ziel dieses Versprechen einzulösen, startete die Session dann mit dem Vortrag von **Jakob Rosinski** (Amazon Web Services, Hamburg), der den Titel „Anwendungsfälle von generativer AI für die audiovisuelle Optimierung sowie Aggregation von Archivmaterial“ trug. Zu Beginn konstatierte Rosinski, dass in Bezug auf die KI-Ausgaben in der Medienbranche für die nächsten fünf Jahre eine durchschnittliche Wachstumsrate von über 30 Prozent prognostiziert wird. Auch Amazon möchte sich an dieser Entwicklung beteiligen und bietet diverse KI-basierte Anwendungen und Services, die in Medienarchiven sehr nützlich eingesetzt werden können – unter anderem zur Produktivitätssteigerung, zur Beschleunigung von Such- und Distributionsprozessen oder aber auch für die Erzeugung neuer bzw. optimierter Produkte.

Einige dieser KI-basierten Services und Entwicklungen stellte der Vortrag im Folgenden dann sehr anschaulich vor. Darunter war unter anderem eine semantische Suchmaschine für Videosequenzen und eine Anwendung zur automatischen Erstellung von Highlight-Schnitten und Video-Zusammenfassungen mit jeweils erstaunlich guten Ergebnissen. Rosinski präsentierte außerdem einen Service für die KI-basierte Erhöhung der Auflösung von alten Videodateien und ein Tool, welches die Möglichkeit bietet, nachträglich sehr hochwertige Zeitlupen-Szenen aus Archivmaterial zu erzeugen.

Für das größte Erstaunen sorgten aber zweifellos die Vorstellungen aus dem Bereich der Synchronisierung.

Der Vortrag präsentierte hier Anwendungen, mit denen Videomaterial in zahlreiche Sprachen übersetzt und dann mit vorher trainierten Stimmen automatisch synchronisiert werden kann. Darüber hinaus ist es mit Hilfe von KI möglich, die Lippenbewegungen an die jeweilige Tonspur automatisch anzupassen. Das originale Videomaterial kann so mit einem minimalen Aufwand für eine Vielzahl an internationalen Nutzer:innen optimal aufbereitet und angepasst werden.

So beeindruckend die technischen Entwicklungen auch sein mögen, wurde insbesondere in Bezug auf die nachträgliche Veränderung von Lippenbewegungen an dieser Stelle auch noch einmal sehr deutlich, dass die enormen Potentiale der Implementierung von KI gerade im Medienkontext immer auch ethische Fragen aufwerfen und eine große gesellschaftliche Verantwortung mit sich bringen. Die vormals „müden, kleinen Augen“ des Publikums waren nach diesem sehr interessanten Vortrag nun in jedem Fall hellwach und bereit für den zweiten Vortrag der Session.



Dieser wurde von **Michael Koltan** (Archiv Soziale Bewegungen, Freiburg) gehalten und trug den Titel „Der digitale Archivar – Lessons learned aus einem KI Erschließungsprojekt“. Der Vortrag startete zunächst mit einigen sehr interessanten Ausführungen zur langen Geschichte der KI, bei denen unter anderem der Unterschied von regelbasierter symbolischer KI und statistischer KI näher erläutert wurde.

Im Anschluss präsentierte Michael Koltan dann den „digitalen Archivar“ – ein KI-Projekt des Archivs Soziale Bewegungen, bei dem die Vorteile von statistischer und symbolischer KI miteinander verknüpft werden sollten, um gedruckte Archivbestände automatisiert erschließen zu können. Der Workflow gestaltete sich hierbei folgendermaßen: Ausgehend von einem Scan sollte mit Hilfe von OCR (optischer Texterkennung) zunächst der Text des jeweiligen Dokuments ermittelt werden. Anschließend sollten dann über Entity Recognition unter anderem relevante Personen, Körperschaften oder Orte aus den Texten extrahiert werden. Im letzten Schritt sollten diese statistisch ermittelten Termini dann wiederum mit Wikidata verknüpft werden.

Das Projekt wurde in den Jahren 2020–2021 vorangetrieben, scheiterte jedoch letztendlich an der Finanzierung. Und auch die Qualität der durch

die OCR extrahierten Texte zeigte sich damals für die Erschließung noch als Problem. Die Texte, die für das Archiv Soziale Bewegungen so typischen Dokumente, wie beispielsweise Flugblätter, konnten nämlich stellenweise nur unzureichend erkannt werden, was dann wiederum negative Auswirkungen auf die Ergebnisse der Entity Recognition hatte. Michael Koltan berichtete jedoch, dass heutzutage durch den Einsatz von ChatGPT und anderen LLMs sowohl die Fehlerbehebung der Texterkennung als auch die Ergebnisse der Entitätenerkennung deutlich verbessert werden können, wenngleich durch die Implementierung stellenweise auch wieder neue Fehler entstehen.

Der Vortrag schloss in diesem Sinne mit dem Resümee, dass LLMs auf dem Weg zu einer automatisierten Erschließung von Archivmaterial wohl auch nicht alle Probleme lösen werden, uns aber mit Sicherheit ein gutes Stück weiter voran bringen können. Wir dürfen also gespannt sein, ob der „digitale Archivar“ in Zukunft noch mal eine zweite Chance bekommt.



Den dritten und letzten Vortrag der Session hielt Dr. Sven Strobel (TIB, Hannover) unter dem Titel „Künstliche Intelligenz trifft auf Mediendokumentation – Erfahrungen und

Perspektiven aus dem TIB AV-Portal“. Dr. Sven Strobel resümierte zunächst noch einmal kurz den Wandel in der Mediendokumentation – ausgehend von der manuellen Indexierung analoger Medien in den 1970er Jahren, über die Anfänge der Digitalisierung und der Computerisierung in den 1980er und 1990er Jahren, sowie das Aufkommen des Internets und die Entstehung von Onlinedatenbanken in den 2000er Jahren bis zu den Automatisierungsprozessen und den Anfängen der KI in den 2010er Jahren, um schließlich bei der verstärkten Integration von KI seit den 2020er Jahren zu enden. Es wurde hier noch einmal sehr deutlich, wie sehr sich unser Berufsfeld in den letzten Dekaden immer wieder gewandelt und weiterentwickelt hat.

Der Vortrag präsentierte im Anschluss den aktuellen Stand der technologischen Entwicklungen des TIB AV-Portals – einer Open-Access-Plattform für wissenschaftliche Videos mit einem besonderen Schwerpunkt auf Naturwissenschaft und Technik mit aktuell circa 45.000 recherchierbaren Videos. Über eine Kombination von KI-Techniken und Automatisierung wird hier aktuell die Erfassung, Indexierung und

Suche der audiovisuellen Inhalte des Portals optimiert. In der Video-Analyse findet so eine automatische Shot Detection, Speech Recognition, Text Recognition, Image Recognition und ein automatisches Named-Entity-Linking statt. Für die Optimierung der Suche werden die erkannten Entitäten außerdem automatisch mit Synonymen und Übersetzungen verknüpft. Die Funktionsfähigkeit der einzelnen Services präsentierte der Vortrag sehr anschaulich in einer Live-Demonstration. Strobel ging dabei zum Schluss auch noch einmal näher auf den Qualitätssprung ein, der durch die Implementierung des KI-basierten Spracherkennungsmodells Whisper (OpenAI) für automatische Video-Untertitelungen, Transkripterstellungen und Übersetzungen erreicht werden konnte.

Die technologische Weiterentwicklung des mediarchivarischen Feldes schreitet also auch hier mit großer Geschwindigkeit voran und verspricht große Potentiale für die Automatisierung der Erschließung und die Verbesserung der Recherchierbarkeit von audiovisuellen Medien.

SESSION 4: KI kills the DaTa Stars – dokumentarische vs. künstliche Intelligenz: Diskussionsveranstaltung



Auch in diesem Jahr fand wieder eine spannende Diskussionsrunde statt. Moderiert wurde sie von Michael Vielhaber (ORF, Wien).

Als Podiumsgäste diskutierten Sarah Ball (Ludwig-Maximilians-Universität, München), Peter Effenberg (jay, Potsdam), Jakob Rosinski (Amazon Web Service, Hamburg) und Eva Wolfangel (Journalistin, Stuttgart) und weitere Tagungsteilnehmer:innen über den Zusammenhang und auch die Zusammenarbeit von KI und den Mediendokumentar:innen als Data Stars.



Inhaltlich fokussierte sich die Debatte auf die Verwendung von Daten als Trainingsdaten sowie die Herausforderungen und Probleme von KI. Dabei waren sich die Diskutant:innen einig, dass KI von qualitativ hochwertigen Daten sehr profitieren kann. Dabei sollte allerdings vor allem darauf geachtet werden, dass die Daten auch rechtlich und ethisch sauber sind. Ein Vorschlag dazu war, dass wir Modelle selber bauen können, um die eigenen Bedürfnisse mit einzubeziehen und mit eigenen rechtlich abgesicherten Daten trainiert zu können.

Ein weiterer Diskussionsimpuls kam aus dem Publikum: Sollten wir historisches Material, das ggf. schlechtere Qualität hat, mit Hilfe von KI optimieren? Daraus resultierend stellte sich den Podiumsgästen die weitere Frage: Was ist der Verwendungskontext des Materials? Will ein:e Nutzer:in das Originaldokument oder ein Symbolbild, um ein inhaltliches Thema zu unterstreichen? Alle waren sich einig, dass das erzeugte KI-Material sehr gut gekennzeichnet werden muss. Ein pro-Aspekt wäre aber die Vorstellung, so eine Diskutantin, in der Zukunft beispielsweise mit Albert Einstein via Zoom über seine Theorien sprechen zu können.

Im Zuge der ethnischen und rechtlichen Aspekte spielte auch die Frage nach der Transparenz von KI-Modellen eine wichtige Rolle. Ein Plädoyer ging dahingehend, dass in diesem Kontext wichtig sei, mit welchen Trainingsdaten die KI trainiert wurde. Dabei muss auch die Frage nach dem Urheber von KI generierten Texten betrachtet werden – diese Frage ist allerdings noch offen.

Eine spannende Frage war diejenige nach aktuellen großen Entwicklungen im Bereich KI. Das Beispiel des AI Agent wurde hervorgehoben, also einer KI, die mit Hilfe von Interaktion mit der Umgebung eine Aufgabe lösen kann. Dabei wird vermutlich auffallen, dass wir eine große Sicherheitslücke haben, so eine Meinung. Dies wiederum wird vermutlich die Möglichkeiten der KI einschränken.

Als Fazit der Diskussionsveranstaltung können einige Zitate der Podiumsgäste hervorgehoben werden:

Data Stars dürfen wir auf keinen Fall töten – Für ein gutes Outcome ist eine Kollaboration zwischen Mensch und Maschine wichtig – Um keine schlaflosen Nächte wegen einer mangelnden Cybersecurity zu haben, müssen wir uns damit befassen und einbringen – Die Welt ist komplex – Wir leben in spannenden Zeiten.

SESSION 5: New Work



„New Work“ verspricht Agilität und Innovativität. Mit diesem Thema des Wandels war die Session überschrieben, durch die **Lea Plattner** (SRF, Basel) leitete und die neue Formen der Zusammenarbeit beinhaltete.



Das Thema New Work griffen **Veronika Förstl** (BR, München) und **Laura Perschon** (BR, München) direkt auf. Denn beide waren vorher in getrennten Arbeitsberei-

chen tätig und haben sich ganz New Work-mäßig für das Kennenlernen einfach mal zum Gassigehen mit Veronikas Hund getroffen. Neben dieser frischen Meeting-Kultur, die die beiden pflegen, hat sich das Projekt ihres Vortrags „Social Listening als Beispiel für zukünftige dokumentarische Arbeitsfelder“ dem relevanten und nicht ganz einfachen Problem des Social Media-Monitorings verschrieben. „Social Listening“ bezeichnet hierbei das „Hineinhören in die sozialen Netzwerke“ und soll aktuelle Themen und Diskussionen ins Programm bringen, die die regionale Zielgruppe des BR bewegt. Das entwickelte Tool ermöglicht es, tatsächlich gestellte Fragen außerhalb der journalistischen Bubble direkt an Expert:innen oder den BR-Faktencheck zu stellen. Diese Themen sind oft niederschwelliger als in der „Bubble“ angenommen oder müssen nochmals behandelt werden, um relevante Fragen und Sorgen ernsthaft zu bearbeiten oder prominente Themen wie beispielsweise Cannabis auf dem Oktoberfest a.k.a. „Smoketober“ oder den neuen Leberkäsedöner aufzugreifen. Technisch realisiert wurde das Tool mit einer cloudbasierten und auf die redaktionelle Zielgruppe zugeschnittene Lösung, das über 460 Millionen Quellen aus Foren, Websites und Blogs KI-basiert nach Sentiment-Analyse und Viralität sortiert für die Suche zur Verfügung stellt. Die Qualität der Sentiment-Analyse, die Vereinheitlichung von keywords und die Priorisierung der Themen stellen hierbei Kernprobleme dar. Die Treffer werden schließlich sortiert, interpretiert (bei einer Regionalsuche nach der Stadt Füssen sollen nun mal keine Fußfetisch-Bilder ausgegeben werden), ggf. mit dem Faktencheck abgeklärt und als Themenanregung an die Redaktionen weitergeleitet oder als allgemeines Themenmonitoring bereitgestellt werden. Die Teamarbeit ist hierbei agil gestaltet, um der schnelllebigen Onlinewelt zu begegnen.

Die Nähe zum Journalismus nimmt das Team nicht als Problem wahr, denn es versteht sich als Schnittstelle zwischen Redaktion, Dokumentation und Publikum. Das Tagungspublikum interessierte sich speziell für die regionalisierende Suchabfrage, die beim BR zum einen in Landespolitik nach Parteien und Politiker:innen und zum anderen in Regionalsuchen nach Städten/Regionen und wichtigsten Themen aufgeteilt ist. Veronika und Laura stellten mit dem Social Listening ein sehr mächtiges Tool vor, das jedoch nur interdisziplinär und mit großem Engagement mit Leben gefüllt und ständig weiterentwickelt werden muss.



Die Referentin des zweiten Session-Vortrags **Stefanie Goldbecker** (RTL News GmbH, Köln) widmete sich in ihrem Vortrag „Neues Anforderungsprofil: Datenauswertung in der Mediendokumentation“ den neuen Bedürfnissen, die von News-Ressorts an die dokumentarische Datenauswertung gerichtet werden. Hierbei gibt es einen großen Bedarf an visuell aufbereiteten Datenauswertungen zu bestimmten Berichterstattungen, Verwendungsmenge von Agenturmaterial, Ausstrahlungen von bestimmten Sendungen oder Berichterstattungen zu einem bestimmten keyword. So wurde ein regelmäßiger Bericht zu Nahost bei ntv, RTL, RTL2 und den VOX-Nachrichten für die Geschäftsleitung realisiert oder eine Abfrage der genauen Zeiten von Berichterstattungen zum Thema „Grillen“ für das Werbemarketing. Außerdem wurden für das Team Datenjournalismus die Promis ermittelt, über die 2023 am meisten bei RTL Exklusiv berichtet wurde – hier an Platz eins Prinz Harry mit seiner Autobiographie. Die dokumentarische Kernkompetenz liegt hierbei in der Kenntnis der Programme und im Zugriff auf verschiedene Datenquellen wie Sendepläne, Schnittprojekte, Webvideos, 13 Mio. gemanagete Videofiles und eine Mio. Onlineartikel. Die Auswertung wird als Pivot-Tabelle oder Power BI-Bericht vom geschulten Expertenteam Data geliefert. Das Team koordiniert die Anfragen, bearbeitet sie mittels SQL-Abfragen und nutzt hierfür das Data Warehouse von RTL, welches Informationen aus verschiedenen Systemen bündelt. Eine Herausforderung und gleichzeitig große Chance war die Notwendigkeit, in einem Team mit verschiedenen Wissensständen Kompetenzen im Bereich Excel, Power BI und SQL zu erlernen oder weiter auszubilden und zu einem produktiven Arbeitsalltag zu gelangen. Das Publikum interes-

sierte sich besonders für die datenjournalistische Komponente der VIP-Auswertung und die Hauptstakeholder des Services, welche bei RTL sehr breit gestreut sind. Und zuletzt noch der Hinweis von Herbert Staub: es sind nach aktuellem Stand noch Plätze frei im VfM-Seminar Arbeiten mit Daten im November.



Für den dritten Vortrag der Session gab **Stefanie Kaiser** (SRF, Zürich) – passenderweise im Kaisersaal des Historischen Kaufhauses – einen Einblick in ein erfolgreiches Kooperationsprojekt innerhalb ihres Senders unter dem Vortragstitel „Der SRF-Newsflash – Recherche und Archive im Zentrum einer abteilungsübergreifenden Newsproduktion“. Der Newsflash ist eine Nachrichtensendung am späteren Abend im Schweizer Fernsehen mit den aktuellsten News des Tages, die seit 2022 in der aktuellen Form im Programm ist. Bis 2018 wurde eine herkömmliche Tagesschau mit Magazinbeiträgen gesendet, die dann 2019 aus Ressourcengründen auf die Moderation verzichtete und noch mit Videonews als „Newsflash“ gesendet wurde, damals noch unter Führung der Tagesschau-Redaktion, aber schon unter Beteiligung des Recherchedesks, der die gesamte Newsproduktion von Radio und TV in einem „Newsroom“ integrierte. In der Folge wurde bei verschiedenen Newsprodukten stärker in Fachredaktionen vernetzt, anstatt herkömmliche Sendungsredaktionen wie die Tagesschau weiterzuführen. Nun war es für Redaktionen, Design, Regie, Recherche und Archive eine große Umstellung und für viele nicht ganz einfach. Die neuen Strukturen boten die Möglichkeit, den Aufwand für einzelne Newsprodukte zu durchleuchten und die Ressourcennutzung nach dem gleichberechtigten Vier-Kräfte-Model (Inhalt, Publikum, Technologie und Distribution) zu optimieren. Ein positiver Effekt der Umstrukturierung war die synergetische Nutzung der Radionews-Inhalte für den Newsflash, die nun noch bebildert werden müssen. Hierfür liefert das Archiv entsprechende Fotos zu, die dann vom TV mit publizistischer Kompetenz geprüft werden. Technisch wurde die Software „Newsbutler“ entwickelt, mit der einfach Audiofiles mit Bildern und Einblendungen versehen und gesendet werden können. Die aus der Not geborene Form der heutigen Sendung entstand durch die Ressourcenverlagerung. Der Prozess war jedoch hilfreich, um die Bereiche IT, Radio, TV und Recherche und Archive nach dem Vier-Kräfte-

Modell zu integrieren. Heute schätzen alle Beteiligten den Austausch mit den anderen News-Bereichen und der entstandene Newsroom steht sinnbildlich für die neue Arbeitsweise des SRF, und Recherche und Archiv ist stolz, mittendrin zu sein. Auf Nachfrage teilte die Referentin mit, dass vor allem Bilder aus der Aktualität und weniger Agenturfotos verwendet werden und dass einige Kompetenzen im Fotobereich erlernt werden mussten.

SESSION 6: Metadatenmanagement



„KI gibt es eigentlich nicht ohne gute Metadaten“, greift **Jörg Wehling** (Deutschlandradio, Berlin) zu Beginn seiner Moderation einen Satz vom Vortrag auf. Herr Wehling be-

schäftigt sich in seinem Haus selbst auch mit Metadaten. Aber nicht nur dort. Er berichtete, dass er seit letztem Jahr ein neues Hobby habe: Er ist in der Rennleitung für Galopprennen. „Man glaubt gar nicht“, so Wehling, „wie viele Metadaten beim Galopprennen wichtig sind“. Und diese komplexe Thematik soll in dieser Session einmal heruntergebrochen werden.



Den Auftakt dieser Session machte **Flurin Planter** (Studio Bundeshaus, Bern) mit seinem Vortrag zum Thema „Metadaten fallen nicht vom Himmel – Wie Politmonitor Dokumente

politischer Akteure verschlagwortet“. Politmonitor ist eine Textdatenbank, die mit Hilfe von automatisierten Prozessen im Internet nach öffentlich zugänglichen Dokumenten sucht, diese herunterlädt, verschlagwortet und über eine Recherchemaske an einem zentralen Ort bereitstellt. Bei diesen Dokumenten handelt es sich u.a. um Medienmitteilungen, Parteizeitungen, Berichte oder Positionspapiere der wichtigsten Akteure der Bundespolitik. Politmonitor ist also eine Mediendatenbank, nicht mit archivierten Zeitungsartikeln, sondern mit öffentlich zugänglichen Publikationen. Planter zeigte in seinem Vortrag an einigen Anwendungsbeispielen anschaulich, wie Politmonitor mittels Web-Crawling funktioniert. Besonders hervorgehoben wurde das Search Tagging, das die Generierung von neuen Metadaten ermöglicht, die durch das Crawling nicht mitgeliefert wurden. Zudem demonstrierte Planter, wie die Recherche funktioniert und wie Dashboards mit bestimmten Abfragen erstellt und über SharePoint abgerufen werden können. Diese Dashboards werden im

Hintergrund automatisch kontinuierlich aktualisiert.

Abschließend betonte Planter die Wichtigkeit der Metadaten für eine effiziente Recherche und Einbettung, da sie die Suchmöglichkeiten verbessern würde und die manuelle Arbeit erleichtere. Eine Herausforderung besteht allerdings darin, dass das Crawling und Tagging an Webseiten stets angepasst werden müsse und dies je nach Einstellung ein kompliziertes Unterfangen sei. Perspektivisch soll in die Datenbank eine KI integriert werden.

Das wichtigste Learning betonte Planter ganz zum Schluss: „Ohne Metadaten ist auch der größte Datensatz wertlos“

Nach einer ausführlichen Fragerunde aus dem Publikum und von den Onlineteilnehmer:innen aus dem Chat, leitete Jörg Wehling den nächsten Vortrag mit dem Titel „Zeitfunk: Exploration automatisierter Inhaltserschließung von historischen Tondokumenten im Deutschen Rundfunkarchiv“ ein.



Claudia Hunold und Alexander Wolff (DRA, Potsdam) präsentierten das innovative Projekt „Zeitfunk“. Dieses Projekt zielt darauf ab, historische Tonbestände durch auto-

matistische Verfahren zu erschließen. Dabei werden in Kombination verschiedene Data-Mining-Tools und Large Language Modellen verwendet. Bei dem verwendeten Bestand handelt es sich um den „Zeitfunk“-Bestand, also aus der Redaktion des Deutschlandsenders (Rundfunk der DDR) aus den 1950er und 1960er Jahren. Eine Herausforderung bestand darin, dass der Bestand nicht ordentlich archiviert wurde. Er beinhaltet verschiedene Sammelbänder mit Weißband, die sich in nicht immer ordentlich beschriftete Bandkartons befinden.

Das DRA setzt bei diesem Projekt auf zwei maßgebliche Tools: DRA-Auto-Cut und DRA-Auto-Doc. Beide Tools wurden von den Referent:innen an Beispielen demonstriert.

Bei dem DRA-Auto-Cut handelt es sich um ein Tool, das automatisch Tonunterschiede zwischen Weißband und Beiträgen auf den Sammelbändern erkennt und diese in einzelne Beiträge schneidet. Diese Cuts können von den Dokumentar:innen in einem nächsten Schritt überprüft werden. Im Anschluss werden daraus einzelne Audiodateien generiert. Für das Tool wurde eine spezielle Schnittstelle zur Open-Source-Software Audacity (Audio-Schnitttool) entwickelt.

Im nächsten Schritt übernimmt der DRA-Auto-

Doc die automatische Generierung von Metadaten. Dieses Tool befindet sich momentan in der Entwicklung. Ein zentrales Element des DRA-Auto-Doc ist das Tool „Wisper“, das Transkripte erstellen kann. Da Wisper allerdings zu sogenannten Halluzinationen neigt, wenn es nicht ausschließlich für gesprochene Texte verwendet wird, werden die Files durch den Speech Activity Detection und Musikdetektion vorverarbeitet. Daraus entsteht dann ein reines File mit Sprechpassagen. Die Metadaten werden dann durch ein weiteres LLM (Google Gemini) generiert und den Dokumentar:innen zur Überprüfung vorgeschlagen. Dann erfolgt noch ein Abgleich mit der Normdatenbank. Wichtig zu erwähnen ist, dass nur Metadaten generiert werden, die im Audiofile tatsächlich erwähnt werden, d.h. beispielsweise keine Sendedaten.

Die Kernherausforderung bei diesem Projekt sei die optimale Kombination von Mensch und Maschine („Human in the loop“), um sicherzustellen, dass automatisiert generierte Daten überprüft und korrigiert werden können. Das DRA zeigte somit, wie modernste Technologien zur Erhaltung und Erschließung von historischen Rundfunkmaterial beitragen können.



„Wir beamen uns jetzt in die Jetzt-Zeit“, so begrüßt Jörg Wehling die beiden Referent:innen **Judith Markert** und **Peter Steltzer** (RTL News GmbH, Köln). In ihrem Werk-

stattbericht „Jedem Anfang wohnt ein Zauber inne“ ... – Metadatenveredelung am Anfang der Prozesskette“ vermittelten die beiden Referent:innen einen eindrucksvollen Einblick in die Herausforderungen und Lösungen der Metadatenveredelung und Metadatenvererbung in den aktuellen TV-Produktionen von RTL News. Ein zentrales Ziel sei es, Orientierung in der täglichen Materialflut zu schaffen und eine Vordokumentation zu ermöglichen.

Bereits im Ingest Prozess liegen Metadaten in unterschiedlicher Qualität vor. Diese werden dann vordokumentiert und veredelt. Ziel ist es, eine schnelle Auffindbarkeit sicherzustellen und ein Überblick über die Rechtelage des Material zu geben.

Ein zentraler Punkt des Vortrags war Vorstellung der Prozesskette und die Frage, wo die Hauptlast der dokumentarischen Arbeit liegen würde. Die Referent:innen betonten, dass nicht jedes Material, das einläuft, dieselbe Priorisierung hat. Dabei stellten sie fest, dass bei der Materialflut vorher schon eine Orientierung geschaffen werden muss, damit

die Kolleg:innen das gewünschte Material effizienter finden können. Unter anderem mit Unterstützung einer KI kann eine Kategorisierung und Verschlagwortung des Materials stattfinden. Aktuell befindet sich eine KI-gestützte Vergabe von Kategorien nach IPTC Standards auf Basis automatisierter Transkripte in der Testung.

Wie wird das einlaufende Material nun priorisiert und welche Tools unterstützen dabei? Der erste Faktor ist der Mensch. Die Kolleg:innen nehmen an Ressortversammlungen teil und kommunizieren die besprochenen Themen. Daraus werden neue Schlagwörter angelegt und diese ebenfalls kommuniziert. Ein zweiter Faktor ist der sogenannte Cutlistenzähler. Dieser zählt die Zugriffe auf das Material und hilft somit bei einer Relevanz-Ermittlung.

Die lizenzrechtlichen Metadaten werden durch die Datenbank Infoclicense verwaltet, die eine Echtzeitvisualisierung der Handlungsempfehlungen auf Basis von maschinenlesbaren Lizenzverträgen bietet. Diese werden in das Redaktionsportal übernommen.

Daneben erkennt das Archivierungstool, welche Sequenzen aus Archivmaterial bestehen und gibt dort dann die ArchivID an, damit immer das Quellmaterial recherchiert und verwendet werden kann.

Wie geht's weiter? Die Genealogie, also eine Rückwärtstraverse zum Quellmaterial und eine Vorwärtstraverse zum Zielmaterial, wodurch u.a. Cluster geschaffen werden können, soll im Redaktionsportal als Empfehlung integriert werden. Eine Gesichtserkennung und ein Videologger mit integrierter KI-Hilfe sollen eingebunden sowie Altsysteme abgelöst werden.

Der Werkstattbericht verdeutlicht, wie wichtig eine durchdachte Metadatenveredelung am Anfang der Prozesskette und eine Metadatenvererbung ist.

SESSION 7: Medienarchive im Produktionsprozess



Die letzte Session der Tagung stand unter dem Titel „Medienarchive im Produktionsprozess“ und versprach noch einmal einige interessante Einblicke in die medienarchivarische Praxis. **Sabrina Bernhöft** (ZDF, Berlin) führte humorvoll in die Session ein und ergänzte den Titel angesichts der riesigen Mengen an täglich zu bewältigendem Material um den kleinen Zusatz „Der Weg ist das Ziel“. Die drei Vorträge der Session sollten zum einen die Bedeutung und Funktionalität

von Medienarchiven im Produktionsprozess aufzeigen und zum anderen noch einmal veranschaulichen, wie die Automatisierung von Prozessen uns einiges an Arbeit abnehmen kann. Dass die technologische Weiterentwicklung jedoch nicht nur Arbeitsprozesse erleichtert, sondern auch neue Probleme produzieren kann, zeigte sich dann allerdings ungeplant an Hand von technischen Problemen bei der Online-Zuschaltung der ersten Referentin.



Der Ablaufplan musste deswegen spontan umgestellt werden. Die Session startete also anders als ursprünglich geplant mit dem Vortrag von **Helena Knuf** (ZDF, Mainz) mit dem Titel „Was wird wie gesucht? – Nutzungsdaten des SPHINX-Dashboards zur Optimierung des Bestandsaufbaus, der Serviceangebote und der User Experience“. Helena Knuf stellte im Folgenden das ZDF-Projekt „SPHINX-Dashboard“ näher vor. Es handelt sich hierbei um ein Analysetool zur Erhebung, Auswertung und Aufbereitung der Bestands- und Nutzungsdaten des Informations- und Archivsystems SPHINX, wobei im Archivbereich aktuell vor allem der Videobereich im Fokus steht. Ziel des Projekts ist es, Nutzungsdaten als Lenkungsinstrument zu nutzen, um die Archivbestände und die Datenbankoberflächen langfristig noch stärker auf die Bedürfnisse der Nutzer:innen auszurichten und Ressourcen gezielter einsetzen zu können. Vor allem Suchanfragen, Sichtungsprozesse, Exportvorgänge und Interaktionen mit der Benutzeroberfläche sollen dafür systematisch erfasst und ausgewertet werden. Für das Tracking der Aktivitäten in der SPHINX-Datenbank wird dabei das Tool Matomo verwendet, und für die Visualisierung der Ergebnisse ist außerdem noch die Anwendung Kibana im Einsatz.

Der Vortrag veranschaulichte sehr praxisnah, wie die Analyse von Bestands- und Nutzungsdaten nicht nur für Fragen der archivinternen Schwerpunktsetzungen und Aufgabenverteilung, sondern auch in Bezug auf Fragen der Programmgestaltung eine wichtige datenbasierte und transparente Diskussionsgrundlage schaffen kann. Es wurde jedoch auch noch einmal ausdrücklich betont, dass die ermittelten Ergebnisse auch weiterhin immer wieder mit den Erkenntnissen der Kolleg:innen aus der Recherche und mit den Erfahrungen aus dem persönlichen Kontakt mit Kund:innen abgeglichen werden sollen. In Bezug auf den aktuellen Projektstatus resümierte Knuf zum Schluss ihres Vortrags: „Wir sind mittendrin

und ganz gespannt, wie es weitergeht mit unserem SPHINX-Dashboard-Projekt.“ Wir sind ebenso neugierig, wie sich das Projekt in den nächsten Jahren entwickeln wird.



Der zweite Vortrag der Session wurde von **Sabrina Ladenburger** (BR, München) gehalten und trug den Titel „Automatisierte Playlisterstellung in der Musikedokumentation – KI-basierte Analyse und Verschlagwortung von Songtexten“. Sabrina Ladenburger stellte in ihrer Präsentation genauer das Projekt PARAS (PlaylistAufbereitung&RechercheArchivService) vor, das sich aktuell in einer Evaluierungsphase befindet. Es handelt sich hierbei um ein Dienstleistungsserviceprojekt der Musikedokumentation des BR, bei dem die Musikredaktionen des Hauses mit Musiksammlungen und Material zur Programmvorbereitung versorgt werden sollen. Dabei reicht das Spektrum der erstellten Musiksammlungen von Ereignisplaylists (z.B. für Jubiläen) über Playlists für dauerhafte größere Themenkomplexe (z.B. zu Klimawandel oder Migration) bis hin zu Playlists für einmalige Ereignisse (z.B. für Nachrufe) mit jeweils begleitendem Zusatzmaterial. Die Präsentation von Playlists und Begleitmaterial soll im Sinne einer einfachen Nutzbarkeit zukünftig über eine eigene Website erfolgen – mit Schnittstellen zu den Datenbanken des BR. Das Musikarchiv will sich auf diese Weise proaktiver in die Programmgestaltung und in den Produktionsprozess einbringen, als wichtiger Partner sichtbar werden und den Kooperationsgrad zwischen Archiv und Redaktion langfristig erhöhen.

Da die Zusammenstellung der Playlists und des Begleitmaterials händisch sehr arbeitsintensiv ist, ist die Implementierung von KI ein zentraler Bestandteil des PARAS-Projektes. Über gutes Prompting und gezielte Anfragen an ChatGPT & Co können die Dokumentar:innen beispielweise schnell neue thematische Kontexte erschließen und sich anschließend themenspezifische Sammlungen von Songs generieren lassen. Um Fehlern entgegenzuwirken, kann der Prompt dabei die Aufforderung enthalten, die vorgeschlagenen Künstler:innen und Musiktitel immer auch direkt mit einem Wikipedia-Link und einem Link zu Genius (Online-Wissensdatenbank für Songtexte) zu verknüpfen. Mit Hilfe von ChatGPT & Co können darüber hinaus aber auch in Windeseile prägnante Zusammenfassungen von Ereignissen, Songtexten und Wirkungsgeschichten erstellt werden, thematische Songsamm

lungen nach Jahrzehnten geordnet werden oder dokumentarische Schlagworte für Songs generiert werden. Die Einsatzmöglichkeiten von KI sind also auch im Projekt PARAS äußerst vielfältig. Sabrina Ladenburger betonte jedoch, dass die teilautomatisch erstellten Inhalte selbstverständlich immer noch von Dokumentar:innen überprüft und verifiziert werden müssen.

Der Vortrag schloss in diesem Sinne mit dem beruhigenden Fazit, dass der Einsatz von KI die intellektuelle Arbeit auch hier wieder nicht ersetzen kann, sie aber wesentlich bereichert. KI kann dazu beitragen, große Themenkomplexe und Datenmengen 'noch schneller zu durchdringen, um Redaktionen noch umfangreicher und effizienter zu arbeiten zu können.



Da die technischen Probleme zwischenzeitlich behoben werden konnten, startete im Anschluss **Olivia Kipic** (F.A.Z., Frankfurt am Main) in die letzte Präsentation der Session.

Ihr Vortrag trug den Titel „Produktentwicklung in der Praxis – erfolgreiche Strategien und Fallbeispiele“ und präsentierte eine beeindruckende Auswahl an digitalen Produkten, die von der Abteilung Archive und Rights Management der F.A.Z. in den letzten Jahrzehnten maßgeblich mitentwickelt wurden. Näher vorgestellt wurden unter anderem GBI-Genios, PMG Presse-Monitor, Media Monitoring Lizenzen, Schullizenzen und das F.A.Z.-Bibliotheksportal (eine digitale Plattform für Bibliotheken für die Recherche im F.A.Z.-Archiv). Auch im Hinblick auf die Fortschritte im Bereich der KI stellte Olivia Kipic Produktentwicklungen vor, wie die vollautomatische Transkribierung von Audio-Podcasts mittels KI und ein aktuelles Projekt zur Entwicklung eines Archiv-Chatbots in Zusammenarbeit mit Aleph Alpha. Die Verschmelzung des Archivbereichs mit der Contentvermarktung und dem Rights-Management hat sich in den letzten Dekaden bei der F.A.Z. also als äußerst produktiv erwiesen.

Der Vortrag ging jedoch nicht nur auf spezifische Produkte ein, sondern auch auf die allgemeineren Strategien zur Produktentwicklung. Dabei wurde ein Bogen gespannt von der Analyse der jeweiligen Zielgruppe (wie z.B. Journalist:innen, Redakteur:innen, Wissenschaftler:innen, Bildungseinrichtungen, Kulturinstitutionen, Unternehmen, Medienkonsument:innen) bis hin zur Analyse von Marktbedürfnissen und Trends (wie z.B. Digitalisierung, On-Demand-Kultur, Personalisierung, Multimedialität, Nostalgie, KI und Data Mining).

Die Präsentation veranschaulichte zum Abschluss der Session noch einmal sehr eindrücklich die Innovationskraft von Medienarchiven und ihre Bedeutung nicht nur für die internen Prozesse in großen Medienunternehmen, sondern auch für externe Nutzer:innen und Kund:innen. Die Möglichkeiten das Spektrum der medienarchivarischen Tätigkeiten immer wieder weiterzuentwickeln sind enorm und in der Abteilung Archive und Rights Management der F.A.Z. sind noch viele weitere Projekte in Planung. Die Zukunft bleibt spannend.

Impuls to go



Der Moderator der Session **Michael Vielhaber** (ORF, Wien) hob in seiner Einleitung hervor, dass sich der Referent **Dr. Ladislaus Ludescher** (Universität Heidelberg), wissenschaftlich mit der Darstellung des globalen Südens in

Hauptnachrichtensendungen Deutschlands, Österreichs und der Schweiz beschäftigt. Vielhaber betonte, dass wir immer intensiver mit den Journalist:innen konstruktiv zusammenarbeiten, um passendes Material anzubieten. Denn wir kennen das Material im Zweifel oft besser.

„Was wir über unsere Gesellschaft, ja über die Welt, in der wir leben, wissen, wissen wir durch die Massenmedien“. (Niklas Luhmann, 1927-1998)

Mit diesem Zitat begann Ludescher seinen Vortrag „Vergessene Welten und blinde Flecken – Die mediale Vernachlässigung des Globalen Südens“. Es soll die zentrale Rolle der Medien in der Gesellschaft unterstreichen.

Als Untersuchungszeitraum seiner wissenschaftlichen Betrachtung nannte Ludescher den Zeitraum zwischen den Jahren 2007 und 2023. Verschiedene Medien zog er als Untersuchungsgegenstand heran: Die Tagesschau, politische Talkshows, Radiosendungen, Printberichte, ausländische Medien und ausländische Zeitungen. Die Auswertung zeigte, dass starke Zentren im Norden dominieren und der globale Süden kaum erwähnt wird. Bei den Topthemen wird der globale Süden sogar noch seltener erwähnt, manche Staaten verschwinden gänzlich aus der Berichterstattung. Während etwa 90% der Berichte den globalen Norden betreffen, entfallen nur rund 10% auf den globalen Süden – ein Muster, das sich durch alle Medien zieht. Die Berichterstattung über die Coronapandemie hat diese Thematik noch verschärft.

Den globalen Hunger als „das größte lösbare Pro-

blem der Welt“ zog Ludescher als Beispiel heran, um zu verdeutlichen, dass dieses Thema im Diskurs und in der Medienwelt kaum eine Rolle spielte, obwohl viele Hilfsorganisationen davor warnten.

Ein Grund für die Vernachlässigung des globalen Südens identifizierte Ludescher in der Platzierung der Auslandskorrespondenten. Das Journalistennetzwerk ist ungleich verteilt. Ein weiterer Grund ist, dass die westlichen Interessen oft bevorzugt betrachtet werden. Dies zeigt sich in den aktuellen Ergebnissen: der tödlichste Krieg des 21. Jahrhunderts in Äthiopien und die schlimmste humanitäre Krise im Jemen wurden kaum erwähnt, bis westliche Interessen betroffen waren.

Diese Vernachlässigung hat erhebliche Auswirkungen auf die gesellschaftlichen Diskurse. Misereor fordert daher mehr Aufmerksamkeit für die Länder des Südens, da mediale Aufmerksamkeit einen politischen Diskurs erwirken kann.

Am Schluss gab Ludescher in einem Satz den Impuls to go mit: „Die größte Katastrophe ist nicht das Vergessen, denn dann muss man es vorher gewusst haben, sondern die größte Katastrophe ist, überhaupt nicht darüber zu berichten.“

Unterm Strich

Was bleibt unterm Strich nach zweieinhalb Tagen zu sagen?

Zunächst einmal können alle Verantwortlichen, Helfer:innen und Vortragende ihr Konfetti aus der vfm-Tasche in die Hand nehmen und einmal in die Luft schmeißen.

Das gewählte Tagungsthema und die Sessions hoben, wie eigentlich schon letztes Jahr in Nürnberg erkennbar, die Wichtigkeit in der Auseinandersetzung von KI-Tools in den Medienarchiven hervor. Es war spannend zu sehen, welche Tools bereits verwendet werden, welche sich in der Testphase befinden und welche Möglichkeiten es sonst noch gibt. Auch das Thema „New Work“ und die damit verbundenen Chancen waren für den Themenschwerpunkt eine Bereicherung.

Alle Vorträge waren auf einem hohen Niveau, der Fokus auf die Mediendokumentation stets erkennbar. Alles in allem waren die Sessions in sich stimmig.

Auch die Diskussionsrunde war in diesem Jahr ein Highlight. Das lag nicht nur am Thema sondern auch an den vom beruflichen Background her sehr unterschiedlichen Podiumsgästen. Auch die regen Anmerkungen und Fragen aus dem Publikum ga-

ben neue Impulse während der Diskussion. Erfreulicherweise waren sich am Ende alle einig: „Data Stars dürfen wir auf keinen Fall töten“.

Generell zeugte die aktive Beteiligung des Publikums und von den Onlineteilnehmer:innen während der gesamten Tagung von einem großen Interesse an der Thematik und einer regen Diskussionskultur.

Am meisten Luft nach oben hatte dieses Jahr jedoch leider die Technik. Neben den Problemen vor Ort bei Live-Demonstrationen, kamen auch Ton- oder Bildprobleme vor.

Einen Vorteil hatten die technischen Probleme allerdings: die Moderator:innen und Vortragenden überbrückten die Zeit wunderbar und gaben beispielsweise kulinarische Empfehlungen für die Zeit in Freiburg zum Besten. Sollten Sie also noch einmal oder zum ersten Mal Freiburg einen Besuch abstatten, gönnen Sie sich eine „Lange Rote“ am Münsterplatz oder einen Käsekuchen aus dem Automaten am Hauptbahnhof.

Die Sponsoren erhielten auch in diesem Jahr einen eigenen Block und konnten ihre Produkte und Neuheiten vorstellen. Zu sehen waren wieder viele bekannte Gesichter: **Markus Stauffiger** von Archipassion by 4eyes (Basel), **Robert Fischer** von DataScan Computersysteme GmbH (Königstein), **Sebastian Skrypka** von FUJIFILM Recording Media GmbH (Kleve), **Marc Jonas** von medialoopster - nachtblau GmbH (Hamburg), **Friederike Munzinger** von Munzinger-Archiv GmbH (Ravensburg) und **Martin Heiter** von startext GmbH (Bonn)

Der Gesellschaftsabend „Zämme kumme“ am ersten Abend war ebenfalls ein Highlight der Tagung. Denn zunächst trafen sich die neuen Tagungsteilnehmer:innen alleine dort, um sich untereinander kennenzulernen. Im Anschluss folgte dann für alle ein toller Abend. Es wurden viele alte Bekannte getroffen und neue Menschen kennengelernt sowie viele interessante Gespräche geführt.

Neben dem Gesellschaftsabend fanden am zweiten Tag wieder verschiedene sehr sehenswerte Führungen statt, die ebenfalls immer sehr positiv in Erinnerung bleiben.

Unterm Strich war es also eine sehr gelungene Tagung, die insgesamt ein sehr bereicherndes Programm bot.

Die Vorfreude auf die Frühjahrstagung 2025 ist groß. ■■■

Eröffnungsrede

der vfm-Frühjahrstagung in Freiburg am 22. April 2024

Vanessa Sautter

Liebe Kolleg:innen

Herzlich willkommen hier in Freiburg und zuhause vor den Bildschirmen zur vfm-Frühjahrstagung. Ich freue mich sehr auf die nächsten drei Tage mit spannenden Präsentationen und Diskussions-Themen.

Das Programmkomitee hat wieder einmal grossartige Arbeit geleistet und mit grossem Engagement ein herausragendes Programm zusammengestellt. Dies wäre jedoch nicht möglich gewesen, ohne die zahlreichen inspirierenden Einreichungen zum Call for Papers. Ich möchte allen danken, die sich die Zeit genommen haben, ihre Ideen und Beiträge einzureichen. Da waren so viele grossartige Beiträge dabei, dass wir die nächsten drei Tage wirklich aus dem Vollen schöpfen können.

Interessant ist hier zu erwähnen, dass laut Forbes Magazine die grösste Angst eines durchschnittlichen Menschen ist, vor anderen Personen zu sprechen. Die zweitgrösste Angst der Menschen ist – der Tod.

Darum bin ich umso glücklicher, dass so viele kompetente Präsentierenden den Mut gefasst haben und hier auf die Bühne kommen.

Sie haben aber nicht nur Mut gefasst, um hier zu präsentieren, sondern zeigen auch auf, dass sie Courage bei ihrer Arbeit gezeigt haben, gehandelt und Veränderungen angestossen und vorangetrieben haben.

Und wir alle sind heute hier, um diese Geschichten zu hören und vielleicht dann auch inspiriert von diesen Präsentationen selbst Mut zu fassen, um etwas zu verändern.

Denn wir leben und arbeiten in einer spannenden Zeit, in der unser Berufsbild, sei es durch Technologie, KI oder New Work etc. transformiert wird. Wir sind nicht nur Zeugen dieser Veränderung, sondern haben auch die Möglichkeit, sie aktiv mitzugestalten und zu lenken.

Zum Inhalt der Frühjahrstagung: Wir haben den Ablauf der Tagung etwas umgestellt, deswegen vergeben wir den Marianne-Englert-Preis an die Newcomer bereits heute Nachmittag. So haben sie auch Zeit beim vfm-connect (früher Gesellschaftsabend) heute Abend die Gewinner:innen noch etwas auszuquetschen und besser kennenzulernen. Die Diskussion rund um das Thema KI findet am Dienstagmorgen statt. Und am Schluss gibt es, wie letztes Jahr auch, den Input to go und am Mittwochnachmittag die Mitgliederversammlung.

Nun wünsche ich allen einen erfolgreichen Start in die Tagung und viele spannende Begegnungen.



Vanessa Sautter
vfm-Vorsitzende
Schweizer Radio und
Fernsehen SRF
Zürich
vanessa.sautter@
vfm-online.de

Von der Mitgliederversammlung des vfm

Am 24.4.2024 fand im Historischen Kaufhaus in Freiburg die ordentliche Mitgliederversammlung des vfm statt. Es gab einige wichtige Ergebnisse:

Mitgliederversammlungen können künftig online oder hybrid stattfinden. Dies wurde mit folgendem Beschluß festgelegt:

„Die Mitgliederversammlung beschließt, dass Mitgliederversammlungen ab sofort auch hybrid und virtuell stattfinden können. Wird eine hybride oder virtuelle Versammlung einberufen, so muss bei der Einberufung auch angegeben werden, wie die Mitglieder ihre Rechte im Wege der elektronischen Kommunikation ausüben können.“

Der Vorstand wurde in Teilen neu gewählt. Zur Wahl standen:

Frank Dürr; Sonja Duschek; Michael Hafner; Gudrun Menze; Uta Rosenfeld; Ralph Schmidt

Alle Kandidaten wurden mit je einer Enthaltung gewählt.

Angesichts der allumfassenden Preiserhöhungen, die sich auch auf die Veranstaltungen des vfm auswirken, muß auch der vfm seine Mitgliedsbeiträge erhöhen. Ab 2025 gilt daher:

„Die Erhöhung des Beitrags für die Einzelmitgliedschaft auf 70€~~7~~ Jahr und die Erhöhung des Beitrags für die korporative Mitgliedschaft auf 500€~~7~~ Jahr wird wie vorgeschlagen einstimmig angenommen.“

Auf Grund guter Bewirtschaftung ist der vfm trotz der zahlreichen Preissteigerungen weiterhin gut aufgestellt, wie der Finanzbericht ausweist.

Auch bei den Fortbildungsseminaren des vfm haben sich neue Organisationsformen etabliert:

Der Wechsel zu Online-Seminaren hat sich etabliert. Für aktuelle Themen bietet sich das Meet-up als Format mit einer schnellen Umsetzung an. Neben online Seminaren, bei denen die Wissensvermittlung im Vordergrund steht, soll es zukünftig auch auf Themen spezialisierte Seminare geben, die vor Ort durchgeführt werden, bei denen der Austausch im Vordergrund steht.

Die üblichen Regularien des Vereins wurden alle satzungsgemäß durchgeführt, wie z.B. die Entlastung des Vorstands und die Kasse Prüfung.

Link zum Geschäftsbericht:

www.info7.de/info7_2024-1_S57-64.pdf

„Zweifellos schafft also der Rundfunk frohe Menschen ...“

Musik aus dem Radio in den Fabrikhallen

Jörg-Uwe Fischer



Dr. Jörg-Uwe Fischer
Deutsches
Rundfunkarchiv
Marlene-Dietrich-
Allee 20
14482 Potsdam
joerg-uwe.fischer@
dra.de

Gemeinschaftsempfangsanlagen in Unternehmen und Betrieben, die Ende der 1920er, Anfang der 1930er Jahre einiges an Zuspruch fanden, übertrugen Rundfunk und andere Darbietungen bei Betriebsveranstaltungen und Werkfeiern. Und sie dienten aber eben auch zur Erleichterung der alltäglichen Arbeit in der Maschinenhalle wie im Büro. Durch die Vermittlung von Musik oder Wort während der Arbeit sollten Produktion und Wirtschaftlichkeit gesteigert und dadurch gleichzeitig widerlegt werden, solche Anlagen als puren Luxus abzutun.

In unseren Tagen ist viel die Rede davon, wie wichtig es ist, dass sich Mitarbeitende am Arbeitsplatz wohl fühlen. Von Arbeitgeberseite wird viel getan, um eine positive Arbeitsumgebung zu schaffen. Mitarbeiterzufriedenheit hat einen hohen Stellenwert. Die Rede ist von individueller Arbeitsplatzgestaltung, flexiblen Arbeitszeiten, Feedbacks, aber auch von angenehmer Beleuchtung oder Farbe im Büro, Obstkörbe, Sportprogramme und Kinderbetreuung. Rundfunk jedoch, der Zugang zu Information, Unterhaltung und Bildung bietet, spielt dabei keine Rolle.

Dies war nicht immer so, gerade in den 1920/30er Jahren, als sich in der Industrie die Fließarbeit und Bandfertigung durchsetzte, die Mechanisierung der Arbeitsleistung beständig zunahm, Arbeitsvereinfachung und -vereinheitlichung Herausforderung war, bei gleichen Unkosten mehr zu leisten war, um die Produktion zu steigern, gab es Überlegungen, wie ein Ausgleich zur zunehmenden Mechanisierung vorgenommen werden konnte. Eine Methode des Ausgleichs waren häufiger eingelegte Pausen, der Arbeitnehmer sollte bei seinem Achtstundentag nicht über Gebühr ermüden. Eine andere, neuartige Art des Ausgleichs war die Beschallung der Arbeits-

stätten mit Rundfunkprogrammen. Das Ziel war, die von eintöniger Arbeit geistig nicht voll ausgelastete Arbeitskraft in einer Form abzulenken, die sich mit der durch die Art der Arbeit bedingten Aufmerksamkeit vereinbaren ließ und die Güte der Arbeit nicht gefährdete. Versuche mit Tieren in Fabrikhallen oder Musikkapellen in Betrieben verfehlten das gewünschte Ziel, die angewandten Mittel waren letztlich auch zu teuer.¹ Als kostengünstiges Zerstreungsmittel, das auch den größten Teil des Tages zur Verfügung stand, kam deshalb der Rundfunk zum Zuge.

Drei Beispiele dazu: aus dem Berliner Kleinbauwerk der Siemens-Schuckert-Werke, der Berliner Schokoladenfabrik Leonhard Monheim (Trumpf) und aus einer nicht näher verortbaren österreichischen Hutfabrik.

Aus dem Berliner Kleinbauwerk der Siemens-Schuckert-Werke berichtet 1927 dessen Direktor Paul Heinrich Perls (1871-1942) in der Wochenschrift *Funk* über einen Versuch von Rundfunkkonzerten in den Arbeitsstätten des Werkes. Die Werksleitung benutzte Radioapparate, um die Arbeitnehmer bei der Bandfertigung, die sehr eintönig war, „ein wenig anzuregen, um über das ewige Einerlei 'spielend' hinwegzukommen.“² Lediglich in den je eine Stunde am Vor- und am Nachmittag blieb der Lautsprecher ausgeschaltet. Für Arbeitnehmer, die in verlärmten Werkstätten ihrer Arbeit nachkamen, wurde während der Mittagspause der Apparat auf die Werkskantine umgeschaltet. Der Zweck der Beschallung in den Fabrikationsräumen war, das Monotone der Arbeit in Abständen zu unterbrechen und eine „gewisse

¹ Vgl. dazu A. Lion, Rundfunk als Freund der Werktätigen. Erleichterung einformiger mechanischer Fabrikarbeit, in: Welt und Wissen, Jg. 16, H. 28, 1927, S. 47-50.

² Paul H. Perls, Wenn gute Reden sie begleiten, dann fließt die Arbeit ..., in: Funk, Wochenschrift des Funkwesens, H. 8, Berlin 1927, S. 63.

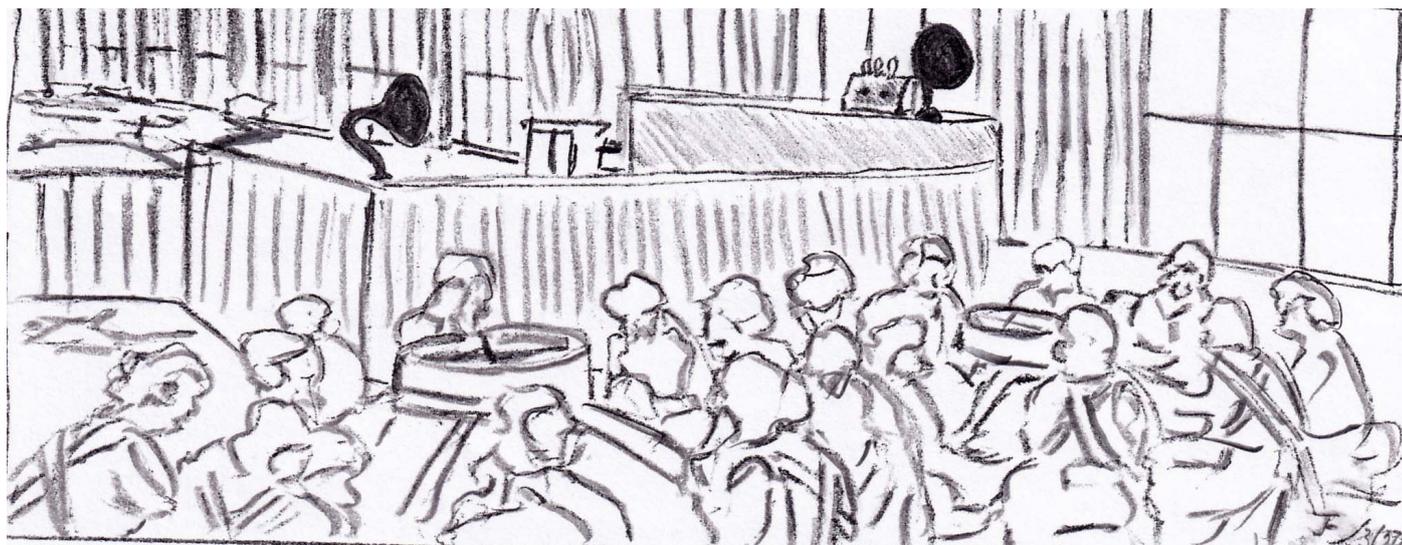


Abb.: Jörg-Uwe Fischer, Bleistiftzeichnung: Arbeiterinnen hören Radio in einer Fabrik

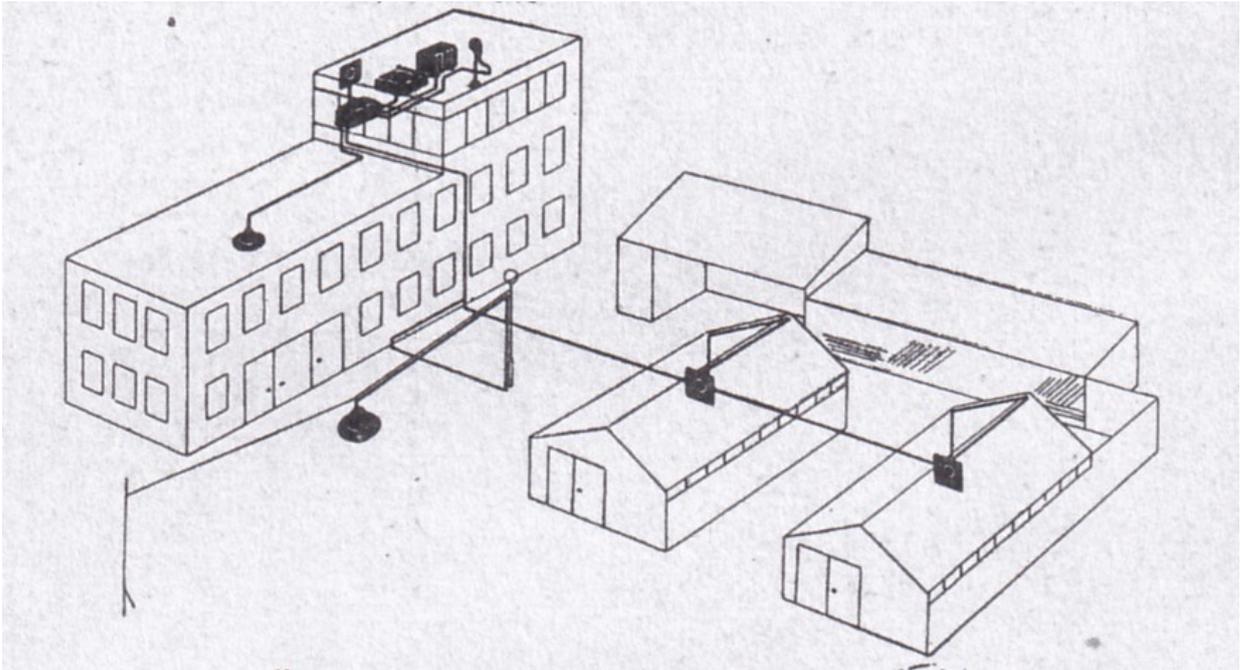
Belebung“ zu erzielen. „Der Rhythmus, durch geeignete Musik erzeugt, verfehlt niemals seine Wirkung auf die Bewegungen der Menschen. (...) Freilich kommt dabei viel auf die geeignete Musik an; in den meisten Fällen wird ein flotter Marsch, werden Lieder, Tanzweisen und ähnliche leichte musikalische Darbietungen am Platze sein.“ Seitens der Belegschaft fanden die Rundfunkkonzerte in den Arbeitsstätten breite Zustimmung.

Rundfunk im Kampf gegen Eintönigkeit und Abstumpfung bei der Arbeit nutzte im selben Jahr auch die Berliner Schokoladenfabrik Leonhard Monheim, einem zur damaligen Zeit modern durchorganisierten Betrieb mit Fließarbeit, die bei der Schokoladefabrikation erforderlich war. Auch hier war das Mittel gegen Einförmigkeit der Arbeit leichte Musik. Eingerichtet wurde eine Zentralempfangsanlage, an die 17 Trichterlautsprecher bzw. an der Decke hängende trichterlose Lautsprecher angeschlossen waren. Die Belegschaft war anfangs skeptisch, zeigte sich dann aber zufrieden: „Die Arbeiterschaft, die sich zuerst dieser Neueinrichtung gegenüber skeptisch und abwartend verhielt, erkennt heute ihren Nutzen an; bei einer Abstimmung (...) darüber, ob die Unterhaltungsmusik eingestellt werden sollte oder nicht, stimmten von 380 Arbeitnehmern 368 für die Fortführung. Andererseits ist auch die Fabrikleitung mit dieser Einrichtung zufrieden, deren Einwirkung, besonders auf die Frauen, außerordentlich günstig ist. Ein Rückgang der Arbeitsleistung infolge dieser Zerstreung, den man der Fabrikleitung prophezeit hatte, ist nicht eingetreten.“³

Auch in der österreichischen Hutfabrik waren gute Erfahrungen mit Rundfunkdarbietungen während der Arbeitszeit gemacht worden. Der Besitzer hatte eine zentrale Rundfunkempfangsanlage einrichten lassen, so dass die Belegschaft ständig Rundfunkdarbietungen hören konnte. Einwendungen, dass sich die Arbeiter nicht auf ihre eigentliche Arbeit konzentrieren können, wenn Musik spielt, hatten Versuche des Fabrikbesizers und seiner wissenschaftlichen Berater widerlegt: „Es hat sich gezeigt, daß die Arbeiter und Angestellten bei Musikbegleitung viel besser und gewissenhafter arbeiten als früher und daß sich seit der Verwendung der Lautsprecheranlage die Qualität der Erzeugnisse wesentlich gebessert hat. Zweifellos schafft also der Rundfunk frohe Menschen, die mit weit größerer Liebe bei ihrer Arbeit sind.“⁴ Die Radioeinrichtung mit angeschlossenenem Grammophon in der Hutfabrik diente indes nicht nur dazu, die Belegschaft bei guter Laune zu halten, sondern sie wurde gleichzeitig als „Befehlsanlage“ der Betriebsleitung genutzt: „Will der Chef einen Auftrag erteilen oder eine bestimmte Person rufen lassen, so geschieht das nicht mehr, wie früher, durch kompliziertes Telephonieren, durch eine Betriebsansprache oder durch die Vermittlung eines Boten. Der Chef spricht einfach durch das Mikrophon zu allen seinen Angestellten.“

³ Siebzehn Lautsprecher in einer Schokoladenfabrik, in: Funk. Wochenschrift des Funkwesens, H. 42, Berlin 1927, S. 347.

⁴ Radio im Fabrikbetrieb, in: Radiowelt, 11. Jg., H. 10, Wien 1934, S. 300.



Skizze einer Übertragungsanlage in einer größeren Fabrik mit mehreren Montagshallen.
Quelle: Radio-Helios, Jg. 42, Nr. 12, 1936, S. 716.

Über die Jahre hatten die Hersteller entsprechend Erfahrungen gesammelt und konnten Verstärker- und Übertragungsanlagen nahezu passgenau für Fabriken, Handwerks- und Handelsbetriebe, Gemeindegänge, Schulen und Institute anbieten: Tonaufnahmegeräte für Rundfunk, Plattenspieler, Mikrofone, Steuerstufen und Leistungsstufen der Vor- und Endverstärker mit integrierten Schall- und Regelvorrichtungen.

Die Propagierung des Gemeinschaftsempfangs durch das NS-Regime ab 1933 hatte zur Folge, dass entsprechende Anlagen unter Verwendung des DAF 1011 aufgebaut wurden, des Deutschen Arbeitsfront-Empfängers, der ab 1935 dem Volksempfänger VE 301 als Standardgerät für den Gemeinschaftsempfang in den Betrieben gefolgt war. „Wenn gerade die Elektroakustik in den letzten Jahren einen so gewaltigen Aufschwung genommen hat, dass man es heute als eine reine Selbstverständlichkeit empfindet, in modernen Betrieben Lautsprecher zu finden, dann ist das nicht zuletzt auf den politischen Einsatz der Lautsprecher zurückzuführen.“⁵ Hier wird der Lautsprecher also missbraucht als Bindeglied zwischen NS-Führung und Bevölkerung.

Für uns Arbeitnehmer heutzutage ist, laut Bundesarbeitsgericht (Az.1 ABR 75/83), das Hören von Musik am Arbeitsplatz grundsätzlich erlaubt, sofern es die Arbeit nicht beeinträchtigt und auch sonst niemanden stört. Der Arbeitgeber kann ein Verbot aussprechen, das aber unverhältnismäßig wäre, wenn weder die konzentrierte Arbeit noch der Kundenverkehr gestört wäre und sich andere Kollegen nicht belästigt fühlen.

Letztlich bedarf es für das Musikhören am Arbeitsplatz grundsätzlich aber der Erlaubnis des Arbeitgebers. Um ein Radioverbot zu verhängen, bedarf es guter Gründe. Für viele ist Radiohören ein Quell der Freude und Inspiration, für andere die pure Qual. Schallberieselung am Arbeitsplatz spaltet.

Dank moderner Technik muss nun aber niemand mehr offen Radio über Lautsprecher hören. Kopfhörer, einerlei ob In-Ear oder Over-Ear, machen es möglich. Doch egal ob Radio oder Kopfhörer, die Musik darf nicht ablenken, die Arbeit in keiner Weise darunter leiden. ■

⁵ R.J.W., Lautsprecher in den Betrieben, in: Der Deutsche Rundfunk, Jg. 16, H. 38, Berlin 1938, S. 42-43; vgl. ferner Musterbeispiele für Übertragungsanlagen, in: Radio-Helios, Jg. 42, Nr. 12, 1936, S. 716.

Glosse

Nostalgische Zeitreise in den Newsroom und ins Medienarchiv von 2024

Eva Krause und Ute Mader

Ach, die guten alten Tage von 2024! Wenn wir heute an die damalige Arbeit in den Medienarchiven und im Newsroom denken, können wir uns ein Schmunzeln kaum verkneifen. Damals, als die Medienwelt noch auf den Spuren der digitalen Steinzeit wandelte und sich mit manuellen Abläufen und mühsamen Prozessen herumschlug. Ja, die Zeiten haben sich glücklicherweise gewaltig geändert.

Stellen Sie sich vor, liebe Leserinnen und Leser, wir befanden uns in einer Ära, in der Journalistinnen und Journalisten tatsächlich mit ihren Fingern auf Tastaturen herumhämmerten, um Nachrichten zu schreiben. Ja, Sie haben richtig gehört – Tastaturen! Die Redakteurinnen und Redakteure verbrachten Stunden damit, Informationen aus verschiedenen Quellen zusammenzusuchen, zu prüfen und zu verifizieren. Zusätzlich gaben sie Anfragen manuell oder sogar persönlich an die Recherneurinnen und Rechercheure der Medienarchive weiter. Und auch

diese führten wiederum die Recherchen fast ohne jegliche Automatisierung durch – Können Sie sich das vorstellen?! Künstliche Intelligenz war zwar schon ein Begriff, aber weit davon entfernt, die allgegenwärtige Informationsmethode zu sein, die sie heute ist.

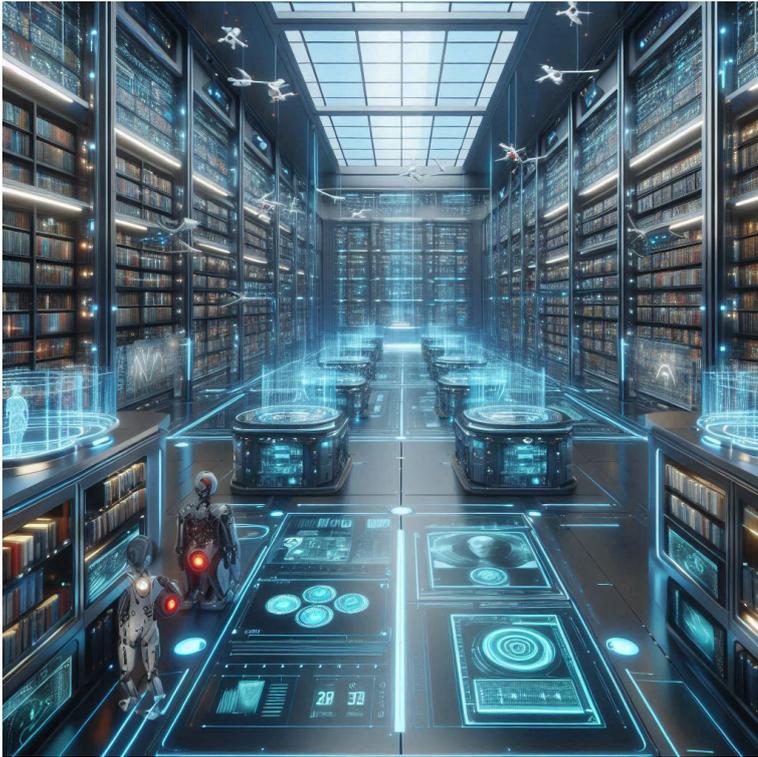
Im Newsroom herrschte damals ein ständiges Summen und Klappern, das von hektischen Gesprächen und dem gelegentlichen Begeisterungsschrei, wenn eine große Story hereinkam, durchdrungen war. Die Redakteurinnen und Redakteure liefen hektisch von Schreibtisch zu Schreibtisch, oft mit zerzausten Haaren und einem Kaffeebecher in der Hand. Kaffee war damals das Lebenselixier der Medienwelt und ein unverzichtbarer Begleiter durch den Nachrichtenalltag.

Und dann waren da noch die Medienarchive. Ein wahrer Hort des Wissens, in dem jede Ausgabe, je-



Ute Mader
ute.mader@info7.de
Eva Krause
eva.krause@info7.de





des Bild, jedes Audio und jedes Video sorgfältig archiviert und dokumentiert wurde. Ach, das alte Handwerk war schon etwas Besonderes. Natürlich, man setzte bereits auf einige KI-Tools, aber diese waren meist isolierte Anwendungen, die selten nahtlos zusammenarbeiteten. Für einige Mitarbeitende war die Technik eine unüberwindbare Hürde und ohne die Unterstützung von Programmieren sowie Informatikerinnen und Informatikern ging oft nichts. Recherchen bedeutete ein stundenlanges Durchsuchen von Daten und Sichten von Material. Selbst dann war man oft auf sich allein gestellt oder musste sogar in persönlichen Kontakt mit einer Kollegin oder einem Kollegen treten, um eine zweite Meinung zu erhalten. Dann gab es aber natürlich auch noch die eher fortschrittsgläubigen Mediendokumentarinnen und Mediendokumentare, die regelmäßig Tagungen veranstalten – auch zum Schwerpunktthema KI. Bei diesen Veranstaltungen wurde sich über neue Arbeitsmethoden, Workflows und Techniken ausgetauscht.

Man stellte sich also der Herausforderung, mit einer Vielzahl von Tools, Programmen und Algorithmen zu arbeiten, die oft nur mit Expertenwissen richtig zu bedienen waren. Man nahm sogar an der ein oder anderen Tool-Party teil. Der ständige Balanceakt zwischen technischer Komplexität und journalistischer Präzision war eine Kunst für sich. Viele Rechercheurinnen und Rechercheure fühlten sich

mehr und mehr wie Techniker, die sich regelmäßig mit neuen Werkzeugen auseinandersetzen und diese bedienen mussten, anstatt ihrer eigentlichen Berufung nachzugehen.

Heute, 50 Jahre später, sieht die Welt ganz anders aus. Die Journalistinnen und Journalisten arbeiten Hand in Hand mit KI-Assistenten, die innerhalb von Sekunden alle relevanten Informationen zusammenstellen und analysieren. Die Tastatur hat dem Gedanken-zu-Text-Interface Platz gemacht, bei dem unsere Gedanken direkt in Worte umgewandelt werden. Der Newsroom ist ein virtueller Ort geworden, dominiert von holographischen Displays und Redaktionskonferenzen, die über die Kontinente hinweg sprachunabhängig stattfinden.

Das Medienarchiv hat sich zu einem selbstlernenden, allwissenden System entwickelt, das jeden noch so kleinen Informationsschnipsel sofort parat hat. Kein mühsames Suchen mehr, keine verlorenen Stunden in endlosen Datenbanken. Alles ist nur einen Gedanken entfernt. Die Tätigkeit der Medienarchivarinnen und Medienarchivare ist die mentale Steuerung der Aufbewahrung und Bewertung des Weltwissens. Man konzentriert sich mehr auf Inhalte und Steuerungsstrategien als auf die händische Dokumentation und Recherche wie früher.

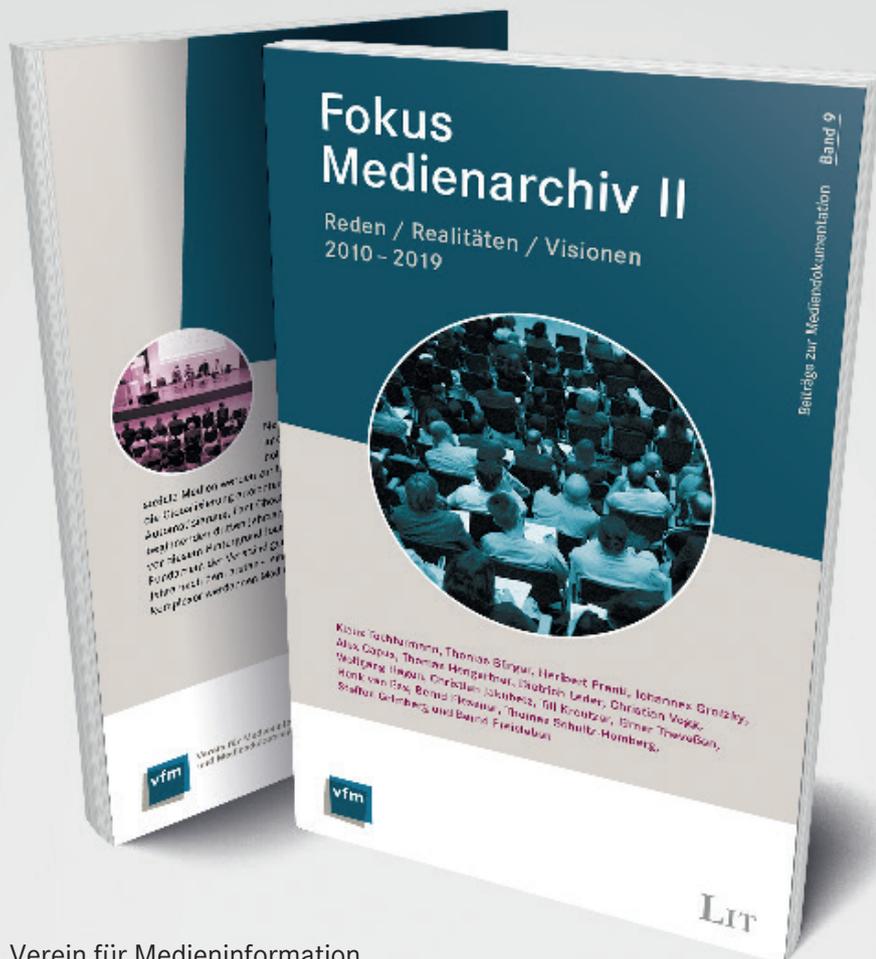
Ja, wenn wir an die Arbeit von damals zurückdenken, streift uns auch ein Hauch von Nostalgie. Es war eine Zeit der Hektik, des Adrenalins und der unglaublichen Befriedigung, eine Sendung erfolgreich abgeschlossen zu haben.

Ach, 2024 – was für eine Zeit! Ein Stück Magie!

Eva und Ute (aus der Zukunft) ■■■



Impressionen von der vfm-Frühjahrstagung 2024 in Freiburg. Die Bilder sind dankenswerter Weise größtenteils von Birgit-Caroline Grill (SWR, Stuttgart) aufgenommen worden. Die vollständige Tagungs- und Fotodokumentation ist für vfm-Mitglieder und Tagungsteilnehmer über die vfm-Homepage erreichbar.



Verein für Medieninformation
und Mediendokumentation (Hg.)

**Fokus Medienarchiv
Reden / Realitäten / Visionen
2010 - 2019**

224 S., 19,90 €, br., ISBN 978-3-643-10697-1

Neue Thesen geistern durch die Medien: Internetaktivitäten befördern wegen des hohen Energieverbrauchs den Klimawandel, soziale Medien werden zur Brutstätte von Populismus und Verrohung, die Globalisierung erleichtert die Ausbreitung von

Pandemien – Automatisierung, Fact Checking und Big Data sind die Stichworte im beginnenden dritten Jahrzehnt der 2000er Jahre. Wo positionieren sich vor diesem Hintergrund Journalismus und Mediendokumentation als Fundament der Verständigung?

Der vorliegende Band leistet – zehn Jahre nach dem ersten – einen Beitrag zur Orientierung in der immer komplexer werdenden Medienwelt.

LIT Verlag Berlin – Münster – Wien – Zürich – London

Auslieferung: LIT Verlag, Fresnostr. 2, D - 48159 Münster, E-Mail: vertrieb@lit-verlag.de

Impressum

info7 Das Magazin für Medien, Archive und Information
Jahrgang 39, Heft 2/2024
ISSN 0930-5483

Herausgeber:

Vorstand des vfm (Verein für Medieninformation und Mediendokumentation e. V.)

Redaktion:

Uta Rosenfeld (Redaktionsleitung, Layout und Satz), Dr. Ute Essegern, Eva Krause, Ute Mader, Dr. Sieglinde Osang, Thimo Kremser, Hans-Gerhard Stülb.
Mail: redaktion@info7.de

Titelbild: Erstellt mit Copilot (KI DALL.E), Ute Mader

Fotos der Frühjahrstagung: Birgit-Caroline Grill

Konzeption Gestaltung:

Carolin Diekmeyer

Ruf nach Autoren:

Manuskripte und Besprechungsstücke an die Redaktion sind erbeten. Für unverlangte Manuskripte wird keine Haftung übernommen.

Rechtshinweis:

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Die Speicherung oder Aufbewahrung einzelner Artikel in Dokumentationsstellen und Archiven zum Zwecke interner Informationsversorgung ist gestattet.

Die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehendung, im Internet oder auf ähnlichem Weg bleiben vorbehalten.

Druck, Verlag und Anzeigenannahme:

LIT Verlag Dr. W. Hopf, Fresnostraße 2, 48159 Münster, Tel: (0251) 6203214, Fax: (0251) 9226099, E-mail: lit@lit-verlag.de.
Bestellungen und Reklamationen zu Abonnements richten Sie bitte per mail an abo@lit-verlag.de. In der Betreffzeile bitte „info7“ angeben.

Bezugsbedingungen:

Erscheinungsweise 3 mal jährlich. Einzelpreis: 7,50€. Das Abonnement (3 Ausgaben pro Jahr): 21,50€. Bestellungen über den Buchhandel und den Verlag. Bestellungen zum Sonderpreis von 3,75€ können beim Verlag von Mitgliedern archivarischer, dokumentarischer, bibliothekarischer und informationswissenschaftlicher Verbände unter Hinweis auf die entsprechende Mitgliedschaft vorgenommen werden. Für diese kostet das Abonnement 10,75 € pro Jahr.