

MARCONI – KI UNTERSTÜTZT RADIOPRODUKTION

WERNER BAILER, ALEXANDRU STAN

Das 2017 gestartete europäische Projekt MARCONI (Multimedia and Augmented Radio Creation: Online, iNteractive, Individual) hat zum Ziel, die Interaktion zwischen Radiogestaltern und Konsumenten zu intensivieren, und gleichzeitig dem Produktionsteam den Überblick über die verschiedenen Kommunikationskanäle sowie die Integration von Konsumentenbeiträgen in die Programmgestaltung zu erleichtern. KI-basierte Technologien spielen eine wichtige Rolle, um den Automatisierungsgrad in diesem Prozess zu erhöhen: Chatbots beantworten Standardanfragen, direkte und Social Media-Nachrichten werden automatisch mit Programmelementen verknüpft, und Bilder sowie Videos können automatisch analysiert und klassifiziert werden. Durch die Integration dieser Dienste in ein Radioredaktionssystem erhöht MARCONI die Effizienz und Qualität der Radioproduktion. Nach erfolgreichen internen Tests bei NPO und VRT und Pilotversuchen bei externen Radiostationen wurde das Projekt im März 2020 abgeschlossen.

► The European project MARCONI (Multimedia and Augmented Radio Creation: Online, iNteractive, Individual), started in 2017, aims at facilitating the integration between radio creators and consumers, while at the same time improving the overview of the production team over the different communication channels and the integration of user feedback into the programme. AI technologies play an important role in improving the level of automation in this process: chatbots help to answer standard questions, messages from direct feedback or social media are automatically associated with programme elements, and images and videos are automatically analysed and classified. With the integration of these services into the radio editorial systems, MARCONI increases the efficiency and quality of radio production. After internal testing at NPO and VRT, and piloting with external radio stations the project has been concluded in March 2020.

Radiogestaltung in Zeiten sozialer Medien

Die Medienlandschaft ist einem erheblichen Wandel unterworfen, der einerseits durch die Konvergenz zwischen verschiedenen Medien und durch zunehmend vielfältige (in

Bezug auf Zeit, Ort, Kontext und Inhalt) und personalisierte Formen des Medienkonsums charakterisiert ist. Diese Veränderungen wirken sich auch auf das Medium Radio aus, das seine Popularität über ein Jahrhundert hinweg trotz heftiger Konkurrenz behaupten konnte. Der Anteil der Europäer, die mindestens einmal pro Woche Radio hören, liegt stabil bei 75 Prozent, und die Zahl der täglichen Konsumenten liegt bei 50 Prozent [1]. Während die primäre Stärke des Rundfunks immer noch die Tatsache ist, dass er eine gemeinsame Live-Erfahrung für die Zuhörer ermöglicht, sind Radiosendungen heute viel mehr als nur ein linearer Audiodatenstrom. Radio dreht sich zunehmend darum, Konsumenten über eine Vielzahl von Kanälen zu erreichen und zu binden, einschließlich der Interaktion auf Mobilgeräten, der direkten Kommunikation (Telefon, E-Mail, Instant Messaging), sozialer Medien und visuellem Radio. Damit vervielfältigen sich nicht nur die Kommunikationskanäle, über die der Kontakt mit den Zuhörern gehalten wird und die als Quellen möglicher Inhalte für die Sendungsgestaltung dienen, sondern wird auch die Bereitstellung von ansprechenden Zusatzinhalten in Form von Text und Multimedia zur Notwendigkeit.

Diese erfordert, dass Radioredaktionen und -präsentatoren eine Vielzahl von Kommunikationskanälen beobachten und verschiedene Plattformen mit passenden Inhalten bespielen müssen. Oft sind dafür verschiedenste, nicht oder nur schlecht miteinander integrierte Systeme (Redaktions- und Playoutsysteme, Social Media Dashboards, Telefonsysteme, etc.) im Einsatz. Die Folge ist einerseits eine hohe Belastung des Teams, und andererseits ein nicht optimales Ergebnis: nicht alle Beiträge von Konsumenten können gesichtet werden, womit unter Zeitdruck eine oberflächliche Auswahl getroffen werden muss, dieselben Inhalte werden wiederverwendet, und viele Konsumenten bekommen keine (zeitnahe) Reaktion auf ihre Interaktion.

Das Projekt MARCONI (Multimedia and Augmented Radio Creation: Online, iNteractive, Individual)¹⁾ hat sich zum Ziel gesetzt, diese Situation zu ändern. MARCONI soll interaktives und personalisiertes Radio durch die Integration von Live-Radio mit digitalen Plattformen ermöglichen, um die Möglichkeiten der Konvergenz zu nutzen. Eine wesentliche Rolle kommt dabei der Unterstützung der redaktionellen Arbeit durch Automatisierung zu, insbesondere durch Technologien der Künstlichen Intelligenz (KI). Das Projekt, das von der Europäischen Union im Rahmen einer Horizon 2020 Innovation Action gefördert wird, wird vom Flämischen Rundfunk (VRT) koordiniert. Weitere Partner sind der Niederländische öffentliche Rundfunk NPO, der Hersteller von Radioredaktionssystemen Pluxbox, der KI-Lösungsanbieter FAKTION, der Content Management Systemanbieter IN2, die Universität Hasselt, die Forschungseinrichtung JOANNEUM

¹⁾ <https://www.projectmarconi.eu/>

RESEARCH und die Universität Wien, die das Projekt in datenschutzrechtlichen Belangen begleitet.

Dieser Artikel beleuchtet zunächst die verschiedenen Einsatzmöglichkeiten von KI im Anwendungsbereich Radio und deren Umsetzung in MARCONI und gibt dann einen Ausblick.

Möglichkeiten zum Einsatz von KI Metadatengenerierung

Eine wesentliche Einsatzmöglichkeit von KI in der Radioproduktion (und in der Medienproduktion im Allgemeinen) ist die automatische Erzeugung von Metadaten für audiovisuelle Inhalte.

Metadaten auf größerer Granularität (zum Beispiel für einen gesamten Clip) und auf höherer semantischer Ebene (zum Beispiel Thema, wesentliche Aussage) lassen sich noch einigermaßen effizient manuell erzeugen. Die große Stärke automatischer Verfahren liegt hingegen in der Erzeugung von detaillierten und segment-bezogenen Metadaten, wie der Detektion von Objekten und deren Position, der Beschreibung von Kamerabewegung, der Erkennung von Gesichtern oder der Transkription des gesprochenen Texts. Detaillierte Metadaten sind beispielsweise notwendig, um interaktive Inhalte zu gestalten, oder existierende Inhalte für verschiedene Plattformen verfügbar zu machen, ohne dass damit eine Vervielfachung der Produktionskosten einhergeht.

Durch den Einsatz von neuronalen Netzwerken mit vielen Layern, von denen vor allem die frühen Layer aus Faltungsoperationen bestehen (sogenannte Deep Convolutional Neural Networks²⁾) wurden in den letzten Jahren insbesondere für Anwendungen in der visuellen Analyse große Fortschritte erzielt. Für viele Aufgaben ist die Qualität der automatischen Ergebnisse in einem Bereich, der die Anforderungen für den produktiven Einsatz erfüllt. Das Training solcher tiefen neuronalen Netzwerke erfordert jedoch große Datenmengen. Diese sind zum Beispiel in der Gesichtserkennung oft nicht verfügbar, insbesondere, wenn es sich um Personen von lokaler oder regionaler Bekanntheit handelt, wie es für Radiostationen vielfach relevant ist. Im Projekt wurde daher ein Ansatz entwickelt, um mit einigen wenigen Beispielbildern eine Erkennung zu ermöglichen, und weitere Personen und Beispielbilder später hinzufügen zu können, und dadurch die Robustheit zu erhöhen.

Während für redaktionelle Inhalte zumindest grundlegende Inhalte verfügbar sind, ist für benutzergenerierte Multimedia-Inhalte nur der Kontext bekannt, in dem sie gepostet wurden (und dieser kann absichtlich oder unabsichtlich nicht zum Inhalt passen). Daher ist für Bilder und Videos, die von Benutzern über direkte Kanäle oder soziale Medien gesandt werden, die automatische Extraktion von Metadaten von großer Bedeutung (siehe Abbildung 1). Dies umfasst die Detektion von Objekten, Orten oder Personen, um eine thematische Zuordnung zu treffen (zum Beispiel ob das Bild eine Reaktion auf einen bestimmten Aufruf einer Radiosendung ist oder zu einem Sendungsthema gehört), oder die Erkennung von Duplikaten bzw. Feststellung der Ähnlichkeit von Bildern, um sie zu gruppieren und der Redaktion die Arbeit zu erleichtern. Ein weiterer wichtiger Aspekt ist die technische Qualität, die manchmal bei benutzergenerierten Inhalten zu wünschen übriglässt. So können beispielsweise unscharfe, aufgrund schlechter Lichtverhältnisse verrauschte Bilder und Videos oder verwackelte Videos aussortiert



Abbildung 1: Beispiel für Objektdetektion in benutzergenerierten Inhalt.
Quelle: JOANNEUM RESEARCH / ICOSOLE Project

werden. Gleichzeitig kann dem Konsumenten Feedback gegeben werden, warum die Inhalte nicht verwendet werden konnten. Im Kontext von Musikfestivals wurde bereits demonstriert, dass Entscheidungen der Redaktion, von Konsumenten zur Verfügung gestelltes Material zu verwenden oder nicht, auch von Maschinen gelernt werden können [3].

Suche und Indizierung

Die Redaktionsteam einer Radiosendung bekommt regelmäßig tausenden Nachrichten über verschiedene Kanäle, wie zum Beispiel Radio-Apps, soziale Medien, Email, etc. Einige dieser Inhalte sind nicht nur zum Zeitpunkt, zu dem sie gesendet werden, relevant, sondern auch für zukünftige Sendungen. Deswegen ist es wichtig, dass all diese Inhalte so indiziert werden, sodass sie für die Redaktion griffbereit sind. Dafür ist es wichtig, dass die Inhalte mit Metadaten angereichert sind (zum Beispiel mittels der oben beschriebenen KI-Methoden). In MARCONI können alle Konversationen, Texte und Multimedia-Beiträge mühelos über eine Google-ähnliche Suchoberfläche durchsucht werden. Bei Bedarf kann die Suche über zusätzliche Parameter, zum Beispiel ein Zeitrahmen, eingeschränkt werden.

Apache Solr und Lucene sind sehr zuverlässige und skalierbare Technologien, die von den meisten Textsuchmaschinen verwendet werden. In MARCONI wurden sie benutzt, um einen Volltextindex und mehrere Indizes zum Filtern und Facettieren von Ergebnissen zu erstellen. Während dieses Prozesses werden einige Textverarbeitungsprozesse (Zerlegung in Wörter, Aufteilen von zusammengesetzten Wörtern, etc.), das Extrahieren von Erwähnungen (zum Beispiel @Benutzername) und Hashtags durchgeführt sowie die Anwendung von Stoppwortlisten, Synonymersetzungen etc. Der Volltextindex wird hauptsächlich für Suchen verwendet, während Facetten das Verfeinern der Ergebnisse durch besondere Merkmale ermöglichen. Der Benutzer kann sowohl mittels einfacher Abfragen (zum Beispiel Schlüsselwörter, Benutzernamen, Hashtags) oder mittels komplexer Abfragen (zum Beispiel Boolesche Operatoren, Verschachtelungen, Platzhalter, Facetten) suchen.

Chatbots

Konversationsbasierte Benutzerschnittstellen gewinnen zunehmend an Bedeutung, sei es für den Support auf Websites, um Kommunikation über Messaging Dienste zu ermöglichen oder für sprachgesteuerte Interaktion wie zum Beispiel mit Smart Speakers. Chatbots ermöglichen es Radiostationen, auf häufig gestellte Fragen automatisiert zu antworten, die

²⁾ https://de.wikipedia.org/wiki/Convolutional_Neural_Network



Quelle: NPO / MARCONI Project

Abbildung 2: Chatbot auf der NPO Radio 5 Website.

sich auf den Sender im Allgemeinen (zum Beispiel Moderatoren) oder auf ein bestimmtes Ereignis beziehen (zum Beispiel Programm, gespielte Musik, Veranstaltungen). Der Dienst ist ständig verfügbar und bietet den Benutzern die Möglichkeit, automatisiert zu Informationen zu gelangen.

Die MARCONI-Chatbots nutzen die KI-basierte Sprachverarbeitung, die derzeit die Sprachen Niederländisch, Französisch, Englisch und Deutsch unterstützt. Mit ihrer Hilfe können Absichten, Gefühle und Entitäten (zum Beispiel Personen, Organisationen, Orte) extrahiert werden. Für die Chatbots müssen Modelle für die Absichtserkennung und optional für die Entitätsextraktion trainiert werden. Eine Absicht bezeichnet eine bestimmte Aktion, die ein Benutzer ausführen kann. Eine Absicht kann zum Beispiel eine Frage, eine Aussage, eine Antwort auf eine Frage oder eine Begrüßung sein. Ausdrücke sind Beispielsätze für eine bestimmte Absicht, mit denen das Modell trainiert wird. Je mehr Ausdrücke zu einer Absicht hinzugefügt werden, desto genauer wird sie erkannt. Wird die Absicht nur mit geringer Zuverlässigkeit erkannt, kann gezielt nachgefragt und als letzte Maßnahme die Kommunikation an einen Menschen übergeben werden.

In typischen Chatbot-Applikationen wie der Bearbeitung von Kundendienst-Anfragen greifen Chatbots auf eine relativ statische Wissensdatenbank zurück. Eine Herausforderung in MARCONI war es, den Chatbot mit dem Redaktionssystem zu verknüpfen, und damit in Echtzeit aktuelle und detaillierte Informationen über das Programm, die DJs und die Playlists der Radiostation liefern zu können. Darüber hi-

naus wurde in MARCONI auch eine Integration für Umfragen geschaffen, wie sie häufig von Radiostationen gemacht werden. Die Umfragen können im Redaktionssystem definiert werden, und Antworten, die über verschiedene Kanäle ankommen und der Umfrage zugeordnet werden können, werden automatisch ausgewertet. Die Chatbots können in die Website (siehe Abbildung 2) oder App einer Radiostation eingebunden werden, aber auch mit anderen Diensten wie zum Beispiel Facebook Messenger integriert werden, um den Konsumenten über deren bevorzugte Kommunikationskanäle zur Verfügung zu stehen.

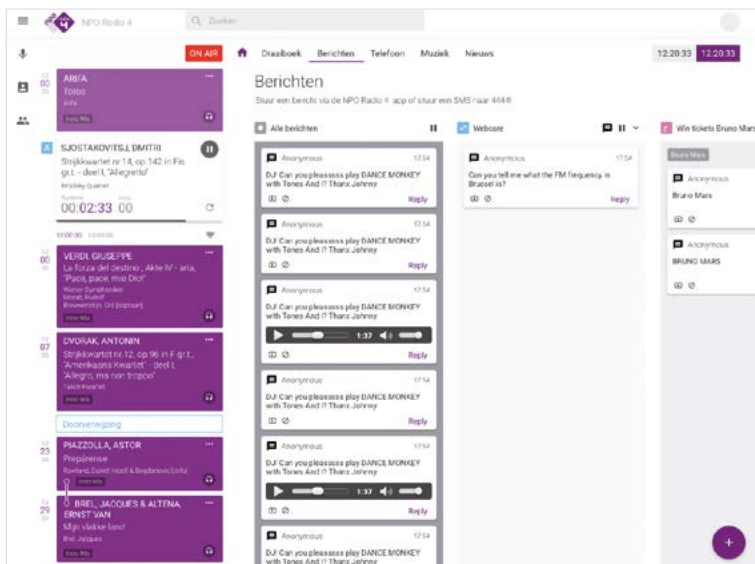
Umsetzung im Projekt MARCONI

Um diese KI-Methoden für die Radioproduktion nutzbar zu machen, wurde eine Microservice-Architektur gewählt. Der Hauptgrund liegt in der Flexibilität: jede Radiostation benötigt eine andere Teilmenge der Services, und die Workflows unterscheiden sich ebenfalls beträchtlich. Das trifft auch auf verschiedene Stationen einer Rundfunkanstalt zu, wie sich bei den Tests mit den VRT Stationen Studio Brussel und MFM sowie NPO 3FM und Radio 5 gezeigt hat. Diese Unterschiede ergeben sich durch unterschiedliche Gewichtung von Programminhalten (zum Beispiel Musikanteil, Diskussionen zu aktuellen Themen mit Hörerbeteiligung), aber auch durch die demographische Zusammensetzung des Publikums (zum Beispiel Konsum über Mobilgeräte, Nutzung von Messaging Diensten).

MARCONI Plattform

Die MARCONI-Plattform dient als Backend für das Redaktionssystem RadioManager³⁾ und steuert die verschiedenen Microservices (KI-basierte und andere). Die Datenspeicherung erfolgt mittels ArangoDB⁴⁾ und RocksDB⁵⁾, wobei GraphQL⁶⁾ als Schnittstelle zu allen anderen Komponenten dient. Das betrifft nicht nur das Einfügen, Aktualisieren und Abfragen von Daten, sondern auch Steuerfunktionen. Über den GraphQL-Subscription-Mechanismus werden Dienste ereignisbasiert (zum Beispiel Eintreffen einer neuen Nachricht über einen bestimmten Kommunikationskanal) aufgerufen.

Die Plattform deckt auch zentrale Funktionalitäten wie Authentifizierung ab. Eine Besonderheit der MARCONI-Plattform ist der PrivVaults-Datencontainer, der die Ablage von benutzerbezogenen Daten nach dem Prinzip „Privacy by Design“ ermöglicht. Für bestimmte Elemente in der Datenbank werden Berechtigungen definiert, die mit den im System definierten erlaubten Verarbeitungsvorgängen für einen Benutzer abgeglichen werden. Die Definition, welche das sind, basiert entweder auf explizit von Benutzern erteilten Zustimmungen (solche Zustimmungen können auch automatisiert über Chatbots eingeholt werden), oder Vorgaben des Datenschutzbeauftragten (zum Beispiel wenn es eine gesetzliche Grundlage dafür gibt). Entwickler von Diensten und Applikationen können dieses Konzept beim Zugriff auf Daten nicht aushebeln, und der Widerruf von Zustimmungen führt automatisch zur Sperre des Zugriffs auf die Daten. Somit bleibt die Kontrolle bei den Datenschutzverantwortlichen und darüber hinaus werden alle Zugriffe auf Benutzerdaten auch protokolliert.



Quelle: Pluxbox / MARCONI Project

Abbildung 3: RadioManager als zentrale Benutzerschnittstelle.

³⁾ <https://pluxbox.com/>

⁴⁾ <https://www.arangodb.com/>

⁵⁾ <https://rocksdb.org/>

⁶⁾ <https://graphql.org/>

Microservices

Die weiter oben beschriebenen KI-Funktionalitäten werden als Microservices bereitgestellt, und über eine Middleware-Komponente an die MARCONI-Plattform angebunden. Diese Middleware stellt die Verbindung zwischen der GraphQL-Schnittstelle und den jeweiligen APIs der KI-Dienste (typischerweise REST-basierte APIs) her.

Die Microservices werden als verteiltes System ausgerollt. Sowohl die Plattform als auch einzelne Services können entweder auf der Infrastruktur der Radiostation als auch in der Cloud bereitgestellt werden. Die auf KI-Algorithmen basierenden Services profitieren dabei stark von hochparallelen Prozessorarchitekturen, wie zum Beispiel Grafikprozessoren (GPUs). Manche der KI-Dienste, wie zum Beispiel die Verarbeitung von Multimedia-Inhalten, die von Konsumenten beigetragen werden, sind starken Auslastungsschwankungen unterworfen. Während die Anzahl solcher Inhalte für viele Radiostationen im Allgemeinen eher gering ist, steigt die Anzahl solcher Nachrichten nach einem Aufruf oder während einer Veranstaltung extrem stark an. Daher ist die Skalierbarkeit von Cloudinfrastrukturen vorteilhaft, um diese kurzfristigen Spitzen abdecken zu können.

Benutzerschnittstelle

Um eine effiziente Bedienung des Systems zu ermöglichen, wurde das Radio-Redaktionssystem RadioManager als Basis verwendet und erweitert. RadioManager ist eine webbasierte Lösung (siehe Abbildung 3), um Planung, Organisation und Ablauf von Radiosendungen zu verwalten, und bietet bestehende Integrationen, zum Beispiel mit Musikdatenbanken und Playoutsystemen. Verschiedenste Kommunikationskanäle können auf Basis der Ergebnisse aus MARCONI integriert werden, die Nachrichten aus verschiedenen Quellen und zusammen mit automatisch extrahierten Metadaten anzeigen. Diese Nachrichten können durchsucht werden, oder nach definierten Filterkriterien organisiert werden. Per Drag&drop können Nachrichten mit Sendungsinhalten assoziiert werden, Multimedia-Inhalte für Visual Radio verwendet werden,

oder Nachrichten gruppiert beantwortet werden. Das Layout kann an verschiedene Rollen im Team angepasst werden, um zum Beispiel der Redaktion einer Sendung den Überblick zu allen mit dem Sendungsthema verbundenen Informationen zu geben, oder dem Moderator den Fokus auf den Sendungsablauf und Nachrichten im Kontext zum aktuellen Beitrag zu geben.

Ausblick

Die im Projekt entwickelten Technologien wurden in etlichen Tests bei den beteiligten Rundfunkanstalten auf ihre Praxistauglichkeit überprüft. So wurden beispielsweise bei NPO positive Erfahrungen mit dem Einsatz von Chatbots bei Radio 5 und mit einem "Studio Messenger" genannten Nachrichten-Dashboard gesammelt. Sowohl bei NPO 3FM als auch bei VRT Studio Brussel kamen Chatbots und Austausch von Multimedia-Inhalten über soziale Medien bei karitativen Kampagnen zum Einsatz. In der finale Pilotphase des Projekts, das im März 2020 abgeschlossen wurde, wurden Technologien auch bei Radiostationen außerhalb des Projektkonsortiums getestet. Die Ergebnisse fließen in die aktuelle Version von RadioManager ein bzw. werden als optionale Erweiterungen angeboten. ➤

Referenzen

- [1] Media use in the European Union. Tech. rep., Standard Eurobarometer 88 - Wave EB88.3 (Nov. 2017)
- [2] Martin Winter, Werner Bailer, "Incremental Training for Face Recognition": In 25th International Conference on MultiMedia Modeling, Thessaloniki, GR, Jan. 2019.
- [3] Werner Bailer, Martin Winter, Stefanie Wechtitsch, "Learning Selection of User Generated Event Videos," In: Workshop on Content-based Multimedia Indexing, Florenz, IT, Jun. 2017
- [4] Werner Bailer, Maarten Wijnants, Hendrik Lievens, Sandy Claes, "Multimedia Analytics Challenges and Opportunities for Creating Interactive Radio Content," In 26th International Conference on MultiMedia Modeling, Daejeon, KR, Jan. 2020.

Bild: JOANNEUM RESEARCH



WERNER BAILER

ist Key Researcher zum Thema Smart Media Solutions bei JOANNEUM RESEARCH in Graz.

➤ <https://www.joanneum.at/>

Bild: IN2



ALEXANDRU STAN

ist Innovation Manager bei IN2 Digital Innovations GmbH in Lindau, Deutschland.

➤ <https://in-two.com>