

Onur Öksüz

**Masterarbeit
im Fach Information Systems**

**Radiohörerlebnis im personalisierten Radio –
Analyse und Simulation klassischer Ausspielreihenfolgen**

Themasteller: Univ.-Prof. Dr. Detlef Schoder

Vorgelegt in der Masterprüfung
im Studiengang Information Systems
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität zu Köln

Köln 2021

Inhaltsverzeichnis

Abbildungsverzeichnis.....	III
Tabellenverzeichnis.....	IV
Abkürzungsverzeichnis.....	V
1 Einleitung	1
1.1 Motivation.....	1
1.2 Zielsetzung.....	3
1.3 Aufbau der Arbeit.....	4
2 Theoretischer Hintergrund.....	5
2.1 Das Radio als Service	5
2.2 Sequenzmusteranalyse	9
3 Qualitative Inhaltsanalyse	11
3.1 Methodik.....	11
3.2 Auswahl und Beschreibung der Daten.....	12
3.3 Aufstellen des Codingframes	15
3.4 Bestimmung der Analyseeinheiten	19
3.5 Kodierung und seine Validität.....	19
3.6 Verrechnung und Resultat	20
3.6.1 Ergebnisse der Klassen	21
3.6.2 Ergebnisse der Bezugsreichweite.....	22
3.6.3 Ergebnisse der Gattungen.....	23
3.6.4 Ergebnisse der Tageszeit	25
3.7 Anforderungen für die Simulation	27
4 Sequenzmusteranalyse.....	29
4.1 Vorbereitung.....	29
4.2 Anwendung der Sequenzmusteranalyse	32
4.2.1 Senderübergreifender Fokus.....	32
4.2.2 Kontextueller Fokus.....	35
5 Simulation von Ausspielreihenfolgen.....	43
5.1 Vorstellung des Algorithmus.....	43
5.2 Evaluierung der Methode	46
6 Diskussion und Handlungsempfehlungen	51
7 Ausblick und Fazit	53

Literaturverzeichnis.....	59
A. Anhang.....	62
B. Anhang.....	74
C. Anhang.....	75
D. Anhang.....	77
Eidesstaatliche Erklärung.....	80
Einverständniserklärung zur Plagiatsprüfung.....	81
Lebenslauf.....	82

Abbildungsverzeichnis

Abbildung 1: Das personalisierte Radio	4
Abbildung 2: Eine sequenzielle Datenbank.....	10
Abbildung 3: Beispiel-Sendeplan für den Sender DLF Nova	15
Abbildung 4: Musteranalyse der Länge 4 nach Klassen (Minsup: 1%)	33
Abbildung 5: Musteranalyse der Länge 4 nach Bezugsreichweite (Minsup: 1%)	34
Abbildung 6: Musteranalyse der Länge 4 nach Gattungen (Minsup: 0.5%)	35
Abbildung 7: Ähnlichkeit der Unterpatterns nach Klassen (Minsup: 1%)	36
Abbildung 8: Klassen-Muster DLF Kultur, tagsüber, Wochentag.....	36
Abbildung 9: Klassen-Muster DLF Nova, morgens, Wochenende	37
Abbildung 10: Ähnlichkeit der Unterpatterns nach Geographika (Minsup: 1%)	38
Abbildung 11: Geographika-Muster DLF, tagsüber, Wochentag	38
Abbildung 12: Geographika-Muster DLF Nova, morgens, Wochenende	39
Abbildung 13: Ähnlichkeit der Unterpatterns nach Gattungen (Minsup 1%).....	39
Abbildung 14: Gattung-Muster DLF, nachts, Wochenende	40
Abbildung 15: Gattung-Muster, DLF Kultur, nachts, Wochentag.....	41
Abbildung 16: Vor- und Nachteile der Oberkategorien für die Simulation	43
Abbildung 17: Pseudocode für die Simulation synthetischer Ausspielreihenfolgen .	46
Abbildung 18: Evaluierung der Methode (Minsup echtes Radio: 0,005).....	49
Abbildung 19: Evaluierung der Methode (Minsup echtes Radio: 0,007).....	51

Tabellenverzeichnis

Tabelle 1: Übersicht der Radiomitschnitte	14
Tabelle 2: Codingframe - Klasse	16
Tabelle 3: Codingframe - Bezugsreichweite	17
Tabelle 4: Codingframe - Gattung.....	18
Tabelle 5: Codingframe - Tageszeit.....	19
Tabelle 6: Die wichtigsten Spalten der Kodierungstabelle.....	20
Tabelle 7: Beobachtete Klassen.....	22
Tabelle 8: Beobachtete Bezugsreichweite	23
Tabelle 9: Beobachtete Gattung	24
Tabelle 10: Beobachtete Tageszeit.....	27
Tabelle 11: Anforderungen an die Simulation	29

Abkürzungsverzeichnis

DLF	Deutschlandfunk
SDB	Sequenzielle Datenbank
SDL	Service Dominant Logic

1 Einleitung

1.1 Motivation

Es ist kein Geheimnis, dass die Digitalisierung in jedem Bereich unseres Lebens immer mehr an Bedeutung gewinnt (Hirschmeier, Beule, & Tilly, 2020). Digitalisierung ist definiert als die Nutzung von Technologie, um die Leistung oder die Reichweite von Unternehmen weitgehend zu verbessern (Hirschmeier, Tilly, & Beule, 2019). Apps für das Smartphone, die früh am Morgen die Frontscheibe des Autos beheizen oder modernste Supermärkte wie Amazon Go, die ein unsichtbares Kassensystem besitzen, wo Kameras mithilfe spezieller Algorithmen erkennen, welche Person welches Lebensmittel dem Regal entnimmt, sind nur kleine Beispiele hierfür. Diese Möglichkeiten waren vor einigen Jahren noch undenkbar. Die Automatisierung und Personalisierung von altbekannten Systemen ermöglicht es uns, Dinge schneller, einfacher und bequemer zu erledigen.

Auch Radiounternehmen sind sich der Wichtigkeit der Digitalisierung bewusst und bekommen vor allem immer mehr Konkurrenz aus anderen Bereichen (Hirschmeier et al., 2019, p. 2). Insbesondere Musik-Streaming-Dienste wie Spotify und Deezer gewinnen immer mehr an Nutzern und sind besonders bei jungen Leuten sehr beliebt (Statista, 2020). Diese Anbieter nutzen die Digitalisierung in vollem Zuge und bieten dem Nutzer automatisierte und personalisierte Inhalte, wie beispielsweise folgende Playlists: „Dein Mix der Woche“, „Dein Song des Jahres“ etc. Diese basieren auf der Interaktion des Nutzers mit dem System. Wie lang und wie oft hat er welchen Song gehört? Welche Songs hat er eher übersprungen? Basierend auf diesen Daten werden ähnliche Songs gesucht, die den Musikgeschmack des Nutzers treffen sollen. Musik-Streaming-Dienste haben hier zwei Vorteile gegenüber dem Radio. Sie sind im digitalen Zeitalter entstanden und haben sich somit schon früh in ihrer Entstehung mit digitalen Themen wie beispielsweise der Interaktion mit dem Nutzer oder den Algorithmen zur Erkennung von Interessen der Nutzer, beschäftigt. Außerdem bieten sie nur homogene Inhalte, nämlich Musik. Umso wichtiger ist es für Radiosender, auf diesen Umschwung zu reagieren und ihr Angebot nachhaltig zu gestalten.

Vom altbekannten linearen Radio, welches man typischerweise im Auto hört, neue und digitalisierte Dienste zu entwickeln, gestaltet sich etwas schwieriger. Denn es ist vom Wesen speziell und unterliegt einer besonderen Struktur. Der Ablauf der

Sendungen ist vorgegeben und somit geplant. Zudem sind die Inhalte heterogen, sprich es existiert eine große Bandbreite an Themen und Gattungen der einzelnen Programmelemente. Außerdem gibt es speziell im Fall Radio noch Gesetze, an die sich öffentlich-rechtliche Sender halten müssen (Ministerium des Innern des Landes Nordrhein-Westfalen, Referat 14, 2020; Pöchhacker, Burkhardt, Geipel, & Passoth, 2017).

Bisher existiert nur wenig Literatur zu dieser Thematik (Hirschmeier et al., 2019, p. 2). Neben einem Mangel in der Literatur sind auch digitalisierungstechnische Umsetzungen von Radiounternehmen kaum oder unzureichend vorhanden. Beispielsweise werden online bereitgestellte Mediatheken, die das nachträgliche Abspielen vergangener Sendungen erlauben, nicht häufig und lang genutzt (Hirschmeier et al., 2020). Zudem sind nur wenige Radios mit einer stark ausgeprägten Personalisierung bekannt. Jedoch ist es für moderne Radiounternehmen wichtig, ihr Angebot stets zu optimieren, um ihre Hörer zu halten oder neue, insbesondere junge Hörer hinzuzugewinnen. Laut Jauert, Ala-Fossi, Föllmer, Lax und Murphy (2017) ist das klassische analoge Radio in Europa nach wie vor weit verbreitet und soll auch 2025 noch genutzt werden. Zu den Analogen zählt man in diesem Fall das AM Radio, welches zu Beginn sehr weit verbreitet war, sowie das UKW-Radio, auch bekannt als das FM Radio. Trotzdem existiert ein wachsender Trend zum digitalen Radio (Jauert et al., 2017, p. 24). Offensichtlich wurde in diese Richtung nicht viel getan, weswegen diese Arbeit den Schritt zur Digitalisierung begünstigt.

Radiounternehmen haben einen entscheidenden Mehrwert, den sonst kein vergleichbares Medium anbietet und welches sie von der Konkurrenz abhebt. Genauer gesagt geht es um den redaktionellen Mehrwert, welcher hier als der menschliche Beitrag zur Gestaltung attraktiver Ausspielreihenfolgen definiert wird. Die Gestaltung, der Ablauf und die Reihenfolge der einzelnen Programmelemente einer Radiosendung sind von Redakteuren durchdachte Konstrukte. Redakteure beschäftigen sich in ihrem gesamten Berufsleben damit, welche Beiträge wann gespielt werden sollen. Das Ergebnis ist das lineare Radio, respektive ein fester Programmablauf basierend auf den langjährigen Erfahrungen der Redakteure. Sie haben Jahrzehnte lange Erfahrung und wissen, wie Sendungen gestaltet werden müssen, um möglichst viele Hörer zu generieren (La Roche & Buchholz, 2016). Daher ist es essenziell, dass beim neuartigen nichtlinearen Radio, was beispielsweise

eine personalisierte Radio App sein könnte, dieser redaktionelle Mehrwert Berücksichtigung findet. Genau hier wird die Analyse angesetzt. Es ergeben sich zwei Forschungsfragen: Welche Muster gibt es im linearen Radio? Wie kann man diese Muster nutzen, um das Radiohörerlebnis im nichtlinearen Radio zu optimieren?

1.2 Zielsetzung

Ziel der Arbeit ist es, Muster im linearen Radio zu erkennen, um diese dann vorteilhaft in neuartigen digitalen Diensten nutzen zu können. Vorteilhaft bedeutet, dass der redaktionelle Mehrwert aus dem linearen Radio auch im nichtlinearen Radio erhalten bleibt. Abbildung 1 zeigt das angestrebte Allgemeinziel. Das lineare Radio beinhaltet den redaktionellen Mehrwert. Durch die Erkennung der Muster ist es möglich, diesen Mehrwert zu erfassen, zu extrahieren und in das neue nichtlineare Radio zu integrieren. Die Kenntnis darüber, welche Art von Beiträgen zu welcher Zeit gespielt werden und welche Reihenfolge typisch ist, ermöglicht es, neue künstliche Sequenzen, basierend auf dieser Mustererkennung, zu erzeugen. Dies geschieht mit dem Radio Pattern Generator. Weiterhin sollte es ein Empfehlungssystem geben, welches basierend auf dem Feedback des Hörers handelt. Sprich wenn der Hörer einen Beitrag überspringt oder öfters wiederholt, wird diese Aktion vom System erkannt und für zukünftige Streams wiederverwendet. Das Empfehlungssystem ist allerdings nicht der Fokus dieser Arbeit. Die Abbildung dient lediglich dazu, zu zeigen, was das große Hauptziel ist. Konzentriert wird sich hier ganz auf die Erkennung von Mustern im linearen Radio und seine Verwendung, damit der redaktionelle Mehrwert im nichtlinearen Radio berücksichtigt wird. Genauer liegt der Fokus auf der Abfolge von einzelnen Beiträgen. Es soll herausgefunden werden, welche Themen und Arten von Beiträgen typischerweise aufeinanderfolgen. Diese Muster werden mithilfe von Algorithmen der Sequenzmusteranalyse erfasst und beim personalisierten Radio berücksichtigt. Diese Herangehensweise stellt einen ersten Schritt in die Digitalisierung des Radios dar. Wie gut der gewählte Ansatz ist, soll am Ende der Arbeit evaluiert werden. Ob andere Ansätze gegebenenfalls besser sind soll ebenfalls zum Schluss diskutiert werden.

Das Ergebnis der Arbeit soll somit die Erkennung von Mustern in Radiosendungen sein. Ebenfalls sollen Auffälligkeiten in den Häufigkeiten und in den Mustern identifiziert werden. Zuletzt soll ein Algorithmus erzeugt werden, der

dazu dient neue künstliche Sequenzen zu generieren. Der Algorithmus soll auf vorab identifizierte Muster, sowie auf einer Nutzereingabe basieren.

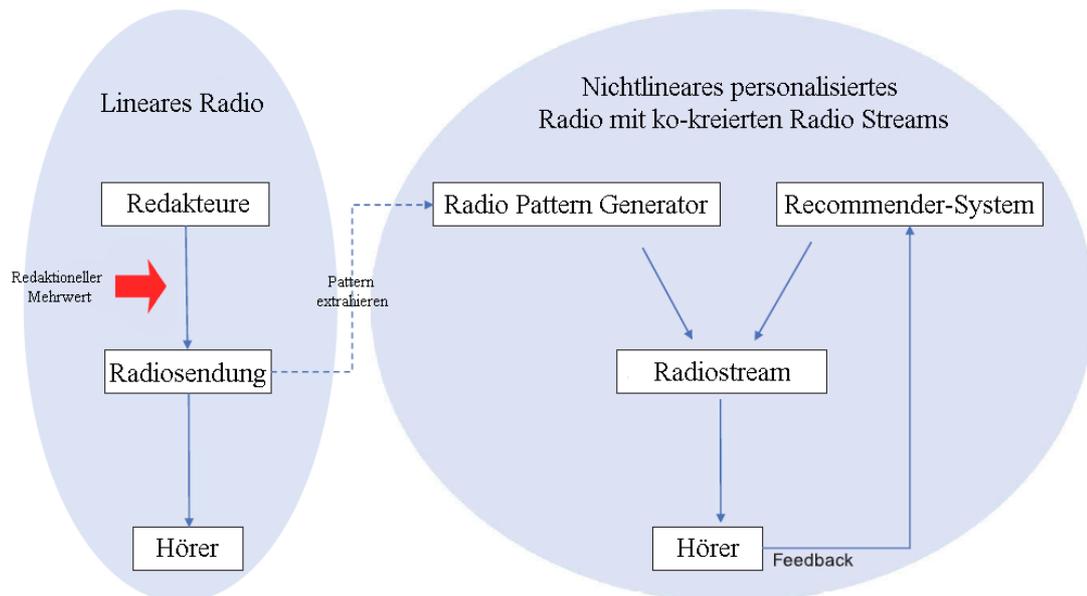


Abbildung 1: Das personalisierte Radio

In Anlehnung an: Hirschmeier et al. (2020)

1.3 Aufbau der Arbeit

Die Arbeit ist in sieben Kapiteln untergliedert. Nach der Einleitung folgt das zweite Kapitel, in dem die theoretischen Grundlagen erklärt werden. Dabei wird auf das Radio im Allgemeinen eingegangen. Außerdem folgt eine Erklärung des theoretischen Hintergrunds der Sequenzmusteranalyse. Das dritte, vierte und fünfte Kapitel bilden den Hauptteil der Arbeit. Nach einer kurzen Erklärung zur Methodik wird im dritten Kapitel basierend auf gegebenen Daten eine qualitative Inhaltsanalyse durchgeführt. Das vierte Kapitel setzt mit einer Sequenzmusteranalyse fort. Die Ergebnisse finden dann Berücksichtigung im fünften Kapitel. Dort folgt eine Simulation von Ausspielreihenfolgen basierend auf die Vorarbeit der qualitativen Inhaltsanalyse sowie der Sequenzmusteranalyse. Hier soll auch eine Evaluierung stattfinden, um die Frage zu beantworten wie gut der Algorithmus für die Simulation funktioniert. Die letzten zwei Kapitel beschäftigen sich mit der Diskussion und dem Fazit, in dem die wichtigsten Ergebnisse noch einmal aufgezeigt und die Grenzen der Arbeit sowie die Beurteilung der Zielerreichung beleuchtet werden.

2 Theoretischer Hintergrund

Vor der Anwendung der qualitativen Inhaltsanalyse sowie der Sequenzmusteranalyse, sollen erst grundlegende Begriffe geklärt werden. Dazu dient dieses Kapitel. Es folgt erst eine Definition des Radios. Dabei werden die Begriffe linear und nichtlinear erklärt. Ebenfalls soll vorhandene Literatur auf die Fragestellung gesichtet werden, ab wann man ein Hörerlebnis als gut betiteln kann. Ferner soll geklärt werden, warum die Thematik eine Relevanz für die Praxis hat. Im letzten Schritt dieses Kapitels wird der theoretische Hintergrund der Sequenzmusteranalyse geschildert.

2.1 Das Radio als Service

Der Begriff Radio hat verschiedene Definitionen und kann daher unterschiedliche Bedeutungen haben. Zu beachten ist auch ob die Rede vom Radiogerät oder vom Service an sich ist. In dieser Arbeit fokussieren wir uns auf das Radio als Service und definieren es als eine Mischung aus gesprochenen Wörtern sowie Musikstücken (Hirschmeier et al., 2020, p. 4). Zu unterscheiden ist das lineare und das nichtlineare Radio. Diese Begriffe hört man oft im Fernsbereich. Linear bedeutet, dass der Programmablauf fest ist. Hingegen bezeichnet nichtlinear einen nicht festen, also beeinflussbaren Programmablauf. Simple Beispiele für linearen Strukturen sind Angebote von öffentlich-rechtlichen Sendern, sowie von Pay-TV Betreibern. Nichtlineare Programmabläufe haben eine Interaktionsstelle, bei der der Nutzer Einfluss darauf hat, was er zu sehen bekommt (Abreu, Nogueira, Becker, & Cardoso, 2017). Ein klassisches Beispiel hierfür ist Netflix. Beim Radio existieren hauptsächlich noch lineare Strukturen. Doch ein Wandel in die Richtung der Digitalisierung ist schon längst im Gange.

Die Interaktionseigenschaft nichtlinearer Medien, lässt sich durch die **Value Co-Creation** erklären, einem Begriff der Service Dominant Logic (SDL) nach Vargo & Lusch (2004). Es beschreibt die Schaffung von Wert durch die Interaktion des Konsumenten und des Anbieters. Durch Kundenfeedback ist es möglich, den Service immer weiter zu verbessern. Service ist definiert als die Anwendung spezifischer Kompetenzen (Kenntnisse und Fähigkeiten) durch Handlungen, Prozesse und Leistungen zum Nutzen einer anderen Entität oder der Entität selbst (Vargo & Lusch, 2004, p. 22). In dieser Arbeit wird das Radio als Service angesehen. Häufig umfasst ein Service in der IS Forschung mehr oder weniger

routinierte Aktivitäten, die eine Interaktion oder eine Beziehung zwischen dem Serviceanbieter und dem Nutzer beinhaltet (Rai & Sambamurthy, 2006). Bisher wurde einem Gut ein Wert durch ihre Produktion zugesprochen. Diese Sichtweise wandelt sich immer mehr zu einer service-zentrierten Perspektive, bei der die Interaktion zwischen dem Konsumenten und dem Anbieter von Bedeutung ist. Diesen Wandel erkennen Vargo & Lusch (2004, p. 5) in ihrer Arbeit. Dabei beschreiben sie die **Leitsätze der SDL** wie folgt:

1. Identifiziere oder entwickle Kompetenzen (Kenntnisse und Fähigkeiten), die einen potenziellen Wettbewerbsvorteil darstellen.
2. Identifiziere potenzielle Kunden, die von diesen Kompetenzen profitieren könnten.
3. Beziehe die Kunden in die Entwicklung mit ein, um ihre spezifischen Bedürfnisse zu erfüllen.
4. Messe den finanziellen Profit von diesem Austausch und lerne wie das Unternehmen Angebote für die Kunden verbessert und den Profit erhöht.

Im klassischen analogen Radio ist dieses Paradigma nicht umgesetzt. Der Hörer kann kein Feedback geben, welches dann verarbeitet werden kann, um personalisierte Inhalte zu spielen oder das Programm zu verbessern. Im nichtlinearen Radio mit variablem Programmablauf ist das allerdings möglich. Im digitalen Bereich sind nichtlineare und personalisierte Programmabläufe umsetzbar. Ein Beispiel wären Apps, bei denen der Hörer gewisse Einstellungen vornehmen kann, wie das Angeben seiner Interessen. An dieser Stelle existieren viele Möglichkeiten, die Nachfrage der Hörer zu identifizieren. Entweder durch eine manuelle Eingabe des Nutzers, in dem er seine Interessen explizit ins System eingibt oder einen Algorithmus, der die Interessen durch Handlungen, wie zum Beispiel das Wiederholen von Beiträgen, das Überspringen dieser oder die Dauer des Hörens, erkennt. Doch die Optimierung der Dienstleistung muss nicht immer ausschließlich durch den Hörer geschehen (Vargo & Lusch, 2004, p. 14).

Insbesondere sind Redakteure in der Optimierung des Programmablaufs involviert. Das Forschungsfeld bezüglich der Frage, was ein **gutes Radiohörerlebnis** ausmacht, ist nicht neu (Hirschmeier et al., 2020, p. 5). Bereits vor 60 Jahren analysierte man die kleinen Unterschiede und Faktoren redaktioneller

Arbeiten, die zum Erfolg führen (Borgers & Koenig, 1960; Crane, Talbot, & Hume, 1961). Viele Radiosender nutzen ein gängiges Verfahren bei ihren Programmen, wobei viel Musik und kurze Nachrichten zwischendurch gespielt werden. Allerdings sorgen andere Faktoren für ein lukrativeres Radioprogramm (Borgers & Koenig, 1960). Die kleinen Details sind die Faktoren, die zum Erfolg führen und den Hörer fesseln. Nach La Roche & Buchholz (2016) gibt es Unmengen an Regeln zu beachten, um dem Hörer das bestmögliche Hörerlebnis zu ermöglichen. Dabei müssen nicht nur grobe Entscheidungen über die Reihenfolge der einzelnen Programminhalte und Beiträge, sondern auch die feingranulare Ebene berücksichtigt werden. Dazu gehören zum Beispiel die Reihenfolge von einzelnen Wörtern in den Sätzen (La Roche 2016, p. 14), die Aussprache und Sprechstimmlage des Sprechers. Einige Radiosender arbeiten mit Uhren, die stündlich wiederkehrende Beiträge, wie etwa Nachrichten, kennzeichnen (Heinrich, 2010). Aufgabe der Redakteure ist es, die Inhalte perfekt aufeinander abzustimmen und die einzelnen Programminhalte nicht beziehungslos da stehen zu lassen. Die Erstellung eines angenehmen Hörerlebnisses ist also komplexer als man zunächst annehmen mag. Doch nicht nur der inhaltliche, sondern auch der technische Aspekt spielt hierbei eine Rolle. So haben Schoeffler, Edler und Herre (2013) bereits erkannt, dass die Audioqualität dazu beitragen kann, das Hörerlebnis zu verbessern. In ihrer Arbeit beziehen sie sich zwar nur auf Musikstücke, trotzdem ist es gegebenenfalls auch beim Radio notwendig, solche Faktoren zu berücksichtigen. Denn besonders im digitalen Radiobereich ist es möglich, den Hörern eine bessere Audioqualität als beim analogen Radio zu bieten.

Grundsätzlich existieren **zwei mögliche Interaktionsmodelle** in einem personalisierten Radio (Hirschmeier & Beule, 2018). Das **Media-Center-Modell** (auch "Pick-One"-Modell genannt) setzt voraus, dass der Hörer mit dem Radio aktiv interagiert. Es folgt somit eine manuelle Eingabe des Nutzers. Zunächst stellt sich die Frage, ob ein Radio mit solch einem Interaktionsmodell bereits existiert. Tatsächlich gibt es Web Radios und auch Radio Apps, jedoch bietet keines oder nur wenige davon einen hohen Grad an Personalisierung. Sappelli, Chu, Cambel, Nortier und Graus (2018) präsentierten eine initiale App, die die Sendungen aus einem Radio in einzelne Audiostücke fragmentiert, um dem Hörer gezieltere Informationen zu bieten. Der Fokus liegt dabei in kurzen und zielgerichteten Audiostücken statt dem Abspielen von ganzen Sendungen. Ebenfalls ist in der App das Anklicken von Tags möglich, um nur bestimmte Audiostücke zu filtern, die den Interessen des Hörers

entsprechen Es wird ausdrücklich hingewiesen, dass die App noch am Beginn steht und großes Verbesserungspotential hat. Insbesondere ist es das Ziel, die App weiter zu personalisieren und dabei mehr Informationen der Nutzer zu verarbeiten. Momentan muss der Hörer jedoch noch nach den gewünschten Tags suchen und manuell selektieren. Hier gibt es keine Berücksichtigung von Mustern, einem redaktionellen Mehrwert oder ähnlichem. Stattdessen werden lediglich Audiostücke gespielt, die der Hörer selbstständig wählen muss. Das ist ein typisches Beispiel für ein Pick-One-Modell. Diese Arbeit fokussiert sich allerdings auf das zweite Interaktionsmodell, nämlich dem **Stream-Modell** (auch "Skip-One"-Modell genannt). Hier wird, ganz wie beim klassischen Radio, das personalisierte Radio abgespielt und die Abspielreihenfolge vom Radio festgelegt. Der Nutzer braucht keine aufwändige, manuelle Eingabe zu tätigen. Er kann lediglich bei Nichtgefallen eines Inhalts, einen Beitrag überspringen. Idealerweise kommt das nie vor, da der Hörer immer das zu hören bekommt, was er wirklich konsumieren möchte. Ein Überspringen wäre somit ein negatives Ereignis. Zum Zeitpunkt dieser Arbeit ist kein personalisiertes und automatisiertes Radio im öffentlich-rechtlichen Bereich bekannt. Häufig gibt es online Mediatheken oder Radios mit personalisierten Ansätzen, die sich jedoch nur auf Musik beschränken (Hayes & Cunningham, 1999; Turnbull et al., 2014).

Die Personalisierung trägt das **Risiko der Entstehung einer Filterblase** mit sich. Diese entsteht, wenn ein Algorithmus versucht vorauszusagen, welche Inhalte der Hörer konsumieren möchte. Dadurch entsteht eine Isolation des Hörers gegenüber anderen Informationen. Diese bleiben ihm dann verwehrt. Dieses Ereignis steht im Gegensatz dazu, was die eigentliche Idee eines öffentlich-rechtlichen Radiosenders ist. Insbesondere existieren große Mengen an Gesetzen, die im Rundfunkvertrag festgelegt sind und den Programmablauf des Radios bestimmen. Dazu gehört beispielsweise ein ausgewogener Mix an Themen. Ebenfalls soll stets ein Überblick über internationale, nationale und regionale Geschehnisse gewährleistet werden. Das sind lediglich zwei Beispiele, jedoch gibt es noch weitere Grundsätze, die stets beachtet werden müssen. Vor allem ergeben sich Schwierigkeiten bei einem zu hohen Grad der Personalisierung. Beispiele hierfür wären ein Nutzer, der nur Hörspiele konsumieren möchte, oder ein Hörer, welcher sich nur für Informationen aus regionaler Umgebung interessiert. Daher muss der Grad der Personalisierung passend bestimmt werden, sodass keine Filterblasen

entstehen und die Verletzung des Rundfunkvertrags vermieden wird. Dazu gibt es bereits Vorschläge, solchen Gegebenheiten auszuweichen, indem beispielsweise eine Compliance-Komponente im Algorithmus berücksichtigt wird. Es wird also deutlich, dass einige Überlegungen notwendig sind, um ein personalisiertes Radio im Sinne des öffentlich-rechtlichen Auftrags zu ermöglichen. Diese Arbeit soll sich erstmal allein auf die Erzeugung synthetischer Ausspielreihenfolgen fokussieren und zeigen, ob es möglich ist, ein künstliches Radio mit personalisierten Funktionen auf einer Beitragsebene anzubieten, welches dem linearen Radio ähnelt. Am Ende soll bewertet werden, ob die Methode zielführend ist, im Hinblick auf die Vermeidung von Filterblasen und Verletzungen des Rundfunkvertrags.

2.2 Sequenzmusteranalyse

Um beim linearen Radio typische Muster herauszufinden, ist die Verwendung von **Pattern Mining Algorithmen** sinnvoll. Dadurch ist es möglich, unerwartete, aber nützliche Muster aus einer Sequenzdatenbank zu identifizieren. Einer der beliebtesten Data Mining Themen ist die Sequenzmusteranalyse, von der verschiedene Anwendungsfelder profitieren können (Fournier-Viger, Lin, Kiran, Koh, & Thomas, 2017, p. 54). Das Suchen sequenzieller Muster zielt darauf ab, Korrelationen zwischen Ereignissen über die Zeit zu entdecken (Plantevit, Choong, Laurent, Laurent, & Teisseire, 2005). Bereits mit der Einführung des Apriori Algorithmus nach Agrawal und Srikant (1995) begann das Forschungsfeld zu dieser Thematik. Ebenfalls basieren viele neuere Algorithmen auf dem Apriori Algorithmus. Für die Musteranalyse ist der Begriff der Sequenz von Bedeutung. Eine Sequenz bezeichnet eine Abfolge von Werten. Zu unterscheiden sind hierbei zwei Typen von sequenziellen Daten, die im Data Mining Feld benutzt werden (Han, Kamber, & Pei, 2011, p. 586). Beim ersten Typ handelt es sich um Zeitfolgen, welche eine geordnete Liste von Zahlen beschreibt. Der zweite Typ sind Sequenzen, sprich geordnete Listen von nominalen Werten beziehungsweise Symbolen. Letztere, sprich Sequenzen, finden in dieser Arbeit Berücksichtigung, da Beiträge im Radio nominale Werte darstellen. Algorithmen, die mit Sequenzen arbeiten, erwarten damit eine sequenzielle Datenbank (SDB) als Input. Ein Beispiel für solch eine Datenbank liefert Abbildung 2. Mathematisch ausgedrückt ist nach Fournier (2017, p.57) eine Sequenz $s_a = A_1, A_2, \dots, A_n$ enthalten in einer weiteren Sequenz $s_b = B_1, B_2, \dots, B_m$, wenn ganze Zahlen $1 \leq i_1 < i_2 < \dots < i_n \leq m$ existieren, sodass $A_1 \subseteq B_{i_1}, A_2 \subseteq$

$B_{i_2}, \dots, A_n \subseteq B_{i_n}$. Beispielsweise ist die Sequenz $\{b\}, \{g\}, \{f\}$ in der Sequenz $\{a, b\}, \{c\}, \{f, g\}, \{g\}, \{e\}$ enthalten. Die Sequenz $\{b\}, \{g\}, \{f\}$ hingegen nicht. Somit ist $sa \subseteq sb$. Zu beachten ist noch der Begriff Itemset, welches eine Menge unsortierter Mengen darstellt. Im obigen Beispiel wäre $\{a, b\}$ ein Itemset. Ob erst a oder b folgt ist für die Logik irrelevant. Trotzdem wird davon ausgegangen, dass die Elemente in einem Itemset in lexikografischer Reihenfolge sortiert sind. Das stellt sicher, dass die SDB einheitlich ist und die entsprechenden Algorithmen mit diesem Input arbeiten können.

Der **Support einer Sequenz** zeigt, wie häufig ein bestimmtes Muster in der SDB vorkommt. Dieser Begriff wird in der Arbeit noch häufig zum Vorschein kommen. Daher soll die Definition im Folgenden geklärt werden. Nach Fournier-Viger et al. (2017, p. 57) bezeichnet der Support einer Sequenz s_a in einer SDB die Anzahl an Sequenzen welche s_a enthalten. Mathematisch ausgedrückt:

$$\text{sup}(s_a) = \{s | s \subseteq s_a \wedge s \in SDB\}$$

Im Beispiel der Abbildung 2 kommt Sequenz $\{a, b\}$ beispielsweise in den Sequenzen 1, 2 und 3 vor. Der Support wäre somit 3 oder absolut ausgedrückt $\frac{3}{4}$. Der Minsup bezeichnet bei der Sequenzmusteranalyse den prozentualen Anteil, den ein Muster in den Sequenzen aufweisen muss, um überhaupt berücksichtigt zu werden.

Tatsächlich gibt es viele verschiedene Algorithmen für die Sequenzmusteranalyse, die allerdings alle ein identisches Ergebnis liefern (Fournier-Viger et al., 2017, p. 58). Der Unterschied liegt lediglich in den Datenstrukturen und im Vorgehen der Algorithmen. Das wirkt sich auch auf die Laufzeit aus. Der Algorithmus CM-Spam ist schneller als andere Algorithmen (Fournier-Viger et al., 2017, p. 61). In dieser Arbeit wird daher auch nur der CM-Spam Verwendung finden.

SID	Sequenz
1	$\{\{a, b\}, \{c\}, \{f, g\}, \{g\}, \{e\}\}$
2	$\{\{a, d\}, \{c\}, \{b\}, \{a, b, e, f\}\}$
3	$\{\{a\}, \{b\}, \{f\}, \{e\}\}$
4	$\{\{b\}, \{f, g\}\}$

Abbildung 2: Eine sequenzielle Datenbank

Quelle: Fournier-Viger et al. (2017)

3 Qualitative Inhaltsanalyse

In diesem Kapitel werden hauptsächlich vorhandene Daten auf Häufigkeiten und Auffälligkeiten untersucht. Dies geschieht in Form einer qualitativen Inhaltsanalyse. Zu Beginn wird die Methodik, sprich die Vorgehensweise erklärt. Daraufhin werden die zu analysierenden Daten genau beschrieben. Die Ergebnisse der qualitativen Inhaltsanalyse werden dargestellt und so aufbereitet, dass sie in den nächsten Kapiteln Verwendung finden können.

3.1 Methodik

Zur Vorbereitung der Daten auf die spätere Anwendung der Sequenzmusteranalyse wird die **qualitative Inhaltsanalyse nach Mayring (2015)** herangezogen. Dieser Ansatz ist sinnvoll, um einen gewissen Grad der Qualitätssicherung für die Analyse sicherzustellen. Im Folgenden werden die einzelnen Schritte dieser Methodik vorgestellt:

1. *Formulierung der Fragestellung*: In diesem ersten Schritt wird festgelegt, welche Fragestellung untersucht werden soll.
2. *Auswahl und Beschreibung der Daten*: Die zu analysierenden Daten werden erfasst und beschrieben.
3. *Aufstellen des Codingframes*: Das sogenannte Codingframe wird festgelegt. Es beinhaltet die Kategorien, nach denen die Datenquelle kodiert werden soll. Eine Definition der einzelnen Kategorien ist ebenfalls vorhergesehen.
4. *Bestimmung der Analyseeinheiten*: Die Bestimmung der Einheit der Kodierung.
5. *Kodierung*: Durcharbeiten der Daten mithilfe des Codingframes.
6. *Verrechnung*: Das Feststellen und Vergleichen der Häufigkeiten einzelner Kategorien.
7. *Resultat*: Die Darstellung und die Interpretation der Ergebnisse.

Zu beachten ist, dass die Formulierung der Fragestellung, sprich der erste Schritt, bereits in den vorherigen Kapiteln ausführlich dargestellt wurde. In den folgenden Kapiteln wird auf die weiteren Schritte genauer eingegangen. Zusammengehörige Schritte sind in einem Unterkapitel zusammengefasst. Die Titel

der Unterkapitel können von den genauen Bezeichnungen der oben definierten Schritte abweichen.

In der Tat handelt es sich um einen datenorientierten Ansatz. Nun könnte man argumentieren, dass es noch weitere valide Ansätze gibt, wie beispielsweise die Befragung von Redakteuren nach gewissen Mustern und Regelmäßigkeiten. Diese Frage ist durchaus berechtigt, allerdings soll diese Arbeit initial mit vorhandenen Daten arbeiten und anschließend diesen Aspekt noch einmal aufgreifen. Dabei soll bewertet werden, ob solch eine Absprache tatsächlich notwendig ist. Weiterhin ist bei einer Befragung von Redakteuren zu erwarten beziehungsweise es ist wahrscheinlich, dass sie unterschiedliche Aussagen von sich geben und auch nicht auf feingranularer Ebene ihren Programmablauf bestimmen, wie es bei einem datenorientierten Ansatz möglich ist. Eine weitere Herangehensweise könnte ein hybrider Ansatz darstellen, welcher die Aussagen der Redakteure und die Analyse der gegebenen Daten berücksichtigt. Nichtsdestotrotz fokussiert sich diese Arbeit erstmal auf vorhandene Radiomitschnitte. Dadurch ist es eben möglich, genau auf dieser feingranularen Ebene zu arbeiten und somit die kleinen und unerwarteten Muster zu identifizieren.

3.2 Auswahl und Beschreibung der Daten

In Deutschland gibt es zwei verschiedene Arten von Radiosendern. Die privaten und öffentlich-rechtlichen Sender. Diese Arbeit fokussiert sich auf letztere. Der Grund hierfür ist, dass öffentlich-rechtliche Sender gewissen Regulierungen unterliegen (Ministerium des Innern des Landes Nordrhein-Westfalen, Referat 14, 2020; Pöchlacker et al., 2017). Dadurch ist ein ausgeglichener Mix an Themen sichergestellt. Diese Ausgeglichenheit ist gewährleistet, da spezielle Programmgrundsätze vorgegeben sind. Darunter fallen beispielsweise Angaben, dass eine angemessene Zeit für bestimmte Themen gewidmet werden muss. Außerdem darf in öffentlich-rechtlichen Programmen keine einseitige Berichterstattung stattfinden. Es darf keiner einzelnen Partei, Gruppe oder Weltanschauung gedient werden. Insbesondere wird darauf geachtet, keine menschenverachtenden Themen zu behandeln, sowie die Meinung, Ansichten oder Glaubenssätze anderer stets zu respektieren. Im Gegensatz dazu kann bei Musik-Streaming-Diensten jeder beliebige Musiker seine Inhalte hochladen. Darunter gehören auch nicht jugendfreie Inhalte oder extreme Meinungen. Beim öffentlich-rechtlichen Radio sind solche

Gegebenheiten ausgeschlossen. Durch das Vorhandensein von Programmgrundsätzen ist eine vielfältige Themenbehandlung sowie eine angemessene Zeit für die Beiträge sichergestellt. Damit einhergehend beinhalten öffentlich-rechtliche Radiosender eine typische und häufig vorkommende Struktur im Radioprogramm. Aufgrund dieser hohen Repräsentativität ist diese Art von Radio für eine Musteranalyse geeignet.

Diese Thesis konzentriert sich auf eine Zusammenarbeit mit dem öffentlich-rechtlichen Sender Deutschlandradio. Die verwendeten Daten für die qualitative Inhaltsanalyse und für die Musteranalyse stammen von diesem Radioanbieter. Gegenwärtig existieren drei verschiedene Programme dieses Senders: Deutschlandfunk (DLF), Deutschlandfunk Kultur (DLF Kultur) und Deutschlandfunk Nova (DLF Nova). Nach eigenen Angaben auf der Homepage ist DLF ein Programm mit einer hohen Informations- und Nachrichtendichte und somit für die breite Masse bestimmt. DLF Kultur ist, wie der Name schon verrät, für ein kulturorientiertes Publikum gedacht und behandelt dabei viele Themen der Literatur sowie der Musik. Der Sender DLF Nova ist an ein jüngeres Publikum gerichtet und unterscheidet sich in seinem Programminhalt stark von den anderen beiden Programmen. Hier wird nämlich viel Musik gespielt und sich oft Themen gewidmet, die in DLF und DLF Kultur keine oder wenig Beachtung finden. Das wird im weiteren Verlauf der Arbeit noch deutlicher. Um möglichst viel Varietät im Tonmaterial sicherzustellen, Unterschiede aufzudecken und möglichst viele Auffälligkeiten zu identifizieren, wurden von jedem dieser Programme jeweils zwei Mitschnitte aufgenommen. Dafür wurde ein Wochenendtag, Samstag der 31.10.2020, sowie ein Wochentag, Montag der 02.11.2020, gewählt. Die Aufnahmen gehen jeweils von 0:00 – 0:00 Uhr. Daraus resultieren sechs ganztägige Aufnahmen und insgesamt etwa 144 Stunden Tonmaterial. Tabelle 1 zeigt eine übersichtliche Darstellung der vorhandenen Mitschnitte.

Weiterhin stehen, neben dem Hörmaterial, für jedes einzelne Programm noch ein Sendeplan zur Verfügung. Die Sendepläne sind online für jeden einzelnen Tag einsehbar. Der Sendeplan wird im nächsten Schritt der Kodierung herangezogen, um Unklarheiten zu vermeiden. Dadurch können Sendungen von Beiträgen klar unterschieden werden. Abbildung 3 zeigt ein Beispiel eines Sendeplans. Zu beachten ist, dass solch ein Sendeplan nur grobe Daten liefert, weswegen auch die angegebenen Zeiten einen großen Abstand voneinander haben. Innerhalb einer Sendung spielen

Beiträge, die im Sendeplan nicht aufgelistet sind. Beispielsweise folgen fast stündlich, manchmal halbstündlich Nachrichten oder auf die Sendungen abgestimmte Beiträge. Die Gattungen der Beiträge differenzieren hierbei stark. Dies wird im späteren Verlauf der Analyse aufgeschlüsselt. Was jedoch bereits im Sendeplan auffällt, ist, dass abendliche und nächtliche Sendungen einen großen zeitlichen Abstand zu der nächsten Tagessendung aufweisen. Außerdem können gewisse Sendungen, die beispielsweise lediglich Musik enthalten, jetzt schon von der Kodierung ausgeschlossen werden. Beispielsweise ist die Sendung „Soundtrack“ des DLF Nova Sendeplans ein Hinweis für solch eine Gegebenheit. Für die folgenden Schritte ist es daher sinnvoll, die Beiträge auch nach der Tageszeit zu klassifizieren.

Tabelle 1: Übersicht der Radiomitschnitte

Spalte1	Sender	Tag	Datum	Zeitraum	Dauer in h
Aufnahme 1	DLF	Samstag	31.10.2020	00:00-00:00	24
Aufnahme 2	DLF	Montag	02.11.2020	00:00-00:00	24
Aufnahme 3	DLF Kultur	Samstag	31.10.2020	00:00-00:00	24
Aufnahme 4	DLF Kultur	Montag	02.11.2020	00:00-00:00	24
Aufnahme 5	DLF Nova	Samstag	31.10.2020	00:00-00:00	24
Aufnahme 6	DLF Nova	Montag	02.11.2020	00:00-00:00	24
					144

Montag	Dienstag	Mittwoch	Donnerstag
0:00 Soundtrack	0:00 Soundtrack	0:00 Soundtrack	0:00 Soundtrack
6:30 Hielscher oder Haase	6:30 Hielscher oder Haase	6:30 Hielscher oder Haase	6:30 Hielscher oder Haase
10:00 Grünstreifen	10:00 Grünstreifen	10:00 Grünstreifen	10:00 Grünstreifen
18:00 Update	18:00 Update	18:00 Update	18:00 Update
20:00 Eine Stunde History	20:00 Eine Stunde Film	20:00 Eine Stunde Talk	20:00 Eine Stunde Was mit Medien
21:00 Ab 21 – Willkommen im Nova-Club	21:00 Ab 21 – Willkommen im Nova-Club	21:00 Ab 21 – Willkommen im Nova-Club	21:00 Ab 21 – Willkommen im Nova-Club
Freitag	Samstag	Sonntag	
0:00 Soundtrack	0:00 Soundtrack	0:00 Soundtrack	
6:30 Hielscher oder Haase	7:00 Early Bird	8:00 Dein Sonntag	
10:00 Grünstreifen	10:00 Endlich Samstag	12:00 Grünstreifen	
18:00 Update	14:00 Grünstreifen	16:00 Einhundert	
20:00 Eine Stunde Liebe	18:00 Hörsaal	18:00 Hörsaal	
21:00 Ab 21 – Willkommen im Nova-Club	19:00 Club der Republik	19:00 Ab 21 – Das Beste aus dem Nova-Club	
		21:00 Soundtrack Lounge	

Abbildung 3: Beispiel-Sendeplan für den Sender DLF Nova

Quelle: Deutschlandfunk Nova (2021)

3.3 Aufstellen des Codingframes

Der **Codingframe** ist ein übergeordnetes Kategoriensystem. Es dient dazu, die Radiobeiträge einzeln zu klassifizieren, das heißt bestimmten Kategorien zuzuordnen. Die Wahl der entsprechenden Codes wurde aufgrund zweier Grundlagen getroffen. Einerseits ist die Wahl inspiriert von vorherigen Arbeiten am Lehrstuhl der Universität zu Köln. Andererseits wurden neue Codes ausgewählt und alte weiter ausgeführt. Dies geschah durch die Zusammenarbeit mit dem Betreuer der Thesis, den Programmexperten und den Redakteuren von Deutschlandradio. Insgesamt resultieren daraus drei Oberkategorien, die den Codingframe bilden: Klasse, Bezugsreichweite und Gattung. Im Folgenden wird jede dieser Oberkategorien definiert. Bei den Kategorien Klasse und Bezugsreichweite sind die einzelnen Elemente selbsterklärend und bedürfen somit keiner genauen Erläuterung. Die Gattungen hingegen sind intransparent und benötigen daher eine Definition. Diese Vorgehensweise verhindert spätere Unklarheiten. Außerdem sorgt es dafür, dass die Kodierung an jeder Stelle der Thesis nachvollziehbar ist. Zu beachten ist,

dass ein Beitrag innerhalb einer Oberkategorie mehreren Elementen zugeordnet sein kann. Sprich es existiert zwischen Beitrag und Kodierung eine 1:n Beziehung. Im Folgenden wird auf jede einzelne Oberkategorie genauer eingegangen:

1. *Klasse*: Diese Oberkategorie beinhaltet 20 Codes (siehe Tabelle 2) und ordnet die Beiträge einem entsprechenden Themenbereich zu. Es wurde darauf geachtet, dass möglichst jeder Beitrag einer Klasse zugeordnet werden kann. Daher wurde auch beispielsweise der Code „Medien“ im Nachhinein ergänzt. Denn medienbezogene Beiträge sind öfter vorgekommen, hatten aber keinen Code, dem sie zugewiesen werden könnten. Eine Zuweisung ist allerdings nicht immer möglich, da tatsächlich auch nichts aussagende Beiträge vorhanden sind. Für diesen Fall existiert der „Keine Klasse“ Code. Für diese Oberkategorie existieren erste Bedenken. Gewisse Themen kommen häufiger vor, welche eigentlich nicht zur Norm gehören. Zum Zeitpunkt dieser Arbeit ist die Pandemie COVID-19 ein ständiges Thema und beeinflusst die Radio- und TV-Programme.

Tabelle 2: Codingframe - Klasse

Klasse	
Biographie	Politik
Erziehung und Bildung	Religion
Wirtschaft	Sport
Freizeit	Technik
Gesellschaft und Soziales	Umwelt
Geschichte	Verkehrswesen
Recht	Wissenschaft und Forschung
Kultur	Katastrophe
Gesundheitswesen und Medizin	Medien
Natur	Keine Klasse
	20 Elemente

2. *Bezugsreichweite*: Beiträge besitzen oft einen geographischen Bezug. Die Unterteilung erfolgt hier in fünf Elemente (siehe Tabelle 3). Da es sich bei der Datenquelle um einen deutschen Radiosender handelt, sind die Codes „Deutschland“ und „Bundeslandebene“ vorhanden. Beiträge, die keinen geographischen Bezug haben, werden mit dem letzten Code „Kein Geographika“ kodiert.

Tabelle 3: Codingframe - Bezugsreichweite

Bezugsreichweite
Bundeslandebene
Deutschland
Europäische Ebene
Außereuropäische Ebene
Kein Geographika
5 Elemente

3. *Gattung*: Diese Klassifizierung beinhaltet insgesamt 108 Codes (siehe Tabelle 4). Im Anhang A befinden sich die Erklärungen der einzelnen Codes, welche aus der Datenbank von Deutschlandradio stammen. Einige Codes beinhalteten keine Erklärung und wurden nachträglich hinzugefügt. Gekennzeichnet sind diese mit runden Klammern. Auch hier existiert für nicht zuordnungsbar Beiträge ein Code „Keine Gattung“.

Tabelle 4: Codingframe - Gattung

Gattung		
Collage	Absage	Hörspiel
Hörspielmusik	Anmoderation	Dokumentarhörspiel
Musiktheater	Ansage	Prozessbericht
Werbespot	Zwischenansage	Reisebericht
Hörspielbearbeitung	Jingle	Betrachtung
Kinderhörspiel	Stationsansage	Debatte
Kurzhörspiel	Teaser	Diskussion
Mitmachhörspiel	Trailer	Rezitation
Mundarthörspiel	Ratgebersendung	Konzertaufführung
Musikalisches Hörspiel	Quiz	Dialog
Originalhörspiel	Rede	Sendespiel
Science Fiction-Hörspiel	Ansprache	Bericht in Interviewform
Interview	Dankesrede	Sportbericht
Umfrage	Predigt	Dokumentarbericht
Büttenrede	Reportage	Programm kennungen
Kabarett	Lesung	Kleinkunst
Sketch	Rezension	Sprechtheater
Witz	Sendung mit Hörerbeteiligung	Theater
Kommentar	Spot	Humoristischer Vortrag
Autorenlesung	Aufklärungs- und Gesundheitsspot	Tilawa
Szenische Lesung	Wahlspot	Tilawa tartitl
Magazin	Sprachkurs	Kriminalhörspiel
Montage	Statement	Originaltonhörspiel
Nachrichtensendung	Tipp	Comedy
Meldung	Vortrag	Revue
Nachrichten	Werkeinführung Musik	Nachruf
Verkehr	Wort-Musik-Sendung	Moderation
Wetter	Monolog	Zwischenmoderation
Öffentliche Veranstaltung	Ars acustica	Rätselsendung
Festakt	Bericht	Feature
Pressekonferenz	Korrespondentenbericht	Amateurrhörspiel
Religiöse Veranstaltung	Textmontage	Essay
Porträt	Filmton	Originaltonfeature
Presseschau	Gespräch	Laudatio
Programmelement	Glosse	Tilawa tagwie
Abmoderation	Hörbild	Keine Gattung
		108 Elemente

4. *Tageszeit*: Aus dem Sendeplan wurde bereits ersichtlich, dass gewisse Sendungen typischerweise zu einer bestimmten Tageszeit laufen. Daher wurde diese Oberkategorie hinzugefügt. Die Tageszeit ist in vier Elemente unterteilt (siehe Tabelle 5). Morgens bezeichnet dabei die Zeit von 05:00 – 10:00 Uhr,

tagsüber 10:00 – 18:00 Uhr, abends 18:00 – 00:00 Uhr und nachts 00:00 – 05:00 Uhr. Diese Klassifizierung ist notwendig, um später genaue Aussagen bezüglich der Häufigkeit von Beiträgen zu bestimmten Tageszeiten tätigen zu können.

Tabelle 5: Codingframe - Tageszeit

Tageszeit
Morgens
Tagsüber
Abends
Nachts
4 Elemente

3.4 Bestimmung der Analyseeinheiten

Der Fokus der Kodierung sowie der späteren Sequenzmusteranalyse liegt auf der Beitragsebene. Das Ziel ist nämlich herauszufinden, welche Beiträge typischerweise aufeinanderfolgen. Auf Sendungsebene würde die Kodierung keine nützlichen Ergebnisse liefern, da eine Sendung mehrere heterogene Beiträge beinhaltet und somit keinem der oben definierten Codes eindeutig zugeordnet werden könnte. Ebenso würde eine feingranulare Ebene, das heißt die Kodierung von jedem kleinsten Detail, vermutlich keine neuen Erkenntnisse liefern. Das liegt zum einen daran, dass sich vorherige Arbeiten am Lehrstuhl sowie ein Forschungspapier bereits damit beschäftigt haben (Hirschmeier et al., 2020). Zum anderen werden häufig innerhalb oder zwischen den Beiträgen Übergänge genutzt, die lediglich die Zeit zum nächsten Beitrag oder zur vollen Stunde überbrücken. Dazu gehören Übergänge wie Musik, Werbespots, Teaser oder Trailer. Diese sind von der Kodierung ausgeschlossen.

3.5 Kodierung und seine Validität

Der Kodierungsprozess bezeichnet in dieser Arbeit das Durchhören der Radiomitschnitte und dem gleichzeitigen Zuordnen der Beiträge zu den im Vorfeld definierten Codes. Dieser Prozess geschah durch eine einzelne Person. Um einen hohen Grad der Validität sicherzustellen, wurde eine erste Kodierung von Programmexperten des Deutschlandradios bereitgestellt. Dies erfolgte durch die Bereitstellung einer initialen Datei. Dort sind bereits einige Beiträge, die sich in der Datenbank befinden, den jeweiligen Codes zugeordnet. Diese initiale

Kodierungstabelle musste dann auf den zuvor definierten Codingframe angepasst werden. Weiterhin mussten Lücken identifiziert und vervollständigt werden. Außerdem wurde mit dem Betreuer der Arbeit Rücksprache gehalten und es wurde Feedback eingeholt, um kleinere Anpassungen durchzuführen. Das Ergebnis ist somit eine lange Excel-Tabelle mit den Beiträgen, welche kodiert sind und keine Übergänge beinhalten, die für die Sequenzanalyse nicht von Bedeutung sind. Diese Excel-Tabelle hat um die 680 Zeilen. Jede Zeile spiegelt einen Beitrag wider. Tabelle 6 zeigt die wichtigsten Spaltennamen dieser Datei. Es folgt eine kurze Erläuterung zu jeder Spalte.

Tabelle 6: Die wichtigsten Spalten der Kodierungstabelle

Spalten
station_name
broadcast_title
audio_title
airtime_start
airtime_end
audio_duration
klassen
geographika
gattungen
tageszeit

1. *station_name*: Name des Radiosenders (DLF, DLF Kultur oder DLF Nova).
2. *broadcast_title*: Name der übergeordneten Sendung.
3. *audio_title*: Name des Beitrags.
4. *airtime_start* & *airtime_end*: Start und Ende des Beitrags.
5. *audio_duration*: Dauer des Beitrags.
6. *klassen*, *geographika*, *gattungen*, *tageszeit*: Die Kodierungsspalten.

3.6 Verrechnung und Resultat

In diesem Kapitel werden die Ergebnisse zur Häufigkeit der jeweiligen Codes vorgestellt. Dieser Schritt ist für die Erkenntnisgewinnung notwendig und insbesondere hilfreich, um Anforderungen für die spätere Sequenzmusteranalyse herzuleiten.

3.6.1 Ergebnisse der Klassen

Einige Klassen dominieren eindeutig das Radioprogramm. Tabelle 7 zeigt, dass weitgehend die Klassen „Politik“, „Gesundheitswesen und Medizin“ sowie „Kultur“ eine hohe Häufigkeit vorweisen. Letztere kommt nicht nur, wie eigentlich erwartet, in DLF Kultur vor, sondern auch in DLF. DLF Nova hingegen besitzt nur sehr wenige Beiträge, die kulturbezogene Themen ansprechen. Eine weitere Auffälligkeit ist, dass am Wochenende deutlich weniger Themen angesprochen werden und das unabhängig vom Programm. An Wochentagen hingegen ist die Anzahl der Themen bei DLF und DLF Nova, verglichen mit den Wochenendtagen, doppelt so hoch. Bei DLF Kultur hingegen macht das keinen großen Unterschied. Es wird deutlich, dass DLF den meisten Informationsgehalt bietet, aufgrund der hohen Anzahl verschiedener Themen.

Betrachtet man die Anteile der Codes, dann lassen sich weitere Erkenntnisse ableiten. Der Anteil eines Codes i ergibt sich wie folgt:

$$\text{Anteil}(i) = \frac{\# \text{ Beiträge in denen Code } i \text{ vorkommt}}{\# \text{ Beiträge}}$$

Der Code „Politik“ hat den größten Anteil in der gesamten Datenquelle. Die Klasse „Gesundheitswesen und Medizin“ ist ebenfalls sehr präsent und weist mit 18,97% den zweithöchsten Anteil auf. Wie bereits vermutet liegt das an der aktuellen Pandemie zum Zeitpunkt der Arbeit. Ebenso verhält es sich mit den Codes „Verkehrswesen“ und „Katastrophe“. Aufgrund aktueller Ereignisse weisen diese beiden Codes einen relativ hohen Anteil auf. Gewisse Codes kommen jedoch überhaupt nicht oder kaum vor. Dazu gehören beispielsweise „Geschichte“ und „Freizeit“. Letztere kommt sogar nur in DLF Nova vor.

Eine weitere Erkenntnis ist, dass einige Themen keinen hohen Anteil besitzen, jedoch in Beiträgen vorkommen, die eine sehr lange Dauer aufweisen. Beispielsweise besitzt der Code „Gesellschaft und Soziales“ einen Anteil von nur 4,28%, jedoch eine Dauer von knapp 14 Stunden und 45 Minuten. Zu beachten ist, dass ein Beitrag oft über mehrere Themen handelt. Das heißt, die Dauer ist nur ein Richtwert und bedeutet nicht immer, dass ein Thema lange gespielt wird, sondern in Beiträgen vorkommt, die eine lange Spieldauer vorweisen.

Für die Sequenzmusteranalyse heißt dies, dass auf häufig zum Vorschein tretende Themen geachtet werden sollte. Ebenfalls sollten Themen genauer untersucht werden, welche in Beiträgen mit höherer Dauer vorkommen. Beispielsweise ist „Politik“ der Code mit dem meisten Anteil. Doch welche Beiträge folgen auf politische Beiträge? Die fünf Codes mit dem höchsten Anteil sind: „Politik“, „Gesundheitswesen und Medizin“, „Kultur“, „Katastrophe“ und „Verkehrswesen“. Ergänzend dazu die fünf Codes mit der höchsten Beitragsdauer: „Kultur“, „Politik“, „Gesundheitswesen und Medizin“, „Gesellschaft und Soziales“ und „Biographie“. Die Beachtung dieser Codes erfolgt in der Sequenzmusteranalyse vermutlich auch schon unbeabsichtigt, da häufige Codes auch in häufigen Mustern vorkommen.

Tabelle 7: Beobachtete Klassen

	DLF		DLF Kultur		DLF Nova		Anteil [%]	Dauer [h:m:s]
	Samstag	Montag	Samstag	Montag	Samstag	Montag		
Biographie	9	8	9	1	7	0	2,24	12:32:00
Erziehung und Bildung	2	11	0	5	1	6	1,65	03:06:16
Wirtschaft	9	35	7	9	1	5	4,35	05:46:25
Freizeit	0	0	0	0	5	5	0,66	00:45:13
Gesellschaft und Soziales	12	14	15	14	4	6	4,28	14:45:49
Geschichte	1	0	0	0	0	2	0,2	01:00:51
Recht	10	26	9	7	3	20	4,94	07:45:04
Kultur	30	39	61	71	8	2	13,9	47:24:07
Gesundheitswesen und Medizin	39	89	30	44	28	58	18,97	26:34:03
Natur	4	2	1	0	6	15	1,84	01:50:52
Politik	64	98	34	46	22	56	21,08	30:40:18
Religion	9	9	8	3	0	6	2,31	04:18:02
Sport	26	15	3	8	12	1	4,28	06:54:29
Technik	5	14	3	1	2	16	2,7	03:44:58
Umwelt	1	7	1	2	6	8	1,65	02:44:59
Verkehrswesen	30	3	21	3	14	8	5,2	06:40:22
Wissenschaft und Forschung	3	5	4	5	9	7	2,17	05:04:17
Katastrophe	19	27	14	14	3	11	5,8	07:04:34
Medien	1	7	4	5	5	4	1,71	02:35:53
Keine Klasse	0	0	0	0	0	1	0,07	00:01:55
Anzahl	274	409	224	238	136	237		

3.6.2 Ergebnisse der Bezugsreichweite

Vor und während dem Kodieren entstand der Eindruck, dass die Beiträge überwiegend „Deutschland“ als geographischen Bezug haben. In der Tat hat sich diese Vermutung bestätigt. Tabelle 8 zeigt, dass „Deutschland“ den höchsten Anteil und die höchste Dauer in den Beiträgen aufweist. Mit dem Code „Europäische Ebene“ verhält es sich ähnlich. Die „Bundeslandebene“ kommt vergleichsweise weniger vor, aber weist dennoch einen hohen Anteil und eine hohe Dauer auf. Dass ein Beitrag keinem Geographika zugeordnet wird, geschieht sehr selten. Fast jeder

Beitrag bezieht sich auf eine bestimmte Region. Im Großen und Ganzen sind die Beiträge, was die Bezugsreichweite angeht, sehr ausgeglichen.

Für die Simulation der Ausspielreihenfolgen heißt dies Folgendes. Die Bezugsreichweite sollte ausgewogen sein. Sprich es sollte darauf geachtet werden, dass nicht ständig nur von einem einzigen Land oder von einer einzigen Region berichtet wird.

Tabelle 8: Beobachtete Bezugsreichweite

	DLF		DLF Kultur		DLF Nova		Anteil [%]	Dauer [h:m:s]
	Samstag	Montag	Samstag	Montag	Samstag	Montag		
Bundeslandebene	32	34	32	15	21	23	12,43	18:33:52
Deutschland	67	121	50	59	25	64	30,56	43:45:48
Europäische Ebene	65	108	45	63	31	67	30,01	36:53:10
Außereuropäische Ebene	43	84	32	49	25	56	22,88	26:12:03
Kein Geographika	11	4	7	8	12	10	4,12	23:21:44
Anzahl	218	351	166	194	114	220		

3.6.3 Ergebnisse der Gattungen

Beim Erstellen des Codingframes wurde eine hohe Anzahl an Gattungen angeführt. Jedoch kommt in den Radiomitschnitten nur ein Bruchteil dieser Codes vor. Diese sind in Tabelle 9 aufgeführt. Die Gattung Bericht weist mit 21,14 % den höchsten Anteil auf. Weiterhin besitzen die Codes „Nachrichten“ und „Wetter“ ebenfalls einen sehr hohen Anteil und eine lange Dauer.

Zu beachten ist, dass gewisse Gattungen typischerweise nur in bestimmten Sendern vorkommen. Die Gattung „Quiz“, beispielsweise, kommt nur in DLF Nova zum Vorschein. In einem Quiz nehmen Hörer bei einem Antwort-Frage-Spiel teil. Für die Simulation der Ausspielreihenfolgen bedeutet dies, dass bei einem jüngeren Publikum spielerische Gattungen wie zum Beispiel ein Quiz mit eingebracht werden sollten. Ein weiteres Beispiel für eine sendertypische Gattung ist der „Tipp“. Dieser kommt am häufigsten in DLF Kultur vor. Das sind meistens Literatur- oder TV Tipps. Auch Hörspiele kommen typischerweise nur in DLF Kultur vor und in sonst keinem Sender. Für die Simulation von Ausspielreihenfolgen heißt das, dass ein kulturorientiertes Publikum auch regelmäßig Tipps oder Hörspiele konsumieren möchte.

Noch eine Auffälligkeit ist der Sender DLF mit dem Code „Bericht“. Dieser hat am Wochentag einen Häufigkeitswert von 87. Außerdem ist auch die Frequenz von „Nachrichten“ vergleichsweise hoch. Die Sequenzmusteranalyse sollte hier

genau auf potenzielle Muster untersuchen. Zudem sollte bei der Simulation von Ausspielreihenfolge darauf geachtet werden, dass bei Hörern, die sonst viel DLF hören, der Informationsgehalt hoch ist.

Es existieren auch Sendungen, die einen geringen Anteil aber eine lange Dauer aufweisen. Der Code „Wort-Musik-Sendung“ ist ein Beispiel hierfür. Diese Gattung spielt oft nachts, wenn kein hoher Informationsgehalt mehr notwendig ist. Die Betrachtung der Tageszeit folgt allerdings im nächsten Schritt genauer. Weiterhin ist auch zu beachten, dass gewisse Gattungen kaum oder gar nicht vorkommen. „Laudatio“ ist ein Beispiel hierfür. Solche Gattungen könnte man im personalisierten Radio auslassen oder als Lückenfüller benutzen.

Tabelle 9: Beobachtete Gattung

	DLF		DLF Kultur		DLF Nova		Anteil [%]	Dauer [h:m:s]
	Samstag	Montag	Samstag	Montag	Samstag	Montag		
Collage	0	0	3	1	0	1	0.57	0:15:56
Kurzhörspiel	0	0	1	0	0	0	0.11	0:05:02
Sketch	1	0	0	0	0	0	0.11	0:02:29
Interview	7	14	4	8	0	2	3.98	5:59:42
Kommentar	6	4	1	1	0	0	1.36	0:58:56
Magazin	3	0	2	1	0	0	0.68	5:20:04
Nachrichtensendung	0	0	1	0	0	0	0.11	0:05:48
Meldung	1	2	1	2	4	14	2.73	1:08:37
Nachrichten	29	40	24	29	14	30	18.86	13:03:41
Wetter	26	38	23	27	14	30	17.95	12:31:03
Porträt	3	1	1	0	0	0	0.57	0:21:45
Presseschau	4	8	1	1	0	0	1.59	1:26:01
Programmelement	0	0	0	0	0	1	0.11	0:03:21
Teaser	0	0	1	0	0	0	0.11	0:00:51
Ratgebersendung	0	1	0	0	0	0	0.11	0:04:20
Quiz	0	0	0	0	4	1	0.57	0:26:21
Dankesrede	0	1	0	0	0	0	0.11	0:24:58
Predigt	1	0	0	0	0	0	0.11	0:54:54
Reportage	5	9	1	4	0	0	2.16	2:38:12
Rezension	2	9	4	8	0	0	2.61	2:44:01
Sendung mit Hörerbeteiligung	0	3	2	0	0	0	0.57	3:22:15
Statement	0	1	0	0	0	0	0.11	0:00:43
Tipp	0	1	2	5	0	0	0.91	0:24:46
Werkeinführung Musik	1	2	0	0	0	0	0.34	2:22:15
Wort-Musik-Sendung	6	0	3	3	2	0	1.59	16:07:49
Bericht	41	87	14	18	6	20	21.14	14:49:12
Korrespondentenbericht	3	9	0	3	1	3	2.16	1:31:40
Gespräch	4	7	18	19	25	16	10.11	15:07:30
Betrachtung	1	3	2	1	0	0	0.8	0:23:22
Diskussion	2	3	0	0	1	0	0.68	4:29:32
Konzertaufführung	0	0	2	1	0	0	0.34	3:06:21
Bericht in Interviewform	1	8	5	6	1	16	4.2	3:03:22
Sportbericht	5	5	0	0	0	0	1.14	0:49:10
Kriminalhörspiel	0	0	0	1	0	0	0.11	0:54:31
Nachruf	2	0	3	0	0	0	0.57	0:27:53
Feature	1	0	2	2	0	0	0.57	2:32:28
Laudatio	0	1	0	0	0	0	0.11	0:24:58
Anzahl	155	257	121	141	72	134		

3.6.4 Ergebnisse der Tageszeit

Bereits vor der Kodierung entstand die Vermutung, dass Beiträge mit niedrigem Informationsgehalt, wie beispielsweise Musik, eher nachts spielen. Während der Kodierung hat sich diese Vermutung bestätigt. Daher wurde die Tageszeit zum Codingframe hinzugefügt. Tabelle 9 zeigt, welche Codes zu welcher Tageszeit und bei welchem Radiosender vorkommen. Das zeigt neue Erkenntnisse.

Was sofort ins Auge fällt, ist, dass DLF Nova nachts keine Codes aufweist. Dies liegt daran, dass zu dem Zeitraum lediglich Musik abgespielt wurde. Diese ist, wie bereits erwähnt, von der Kodierung ausgeschlossen. Ebenfalls kommen untypische Beiträge, das heißt Beiträge, die eine geringe Häufigkeit aufweisen, eher abends oder nachts vor. Beispiele hierfür sind die Gattungen „Quiz“, „Dankesrede“, „Wort-Musik-Sendung“ oder „Werkeinführung Musik“.

Tendenziell ist der Informationsgehalt aller Radiosender tagsüber und abends höher. Dieser Fakt soll aber nicht verwirren, denn die Zeiträume haben laut Definition eine unterschiedliche Dauer. Morgens ist beispielsweise nur ein Zeitraum von fünf Stunden (05:00 – 10:00 Uhr), während tagsüber ganze acht Stunden (10:00 – 18:00 Uhr) sind. Um die Werte vergleichbarer zu machen, führen wir den Begriff der Informationsdichte ein. Die Informationsdichte beschreibt den relativen Informationsgehalt bezüglich der Dauer. Der Informationsgehalt ist einfach die Anzahl der Codes. Je höher diese Anzahl ist, desto höher ist auch der Informationsgehalt. Der Wert der Informationsdichte berechnet sich durch den Informationsgehalt einer Tageszeit bezogen auf eine stündliche Basis. Genauer ergibt er sich aus:

$$\text{Informationsdichte (Tageszeit)} = \frac{\text{Informationsgehalt (Tageszeit)}}{\text{Dauer der Tageszeit in std}}$$

Durch diese Formel und Tabelle 10 ergibt sich exemplarisch für DLF folgende Werte. Nachts: 6. Morgens: 6,8. Tagsüber: 5,6. Abends: 6. Ähnlich verhält es sich bei den anderen Programmen. Die Werte verdeutlichen, dass der relative Informationsgehalt morgens am stärksten ist. Dies ist nicht verwunderlich, da die meisten Menschen um diese Uhrzeit aufstehen, zur Arbeit fahren und eine hohe Nachfrage an neuen Nachrichten beziehungsweise Informationen fordern. Tagsüber sinkt die Informationsdichte dann und abends steigt sie wieder an. Das ist vermutlich

an den Feierabend geknüpft, denn auch auf dem Weg nach Hause hören viele Menschen Radio und wünschen sich kurze und prägnante Informationen. Nachts bleibt der Informationsgehalt, je nach Programm, auf einem ähnlichen Niveau wie abends, sinkt rapide ab oder ist gar nicht erst erhalten.

Die Simulation von Ausspielreihenfolgen sollte die Tageszeit berücksichtigen. Morgens sollten eher Sendungen mit höherem Informationsgehalt spielen. Tagsüber darf der Informationsgehalt auch etwas niedriger sein, aber dennoch auf einem ähnlich hohen Niveau. Abends sollte der Informationsgehalt wieder ansteigen. Nachts ist ein hoher Informationsgehalt nicht von hoher Wichtigkeit, insbesondere bei einem jüngeren Publikum. Die Sequenzmusteranalyse und die Simulation sollten einzelne Kontexte berücksichtigen können, um solche Unterschiede in der Tageszeit abbilden zu können.

Tabelle 10: Beobachtete Tageszeit

	DLF				DLF Kultur				DLF Nova				Summe
	Nachts	Morgens	Tagsüber	Abends	Nachts	Morgens	Tagsüber	Abends	Nachts	Morgens	Tagsüber	Abends	
Biographie	X	X	X	X		X	X	X		X	X		9
Erziehung und Bildung		X	X	X			X	X		X	X	X	8
Wirtschaft	X	X	X	X		X	X	X		X	X		9
Freizeit											X	X	2
Gesellschaft und Soziales	X	X	X	X	X	X	X	X			X	X	10
Geschichte		X									X	X	2
Recht	X	X	X	X	X	X	X	X		X	X	X	11
Kultur	X	X	X	X	X	X	X	X			X	X	10
Gesundheitswesen und Medizin	X	X	X	X	X	X	X	X		X	X	X	11
Natur			X	X			X			X	X	X	6
Politik	X	X	X	X	X	X	X	X		X	X	X	11
Religion		X	X	X		X	X	X		X	X	X	9
Sport	X	X	X	X	X	X	X	X		X	X		10
Technik		X	X	X			X			X	X	X	7
Umwelt	X	X	X	X			X	X		X	X	X	8
Verkehrswesen	X	X	X	X	X	X	X	X		X	X	X	11
Wissenschaft und Forschung	X		X	X	X	X	X	X			X	X	9
Katastrophe	X	X	X	X	X	X	X	X		X	X	X	11
Medien			X			X	X	X		X	X		6
Keine Klasse												X	1
Bundeslandebene	X	X	X	X	X	X	X	X		X	X	X	11
Deutschland	X	X	X	X	X	X	X	X		X	X	X	11
Europäische Ebene	X	X	X	X	X	X	X	X		X	X	X	11
Außereuropäische Ebene	X	X	X	X	X	X	X	X		X	X	X	11
Kein Geographika	X	X	X	X	X		X	X		X	X	X	10
Collage							X			X			2
Kurzhörspiel							X						1
Interview	X	X	X	X		X	X					X	7
Sketch			X										1
Kommentar		X	X	X		X							4
Magazin		X	X	X	X								4
Nachrichtensendung						X							1
Meldung		X	X				X	X			X		5
Nachrichten	X	X	X	X	X	X	X	X		X	X	X	11
Wetter	X	X	X	X	X	X	X	X		X	X	X	11
Porträt	X	X				X							3
Presseschau	X	X	X	X				X					5
Programmelement										X			1
Teaser						X							1
Ratgebersehung			X										1
Quiz												X	1
Dankesrede	X												1
Predigt			X										1
Reportage	X	X	X			X	X	X					6
Rezension			X	X		X	X	X					5
Sendung mit Hörerbeteiligung			X			X	X						3
Statement			X										1
Tipp			X		X	X	X	X					5
Werkeinführung Musik				X									1
Wort-Musik-Sendung	X				X			X					3
Bericht	X	X	X	X	X	X	X	X			X	X	10
Korrespondentenbericht		X	X	X	X	X	X	X		X	X		8
Gespräch	X		X	X	X	X	X	X		X	X	X	10
Betrachtung		X	X	X		X							4
Diskussion			X	X									2
Konzertaufführung								X					1
Bericht in Interviewform	X	X	X		X	X	X	X		X	X	X	10
Sportbericht	X	X		X									3
Kriminalhörspiel								X					1
Nachruf			X	X			X	X					4
Feature			X		X			X					4
Laudatio	X												1
Summe	30	34	45	36	24	32	38	36	0	26	30	28	

3.7 Anforderungen für die Simulation

Die Einleitung, der Theorieteil und die qualitative Inhaltsanalyse dieser Arbeit haben einen Einblick in die vorhandenen Daten ermöglicht. Dieses Unterkapitel wird nun erste Anforderungen an die Simulation künstlicher Ausspielreihenfolgen stellen. Die Anforderungen sind größtenteils selbst postuliert, wobei unterschiedliche Quellen zur Inspiration beigetragen haben. Dazu gehören vergangene Arbeiten am Lehrstuhl, vorhandene Literatur zu dem Thema, sowie bereits existierende digitale Radios mit Personalisierung. Tabelle 11 zeigt diese Anforderungen auf einem Blick. Es folgt nun eine genauere Erklärung dazu. Die **Personalisierbarkeit** ist hierbei eine erste Anforderung an die Simulation. Der Hörer des nichtlinearen Radios soll entscheiden können, welcher Beitrag als erstes

gespielt wird. Das ist der sogenannte seed point. In diesem Fall ist dieser eine Nutzereingabe. Das stellt sicher, dass zu Beginn ein Thema angesprochen wird, welches ganz sicher den Hörer anspricht. So wird der Hörer sofort an das Medium gebunden.

Weiterhin soll die **Angabe der Sequenzlänge** für das künstliche Radio berücksichtigt werden. Damit kann auch der Nutzer entscheiden, wie viele Beiträge er hören möchte. Weiterhin ist eine variable Sequenzlänge für eine spätere Evaluierung der Methode von Bedeutung. Für die Simulation bedeutet dies, dass ein weiterer Input die Länge der künstlichen Sequenz sein sollte. Ein Input von fünf würde beispielsweise bedeuten, dass eine künstliche Sequenz mit fünf Beiträgen gespielt werden soll.

Eine nächste Anforderung ist das Einsetzen der **Musteranalyse**. Dabei werden die wichtigsten Muster, das heißt Muster mit einem hohen Support, genutzt. Dazu ist noch eine Sequenzmusteranalyse nötig, welches im nächsten Kapitel folgt.

Der **Kontext der Beiträge** wird ebenfalls berücksichtigt, was demnach die nächste Anforderung darstellt. Das kann ein gewünschtes Programm sein oder auch der Tag, an dem das Radio gespielt wird. Wie bereits in der qualitativen Inhaltsanalyse gezeigt wurde, werden im linearen Radio an Wochentagen mehr Themen behandelt als am Wochenende. Weiterhin bestimmt die Tageszeit, wie hoch die Informationsdichte ist. Morgens ist die höchste Informationsdichte gegeben. Nachts hingegen ist die Informationsdichte niedrig. Um diese Anteile sicher zu stellen und um die Kontextsensitivität der Simulation zu gewährleisten, sollten die künstlichen Sequenzen den Kontext berücksichtigen und dementsprechend geeignete Beiträge beziehungsweise Muster auswählen.

Zuletzt ist der **Rundfunkvertrag** ein Aspekt der nicht zu vernachlässigen sein sollte. Dieser schreibt bestimmte Programmgrundsätze vor. In der qualitativen Inhaltsanalyse wurde beispielsweise schon erkannt, dass Politik ein am meisten gespieltes Themengebiet im linearen Radio ist. Doch was ist, wenn ein Hörer sich überhaupt nicht für dieses Thema interessiert? Damit der Hörer sich so lange wie möglich mit dem Medium beschäftigt, ist es nahliegend, wenn er das zu hören bekommt, was er wirklich möchte. Um jedoch die Gesetze im Rundfunkvertrag nicht zu verletzen ist hier ein Trade Off die Lösung. Die Nutzereingabe zu Beginn sorgt dafür, dass die Interessen des Hörers vertreten sind. Die danach verwendeten Muster stellen sicher, dass dennoch ein ausgewogener Mix an Themen und Gattungen

gegeben ist. Ebenso verhält es sich mit der Bezugsreichweite. Diese kann für den Beginn, durch Wunsch des Hörers, einseitig sein, aber sollte sich nach einiger Zeit wieder auf einen ausgeglichenen Zustand einpendeln.

Tabelle 11: Anforderungen an die Simulation

Personalisierbarkeit	Die Simulation sollte eine Interaktion mit dem Hörer ermöglichen durch beispielsweise einer Eingabe, welches Element zu Beginn gespielt werden soll (= seed point).
Angabe der Sequenzlänge	Die Simulation sollte variable Sequenzlängen (= Anzahl an Beiträgen) ermöglichen.
Musteranalyse	Die Elemente der künstlichen Sequenz sollten auf einer vorab durchgeführten Musteranalyse basieren um eine hohe Ähnlichkeit zum linearen Radio zu gewährleisten.
Kontext der Beiträge	Die erzeugte Sequenz sollte Beiträge beinhalten, die dem gewählten Kontext entsprechen. Das könnte beispielsweise eine bestimmte Tageszeit sein.
Rundfunkvertrag	Die Regelungen im Rundfunkvertrag sollten an keiner Stelle verletzt werden.

4 Sequenzmusteranalyse

Dieses Kapitel widmet sich der Sequenzmusteranalyse. Es bezeichnet die Suche nach den am häufigsten vorkommenden Mustern in einer Sequenzdatenbank. Im Kontext dieser Arbeit wird nach Mustern in der Abfolge von Radiobeiträgen geschaut. Das erste Unterkapitel widmet sich den Annahmen und den möglichen Dimensionen der Sequenzmusteranalyse. Diese basieren auf der gegebenen Datenquelle, die in der qualitativen Inhaltsanalyse bereits Verwendung fand. Im zweiten Schritt folgt die Anwendung der Sequenzmusteranalyse auf die bereits kodierten Beiträge. Im letzten Unterkapitel wird dann erläutert, was diese Ergebnisse für die Forschung und die Simulation der Abspielreihenfolgen bedeuten.

4.1 Vorbereitung

Um die Sequenzmusteranalyse anwenden zu können müssen erst einige grundlegende Entscheidungen getroffen werden. Insgesamt drei Punkte sind zu beachten:

- a. Festlegen einer Sequenzlänge
- b. Wahl einer Visualisierungsmethode
- c. Wahl des Kontexts

Die erste Entscheidung betrifft die **Sequenzlänge** der Muster. Die Entscheidung über die Sequenzlänge liegt im Ermessen der jeweiligen Person, die sie ausführt. In dieser Arbeit wird eine Sequenzlänge von vier verwendet. Damit sind die zu betrachtenden Pattern nicht zu kurz, aber auch nicht zu lang. Die spätere Visualisierung der Muster bleibt somit übersichtlich.

Als zweites stellt sich die Frage, welche **Visualisierungsmethode** für die Veranschaulichung der sequenziellen Muster geeignet ist. Besonders bei großen Datenmengen können visuelle Hilfsmittel dazu beitragen, neue Erkenntnisse zu gewinnen. Diese Arbeit nutzt Sankey-Diagramme, welche ähnlich zu Flussdiagrammen sind. Der Unterschied zu Flussdiagrammen ist, dass die Dicke des Flusses, die Mengen beziehungsweise in diesem Fall den Support der Muster darstellt. Dadurch erkennt man auf Anhieb, wie wichtig ein Muster ist, das heißt wie häufig es vorkommt. Ein weiterer Vorteil von Sankey-Diagrammen ist, dass man erkennt an welcher Stelle ein bestimmtes Itemset im Muster steht. In den Grafiken ist dies durch folgende Syntax gekennzeichnet: <Itemset A (1)>. In diesem Beispiel bedeutet die 1, dass das Itemset A im Muster an erster Stelle vorkommt. Ein Itemset kann auch mehrere Items besitzen. Das bedeutet dann, dass diese beiden Codes häufig zusammen vorkommen, vorausgesetzt die Dicke des Flusses ist für dieses Itemset ausgeprägt.

Die letzte Entscheidung betrifft die Dimensionen beziehungsweise den **Kontext** der bereits kodierten Daten. Momentan existiert nur eine kodierte Tabelle mit verschiedenen Programmen, Tagen und Uhrzeiten. Somit gibt es viele Filtermöglichkeiten in Bezug auf den Kontext, den man sich wünscht. Im Prinzip existiert ein Datenwürfel mit folgenden Dimensionen:

- a. *Sender*: DLF, Kultur, Nova
- b. *Tag*: Wochentag, Wochenende
- c. *Tageszeit*: 0-5, 5-10, 10-18, 18-24 Uhr
- x. *Gattung*: Hörspiel, Interviews, ...
- y. *Bezugsreichweite*: Deutschland, Bundeslandebene, Europäische Ebene, ...
- z. *Themen-Klasse*: Politik, Kultur, ...

Die letzten drei (Gattung, Bezugsreichweite und Themen-Klasse) sind die Elemente, deren Sequenzen analysiert werden. Die ersten drei (Sender, Tag und Tageszeit) sind variable Ausprägungen. Daraus resultieren insgesamt $3 \times 4 \times 2 \times 3 = 72$ Kontexte (siehe Anhang B). Es ergeben sich somit mannigfache Möglichkeiten, da diese Patterns nicht nur einzeln, sondern auch paar- oder gruppenweise analysiert werden können. Diese Arbeit widmet sich zwei Untersuchungen. Zum einen werden die Kontexte kumuliert, das heißt als Gesamtes betrachtet. Dadurch wird die Sequenzmusteranalyse senderübergreifend durchgeführt, wobei keine Unterscheidung zwischen Wochentag/Wochenende oder Tageszeit folgt. Zum anderen wird für jeden einzelnen Kontext eine Sequenzmusteranalyse durchgeführt. Die Ergebnisse dieser Kontexte werden dann miteinander verglichen. Dazu wird im Folgenden eine neue Metrik eingeführt, welche von Hirschmeier et al. (2020) inspiriert ist. Diese Metrik dient dazu, herauszufinden welche der Kontexte sich ähnlich oder unähnlich sind. Dadurch werden die Ergebnisse der Sequenzmusteranalysen vergleichbar und es können Unterschiede festgemacht werden. Interessiert man sich beispielsweise für die Sequenzen der Klassen in den Kontexten „Deutschlandfunk, Abends, Wochentag“ und „Deutschlandfunk, Abends, Wochenende“, so lässt sich die Ähnlichkeit oder Unähnlichkeit wie folgt bestimmen. Seien A und B zwei Ergebnisse der Sequenzmusteranalyse zweier Unterpatterns und sei m die Anzahl der Top m Muster, das heißt solche mit dem höchsten Support, dann lässt sich die **Unähnlichkeit** dieser Mengen anhand dieser gewichteten Formel bestimmen:

$$\text{Unähnlichkeit}(A_i, B_i) = \frac{\sum_{n=1}^m |\text{sup}_n(A_i) - \text{sup}_n(B_i)|}{\sum_{n=1}^m \max(\text{sup}_n(A_i), \text{sup}_n(B_i))}$$

Die Formel für die **Ähnlichkeit** ist demnach:

$$\text{Ähnlichkeit}(A_i, B_i) = |1 - \text{Unähnlichkeit}(A_i, B_i)|$$

Auffällige Unterpattern, also solche die eine hohe Ähnlichkeit oder eine hohe Unähnlichkeit aufweisen, sollen mithilfe von Sankey-Diagrammen genauer betrachtet werden.

4.2 Anwendung der Sequenzmusteranalyse

Dieses Kapitel ist eine Anwendung des CM-SPAM Algorithmus auf die kodierten Daten. Im Vorkapitel wurde bereits erwähnt, dass zwei Kontexte Anwendung finden werden:

- a. Senderübergreifender Fokus
- b. Kontextueller Fokus

Der senderübergreifende Fokus betrachtet hierbei die gesamten kodierten Daten und sucht dort nach Mustern. Der zweite Fokus ist kontextuell, betrachtet alle 72 Dimensionen und vergleicht diese hinsichtlich ihrer Ähnlichkeit in den Mustern. Ziel ist es hierbei drei Tabellen, jeweils für die Kategorien Klassen, Bezugsreichweite und Gattungen, zu erstellen. Aus den Ergebnissen sollen weitere Anforderungen an die Simulation synthetischer Ausspielreihenfolgen abgeleitet werden.

4.2.1 Senderübergreifender Fokus

Bevor die Arbeit sich mit dem Vergleich der einzelnen Unterpatterns beschäftigt, folgt hier eine Betrachtung des gesamten Datensatzes. Das heißt, die Sequenzmusteranalyse bekommt als Input die komplette Kodierungsliste. Es folgt keine Unterscheidung zwischen Sender, Tag oder Tageszeit. Abbildung 4 zeigt Muster der Länge 4, senderübergreifend und nach den Klassen. Die Wahl der Sequenzlänge soll die Visualisierung anschaulich machen. Weiterhin existieren bei der Wahl einer zu hohen Länge wenige Muster. Eine Visualisierung bringt dann keine Erkenntnisse. Für die vorhandene Datenquelle ist daher eine Länge von 4 festgelegt worden. Man erkennt deutliche Schwerpunkte, die die Sender gemeinsam haben. Bereits bei der Kodierung war auffällig, dass Beiträge, die die Klasse „Kultur“ beinhalten, häufig aufeinanderfolgend aufgetreten sind. Dieses Muster zieht sich über mehrere Itemsets hinweg. Beispielsweise ist „Kultur“ als zweites, drittes oder viertes Element im Muster ebenfalls noch stark ausgeprägt. Weitere häufige Muster beinhalten die Codes „Politik“ und „Gesundheitswesen und Medizin“. Das ist nicht wegen den Nachrichten der Fall. Denn es gibt auch viele Gespräche, Diskussionen oder längere Berichte zu diesen Themen. Die Codes „Gesundheitswesen und Medizin“ sowie „Politik“ kommen auch häufig zusammen

beziehungsweise aufeinanderfolgend vor. Oft wird nämlich über die politischen Maßnahmen und Events geredet, die die aktuelle Pandemie betreffen. Als Minsup wurde 1% gewählt, da ein höherer Wert keine Muster generiert. Ein zu geringer Wert macht die Visualisierung unübersichtlich.

Für die Simulation synthetischer Ausspielreihenfolgen bedeutet dies, dass der Minsup einen kritischen Wert darstellt. Ist dieser zu hoch gewählt, findet man keine Muster. Ist er zu gering, dann ist die Laufzeit der Mustererkennung deutlich länger. Zudem bietet die Sequenzmusteranalyse weitere Parametereingaben. Eine Berücksichtigung von verschiedenen Eingaben ist somit möglich. Dies geschieht, um ein Beispiel zu nennen, indem der Hörer angibt, welche Themen er hören möchte. Daraufhin werden nur die Muster gefiltert, die die entsprechenden Codes beinhalten. Gewählt wird dann das Muster mit dem höchsten Support. Ist der Hörer gleichermaßen interessiert an allen Sendern, dann kann die senderübergreifende Sequenzmusteranalyse eine gute Übersicht über die am häufigsten auftretenden Themen und Mustern bieten.

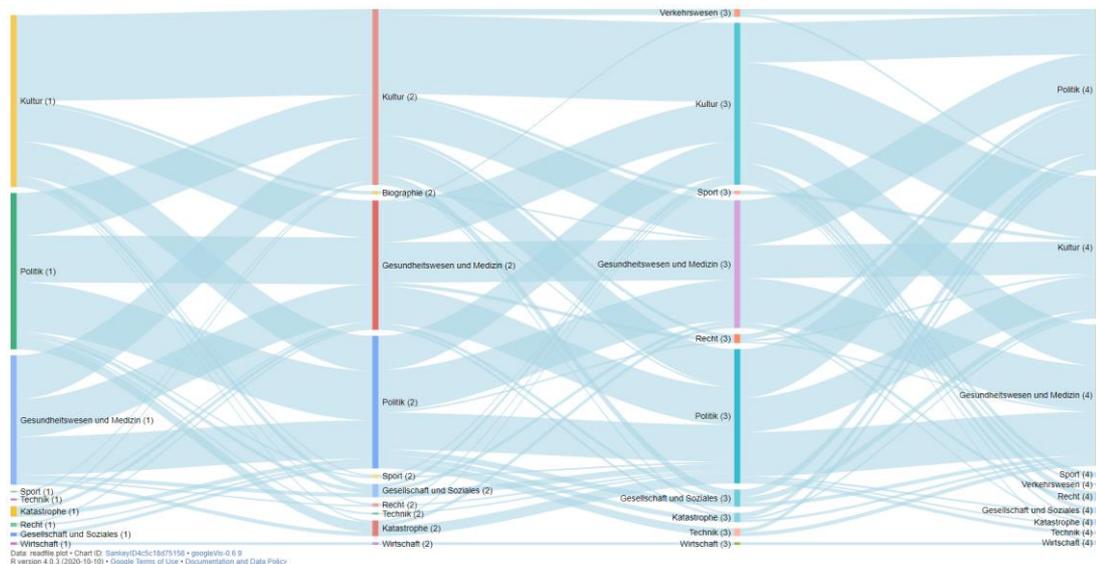


Abbildung 4: Musteranalyse der Länge 4 nach Klassen (Minsup: 1%)

Im nächsten Schritt erfolgt eine Betrachtung der Muster, senderübergreifend und nach der geographischen Bezugsreichweite (siehe Abbildung 5). Logischerweise ist hierbei Deutschland an jeder Stelle der häufigsten Muster stark vertreten. Die Datenquelle richtet sich hauptsächlich an ein deutsches Publikum. Allerdings zeigt die qualitative Inhaltsanalyse, dass „Deutschland“ und „Europäische Ebene“ eine ähnliche Häufigkeit in den Beiträgen aufweisen. Jedoch kommt in der Abbildung die

europäische Ebene, im Verhältnis zu Deutschland, weniger zum Vorschein. Das liegt daran, dass Beiträge mit dem Code „Europäische Ebene“ in der Datenquelle nicht häufig aufeinander folgend vorkommen, sondern oft mit gewissen Abständen. Eine weitere Erkenntnis ist, dass die Codes „Außereuropäische Ebene“ sowie „Europäische Ebene“, wenn sie zuerst spielen, danach meistens wieder zum Code „Deutschland“ umschwenken.

Für die Simulation heißt dies, dass ein Hörer, welcher sich gleichermaßen für alle drei Sender interessiert, Ausspielreihenfolgen erhalten wird, welche einen geographischen Bezug zu Deutschland haben. Selbst wenn eine andere Bezugsreichweite gewünscht ist, ist ein umschwenken zu deutschen Themen sehr wahrscheinlich, es sei denn diese werden vom Hörer explizit ausgeschlossen.

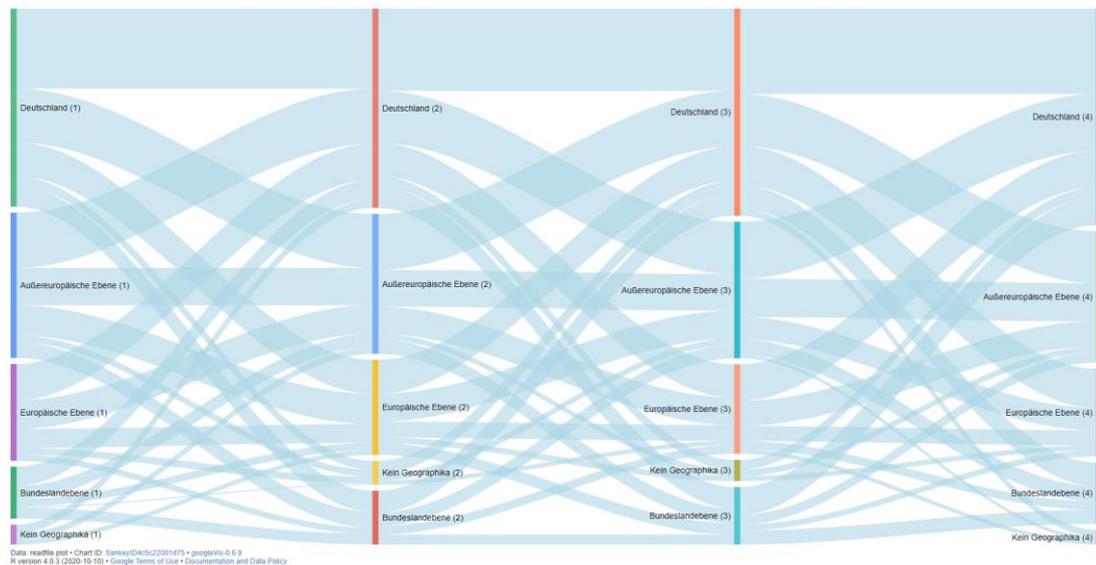


Abbildung 5: Musteranalyse der Länge 4 nach Bezugsreichweite (Minsup: 1%)

Betrachtet man die Muster der Gattungen in Abbildung 6, so fällt die Häufigkeit des Codes „Nachrichten“ auf. Dieser Code kommt besonders an erster Stelle der Muster vor. Nach den Nachrichten folgt typischerweise ein Bericht. Diese Erkenntnis war schon beim Kodieren auffällig. Doch zusätzlich werden mit der Visualisierung Muster sichtbar, die bei dem Kodierungsprozess nicht aufgefallen sind. So ist es beispielsweise ein gängiges Muster, dass auf einen Bericht ein weiterer Bericht folgt. Tatsächlich gibt es häufig Sendungen, die hauptsächlich nur aus Beiträgen der Gattung „Bericht“ bestehen. Weiterhin ist auffällig, dass viele Muster nicht nur mit den Nachrichten anfangen, sondern auch mit ihnen enden.

Für die Simulation synthetischer Ausspielreihenfolgen heißt dies, dass bei Erzeugung längerer Sequenzen (ab Sequenzlänge 3), Nachrichtensendungen mit den Nachrichten anfangen, aber nach einer gewissen Folge von Beiträgen auch abschließen sollten. Zudem sollte Vorsicht geboten sein, denn manche Muster bestehen an jeder Stelle des Musters aus demselben Code. Um zu viele Wiederholungen zu vermeiden, ist eine neue Anforderung an die Simulation, dass mit Wahrscheinlichkeiten bei der Auswahl der Muster gearbeitet werden sollte. Wiederholungen vermeiden heißt hier aber nicht, dass sie niemals vorkommen dürfen, denn auch im linearen Radio kommt dies gelegentlich vor.

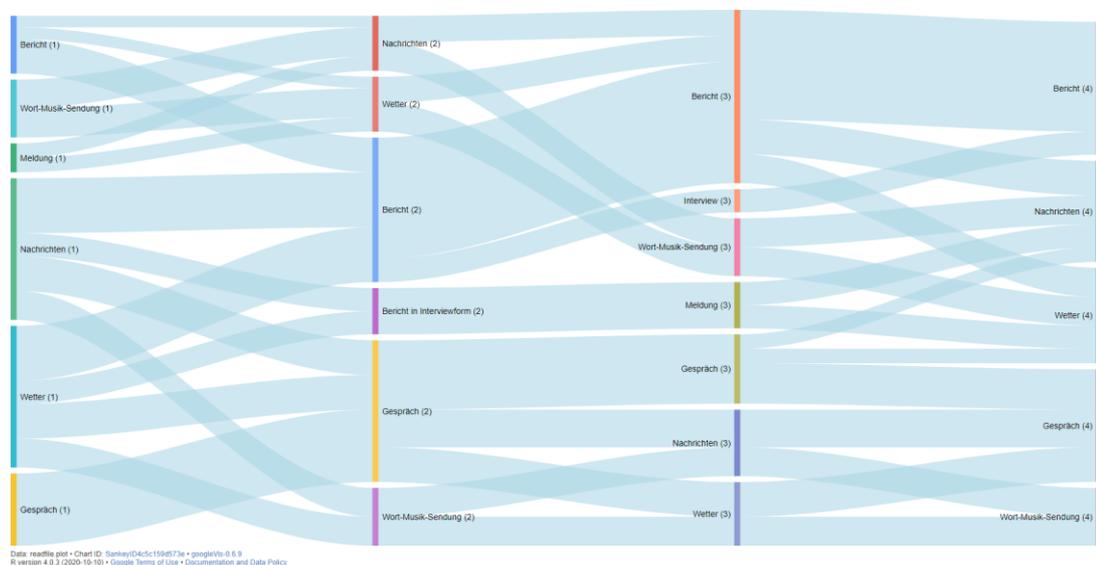


Abbildung 6: Musteranalyse der Länge 4 nach Gattungen (Minsup: 0.5%)

4.2.2 Kontextueller Fokus

In diesem Kapitel fällt der Fokus auf den Vergleich der einzelnen Unterpatterns. Durch die bereits vorgestellt Formel für die Ähnlichkeit dieser Patterns lassen sich drei Dreieckstabellen bilden, jeweils für die Oberkategorien Klasse, Bezugsreichweite und Gattung. Durch die Dreieckstabellen sind alle Dimensionen abgedeckt und auf einen Blick vergleichbar. Werte, die besonders auffällig sind, sollen genauer betrachtet werden.

Abbildung 7 zeigt einen Ähnlichkeitsvergleich nach den Klassen. Werte größer als 0.4 sind grün hervorgehoben. Auffällig sind die niedrigen Werte bei DLF Nova. Nicht verwunderlich, da dieses Programm von den anderen abweicht und eher einem jüngeren Publikum dient. Weiterhin sind besonders nächtliche oder abendliche Programme am Wochenende relativ unähnlich zu den anderen Kontexten. Im

Folgenden bezeichnet die **kumulierte Ähnlichkeit** die aufsummierte Ähnlichkeit eines Kontexts zu allen anderen Kontexten der Datenquelle. Die höchste kumulierte Ähnlichkeit besitzt DLF Kultur, tagsüber und an einem Wochentag.

Unterpattern	[DLF, Wochentag, Nachts]	[DLF, Wochentag, Morgen]	[DLF, Wochentag, Tagüber]	[DLF, Wochentag, Abends]	[DLF, Wochenende, Nachts]	[DLF, Wochenende, Morgen]	[DLF, Wochenende, Tagüber]	[DLF, Wochenende, Abends]	[DLF Kultur, Wochentag, Nachts]	[DLF Kultur, Wochentag, Morgen]	[DLF Kultur, Wochentag, Tagüber]	[DLF Kultur, Wochentag, Abends]	[DLF Kultur, Wochenende, Nachts]	[DLF Kultur, Wochenende, Morgen]	[DLF Kultur, Wochenende, Tagüber]	[DLF Kultur, Wochenende, Abends]	[DLF Noviz, Wochentag, Nachts]	[DLF Noviz, Wochentag, Morgen]	[DLF Noviz, Wochentag, Tagüber]	[DLF Noviz, Wochentag, Abends]	[DLF Noviz, Wochenende, Nachts]	[DLF Noviz, Wochenende, Morgen]	[DLF Noviz, Wochenende, Tagüber]	[DLF Noviz, Wochenende, Abends]
[DLF, Wochentag, Nachts]	1	0.1624	0.3636	0.2075	0	0.0424	0.1106	0.0365	0.1186	0.3846	0.5473	0.4706	0.0174	0.3049	0.4257	0.1282	0.0742	0.1533	0.0562	0	0.068	0.0577		
[DLF, Wochentag, Morgen]	0.1624	1	0.6597	0.0922	0	0.024	0.3502	0	0	0.0661	0.0674	0	0	0	0	0	0.3824	0.3344	0.0622	0.0108	0.1244	0		
[DLF, Wochentag, Tagüber]	0.3636	0.6597	1	0.1551	0	0.0426	0.3629	0	0	0.1837	0.3512	0.1671	0	0.1317	0.2438	0.0643	0.4427	0.3995	0.1091	0	0.1368	0.0367		
[DLF, Wochentag, Abends]	0.2075	0.0922	0.1551	1	0.0161	0	0.1714	0.0411	0	0.1222	0.6078	0.2297	0	0.1466	0.4498	0.1711	0.067	0.0673	0	0.008	0.0506	0.0364		
[DLF, Wochenende, Nachts]	0	0	0	0.0161	1	0.025	0	0.0252	0.0473	0	0	0	0	0	0	0.0364	0	0	0	0	0	0	0	0
[DLF, Wochenende, Morgen]	0.0424	0.024	0.0426	0	0.025	1	0.0311	0	0	0.093	0	0.0172	0	0.1161	0	0	0	0	0	0	0	0	0	0.0274
[DLF, Wochenende, Tagüber]	0.1106	0.3502	0.3629	0.1714	0	0.0311	1	0.0133	0	0.0642	0.4498	0.1121	0	0.0821	0.3769	0.25	0.1263	0.0425	0.0075	0	0	0	0	0.0174
[DLF, Wochenende, Abends]	0.0365	0	0	0.0411	0.0252	0	0.0133	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[DLF Kultur, Wochentag, Nachts]	0.1186	0	0	0	0.0473	0	0	0	1	0.0957	0	0.2605	0.1761	0.0932	0	0.022	0	0	0	0	0	0	0	0.1039
[DLF Kultur, Wochentag, Morgen]	0.3846	0.0661	0.1837	0.1222	0	0.093	0.0642	0	0.0957	1	0.3588	0.3951	0.0265	0.2864	0.2574	0.0778	0.067	0.0429	0	0	0.0216	0.0746		
[DLF Kultur, Wochentag, Tagüber]	0.5473	0.0674	0.3512	0.6078	0	0	0.4498	0	0	0.3588	1	0.607	0	0.4848	0.7888	0.4238	0.0351	0.0313	0	0	0.0193	0.0955		
[DLF Kultur, Wochentag, Abends]	0.4706	0	0.1671	0.2297	0	0.0172	0.1121	0	0.2605	0.3951	0.607	1	0.0298	0.396	0.4448	0.1623	0	0	0	0	0	0	0.1429	
[DLF Kultur, Wochenende, Nachts]	0.0174	0	0	0	0	0	0	0	0.1761	0.0265	0	0.0298	1	0	0	0.0275	0	0	0.0196	0	0	0	0.0238	
[DLF Kultur, Wochenende, Morgen]	0.3049	0	0.1317	0.1466	0	0.1161	0.0821	0	0.0932	0.2864	0.4848	0.396	0	1	0.3929	0.0795	0	0	0	0	0	0	0	0.1135
[DLF Kultur, Wochenende, Tagüber]	0.4257	0	0.2438	0.4498	0	0	0.3769	0	0	0.2574	0.7888	0.4448	0	0.3929	1	0.3744	0	0	0	0	0	0	0	0.0612
[DLF Kultur, Wochenende, Abends]	0.1282	0	0.0643	0.1711	0.0364	0	0.25	0	0.022	0.0778	0.4238	0.1623	0.0275	0.0795	0.3744	1	0	0	0	0	0	0	0	0.0206
[DLF Noviz, Wochentag, Nachts]	0.0742	0.3824	0.4427	0.067	0	0	0.1263	0	0	0.067	0.0351	0	0	0	0	0	1	0.3213	0	0	0	0.107	0	
[DLF Noviz, Wochentag, Morgen]	0.1533	0.3344	0.3995	0.0673	0	0	0.0425	0	0.0429	0.0313	0	0	0	0	0	0	0.3213	1	0.0543	0	0.0841	0		
[DLF Noviz, Wochentag, Tagüber]	0.0562	0.0622	0.1091	0	0	0	0.0075	0	0	0	0	0.0196	0	0	0	0	0	0.0543	1	0	0.0148	0		
[DLF Noviz, Wochentag, Abends]	0	0.0108	0	0.008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0		
[DLF Noviz, Wochenende, Nachts]	0.068	0.1244	0.1368	0.0506	0	0	0	0	0.0216	0.0193	0	0	0	0	0	0	0.107	0.0841	0.0148	0	1	0		
[DLF Noviz, Wochenende, Morgen]	0.0577	0	0.0367	0.0364	0	0.0274	0.0174	0	0.1039	0.0746	0.0955	0.1429	0.0238	0.1135	0.0612	0.0206	0	0	0	0	0	0	1	
Summe	3.7296	2.3361	3.8507	2.64	0.15	0.4187	2.5683	0.1161	0.9173	2.6176	4.868	3.4352	0.3207	2.6276	3.8157	1.838	1.623	1.5309	0.3237	0.0188	0.6267	0.8116		

Abbildung 7: Ähnlichkeit der Unterpatterns nach Klassen (Minsup: 1%)

Abbildung 8 zeigt das entsprechende Sankey-Diagramm. Auffällig ist, dass dieses ähnlich zu dem senderübergreifenden Diagramm ist. Dementsprechend ist es auch ähnlich zu den meisten anderen Unterpatterns.

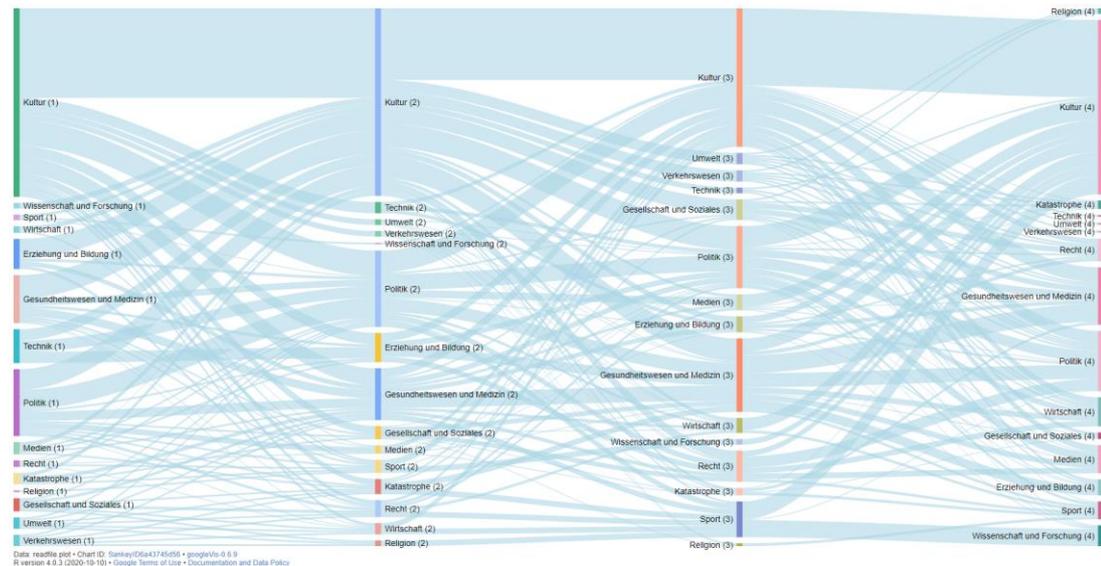


Abbildung 8: Klassen-Muster DLF Kultur, tagsüber, Wochentag

Die geringste kumulierte Ähnlichkeit hat DLF Nova, morgens und am Wochenende. Abbildung 9 veranschaulicht dieses Untermuster. Dort ist besonders auffällig, dass jedes der Muster mit den Codes „Biographie“ oder „Sport“ endet. Das ist höchstwahrscheinlich nie oder kaum der Fall bei den anderen Untermustern. So ergibt sich eine geringe Ähnlichkeit. Zu beachten ist, dass dieses Untermusters nur eine geringe Anzahl an Beiträgen enthält. Wenn mehrere Daten zur Verfügung stehen kann das Ergebnis anders aussehen. Für die vorhandene Menge an Daten ist dieser Kontext aber sehr verschieden zu den anderen.

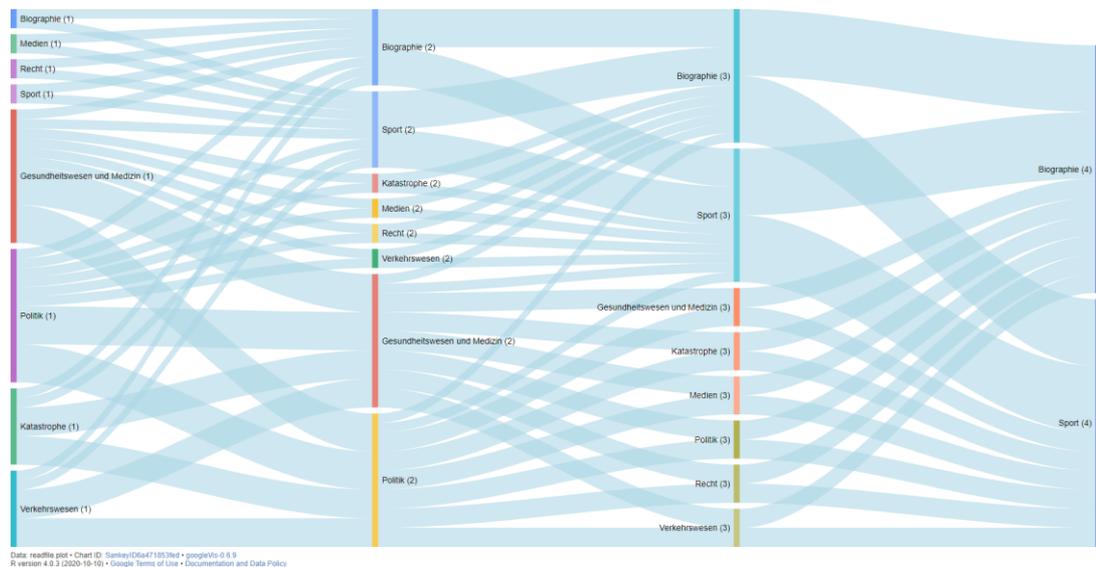


Abbildung 9: Klassen-Muster DLF Nova, morgens, Wochenende

Abbildung 10 zeigt die Ähnlichkeit der Untermustern nach der Bezugsreichweite. Werte größer als 0.6 sind wieder in grüner Farbe hervorgehoben. Geringe Werte sind auch hier meistens abends, nachts oder bei DLF Nova.

Unterpattern	[DLF, Wochentag, Nacht]	[DLF, Wochentag, Morgen]	[DLF, Wochentag, Tagüber]	[DLF, Wochentag, Abends]	[DLF, Wochenende, Nacht]	[DLF, Wochenende, Morgen]	[DLF, Wochenende, Tagüber]	[DLF, Wochenende, Abends]	[DLF Kultur, Wochentag, Nacht]	[DLF Kultur, Wochentag, Morgen]	[DLF Kultur, Wochentag, Tagüber]	[DLF Kultur, Wochentag, Abends]	[DLF Kultur, Wochenende, Nacht]	[DLF Kultur, Wochenende, Morgen]	[DLF Kultur, Wochenende, Tagüber]	[DLF Kultur, Wochenende, Abends]	[DLF Nova, Wochentag, Morgen]	[DLF Nova, Wochentag, Tagüber]	[DLF Nova, Wochentag, Abends]	[DLF Nova, Wochenende, Morgen]	[DLF Nova, Wochenende, Tagüber]	[DLF Nova, Wochenende, Abends]
[DLF, Wochentag, Nacht]	1	0.7896	0.8022	0.7952	0	0.2679	0.6787	0.6154	0	0.7677	0.5352	0.1119	0	0.5556	0.4638	0.032	0.875	0.5183	0.415	0	0.1674	0
[DLF, Wochentag, Morgen]	0.7896	1	0.7451	0.5974	0	0.2637	0.6397	0.5813	0	0.6398	0.6214	0.1579	0	0.5279	0.4876	0.0376	0.8654	0.5516	0.3146	0	0.382	0
[DLF, Wochentag, Tagüber]	0.8022	0.7451	1	0.7848	0	0.3203	0.675	0.6667	0	0.7173	0.5269	0.1244	0	0.5396	0.5714	0.2126	0.7225	0.4552	0.405	0.009	0.1795	0
[DLF, Wochentag, Abends]	0.7952	0.5974	0.7848	1	0	0.3301	0.7256	0.6353	0	0.7047	0.33	0.0851	0	0.4294	0.4583	0.1587	0.6222	0.3876	0.3986	0	0.0917	0
[DLF, Wochenende, Nacht]	0	0	0	0	1	0.0699	0	0	0.2111	0	0	0.0222	0.1786	0	0.0342	0	0	0.1321	0.0417	0	0	0
[DLF, Wochenende, Morgen]	0.2679	0.2637	0.3203	0.3301	0.0699	1	0.3568	0.2166	0	0.2295	0.1733	0.0131	0	0.2428	0.2186	0.0719	0.2341	0.1493	0.1613	0	0.0601	0
[DLF, Wochenende, Tagüber]	0.6787	0.6397	0.675	0.7256	0	0.3568	1	0.6579	0	0.6533	0.438	0.0774	0	0.5156	0.487	0.0955	0.6223	0.4169	0.4068	0.0353	0.0941	0
[DLF, Wochenende, Abends]	0.6154	0.5813	0.6667	0.6353	0	0.2166	0.6579	1	0	0.6512	0.3423	0.1364	0	0.3777	0.5086	0.1027	0.4527	0.2591	0.3412	0	0.046	0
[DLF Kultur, Wochentag, Nacht]	0	0	0	0	0.2111	0	0	0	1	0	0	0.0297	1	0	0.0229	0	0	0	0.0636	0	0	0.1765
[DLF Kultur, Wochentag, Morgen]	0.7677	0.6398	0.7173	0.7047	0	0.2295	0.6533	0.6512	0	1	0.3222	0.1333	0	0.4795	0.4512	0.0531	0.6725	0.4515	0.3016	0	0.0931	0
[DLF Kultur, Wochentag, Tagüber]	0.5352	0.6214	0.5269	0.33	0	0.1733	0.438	0.3423	0	0.3222	1	0.0942	0	0.3758	0.3228	0.0538	0.6132	0.3655	0.1579	0	0.2287	0.0088
[DLF Kultur, Wochentag, Abends]	0.1119	0.1579	0.1244	0.0851	0.0222	0.0131	0.0774	0.1364	0.0297	0.1333	0.0942	1	0.0316	0.0776	0.0732	0.0353	0.1783	0.0833	0.0286	0	0.092	0
[DLF Kultur, Wochenende, Nacht]	0	0	0	0	0.1786	0	0	0	0	0	0	0.0316	1	0	0.024	0	0	0	0.0577	0	0	0.1935
[DLF Kultur, Wochenende, Morgen]	0.5556	0.5279	0.5396	0.4294	0	0.2428	0.5156	0.3777	0	0.4795	0.3758	0.0776	0	1	0.3061	0.0283	0.529	0.2724	0.2957	0	0.2151	0
[DLF Kultur, Wochenende, Tagüber]	0.4638	0.4876	0.5714	0.4583	0.0342	0.2186	0.487	0.5086	0.0229	0.4512	0.3228	0.0732	0.024	0.3061	1	0.1101	0.3538	0.108	0.2519	0.0323	0.0539	0.0549
[DLF Kultur, Wochenende, Abends]	0.032	0.0376	0.2126	0.1587	0	0.0719	0.0955	0.1027	0	0.0531	0.0538	0.0353	0	0.0283	0.1101	1	0.0117	0.0534	0.0213	0	0.0238	0
[DLF Nova, Wochentag, Morgen]	0.875	0.8654	0.7225	0.6212	0	0.2341	0.6213	0.4517	0	0.6725	0.6132	0.1783	0	0.529	0.3538	0.0117	1	0.6427	0.3014	0	0.4302	0
[DLF Nova, Wochentag, Tagüber]	0.5183	0.5516	0.4552	0.3876	0.1321	0.1493	0.4169	0.2591	0	0.4515	0.3655	0.0833	0	0.2724	0.108	0.0534	0.6427	1	0.1486	0.0413	0.1838	0
[DLF Nova, Wochentag, Abends]	0.415	0.3146	0.405	0.3986	0.0417	0.1613	0.4068	0.3412	0.0636	0.3016	0.1579	0.0286	0.0577	0.2957	0.2519	0.0213	0.3014	0.1486	1	0.0481	0.0656	0.0704
[DLF Nova, Wochenende, Morgen]	0	0	0.009	0	0	0	0.0353	0	0	0	0	0	0	0	0.0323	0	0	0.0413	0.0481	1	0.0112	0.0303
[DLF Nova, Wochenende, Tagüber]	0.1674	0.382	0.1795	0.0917	0	0.0601	0.0941	0.046	0	0.0931	0.2287	0.092	0	0.2151	0.0538	0.0328	0.4302	0.1838	0.0656	0.0112	1	0
[DLF Nova, Wochenende, Abends]	0	0	0	0	0	0	0	0	0.1765	0	0.0088	0	0.1935	0	0.0549	0	0	0.0704	0.0303	0	0	1
Summe	8.3908	8.2025	8.4573	7.5349	0.6898	3.3793	7.5759	6.5911	1.5038	7.3215	5.5102	1.5854	1.4854	5.7678	5.3945	1.1019	8.127	5.2206	4.2963	0.2074	2.4182	0.5345

Abbildung 10: Ähnlichkeit der Unterpatterns nach Geographika (Minsup: 1%)

Die höchsten kumulierten Ähnlichkeitswerte hat DLF, tagsüber und an einem Wochentag. Abbildung 11 zeigt das dementsprechende Sankey-Diagramm. Auffällig ist die Ähnlichkeit zur senderübergreifenden Abbildung. Bei beiden Visualisierungen ist Deutschland am stärksten und auch an jeder Stelle der Muster vertreten.

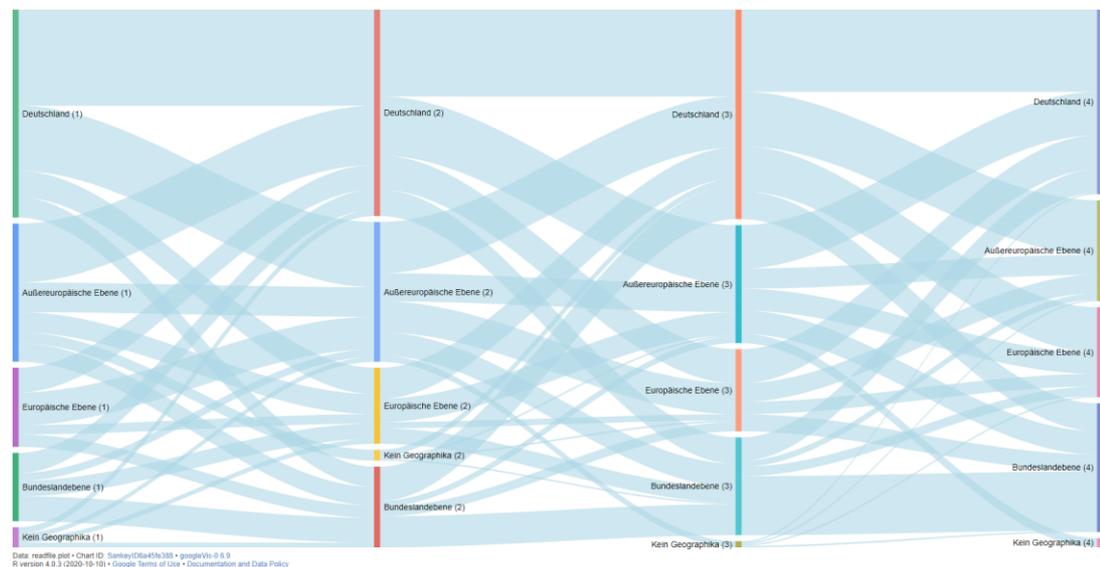


Abbildung 11: Geographika-Muster DLF, tagsüber, Wochentag

Die niedrigsten kumulierten Ähnlichkeitswerte weist DLF Nova, morgens und am Wochenende auf. Abbildung 12 zeigt, dass jedes Muster mit dem Code „Kein Geographika“ endet. Dieser Code kommt in der gesamten Datenquelle selten

vor. Daher ist ein niedriger Wert der Ähnlichkeit keine Überraschung. Auch hier gilt, dass weitere Daten gegebenenfalls zu einem anderen Wert der Ähnlichkeit führen können. Der betrachtete Kontext ist klein, da es nur wenige Beiträge beinhaltet.

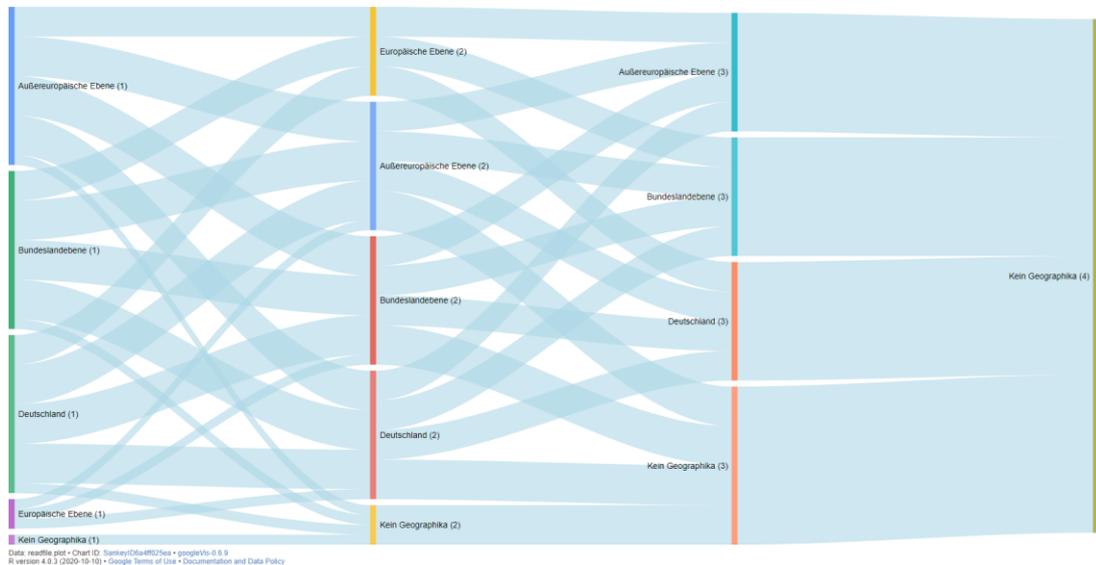


Abbildung 12: Geographika-Muster DLF Nova, morgens, Wochenende

Zuletzt zeigt Abbildung 13 die Ähnlichkeit der Unterpatterns nach den Gattungen. Werte größer als 0.1 sind grün hervorgehoben. Auf dem ersten Blick ist zu erkennen, dass alle Werte gering sind. Die Abspielreihenfolge der Gattungen ist somit in jedem Kontext sehr eigen.

Unterpattern	[DLF, Wochenende, Nacht]	[DLF, Wochenende, Morgens]	[DLF, Wochenende, Tagüber]	[DLF, Wochenende, Abend]	[DLF, Wochenende, Nacht]	[DLF, Wochenende, Morgens]	[DLF, Wochenende, Tagüber]	[DLF, Wochenende, Abend]	[DLF Kultur, Wochenende, Nacht]	[DLF Kultur, Wochenende, Morgens]	[DLF Kultur, Wochenende, Tagüber]	[DLF Kultur, Wochenende, Abend]	[DLF Kultur, Wochenende, Nacht]	[DLF Kultur, Wochenende, Morgens]	[DLF Kultur, Wochenende, Tagüber]	[DLF Kultur, Wochenende, Abend]	[DLF Nova, Wochenende, Tagüber]	[DLF Nova, Wochenende, Abend]	[DLF Nova, Wochenende, Morgens]	[DLF Nova, Wochenende, Nacht]		
[DLF, Wochenende, Nacht]	1	0	0	0	0	0.0909	0.0101	0.0222	0	0	0	0.022	0	0	0.0412	0	0	0	0	0.009	0	
[DLF, Wochenende, Morgens]	0	1	0	0	0	0.0196	0.0101	0	0	0	0	0	0	0	0.01	0	0	0	0	0	0	
[DLF, Wochenende, Tagüber]	0	0	1	0.1742	0	0	0.0294	0.0315	0	0	0	0	0	0	0.0292	0	0	0.0054	0	0	0	
[DLF, Wochenende, Abend]	0	0	0.1742	1	0	0	0.0882	0.0316	0	0	0	0	0	0	0.0566	0	0	0	0	0	0.0172	
[DLF, Wochenende, Nacht]	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
[DLF, Wochenende, Morgens]	0.0909	0.0196	0	0	0	1	0	0	0	0	0	0	0	0	0.0096	0	0	0.0265	0	0	0	
[DLF, Wochenende, Tagüber]	0.0101	0.0101	0.0294	0.0882	0	0	1	0	0	0	0	0	0	0	0.0202	0	0	0	0	0	0	
[DLF, Wochenende, Abend]	0.0222	0	0.0315	0.0316	0	0	0	1	0	0	0	0	0	0	0.022	0	0	0	0	0	0	
[DLF Kultur, Wochenende, Nacht]	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0.3077	0	0	0	0	0	0.1333	
[DLF Kultur, Wochenende, Morgens]	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0.0204	0	0	0	0	0	0	
[DLF Kultur, Wochenende, Tagüber]	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.0707	0	0	0	0.0106	0.0469	0.0648	
[DLF Kultur, Wochenende, Abend]	0.022	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
[DLF Kultur, Wochenende, Nacht]	0	0	0	0	0	0	0	0	0.3077	0	0	0	1	0	0	0	0	0	0	0	0.1111	
[DLF Kultur, Wochenende, Morgens]	0	0	0	0	0	0	0	0	0.0204	0	0	0	0	1	0	0	0	0	0	0	0	
[DLF Kultur, Wochenende, Tagüber]	0.0412	0.01	0.0292	0.0566	0	0.0096	0.0202	0.022	0	0	0.0707	0	0	0	1	0	0	0.0136	0	0.0469	0.1068	
[DLF Kultur, Wochenende, Abend]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
[DLF Nova, Wochenende, Tagüber]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.0135	0	0	0	
[DLF Nova, Wochenende, Abend]	0	0	0.0054	0	0	0.0265	0	0	0	0	0	0	0	0	0.0136	0	0.0135	1	0	0	0.0127	
[DLF Nova, Wochenende, Morgens]	0	0	0	0	0	0	0	0	0	0.0106	0	0	0	0	0	0	0	0	1	0	0.0693	
[DLF Nova, Wochenende, Nacht]	0.009	0	0	0.0172	0	0	0	0	0	0.0648	0	0	0	0	0.1068	0	0	0.0127	0.0693	0.0267	1	
[DLF Nova, Wochenende, Tagüber]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
[DLF Nova, Wochenende, Abend]	0	0	0	0	0	0	0	0	0.1333	0	0	0	0	0	0.1111	0	0	0	0	0	0	
Summe	0.1955	0.0397	0.2697	0.3679	0	0.1466	0.1581	0.1073	0.441	0.0204	0.193	0.022	0.4188	0.0204	0.4268	0	0.0135	0.0717	0.0799	0.1204	0.3065	0.2444

Abbildung 13: Ähnlichkeit der Unterpatterns nach Gattungen (Minsup 1%)

Abbildung 14 zeigt das Sankey-Diagramm vom Kontext DLF, nachts und an einem Wochenende. Dieser hat eine kumulierte Ähnlichkeit von 0, weist also keine Ähnlichkeit zu den anderen Unterpatterns auf. Bei Betrachtung des Diagramms ist auf den ersten Blick eine solche Vermutung nicht sichtbar. Doch vergleicht man die vorkommenden Codes mit den Ergebnissen der qualitativen Inhaltsanalyse, so wird dieses Ergebnis nachvollziehbar. Die Codes „Porträt“ oder „Wort-Musik-Sendung“ kommen im gesamten Datensatz relativ selten vor. In jedem der Muster des betrachteten Kontexts ist eins dieser beiden Codes vorhanden. Diese seltenen Codes intervenieren somit in die Muster und sorgen dafür, dass keine Ähnlichkeit zu den anderen Unterpatterns vorhanden ist. Zudem ist auch hier zu erwähnen, dass der Kontext relativ klein ist, also wenige Beiträge beinhaltet.

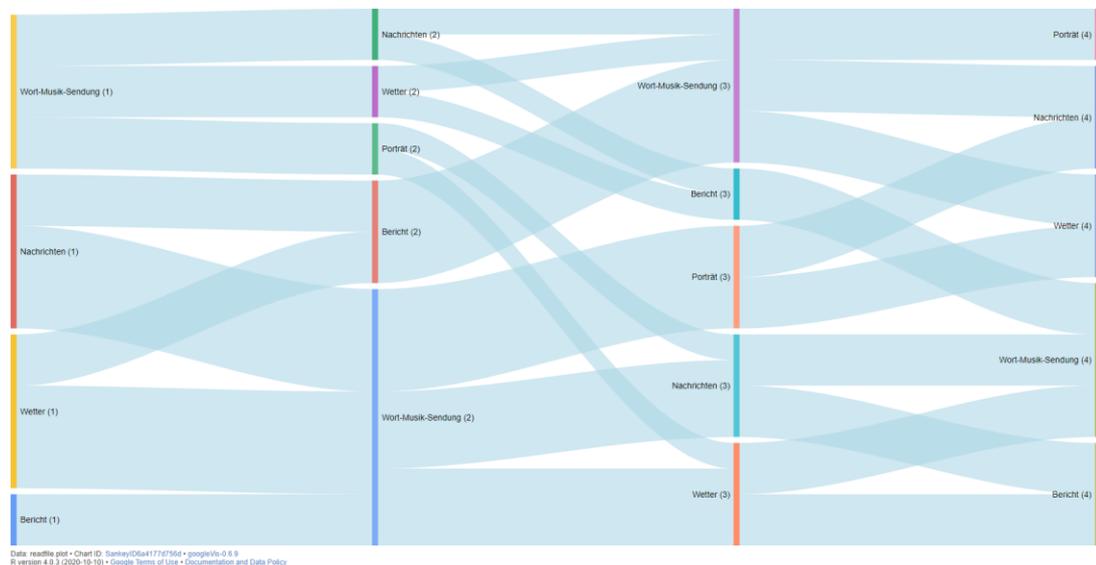


Abbildung 14: Gattung-Muster DLF, nachts, Wochenende

Abbildung 15 zeigt ein Sankey-Diagramm des Kontexts DLF Kultur, nachts und an einem Wochentag. Dieser Kontext weist die höchste kumulierte Ähnlichkeit auf. Doch bei Betrachtung dieses Werts ist Vorsicht geboten, denn dieser Kontext hat nur relativ starke Ähnlichkeiten zu zwei weiteren Unterpatterns. Zu den restlichen Unterpatterns ist keine Ähnlichkeit vorhanden. Sogas kann vorkommen, wenn ganze Beiträge oder Themengebiete für andere Tage übernommen werden.

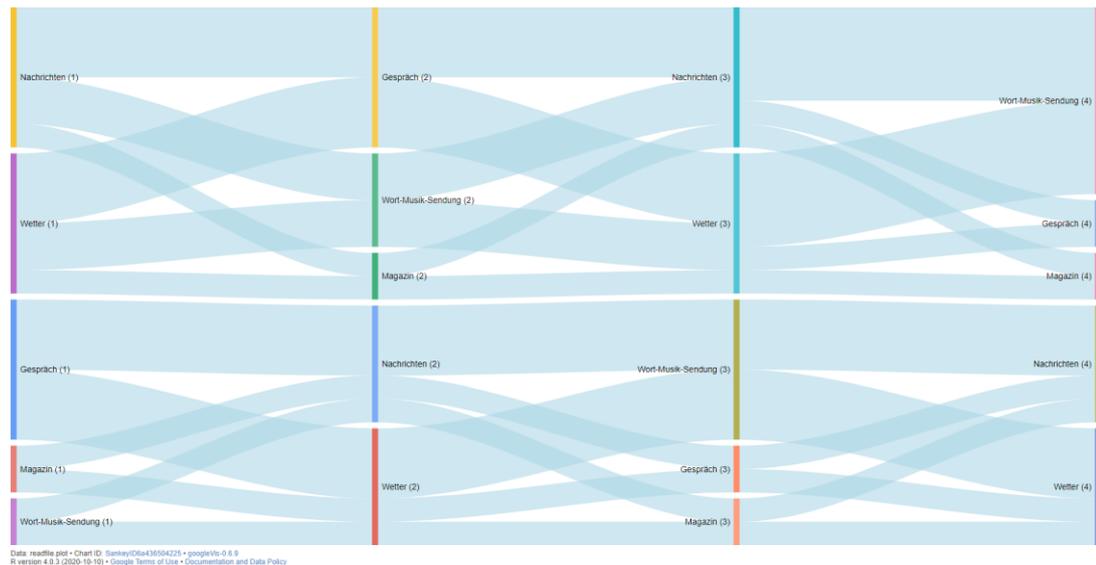


Abbildung 15: Gattung-Muster, DLF Kultur, nachts, Wochentag

Beim Vergleich der Werte in den Dreiecksmatrizen fallen noch einige Aspekte auf. Zunächst würde man meinen, dass sich die Kontexte an den Wochentagen und Wochenenden grundsätzlich unähnlich sind. Doch bei genauer Betrachtung wird ersichtlich, dass einige Wochenendprogramme sogar in höherem Maße den Wochentagprogrammen ähneln. Das bedeutet, dass sich Kontexte am Wochenende und am Wochentag tatsächlich oft ähnlich sind. Bei der Kodierung war bereits auffällig, dass einige Beiträge oder Sendungen sich gedoppelt haben. Beispielsweise ist die Sendung „Informationen am Morgen“ eine Sendung, die bei dem Radioprogramm DLF am Wochenende und auch am Wochentag vorkommt. Dadurch werden identische oder sehr ähnliche Beiträge gespielt, was in einer hohen Ähnlichkeit der verschiedenen Kontexte resultiert. Weiterhin ist auffällig, dass bei Nacht die Beiträge der Kontexte eher eigen sind und dadurch grundsätzlich eine geringe Ähnlichkeit haben.

Für die Simulation bedeuten diese Ergebnisse, dass ein kleinerer Kontext gegebenenfalls wenig Beiträge und somit wenige Muster enthält. Das kann die Simulation beeinträchtigen, da schon eine gewisse Menge an Mustern vorhanden sein sollte, um daraus synthetische Sequenzen zu generieren und diese ausreichend zu evaluieren. Weiterhin sind Informationen über den Hörer hilfreich. Ist dieser eher eine junge Person und hört lieber das DLF Nova Programm, so sollten nur die Muster aus diesem Kontext verwendet werden. Auch sollte der Kontext der

Tageszeit berücksichtigt werden, da beispielsweise abends und nachts besondere und seltene Themen gespielt werden.

Abschließend stellt sich die Frage nach welcher Kategorie (Klasse, Geographika oder Gattung) die Simulation Sequenzen generieren soll. Abbildung 16 fasst die Vor- und Nachteile der jeweiligen Oberkategorien zusammen. Es ist klar, dass für diese Frage keine pauschale Antwort existiert. Es kommt drauf an, wie die Datenquelle aussieht und was die Interessen der Hörer sind. Intuitiv würde man meinen, dass Geographika aufgrund ihrer geringen Anzahl an Codes und ihrer geringen Aussagekraft bezüglich der Beiträge nicht in Frage kommen. Jedoch ist der Fall nicht ausgeschlossen, dass einem Hörer die Klasse und die Gattung gleichgültig ist und er sich tatsächlich nur Beiträge über ein bestimmtes geographisches Gebiet zu hören wünscht. Ein anderer Hörer interessiert sich vielleicht mehr für Hörspiele, wobei hier eine Generierung nach der Gattung sinnvoll wäre. Um den Stundenuhren und der Ähnlichkeit des linearen Radios gerecht zu werden ist, wahrscheinlich eine Generierung nach den Gattungen am naheliegendsten. Nachrichten sind hierbei ein wichtiger Bestandteil des echten Radios. In jedem Falle ist es aber von Bedeutung den Hörer entscheiden zu lassen, was er hören möchte, bei gleichzeitiger Erhaltung des redaktionellen Mehrwerts im linearen Radio. Vorsicht sollte bei der Generierung durch Klassen geboten sein. Wenn der Hörer sich eventuell die Klassen Gesundheitswesen und Politik wünscht, dann ist es möglich, dass bei einer klassenorientierten Generierung der Sequenzen mehrere Beiträge derselben Gattung hintereinander spielen. So ist es gegebenenfalls möglich, dass zwei oder mehr Nachrichtenbeiträge aufeinanderfolgen. Das würde eine nur geringe Ähnlichkeit zum linearen Radio wider spiegeln. Doch auch hier gilt, dass wenn es der Hörer wünscht, es der richtige Ansatz sein könnte. Ein weiterer Nachteil der Oberkategorien Klassen und Bezugsreichweite ist die mehrfache Zuordnung der Codes zu einem Beitrag. Wenn der Hörer beispielsweise nur politische Themen wünscht, so kann es sein, dass ihm Beiträge gespielt werden, die auch andere, für ihn gegebenenfalls uninteressante Themen beinhaltet. Ein Extrembeispiel wäre ein einstündiger Beitrag, der größtenteils über Medien handelt und verhältnismäßig gering auch über Politik. Dann würde ein Hörer, der sich nur politische Beiträge wünscht, auch lange Inhalte konsumieren, die ihn eigentlich nicht interessieren. Dadurch steigt die Gefahr, dass der Nutzer das Radio abschaltet. Hingegen hat die Oberkategorie Gattung größtenteils eine eindeutige Zuordnung. Wünscht der Hörer also eine bestimmte

Gattung, so wird ihm dieser Wunsch komplett gewährt. Allgemein überwiegen die Vorteile dieser Oberkategorie. Im folgenden Kapitel wird daher eine Simulation anhand der Kategorie der Gattungen durchgeführt.

Klassen	Vorteile	⊕ Generierung nach Themengebieten
	Nachteile	⊖ Keine Berücksichtigung von Stundenuhren ⊖ Gewünschte Klasse kann in einem Beitrag verhältnismäßig nur gering enthalten sein
Bezugsreichweite	Vorteile	⊕ Generierung nach Geographika
	Nachteile	⊖ Wenige Ausprägungen ⊖ Keine Berücksichtigung von Stundenuhren ⊖ Gewünschte Bezugsreichweite kann in einem Beitrag verhältnismäßig nur gering enthalten sein
Gattungen	Vorteile	⊕ Generierung nach Gattungen ⊕ Hohe Anzahl an Codes ⊕ Stundenuhren werden am ehesten berücksichtigt ⊕ Eindeutige Zuordnung d.h. ein Beitrag hat meistens nur eine Gattung
	Nachteile	⊖ Keine Berücksichtigung von Themen oder Geographika

Abbildung 16: Vor- und Nachteile der Oberkategorien für die Simulation

5 Simulation von Ausspielreihenfolgen

Die qualitative Inhaltsanalyse sowie die Sequenzmusteranalyse haben Ergebnisse geliefert, die nun Berücksichtigung in der Simulation von künstlichen Ausspielreihenfolgen finden sollen. Dieses Kapitel beginnt mit der Vorstellung des Algorithmus. Dabei wird die Vorgehensweise geschildert und noch einmal aufgefasst, welche Faktoren in der Methode berücksichtigt werden müssen. Gleichzeitig wird erläutert, welche Eingabeparameter die Simulation benötigt und wie die jeweiligen Funktionen der Methode arbeiten. Im selben Kapitel folgt die Vorstellung der Methode in Form eines Pseudocodes. Letztlich folgt im zweiten Kapitel eine Evaluierung der Simulation. Dort werden beispielhafte, künstlich erzeugte Sequenzen hinsichtlich ihrer Ähnlichkeit zu den echten Sequenzen untersucht.

5.1 Vorstellung des Algorithmus

Der Algorithmus verfolgt einen sogenannten *stateful approach*, betrachtet also welchen Beitrag der Hörer gerade hört (Hirschmeier et al., 2020, p. 10). Basierend auf dem aktuellen Beitrag und den vorher gespielten, werden passende Muster gefunden. Diese Muster entscheiden schließlich, welcher Beitrag als nächstes

gespielt werden soll. Um die Personalisierbarkeit zu gewährleisten, wird als Input eine Nutzereingabe erwartet. Der Hörer entscheidet zu Beginn, was für einen Beitrag er als Erstes hören möchte. Ein Beispiel hierfür wäre der Befehl: „Spiele mir die aktuellen Nachrichten“. Das ist auch der *seed point* beziehungsweise der Startpunkt der Methode. Weiterhin benötigt der Algorithmus den jeweiligen Kontext, sprich den Tag, die Tageszeit und den präferierten Sender. Die Betrachtung des Kontextes bietet den Vorteil, dass die passenden Muster zur passenden Zeit gespielt werden. Dadurch wird auch das Konzept des Informationsgehalts umgesetzt, welches in der qualitativen Inhaltsanalyse behandelt wurde. Denn wenn beispielsweise morgens viele Nachrichten und Berichte gespielt werden, dann ergeben sich höchstwahrscheinlich dementsprechende Muster mit einer hohen Informationsdichte. Im ersten Schritt der Methode werden alle Muster identifiziert, die in diesem Kontext enthalten sind. Basierend darauf werden diese, der Nutzereingabe entsprechend, gefiltert. Im Beispiel für die Nachrichten würde die Methode alle Muster filtern, die dieses Element an erster Stelle beinhalten. Diese gefilterten Muster haben, je nach Supportwert, eine unterschiedliche Wahrscheinlichkeit, gewählt zu werden. Das gewählte Muster wird in die synthetische Sequenz aufgenommen und bildet die Grundlage für die weitere Abspielreihenfolge. Die initiale Sequenz könnte beispielsweise so aussehen: abcd, wobei a der Beitrag ist, den der Hörer sich gewünscht hat und d der Beitrag, bei dem der Hörer sich momentan befindet. Im nächsten Schritt betrachtet der Algorithmus das maximal mögliche Zeitfenster. Das wäre die Sequenzlänge der Muster minus eins. Dem Beispiel nach ist das bcd. Nun werden wieder aus allen verfügbaren Mustern diejenigen gefiltert, die bcd am Anfang vorweisen, wie zum Beispiel bcde und bcdf. Aus diesen Mustern wird wieder, abhängig vom Supportwert, das wahrscheinlichste Muster gewählt. Das letzte Element dieses Musters ist dann das nächste Element, welches in die synthetische Sequenz aufgenommen wird. Wird beim Filtern der Muster kein einziges gefunden, dann wird das Zeitfenster um 1 verkürzt und wieder nach Mustern gefiltert. Das geht solange bis entweder ein Muster gefunden wurde oder das Zeitfenster 0 ist. Bei Eintreten des letzten Falls wird ein neues und beliebiges Muster gewählt und in die künstliche Sequenz aufgenommen. Um dem Zufall entgegenzuwirken, wird auch hier der Supportwert herangezogen, um stets die Wahrscheinlichkeiten bei der Auswahl zu beachten. Die Methode läuft solange bis die gewünschte Ziellänge der synthetischen Sequenz erreicht ist. Zu erwähnen ist

noch, dass nach dem Wählen eines Musters die Wahrscheinlichkeit sinkt, dass es noch mal verwendet wird. Nachdem die Wahrscheinlichkeiten für alle Muster bei 0 sind, werden diese wieder auf den ursprünglichen Zustand zurückgesetzt. Dann sind die Wahrscheinlichkeitswerte wieder auf dem initialen Zustand. Das soll sicherstellen, dass bei einer sehr langen gewünschten Sequenzlänge alle Muster Berücksichtigung finden. Es wird also deutlich, dass diese Methode möglichst wenig dem Zufall überlässt und stets den Zusammenhang zwischen den einzelnen Mustern in der Sequenz abzubilden versucht.

Die künstlich erzeugte Sequenz basiert auf Mustern, also Aneinanderreihungen von Codes, des echten Radios. Es erfolgt noch keine Zuordnung von tatsächlich existierenden Beiträgen zu dieser künstlichen Sequenz. Sie ist noch eine leere Hülle, die noch befüllt werden muss mit echten Inhalten, sprich Beiträgen. Soll beispielsweise der Code Nachrichten gespielt werden, so stellt sich die Frage, welcher exakte Beitrag denn gespielt werden soll. Es ergibt sich eine Reihe von Möglichkeiten. Der Hörer könnte selbst entscheiden, welche Nachrichten er sich wünscht, indem er beispielsweise zusätzlich einen Tag angibt. Eine alternative Möglichkeit ist, dass der Algorithmus schlicht die neusten Nachrichten abspielt. Eine weitere Möglichkeit ist die Wahl zufälliger Beiträge oder die Berücksichtigung der Tageszeit. Es existieren also denkbar viele Ansätze. Um die Aktualität des künstlichen Radios zu gewährleisten und keine älteren Beiträge als im nichtlinearen Radio zu spielen, liegt es nahe, möglichst aktuelle Beiträge zur Befüllung der künstlichen Sequenz zu verwenden. Für den Algorithmus bedeutet dies, dass der Kontext auch die genaue Uhrzeit und das Datum erfassen muss, an welchem der Nutzer das Radio hört. Wünscht der Hörer also zu Beginn einen Beitrag über Nachrichten, dann sollten die letzten Nachrichten abgespielt werden. Analog werden die nächsten Codes ebenfalls mit den neusten Beiträgen befüllt. Genauer gesagt erfolgt das Befüllen der Elemente durch tatsächliche Beiträge in dieser Arbeit, indem der aktuelle Zeitpunkt herangezogen wird. Von dort aus wird die kodierte Tabelle rückwärts durchlaufen. Anschließend wird immer der erstgefundene Beitrag dem entsprechenden Sequenzelement zugeordnet. Dadurch ist die Aktualität der Beiträge gewährleistet. Aufgrund der vielfältigen Möglichkeiten und einer Verkomplizierung der Methode erfolgt die Evaluierung der Methode ohne befüllte Codes.

Abbildung 17 zeigt die Vorgehensweise der Methode als Pseudocode. Genauer gesagt wird hier der Kontext definiert als eine Zusammensetzung aus den Informationen Wochentag/Wochenende, Tageszeit und Rundfunkprogramm. Die Nutzereingabe ist das erste Element der Sequenz. Daher muss sie als Element in den Mustern vorkommen, sonst wird keine synthetische Sequenz erzeugt. Programmiert wurde das Ganze mit Python. Als Entwicklungsumgebung wurde PyCharm benutzt. Die Sequenzmusteranalyse zur Erstellung der Regeln wurde vorab durchgeführt. Dafür wurde der Python Wrapper für SPMF vom Autor Lorenz Leitner verwendet. Sowohl für die Sequenzmusteranalyse als auch für die Simulation wurde insbesondere die Bibliothek Pandas zur Verwaltung und Analyse der Daten verwendet. Die Nutzereingabe erfolgt durch eine Spracherkennung, weil dies ein wahrscheinliches Szenario für das neue personalisierte Radio darstellt.

```

FUNCTION generateSequence (userInput, weekday, timeOfDay, broadcastName)
  INIT sequence
  Add userInput To sequence
  rules = getRules(weekday, timeOfDay, broadcastName)
  initialItems = findPattern(sequence, rules, None)
  FOR 1 To sequenceLength DO
    nextPattern = findPattern(sequence, rules, historyWindow)
    nextItem = lastItemOfNextPattern
    Add nextItem To sequence
  END FOR
  RETURN sequence

FUNCTION findPattern(sequence, rules, historyWindow)
  INIT pattern
  IF historyWindow = None
    pattern = choosePatternThatBeginWithInitialItemOfSequence
  ELSE IF historyWindow == 0
    pattern = chooseAnyPattern
  ELSE
    pattern = choosePatternThatContainsLastItemsOfSequence
    IF pattern == NULL
      findPattern(requiredItem, rules, historyWindow-1)
    END IF
  END IF
  rules.updateProbabilities()
  RETURN pattern
END FUNCTION

```

Abbildung 17: Pseudocode für die Simulation synthetischer Ausspielreihenfolgen

5.2 Evaluierung der Methode

Die Evaluierung des Algorithmus wird mithilfe eines Vergleichs zwischen der Ähnlichkeit echter Abspielreihenfolgen zu synthetischen Sequenzen

durchgeführt. Das geschieht mit der Ähnlichkeitsformel, die auch bei der Sequenzmusteranalyse verwendet wurde. Die Methode performte bei den ersten Versuchen nicht zufriedenstellend und wies nur eine geringe Ähnlichkeit zum echten Radio vor. Der Grund lag darin, dass die Methode die falschen Muster wählte. Beispielsweise wurde im ersten Programm stets das Muster mit dem höchsten Support gewählt, ohne dass auch andere Muster eine Chance zur Auswahl hatten. Dadurch kam es vor, dass teilweise immer dasselbe Muster hintereinander gespielt wurde. Daher wurde der Algorithmus verfeinert, indem Wahrscheinlichkeiten für jedes Muster eingeführt wurden. Jedoch waren die Ergebnisse ebenfalls nicht zufriedenstellend, da die Wahrscheinlichkeiten stets konstant blieben. Demzufolge war hier ebenfalls eine Verfeinerung notwendig. Die überzeugendste, finale Methode wählt nicht automatisch das Muster mit dem höchsten Support, sondern betrachtet stets die Wahrscheinlichkeiten und aktualisiert diese, nachdem ein Muster gewählt wurde. Dadurch wird gewährleistet, dass jedes Muster gewählt werden kann und gleichzeitig wird auch der Supportwert berücksichtigt, in dem Muster mit einem hohen Support eine höhere Wahrscheinlichkeit haben, gewählt zu werden.

Die Evaluierung erfolgte zu Beginn mit beliebigen Kontexten. Die Ergebnisse hierfür waren aber nicht wirklich bedeutungsvoll. Das liegt daran, dass die Unterpatterns aus relativ wenigen Beiträgen bestehen. Es existieren häufig nur wenige Muster, die nicht sehr aussagekräftig sind, da der Supportwert nahezu aller Muster bei lediglich 1 liegt. Demnach ist für die Evaluierung ein Input mit einer ausreichenden Anzahl an Beiträgen notwendig. Daher wurde die Evaluierung für einen größeren Kontext durchgeführt. Es wurden alle Tage des Senders Deutschlandfunk herangezogen. Das stellt sicher, dass der Input vielfältig und realistisch ist. Festzuhalten ist also, dass für die Evaluierung der Methode ein geeigneter Input erforderlich ist.

Weiterhin sollte die künstlich erzeugte Sequenz nicht zu kurz sein. Denn bei der synthetischen Sequenz wird ebenfalls eine Sequenzmusteranalyse vollzogen. Dafür ist vorausgesetzt, dass genügend Muster vorhanden sein müssen. Das Argument, dass auch kurze synthetische Sequenzen vergleichbar sein müssen, ist durchaus valide. Jedoch wird bei der vorgestellten Methode das erste Element der Sequenz vom Nutzer gewählt und nicht vom Algorithmus. Dies lässt sich gut anhand eines Beispiels erklären. Man nehmen an, die Musteranalyse des echten Radioprogramms habe ergeben, dass auf Nachrichten häufig ein Bericht folgt. Der

Nutzer wählt zu Beginn jedoch ein eher seltenes Element, wie beispielsweise ein Hörspiel. Dann werden mit hoher Wahrscheinlichkeit erst untypische Muster gespielt und die Top m Muster, die im echten Radio häufig vorkommen, werden in der synthetischen Sequenz erst später gespielt.

Sind diese beiden Bedingungen gegeben, sprich ein geeigneter Input sowie eine künstliche Sequenz mit einer ausreichenden Länge, dann ist eine Evaluierung sinnvoll. Was ausreichend ist, lässt sich allerdings nicht einfach so sagen. Daher werden unterschiedliche Längen bei der Evaluierung betrachtet.

Der finale Ähnlichkeitsvergleich zwischen dem echten Radio und der synthetischen Sequenz bestätigt diese Beobachtungen. Abbildung 18 ist ein exemplarischer Ähnlichkeitsvergleich zwischen einer synthetischen Sequenz und den Mustern des DLF Programms in Bezug auf die Gattungen. Für die Sequenzlänge der Muster wurde 4 gewählt. Auffällig ist, dass bei einem Vergleich der Top 10 Muster die beiden Sequenzen eine geringe Ähnlichkeit aufweisen. Das passiert aufgrund der Nutzereingabe, die zwanghaft dafür sorgt, dass zu Beginn der synthetischen Sequenz nicht hauptsächlich die Muster mit dem höchsten Support im echten Radio gespielt werden. Betrachtet man nun die Top 50 oder Top 100 Muster, wird deutlich, dass die Ähnlichkeit zum echten Radio auch für kürzere Sequenzlängen deutlich höher ist. Angenommen, der Hörer wünscht sich die Gattung „Hörspiel“ zu Beginn der Sequenz. Weiterhin sei angenommen, dass die Muster dieser Gattung und auch die darauffolgenden Muster allesamt einen geringen Support aufweisen. Im Gegensatz dazu, soll die Gattung „Nachrichten“ Muster mit einem hohen Support enthalten. Bei diesen Rahmenbedingungen startet das Radio mit Mustern eines geringen Supportwerts. Erst später folgt der Umschwung zu beliebteren Mustern mit höherem Supportwert. Dieser Umschwung geschieht in jedem Fall, denn unendliche Loops sind in der Methode ausgeschlossen. Das liegt daran, dass der Algorithmus Wahrscheinlichkeiten aktualisiert, nachdem ein Muster für die Sequenzaufnahme verwendet wurde. Wahrscheinlichkeiten können hierbei auch einen Wert von 0 annehmen. Eine weitere Auffälligkeit ist, dass die Ähnlichkeit mit der Länge des künstlichen Radioprogramms zunimmt. Das liegt daran, dass mehr Muster aus dem echten Radio verwendet werden, je länger das künstliche Radio ist. Erwähnenswert ist, dass diese Ergebnisse aber auch abhängig von den Modellen (Top 10, Top 50 oder Top 100) sind.

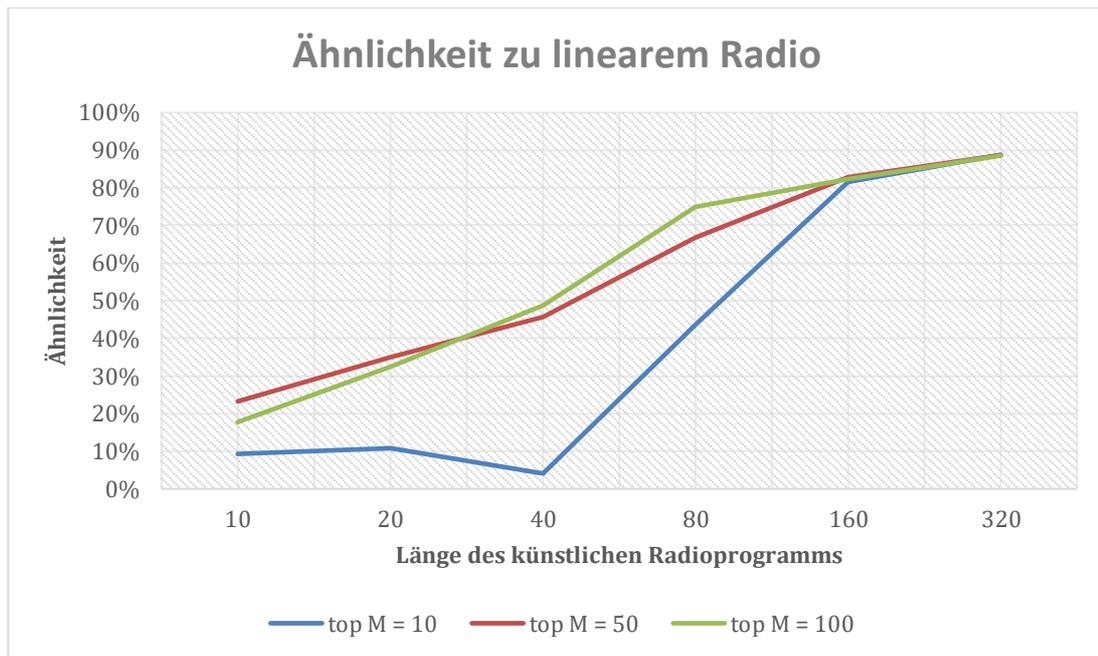


Abbildung 18: Evaluierung der Methode (Minsup echtes Radio: 0,005)

Die Evaluierung erfolgte auch durch Betrachtung längerer Sequenzlängen für die Muster. Genauer gesagt wurden die Längen 5 und 7 betrachtet. Auch hier zeigten sich ähnliche Ergebnisse, wie bei der Sequenzlänge von 4. Ein niedriger Minsup-Wert sorgt hier ebenfalls für eine geringere Ähnlichkeit zum echten Radio. Bei einem höheren Minsup-Wert werden nur die wichtigsten Muster gefiltert und für die Simulation genutzt. Dadurch werden nur die aussagekräftigsten Muster gefiltert und bei der Simulation verwendet, was zu einer höheren Ähnlichkeit zum echten Radio führt. Jedoch ist eine Beobachtung, dass sich bei hohen Sequenzlängen wenig verwendbare und aussagekräftige Muster ergeben. Im Beispiel der Sequenzlänge 5 (siehe Anhang C), existiert für einen höheren Minsup-Wert kein Muster, welches mit der Gattung Nachrichten anfängt und einen hohen Support aufweist. Dennoch wurde eine Evaluierung durchgeführt, um zu untersuchen, ob die bisherigen Ergebnisse auch für längere Sequenzlängen übereinstimmen oder ob sich neue Erkenntnisse ergeben. Dabei erfolgt die Evaluierung im Hinblick auf eine Variation von Nutzereingaben. Auch hier bleiben die Ergebnisse kaum verändert. Der Minsup-Wert beeinflusst nach wie vor den Ähnlichkeitsvergleich. Jedoch ergibt sich noch eine weitere Beobachtung. Die Wahl der Nutzereingabe kann nämlich dafür sorgen, dass zu Beginn Muster mit einem hohen Supportwert gespielt werden. Dann ist die Ähnlichkeit bereits bei kurzen synthetischen Sequenzlängen auf einem hohen Niveau

und sinkt danach für höhere synthetische Sequenzlängen wieder ab. Das liegt daran, dass im späteren Verlauf der synthetischen Sequenz Muster mit einem geringeren Support gewählt werden. Danach steigt die Ähnlichkeit wieder an. Das liegt daran, dass lange synthetische Sequenzlängen dafür sorgen, dass viele Muster des echten Radios genutzt werden. Diese Erkenntnis wurde bereits bei einer Sequenzlänge von 4 gezeigt und wurde hier noch einmal bestätigt. Weiterhin wurden Muster der Länge 7 bestimmt. Hierbei werden bei einem niedrigen Minsup-Wert hauptsächlich Muster mit einem Supportwert von 1 gefiltert. Bei einem hohen Minsup-Wert resultiert nur ein einziges Muster. Beide Fälle beinhalten daher keine verwendbaren beziehungsweise aussagekräftigen Muster. Deswegen wurde für diese Sequenzlänge auf eine Evaluierung verzichtet. Letztlich ist für die Generierung des Inputs der Simulation also auch die betrachtete Sequenzlänge der Muster ein entscheidendes Kriterium hinsichtlich der Ähnlichkeit zum echten Radio. Die verwendeten Muster sollten die entsprechenden Nutzereingaben beinhalten und aussagekräftig sein. Das bedeutet, dass die Muster, oder ein Teil davon, höhere Supportwerte enthalten sollten. Weiterhin wurde die Evaluierung auch für die anderen Oberkategorien Bezugsreichweite und Klassen durchgeführt (siehe Anhang D). Die Ergebnisse sind analog zu den der Gattungen. Je höher der Minsup-Wert gewählt ist, desto schneller erreicht man eine hohe Similarität. Bei einem zu hohen Wert existieren überhaupt keine Muster. Bei einem zu niedrigen Wert existieren zu viele. Ebenso verhält es sich mit den Ergebnissen der Generierung durch Klassen.

Bereits in der Sequenzmusteranalyse wurde deutlich, dass der Minsup ein kritischer Wert ist. Diese Beobachtung trifft auch die Simulation der Abspielreihenfolgen. Durch die Variation des Minsup bei der echten Radiosendung ergeben sich unterschiedlich viele Muster. Verändert sich also der Minsup-Wert des echten Radios, dann verändert sich auch der Input der Simulation. Abbildung 19 zeigt noch einmal den Ähnlichkeitsvergleich, dieses Mal allerdings mit einem größeren Minsup-Wert. Es wird deutlich, dass bereits kürzere Sequenzlängen des synthetischen Radios eine hohe Ähnlichkeit zum echten Radio aufweisen. Allgemein befindet sich die Ähnlichkeit auf einem hohen Niveau, unabhängig davon, welche Top M Muster betrachtet werden. Demnach sollte bei der Sequenzmusteranalyse der Minsup-Wert Beachtung finden. Das stellt sicher, dass die Simulation, insbesondere für kurze Sequenzlängen, eine hohe Ähnlichkeit erzielt.

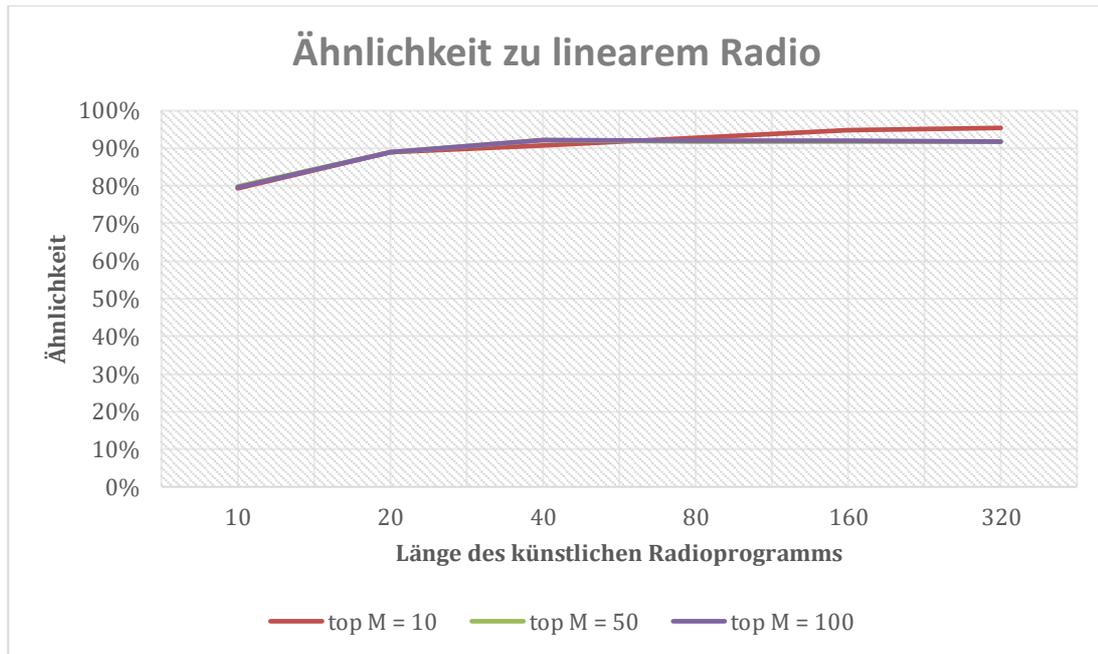


Abbildung 19: Evaluierung der Methode (Minsup echtes Radio: 0,007)

Zusammenfassend lässt sich also sagen, dass Methoden für die Simulation synthetischer Abspielreihenfolgen darauf achten sollten, vorab eine ausführliche und durchdachte Sequenzmusteranalyse durchzuführen. Diese ist nämlich der Input für den Algorithmus. Es sollten tatsächliche, aussagekräftige Muster identifiziert werden. Sehr hohe Musterlängen als Input der Simulation sollten mit Vorsicht behandelt werden. Der betrachtete Kontext sollte nicht zu klein gewählt werden, es sei denn es stehen mehr Daten und dementsprechend mehr Muster zur Verfügung. Für kleinere Kontexte können Muster über mehrere Tage gesammelt und anschließend kumuliert werden. Weiterhin sollte möglichst wenig dem Zufall überlassen werden. Die Sequenz sollte soweit es geht aus Mustern des echten Radios bestehen. Zuletzt sollte das Zeitfenster groß gewählt werden. Falls dennoch keine Muster vorhanden sind, ist eine stetige Reduzierung des Zeitfensters sinnvoll, bis wieder welche gefunden werden.

6 Diskussion und Handlungsempfehlungen

Ziel der Arbeit war es, Muster und kleine Unterschiede im linearen Radio zu erkennen. Dieses Ziel wurde erfüllt, indem eine umfassende qualitative Inhaltsanalyse durchgeführt wurde, um die Häufigkeiten einzelner Kategorien zu bestimmen. Dabei erfolgte auch eine Unterscheidung zwischen verschiedenen

Tageszeiten. Es wurde festgestellt, dass einige Kategorien sendertypisch sind oder nur zu bestimmten Tageszeiten vorkommen. Weiterhin wurde eine Sequenzmusteranalyse für den gesamten Datensatz durchgeführt. Analog erfolgte auch ein Vergleich der Muster für einzelne Kontexte. Die Arbeit hat somit gezeigt, dass durch die verwendeten Werkzeuge Muster in der Datenquelle identifiziert werden können. Ein weiteres Ziel der Arbeit war das Ausarbeiten einer Methodik für die Simulation synthetischer Ausspielreihenfolgen. Dieses Vorhaben wurde ebenfalls erreicht. Die Arbeit hat gezeigt, dass einige Faktoren eine Rolle spielen, um eine hohe Ähnlichkeit der künstlichen und echten Sequenzen zu erreichen. Beispielsweise dürfen Sequenzelemente nicht zufällig ausgewählt werden, weswegen mit Wahrscheinlichkeiten gearbeitet werden sollte. Weiterhin verändert die Wahl der Parameter der Sequenzmusteranalyse, den Input für die Simulation und beeinflusst somit den Ähnlichkeitsvergleich.

Aus den Ergebnissen dieser Arbeit lassen sich einige **Handlungsempfehlungen** ableiten. Die qualitative Inhaltsanalyse sowie die Sequenzmusteranalyse liefern interessante Einblicke in das Medium Radio. Dabei wurde ersichtlich, wie das Radio aufgebaut ist. Vor allem wurde deutlich, welche Themen besonders häufig vorkommen und wie diese aufeinander folgen. Zukünftige Forschungsarbeiten oder Projekte können auf diese Vorgehensweise zurückgreifen, insbesondere mit dem Ziel, synthetische Sequenzen zu erzeugen. Dafür ist eine intensive Beschäftigung mit der Mustererkennung hilfreich. Diese Arbeit bietet hierfür eine entsprechende Grundlage.

Die, in dieser Arbeit vorgestellte Methode zur Erzeugung synthetischer Sequenzen, weist eine hohe Ähnlichkeit zum echten Radio auf. Die Anwendung der Methode, sowie die Evaluierung, erfolgte exemplarisch auf die Oberkategorie Gattungen und die Sequenzlänge von 4. Die Anwendung erfolgt für andere Kontexte analog. Um nicht den Rahmen der Arbeit zu überschreiten und die wesentlichen Faktoren intensiv zu erarbeiten, lag der Fokus auf einem Kontext. Dabei zeigte die Arbeit, dass für eine Simulation gewisse Faktoren zu beachten sind, um eine hohe Güte der synthetischen Sequenzen zu erreichen. Dabei spielt vor allem der Input und die damit einhergehende Wahl des Minsup eine Rolle. Weiterhin wurden bereits in den ersten Kapiteln der Arbeit Anforderungen an die Simulation definiert. Eine Nutzereingabe sowie eine Sequenzmusteranalyse stellen hierbei zwei essenziellen Anforderungen dar. Diese stellen sicher, dass ein Grad der Personalisierung beim

nichtlinearen Radio vorhanden ist, bei gleichzeitiger Erhaltung des redaktionellen Mehrwerts vom linearen Radio. Die Arbeit bietet ein Grundgerüst für zukünftige Projekte. Das Übernehmen der gesamten Methodik oder einem Teil davon ist möglich. Die Ergebnisse sind nicht nur auf den Kontext Radio beschränkt. Weitere Themengebiete, wie zum Beispiel das Fernsehen, sind ebenfalls Kandidaten für die Erkennung von Mustern und deren künstlicher Erzeugung.

7 Ausblick und Fazit

Das personalisierte Radio befindet sich noch im Anfangsstadium seiner Entwicklung. Arbeiten wie diese bieten Einblicke in die Methoden, sowie auffällige Ergebnisse, die bei der Implementierung solcher automatisierten Radiodienste Berücksichtigung finden sollten.

Eine zukünftige Forschungsfrage könnte sich dem **Grad der Personalisierung** widmen. Mit einer einzelnen Nutzereingabe, die den Startpunkt bestimmt, ist es möglich, das nichtlineare Radio ähnlich wie das lineare Radio zu gestalten. Doch was ist, wenn der Hörer einen einseitigen Wunsch äußert und beispielsweise nur Beiträge zu einem bestimmten Thema hören will? Für dieses Szenario existieren gegebenenfalls keine Muster. Die Ähnlichkeit zum linearen Radio sinkt in diesem Fall. Zudem stellt sich die Frage, ob für solche Fälle der Rundfunkvertrag nicht verletzt wird. Weiterhin verhindert ein maximal personalisiertes Radio die Entdeckung neuer Themen durch den Hörer, da ein Großteil der Themen ausgeschlossen wird. Zweifelsohne ist es für jedes Radiounternehmen profitabel, wenn der Hörer das gespielt bekommt, was er möchte und somit sich möglichst lange mit dem Medium beschäftigt. Doch der Grad der Personalisierung ist unabdingbar ein Aspekt, der weitere Aufklärung benötigt.

Ein weiterer Aspekt ist die Frage nach dem **Grad der Automatisierung**. In dieser Arbeit wurde eine Methodik vorgestellt, die es erlaubt, die Sequenzen des nichtlinearen Radios komplett zu automatisieren. Die Nutzereingabe hingegen erfolgt manuell und bestimmt das zu Beginn gespielte Element. Doch ist das überhaupt ein zielführendes Vorgehen? Wieviel Aufwand soll oder darf dem Hörer aufgespielt werden, sodass ihm nicht langweilig wird und er das Medium als anstrengend empfindet? Die Nutzereingabe kann verschiedene Formen annehmen. Beispiele hierfür sind eine manuelle Eingabe per Knopfdruck oder eine

Spracherkennung. Weiterhin sind mögliche Szenarien das Anlegen von Profilen für jeden einzelnen Hörer, bei dem eine Angabe von Interessen erfolgt. Eine automatische Feedback-Erkennung ist ebenfalls möglich durch die Verarbeitung von Skip-Vorgängen. Überspringt der Hörer einen Beitrag, so kann das System automatisch erkennen, welche Beiträge der Hörer ungerne hört. Das Überspringen von Beiträgen ist aus Sicht des Hörers ein deutlich geringerer Aufwand als die Erstellung eines Profils, bei dem manuelle Eingaben erwartet werden. Doch die Verarbeitung von Feedback sollte nicht nur direkt am System erfolgen.

Ohne Frage ist es das Ziel, dass der Hörer sich solange wie möglich mit dem Radio beschäftigt. Gegebenenfalls wäre ein **Hörer-orientierter Ansatz** geeignet, um die grundlegenden Bedürfnisse der Kunden zu erfahren und daraufhin das Radio von morgen zu gestalten. Demnach sollte untersucht werden, welche Faktoren zur Akzeptanz des Radios führen. Sogenannte Akzeptanzmodelle sind hierfür geeignet. Daraus ist es gegebenenfalls möglich, neue Anforderungen an die Sequenzmusteranalyse oder die Simulation abzuleiten. In der Arbeit von Lahabou und Wok (2011) wurde beispielsweise das Campus Radio einer Universität hinsichtlich der Akzeptanz untersucht. Dabei wurde festgestellt, dass das Radio weniger genutzt wird, wenn die Studenten eine schlechte Einstellung gegenüber dem Radio haben. Diese Gegebenheit reicht schon, dass die wahrgenommene Nützlichkeit des Radios leidet und die Studenten das Medium letztendlich kaum nutzen. Solche Erkenntnisse sind hilfreich und können im personalisierten Radio Berücksichtigung finden. Denn es wird mit hoher Wahrscheinlichkeit potenzielle Hörer geben, die erst einmal an das Medium gebunden werden müssen. Falls Hörer zu oft Beiträge überspringen, könnte das Radio diese Aktivitäten erkennen und dementsprechend entgegenwirken und andere Beiträge spielen. Um die Einstellung gegenüber dem personalisierten Radio zu beeinflussen, könnte bei den synthetischen Sequenzen gelegentlich Eigenwerbung geschaltet werden, sowie es andere Audiomedien bereits tun. Spotify ist ein Beispiel hierfür und macht, nachdem einige Songs gespielt wurden, Eigenwerbung, um sich besser als andere Medien darzustellen. Auch im Radio könnte dies Verwendung finden, indem die Vorteile des neuen personalisierten Radios kommuniziert werden. Ebenfalls ist es mithilfe von Akzeptanzmodellen, möglich den Grad der Automatisierung zu untersuchen. Eine manuelle Profilerstellung könnte von Hörern als zu aufwendig empfunden werden, sodass die wahrgenommene Benutzerfreundlichkeit darunter leidet. Dem könnte man

entgegenwirken, indem das Radio so simpel wie möglich gehalten wird. Untersuchungen bezüglich der Akzeptanzmodelle von Musik-Streaming-Diensten haben ergeben, dass die soziale Präsenz des Services das Serviceerlebnis positiv beeinflusst (Wang, Huang, & Tai, 2017). Soziale Präsenz beschreibt im Kontext von digitalen Systemen das Gefühl, mit einer natürlichen Person zu kommunizieren. Auch im Kontext vom Radio könnten solche Untersuchungen hilfreich sein, um das personalisierte Radio zu optimieren, denn solche Erkenntnisse verändern gegebenenfalls die Art und Weise der Nutzer-Anbieter-Kommunikation oder aber auch die Methode selbst. Festzuhalten ist also, dass Faktoren identifiziert werden sollten, die für eine Nutzung des personalisierten Radios sprechen. Diese sollten dann bei der Implementierung von Algorithmen zur Erstellung synthetischer Sequenzen Berücksichtigung finden.

Weiterhin sollte man sich mit der Frage beschäftigen, inwiefern die **Komplexität des linearen Radios** im nichtlinearen Radio berücksichtigt werden soll. Das Erkennen von Mustern und das reine Abspielen dieser sorgt zwar für eine hohe Ähnlichkeit, aber reicht gegebenenfalls nicht aus, um zum Beispiel den Informationsgehalt der Kontexte abzubilden. Beispielsweise werden im linearen Radio morgens kürzere Beiträge, wie Nachrichten oder Berichterstattungen, abgespielt. Die präsentierte Simulation bildet diesen Informationsgehalt zwar durch die Musteranalyse ab, aber möglicherweise nicht ganz zufriedenstellend. Um ein Beispiel zu nennen, sind Nachrichten größtenteils Beiträge mit einer kurzen Spieldauer. Allerdings existieren in Ausnahmefällen auch welche mit einer langen Spieldauer. Damit der Informationsgehalt zu jeder Tageszeit angemessen ist, könnte in Zukunft auch vermehrt die Dauer der Beiträge berücksichtigt werden. Das könnte man beispielsweise durch die Kodierung der Beiträge nach der Beitragsdauer in kurz, mittel und lang erfolgen. Ein weiterer Punkt ist die Berücksichtigung von Stundenuhren. Nachrichten werden im linearen Radio häufig zur vollen Stunde gespielt. In der vorgestellten Methode ist kein Beitrag an einer Stundenuhr gebunden und wird erst dann gespielt, wenn ein Muster dafür gefunden ist. Ob und wie Stundenuhren im personalisierten Radio Berücksichtigung finden sollten, bedarf weiterer Forschung.

Ein weiterer Faktor ist die **Berücksichtigung von Redakteuren und Programmexperten**. Die qualitative Inhaltsanalyse sowie die Sequenzmusteranalyse finden kleine Unterschiede in der Häufigkeit und in den Mustern des linearen

Radios. Jedoch wurde im Gespräch mit dem Lehrstuhl und den Programmexperten deutlich, dass Redakteure die Abfolge der Beiträge nicht immer so feingranular planen, wie es der Algorithmus tut. Daher gibt es in der Musteranalyse auch gewisse Muster, die unbeabsichtigt berücksichtigt werden und gar nicht so wichtig ist, wie man zunächst annimmt. Daher ist für zukünftige Arbeiten auch die Zusammenarbeit mit erfahrenen Redakteuren von Bedeutung. Ein hybrider Ansatz wäre eine Möglichkeit, die beide Blickpunkte vereinbart. Einmal die mathematisch algorithmische Methodik, aber auch die Gedanken der Redakteure. Dadurch könnten weitere Erkenntnisse gewonnen werden, um die Simulation noch weiter zu optimieren und dem Ziel, ein personalisiertes Radio zu implementieren, näher zu kommen.

Darüber hinaus soll auf die **Grenzen dieser Arbeit** hingewiesen werden. Zum einen basieren die Ergebnisse auf einer spezifischen, für diese Arbeit bereitgestellten, Datenquelle. Dabei sind besonders die betrachteten Kontexte oftmals zu klein. Der Kontext bezeichnet hier eine Teilmenge der Datenquelle basierend auf gegebenen Parametern wie Sender, Tag (Wochentag/Wochenende) und Tageszeit. Je mehr Parameter gesetzt werden, desto kleiner sind die Teilmengen und dementsprechend die Muster, welche daraus resultieren. Die Evaluierung des vorgestellten Algorithmus war mit solch kleinen Kontexten nicht möglich, sodass auf einige Parameter verzichtet werden musste. In Zukunft könnte man für jeden dieser Kontexte mehr Daten sammeln, was dann eine intensivere Analyse ermöglicht. Bezüglich der Darstellung der Muster könnte man zusätzliche Visualisierungsmethoden verwenden, die auch eine Betrachtung längerer Sequenzlängen ermöglichen. Die gewählte Visualisierung durch Sankey-Diagramme verschlechtert aber nicht den vorgestellten Ansatz dieser Arbeit. Weiterhin weist die in dieser Arbeit vorgestellte Personalisierung Grenzen auf. Durch die Verwendung von Mustern bei der Erstellung synthetischer Sequenzen, wird die Personalisierung zum einen ermöglicht, zum anderen wird sie aber auch eingeschränkt. Eine Erzeugung synthetischer Sequenzen durch Muster kann beispielsweise dafür sorgen, dass der nächste Beitrag ein Thema beinhaltet, wofür der Hörer sich gar nicht interessiert. Momentan ist lediglich die Nutzereingabe, welches das erste Element der synthetischen Sequenz angibt, ein Aspekt der Personalisierung. Daher soll auf die Erweiterbarkeit der in dieser Arbeit vorgestellten Methode hingewiesen werden. Gegebenenfalls muss der Algorithmus verändert beziehungsweise ergänzt werden,

sobald Empfehlungssysteme oder weitere Nutzereingaben integriert werden. Falls weitere Interessen des Hörers berücksichtigt werden sollen, dann ist dies technisch unkompliziert realisierbar. In der Sequenzmusteranalyse müssen dazu lediglich die Themen beziehungsweise Gattungen angegeben werden, die im Interessenfeld des Hörers sind. Dann werden auch nur Muster gefiltert, die diese Elemente beinhalten und können dementsprechend genutzt werden. Zu beachten ist hier aber, dass es zu Komplikationen kommen kann, wenn keine oder wenige Muster gefunden werden. Eine weitere Personalisierung wird durch das Füllen der Muster mit realen Beiträgen erzielt. Im gesamten Prozess sollte darauf geachtet werden, dass der Rundfunkvertrag nicht verletzt wird. Weiterhin soll auf die Grenzen hingewiesen werden, die bei der Evaluierung des Algorithmus zum Vorschein kamen. Die grobe Betrachtung der Top 10, Top 50 oder Top 100 Muster führt zu unterschiedlichen Ergebnissen bei der Ähnlichkeit der synthetischen Sequenz zum linearen Radio. Die Wahl dieser Zahlen erfolgte aufgrund einer Inspiration durch die vorhandene Literatur. Gegebenenfalls kann hier eine andere Vorgehensweise verschiedene Ergebnisse erzielen. Weiterhin stellt sich die Frage wie valide die verwendete Ähnlichkeits-Metrik ist. Zum einen wurde diese im Vergleich zwischen den jeweiligen Kontexten verwendet, sprich bei der Erstellung der Dreiecksmatrizen. Zum anderen wurde sie beim Vergleich der Muster der Simulation und der Muster des linearen Radios verwendet. Der Begriff Ähnlichkeit kann hierbei subjektiv sein. Für jemand anderen kann das lineare Radio ähnlich zum synthetischen Radio sein, wenn diese eine ähnliche Dauer bei den Beiträgen beinhaltet oder wenn der Sprecher dieselbe Person ist etc. Dabei gibt es unzählige Faktoren, die berücksichtigt werden können. Hier basiert die Ähnlichkeits-Metrik nur auf dem Vergleich zwischen der Häufigkeit identifizierter Muster in zwei unterschiedlichen Mengen. Weiterhin gibt es mathematisch verschiedene Ansätze, eine Ähnlichkeit zu bestimmen. Ein weiterer Gedanke ist, ob die gewählte Kodierung für das Vorhaben valide ist. In der Tat ist die Wahl der Codes abhängig von der Datenquelle. Dabei existieren jedoch viele Möglichkeiten hinsichtlich der Feingranularität. Beispielsweise beinhaltet die Oberkategorie Geographika nur eine geringe Anzahl an Codes. Jedoch ist hier auch eine weitgehendere Kodierung möglich, indem zum Beispiel die Kodierung nach Ländern oder sogar Städten erfolgt. Auch können beispielsweise die Klassen verfeinert werden, sodass die Klasse „Sport“ auch zwischen Handballsport oder Fußballsport differenziert werden kann. Außerdem stellt sich die Frage, wie genau

der Prozess der Kodierung ist. Die Kodierung wurde von einer Person durchgeführt. Bei Durchführung mit mehreren Personen würden gegebenenfalls andere Ergebnisse resultieren. Angenommen, es existiert ein einstündiger Beitrag, der lediglich fünf Minuten über die Klasse „Politik“ handelt. Eine Person würde die Klasse ignorieren und sie nicht dem Beitrag zuordnen, mit der Begründung, dass die relative Dauer zu gering ist. Eine andere Person würde diese Zuordnung allerdings mit aufnehmen. Weiterhin ergaben sich bei der Kodierung Schwierigkeiten, wenn man sich unsicher ist, zu welcher Gattung ein Beitrag gehört. Beispielsweise ist die Unterscheidung zwischen einem Bericht und einem Korrespondentenbericht nicht immer klar gewesen. Der Kodierungsprozess ist also noch teilweise subjektiv und benötigt feste Regeln.

Zusammenfassend bietet diese Arbeit ein gutes Fundament und beleuchtet eine andere Sichtweise der Sequenzmusteranalyse, nämlich durch Betrachtung einer Beitragsebene. Die Arbeit hat gezeigt, dass eine qualitative Inhaltsanalyse und eine Sequenzmusteranalyse sinnvolle Werkzeuge sind, um Sequenzen für das personalisierte Radio zu erzeugen. Das allein reicht allerdings nicht aus, da noch weitere Aspekte untersucht werden müssen, um das personalisierte Radio implementierungstauglich zu machen. Dabei kann eine Absprache mit erfahrenen Redakteuren bezüglich der Bedürfnisse der Hörer sowie der Gestaltung der künstlich erzeugten Sequenzen hilfreich sein. Ein weiterer Forschungsvorschlag ist die Untersuchung der Faktoren, die zur Akzeptanz eines personalisierten Radios führen. Auch wenn sich das personalisierte Radio noch im Anfangsstadium seiner Entwicklung befindet, leistet diese Arbeit einen Beitrag zur Digitalisierung dieses altbekannten Systems. Es zeigt eine Möglichkeit, die Vorteile des linearen Radios ins nichtlineare, personalisierte Radio zu übertragen.

Literaturverzeichnis

- Abreu, J., Nogueira, J., Becker, V., & Cardoso, B. (2017). Survey of Catch-up TV and other time-shift services: a comprehensive analysis and taxonomy of linear and nonlinear television. *Telecommunication Systems*, 64(1), 57–74.
<https://doi.org/10.1007/s11235-016-0157-3>
- Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the Eleventh International Conference on Data Engineering*.
- Borgers, E. W., & Koenig, A. E. (1960). The structure and content of music continuity. *Journal of Broadcasting*, 4(3), 269–279.
<https://doi.org/10.1080/08838156009385926>
- Crane, E., Talbott, A., & Hume, R. (1961). Time use profiles and program strategy. *Journal of Broadcasting*, 5(4), 335–343.
<https://doi.org/10.1080/08838156109385981>
- Deutschlandfunk Nova (2021). Programmschema. Retrieved from
<https://www.deutschlandfunknova.de/programmschema>
- Fournier-Viger, P., Lin, J. C.-W., Kiran, R. U., Koh, Y. S., & Thomas, R. (2017). A survey of sequential pattern mining. *Data Science and Pattern Recognition*, 1(1), 54–77.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques: Concepts and Techniques* (3. Aufl.). s.l.: Elsevier professional. Retrieved from <http://lib.myilibrary.com/detail.asp?id=317117>
- Hayes, C., & Cunningham, P. (1999). *Smart Radio-a proposal*.
- Heinrich, J. (2010). *Mediensystem, Zeitung, Zeitschrift, Anzeigenblatt* (3. Aufl.[unveränd. Nachdr.]). *Medienökonomie: / Jürgen Heinrich ; 1*. Wiesbaden: Westdeutscher Verlag.
- Hirschmeier, S., & Beule, V. (2018). *Compliance of personalized radio with public-service remits*. Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization.
- Hirschmeier, S., Beule, V., & Tilly, R. (2020). *Translating Editorial Work into Algorithms for Personalized Radio Streams*.
<https://doi.org/10.1080/19376529.2019.1704760>

- Hirschmeier, S., Tilly, R., & Beule, V. (2019). *Digital Transformation of Radio Broadcasting: An Exploratory Analysis of Challenges and Solutions for New Digital Radio Services*. <https://doi.org/10.24251/HICSS.2019.602>
- Jauert, P., Ala-Fossi, M., Föllmer, G., Lax, S., & Murphy, K. (2017). The Future of Radio Revisited: Expert Perspectives and Future Scenarios for Radio Media in 2025. *Journal of Radio & Audio Media*, 24(1), 7–27. <https://doi.org/10.1080/19376529.2017.1310574>
- La Roche, W. von, & Buchholz, A. (2016). *Radio-Journalismus: ein Handbuch für Ausbildung und Praxis im Hörfunk*: Springer-Verlag.
- Lahabou, M., & Wok, S. (2011). *Students' Acceptance of an Internet Campus Radio: A Test of the Technology Acceptance Model*. Retrieved from [http://irep.iium.edu.my/15076/1/iium.fm\[1\]_-_mention_mahaman.pdf](http://irep.iium.edu.my/15076/1/iium.fm[1]_-_mention_mahaman.pdf)
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarbeitete Auflage). Weinheim, Basel: Beltz Verlag. Retrieved from <http://dnb.info/1063369835/04>
- Ministerium des Innern des Landes Nordrhein-Westfalen, Referat 14 (2020, October 28). Historische Gesetze | Landesrecht NRW. Retrieved from https://recht.nrw.de/lmi/owa/br_bes_text?anw_nr=2&gld_nr=2&ugl_nr=2251&bes_id=5506&aufgehoben=J&menu=1&sg=0
- Plantevit, M., Choong, Y. W., Laurent, A., Laurent, D., & Teisseire, M. (2005). *M 2 SP: Mining sequential patterns among several dimensions*.
- Pöchhacker, N., Burkhardt, M., Geipel, A., & Passoth, J.-H. (2017). *Interventionen in die Produktion algorithmischer Öffentlichkeiten: Recommender Systeme als Herausforderung für öffentlich-rechtliche Sendeanstalten*. DEU.
- Rai, A., & Sambamurthy, V. (2006). Editorial Notes—The Growth of Interest in Services Management: Opportunities for Information Systems Scholars. *Information Systems Research*, 17(4), 327–331. <https://doi.org/10.1287/isre.1060.0108>
- Sappelli, M., Chu, D. M., Cambel, B., Nortier, J., & Graus, D. (2018). SMART radio: personalized news radio. *Anne Dirkson, Suzan Verberne, Gerard Van Oortmerssen & Wessel Kraaij*, 27.

- Schoeffler, M., Edler, B., & Herre, J. (2013). *How much does audio quality influence ratings of overall listening experience*. Proc. of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR).
- Statista (2020, October 28). Musik-Streaming - Deutschland | Statista Marktprognose. Retrieved from <https://de.statista.com/outlook/209/137/musik-streaming/deutschland>
- Turnbull, D. R., Zupnick, J. A., Stensland, K. B., Horwitz, A. R., Wolf, A. J., Spigel, A. E., . . . Joachims, T. (2014). Using personalized radio to enhance local music discovery. *CHI'14 Extended Abstracts on Human 2014*, 2023–2028.
- Vargo, S., & Lusch, R. (2004). Evolving to a New Dominant Logic. *Journal of Marketing*, 68, 1–17. <https://doi.org/10.1509/jmkg.68.1.1.24036>
- Wang, K., Huang, S.-T., & Tai, C.-F. (Eds.) (2017). *Role of service experience and value co-creation in determining music streaming service continuance*: Association For Information Systems.

A. Anhang

Gattung	Erklärung
Collage Hörspielmusik	Mischung verschiedener Stilelemente, oft in Verbindung mit elektronischen Hilfsmitteln/ Geräuschen. Wort: Seit den 20er Jahren bis heute verwendete künstlerische Form mit in der Regel erkennbarem Thema und unverdeckten Schnittstellen. Die Collage verknüpft verschiedenartige, vorgefundene Gestaltungselemente ohne bzw. mit wenig erklärendem oder verbindendem Text. Die Materialgrundlage bildet (Musik, die im Hörspiel zum Einsatz kommt)
Musiktheater	darunter: alle Formen der Bühnenmusik wie Oper, Operette, Musical, Ballett, Schauspielmusik, Melodram
Werbespot	Klassische Werbeform im Rundfunk. Übliche Spotlängen sind 7, 10, 15, 20, 30, 45 und 60 Sekunden.
Hörspielbearbeitung	Hörspiel, dessen Manuskript auf einer (meist literarischen) Vorlage beruht. (Gegenteil von Originalhörspiel)
Kinderhörspiel	Formal nicht näher bestimmtes Hörspiel für die Zielgruppe Kind
Kurzhörspiel	Hörspiel bis zu einer Länge von ca. 30 Minuten
Mitmachhörspiel	Hörspiel mit Hörerbeteiligung
Mundarthörspiel	Hörspiel im Dialekt
Musikalisches Hörspiel	Hörspiel, in dem die Musik eine konstruktive Rolle spielt. Paradigma einer sinnvollen Synthese von Literatur und Musik: Brechts ‚Lindberghflug‘
Originalhörspiel	Hörspiel, das auf einem eigens für den Hörfunk verfassten Manuskript beruht. (Gegenteil von Hörspielbearbeitung)
Science Fiction-Hörspiel	Hörspiel, das den Gesetzen des Genres Science Fiction folgt (Schilderungen von Abenteuern und Erlebnissen in einer auf naturwissenschaftlich-technischer Grundlage phantasievoll ausgemalten Zukunft)

Interview	<p>Im Gegensatz zum eher dialogischen Gespräch handelt es sich bei einem Interview um die Befragung einer Person, die der Ermittlung von Informationen und/oder Meinungen zu einem bestimmten Themenbereich dient. Charakteristisch sind kurze, präzise Fragen; die Antworten können ebenfalls kurz, aber auch länger und ausführlicher sein. Das Interview kann einen eigenen Sendeplatz einnehmen.</p>
Umfrage	<p>Eine spezielle Form des Interviews, bei der eine Frage an mehrere Personen gestellt wird, um ein Meinungsbild zu erhalten, z.B. eine Passantenbefragung oder Hörerumfrage etc.</p>
Büttenrede	<p>Bestandteil einer karnevalistischen Großveranstaltung. Publikumsanrede (Monolog) von einer Bühne unter direkter Einbeziehung der Zuschauer. Menschliche Schwächen und politisch-gesellschaftliche Verhältnisse werden in humoristischer Form der Lächerlichkeit preisgegeben.</p>
Kabarett	<p>Kleinkunst, die kleine Formen der darstellenden Kunst (Szene, Monolog, Dialog, Pantomime, Sketch), der Literatur (Gedicht, Ballade) und der Musik (Lied, Chanson, Couplet) verknüpft und (auf einer Bühne) als humoristisch-satirisches oder politisch-gesellschaftskritisches Programm zur Aufführung bringt.</p>
Sketch	<p>Kleine Form der Unterhaltung. Existiert eigenständig oder als Teil eines Nummernprogramms. Meist dramatische Kurzszene (ca. 3-5 Minuten), ironisch-witziger Inhalt; auch satirisches Bühnenwerk mit mehreren Szenen und durchgängiger Handlung in Art des Einakters. Auch Einpersonennummer.</p>
Witz	<p>Wortkomik, kleinste Darstellungsform, vornehmlich erzählend. Meist nur aus wenigen Sätzen bestehend, in deren Pointe die durch das Erzählen gesteigerte Spannung in einer völlig unerwarteten, logisch nicht vorhersehbaren Richtung gelöst wird.</p>
Kommentar Autorenlesung	<p>S,B / Der Kommentar ist die Ausdrucksform für Meinung und Interpretation. Er wird in der Regel vom Autor selbst gesprochen. vgl. im Unterschied dazu Glosse Vom Autor selbst vorgetragener Text</p>

Szenische Lesung	Lesung mit verteilten Rollen
Magazin	S / Sendung mit mehreren Beiträgen, die von einem Moderator präsentiert werden. Die Beiträge können sich auf ein gemeinsames Thema beziehen oder unterschiedliche Themen haben. Wird ein Ausschnitt eines Magazins dokumentiert, so kann die Bezeichnung Magazin zusätzlich vergeben werden.
Montage	Rein technische Aneinanderreihung von akustischen Materialien (O-Ton oder Geräusch oder Musik oder Text). Vgl. im Unterschied dazu „Collage
Nachrichtensendung	Aktuelle Information in knapper, prägnanter Form
Meldung	Aktuelle Meldung außerhalb der vorgesehenen Nachrichtenzeit, z.B. Katastrophenmeldung, Warnhinweis etc.
Nachrichten	Sendung mit tagesaktueller, in der Regel multithematischer Information, die in regelmäßigen Zeitabständen auf festen Sendeplätzen ausgestrahlt wird. Nachrichten können auch O-Ton-Einspielungen und eigenständige Kurzberichte enthalten.
Verkehr	Aktueller Verkehrsstand, Staumeldungen, Warnung vor Verkehrskontrollen etc.
Wetter	Wettervorhersage, Wetterbericht
Öffentliche Veranstaltung	Live-Übertragung oder kompletter Mitschnitt bzw. Teil eines Mitschnittes einer Darbietung vor Publikum, welches nicht primär Rundfunkpublikum ist. Das Attribut „öffentlich“ soll hier auch auf Veranstaltungen angewendet werden, die nicht unbedingt der gesamten Öffentlichkeit zugänglich sind, wie z.B. Staatsempfänge, Eröffnungsveranstaltungen etc. Im Falle von künstlerischen Darbietungen i
Festakt	Hierunter fallen beispielsweise Preisverleihungen, Einweihungen, Eröffnungen, Empfänge, Gedenkstunden sowie Feierstunden, Feiern, Festivals, Staatsakte etc. Die spezifische Bezeichnung wird im Titel angesetzt.

	Von Organisationen, Unternehmen, Interessengruppen oder einzelnen Personen (meist aus Politik, Wirtschaft, Kultur, Sport) organisierte Informationsveranstaltung mit dem Ziel, eingeladene Journalisten über Vorgänge zu unterrichten. Diese Mitteilungen werden häufig durch mögliche Fragen und Antworten ergänzt. Wird aus solchen Mitteilungen oder Stellungnahmen eine Sequenz isoliert, so handelt es sich um eine liturgische Handlung, darunter auch der Gottesdienst. Vgl. im Unterschied dazu „Betrachtung“, die das reflektierende Wort in den Vordergrund stellt.
Pressekonferenz	
Religiöse Veranstaltung	Radiophone Darstellung einer Person, einer Körperschaft, einer Stadt, eines Landes oder einer Landschaft in ihren wesentlichen Eigenschaften. Zur hinreichenden Formalkategorisierung eines Porträts bedarf es in der Regel weiterer Präsentationsformen (vgl. Feature, Bericht, Lesung, Gespräch)
Porträt	Lesung von Presseartikeln/-kommentaren aus aktuellen Zeitungen, oft auch mit knappen Sprechertexten eingeleitet oder verbunden
Presseschau	Hat den Zweck, die Hörer/Zuschauer auf Programminhalte aufmerksam zu machen. Über einen Hinweis hinausgehende Programmankündigung
Programmelement	Schlußmoderation nach Ende eines Beitrags oder einer Sendung
Abmoderation	Absage einer soeben beendeten konkreten Sendung
Absage	Thematische Einleitung einer Sendung oder eines Einzelbeitrags innerhalb einer Sendung durch Hintergrundinformationen oder kommentierende Information
Anmoderation	Ansage einer konkreten Sendung, die sofort folgt
Ansage	Ansage innerhalb einer konkreten Sendung
Zwischenansage	Vor- und Nachspann, Indikativ und Abdikativ, (Er)Kennung einer Sendung
Jingle	Ansage eines Senders bzw. einer Welle, evtl. mit Nennung der Frequenz
Stationsansage	Kurzer Aufreißer, Anriss eines Themas oder einer Sendung
Teaser	

Trailer	<p>S,B / Hinweise auf andere Sendungen. Ankündigungsspot für bestimmte Programmformate oder einen Sender. Wichtiger Bestandteil der On-Air-Promotion. Dienstleistungsorientierte Sendung/Produktion mit dem Ziel, dem Hörer Hilfestellungen, Ratschläge und Informationen für verschiedene Lebensbereiche (Gesundheit, Recht, Soziales u.a.) zu vermitteln</p>
Ratgebersendung	<p>Frage-Antwort-Spiel zwischen einem Quizmaster und einem oder mehreren Kandidaten (meist Hörern), das nach bestimmten Regeln abläuft (z.B. die Beantwortung der Fragen innerhalb einer festgelegten Zeit). Die Quizfragen sind in der Regel reine Wissensfragen.</p>
Quiz	<p>Mündliche Darlegung von Gedanken vor einem Publikum über ein oder mehrere Themen. Ziel ist es, neben der sachlichen auch eine emotionale Wirkung zu erzeugen. Im Vergleich zum Vortrag gewinnt hier die rhetorische Gestaltung, z.B. durch Variation der Stimmlage und der Dynamik eine größere Bedeutung.</p>
Rede	<p>Kurze Rede zu besonderen Anlässen, z.B. zu Fest- und Feierlichkeiten, Geburtstagen, Jubiläen, Trauerfeiern und zu besonderen politischen Ereignissen</p>
Ansprache	<p>Im Rahmen eines Festaktes gehaltene feierliche Rede, in welcher sich der Redner für eine Auszeichnung bzw. Würdigung bedankt.</p>
Dankesrede	<p>Bezeichnung für die während des Gottesdienstes/Messe gehaltene Kanzelrede, auch für den in der Katholischen Kirche selbständigen Predigt-Gottesdienst (z.B. Fasten-Predigt)</p>
Predigt	<p>Aus unmittelbarem Erleben hervorgehende und die Atmosphäre einbeziehende Berichterstattung. Der Reporter schildert vor Ort, was er sieht und erfährt. Er ist häufig im Bild zu sehen, spricht mit Akteuren, sucht Schauplätze auf etc. vorproduziert/nicht live/geschnitten.</p>
Reportage	

Lesung	S / Rezitation eigener oder fremder literarischer Texte (wie Dichterlesung, Lesebühne). Vom Autor selbst oder von einem Sprecher/Schauspieler gesprochener literarischer Text. Unter einem "literarischen Text" wird hier nicht nur die belletristische, sondern auch die Sachliteratur verstanden.
Rezension	Beschreibende und/oder wertende Betrachtung öffentlicher, künstlerischer Darbietungen (Theater, Konzert, Ausstellung, Film) und Medienpublikationen jeder Art (Presse, Buch, CD, DVD, Video, MC etc.). Sendeform, bei der die aktive Beteiligung von Hörern wesentlicher inhaltlicher Bestandteil der Sendung ist. Diese Beteiligung findet statt im Studio, „vor Ort“ oder durch telefonische Mitwirkung. Die bloße Anwesenheit von Publikum macht noch keine „Sendung mit Hörerbeteiligung“ aus; dies wäre eine „Öffentliche Veranstaltung“ (siehe dort).
Sendung mit Hörerbeteiligung	S / Kurzpräsentation eines Objekts oder Sachverhalts. Klassische Werbeform im Rundfunk. Übliche Spotlängen sind 7, 10, 15, 20, 30, 45 und 60 Sekunden
Spot Aufklärungs- und Gesundheitsspot	(Die kurze Aufklärung über einen bestimmten gesundheitlichen Sachverhalt)
Wahlspot	(Ein kurzer Hinweis beziehungsweise ein Aufruf zum Thema Wählen)
Sprachkurs	Systematische Anleitung zum Erlernen und Trainieren einer Fremdsprache mit radiophonen Mitteln
Statement	Ein Statement ist in der Regel eine von einem O-Ton-Geber frei formulierte oder verlesene Aussage, die als reine Stellungnahme entstanden sein kann (z.B. während einer Pressekonferenz) oder aus einem Interview oder Gespräch isoliert wurde.
Tipp	(Positive) Empfehlung des Journalisten/Reporters/Moderators für die Hörer wie zum Beispiel Buchtipp, Kinotipp, CD-Tipp (auch Hörbuch), TV-Tipp, Veranstaltungstipp, Verbrauchertipp, Ausflugstipp u.a.

Vortrag	S,B / Frontalvermittlung von Themen und Sachinhalten durch Personen. Der Vortrag wird vom Autor selbst über ein oder mehrere Fachthemen zum Zweck der Wissens- und Informationsvermittlung gehalten. Ihn kennzeichnen eine sachliche Gedankenführung und Vortragsweise. vgl. im Unterschied dazu Rede
Werkeinführung Musik	Sprachliche Beschreibung eines musikalischen Werkes, die sich auf geschichtliche, formale, inhaltliche und interpretatorische Aspekte desselben beziehen kann und eher vortragenden Charakter hat. Die Werkeinführung ist vorrangig textuell, kann aber auch musikalische Beispiele/Zitate enthalten, die auch live aufgeführt werden können. Sendung, bei der Wort und Musik in engem inhaltlichen Bezug stehen, z.B. Komponisten- und Musikerporträts, Interpretationsvergleich, Schallplattenprisma u.ä. Es kann sich dabei auch um eine Vorstellung der Entwicklung neuer Musikstile/-Richtungen und deren Formensprache handeln, wie auch um Produktionen, in denen Musik und Poesie eine wechselseitige Wirkung entfalten
Wort-Musik-Sendung	(gelesene Lyrik im W
Monolog	(Ein Selbstgespräch beziehungsweise eine Rede, die keinen direkten Zuhörern gewidmet ist)
Ars acustica	Bezeichnung für ein Werk, in dem die klassischen Abgrenzungen von Musik, Literatur und bildender Kunst aufgehoben sind. Klang, Wort, Geräusch, Gesang, Monolog, Komposition und Gespräch werden zu einem zeitgenössischen Radiokunstwerk verwoben. Die Ars acustica entwickelte sich in den späten 70er/Anfang 80er Jahre aus dem „Neuen Hörspiel“. Im Gegensatz zum traditionellem Hörspiel („Spiel z

	<p>Darlegung von Gelesenem, Gehörtem, Erlebtem oder Gesehenem. Der Bericht berücksichtigt Zusammenhänge, Vorgeschichte und andere wichtige Aspekte eines Ereignisses. Er kann O-Ton-Einblendungen enthalten und wird meist vom</p> <p>Berichterstatter/Korrespondenten/Reporter selbst gesprochen. vgl. im Unterschied dazu Feature und Reportage</p>
Bericht	
Korrespondentenbericht	<p>Bericht eines (Auslands)Korrespondenten aus seinem Zuständigkeitsgebiet</p> <p>Feature, das auf heterogenen Texten basiert, die im Sinne eines Themas kompiliert werden.</p>
Textmontage	
Filmton	<p>"Szenischer" Ausschnitt aus einem Film mit Original-Klängen und Dialogen; eigenständiges Klangbild, das sich aus einer Komposition von Tönen, insbesondere Klängen und Geräuschen zusammensetzt und so eine eigenständige Ebene des Films schafft.</p>
Gespräch	<p>Das Gespräch ist thematisch breit und umfassend angelegt. Beim Gespräch entwickeln sich die Fragen oft aus den Antworten, d.h. es spielt sich nicht nur ein Frage- und Antwortspiel ab. Das Gespräch ist in Abgrenzung zum Interview eher dialogisch. Die Gesprächsform wird häufig bei längeren Befragungen über Biographie und Tätigkeit von Personen verwendet, z.B. bei Lebenserinnerungen.</p>
Glosse	<p>Spöttische Randbemerkung. Pointierter, persiflierender Kommentar zu politischen und gesellschaftlichen Themen und Ereignissen. Wichtige Stilmittel der Glosse sind Ironie, Satire und Schlagfertigkeit, um Schwächen und Missstände bloß zu stellen. Unterscheidet sich vom Kommentar durch ihren zugespitzten, polemischen Stil.</p>
Hörbild	<p>Szenische Rekonstruktion eines Ereignisses (historisch, zeitgenössisch) oder Inszenesetzung eines Ortes (z.B. einer Landschaft, einer Stadt etc.) oder einer Situation, die Authentizität vermitteln soll. Der Situation kann auch eine vermutete Realität zugrunde liegen. Vgl. im Unterschied dazu „Hörspiel“ und „Feature</p>

Hörspiel	<p>Radiophone Kunstform, die originale Hörspieltex te oder auch literarische Tex te nach funkdramaturgischen Gesichtspunkten umsetzt.</p> <p>Hörspiel, das unter Verwendung von dokumentarischen Material, auch O-Ton, ein/e tatsächliche/s historisches Ereignis/Person künstlerisch-ästhetisch, meistens in schauspielerischer Inszenierung gestaltet. Dies geschieht (durchaus) auch mit eindeutiger Tendenz und ist in der Regel unter Verzicht auf einen kommentierenden Sprecher gestaltet.</p>
Dokumentarhörspiel	<p>Bericht (eines Gerichtsreporters) über einen juristischen Prozess</p>
Prozessbericht	<p>Journalistisch verfasster Bericht über eine Reise. Die Lesung eines literarischen Reiseberichts ist hier nicht gemeint.</p>
Reisebericht	<p>Reflektierende Erörterung eines Gedankens aus der Sicht des Autors zu religiösen, philosophischen oder sonstigen Themen. Die Betrachtung soll zum Innehalten, Nachdenken oder Handeln anregen und wird in der Regel vom Autor selbst gesprochen. Musikelemente und/oder O-Töne können als Stilmittel darin enthalten sein.</p>
Betrachtung	<p>Die Erörterung (meist im Parlament oder Parteitag) als öffentliche und in der Form geregelte Aussprache in Rede und Gegenrede dient der Darlegung, Klärung und ggf. Korrektur von Standpunkten</p>
Debatte	<p>Meinungsaustausch zu einem vorher bestimmten Thema unter Leitung eines Diskussionsleiters. Diskussionssendungen haben vorrangig das Ziel, unterschiedliche, kontroverse Standpunkte zu einem Thema herauszuarbeiten und dem Zuhörer die Möglichkeit zur Bildung einer eigenen Meinung zu geben. S,B / Erörterung von Sachthemen zwischen Experten aus unterschiedlichen Perspektiven mit oder ohne Diskussionsleiter</p>
Diskussion	<p>Künstlerischer, expressiver Vortrag eines literarischen Werkes</p>
Rezitation	<p>(Ein musikalischer Auftritt vor einem Publikum)</p>
Konzertaufführung	<p>(Ein Gespräch zwischen zwei oder mehreren Personen)</p>
Dialog	

Sendespiel	<p>Als Sendespiel galt in der Weimarer Republik das für den Rundfunk eingerichtete Bühnendrama klassischer oder zeitgenössischer Literatur. In der Regel wurden die Stücke auf rundfunkeigenen Sendebühnen live vor dem Mikrofon inszeniert unter Anwendung von Methoden der Geräuschvermittlung (Instrumente/Gesang/Chor – „akustische Kulisse“) und/oder der Interpretation des Bühnengeschehens durch Ansager oder Erzähler. Die mehr oder weniger deutliche funkdramaturgische Umarbeitung der vorhandenen Bühnenliteratur im Sendespiel unterscheidet dieses Genre von der reinen Rezitation von Texten durch einen oder mehrere Sprecher, wie sie in den Anfängen des Rundfunks noch üblich war. Vom Sendespiel wurde mit der Zeit das Hörspiel (obwohl zunächst oft synonym verwendet) als originäre künstlerischer Schöpfung für den Rundfunk ohne Rückgriff auf vorhandene dramatische oder epische Vorlagen abgegrenzt.</p>
Bericht in Interviewform	<p>Der Berichterstatter/Korrespondent/Reporter berichtet in Interviewform über ein Thema. Der Unterschied zum Interview besteht darin, daß im Bericht in Interviewform beide beteiligten Personen in ihrer Funktion als Rundfunkmitarbeiter, nicht als O-Ton-Geber, auftreten. Vgl. im Unterschied dazu „Interview“</p>
Sportbericht	<p>(Siehe Gattung "Bericht", ausschließlich über das Thema Sport)</p>
Dokumentarbericht	<p>Zusammenhängender, längerer Bericht bzw. Schilderung von Geschehnissen und/oder Umständen (Dokumentarfilm, Porträt, Video-Protokoll, Langzeitbeobachtung, O-Ton-Dokumentation ...)</p>
Programmkennungen	<p>S,B / Für das Sendegeschehen produzierte Programmteile mit Wiedererkennungsfunktion (Design, Senderkennungen, Vorspanne, Nachspanne)</p>

	Gattung der darstellenden Kunst. Vereint auch in Mischform Formen des Theaters (Szene, Monolog, Dialog) wie der Literatur (Lyrik, Prosa) und der Musik (Lied, Chanson, Tanz, Couplet). Sie bedient sich eigener Mittel (Irreführung, Auslassung, Sprachspielereien, Satire) und Methoden (Travestie, Artistik, Parodie, Karikatur).
Kleinkunst	(Theater mit ausschließlich gesprochenen Stücken)
Sprechtheater	(Szenische Darstellung eines Ereignisses, Kommunikation zwischen den Darsteller und Publikum)
Theater	Humorvoller Sprechgesang mit Musik, oft im Dialekt,
Humoristischer Vortrag	Koranrezitation
Tilawa	Koranrezitation ohne Verzierungen, Pausen und Wiederholungen
Tilawa tartitl	Hörspiel, das den Gesetzen des Genres "Krimi" folgt (Aufklärung eines Kriminalfalles)
Kriminalhörspiel	Hörspiel, das überwiegend oder ausschließlich mit O-Ton-Material arbeitet.
Originaltonhörspiel	
	Gehört zur großen, nicht immer abendfüllenden Form der Unterhaltung. Urboden der Comedy sind schräge Abwandlungen (i.S. Parodie) ebenso wie der geballte Unsinn, jede Menge Situations- und Wortkomik, aber ohne gesellschaftsveränderndes Leitmotiv. Teilweise drastisch-derbe, respektlose Sprache.
Comedy	S / Große, meist abendfüllende Form der leichten Unterhaltung. Dramaturgisch offene Form des Theaters, die sich der verschiedenen Bühnengenres (Gesang, Instrumentalmusik, Tanz, Wort und Spiel) bedient. Meist Nummernprogramm mit übergreifender Idee oder Rahmen aus aufeinander abgestimmten Bildern. (Im Hörfunk eher selten, da auch auf optische Effekte aufbauend: Tanz, Akrobatik, Zauberei).
Revue	Versuch, das Andenken an einen kürzlich verstorbenen Menschen zu wahren, indem in der Regel auf würdigende Art und Weise an das Lebenswerk erinnert wird.
Nachruf	Hinführung zu einem Programminhalt
Moderation	

Zwischenmoderation	Einzelbeiträge (innerhalb einer Sendung) verbindende Moderation oder von einer Sendung zur nächsten überleitende Moderation
Rätselsendung	Als Spiel konzipierte Sendung, z.B. eine Sendung, in der den Hörern oder einem meist aus Prominenten und/oder Experten bestehendem Rateteam eine Denkaufgabe gestellt wird. Der zu erratende Begriff (meist eine Person, ein Gegenstand, ein Text oder ein Ereignis) wird dabei umschrieben oder verschlüsselt. Die Rätselfragen können während der Sendung oder (von den Hörern) nach der Sendung sch
Feature	(engl. = profilieren, herausstellen) Gestaltung eines journalistischen Themas mittels verschiedener oder aller radiophoner Ausdrucksformen. Die Dauer eines Features kann mehrere Stunden oder wenige Minuten betragen. Der Begriff Hörfolge wird meist synonym für Feature verwendet. Spezialformen des Feature sind (abweichend von der Definition) das Originaltonfeature und die Textmontage.
Amateurrhörspiel	Der Produzent, Urheber oder die Ausführenden (z.B. die Sprecher) oder der/die Urheber von einmontiertem Tonmaterial sind Amateure.
Essay	(engl.= Versuch, Probe) Prosatext über ein beliebiges Thema. Der Essay umkreist sein Thema nicht systematisch, sondern locker, assoziativ, kritisch, spielerisch und unterhaltend, stilistisch anspruchsvoll
Originaltonfeature	Feature, das ausschließlich oder überwiegend aus O-Tönen zusammengestellt ist.
Laudatio	Lobrede. Im Rahmen eines Festaktes gehaltene feierliche Rede, in welcher der Redner die Leistungen und Verdienste einer Person würdigt.
Tilawa tagwie	Koranrezitation mit Verzierungen, Pausen und Wiederholungen
Keine Gattung	/

108 Elemente

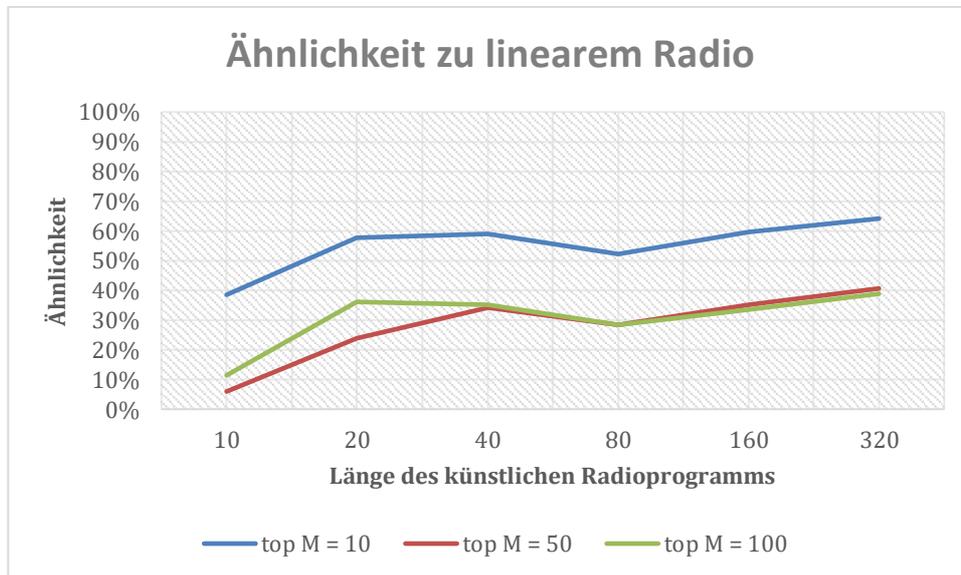
B. Anhang

Klasse: DLF, Wochentags, 0-6 Uhr	Gattung: DLF, Wochentags, 0-6 Uhr	Geographika: DLF, Wochentags, 0-6 Uhr
Klasse: DLF, Wochentags, 6-10 Uhr	Gattung: DLF, Wochentags, 6-10 Uhr	Geographika: DLF, Wochentags, 6-10 Uhr
Klasse: DLF, Wochentags, 10-18 Uhr	Gattung: DLF, Wochentags, 10-18 Uhr	Geographika: DLF, Wochentags, 10-18 Uhr
Klasse: DLF, Wochentags, 18-24 Uhr	Gattung: DLF, Wochentags, 18-24 Uhr	Geographika: DLF, Wochentags, 18-24 Uhr
Klasse: DLF, Wochenende, 0-6 Uhr	Gattung: DLF, Wochenende, 0-6 Uhr	Geographika: DLF, Wochenende, 0-6 Uhr
Klasse: DLF, Wochenende, 6-10 Uhr	Gattung: DLF, Wochenende, 6-10 Uhr	Geographika: DLF, Wochenende, 6-10 Uhr
Klasse: DLF, Wochenende, 10-18 Uhr	Gattung: DLF, Wochenende, 10-18 Uhr	Geographika: DLF, Wochenende, 10-18 Uhr
Klasse: DLF, Wochenende, 18-24 Uhr	Gattung: DLF, Wochenende, 18-24 Uhr	Geographika: DLF, Wochenende, 18-24 Uhr
Klasse: DLF Kultur, Wochentags, 0-6 Uhr	Gattung: DLF Kultur, Wochentags, 0-6 Uhr	Geographika: DLF Kultur, Wochentags, 0-6 Uhr
Klasse: DLF Kultur, Wochentags, 6-10 Uhr	Gattung: DLF Kultur, Wochentags, 6-10 Uhr	Geographika: DLF Kultur, Wochentags, 6-10 Uhr
Klasse: DLF Kultur, Wochentags, 10-18 Uhr	Gattung: DLF Kultur, Wochentags, 10-18 Uhr	Geographika: DLF Kultur, Wochentags, 10-18 Uhr
Klasse: DLF Kultur, Wochentags, 18-24 Uhr	Gattung: DLF Kultur, Wochentags, 18-24 Uhr	Geographika: DLF Kultur, Wochentags, 18-24 Uhr
Klasse: DLF Kultur, Wochenende, 0-6 Uhr	Gattung: DLF Kultur, Wochenende, 0-6 Uhr	Geographika: DLF Kultur, Wochenende, 0-6 Uhr
Klasse: DLF Kultur, Wochenende, 6-10 Uhr	Gattung: DLF Kultur, Wochenende, 6-10 Uhr	Geographika: DLF Kultur, Wochenende, 6-10 Uhr
Klasse: DLF Kultur, Wochenende, 10-18 Uhr	Gattung: DLF Kultur, Wochenende, 10-18 Uhr	Geographika: DLF Kultur, Wochenende, 10-18 Uhr
Klasse: DLF Kultur, Wochenende, 18-24 Uhr	Gattung: DLF Kultur, Wochenende, 18-24 Uhr	Geographika: DLF Kultur, Wochenende, 18-24 Uhr
Klasse: DLF Nova, Wochentags, 0-6 Uhr	Gattung: DLF Nova, Wochentags, 0-6 Uhr	Geographika: DLF Nova, Wochentags, 0-6 Uhr
Klasse: DLF Nova, Wochentags, 6-10 Uhr	Gattung: DLF Nova, Wochentags, 6-10 Uhr	Geographika: DLF Nova, Wochentags, 6-10 Uhr
Klasse: DLF Nova, Wochentags, 10-18 Uhr	Gattung: DLF Nova, Wochentags, 10-18 Uhr	Geographika: DLF Nova, Wochentags, 10-18 Uhr
Klasse: DLF Nova, Wochentags, 18-24 Uhr	Gattung: DLF Nova, Wochentags, 18-24 Uhr	Geographika: DLF Nova, Wochentags, 18-24 Uhr
Klasse: DLF Nova, Wochenende, 0-6 Uhr	Gattung: DLF Nova, Wochenende, 0-6 Uhr	Geographika: DLF Nova, Wochenende, 0-6 Uhr
Klasse: DLF Nova, Wochenende, 6-10 Uhr	Gattung: DLF Nova, Wochenende, 6-10 Uhr	Geographika: DLF Nova, Wochenende, 6-10 Uhr
Klasse: DLF Nova, Wochenende, 10-18 Uhr	Gattung: DLF Nova, Wochenende, 10-18 Uhr	Geographika: DLF Nova, Wochenende, 10-18 Uhr
Klasse: DLF Nova, Wochenende, 18-24 Uhr	Gattung: DLF Nova, Wochenende, 18-24 Uhr	Geographika: DLF Nova, Wochenende, 18-24 Uhr

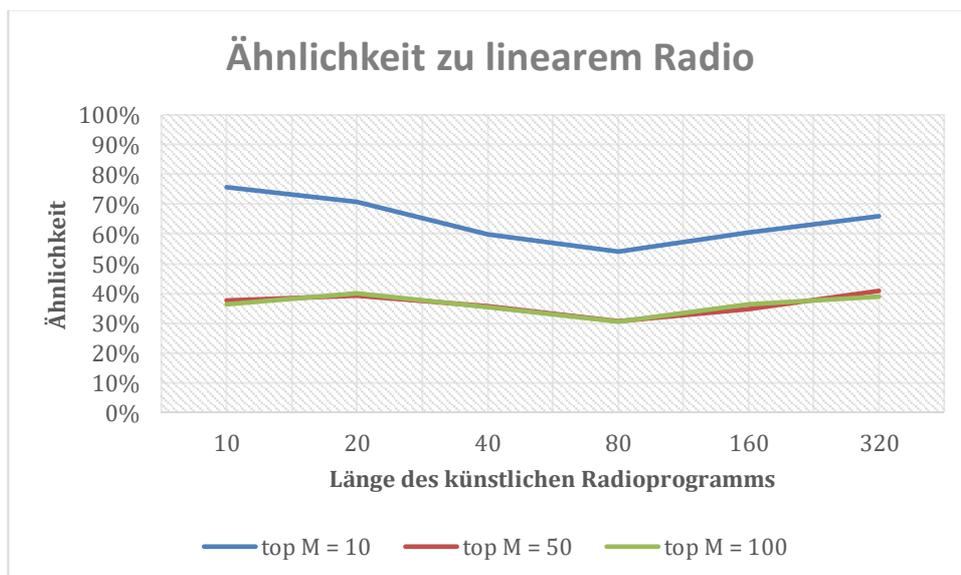
C. Anhang

Sequenzlängen der Länge 5 und Minsup von 0,005:

1) Für die Gattung Nachrichten:

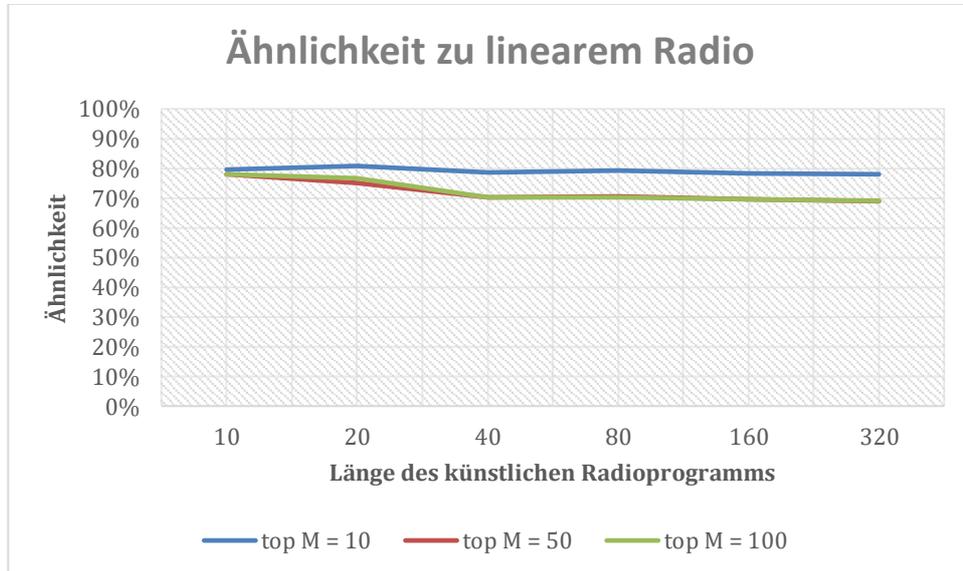


2) Für die Gattung Interview:

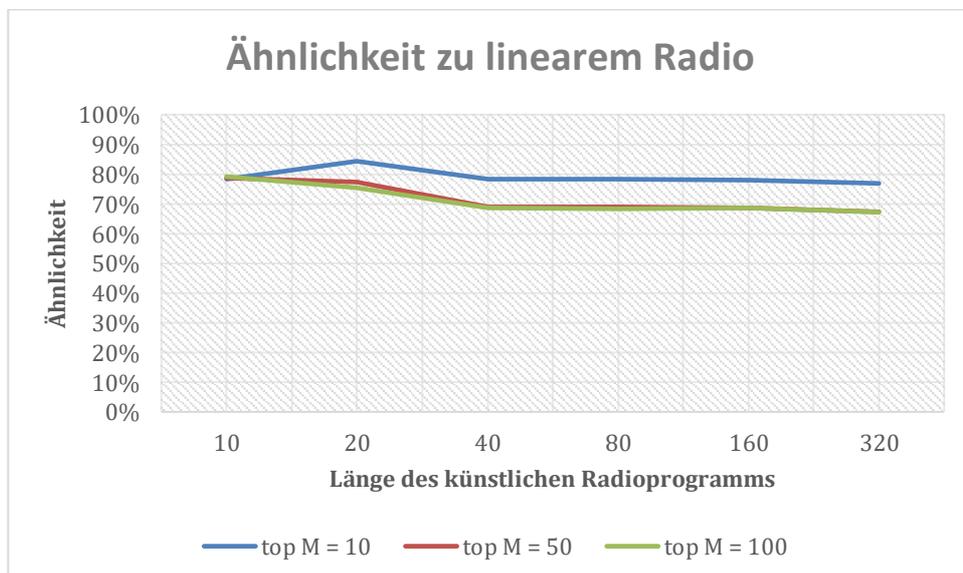


Sequenzlängen der Länge 5 und Minsup von 0,007:

1) Für die Gattung Nachrichten (Achtung, die war hier nicht vorhanden. Daher wurde ein initiales Muster für die Gattung Nachrichten aus der Musteranalyse mit Minsup von 0,005 übernommen):



2) Für die Gattung Interview:

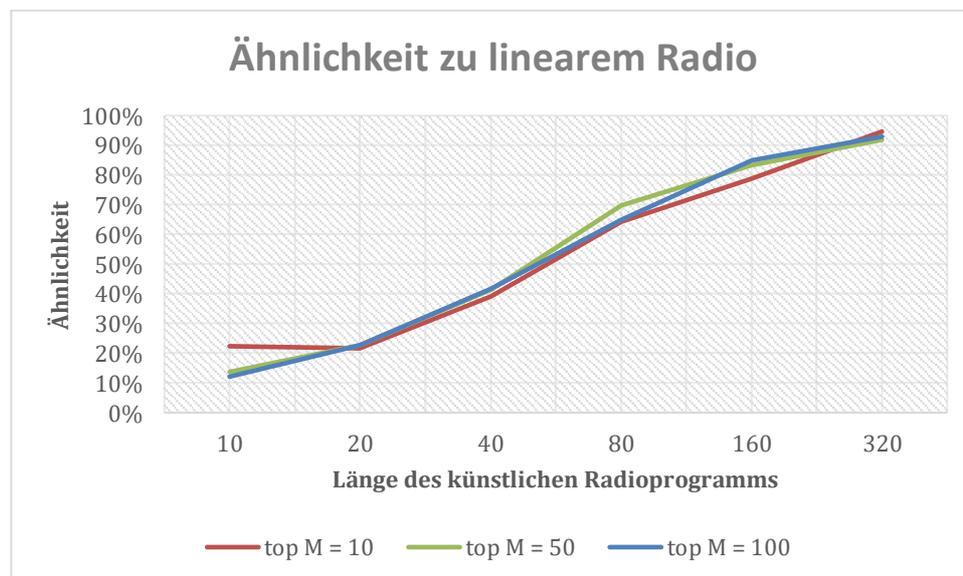


D. Anhang

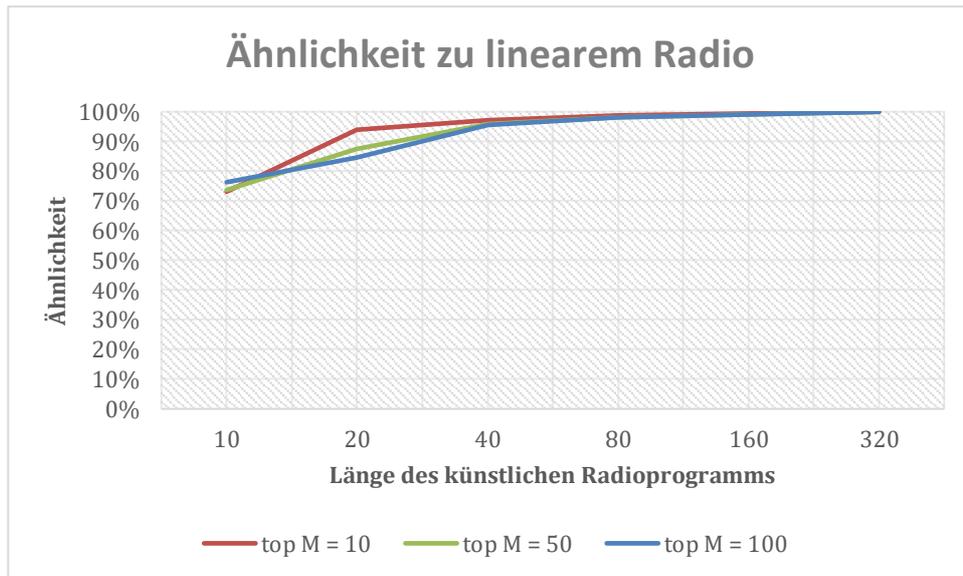
Hinweis: Es folgt eine Evaluierung durch die Generierung der synthetischen Sequenzen basierend auf den anderen Oberkategorien Bezugsreichweite und Klassen. Es werden lediglich die Top 100 Muster präsentiert, also auf eine Darstellung der Top 10 oder Top 50, wie es bei den Gattungen der Fall war, verzichtet. Denn es soll lediglich gezeigt werden ob die Erkenntnisse dieser Oberkategorien mit den Ergebnissen der Gattungen übereinstimmen.

Evaluierung der synthetischen Reihenfolgen basierend auf der Oberkategorie Geographika; Betrachtung: Top 100 Muster; Sequenzlänge: 4

1) 0.007 Minsup, 4er Länge:

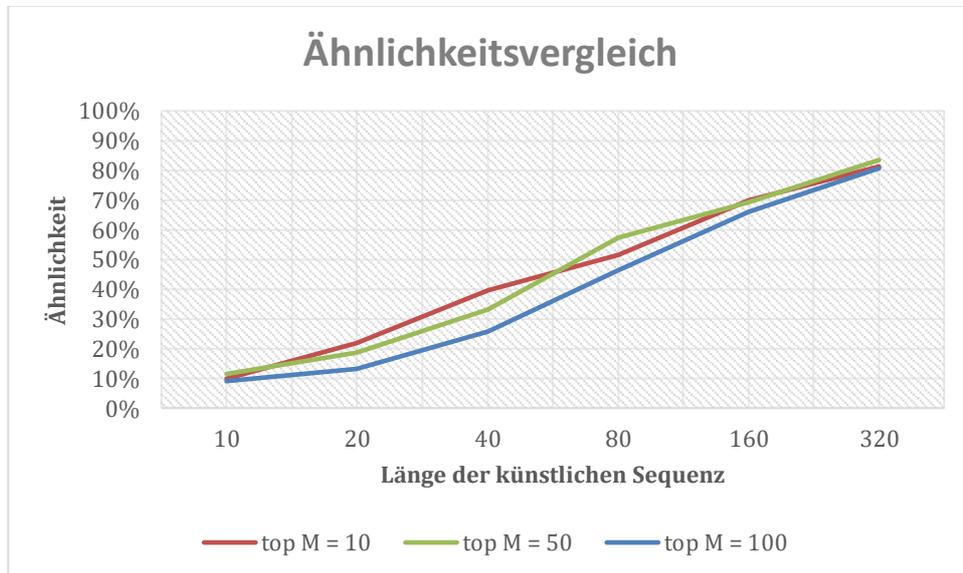


2) 0.05 Minsup, 4er Länge:



Evaluierung der synthetischen Reihenfolgen basierend auf der Oberkategorie Klassen; Betrachtung: Top 100 Muster; Sequenzlänge: 4

1) 0.007 Minsup, 4er Länge:



2) 0.05 Minsup, 4er Länge:

