



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Masterthesis

Max Bahne

Konzept eines Recommender Systems für die
NDR Mediathek

Max Bahne

Konzept eines Recommender Systems für die NDR Mediathek

Masterarbeit eingereicht im Rahmen der Masterprüfung

im Studiengang Next Media
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften

Betreuender Prüfer: Prof. Dr. Kai von Luck
Zweitgutachterin: Dr. Susanne Draheim

Eingereicht am: 04.08.2016

Max Bahne

Thema der Arbeit

Konzept eines Recommender Systems für die NDR Mediathek

Stichworte

Recommender System, Mediennutzung, NDR

Kurzzusammenfassung

Mit dem veränderten Mediennutzungsverhalten gerade jüngerer Rezipienten stehen die traditionellen Medienanbieter vor großen Herausforderungen. Diese Masterarbeit beschäftigt sich mit der Personalisierung der NDR Mediathek mit Hilfe von Recommender Systems. Dazu werden Mediennutzungsverhalten der deutschen Gesellschaft, die Struktur des öffentlich-rechtlichen Rundfunks sowie die Grundlagen der Recommender Systems untersucht, um aus den in diesen Bereichen gewonnenen Erkenntnissen ein Konzept für ein Recommender System für die NDR Mediathek zu erstellen.

Max Bahne

Title of the paper

Concept of a Recommender System for the NDR Media Library

Keywords

Recommender System, Media Use, NDR

Abstract

The media use – especially of the younger recipients – has changed a lot in the recent years and has challenged the traditional media contributors. This Masterthesis discusses the personalization of the NDR Media Library with the help of Recommender Systems.

This thesis dives into the media use of the German society, the structure of the NDR and other broadcasting corporations and the basics of Recommender Systems. The concept of a Recommender System for the NDR Media Library should be derived from the insights of the examined domains.

Inhaltsverzeichnis

| | |
|---|----|
| 1. Einleitung | 8 |
| 1.1 Motivation..... | 10 |
| 1.2 Thesis..... | 11 |
| 1.3 Thematische Eingrenzung..... | 11 |
| 1.4 Gliederung | 12 |
| 1.5 Begriffsdefinition des Recommender Systems..... | 13 |
| 1.6 Verwendete Methoden..... | 13 |
| 2. Veränderung des Mediennutzungsverhaltens innerhalb der Gesellschaft in Deutschland..... | 14 |
| 2.1 Überblick zur Mediennutzung | 15 |
| 2.1.1 Was bedeutet Mediennutzung? | 15 |
| 2.1.2 Motivation zur Mediennutzung | 16 |
| 2.1.2.1 Entstehung von Gewohnheit bei der Mediennutzung..... | 18 |
| 2.2 Veränderungen in der Mediennutzung..... | 20 |
| 2.2.1 Voraussetzungen für die Mediennutzung in Deutschland..... | 20 |
| 2.2.1.1 Zeitungen | 20 |
| 2.2.1.2 Radio | 21 |
| 2.2.1.3 Fernsehen | 22 |
| 2.2.1.4 Internet..... | 22 |
| 2.3 Veränderungen in der Mediennutzung..... | 24 |
| 2.3.1 Mediennutzung 1964 | 24 |
| 2.3.2 Mediennutzung 2015 | 27 |
| 2.3.2.1 Mediennutzung 2015 – Standarddaten | 27 |
| 2.3.2.2 Mediennutzung 2015 – Weitere erhobene Daten..... | 29 |
| 2.3.2.3 Nutzung der Mediatheken und Online-Videoangebote | 34 |
| 2.3.3 Veränderungen in der Mediennutzung | 36 |
| 2.4 Zwischenfazit | 41 |
| 3. Die Grundlagen eines Recommender Systems..... | 42 |
| 3.1 Definition eines Recommender Systems..... | 43 |
| 3.2 Zielsetzung und Zweck der Recommender Systems..... | 44 |
| 3.3 Implementierung eines Recommender Systems in eine Webseite..... | 45 |
| 3.4 Filter-Methoden für ein Recommender System..... | 47 |
| 3.4.1 Die Collaborative Filtering-Methode..... | 48 |
| 3.4.2 Die Methodik des Content-based Recommender Systems | 51 |
| 3.4.3 Die Methodik des Knowledge-based Recommender Systems | 52 |
| 3.4.4 Hybride Recommender Systems | 54 |
| 3.5 Die Entstehung einer Empfehlung im Recommender System..... | 55 |
| 3.5.1 Voraussetzungen für die Empfehlungserstellung..... | 55 |
| 3.5.2 Die Empfehlungserstellung beim Item-based Collaborative Filtering | 56 |
| 3.5.3 Die Empfehlungserstellung im Content-based Recommender System..... | 58 |
| 3.5.4 Die Empfehlungserstellung im Constraint-based Recommender System | 58 |
| 3.5.5 Zusammenkommen der Empfehlungen im hybriden Recommender System .. | 61 |
| 3.5.6 Evaluation des Recommender Systems | 61 |
| 3.6 Netzwerkeffekte und Recommender Systems..... | 62 |
| 3.7 Herausforderungen bei der Einführung eines Recommender Systems | 63 |
| 3.8 Zwischenfazit | 67 |
| 4. Rechtliche und organisatorische Struktur der öffentlich-rechtlichen Mediatheken..... | 67 |

| | |
|---|-----|
| 4.1 Auftrag der öffentlich-rechtlichen Mediatheken | 68 |
| 4.1.1 Problematiken bei der Ausgestaltung des Auftrags | 70 |
| 4.2 Der Drei-Stufen-Test als Legitimierung der Online-Aktivität | 73 |
| 4.2.1 Die drei Stufen des Tests | 75 |
| 4.2.2 Ablauf des Drei-Stufen-Tests | 77 |
| 4.3 Der Drei-Stufen-Test der NDR Mediathek | 78 |
| 4.3.1 Das Telemedienkonzept der NDR Mediathek..... | 78 |
| 4.3.2 Ergebnis des Drei-Stufen-Tests | 81 |
| 4.4 Arbeitsprozess in der NDR Mediathek..... | 82 |
| 4.4.1 Was ist ein CMS?..... | 82 |
| 4.4.2 Die Datenstruktur in der NDR Mediathek..... | 85 |
| 4.4.2.1 In Sophora erfasste Daten und Metadaten..... | 86 |
| 4.4.2.2 Durchs Tracking gesammelte Nutzungsdaten..... | 90 |
| 4.4.3 Arbeitsablauf bei der Produktion und Veröffentlichung eines Videos..... | 95 |
| 4.4.4 Inhalte der NDR Mediathek..... | 98 |
| 4.5 Datenschutz beim NDR | 102 |
| 4.6 Der NDR und die Disruption der Medienlandschaft | 104 |
| 4.7 Zwischenfazit | 105 |
| 5. Konzept eines Recommender Systems für die NDR Mediathek..... | 107 |
| 5.1 Zielsetzung des Recommender Systems für die NDR Mediathek..... | 107 |
| 5.2 Spezielle Herausforderungen für ein Empfehlungssystem in der NDR Mediathek | 109 |
| 5.3 Elemente und Struktur des Recommender Systems für die NDR Mediathek..... | 109 |
| 5.3.1 Implementierung des Recommender Systems in die ndr.de-Seite | 110 |
| 5.3.2 Der Empfehlungsprozess in der NDR Mediathek | 111 |
| 5.3.3 Im Recommender System verwendete Kontext-Informationen | 113 |
| 5.4 Zwischenfazit | 113 |
| 6. Zusammenfassung und Ausblick..... | 114 |
| 6.1 Zusammenfassung | 114 |
| 6.2 Ausblick | 115 |
| 7. Literaturverzeichnis & Quellenverzeichnis..... | 117 |
| 8. Abbildungsverzeichnis | 122 |

Danksagung

Das Kapitel über die Struktur der NDR Mediathek hätte niemals in der Form entstehen können, wie es entstanden ist, hätte ich mit Alexander Svensson, dem Leiter der Mediathek, Maïke Lahmann, die sich sehr gut mit den Arbeitsprozessen innerhalb der Mediathek auskennt, Max Nellen, der sich in der Medienforschung um das Web-Tracking kümmert, und Horst Brendel, der als stellvertretender Justitiar und Datenschutzbeauftragter Rede und Antwort stand, nicht sehr viele Leute beim NDR gefunden, die mir in den Experteninterviews ausführlich Auskunft über die Rundfunkanstalt gegeben haben. Dafür möchte ich mich herzlich bei ihnen bedanken.

1. Einleitung

„Like the physical universe, the digital universe is large – by 2020 containing nearly as many digital bits as there are stars in the universe. [...] But unlike the physical universe, the digital universe is created and defined by software – a man-made construct. It is defined by software that analyzes this ever-expanding universe of digital data, finding the hidden value and new opportunities to transform and enhance the physical world [...] And it is software that will both create new opportunities and new challenges for us as we try to extract value from the digital universe that we created.“¹

Diese Worte stammen aus der Feder von John F. Gantz, dem Senior Vice President der International Data Corporation (IDC), und zeichnen ein sehr treffendes Bild des World Wide Webs (WWW) der heutigen Zeit. Jeder User, der sich durchs Web bewegt, hinterlässt zwangsläufig einen immer größer werdenden Datenberg.

Parallel zu den Datenspuren, die man im Netz hinterlässt, steigt auch die Dichte des angebotenen Contents immer weiter an – seien es per Streaming abrufbare Serien oder Blockbuster, News-Angebote oder auch E-Commerce-Angebote.

Zusammengenommen nimmt der Datenberg, der schon heute im Netz mehr oder weniger öffentlich zugänglich ist, riesige Ausmaße an und wird in Zukunft exponentiell weiterwachsen².

Gerade der letztgenannte Punkt – das nahezu exponentielle Wachstum von Angeboten von Inhalten im Web – machte es notwendig, eine Art Filter zu erschaffen, der es Nutzern ermöglicht, das große Angebot des Webs zu sortieren und nur die Dinge zu finden, die für ihn relevant sind. Einen ersten Schritt in diese Richtung machten schon kurz nach dem Start des öffentlich zugänglichen WWW zu Beginn der 1990-er Jahre die Suchmaschinen³. Die Suchmaschinen ermöglichten

¹ S. Gantz, John u.a.: The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, Framingham, 2014. Die Autoren der Studie berechneten, dass im Jahr 2013 rund 4,4 Zettabyte an Daten produziert wurden und sich dieser Wert bis zum Jahr 2020 auf 44 Zettabyte verzehnfachen wird. Würde man diese Datenmenge auf iPads mit einer Speicherkapazität von 128 Gigabyte und einer Dicke von 6,1 Millimetern speichern und diese stapeln, würde der Stapel fast sieben Mal von der Erde bis zum Mond reichen.

² Vgl. Ebd.

³ Vgl. Conolly, Dan: A Little History of the World Wide Web, Princeton, 2000. An der Forschungseinrichtung CERN in der Schweiz entwickelte Tim Berners-Lee das Konzept

es dem Anwender, das Web nach für ihn relevanten Informationen zu durchsuchen, und wurden so zu einer der wichtigsten Anlaufstellen im Web. Google war am erfolgreichsten damit, Geld mit den Suchanfragen der Nutzer zu verdienen und löste vor Kurzem Apple als wertvollstes Unternehmen der Welt ab⁴.

Beim Blick unter die simple Benutzeroberfläche von Google oder anderer Suchmaschinen offenbart sich das recht einfache Konzept. Über einen Webcrawler werden nahezu alle Webseiten in einer Datenbank indiziert, ihre heuristische Struktur von der Software „verstanden“ und im Fall von Google mit dem PageRank-Algorithmus bewertet. Bei dieser Bewertung fließen eine Menge Faktoren ein, wie zum Beispiel die Anzahl der Links, die auf diese Seite verweisen, wie viele Nutzer über Google auf diese Seite klickten und wie lange sie dort verweilten, bis sie zur Google-Suche zurückkehrten. Durch den PageRank-Algorithmus entsteht zu großen Teilen das Suchergebnis. Zusätzlich fließt aber auch noch ein anderer wesentlicher Faktor mit ein: Die Vorlieben des Nutzers. Beliebt ist das Java-Beispiel: Menschen, die oft nach Reisen suchten oder in der Vergangenheit über Google Urlaubsportale besuchten, bekommen als erstes Suchergebnis Informationen zur Insel Java angezeigt, während ein Entwickler eher Suchergebnisse zur Programmiersprache Java erhält. Ein großer Teil des Erfolgs der Suchmaschinen wie Google und Co. basiert folglich auf der Personalisierung der Suchergebnisse für jeden individuellen Nutzer. Eine Option zur Personalisierung einer Webseite liefern dabei Recommender Systems. Dabei ist die Idee der Empfehlung keine Idee des Internets. Mit den ersten Zeitungen im 19. Jahrhundert etablierte sich bis heute eine Kritikerschaft, die Bücher, Filme und Videospiele bespricht und den Lesern empfiehlt oder nicht empfiehlt. Diese Art des „Empfehlungssystems“ fand seinen Weg auch in die Internetfirmen der späten 1990-er Jahre. So leistete sich Amazon zum Beispiel eine Redaktion, die Bücher rezensierte und den Usern weiterempfahl. James Marcus, Mitarbeiter Nummer 55 bei Amazon, beschreibt in „Amazonia“ den Alltag dieser

für das WWW, in den Folgejahren fand das Konzept immer mehr private Anwender außerhalb der Forschung.

⁴ Vgl. <http://www.zeit.de/wirtschaft/unternehmen/2016-02/google-alphabet-quartalszahlen-rekord>, zuletzt besucht am 06.03.2016 um 15:30 Uhr. Google selbst ist der Muttergesellschaft Alphabet eingegliedert, die Apple überholte und zum wertvollsten Unternehmen der Welt wurde.

Redaktion, bis sie schließlich durch ein digitales Recommender System, also einen Empfehlungsalgorithmus ersetzt wird⁵.

Bis zum heutigen Tag sind die Recommender Systems wesentlich komplexer und ausgefeilter geworden und in nahezu allen größeren Plattformen integriert. So nutzen zum Beispiel neben Amazon auch die Videostreamingplattform Netflix, das Auktionshaus eBay, die Musikplattform iTunes oder das Bewertungsportal Yelp eine Recommendation Engine⁶. Ein großer Teil des Webs wie wir es kennen ist also mittlerweile über ein Recommender System den individuellen Vorlieben des Nutzers angepasst.

1.1 Motivation

Obwohl der Journalismus in gewisser Art und Weise ein Mitbegründer des „Empfehlungssystems“ ist, gibt es in den News-Redaktionen bisher noch kein überzeugendes oder nur ein sehr rudimentäres Empfehlungssystem⁷. Deshalb soll in der vorliegenden Masterarbeit das Konzept eines Recommender Systems für die Mediatheken des öffentlich-rechtlichen Rundfunks, speziell der NDR Mediathek, entwickelt werden.

Die Motivation zu dieser Arbeit speist sich aus zwei Quellen. Zum einen arbeitete der Autor der vorliegenden Arbeit in der NDR Mediathek und kennt sich mit den Arbeitsabläufen und der Datenstruktur der dort verarbeiteten Videos aus. Zum anderen bestand seitens des NDR – und übergeordnet der ARD – der Wunsch, ein Empfehlungssystem für die eigenen Mediatheken zu entwickeln. Diese Planungen steckten jedoch noch in der Planungsphase, als der Autor dieser Masterthesis mit der Arbeit begann. Daher bot es sich an, die Konzeptarbeit parallel zu den ersten Planungen der ARD-Verantwortlichen zu erstellen.

Aus diesen beiden Motivationsgründen heraus entwickelte sich die Kooperation, die zur Entstehung der vorliegenden Arbeit führte.

⁵ Vgl. Marcus, James: Amazonia. Fünf Jahre im Zentrum der E-Commerce-Revolution, Berlin, 2006, S. 248.

⁶ Ricci, Francesco; Rokach, Lior; Shapira, Bracha: Introduction to Recommender Systems Handbook, In: Recommender Systems Handbook, hrsg. von ebendiesen, New York, 2011, S. 3.

⁷ Das Spektrum reicht hier von den Medien, die gar kein Empfehlungssystem nutzen, bis hin zu Medien, die aufgrund vom Thema oder Beliebtheit des Contents Vorschläge machen. Zu letztgenannten gehören zum Beispiel Spiegel Online oder auch Zeit Online.

1.2 Thesis

Diese Arbeit soll also einen ersten Ansatzpunkt für den öffentlich-rechtlichen Rundfunk bieten, um mit der Einführung eines Empfehlungssystems in den eigenen Mediatheken beginnen zu können. Dafür soll erarbeitet werden, welche Probleme bei der Einführung eines solchen Systems zu beachten sind und welche Lösungsansätze in der aktuellen Forschung bisher bereitgestellt werden. Hinzu kommen noch spezifische Problemstellungen des öffentlich-rechtlichen Rundfunks, die bei der Konzeption eines Recommender Systems beachtet werden müssen. Dazu zählen zum Beispiel Datenschutzfragen und fehlende Nutzerprofile. Dementsprechend lautet die Leitfrage der vorliegenden Arbeit: „Welche Probleme existieren bei der Konzeption eines Recommender Systems für die öffentlich-rechtlichen Mediatheken und welche Lösungsansätze stehen dafür bereit?“. Mit der Antwort auf diese Leitfrage soll idealerweise ein Grundgerüst errichtet werden, worauf weitere Überlegungen des öffentlich-rechtlichen Rundfunks zur Einführung eines Empfehlungssystems für die eigenen Mediatheken aufbauen können.

1.3 Thematische Eingrenzung

Wie bereits im vorhergehenden Kapitel aufgeführt, soll diese Arbeit ein Konzept für ein Recommender System für die öffentlich-rechtlichen Mediatheken erarbeiten. Darunter fällt vor allem die Recherche zu den generellen Problematiken der Recommender Systems und den spezifischen Problemen der öffentlich-rechtlichen Mediatheken.

Was diese Arbeit nicht leistet, ist die Erschaffung einer lauf- und einbindungsfähigen Recommendation Engine für die öffentlich-rechtlichen Mediatheken auf der Basis von Empfehlungs-Algorithmien. Sie soll vielmehr die Vorarbeit liefern, um solch eine Engine bauen zu können.

Eine weitere – zumindest leichte - thematische Eingrenzung ist die „Beschränkung“ dieser Arbeit auf das Content-Management-System (CMS) Sophora⁸. Zwar soll ein Kapitel der Masterthesis generell das Thema CMS behandeln, dennoch baut das Konzept für das Recommender System der Mediatheken auf Sophora und dessen

⁸ Sophora ist ein von dem Hamburger Unternehmen subshell entwickeltes Content-Management-System, das laut Aussage des Unternehmens „dank seiner flexiblen Möglichkeiten für das multimediale Publikation- und Verwaltungsmanagement [...] die Antwort auf die steigenden Anforderungen moderner Medienunternehmen“ ist. S. <http://www.subshell.com/de/subshell/index.html>; zuletzt besucht am 10.03.2016 um 10:55 Uhr.

Datenstruktur auf, da es mittlerweile bei fast allen Rundfunkanstalten in Verwendung ist⁹. Weil der Autor dieser Arbeit keinen Einblick in die Systeme abseits von Sophora gewinnen kann, könnte es bei der Übertragbarkeit des Konzepts auf andere Systeme zu Komplikationen kommen, die vor allem im Bereich der Konsistenz der verfügbaren Daten-Basis liegen.

1.4 Gliederung

Um die Leitfrage dieser Arbeit sinnvoll bearbeiten zu können, ist sie in die folgende Gliederung unterteilt. Auf die Einleitung, die als Einführung ins Thema dienen und ebenso Fragen zu Methoden und grundlegenden Begriffen klären soll, folgt ein Kapitel zur Veränderung der Mediennutzung innerhalb der deutschen Gesellschaft. Hier soll vor allem die Frage beantwortet werden, wie sich die Mediennutzung der Menschen in Deutschland über die Jahre von einer getakteten Nutzung zu einer enttakteten entwickelte und welche Faktoren in diesem Prozess eine wichtige Rolle spielten. Außerdem soll dieses Kapitel auch beschreiben, wie sich der ausgespielte Content selbst über die Jahre verändert hat.

Im nächsten Kapitel soll es um die Struktur der öffentlich-rechtlichen Mediatheken gehen. Neben dem Auftrag der Mediatheken sollen hier auch die Voraussetzungen sowie die Dateninfrastruktur der Mediathek erläutert werden. Einen großen Teil in diesem Kapitel nimmt die Beschreibung des Arbeitsprozesses in der Mediathek sowie der Inhalte der Mediathek ein.

Das nächste Kapitel soll auf die Anforderungen der öffentlich-rechtlichen Mediatheken für ein Recommender System eingehen. Hier sollen kurz die Grundlagen dieser Systeme erläutert werden, um zu schauen, welche Arten von Empfehlungssystemen sich überhaupt für die öffentlich-rechtlichen Mediatheken eignen.

Im anschließenden Kapitel sollen die Grundlagen der Recommender Systems genauer beschrieben werden. Hier soll beschrieben werden, welche Elemente ein Recommender System enthält und wie solch ein System in eine Webseite implementiert werden kann. Zusätzlich soll es in diesem Kapitel auch um die verschiedenen Filter-Methoden gehen, die zur Empfehlungserstellung verwendet werden.

⁹ Lediglich RBB, SWR und SR nutzen Sophora noch nicht.

Im letzten Kapitel dieser Arbeit sollen die Erkenntnisse der vorhergehenden Kapitel zusammengefasst werden und auf ein Recommender System für die NDR Mediathek übertragen werden.

Zum Ende der Arbeit soll außerdem noch ein Schlussfazit gezogen werden, das noch einmal auf die weiter oben formulierte Leitfrage eingeht. Interessante Themenaspekte, für die in dieser Arbeit kein Platz mehr war, sollen in einem Ausblick aufgeführt werden – ebenso wie auf diese Arbeit folgende Arbeitsschritte.

1.5 Begriffsdefinition des Recommender Systems

In der Literatur finden sich zwei verschiedene Schreib- und Sprecharten zu den Recommender Systems. Vor allem in der etwas älteren Literatur fällt oft der Begriff „Recommendation System“, der sich wohl als eine Mischform der „Recommendation Engine“ und des „Recommender Systems“ entwickelte. Demgegenüber steht die Bezeichnung „Recommender System“, auf die im Großteil der aktuellen Literatur zurückgegriffen wird. Da es noch keine allgemeingültige Definition des Begriffes gibt, entscheidet sich der Autor dieser Arbeit, durchgängig die aktuell verwendete Bezeichnung „Recommender System“ zu nutzen.

1.6 Verwendete Methoden

Für das Kapitel über die Struktur des öffentlich-rechtlichen Rundfunks greift diese Arbeit auf die Methode des Experteninterviews zurück, da es zu dieser Thematik bislang sehr wenig Literatur gibt. Die Methode des Experteninterviews kommt aus der Sozialforschung und musste lange Zeit mit dem Ruf einer „gewissen Hemdsärmlichkeit“¹⁰ leben. Über die Jahre entwickelte sie sich aber zu einer anerkannten und häufig verwendeten Methode. Im Grunde geht es beim Experteninterview um eine Analyse der „Strukturen und Strukturzusammenhänge von Expertenwissens und –handelns“¹¹. Idealerweise wird vorher vorhandenes Wissen über eine gewisse Thematik – in diesem Fall das Wissen des Autors über die NDR Mediathek – von einem Experten gestützt und mit zusätzlichen Informationen versehen, die sich in der Arbeit verwenden lassen. Als Experten gelten im Fall dieser Arbeit vor allem Mitarbeiter des NDR, die Kontextwissen über die Arbeitsprozesse und Strukturen innerhalb der Mediathek

¹⁰ S. Meuser, Michael; Nagel, Ulrike: Das Experteninterview – konzeptionelle Grundlagen und methodische Anlage, In: Methoden der vergleichenden Politik- und Sozialwissenschaft, hrsg. von Detlef Jahn, Wiesbaden, 2009, S. 465.

¹¹ S. Bogner, Alexander [u.a.]: Das Experteninterview, Opladen, 2005, S. 73.

haben. Als „Kontextwissen“ gilt Wissen, das von den Experten durch „ihre Tätigkeit als spezialisiertes Sonderwissen erworben“¹² wurde.

Ziel der geführten Experteninterviews soll ein tiefergreifendes Verständnis der Funktionsweisen innerhalb der öffentlich-rechtlichen Mediatheken sein.

Ansonsten wird in der vorliegenden Arbeit sehr stark die Literaturrecherche genutzt, da zu den Themen Mediennutzungsverhalten und Recommender Systems sehr viel Literatur vorhanden ist.

2. Veränderung des Mediennutzungsverhaltens innerhalb der Gesellschaft in Deutschland

„Die dynamische und technische Entwicklung von Formaten und Angebotsformen im Internet führt in der näheren Zukunft zu sich stetig ändernden Anforderungen seitens der Mediennutzer. Beispiele für diesen permanenten Prozess sind die mittlerweile zum Standardrepertoire publizistischer Internetangebote gehörenden Partizipationsmöglichkeiten, die unter dem Schlagwort ‚Web 2.0‘ zusammengefasst werden. Telemedien müssen – wie die klassischen Medien auch – auf diese Veränderungen reagieren und neue Funktionalitäten in ihr Repertoire aufnehmen, um am Markt bestehen zu können.“¹³

In diesem Kapitel geht es um das veränderte Mediennutzungsverhalten der Menschen in Deutschland, an dessen Ende der Aufstieg von Netflix und anderer Streamingdienste steht – und damit auch der Aufstieg der Algorithmus-basierten Empfehlungssysteme.

Zu Beginn soll ein Überblick über das Thema Mediennutzung gegeben werden: Was bedeutet der Begriff überhaupt und – noch interessanter – welche Motive und Motivationen haben die Menschen für die Mediennutzung? Im zweiten Teil dieses Kapitels sollen explizit die Veränderungen im Mediennutzungsverhalten der Deutschen erarbeitet werden. Hier wird die Mediennutzung von 1964 mit der von 2015 verglichen, um daraus Schlüsse zu ziehen, in welche Richtung sich die Nutzung veränderte.

Am Ende dieses Kapitels soll ein Zwischenfazit gezogen werden, das auf die Bedeutung der Recommender Systems für die veränderte Mediennutzung eingeht.

¹² S. Bogner: Das Experteninterview, S. 73.

¹³ Vgl. Telemedienkonzepte des Norddeutschen Rundfunks, Angebotsbeschreibung für tagesschau.de und eins-extra.de (überarbeitete Fassung – Stand 25. Juni 2010), S. 41.

2.1 Überblick zur Mediennutzung

2.1.1 Was bedeutet Mediennutzung?

Eine allgemeine Definition des Begriffs Mediennutzung ist in der Literatur nicht zu finden. Vielmehr wird das Wort auf den Begriff „Medien“ reduziert und unter Mediennutzung eine Interaktion mit ebendiesen verstanden, die durch das Sehen, Hören oder seltener auch Fühlen erfolgen kann. Daher ist es sinnvoll, die Ausprägungen des Begriffs der „Medien“ zu verstehen und sich darauf mit der Mediennutzung auseinanderzusetzen. Der Medien- und Kommunikationswissenschaftler Thomas Mock stellte 2006 eine sehr treffende Definition des Begriffs „Medium“ auf. Er unterteilte ihn in vier verschiedene Teile¹⁴:

1. Der *physikalische* Medienbegriff. Dieser Medienbegriff schließt das Mittel der Wahrnehmung mit ein und damit auch die Voraussetzung zur Kommunikation, zum Beispiel Übertragungswege wie Glasfaserkabel oder Mobilfunknetze.
2. Der *codebezogene* Medienbegriff. Dieser Medienbegriff umkreist das Mittel der Verständigung. Dazu zählen zum Beispiel semiotische Elemente wie Zahlen und Zeichen, aber ebenso mimische und gestische Ausdrücke.
3. Der *technische* Medienbegriff. Er umschließt das Mittel der Verbreitung und entspricht daher dem klassischen Verständnis eines Mediums. Dazu zählen zum Beispiel Papier, digitale Datenträger, aber auch Computer, Fernseher oder Radios.
4. Der *soziale* Medienbegriff. Er bezeichnet Formen der Kommunikation wie zum Beispiel das Gespräch zwischen zwei Menschen, Briefe, Telefongespräche oder auch E-Mails.

Wie bereits erwähnt bedeutet Mediennutzung in diesem Zusammenhang, dass ein Individuum auf beliebige Art und Weise mit einem beliebigen technischen Medium interagiert. Deutlich besser in der Literatur repräsentiert ist die Ambivalenz des Begriffs „Mediennutzung“: Schon im ausgehenden 17. Jahrhundert sorgte die „Zeitungslust“ mancher Zeitgenossen für Kritik¹⁵. Mit dem Aufkommen neuer Medien wandelte sich auch das Bild der Kritik: Heute ist das Klischee des übergewichtigen Kindes, das große Teile seiner Freizeit mit Videospielen und Fernsehen verbringt, omnipräsent – „Zeitungslust“ ist hingegen gesellschaftlich schon lange anerkannt.

¹⁴ Vgl. Mock, Thomas: Was ist ein Medium? Eine Unterscheidung kommunikations- und medienwissenschaftlicher Grundverständnisse eines zentralen Begriffs, In: Publizistik (Band 51), Leipzig, 2006, S. 195.

¹⁵ Vgl. Schweiger, Wolfgang: Theorien der Mediennutzung – Eine Einführung, Wiesbaden, 2007, S. 11.

Auf der anderen Seite ist die Mediennutzung ein Schlüssel zur Teilhabe am gesellschaftlichen Leben und indirekt auch ein Standbein der Demokratie. Nicht umsonst schwingt mit der Verbreitung des Internets auch die Hoffnung mit, mit genügend Netzwerkanschlüssen und der nötigen Medienkompetenz die Bedingungen für die gelebte Demokratie auch in Ländern ohne demokratische Regierungen zu verbessern¹⁶. So spielt die Mediennutzung trotz ihrer ambivalenten Bedeutung im alltäglichen Leben jedes Menschen eine zentrale Rolle, weshalb sie auch ein zentraler Punkt vieler wissenschaftlicher Arbeiten ist.

2.1.2 Motivation zur Mediennutzung

Bei der Frage, was die Menschen zur Mediennutzung antreibt, kann grob zwischen einer willkürlichen, nicht-intentionalen Mediennutzung und der gewohnheitsmäßigen, intentionalen Mediennutzung unterschieden werden. Die Kommunikationswissenschaft hat allerdings herausgefunden, dass beiden Nutzungsarten dieselbe Motivation zugrunde liegt: Nämlich die Bedürfnisbefriedigung der einzelnen Nutzer. Grundlage für diese Annahme ist der Nutzen-Belohnungs-Ansatz, der ein Konstrukt der Mediennutzungsforschung ist, der „die Motive der Rezipienten und die Funktionen der Medien für ihr Publikum in den Blick nimmt“¹⁷. Der Ansatz setzt ein absichtsvolles Handeln der Nutzer voraus, der die Medien zielgerichtet nutzt. Er stellt heraus, dass der Nutzer die Medien zur Bedürfnisbefriedigung nutzt. Das wiederum bedeutet, dass die persönlichen Bedürfnisse jedes einzelnen Nutzers zum Schlüssel seiner Mediennutzung werden¹⁸. Von dieser Grundlage ausgehend lassen sich verschiedene Motive für die Mediennutzung erarbeiten.

In ihrer Publikation zur Fernsehnutzung legt die Sozialwissenschaftlerin Teresa Naab den Schwerpunkt auf das Medium Fernsehen, was in Anbetracht des Themas der vorliegenden Arbeit ein passender Fokus ist. Auf Basis des Nutzen-Belohnungs-Ansatzes stellt Naab vier verschiedene Bedürfnisse heraus, die durch die Fernsehnutzung befriedigt werden¹⁹:

1. Das Bedürfnis nach Zeitvertreib. Der Fernseher macht es dem Nutzer leicht, seine freie Zeit zu verbringen. Durch die bewegten Bilder und immer wieder neuen Eindrücke verliert der Nutzer leichter sein Gefühl für Zeit.

¹⁶ Vgl. Schweiger: Theorien zur Mediennutzung, S. 11.

¹⁷ S. Naab, Teresa: Gewohnheiten und Rituale der Fernsehnutzung – Theoretische Konzeption und methodische Perspektiven, In: Reihe Rezeptionsforschung (Band 27), hrsg. von Marco Dohle u.a., Baden-Baden, 2013, S. 50.

¹⁸ Vgl. Ebd., S. 50.

¹⁹ Vgl. Ebd., S. 51.

2. Das Bedürfnis nach Gesellschaft. Auch wenn das Fernsehen eine einseitig ausgerichtete Kommunikation ist, vermittelt das Fernsehen dem Nutzer durch audiovisuelle Kommunikation stärker als andere Medien das Gefühl, nicht alleine zu sein.
3. Das Bedürfnis nach Eskapismus. Der Fernseher macht es dem Nutzer leicht, den Alltag um ihn herum zu vergessen und in eine eigene Welt einzutauchen.
4. Das Bedürfnis nach Bequemlichkeit. Das Medium Fernsehen zu nutzen stellt den Nutzer vor nicht allzu hohe Hürden. Sind Fernseher, Strom und ein Fernsehanschluss vorhanden, genügt ein Knopfdruck, um das Medium zu nutzen.

Diese vier Bedürfnisse setzt Naab gleich mit den Motiven für die Fernsehnutzung gleich: Um die aufgeführten Bedürfnisse zu befriedigen, nutzt der Nutzer das Medium Fernsehen. Naab stellt außerdem noch heraus, dass die Motive für die Fernsehnutzung kontextabhängig sind²⁰: So kann an Werktagen zum Beispiel das Motiv Eskapismus als Antrieb zur Fernsehnutzung dienen, da der Nutzer nach dem anstrengenden Job vor dem Fernseher abschalten und keine anspruchsvollen Denkvorgänge mehr durchführen möchte. Am Wochenende könnte derselbe Nutzer aber vom Motiv Gesellschaft zur Fernsehnutzung getrieben werden und alleine oder mit Freunden die Bundesliga-Spiele verfolgen, um am Montag mit den Arbeitskollegen reden zu können. Laut Naab biete das Fernsehen von allen Medien das breiteste Spektrum an Zuwendungsmotiven²¹. Der Rezipient „darf sich darauf verlassen, dass das Fernsehen ‚schon zu irgendetwas gut sein wird.‘²²“.

Zudem sind die Motive zur Mediennutzung auch von den eingeschliffenen Gewohnheiten des Nutzers beim Fernsehen abhängig. Das zeigt eine Umfrage, die Naab in ihrer Publikation anführt: So vermissten viele Personen nach einem Stromausfall am meisten, dass sie eine Fernsehsendung nicht mehr zur gewohnten Uhrzeit anschauen konnten²³. Außerdem gaben zwischen 10 und 20 Prozent der Befragten an, das Sicherheitsgefühl bei der Fernsehnutzung zu vermissen²⁴. Diesen Umfrageergebnissen versucht Naab auf den Grund zu gehen und führt zwei Begründungen an: Grund für das Vermissen könnte der nun deutlich erhöhte Aufwand zur Informationsgewinnung oder zur Unterhaltung sein, da die gewohnheitsmäßige Fernseh- und Mediennutzung immer mit einem geringen Aufwand verbunden und selbstverständlich sei²⁵. Als Begründung für das fehlende Sicherheitsgefühl gibt Naab die unterbrochene Wiederholung der Handlung der Mediennutzung an. So sei

²⁰ Vgl. Naab: Gewohnheiten und Rituale der Fernsehnutzung, S. 53.

²¹ Vgl. Ebd.

²² S. Ebd.

²³ Vgl. Ebd., S. 98.

²⁴ Vgl. Ebd.

²⁵ Vgl. Ebd.

das allabendliche Schauen der Nachrichten trotz Meldungen zu Terror und Katastrophen ein Faktor zum Wohlfühlen und Fühlen von Sicherheit, weil es die eigene Welt konstant und kontrollierbar erscheinen lasse – nach dem Motto: Wenn die Nachrichten zu schlimm werden, schalte ich den Fernseher einfach aus²⁶.

Zum Schluss nennt Naab noch die Informationssuche als Motiv für die Mediennutzung, die ebenfalls zur Bildung von Gewohnheiten beitrage und aus dem täglichen Tagesablauf nicht mehr wegzudenken sei²⁷.

2.1.2.1 Entstehung von Gewohnheit bei der Mediennutzung

Die Frage nach der Entstehung von Gewohnheit bei der Mediennutzung ist für diese Arbeit interessant, da die Gewohnheit ein wichtiger Bindungsfaktor zwischen Medium und Nutzer ist. Grundsätzlich ist die Fernsehnutzung laut Naab in zwei verschiedene Grundstadien aufzuteilen: Ein aktives und ein passives Stadium²⁸. Beim passiven Stadium läuft der Fernseher neben einer anderen Tätigkeit – der Fernseher ist also ein Begleitmedium. Hier liegt der Fokus des Nutzers auf der anderen Tätigkeit, also der Haupttätigkeit, der er neben dem Fernsehen nachgeht.

Daher ist das aktive Stadium der interessantere Untersuchungsgegenstand bei der Frage, wie Gewohnheiten bei der Fernsehnutzung entstehen. Hier geht Naab auf das Modell des Soziologen Hartmut Esser ein, der den Begriff „Skripte“ bei der Mediennutzung ins Spiel bringt. „Skripte“ sind demnach eine Art mentale Wissensstruktur, in der ein Handeln für eine gewisse Situation festgelegt ist²⁹. So existiere zum Beispiel ein Nachrichtenskript, das wie folgt ablaufen könnte: Der Nutzer verspürt das Bedürfnis nach Information; das ist der Auslöser für das Skript. Es beinhaltet zunächst das Einschalten des Fernsehers und anschließend die Wahl eines Programms, auf dem Nachrichten laufen. Dabei muss der Nutzer abwägen, ob die Kosten für den Aufwand der Suche³⁰ den Nutzen rechtfertigen. Außerdem gehört zum Skript noch das Wissen, dass die Nachrichtensendung 15 Minuten dauert, sowie die Erwartung, danach die gesuchten Informationen bekommen zu haben.

²⁶ Vgl. Naab: Gewohnheiten und Rituale der Fernsehnutzung, S. 99.

²⁷ Vgl. Ebd.

²⁸ Vgl. Ebd., S. 36.

²⁹ Vgl. Esser, Hartmut: Die Definition der Situation, In: Kölner Zeitschrift für Soziologie und Sozialpsychologie (Band 48.1), Köln, 1996, S. 12 f.

³⁰ Die „Kosten“ einer Suche können monetärer (zum Beispiel der Kauf einer Programmzeitschrift), zeitlicher (das Suchen nach einer passenden Sendung) oder kognitiver (Denkleistung, um sich an positiv im Gedächtnis gebliebene Sendungen zu erinnern) Natur sein.

Im Mittelpunkt bei der Entstehung von Gewohnheiten steht beim Skript-Modell vor allem die Suche nach dem geeigneten Programm. So durchläuft jeder Rezipient zunächst eine Scanningphase, in der er das verfügbare Programm durchläuft. Das kann entweder mithilfe einer Programmzeitschrift, dem TV-Guide im Netz oder durch das Durchzappen aller Programme erfolgen³¹. Stößt der Nutzer nun auf einen Inhalt, der sein Interesse weckt, wird ein Prozess in Gang gesetzt, der meist unterbewusst abläuft. Er muss entscheiden: Erfüllt das gefundene Programm meine Erwartungen und Bedürfnisse oder nehme ich weitere „Kosten“ in Kauf und suche weiter? Fällt die unterbewusste Entscheidung zugunsten des erstgenannten Punkts, bleibt der Nutzer beim Programm, tendiert die Entscheidung zu letztgenanntem, sucht er weiter – bis er schließlich bei einer passenden Sendung landet oder den Fernseher ergebnislos abschaltet.

Genau diese Entscheidung ist nun kritisch bei der Entstehung von Gewohnheiten. Zentraler Punkt ist dabei die positive Verstärkung³². Trifft das ausgewählte Programm die Erwartung und das Bedürfnis des Rezipienten, wird er sich bei einer erneuten Suche unterbewusst daran erinnern und wieder bei diesem Programm bleiben. Durch erneute Wiederholung dieses Prozesses entsteht die immer stärker werdende Gewohnheit beim Rezipienten, immer wieder zu diesem Programm zurückzukehren.

Das Festhalten an Gewohnheiten hat für den Nutzer zwei große Vorteile: Durch die Gewohnheit spart er Ressourcen bei der Suche nach dem richtigen Programm, die er für andere Dinge aufwenden kann. Außerdem bringt die Gewohnheit eine große Erwartungssicherheit für den Nutzer mit sich: Schaltet er ein angewöhntes Programm ein, weiß er, was er zu erwarten hat, was ihm ein Gefühl von Sicherheit und Struktur gibt. Dass Gewohnheiten überhaupt entstehen können, führt Naab vor allem die Konstanz zurück: Zum einen ist Konstanz im Fernsehprogramm³³ wichtig, damit der Nutzer zur angewöhnten Zeit seine angewöhnten Sendungen sehen kann³⁴. Zum anderen ist auch Konstanz auf Seiten der Rezipienten wichtig³⁵: Bei einem geregelten Tagesablauf prägen sich Gewohnheiten bei der Fernsehnutzung leichter aus, als bei einem unregelmäßigen Tagesablauf.

³¹ Vgl. Naab: Gewohnheiten und Rituale der Fernsehnutzung, S. 36 f.

³² Vgl. Ebd., S. 40.

³³ Konstanz im Fernsehprogramm wird unter dem Begriff „Stripping“ zusammengefasst. Er beinhaltet, dass gewisse Programmelemente immer zur selben Zeit und am selben Wochentag ausgestrahlt werden. Stripping umfasst aber nicht nur spezifische Sendungen, sondern auch Sendeblocke wie zum Beispiel den Sonntags-Blockbuster auf ProSieben.

³⁴ Vgl. Naab: Gewohnheiten und Rituale der Fernsehnutzung, S. 74.

³⁵ Vgl. Ebd., S. 74.

2.2 Veränderungen in der Mediennutzung

2.2.1 Voraussetzungen für die Mediennutzung in Deutschland

In Deutschland sind die Voraussetzungen innerhalb der Gesellschaft für die Verbreitung von Medieninhalten und die Mediennutzung sehr unterschiedlich ausgeprägt. Unter den Begriff „Voraussetzungen“ fallen in dieser Betrachtung vor allem die infrastrukturellen Möglichkeiten der Menschen, Medien zu nutzen. Das Spektrum reicht hier vom Übertragungsweg bis hin zum Endgerät des Nutzers – egal ob es ein Smartphone, ein Radio oder ein Fernseher ist. Zudem lohnt es sich an dieser Stelle, die verschiedenen Übertragungswege bzw. Endgeräte spezifisch zu betrachten. Aufgrund des recht breiten thematischen Spektrums beschränkt sich diese Untersuchung auf Zeitungen, Radio, Fernsehen und Internet.

2.2.1.1 Zeitungen

Zeitungen nehmen schon seit sehr langer Zeit eine wichtige Rolle im Medienspektrum ein. Daher ist der Markt für Zeitungen auch in Deutschland sehr stark ausgeprägt³⁶. In Abbildung 1 sind die Auflagen und die Anzahl aller Zeitungen aufgeführt, die Anfang 2016 mit einer Gesamtauflage von insgesamt 16,08 Millionen Exemplaren pro Erscheinungstag in Deutschland auf den Markt kamen³⁷.

| Zeitungen | Auflage |
|--|------------------------|
| 329 lokale und regionale Abonnementzeitungen | 12,2 Millionen |
| 7 überregionale Zeitungen | 1,1 Millionen |
| 8 Straßenverkaufszeitungen | 2,8 Millionen |
| 344 Tageszeitungen gesamt | 16,08 Millionen |
| 20 Wochenzeitungen | 1,7 Millionen |
| 7 Sonntagszeitungen | 2,7 Millionen |

Abbildung 1: In Deutschland erschienene Zeitungen 2016

Die Abbildung zeigt, dass das Angebot an Zeitungen im deutschen Markt trotz sinkender Reichweiten immer noch sehr weit gefächert ist: So erscheinen im gesamten Bundesgebiet noch insgesamt 344 Tageszeitungen und 27 Wochen- und Sonntagszeitungen. Von dieser Auflage werden rund 11,3 Millionen Exemplare

³⁶ Laut Bund Deutscher Zeitungsverleger (BDZV) ist der deutsche Zeitungsmarkt der größte Europas und der fünfgrößte der Welt.

³⁷ Vgl. Pasquay, Anja: Die deutschen Zeitungen in Zahlen und Daten 2016, hrsg. vom Bund Deutscher Zeitungsverleger, Berlin, 2016, S. 2.

über Abonnements verkauft, die restlichen Exemplare wandern in einer der rund 113.000 Verkaufsstellen in Deutschland über den Tresen. Das ergibt eine „Händlerdichte“ von 1,4 Einzelhändlern pro 1000 Einwohner³⁸.

Neben der traditionellen gedruckten Zeitung bieten auch immer mehr Verlage ihre Produkte als e-Paper an. Seit 2005 steigen die Verkaufszahlen dieser Distributionsart stetig an, im zweiten Quartal verkauften die Verlage über 780.000 e-Paper ihrer gedruckten Ausgaben³⁹.

Eine weitere wichtige Einnahmequelle der Verlage ist die Webseite der Zeitung. Da dieser Punkt aber unter das Kapitel „Internet“ fällt, soll er hier nicht näher betrachtet werden. Hier sei nur noch erwähnt, dass bisher über 660 verschiedene Zeitungswebsites in Deutschland existieren⁴⁰.

2.2.1.2 Radio

Die ARD/ZDF-Onlinestudie zeigt, dass 2015 in nahezu allen Haushalten auch ein Radio zu finden war: Rund 96 Prozent aller Haushalte in Deutschland verfügen demnach über ein Radio⁴¹. Somit kann das Medium Radio theoretisch einen sehr großen Teil der Bevölkerung erreichen.

Die Verbreitung des Radioprogramms funktioniert über unterschiedliche Wege: Entweder klassisch über die verschiedenen Frequenzen UKW, MW oder KW, über den terrestrischen Empfang von Digitalradio (Digital Audio Broadcasting, kurz DAB), oder über das Internet.

Die Distributoren des Mediums Radio lassen sich grob in zwei verschiedene Gruppen aufteilen: Zum einen in öffentlich-rechtliche Anbieter, die in einzelne Landesrundfunkanstalten gegliedert sind, und zum anderen in private Anbieter. Bei der Verteilung der Sender liegen die privaten deutlich vor den öffentlich-rechtlichen.

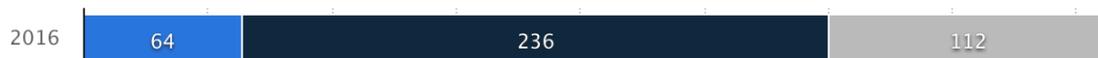


Abbildung 2: Verteilung der Radiosender in Deutschland (hellblau: öffentlich-rechtlich, dunkelblau: privat, grau: Sonstige)

³⁸ Vgl. Pasquay, Anja: Zeitungen in Zahlen und Daten 2016, S. 16.

³⁹ Vgl. Ebd., S. 20. Immer beliebter wird auch die Ausspielung per App, derzeit kann sich der Konsument zwischen knapp 600 verschiedenen Apps der Verlage entscheiden.

⁴⁰ Vgl. Ebd., S. 18.

⁴¹ Vgl. Engel, Bernhard; Breunig, Christian: Massenkommunikation 2015 – Mediennutzung im Intermediavergleich, hrsg. von der ARD Werbung Sales & Services, Frankfurt, 2015, S. 311.

Wie die obige Grafik verdeutlicht, existieren in Deutschland fast vier Mal so viele private Radiosender wie öffentlich-rechtliche. Unter „Sonstige“ fallen unter anderem Anbieter von Internetradios, die fast doppelt so stark vertreten sind wie die öffentlich-rechtlichen Sender.

2.2.1.3 Fernsehen

Die Situation beim Fernsehen ist ähnlich wie beim Radio: Auch hier sind mit 98 Prozent nahezu alle Haushalte in Deutschland mit einem Endgerät ausgestattet, weshalb auch über den Fernseher theoretisch fast jeder Teil der Bevölkerung erreicht werden kann⁴².

Ebenfalls mit dem Radio vergleichbar sind die Übertragungswege des Fernsehens: Das Programm kann entweder über hochfrequente oder terrestrische Verbindungen ausgestrahlt werden. Dazu kommen als Übertragungswege noch das Kabelfernsehen, der Empfang via Satellit oder über das Internet als IPTV.

Auch die Distributoren sind mit denen des Radios zu vergleichen: Auch den Fernsehmarkt teilen sich die öffentlich-rechtlichen und die privaten Sender auf. Von den rund 145 TV-Sendern in Deutschland sind durchschnittlich aber nur 74 Sender in den deutschen Wohnzimmern zu empfangen⁴³.

2.2.1.4 Internet

Die ARD/ZDF-Onlinestudie gibt zwar an, wie viele internetfähige Geräte in den deutschen Haushalten verfügbar sind, allerdings gibt sie keinen Aufschluss darüber, ob alle diese Geräte tatsächlich mit dem Netz verbunden sind. Insgesamt befinden sich in 83 Prozent aller deutschen Haushalte internetfähige Geräte, egal ob es ein Smartphone, ein Tablet, ein Desktop-Rechner, eine Spielekonsole oder ein Smart-TV ist⁴⁴. An dieser Stelle ist aber noch eine andere Zahl interessant: Die Haushalte haben nicht nur ein Gerät im Haus, sondern durchschnittlich 2,7 – wenn Smartphones mit in diese Kennziffer einbezogen werden, in der jüngeren Altersgruppe zwischen 14 und 29 Jahren sogar 3,7⁴⁵. Diese Werte unterstreichen

⁴² Vgl. Engel, Breunig: Massenkommunikation 2015, S. 311.

⁴³ Vgl. ARD-Werbung Sales & Services. n.d. Anzahl der durchschnittlich pro Haushalt empfangbaren TV-Sender in Deutschland in den Jahren 1988 bis 2016 (jeweils 1. Januar). Statista. Zugriff am 17. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/160407/umfrage/frei-empfangbare-tv-sender-in-deutschland-seit-1988/>.

⁴⁴ Vgl. Engel, Breunig: Massenkommunikation 2015, S. 311.

⁴⁵ Vgl. Ebd.

zum einen die enorm wichtige Stellung, die internetfähige Geräte im Alltagsleben mittlerweile eingenommen haben, zum anderen verdeutlichen sie die flexiblen Einsatzmöglichkeiten der Geräte. Denn würde ein Gerät alle Funktionen bieten, wäre die Gesamtanzahl der Geräte in den Haushalten deutlich geringer.

Die Übertragungswege für das Internet müssen auf zwei verschiedenen Ebenen gesehen werden – zum einen die drahtlosen und zum anderen die kabelgebundenen Übertragungswege.

Generell ist bei der Betrachtung des Internets wichtig zu wissen, dass es sich technisch um ein globales Netzwerk aus miteinander kommunizierenden Computern handelt, die unterschiedlich wichtige Knotenpunkte einnehmen. Grundlage jeglicher Kommunikation ist das Internet Protocol und das Transmission Control Protocol. Über diesen Standard werden Datenpakete ausgetauscht und empfangen.

Die Übertragung der Datenpakete erfolgt beim kabelgebundenen Übertragungsweg über Kabelnetze, die aus unterschiedlichen Metallen bestehen. In letzter Zeit etablierten sich Glasfasernetze, die deutlich höhere Übertragungsraten ermöglichen. Die Datendurchsatzrate der Netze spielt dabei eine wesentliche Rolle, welche Medien über das Internet genutzt werden können. So gibt es aktuell in Deutschland immer noch Haushalte, die YouTube-Videos aufgrund der langsamen Netzanbindung nicht abspielen können⁴⁶. Erst vor Kurzem gab Wirtschaftsminister Sigmar Gabriel bekannt, dass der Großteil der deutschen Haushalte bis 2025 ans Gigabit-Glasfasernetz angeschlossen sein soll⁴⁷.

Neben der Übertragung über Kabelnetze lassen sich die oben beschriebenen Datenpakete auch drahtlos über das Mobilfunknetz versenden. Dieser Übertragungsweg wird immer beliebter, im vergangenen Jahr nutzten fast 36 Prozent der Bevölkerung das Internet über das Mobilfunknetz⁴⁸. In diesem Bereich

⁴⁶ Vgl. <http://www.abendblatt.de/region/stormarn/article205522101/Doerfer-in-Stormarn-wollen-das-schnelle-Internet.html>, zuletzt besucht am 17.04.2016 um 17:37 Uhr.

⁴⁷ Vgl. <http://www.heise.de/newsticker/meldung/Gabriel-will-Gigabit-Glasfasernetz-bis-2025-3134766.html>, zuletzt besucht am 17.04.2016 um 17:42 Uhr.

⁴⁸ Vgl. IfD Allensbach. Anzahl der Internetnutzer in Deutschland, die das Internet über das Handy oder Smartphone nutzen, von 2013 bis 2015 (in Millionen). <http://de.statista.com/statistik/daten/studie/170557/umfrage/internutzung-ueber-das-handy-oder-smartphone/> (zugegriffen am 17. April 2016).

besteht allerdings noch Verbesserungspotential: In einer Umfrage gaben 70 Prozent der Befragten an, das Mobilfunknetz für verbesserungswürdig zu halten⁴⁹.

Die Frage der Distributoren im Netz ist ebenfalls auf mehreren Ebenen zu beantworten: Besitzerin der unterirdischen Übertragungskabel und der Mobilfunkfrequenzen in Deutschland ist die Telekom, die aber Nutzungsverträge mit anderen Anbietern vereinbaren kann. Besitzer der angebotenen Medieninhalte sind die einzelnen Anbieter im Netz. Für den Inhalt der schon erwähnten Zeitungswebseiten sind also die einzelnen Verlage verantwortlich, ebenso wie die Radiosender für ihre Internetstreams oder die Fernsehsender für ihre Mediatheken.

2.3 Veränderungen in der Mediennutzung

Im folgenden Abschnitt soll nun die Veränderung der Mediennutzung in Deutschland näher untersucht werden. Dafür bietet es sich an, soweit in der Zeit zurückzugehen, wie es die Aufzeichnungen zur Mediennutzung erlauben, und diese mit den aktuellen Zahlen zur Nutzung zu vergleichen. Die ersten Aufzeichnungen zur Mediennutzung kommen aus dem Jahr 1964. Damals beschäftigte sich eine ARD/ZDF-Studie zum ersten Mal mit dem Mediennutzungsverhalten der Deutschen. Leider wurde 1964 die Mediennutzung noch nicht in dem Umfang abgefragt wie 2015, dennoch gibt es einige Grundpfeiler, die sich miteinander vergleichen lassen. Größter Unterschied im Vergleich der Studien ist natürlich das fehlende Internet, das erstmals in der Studie Massenkommunikation von 2000 eine Rolle spielte.

2.3.1 Mediennutzung 1964

Die erste Studie Massenkommunikation zur Mediennutzung der Deutschen stammt aus dem Jahr 1964 und wurde schon damals von der ARD in Auftrag gegeben⁵⁰. Die erste Studie bestand aus zwei Teilstudien, von denen die erste eine repräsentative Bevölkerungsumfrage zur Nutzung des Angebots an politischen Informationen und die zweite eine Umfrage zu den Einstellungen und Erwartungshaltungen gegenüber verschiedenen Medien war. Zusätzlich beinhaltete die Studie noch eine Inhaltsanalyse der

⁴⁹ Vgl. Goodnity. n.d. Findest du das mobile Datennetz in Deutschland verbesserungswürdig?. Statista. Zugriff am 17. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/533641/umfrage/umfrage-zur-zufriedenheit-mit-dem-mobilen-datennetz-in-deutschland/>.

⁵⁰ Vgl. Breunig, Christian; van Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 505.

politischen Informationsangebote von Fernsehen und Tageszeitung, auf die bei den Folgeuntersuchungen allerdings verzichtet wurde⁵¹.

Ähnlich wie heute befand sich der deutsche Medienmarkt auch 1964 mitten in einer disruptiven Innovationsphase, die bestehende Medienangebote zu verdrängen oder gar zu ersetzen drohte. Diese disruptive Innovation war damals der Fernseher, der zu Beginn der 1960er-Jahre erstmals zu verbraucherfreundlichen Preisen auf den Markt kam und so für große Bevölkerungsteile erschwinglich wurde. Damals fürchteten die großen Verlage einen Verdrängungswettbewerb für ihre eigenen Printprodukte, was einer der Gründe für die Erstellung der Studie Massenkommunikation war⁵².

Die Studie zeigte, dass das durchschnittliche Zeitbudget für die Mediennutzung bei der deutschen Gesamtbevölkerung bei 3:14 Stunden oder umgerechnet bei 194 Minuten lag. Zur Verfügung standen dabei lediglich die drei Medien Zeitung, Fernsehen und Radio⁵³.

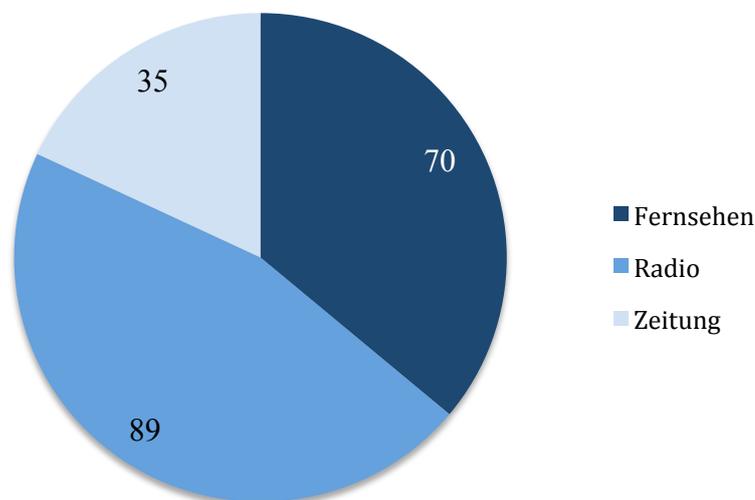


Abbildung 3: Nutzungsdauer der einzelnen Medien 1964 in Minuten

Die Abbildung 3 zeigt, wie viel Zeit die Nutzer für die einzelnen Medien im Jahr 1964 aufgewendet haben. Erstaunlich ist dabei, dass das relativ junge Medium Fernsehen schon 70 Minuten oder rund 36 Prozent des Medienbudgets einnimmt. Deutlich an der Spitze steht das Radio mit rund 46 Prozent Anteil am Medienbudget, abgeschlagen auf dem dritten Platz landet die Zeitung mit einem Anteil von rund 18 Prozent.

Stellt man gegen diese Zahlen die Reichweite der drei Medien, ergibt sich ebenfalls ein interessantes Bild: Trotz des geringsten Anteils am Medienzeitbudget erreicht die Zeitung

⁵¹ Vgl. Breunig; van Eimeren: 50 Jahre Massenkommunikation, S. 505.

⁵² Vgl. Ebd., S. 506. Die Studie sollte damals nachweisen, ob ein Verdrängungswettbewerb zwischen Zeitungen und Zeitschriften auf der einen und dem Fernsehen auf der anderen Seite nachgewiesen werden kann.

⁵³ Vgl. Ebd.

mit 69 Prozent den größten Teil der Bevölkerung⁵⁴. Dicht dahinter folgt das Radio mit 68 Prozent der Gesamtbevölkerung und relativ abgeschlagen landet das Fernsehen mit 47 Prozent auf dem dritten Rang⁵⁵. Grund für die geringe Reichweite des Fernsehers ist die zu diesem Zeitpunkt vergleichsweise kleine Anzahl an Geräten in deutschen Wohnzimmern. Zum Zeitpunkt der Erhebung standen in 55 Prozent der deutschen Wohnzimmer Fernseher, Radios waren 1964 in 95 Prozent aller Haushalte vertreten⁵⁶. Die Zahl der verkauften Auflage der deutschen Zeitungsverlage lag im Jahr 1960 bei insgesamt 15,5 Millionen – 4,2 Millionen Zeitungen wurden im Rahmen eines Abonnements verkauft, 11,3 Millionen im Straßenverkauf⁵⁷. Auf Basis der rund 20 Millionen Privathaushalte im Jahr 1961 ergibt sich, dass zu Beginn der 1960er-Jahre in ungefähr 75 Prozent der deutschen Haushalte jeden Tag eine Zeitung vorhanden war. Damit ordnet sich die Verfügbarkeit einer Zeitung zwischen der des Radios und der des Fernsehers ein.

Zudem wurde in der ersten Studie Massenkommunikation gefragt, wie sehr die Teilnehmer ein bestimmtes Medium in einer simulierten Grenzsituation vermissen würden. 36 Prozent der Befragten gaben an, das Fernsehen sehr stark zu vermissen, 46 Prozent würden das Radio vermissen und 49 Prozent die Tageszeitung⁵⁸. Mit dieser Art der Befragung lässt sich eine Bindung an die verschiedenen Medien ermitteln. Die Untersuchung zeigte, dass aufgrund der noch geringen Verbreitung die Bindung zum Fernsehen noch nicht sehr ausgeprägt ist, jedoch nahezu jeder zweite Befragte Radio und Zeitung in einer Grenzsituation vermissen würde.

Zusammenfassend lässt sich zur Studie Massenkommunikation 1964 sagen, dass sich die Menschen nur an einem sehr kleinen Teil des Tages mit Medien beschäftigten. Das dürfte vor allem dem Umstand geschuldet sein, dass die Geräte zur Mediennutzung (abgesehen von der Zeitung) stationär waren und sie nur an bestimmten Orten genutzt werden konnten. Aus heutiger Sicht erstaunlich ist die hohe Reichweite der Zeitung, die knapp vor dem Radio die größte Reichweite aufweist. Neben der Reichweite bietet auch die Umfrage zum Vermissen von Medien in Grenzsituationen einen guten Anhaltspunkt zur Stellung der Medien in der Gesellschaft: Hier gab fast jeder Zweite an, die Zeitung zu vermissen, was ihr eine wichtige Stellung als Informationsquelle innerhalb der Gesellschaft attestiert –

⁵⁴ Vgl. Breunig; van Eimeren: 50 Jahre Massenkommunikation, S. 510.

⁵⁵ Vgl. Ebd.

⁵⁶ Vgl. Ebd., S. 518.

⁵⁷ Vgl. Pasquay, Anja: Die deutschen Zeitungen in Zahlen und Daten 2012, hrsg. vom Bund Deutscher Zeitungsverleger, Berlin, 2012, S. 9.

⁵⁸ Vgl. Breunig; van Eimeren: 50 Jahre Massenkommunikation, S. 520.

obwohl sie gemäß des Medienzeitbudgets das Medium ist, mit dem die Rezipienten pro Tag am wenigsten Zeit verbrachten.

2.3.2 Mediennutzung 2015

Da für die Mediennutzung 2015 deutlich mehr Daten verfügbar sind, als es sie es noch bei der Erhebung 1964 waren, ist dieses Kapitel in zwei Unterkapitel geteilt. Das erste führt die Daten auf, die auch schon 1964 verfügbar waren, das zweite führt noch einige zusätzliche Kennzahlen mit ein, die ebenfalls für diese Arbeit relevant sind.

2.3.2.1 Mediennutzung 2015 – Standarddaten

Schon deutlich länger als 1964 steckt der globale Medienmarkt zurzeit in einer disruptiven Innovation, deren Auswirkungen sich in den letzten Jahren deutlich zeigten. Einige traditionelle Medienmarken mussten kapitulieren, das prominenteste Beispiel dafür dürfte die Insolvenz des Verlags der „Frankfurter Rundschau“ im Jahr 2012 sein⁵⁹, da er den Sprung vom Print- zum Digital-Verlag nicht rechtzeitig schaffte.

Aus diesem Grund hat die Studie Massenkommunikation auch 2015 noch nicht an Aktualität eingebüßt und kann Hinweise darauf geben, wie die Mediennutzer mit der vom Internet ausgelösten Disruption umgehen.

Im Jahr 2015 werden in der Studie acht verschiedene Medien untersucht: Zu den bereits 1964 vertretenen Medien Fernsehen, Tageszeitung und Radio gesellten sich über die Jahre noch die auditiven Medien (CD/LP/MP3), Bücher, Zeitschriften, Video/DVD-Blu-Ray sowie das Internet⁶⁰. Mit der Anzahl der untersuchten Medien erhöhte sich auch das Medienzeitbudget der Rezipienten deutlich: 2015 wendeten die Deutschen pro Tag 9:26 Stunden, also umgerechnet 566 Minuten für die Mediennutzung auf⁶¹.

⁵⁹ Vgl. <http://www.faz.net/aktuell/wirtschaft/unternehmen/medien-frankfurter-rundschau-meldet-insolvenz-an-11959199.html>, zuletzt besucht am 19.04.2016 um 13:21 Uhr.

⁶⁰ Vgl. Breunig; van Eimeren: Studie Massenkommunikation 2015, S. 510.

⁶¹ Vgl. Ebd., S. 506.

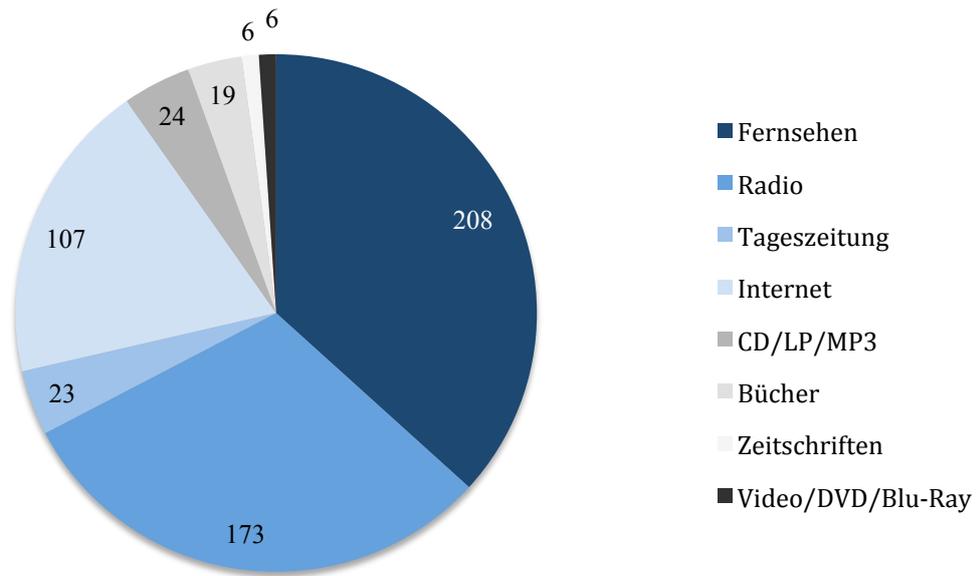


Abbildung 4: Nutzungsdauer der einzelnen Medien 2015 in Minuten

Das Kreisdiagramm in Abbildung 4 zeigt die Verteilung der zur Verfügung stehenden Zeit auf die einzelnen Medien: Fernsehen und Radio nehmen mit 208 und 173

Nutzungsminuten den größten Teil des Medienzeitbudgets ein, auch das Internet hat mit 107 Minuten einen recht großen Anteil. Für Tageszeitungen, Zeitschriften, Bücher, Musik und Videos wenden die Deutschen vergleichsweise wenig Zeit aus ihrem Budget auf.

Auch 2015 wurden die Reichweiten der unterschiedlichen Medien erhoben. Aufgrund der größeren Anzahl dient hier eine Tabelle der besseren Übersicht:

| | | | |
|---------------|----|-------------------|---|
| Fernsehen | 80 | Video/DVD/Blu-Ray | 4 |
| Radio | 74 | | |
| Tageszeitung | 33 | | |
| Internet | 46 | | |
| CD/LP/MP3 | 16 | | |
| Bücher | 18 | | |
| Zeitschriften | 6 | | |

Abbildung 5: Reichweite der einzelnen Medien in Deutschland 2015 in Prozent

In Abbildung 5 wird ersichtlich, dass das Fernsehen mit 80 Prozent die höchste Reichweite aller Medien aufweist. Kurz dahinter folgt das Radio, das Internet liegt – wie schon bei der Nutzungsdauer – auch bei der Reichweite auf dem dritten Platz. Die Tageszeitung liegt noch recht nah am Internet, die übrigen Medien sind in ihrer Reichweite eher begrenzt.

Auch an dieser Stelle ist es interessant zu überprüfen, wie gut die Haushalte in Deutschland mit den untersuchten Medien ausgestattet sind. Auch hier dient eine Tabelle der besseren Übersichtlichkeit:

| | | | |
|-------------------------------|----|--|----|
| Fernseher | 98 | DVD-Player | 64 |
| Radio (inkl. Autoradio) | 96 | Blu-Ray-Player | 27 |
| PC | 83 | Zeitungen (Abonnements + Straßenverkauf) | 46 |
| Internetanschluss (stationär) | 67 | | |
| Internetanschluss (mobil) | 48 | | |
| Smartphone | 61 | | |
| Zeitschriften | 6 | | |

Abbildung 6: Ausstattung der deutschen Haushalte mit verschiedenen Medien 2015 (Ausnahme: Daten zu den Zeitungen sind aus dem Jahr 2011)

Die Übersicht zeigt, dass Fernseher und Radios 2015 in nahezu allen deutschen Haushalten verfügbar sind. Gleiches trifft auf Computer (Desktop und Laptop) zu. Ebenso haben viele Haushalte in Deutschland einen Internetanschluss, ein Smartphone und einen DVD-Player im Haus. Außerdem nutzt jeder zweite Haushalt in Deutschland das Mobiltelefon, um ins Internet zu gehen. Eine gedruckte Zeitung erreicht über den Straßenverkauf oder ein Abonnement ebenfalls fast jeden zweiten Haushalt.

Auch 2015 wurden die Teilnehmer der Studie befragt, welche Medien sie in einer simulierten Grenzsituation am meisten vermissen würden. Hier erklärten 45 Prozent, dass sie das Fernsehen stark vermissen würden, 50 Prozent, dass sie das Radio vermissen würden, 36 Prozent entschieden sich für die Tageszeitung und 40 Prozent für das Internet⁶². Somit ist laut der Studie Massenkommunikation 2015 die Bindung der Deutschen zum Radio tendenziell am größten und zur Tageszeitung am geringsten.

2.3.2.2 Mediennutzung 2015 – Weitere erhobene Daten

Da sich die Studie über die Jahre immer weiterentwickelte, blieb es nicht bei den recht rudimentären Daten, die noch 1964 erhoben wurden. In der aktuellen Studie

⁶² Vgl. Breunig; van Eimeren: Studie Massenkommunikation 2015, S. 520.

Massenkommunikation werden noch sehr viel mehr Daten abgefragt, bei denen es sich lohnt, etwas genauer hinzuschauen.

Dazu gehört vor allem die Betrachtung der Mediennutzungsdauern im Tagesverlauf. Für 2015 ergibt sich folgendes Bild:

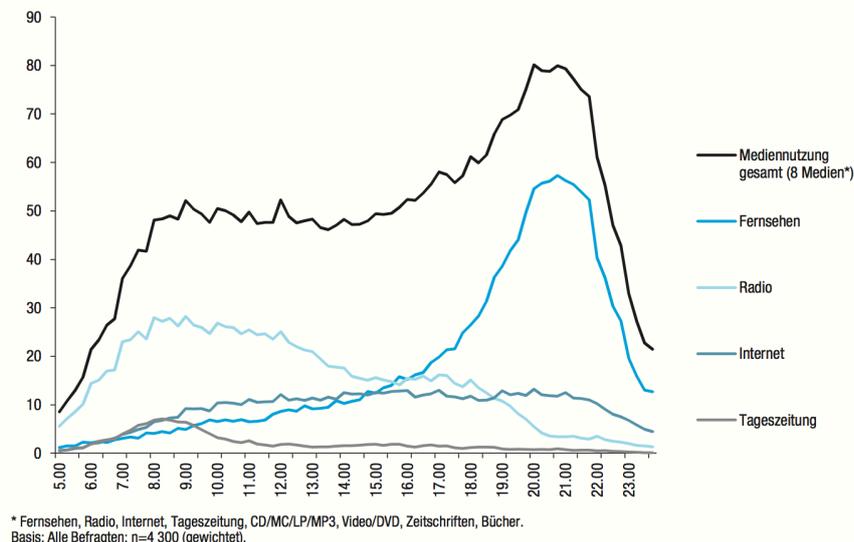


Abbildung 7: Mediennutzung der deutschen Gesamtbevölkerung 2015 im Tagesverlauf

Die Abbildung 7 zeigt sehr deutlich, dass die Mediennutzung der Deutschen zum Abend hin immer weiter zunimmt. Von fünf bis sieben Uhr morgens steigt die Nutzungskurve rasant an, was vor allem auf die intensive Radionutzung am Morgen zurückzuführen ist. Hier hören rund 25 Prozent der deutschen Bevölkerung Radio. Auch die Kurve der Tageszeitung hat in dieser Zeitspanne einen Peak, der allerdings deutlich kleiner ausfällt als beim Radio. Das Internet, das als viertes großes Medium in dem Diagramm berücksichtigt wird, wird zu Tagesbeginn ähnlich stark genutzt wie die Tageszeitung. Über den weiteren Tagesverlauf, also wenn der Großteil der Bevölkerung auf der Arbeit, in der Schule oder in der Universität sitzt, nehmen nahezu alle Medien in ihrer Nutzung ab. Das Radio hält sein Niveau bis zwölf Uhr annähernd, danach sinkt die Radionutzung stetig ab. Die Zeitung hat ihren Nutzungsspeak schon um acht Uhr morgens erreicht und nähert sich danach immer näher der x-Achse an. Die Internetnutzung erreicht gegen elf Uhr einen Nutzungshöhepunkt und kann dieses Niveau über den Tag verteilt halten. Die Nutzung des Fernsehens nimmt über den Tag immer mehr zu.

Ab 16 Uhr steigt die Nutzung des Fernsehers rasant an und erreicht gegen 21 Uhr ihren Höhepunkt. Hier sehen über 55 Prozent der deutschen Bevölkerung fern. Nach dem Nutzungsspeak sinkt die Nutzung des Fernsehers aber genauso rapide ab, wie sie zuvor angestiegen ist. Gegen halb elf Uhr abends erreicht die Nutzungskurve des Fernsehens ihren Tiefpunkt und stabilisiert sich. Mit der Erhöhung der Fernsehnutzung geht eine

Abnahme bei der Nutzung der anderen Medien einher: Ab 20 Uhr liegt die Nutzung des Radios bei fünf Prozent der Bevölkerung, die Tageszeitung vom Morgen ist um diese Zeit nicht mehr relevant. Lediglich die Internetnutzung bleibt bis 22 Uhr nahezu konstant bei über zehn Prozent der Bevölkerung bestehen und sinkt erst am späten Abend leicht ab. Die Gesamtnutzungskurve der Medien steigt am Abend ähnlich an wie die Fernseh-Kurve. Auf ihrem Höhepunkt gegen 21 Uhr nutzen 80 Prozent der deutschen Bevölkerung ein beliebiges Medium.

Die Auswertung des Diagramms zeigt: Am Morgen werden vor allem Radio und Zeitung genutzt, am Abend eher der Fernseher. Erstaunlich ist allerdings die konstante Nutzung des Internets über den gesamten Tagesverlauf. Hier zeichnet sich kein Peak wie beim Radio oder beim Fernsehen ab, sondern die Kurve verläuft fast gleichmäßig über den gesamten Tag hinweg. Die Autoren der Studie Massenkommunikation führen diesen Umstand auf die breite Verfügbarkeit des Internets zurück: Es ist sowohl am Arbeitsplatz, als auch auf dem Weg zur und von der Arbeit verfügbar, ebenso wie zuhause bei den Rezipienten⁶³. Außerdem sei der multifunktionale Charakter des Internets ein wichtiger Faktor für die konstante Nutzung über den gesamten Tagesverlauf: Weil im Internet nicht nur Text, sondern auch multimediale Inhalte abgerufen werden könnten und es darüber hinaus Plattform für Kommunikation, Transaktion und andere Aktivitäten sei, sei eine gleichmäßige Verteilung über den ganzen Tag plausibel⁶⁴.

Des Weiteren untersucht die Studie Massenkommunikation genauer, wie das Internet von der Bevölkerung genutzt wird. Die Ergebnisse sind in folgender Abbildung veranschaulicht:

⁶³ Vgl. Engel; Breunig: Massenkommunikation 2015, S. 314.

⁶⁴ Vgl. Ebd.

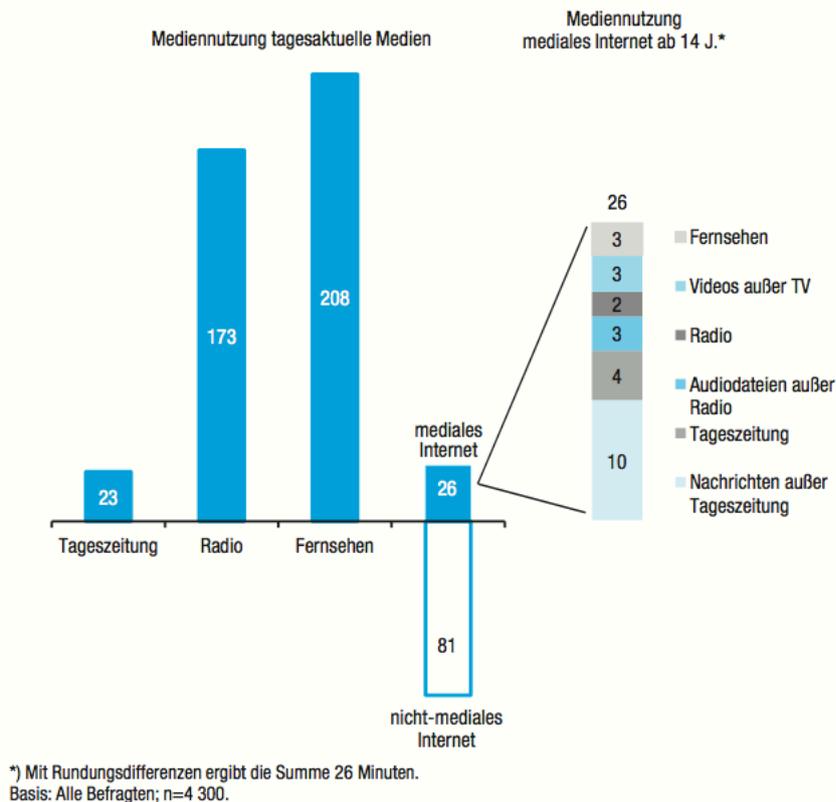


Abbildung 8: Tagesaktuelle Nutzung der Medien in Minuten 2015; Aufschlüsselung des Internets in mediale und nicht-mediale Nutzung

Abbildung 8 zeigt, dass 2015 nur jeder vierte Nutzer tagesaktuell das mediale Internet nutzt. Darunter fallen Interaktionen wie das Abrufen von Mediatheken, das Abrufen von Nachrichten-Websites und Audio-Dateien sowie Livestreams des Fernsehprogramms. Ein Großteil der Internetnutzung fällt unter den nicht-medialen Aspekt. Dazu gehört zum Beispiel die Benutzung von Google zur Informationsgewinnung, die Nutzung von Social Media oder auch das Einkaufen auf E-Commerce-Seiten. Die insgesamt 26 Minuten, die zur Nutzung des medialen Internets aufgewendet werden, schlüsselt das Diagramm noch genauer auf: Den Großteil der Nutzungszeit verwenden die Rezipienten darauf, Nachrichten abseits von den Webseiten der Tageszeitungen zu konsumieren. Darunter fallen zum Beispiel Portale wie T-Online oder Web.de. Nahezu ausgeglichen verteilen sich die übrigen 16 Minuten auf die übrigen medialen Nutzungsarten: Nachrichtenrecherche auf Websites, die zu Tageszeitungen oder Magazinen gehören, dauert durchschnittlich vier Minuten, Fernseh-, Video und Audiostreams, die abseits vom Radio laufen, belegen jeweils drei Minuten und das Radio selbst wird im Netz durchschnittlich zwei Minuten pro Tag genutzt.

Das von der Studie gezeichnete Bild untermalt auch die Statistik der meistabgerufenen Websites für das Jahr 2015 in Deutschland. So liegt Google.de hier auf dem ersten Platz,

dahinter folgen Plattformen wie Facebook und YouTube sowie E-Commerce-Angebote wie Amazon oder Ebay. Erst dann folgen mediale Angebote - mit Web.de, T-Online und Gmx.de in diesem Fall tatsächlich zuerst die Nachrichtenwebsites abseits der Tageszeitungen und Magazine. Knapp dahinter liegen mit Bild.de und Spiegel.de die ersten Websites klassischer Medienunternehmen⁶⁵.

Speziell zur Internetnutzung bietet die Studie Massenkommunikation noch einige weitere interessante Aufstellungen, die für die Leitfrage dieser Arbeit durchaus erwähnenswert sind. Dazu gehört zum Beispiel die Übersicht, über welche Geräte ins Netz gegangen wird⁶⁶. Sie zeigt, dass an einem Tag von den 107 durchschnittlich gemessenen Minuten 49 Minuten über einen stationären PC und 26 Minuten über einen Laptop im Internet verbracht werden. Der dritte große Posten ist mit 25 Minuten die Internetnutzung über das Smartphone. Von den übrig gebliebenen Minuten der Internetnutzung nimmt der Tablet-PC noch einen größeren Teil ein, die Nutzung per Smart-TV oder Spielekonsole fällt nicht ins Gewicht.

Auch zur Bindung ans Medium Internet stellt die Studie Massenkommunikation eine Untersuchung zur Verfügung. Hier wird wieder gefragt, welche Funktion des Internets die Befragten in einem simulierten Grenzfall am meisten vermissen würden. Sie liefert folgendes Ergebnis:

| Tab. 13 Bindung an das Internet 2015 | |
|---|----|
| Personen ab 14. J., in % | |
| Aus dem Internet würden sehr stark/stark vermisst: ¹⁾ | |
| Fernsehsendungen, z.B. Nachrichten oder Serien, oder Teile von Sendungen live im Internet ansehen | 25 |
| Fernsehsendungen, z.B. Nachrichten oder Serien oder Teile von Sendungen zeitversetzt im Internet ansehen | 28 |
| andere Videos im Internet ansehen | 28 |
| Radioprogramme live im Internet anhören | 19 |
| Radiosendungen wie z.B. Nachrichten oder Beiträge zeitversetzt im Internet anhören | 17 |
| andere Musikangebote über das Internet anhören | 32 |
| Nachrichten auf Homepages von Tageszeitungen lesen | 42 |
| Nachrichten auf Homepages von Nachrichtenmagazinen lesen | 33 |
| News bei Suchmaschinen oder Internet-Providern lesen | 40 |
| über das Internet kommunizieren (E-Mails versenden, posten, chatten, Instant Messaging, wie z.B. WhatsApp, Facebook oder andere Netzwerke nutzen) | 77 |
| im Internet spielen | 21 |
| einfach so im Internet surfen | 56 |

1) Basis: Internetnutzer ab 14 Jahren, die das Internet sehr stark/stark vermissen würden (n=1 737; gewichtet).

Abbildung 9: Bindung an das Internet 2015 in Prozent

Die Untersuchung bestätigt die bereits zuvor untersuchte Statistik zur generellen Bindung an die verschiedenen Medien: Die Bindung der Menschen ans Internet ist noch nicht sehr

⁶⁵ Vgl. <http://www.alexa.com/topsites/countries/DE>, zuletzt besucht am 21.04.2016 um 10:25 Uhr.

⁶⁶ Vgl. Breunig; van Eimeren: Massenkommunikation 2015, S. 517.

fest gewachsen. So liegt ein Großteil der ermittelten Bindungswerte weit unter 40 Prozent. Lediglich die Kommunikationsfunktion vermissen vier von fünf Befragten. Ebenso würden die Befragten die Informationsgewinnung und das einfache Surfen im Internet stark vermissen. Die für die vorliegende Masterarbeit wichtigste Nutzungsart, Videodateien im Netz abzurufen, würde durchschnittlich nur jeder vierte Befragte stark vermissen. Trotzdem scheint die Bindung an diese Nutzung des Internets stärker zu sein, als die zum Radio- und Audiostreaming.

2.3.2.3 Nutzung der Mediatheken und Online-Videoangebote

Die Nutzung der Mediatheken sowie der Online-Videodienste ist in der Studie Massenkommunikation nicht als einzelnes Segment vertreten. Grund dafür scheint die relativ geringe Zeit zu sein, die Internetnutzer für diese Portale aufwenden. Wie aus Abbildung 8 ersichtlich wird, entfallen lediglich sechs Minuten der Internetnutzungszeit von insgesamt 107 Minuten auf das Livestreamen des Fernsehprogramms oder die Nutzung von Videoportalen wie YouTube oder Clipfish. Somit macht diese Art der Internetnutzung gerade etwas über fünf Prozent der Gesamtnutzung aus.

Vor dem Hintergrund der Leitfrage dieser Arbeit lohnt sich eine genauere Betrachtung dieses kleinen Segments trotzdem, um zu überprüfen, wie intensiv die Plattformen genutzt werden⁶⁷: So zeigt eine Auswertung der ARD/ZDF-Onlinestudie, dass 82 Prozent der Gesamtbevölkerung und sogar 98 Prozent der 14-29 Jährigen Videoangebote im Internet „zumindest selten“ nutzen. Aus der tiefergehenden Aufschlüsselung der Nutzung wird allerdings deutlich, dass es große Unterschiede zwischen den Portalen gibt: So werden Videoportale wie YouTube oder Clipfish von 61 Prozent der Gesamtbevölkerung und 86 Prozent der 14-29 Jährigen „zumindest selten“ genutzt, Video-Streamingdienste wie Netflix oder Maxdome aber nur von 15 und 28 Prozent. Demzufolge schaut 2015 nicht einmal jeder Dritte Jugendliche oder junge Erwachsene Inhalte eines Streamingdienstes. Hier stehen die Mediatheken der Fernsehsender etwas besser da: 36 Prozent der Gesamtbevölkerung und 45 Prozent der jungen Bevölkerung schaute „zumindest selten“ in den Mediatheken vorbei. Ähnliche Werte erreichte das klassische Fernsehprogramm im Internet: Live oder zeitversetzt schauten 43 Prozent der gesamten Bevölkerung und 57 Prozent der 14-29 Jährigen fern.

Neben der allgemeinen Nutzung stellt die Onlinestudie auch die regelmäßige Nutzung heraus. Gefragt wurde hier, wer täglich Video-Content aus einer Online-Quelle abrufen.

⁶⁷ <http://www.ard-zdf-onlinestudie.de/index.php?id=522>, zuletzt besucht am 21.04.2016 um 15:26 Uhr.

Insgesamt nutzen 26 Prozent der Gesamtbevölkerung und 54 Prozent der 14-29 Jährigen täglich Video-Content im Netz. Auch in der täglichen Nutzung liegen Videoportale wie YouTube auf dem ersten Platz: 14 Prozent der Gesamtbevölkerung und 32 Prozent der Jüngeren schauen dort täglich vorbei. Live- oder zeitversetztes Fernsehen sehen 3 Prozent und 6 Prozent, während Mediatheken der Sender sowie die Video-Streamingdienste täglich von 2 Prozent der Gesamtbevölkerung und 3 Prozent der 14-29 Jährigen genutzt werden. Das Ergebnis der Onlinestudie unterstreicht auch eine Bitkom-Umfrage zur Art der Sendungen, die online abgerufen werden: 60 Prozent der Befragten schauen Videos auf YouTube und anderen Portalen, 54 Prozent rufen Beiträge und Sendungen aus den Mediatheken der Fernsehsender ab, 36 Prozent streamen das laufende Fernsehprogramm, während 22 Prozent der Befragten Serien und Spielfilme von Streamingdiensten nutzen⁶⁸. Wirklich konkret wird eine Stichprobe zur Mediatheknutzung, die genau aufzeigt, welche Sendungen die meisten Abrufe in einer Woche erreichten⁶⁹: Hier liegt ein ProSieben-Clip zu Miley Cyrus mit über 300.000 Abrufen auf dem ersten Platz, gefolgt von einer Episode einer ARD-Soap. Knapp dahinter liegt der Tatort vom Sonntagabend sowie die Satiresendung „Heute Show“ vom ZDF. Darauf folgt eine Überraschung: Die Plätze fünf bis neun belegen alle Folgen der vergangenen Woche der ARD-Soap „Sturm der Liebe“, die zwischen 180.000- und 130.000-mal abgerufen wurden. Hier scheint es eine konstant große Gruppe an interessierten Fans zu geben, welche die Sendungen regelmäßig online abrufen.

Zum Schluss dieses Unterkapitels soll noch untersucht werden, welche Geräte zum Abruf von Online-Video-Content genutzt werden. Hier hilft eine Untersuchung von Februar 2016 des Online-TV-Anbieters Zattoo⁷⁰: Sie zeigt, dass ein sehr großer Teil der Befragten ihren Laptop oder PC nutzen, um Video-Content aus dem Netz abzurufen. Rund drei Viertel

⁶⁸ Vgl. Bitkom. n.d. Welche Art von Sendungen oder Videos schauen Sie sich als Stream im Internet an?. Statista. Zugriff am 24. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/316608/umfrage/art-der-nutzung-von-videostreaming-in-deutschland/>.

⁶⁹ Vgl. AGF. n.d. Ranking der Sendungen mit den meisten Abrufen bei den Videostreaming-Angeboten der TV-Sender in Deutschland in der Kalenderwoche 47 im Jahr 2015 . Statista. Zugriff am 21. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/475379/umfrage/sendungen-nach-abrufen-bei-den-videostreaming-angeboten-der-tv-sender/>.

⁷⁰ Vgl. Zattoo. n.d. Auf welchen Geräten nutzen Sie Internet-TV Services?. Statista. Zugriff am 24. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/528875/umfrage/nutzung-von-internet-fernsehen-nach-endgeraeten-in-deutschland/>.

aller Befragten greifen auf diese Geräte zurück. Recht weit abgeschlagen sind die Smart-TVs. Sie werden von jedem dritten Befragten genutzt, um online Videos abzurufen. Knapp dahinter folgen mit 24 Prozent und 20 Prozent Tablets und Smartphones. Mit unter zehn Prozent landen Streamingboxen und Spielekonsolen auf dem letzten Platz bei der Online-Videonutzung. Das Ergebnis der Umfrage zeigt recht deutlich, dass die mobilen Geräte bei der Online-TV-Nutzung immer noch eine recht kleine Rolle neben den stationären PCs sowie Laptops spielen.

2.3.3 Veränderungen in der Mediennutzung

Gewählt wurde für den Vergleich der Mediennutzung ein recht großer Zeitraum von mehr als fünfzig Jahren. Trotz des recht großen Zeitunterschieds wird allerdings deutlich, dass die Situation 1964 mit der von heute in einigen Punkten vergleichbar ist. Wie bereits erwähnt, drohte schon damals ein neues Medium die Medienlandschaft umzukrempeln, was den etablierten Medien Sorgen bereitete. Auch bei Erstellung der Studie von 2015 steckt die Medienlandschaft aufgrund der Digitalisierung mitten in einem Transformationsprozess, der sich allerdings in deutlich mehr Bereiche des öffentlichen Lebens auswirkt, als damals das Fernsehen.

Stark von diesem Umschwung betroffen ist natürlich auch die Mediennutzung jedes Einzelnen. Das zeigt sich vor allem im Medienzeitbudget der Menschen:

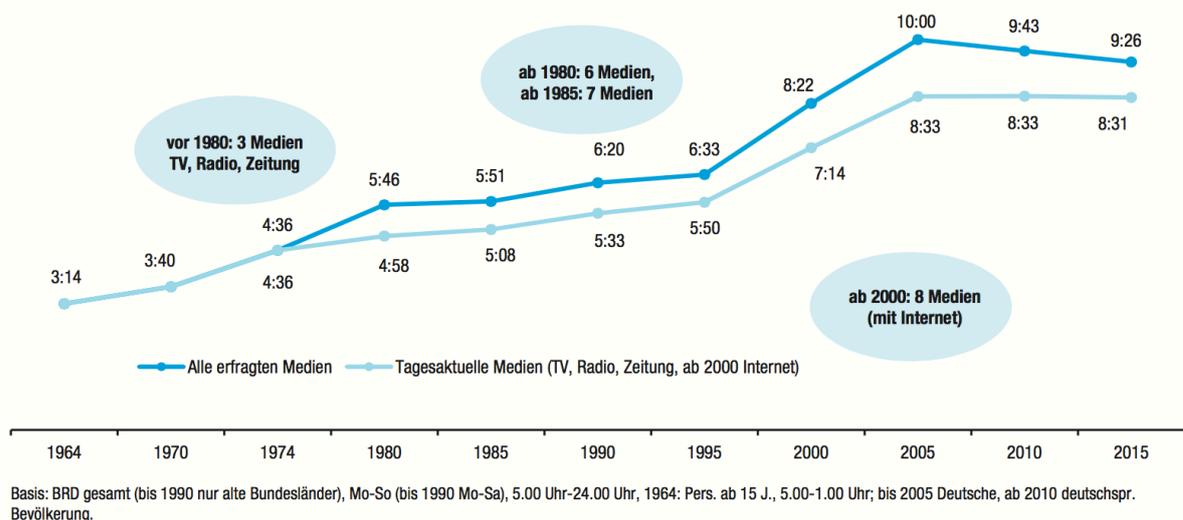


Abbildung 10: Medienzeitbudget eines Deutschen durchschnittlich pro Tag in Stunden

Im Vergleich zu 1964 hat sich das Budget fast verdreifacht: Nutzten die Menschen damals noch knapp drei Stunden drei verschiedene Medien, liegen sie heute schon bei neuneinhalb Stunden und 8 unterschiedlichen Medien pro Tag. Außerdem zeigt die Abbildung deutlich, dass sich das Budget über die Jahre immer weiter erhöhte bis es 2005 bei genau zehn

Stunden lag. Seitdem hat es sich wieder um rund eine halbe Stunde reduziert. Zudem veranschaulicht die Abbildung, dass das Verhältnis zwischen tagesaktuellen Medien und allen erfragten Medien über die Jahre annähernd gleich geblieben ist. Dazu ist allerdings noch wichtig anzumerken, dass vor 1974 ausschließlich tagesaktuelle Medien im Zeitbudget der Studie erfasst wurden, es aber auch schon zuvor nicht-tagesaktuelle Medien in Form von Büchern oder Schallplatten gab.

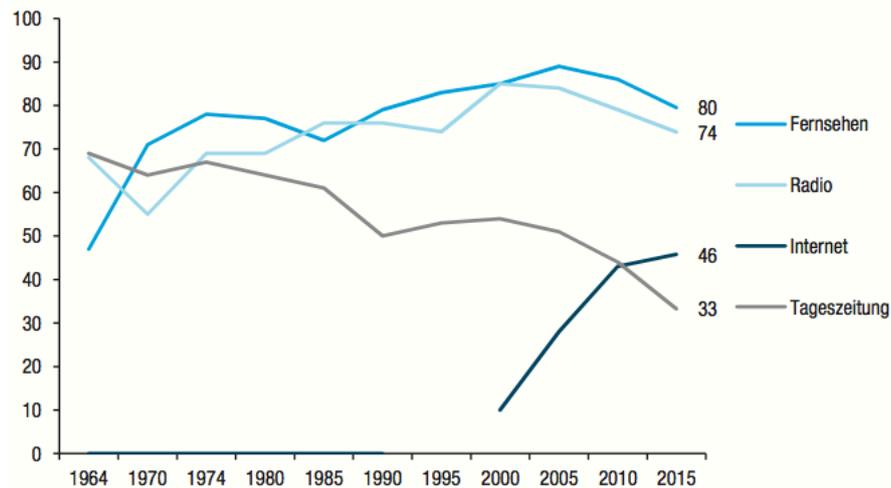


Abbildung 11: Tages-Reichweiten unterschiedlicher Medien in Prozent

Betrachtet man die Reichweitenentwicklung im Zeitverlauf, wird vor allem eines deutlich: Die Annahme, dass eine disruptive Innovation die bestehenden, traditionellen Medien verdrängt, trifft nicht immer zu. Zwar nimmt die Reichweite des Fernsehens von 1964 bis 1974 rapide zu, dafür büßen die anderen Medien nur kleine Teile ihrer Reichweite ein. Das Radio verliert von der Studie 1964 zur Studie 1970 rund 15 Prozent seiner Reichweite, erholt sich aber bereits in der Studie 1974 vom Reichweitenrückgang. Die Zeitung verliert im Vergleich zum Radio anfänglich nur kleine Teile ihrer Reichweite, doch sie kann sich vom steten Rückgang nicht mehr erholen. Die Abbildung zeigt somit, dass die Innovation Fernsehen 1964 die bestehenden Medien eher ergänzte, aber sie nicht in ihrer Existenz oder ihrer Reichweite bedrohte. Grund dafür dürfte der Umstand sein, dass mit dem Fernsehen ein weiteres klassisches Medium im Sinne der Sender-Empfänger-Theorie die Medienlandschaft ergänzte – es war also eher eine technische und keine „inhaltliche“ Disruption.

Hier ist ein deutlicher Unterschied zum Internet zu finden, das erstmals 2000 in der Studie Massenkommunikation auftauchte. In Abbildung 11 ist sehr deutlich zu erkennen, dass mit dem rasanten Aufstieg des Internets alle übrigen Medien an Tagesreichweite verloren. Das Fernsehen verlor rund zehn Prozent an Reichweite, das Radio knapp sechs Prozent.

Besonders große Anteile verlor die Tageszeitung: Seit der Erhebung 2000 fiel sie von rund 55 Prozent Tagesreichweite auf 33 Prozent.

Gerade bei der Internetnutzung gibt es einen großen Unterschied zwischen den Altersklassen. Das zeigt sich in den Ergebnissen der ARD/ZDF-Onlinestudie zur Videonutzung im Netz sehr deutlich. Während von der Gesamtbevölkerung nur jeder Vierte täglich Video-Content im Netz abrufen, ist es bei den 14-29-Jährigen schon jeder zweite. Große Unterschiede finden sich an dieser Stelle auch beim Nutzungsverhalten der Online-Plattformen. Auch hier besetzt die jüngere Altersgruppe immer deutlich größere Anteile als die Bevölkerung insgesamt. Vor diesem Hintergrund dürfte es sehr spannend werden, die Entwicklungen der nächsten Studien zu beobachten, denn trotz der wenigen bisher verfügbaren Datenpunkte lässt sich bereits ein Trend erkennen und eine Hypothese formulieren: Mit dem Erstarken und der weiteren Verbreitung des Internets nehmen die Tagesreichweiten der klassischen Medien immer weiter ab. Das Internet ergänzt also das klassische Medium nicht nur um einen weiteren Ausspielkanal, sondern scheint die klassischen Medien teilweise abzulösen. Die Studie Massenkommunikation bietet dafür folgende Erklärung: Im Internet können „Individual- und Massenkommunikation methodisch und inhaltlich immer weniger trennscharf voneinander unterschieden werden“⁷¹.

Außerdem offenbaren die zahlreichen Untersuchungen zur Mediennutzung ein weiteres Phänomen: Die so genannte Medienkonvergenz. Unter diesem Begriff versteht man die immer weiter fortschreitende Verschmelzung diverser Medien. Die Medienkonvergenz wird durch ein Problem begünstigt: „Während die Medienangebote für die Konsumenten zunehmend attraktiver werden, bleibt die Zeit für die Nutzung begrenzt.“⁷² Dieses Problem zieht nun zwei Effekte nach sich: Entweder konsumieren die Rezipienten weniger Medien, in die sie dann mehr Zeit investieren (Substitutionseffekt) oder sie erhöhen die Nutzung, indem sie mehrere Medien parallel nutzen (Parallelisierungseffekt)⁷³. Letzterer Effekt begünstigt die Medienkonvergenz. Klassisches Beispiel ist hier das Radio: Schon 1935 gingen zwei Drittel der Radiohörer einer Nebentätigkeit nach - wie zum Beispiel dem Zeitunglesen⁷⁴. Die Verbreitung des Internets begünstigt diese Entwicklung, da die Möglichkeiten der Parallelnutzung von einem beliebigen klassischen Medium und dem

⁷¹ S. Breunig; van Eimeren: Studie Massenkommunikation 2015, S. 506.

⁷² S. Goldhammer, Klaus u.a.: Social TV – Aktuelle Nutzung, Prognosen, Konsequenzen, In: Schriftenreihe Medienforschung der Landesanstalt für Medien Nordrhein-Westfalen (Band 76), hrsg. von Jürgen Brautmeier, Düsseldorf, 2015, S. 47.

⁷³ Vgl. Ebd.

⁷⁴ Vgl. Ebd., S. 44.

Internet schier unbegrenzt sind. Ein wichtiges Stichwort ist zum Beispiel der Second Screen, der parallel zum Fernsehen zusätzliche Informationen auf einem zweiten Display anzeigt. Dank dieser neuartigen Technologien steigt die parallele Nutzung immer weiter an: 2000 wurden 16 Minuten lang Medien parallel genutzt, 2010 waren es bereits 37 Minuten – Tendenz weiter steigend⁷⁵. Auch in Zukunft dürfte zunehmende Medienkonvergenz also ein wichtiges Thema sein.

Speziell bei der Videonutzung im Netz lassen sich über den vergleichsweise kleinen Zeitraum ebenfalls erste Veränderungen feststellen. Sie werden in Abbildung 12 recht deutlich dargestellt:

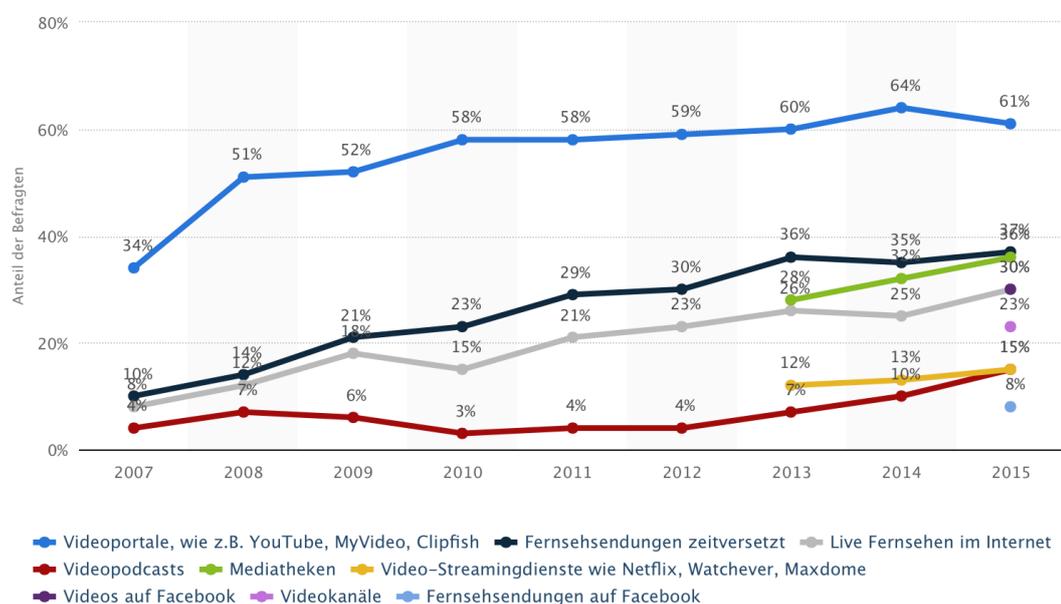


Abbildung 12: Nutzung von Online-Videoanwendungen in den Jahren 2007 bis 2015

Die Abbildung zeigt, dass die Bereitschaft innerhalb der Befragten, Videocontent im Netz abzurufen, über die Jahre immer weiter angestiegen ist. Nahezu alle Anwendungen verbuchten seit 2007 konstante Nutzungsanstiege. Ein großer Teil der Befragten nutzt bereits Videoportale wie YouTube, doch auch die übrigen Anwendungen stiegen in der Häufigkeit der Nutzung immer weiter: So entwickelte sich die Nutzerbasis für das zeitversetzte Fernsehen und die Mediatheken der Fernsehsender rasant. Auch den Livestream des Fernsehens nutzten immer größere Anteile der Befragten. Relativ gering scheint das Wachstum bei den Videostreamingdiensten wie Watchever oder Netflix zu sein. Doch in diesem Bereich ist zu erwarten, dass die Zahlen noch deutlich ansteigen. Das zeigt eine Betrachtung der Visits auf dem deutschen Netflix-Portal: Laut

⁷⁵ Vgl. Goldhammer: Social TV, S. 45.

SimilarWeb lagen die monatlichen Visits auf Netflix 2015 noch bei rund 15 Millionen, 2016 allerdings schon bei über 30 Millionen – ein Wachstum von rund 100 Prozent⁷⁶. Außerdem zeigen die Untersuchungen noch eine weitere Tendenz: Die Mediennutzung der jüngeren und älteren Konsumenten scheint sich immer weiter auseinander zu entwickeln. Das zeigt sich zum einen in der Konzeption der Studie Massenkommunikation: Sie unterteilt nahezu alle Betrachtungen in eine jüngere Nutzergruppe der 14 bis 29-Jährigen und in eine Gesamtbetrachtung der Bevölkerung. Zum anderen wird der Nutzungsunterschied sehr deutlich, wenn man die tagesaktuellen Reichweiten verschiedener Medien für beide Gruppen gegenüberstellt:

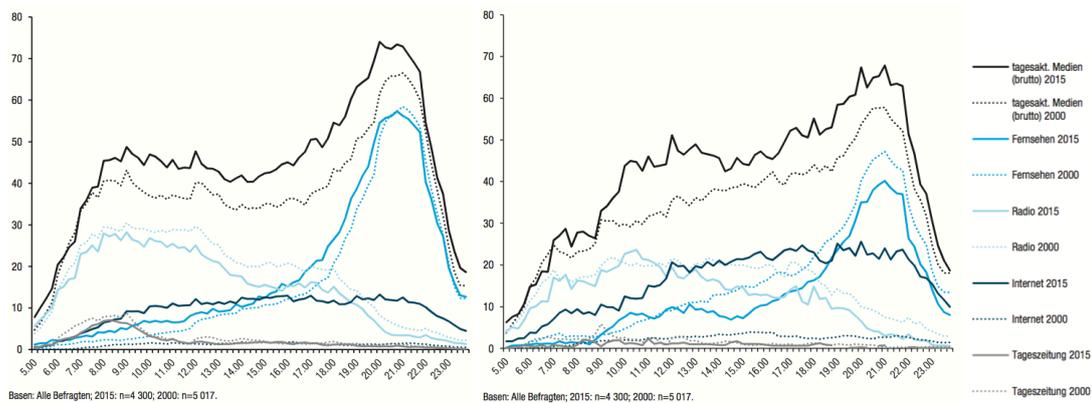


Abbildung 13: Tagesaktuelle Reichweite verschiedener Medien 2000 und 2015 nach Altersgruppen, links die Gesamtbevölkerung, rechts die 14-29-Jährigen

Die beiden Diagramme zeigen recht deutlich, wie unterschiedlich die verschiedenen Altersgruppen⁷⁷ erreicht werden. Bei den 14 bis 29-Jährigen ist der Tiefpunkt in der Medienreichweite zum Mittag deutlich weniger ausgeprägt als in der Gesamtbevölkerung. Außerdem verlaufen auch die Reichweitenlinien beim Diagramm der Gesamtbevölkerung deutlich konstanter als beim Diagramm der jüngeren Bevölkerung. Daraus lässt sich schließen, dass die Mediennutzung dort weniger konstant ist als bei der Gesamtbevölkerung, die Jüngeren also etwas sprunghafter bei der Wahl der Medien sind. Ebenso zeigt sich, dass die jüngere Bevölkerung das Internet deutlich intensiver nutzt als im Vergleich die Gesamtbevölkerung. Bemerkenswert ist die minimale Differenz zwischen

⁷⁶ Vgl. <http://meedia.de/2016/04/27/erstmal-mehr-visits-als-kinox-to-der-spannende-siegeszug-von-netflix-in-deutschland/>, zuletzt besucht am 29.04.2016 um 10:03 Uhr. Die Daten von SimilarWeb sind nur eine ungefähre Schätzung, genaue Abbildungen der realen Visits liefert der Dienst natürlich nicht. Grund dafür ist vor allem, dass SimilarWeb keine Abrufe über die Netflix-App erfassen kann. Trotzdem ermöglicht der Dienst zuverlässig das Ablesen von Wachstumstendenzen.

⁷⁷ Hier lässt sich von Altersgruppen sprechen, da das Durchschnittsalter der Gesamtbevölkerung laut Bundesinstitut für Bevölkerungsforschung bei rund 44 Jahren liegt – Tendenz steigend.

Internet- und Fernsehreichweite bei den 14-29-Jährigen am Abend. Bei der Gesamtbevölkerung ist diese Differenz deutlich intensiver ausgeprägt. Den hohen Stellenwert des Internets für die 14-29-Jährigen unterstreicht auch die Frage nach der Medienbindung dieser Altersgruppe: 65 Prozent gaben hier an, das Internet in einer simulierten Grenzsituation stark zu vermissen, die Bindungen zu Fernsehen (36 Prozent), Radio (36 Prozent) und Tageszeitung (16 Prozent) waren 2015 deutlich weniger ausgeprägt⁷⁸. Bei der Befragung aller Altersgruppen gaben nur 40 Prozent an, das Internet in einer simulierten Grenzsituation zu vermissen – hier lag das Radio mit 50 Prozent Anteil bei den Befragten vorn.⁷⁹.

2.4 Zwischenfazit

Die Untersuchung zur Mediennutzung der Deutschen und der Veränderung dieser Mediennutzung über die Jahre förderte einige interessante Erkenntnisse zutage, die für die Leitfrage dieser Arbeit relevant sind:

- Mit dem Internet steht den etablierten Medien nicht nur ein neuer Ausspielkanal für deren Content zur Verfügung, vielmehr scheint das Internet die etablierte Medien langsam abzulösen. Das zeigte sich vor allem an der täglichen Reichweite der Zeitung, die seit dem Aufstieg des Internets deutlich abnahm. Auch die Reichweiten des Fernsehens und des Radios sind betroffen, allerdings nicht in dem Ausmaß wie die der Tageszeitungen. Bezogen auf die Leitfrage dieser Arbeit ist diese Erkenntnis elementar wichtig, bildet diese Verschiebung ins Internet sowie eine hohe Reichweite des Contents im Netz – speziell des Videocontents im Netz – überhaupt erst die Grundlage sowie die Notwendigkeit für den Einsatz eines Recommender Systems.
- Der neue Ausspielkanal „Internet“ begünstigt noch ein weiteres Phänomen: Die Medienkonvergenz. Unter diesem Begriff versteht man die immer weiter fortschreitende Annäherung von Medien- und Medieninhalten. Das Internet stellt den etablierten Medien immer einfacher zu bedienende Technologien zur Verfügung, die es ihnen ermöglicht, ihre Inhalte immer weiter zu verbinden. So ist es durch Einbindung von Videos oder Audios problemlos möglich, aus einer reinen Text-Reportage eine Multimedia-Dokumentation zu erstellen, die für den Nutzer intuitiv zu bedienen ist. Auch diese Erkenntnis ist wichtig im Bezug auf die Leitfrage dieser Arbeit: Ein Recommender System kann hier einen deutlichen

⁷⁸ Vgl. Breunig; van Eimeren: Studie Massenkommunikation 2015, S. 521.

⁷⁹ Vgl. Ebd., S. 520.

Mehrwert für den Konsumenten bieten, indem es ihm den Content liefert, der ihn interessiert. So kann das Recommender System dazu beitragen, die User Experience eines Medienangebots deutlich attraktiver zu machen.

- Gerade bei den jüngeren Menschen in Deutschland geht die Mediennutzung immer mehr vom getakteten auf den enttakteten Content über. Statt jeden Abend um 20 Uhr die Tagesschau vor dem Fernseher zu verfolgen, nutzen immer mehr junge Nutzer die Mediatheken, um die Sendung dann anzusehen, wenn es ihr Zeitbudget zulässt. Immer wichtiger werden auch die Videostreaminganbieter wie Netflix sowie Videoportale wie YouTube, die nicht nur tagesaktuelle Inhalte anbieten, sondern auch unterhaltenden Content in Form von Serien, Video-Blogs (Vlogs) oder Blockbustern. Auch hier zeigt sich der Trend, diese Inhalte nicht zur fest vorgegebenen, getakteten Zeit über den Fernseher abzurufen, sondern zur frei wählbaren, enttakteten Zeit über den PC, das Smartphone oder die Streamingbox. Diese Erkenntnis ist ebenfalls für die Leitfrage dieser Arbeit interessant, da sie für die Konzeption eines Recommender Systems - hier speziell für den Punkt „Kontext“ - höchst relevant ist. Nur wenn das Recommender System genügend Informationen zum zeitlichen Mediennutzungsverhalten der Konsumenten zur Verfügung hat, kann es in den passenden Situationen die richtigen Inhalte empfehlen.

Wie im Zitat zu Beginn des Kapitels bereits erwähnt, stellt das veränderte Mediennutzungsverhalten der Rezipienten auch die öffentlich-rechtlichen Anbieter vor Probleme. Das Kapitel zeigt, dass ein Recommender System in jedem Fall eine der neuen „Funktionalitäten“ sein kann, die es den Anstalten in Zukunft ermöglicht, weiterhin erfolgreich „am Markt bestehen zu können.“⁸⁰ Außerdem zeigt sich in diesem Kapitel das Phänomen der Disruption innerhalb der Medien, was in Kapitel 4.6 näher beschrieben wird.

3. Die Grundlagen eines Recommender Systems

Das folgende Kapitel soll einen Einblick in die Welt der Recommender Systems vermitteln. Dazu soll zunächst untersucht werden, was diese Systeme überhaupt sind und wozu sie genutzt werden. Darauf folgt eine Beschreibung des Aufbaus eines Recommender Systems, in dem die unterschiedlichen Filter-Methoden aufgeführt werden

⁸⁰ Vgl. Telemedienkonzepte des Norddeutschen Rundfunks, Angebotsbeschreibung für tagesschau.de und eins-extra.de (überarbeitete Fassung – Stand 25. Juni 2010), S. 41.

sollen, die derzeit für diese Systeme existieren. Außerdem soll in diesem Kapitel untersucht werden, wie eine Empfehlung in solch einem System überhaupt zustande kommt und ob Netzwerkeffekte bei einem Recommender System zum Tragen kommen. Nicht zuletzt soll dieses Kapitel auch einen Überblick über die Herausforderungen geben, die es bei der Einführung eines Recommender Systems geben kann.

3.1 Definition eines Recommender Systems

Eine einheitliche, allgemein akzeptierte Definition eines Recommender Systems existiert bis heute nicht. Das liegt vor allem an den unterschiedlichen Anforderungen, die an ein Recommender System gestellt werden⁸¹. Dennoch lassen sich aus den verschiedenen Definitionen einige zentrale Elemente ableiten:

- *Recommender Systems sind Software-Tools und Techniken, die dem Nutzer einer Dienstleistung, einer Webseite oder eines Portals Items⁸² dieses Dienstes empfehlen⁸³.*
- *Recommender Systems können sowohl für personalisierte, als auch für nicht-personalisierte Empfehlungen eingesetzt werden.⁸⁴*
- *Recommender Systems nutzen die Ratings⁸⁵ einer Community von Usern, um Individuen in dieser Community besser dabei helfen zu können, in einer sehr großen Masse an Content die für sie relevanten Inhalte zu finden.⁸⁶*
- *Recommender Systems sind auch „Decision Support Systems“, die für den Nutzer große Angebote an Items durchsuchen, sortieren, klassifizieren, filtern und zuletzt die Ergebnisse mit dem Nutzer teilen.⁸⁷*

⁸¹ So unterscheiden sich die Zielvorgaben für ein Recommender System, das im Bereich E-Commerce zum Einsatz kommt, von Systemen, die auf Videoportalen zur Empfehlung genutzt werden. Vgl. hierzu: Ahrens, Sophie Charlotte: Recommender Systems – Relevance in the Consumer Purchasing Process, München, 2010, S. 9 ff. sowie Aggarwal, Charu: Recommender Systems – The Textbook, New York, 2016, S. 3 ff.

⁸² Der Begriff „Item“ ist in der Literatur gängig und beschreibt alle Objekte, die von einem Recommender System empfohlen werden, egal ob es sich um eine Produkt auf Amazon, ein Artikel auf einer News-Seite oder ein Video auf einem Videoportal handelt. Dieser Begriff wird wegen einer passenden deutschen Übersetzung in dieser Arbeit übernommen.

⁸³ Vgl. Ricci: Recommender Systems Handbook, S. 1.

⁸⁴ Vgl. Ebd., S. 2.

⁸⁵ Wie schon die „Items“ gibt es auch für den Begriff „Rating“ keine adäquate deutsche Übersetzung, weshalb auch dieser Begriff in dieser Arbeit verwendet wird. Unter Ratings versteht man die Bewertungen der Nutzer für ein einzelnes Item. In Kapitel 3.4 sollen die Ratings genauer beschrieben werden.

⁸⁶ Vgl. Herlocker, Jonathan [u.a.]: Evaluating Collaborative Filtering Systems, In: ACM Transactions on Information Systems, New York, 2004, S. 5.

- *Recommender Systems basieren auf Algorithmen, von denen die meisten in der Lage sind aus dem Konsumverhalten und anderen Nutzer- und Nutzungsdaten zu lernen und die Empfehlungen laufend zu verbessern.*⁸⁸
- *Recommender Systems können sowohl passiv als auch aktiv sein. Passive Systeme unterstützen den Nutzer während des Kaufprozesses, zum Beispiel wie bei Amazon's „Kunden, die dieses Produkt kauften, kauften auch“-Anzeige. Die passiven Recommender Systems sind ein organischer Teil der Webseite. Aktive Systeme kommen auf den Nutzer mit Empfehlungen zu, zum Beispiel mit einer Übersicht über interessante Angebote per E-Mail. Die aktiven Recommender Systems gehen über die Webseite hinaus.*⁸⁹

3.2 Zielsetzung und Zweck der Recommender Systems

Die Zielsetzung für ein Recommender System lässt sich für alle bestehenden Systeme verallgemeinern. Sie besteht aus mehreren Punkten⁹⁰: 1. *Relevanz*. Das Recommender System soll nur Items anzeigen, die auch für den Nutzer relevant sind. Dies ist laut Aggerwal die wichtigste Zielsetzung für ein Recommender System. 2. *Neuartigkeit*: Das Recommender System soll den Nutzer auf Items aufmerksam machen, die er noch nicht kennt. Immer wieder neue Vorschläge aus bereits bekannten Items zu generieren stört die Nutzer. 3. *Zufallsfunde/“Serendipity“*: Im besten Fall schlägt das Recommender System dem Nutzer nicht nur neue Items vor, sondern überrascht den Nutzer positiv mit Vorschlägen, die er nicht erwartet hätte. 4. *Vielfalt*: Die Empfehlungen des Recommender Systems beschränken sich nicht nur auf die Items eines Typs, sondern enthalten Items mehrerer Typen, die bestenfalls alle für den Nutzer relevant sind. So wird sichergestellt, dass der Nutzer immer etwas Neues sieht, aber trotzdem immer ein relevantes Item für ihn dabei ist. Das stärkt auch das Vertrauen des Nutzers in das Empfehlungssystem. Diese vier Ziele sollten bei der Konzeption eines Recommender Systems berücksichtigt werden, um dem Nutzer nach der Einführung des Systems qualitativ hochwertige Empfehlungen machen zu können.

Wie weiter oben bereits erwähnt, sind die Anforderungen an ein Recommender Systems nicht bei allen Webseiten oder Portalen gleich, sondern orientieren sich an der Ausrichtung ebendieser. Die verschiedenen Anforderungen lassen sich dennoch in zwei Gruppen

⁸⁷ Vgl. Ahrens: Recommender Systems, S. 9.

⁸⁸ Vgl. Ebd.

⁸⁹ Vgl. Ebd., S. 27 f.

⁹⁰ Vgl. Aggerwal: Recommender Systems, S. 3 ff.

aufteilen: Zum einen muss das Recommender System für den Anbieter der Webseite bestimmte Zwecke erfüllen, zum anderen hat auch der Nutzer der Webseite Erwartungen an solch ein System. Ahrens hat diese Anforderungen gebündelt⁹¹:

| Anforderungen der Anbieter | Anforderungen der Nutzer |
|--|---|
| Erhöhen der Absatzzahlen, Cross-Selling von Items, Beeinflussen der Produktwahl der Nutzers, Erhöhen der Aufmerksamkeit für Items, Unterstützen neuer Nutzer mit generellen Empfehlungen, Werbe- und Marktrecherche, Reduzieren der Service-Kosten beim Nutzer-Support, Effiziente Informations-Produktion, Erhöhen der Nutzerloyalität, Erhöhung der Glaubwürdigkeit der Webseite, Reduzieren der Informations-Asymmetrie | Verbessern der Entscheidungsqualität, Anzeige von neuen oder bisher nicht wahrgenommenen Items, Erhalten von zusätzlichen „Meinungen“ zu den Items, Reduzieren der Masse an Informationen, Erleichtern der Entscheidungsfindung, Reduzieren des Suchaufwands, Anzeigen von Produkt-Alternativen, Reduzieren des Entscheidungsaufwands, Reduzieren von Unsicherheit beim Kauf, Erhöhen der Transparenz beim Kauf |

Abbildung 14: Anforderungen der Anbieter und Nutzer an ein Recommender System

Inwieweit ein Recommender System sowohl die gestellten Anforderungen und die Zielsetzungen erfüllt, kann durch eine Evaluation herausgefunden werden. Dieser Punkt soll im Kapitel 3.4 etwas genauer beschrieben werden.

Sehr treffend vereint Jeff Bezos, der CEO von Amazon, sowohl die Zielsetzung als auch die Anforderungen der Anbieter und Nutzer an ein Recommender System: „If I have three million customers on the web, I should have three million stores on the web.“⁹²

3.3 Implementierung eines Recommender Systems in eine Webseite

Den exakten Aufbau eines Recommender Systems zu beschreiben gestaltet sich schwierig, da es sehr viele unterschiedliche Typen von Recommender Systemen gibt – wie die folgenden Kapitel zeigen werden. Dieser Umstand macht es nicht möglich, ein allgemeingültiges Muster für den Aufbau eines Empfehlungssystems zu zeichnen.

Vielmehr kann ein Recommender System als Baukasten verstanden werden, der sich aus mehreren Werkzeugen zusammensetzt und auf Teile zurückgreift, die teilweise exklusiv für das System in die Website integriert wurden und teilweise schon vor der Einführung des Systems vorhanden sind.

⁹¹ Vgl. Ahrens: Recommender Systems, S. 12 ff.

⁹² S. Ebd., S. 30.

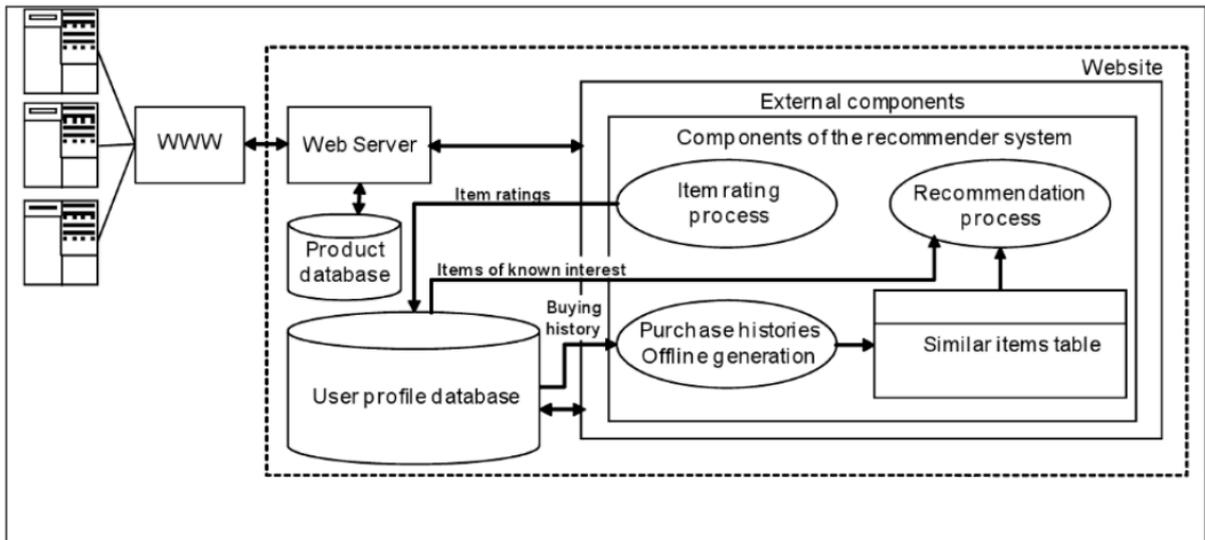


Abbildung 15: Implementierung eines Recommender Systems in eine Webseite

Abbildung 15 macht deutlich, wie ein Recommender System in eine Webseite implementiert ist und welche Komponenten aufeinander zugreifen. Zentrale Sammelstelle der User-Anfragen auf der Webseite ist der Web Server, der die Anfragen weiterverarbeitet. Über ihn bekommen Nutzer Zugriff auf die Produktdatenbank, in der relevante Daten zu den Items und die Items selbst gespeichert werden. Über den Web Server hat auch das angeschlossene Recommender System Zugriff auf die Produktdatenbank. Je nach den Arbeitsweisen der Empfehlungs-Algorithmen, die im folgenden Unterkapitel genauer vorgestellt werden sollen, ist diese Verbindung zur Produktdatenbank elementar wichtig zur Empfehlungsgenerierung. Daneben existiert noch eine User Profile-Datenbank. In dieser Datenbank werden die weiter oben bereits erwähnten Logfiles zu jedem einzelnen angemeldeten User der Webseite gespeichert. Sie enthält dementsprechend nicht nur Mailadressen und Kontextinformationen wie Alter oder Geschlecht des Users, sondern auch seine Kaufhistorie und seine Ratings für die Items des Anbieters. Der Prozess der Rating-Erstellung ist in der Abbildung als alleinstehende Komponente des Recommender Systems abgebildet. Die im Laufe dieses Prozesses erstellten Nutzer-Ratings werden in der User Profile-Datenbank zwischengespeichert und kommen – bei Bedarf – im späteren Empfehlungsprozess wieder zum Einsatz. Über die Verkaufshistorien der einzelnen Nutzer werden Ähnlichkeitstabellen erstellt, die im Kapitel 3.4 näher erläutert werden sollen. Über diese Tabellen lassen sich Ähnlichkeiten zwischen verschiedenen Items oder Usern errechnen, die im direkt angeschlossenen Empfehlungsprozess zusammen mit den Ratings schließlich zu einer Empfehlung führen. Somit ist der Aufbau eines Recommender Systems am ehesten mit Komponenten zu beschreiben, die zum einen aus Prozessen und zum anderen aus mathematischen sowie

algorithmischen Modellen bestehen. Hier gibt es aber keine Vorgaben dazu, welche Prozesse und Modelle Teil eines Recommender Systems sein *müssen*; so bildet zum Beispiel die Abbildung 15 ein vergleichsweise rudimentäres Recommender System mit wenigen Komponenten ab⁹³. Eine zusätzliche Komponente dazu könnte zum Beispiel ein Feedback-Kanal sein, der es dem User erlaubt, auf die angezeigten Empfehlungen zu reagieren⁹⁴.

3.4 Filter-Methoden für ein Recommender System

Ein zentrales Element jedes Recommender Systems ist die Filter-Methode, die aus den vorhandenen Daten die Empfehlungen für jeden einzelnen Nutzer eines Portals generiert. Dabei kommen verschiedene Methoden zum Einsatz, die in diesem Unterkapitel näher vorgestellt werden sollen.

Welche Methode auf welchem Portal eingesetzt wird, hängt auf der einen Seite von den Anforderungen an das Recommender System ab, weil jede Filter-Methode individuelle Vor- und Nachteile bietet, und zum anderen von den Voraussetzungen, die das Recommender System auf der Webseite vorfindet. Hier spielt zum Beispiel der Umfang und die Qualität der zur Generierung einer Empfehlung verfügbaren Daten eine wichtige Rolle, da auch auf diesem Feld die verschiedenen Filter-Methoden unterschiedliche Vor- und Nachteile haben.

Bei der Überlegung, welche Filter-Methode die richtige für das Recommender System der Webseite ist, darf nicht außer Acht gelassen werden, dass der Trend immer mehr in die Richtung hybrider Systeme geht. Diese vereinen unterschiedliche Filter-Methoden in einem Recommender System, um die Nachteile der einen Methode mit den Vorteilen der anderen auszugleichen. Auf diese Weise lassen sich die generierten Empfehlungen stetig verbessern.

Zentraler Bestandteil fast aller Filter-Methoden sind die Ratings der User, die auf fünf unterschiedlichen Wegen erhoben werden können⁹⁵:

⁹³ Das Recommender System von Amazon besteht zum Beispiel aus 18 verschiedenen Komponenten. Vgl. *Hansen, Hans-Robert [u.a.]*: Empfehlungssysteme zur Verkaufsberatung im Internet State-of-the-Art und Konsumentenakzeptanz, In: *Wirtschaftsinformatik (Band 49)*, Wiesbaden, 2007, S. 52.

⁹⁴ Vgl. Ahrens: *Recommender Systems*, S. 10.

⁹⁵ Vgl. Aggarwal: *Recommender Systems*, S. 31 f.

1. „Continuous Ratings“: Diese Ratings werden über eine Skala gesammelt, die sich über eine beliebige Größe erstrecken kann. Prominent ist zum Beispiel das „Jester Joke Rating“, das über eine Skala von -10 bis 10 ermittelt wird.
2. Intervall-basierte Ratings: Ähnlich wie die „Continuous Ratings“ werden auch diese Ratings über eine Skala erstellt. Klassisch sind hier 7-Punkt- oder 5-Punkt-Skalen, letztere werden zum Beispiel auf Amazon oder in der NDR Mediathek eingesetzt.
3. „Ordinal Ratings“: Diese Ratings sind ähnlich wie die beiden vorhergehenden, werden aber nicht mit numerischen Werten gefüllt. Stattdessen enthalten sie einordnende Bezeichnungen wie zum Beispiel „Gefällt mir sehr gut“, „Gefällt mir gut“, „Gefällt mir weniger gut“ oder „Gefällt mir nicht“.
4. Binäre Ratings: Diese Ratings können entweder positiv oder negativ sein, eine genauere Einordnung erlauben sie nicht. Ein Beispiel für binäre Ratings sind die „Mag ich“/„Mag ich nicht“-Buttons auf YouTube.
5. Unäre Ratings: Die unären Ratings geben lediglich positive Interaktionen des Nutzers auf der Webseite sein. Das kann zum Beispiel der Aufruf oder der Kauf eines Items auf der Webseite sein. Vor der Einführung der Reactions auf Facebook war zum Beispiel auch der „Gefällt mir“-Button ein unäres Rating.

Neben den unterschiedlichen Wegen, wie Nutzer-Ratings erhoben werden können, können sie noch in implizite und explizite Ratings unterschieden werden. Implizit sind in den meisten Fällen unäre Ratings, die der Nutzer unabsichtlich bei der Nutzung der Seite abgibt. Wie weiter oben beschrieben fällt darunter zum Beispiel der Aufruf einer Produktseite oder der Kauf eines Items⁹⁶. Die übrigen oben aufgeführten Ratings sind explizit: Damit sie zustande kommen, muss der Nutzer selbst aktiv werden und ein Rating für ein Item abgeben⁹⁷. Implizite Ratings gibt es in deutlich größerem Umfang als explizite.

Die folgenden Unterkapitel sollen die unterschiedlichen Methoden lediglich beschreiben, während das folgende Kapitel 3.4 genauer beschreibt, wie die Empfehlungen mit den Methoden zustande kommen.

3.4.1 Die Collaborative Filtering-Methode

Bei der Methode des kollaborativen Filterns werden die unterschiedlichen Ratings der Community für die Items genutzt, um dem einzelnen Nutzer auf dieser Basis Items

⁹⁶ Vgl. Aggarwal: Recommender Systems, S. 32.

⁹⁷ Vgl. Ebd.

empfehlen zu können⁹⁸. Das kollaborative Filtern kann auf zwei unterschiedlichen Wegen umgesetzt werden: Entweder über Neighborhood-based Collaborative Filtering oder über Model-based Collaborative Filtering. Zunächst soll die erste Filtermethode näher erläutert werden.

Das **Neighborhood-based Collaborative Filtering** basiert auf den ältesten Algorithmen, die für das kollaborative Filtern entwickelt wurden⁹⁹. Sie basieren auf der Annahme, dass User mit vergleichbaren Interessen Items annähernd gleich bewerten. So lassen sich auf Basis der historischen Ratings der User-Community Vorhersagen für die Ratings jedes einzelnen Nutzers treffen. Auf diese Weise lassen sich gewissermaßen seine Interessen vorhersagen und auf dieser Basis Items empfehlen.

Bei dieser Methode gibt es zwei unterschiedliche Ansätze, die Nachbarschaften zwischen Ratings zu bestimmen. Der erste Ansatz stellt die User in den Mittelpunkt und wird demnach User-based genannt, während der zweite die Items in den Mittelpunkt stellt und deshalb Item-based genannt wird¹⁰⁰. Beim User-based-Ansatz werden die Ratings der User genutzt, um Ähnlichkeiten zwischen ihnen zu bestimmen und auf dieser Basis Empfehlungen auszusprechen. Beim Item-based-Ansatz werden die Ratings des Users für Items genutzt, um Ähnlichkeiten zwischen den verschiedenen Items zu bestimmen und so dem User Items vorzuschlagen, die er aufgrund seiner Rating-Vergangenheit bereits ähnlich gut bewertet hat. Der Unterschied zwischen den beiden Ansätzen liegt folglich darin, dass beim User-based-Ansatz die Ratings der User mit ähnlichen Interessen zur Empfehlung eingesetzt werden, während beim Item-based-Ansatz die *eigenen* Ratings des Users zur Empfehlung genutzt werden.

Im Vergleich bringt der Item-based-Ansatz mehr für den Nutzer relevante Empfehlungen, weil diesem Ansatz die eigenen Ratings des Users zugrunde liegen¹⁰¹. Beim User-based-Ansatz werden zwar auch die Ratings anderer User zur Empfehlungserstellung genutzt, trotzdem können die Interessen der Community und die des einzelnen Nutzers aber voneinander abweichen, weshalb die Empfehlungen dieses Ansatzes meist weniger relevant für den User sind¹⁰². Großer Nachteil des Item-based-Ansatzes ist allerdings die fehlende Vielfalt der Empfehlungen, die ebenfalls aus der Erstellungsweise der Empfehlungen resultiert: Weil die Ratings der für den Nutzer relevanten Items die

⁹⁸ Vgl. Ricci: Recommender Systems Handbook, S. 12.

⁹⁹ Vgl. Aggarwal: Recommender Systems, S. 29.

¹⁰⁰ Vgl. Ebd.

¹⁰¹ Vgl. Ebd., S. 42.

¹⁰² Vgl. Ebd.

Grundlage für die Berechnung bilden, sind die Empfehlungen meist sehr vorhersehbar und die weiter oben beschriebenen Zufallsfunde sehr selten. Auf diesem Feld schlägt sich der User-based-Ansatz deutlich besser¹⁰³.

Wie akkurat die Ergebnisse der beiden Ansätze des Neighborhood-based Collaborative Filtering sind, hängt stark davon ab, wie groß das Datenset ist, auf das der Algorithmus zugreifen kann. Erst bei wirklich großen Sets bringt diese Methode gute Ergebnisse. Soll zum Beispiel das Rating des Users John für das Item Terminator vorhergesagt werden und hat noch kein ähnlicher User wie John das Item bewertet, ist die Vorhersage des Recommender Systems nicht sehr genau. Mit der Voraussetzung des großen Datensets kommt ein weiterer Nachteil dieser Methode ins Spiel: Weil sehr viele Daten gebraucht und verarbeitet werden, ist die Performance bei dieser Methode oftmals ein Problem. Deshalb sollte bei dieser Methode darauf geachtet werden, dass sie auf einem System eingesetzt wird, das über genügend Speicherkapazität und Rechenpower verfügt. Dennoch bietet diese Methode auch einige große Vorteile: So ist solch ein Recommender System vergleichsweise einfach in ein Umfeld zu implementieren. Außerdem ist es bei dieser Methode recht einfach herauszufinden, wie ein vorhergesagtes Rating für ein Item oder einen User zustande gekommen ist.

Das **Model-based Collaborative Filtering** ist die deutlich komplexere Filter-Methode und basiert auf unterschiedlichen Data Mining- sowie Machine Learning-Modellen. Dazu gehören zum Beispiel Decision Trees, Neural Networks, Regel-basierte Methoden oder auch Regressionsmodelle¹⁰⁴. Diese Methode bietet Lösungen für die Probleme, die durch das Neighborhood-based Collaborative Filtering nicht gelöst werden können. Dazu gehört zum Beispiel das Fehlen von unzähligen Ratings von Usern für Items, also die Mängel am Datenset, die bei der zuerst beschriebenen Methode zu wenig relevanten Empfehlungsergebnissen führen. Durch den sinnvollen Einsatz der weiter oben beschriebenen Modellierungen können die fehlenden Ratings beim Model-based Collaborative Filtering mit deutlich besseren Vorhersagen ausgefüllt werden. Die effektivsten Vorhersage-Modelle sind derzeit die Latent Factor Models¹⁰⁵.

¹⁰³ Vgl. Aggarwal: Recommender Systems, S. 42.

¹⁰⁴ Vgl. Klahold, André: Empfehlungssysteme – Grundlagen, Konzepte, Lösungen, Wiesbaden, 2009, S. 65.

¹⁰⁵ Vgl. Aggarwal: Recommender Systems, S. 134.

3.4.2 Die Methodik des Content-based Recommender Systems

Während die Collaborative Filtering-Methoden die Korrelationen zwischen den Ratings der Usern zur Empfehlungserstellung nutzen, setzen die Content-based Recommender Systems auf Ähnlichkeiten zwischen den Attributen von Items, um Empfehlungen für den Nutzer geben zu können¹⁰⁶. Auch hier spielen die historischen Ratings des Nutzers wieder eine zentrale Rolle, da sie Items mit ähnlichen Attributen zugeordnet werden und letzten Endes wieder ein Indikator dafür sind, wie relevant ein Item für den Nutzer ist.

Die Content-based Recommender Systems stützen sich daher auf zwei verschiedene Datenquellen¹⁰⁷: 1. Die Attribute eines Items. Diese können zum Beispiel aus der Produktbeschreibung eines Items auf Amazon oder ähnlichen Portalen extrahiert werden. Die Attribute bestehen in den meisten Fällen aus Text-Keywords. 2. Die Logfiles des Users. Die Logfiles enthalten wie weiter oben beschrieben spezifische Nutzungs- und Nutzerdaten, die sowohl implizit als auch explizit sein können. Bei den Content-based Recommender Systems stehen also wie beim Item-based-Ansatz nicht die Ratings der Community im Mittelpunkt der Empfehlungserstellung, sondern die des einzelnen Users. Damit bietet diese Methode einige Vor- und Nachteile. So ist sie in manchen Kaltstart-Szenarien, die weiter unten in Kapitel 3.6 beschrieben werden sollen, sehr effektiv. So ist zum einen genug über Interessen der Nutzer bekannt, zum anderen lassen sich neue Items über die Attribute schon kurz nach Veröffentlichung sinnvoll ähnlichen Items zuordnen und somit an Nutzer empfehlen¹⁰⁸. Das Kaltstart-Problem für neue Nutzer löst diese Methode allerdings nicht.

Ein weiterer Nachteil resultiert wie beim Item-based-Ansatz aus der Methode selbst: Weil über die Attribute Ähnlichkeiten zwischen den Items errechnet werden, werden bei dieser Methode in den meisten Fällen offensichtliche Empfehlungen gegeben, die den Nutzer selten überraschen¹⁰⁹.

Eine wichtige Voraussetzung für den Einsatz der Content-based Recommender Systems ist eine ausreichend große und vor allem gut gepflegte Produktdatenbank, damit den Items möglichst viele und bestenfalls konsistente Attribute zugeordnet werden können¹¹⁰. So lassen sich die Empfehlungen dieser Methode optimieren.

¹⁰⁶ Vgl. Aggarwal: Recommender Systems, S. 139.

¹⁰⁷ Vgl. Lops, Pasquale [u.a.]: Content-based Recommender Systems – State of the Art and Trends, In: Recommender Systems Handbook, hrsg. von Francesco Ricci [u.a.], New York, 2011, S. 75 f.

¹⁰⁸ Vgl. Aggarwal: Recommender Systems, S. 139.

¹⁰⁹ Vgl. Ebd.

¹¹⁰ Vgl. Ebd.

3.4.3 Die Methodik des Knowledge-based Recommender Systems

Die Knowledge-based Recommender Systems kommen dort zum Einsatz, wo die übrigen Filter-Methoden aufgrund ihrer Nachteile für wenig akkurate Empfehlungen sorgen würden¹¹¹. Das ist vor allem auf Domains der Fall, auf denen sehr komplexe Items angeboten werden. Darunter fallen zum Beispiel Immobilien, Autos oder Finanzdienste. Diese Items weisen meist eine sehr komplexe Struktur auf. Zum Beispiel eine Immobilie: Sie hat eine bestimmte Anzahl an Zimmern, die sich noch in Schlafzimmer, Badezimmer oder Arbeitszimmer unterteilen, außerdem hat jede Immobilie eine individuelle Lage, ein Nachbarschaftsumfeld und natürlich auch einen Kaufpreis. Oder ein Auto: Hier gibt es verschiedene Ausstattungsvarianten, von denen einige nicht miteinander kombiniert werden können oder eine Ausstattungsvariante erst Voraussetzung für eine andere ist. Diese komplexen Items – oder im Fall von Luxusuhren oder Luxuskleidung sehr teuren Items – sorgen dafür, dass sie nicht „on-the-go“ gekauft werden, weil zum Beispiel umfassende Kaufverträge geschlossen werden oder die Items schlichtweg zu teuer sind, um massenweise verkauft zu werden. Das führt dazu, dass für diese Items nur sehr wenige Ratings und Verkaufshistorien von Nutzern zur Verfügung stehen, die ein Empfehlungssystem nutzen könnte, um dem Nutzer Empfehlungen vorzustellen¹¹². Diesbezüglich stellt die zeitliche Abhängigkeit der Ratings für diese Items ein weiteres Problem dar¹¹³: So kann zum Beispiel ein Rating für ein Haus nicht mehr aktuell sein, weil es um einen Anbau ergänzt oder unterkellert wurde. Dort setzen die Knowledge-based Recommender Systems an, indem sie die Interessen des Nutzers in den Mittelpunkt der Empfehlung rücken – allerdings mit anderen Ansätzen als die in den vorhergehenden Unterkapiteln beschriebenen Methoden. Die Nutzer-Vorlieben errechnet das System beim Knowledge-based-Ansatz nicht auf passivem Weg aus der Einkaufs- und Ratinghistorie des Users, sondern es fragt den Nutzer aktiv nach seinen Vorlieben und ermöglicht es ihm, Feedback zu den ausgespielten Empfehlungen zu geben¹¹⁴. Zudem versucht das Recommender System den Nutzer zu führen und die Komplexität des Items durchschaubarer zu machen. Dazu nutzt das System bei dieser Methode drei unterschiedliche Datenquellen¹¹⁵: 1. Die vom Nutzer angegebenen

¹¹¹ Vgl. Aggarwal: Recommender Systems, S. 167.

¹¹² Vgl. Felferning, Alexander [u.a.]: Developing Constraint-based Recommender Systems, In: Recommender Systems Handbook, hrsg. von Francesco Ricci [u.a.], New York, 2011, S. 187.

¹¹³ Vgl. Ebd.

¹¹⁴ Vgl. Aggarwal: Recommender Systems, S. 168.

¹¹⁵ Vgl. Ebd.

Spezifikationen. 2. Die „Domain Knowledge“¹¹⁶. 3. Die Attribute der Items. In letzterem Punkt unterscheidet sich die Knowledge-based-Methode aber deutlich von der Content-based-Methode. Während bei letztgenannter Methode hauptsächlich Text-Keywords aus Produktbeschreibungen extrahiert werden, können die Attribute von Items in Knowledge-based Recommender Systems deutlich umfangreicher und vielschichtiger sein¹¹⁷. So kann ein Immobilien-Item zum Beispiel die Elemente Anzahl und Spezifikation der Zimmer, Lokation, Typ der Architektur, Quadratmeteranzahl und Preis enthalten.

| Item-Id | Beds. | Baths. | Locality | Type | Floor Area | Price |
|---------|-------|--------|-----------|-------------|------------|---------|
| 1 | 3 | 2 | Bronx | Townhouse | 1600 | 220,000 |
| 2 | 5 | 2.5 | Chappaqua | Split-level | 3600 | 973,000 |
| 3 | 4 | 2 | Yorktown | Ranch | 2600 | 630,000 |
| 4 | 2 | 1.5 | Yorktown | Condo | 1500 | 220,000 |
| 5 | 4 | 2 | Ossining | Colonial | 2700 | 430,000 |

Abbildung 16: Attribute eines komplexen Items

Knowledge-based Recommender Systems unterscheiden sich darin, wie sie mit dem User interagieren. Grundsätzlich kann zwischen zwei Ansätzen unterschieden werden¹¹⁸: 1. Constraint-based Recommender Systems: In diesen Recommender Systems geben die Nutzer ihre Anforderungen und Beschränkungen (constraints) für die Attribute der Items an. In den meisten Systemen existieren außerdem domain-spezifische Regeln wie zum Beispiel „Autos mit dem Baujahr vor 1970 haben keine Start/Stop-Automatik“. Zudem können die meisten Systeme aus den Vorgaben der User und den Attributen der Items zusätzliche Regeln ableiten, wie zum Beispiel: „Ältere Investoren investieren nicht in Wertpapiere mit sehr hohem Risiko“. Das System ermöglicht es dem User außerdem, die angezeigten Ergebnisse weiter zu modifizieren. 2. Case-based Recommender Systems: In diesen Systemen werden von den Nutzern spezifische Fälle (cases) vorgegeben, die als Ziele oder Ankerpunkte für die ausgegebenen Empfehlungen dienen. Auf Basis der Item-Attribute werden nun Ähnlichkeiten zwischen dem vorgegebenen Fall des Users und den Items bestimmt und eine Empfehlung ausgegeben, auf die der Nutzer erneut Einfluss nehmen kann – bis ihn das Ergebnis zufrieden stellt.

¹¹⁶ Unter „Domain Knowledge“ versteht man das Domain-spezifische Wissen, zum Beispiel zu den enthaltenen Items. Dieses Wissen kann von Domain zu Domain sehr unterschiedlich sein.

¹¹⁷ Vgl. Aggarwal: Recommender Systems, S. 172.

¹¹⁸ Vgl. Ebd., S. 169.

Neben den beiden Interaktionsansätzen kann die Art der Führung des Nutzers durch die Empfehlungen ebenfalls voneinander unterschieden werden¹¹⁹: 1. Conversational Systems: Diese Systeme generieren die Vorlieben des Nutzers aus einer Feedbackschleife. 2. Search-based Systems: Diese Systeme nutzen vordefinierte Fragen, um die Vorlieben des Nutzers herauszufinden. Ein Beispiel könnte folgende Frage sein: „Wohnst du lieber auf dem Land oder in der Stadt?“ 3. Navigation-based Systems: In diesen Systemen kann der Nutzer mittels iterativer Sets von Änderungswünschen die ausgegebenen Ergebnisse weiter verfeinern. Ein Beispiel wäre hier: „Ich hätte gerne ein Haus 5 Kilometer westlich von der jetzigen Empfehlung“, was zum Beispiel über Dropdown-Menüs generiert werden könnte. Egal welche Art der Nutzerführung in der Knowledge-based-Methode zum Einsatz kommt, oberstes Ziel sollte immer folgendes sein: „This interactive process is used to guide the user towards the final recommendation.“¹²⁰

Ihre Herangehensweise macht die Knowledge-based Recommender Systems sehr effektiv bei der Lösung des Kaltstart-Problems, dafür hat die Methode aber Probleme bei der Vorhersage von Ratings, da sie nicht auf historische Daten zurück greift¹²¹.

3.4.4 Hybride Recommender Systems

In vielen Einsatzszenarien ist es sinnvoll, die unterschiedlichen Methoden der jeweiligen Filter-Methoden miteinander zu kombinieren, um möglichst akkurate Empfehlungen für die Nutzer treffen zu können. In ein hybrides Recommender System können allerdings nicht nur die weiter oben genannten Methoden einfließen, sondern auch noch viele weitere. Wichtige Bestandteile vieler hybrider Systeme sind zum Beispiel auch Demographic Recommender Systems, die demografische Daten der Nutzer als Grundlage für Empfehlungen nehmen¹²². Diese Systeme ergänzen in hybriden Anwendungen oftmals wichtige Kontext-Informationen zu den Nutzern und haben somit einen erheblichen Einfluss auf die Wertigkeit der ausgegebenen Empfehlungen.

Hybride Recommender Systems sind auf drei unterschiedlichen Wegen zu erschaffen¹²³: 1. Ensemble Design. In diesem Design werden die Ergebnisse der jeweiligen Empfehlungsmethoden gesammelt und als eine einzige Empfehlung ausgegeben. Dies geschieht in den meisten Fällen durch eine unterschiedliche Gewichtung der einzeln gemachten Empfehlungen, die schließlich zu einer Empfehlung kombiniert werden. 2. Monolithic

¹¹⁹ Vgl. Aggarwal: Recommender Systems, S. 170.

¹²⁰ S. Ebd., S. 169.

¹²¹ Vgl. Ebd., S. 194.

¹²² Vgl. Ebd., S. 19.

¹²³ Vgl. Ebd., S. 200.

Design: Hier werden die Ergebnisse der unterschiedlichen Methoden nicht gesammelt, sondern es greift ein Algorithmus, der Elemente aller Methoden enthält, direkt auf die vorhandenen Datenquellen zu. Daraus errechnet er wiederum eine einzige Empfehlung. 3. Mixed Systems: Dieses System setzt auf die verschiedenen Filter-Methoden, die jeweils in voneinander unabhängigen Black-Boxen ihre Empfehlungen errechnen. In den Mixed Systems werden sie allerdings nicht zusammengefasst, sondern jede einzelne Methode hat ihren eigenen Output. „It is the combination [...] that creates the recommendation.“¹²⁴ Der Sinn der hybriden Systeme ist, die Vorteile der unterschiedlichen Methoden miteinander zu kombinieren und so die Nachteile der verschiedenen Methoden auszugleichen oder im besten Fall sogar aufzuheben.

3.5 Die Entstehung einer Empfehlung im Recommender System

Das vorhergehende Unterkapitel hat aufgezeigt, wie viele unterschiedliche Methoden zur Verfügung stehen, um eine Empfehlung zu generieren. In diesem Unterkapitel soll sich nun mit der Frage beschäftigt werden, wie eine Empfehlung in einem Recommender System entsteht. Dies soll anhand der Methoden des Item-based Collaborative Filtering, den Content-based Recommender Systems, der Constraint-based-Methode sowie den hybriden Recommender Systems näher untersucht werden.

3.5.1 Voraussetzungen für die Empfehlungserstellung

Bei der Untersuchung der Voraussetzungen für die Empfehlungserstellung der beiden Methoden spielen vor allem die Datenquellen, auf die die jeweiligen Algorithmen zugreifen können, eine wichtige Rolle. Dabei greifen sie beim Content-based Recommender System auf zwei Quellen zu¹²⁵: Die Ratings der User und die Attribute der Items. Diese liegen, wie weiter oben in Abbildung 15 zu sehen, zum einen in der Produktdatenbank und zum anderen in der User-Datenbank der Domain. Beim Item-based Collaborative Filtering kommen lediglich die historischen Ratings des Nutzers zum Einsatz¹²⁶. Bei der Knowledge-based-Methode werden drei Datenquellen genutzt¹²⁷: Die Spezifikationen der Nutzer, die Attribute der Items sowie die Domain Knowledge. Dabei liegen die Informationen zum Nutzer in der User-Datenbank und die Informationen zu den Items und zur Domain Knowledge in der Produktdatenbank.

¹²⁴ Vgl. Aggarwal: Recommender Systems, S. 200.

¹²⁵ Vgl. Ebd., S. 169.

¹²⁶ Vgl. Ebd.

¹²⁷ Vgl. Ebd.

Um möglichst gute Empfehlungen zu ermöglichen, muss das Recommender System folglich auf alle verfügbaren Datenquellen zugreifen können. Ist der Zugriff gewährleistet, startet der Recommendation Process.

3.5.2 Die Empfehlungserstellung beim Item-based Collaborative Filtering

Die Grundlage dieser Methode, die Empfehlungen auf Basis der Ähnlichkeit zwischen zwei Items gibt, ist eine $m \times n$ -Matrix, die anhand der von verschiedenen Usern m erstellten Ratings für die Items n aufgestellt wird. Ein Beispiel dafür könnte folgende Matrix sein:

| Item-Id ⇒ | 1 | 2 | 3 | 4 | 5 | 6 | Mean Rating |
|-----------|---|---|---|---|---|---|-------------|
| User-Id ↓ | | | | | | | |
| 1 | 7 | 6 | 7 | 4 | 5 | 4 | 5.5 |
| 2 | 6 | 7 | ? | 4 | 3 | 4 | 4.8 |
| 3 | ? | 3 | 3 | 1 | 1 | ? | 2 |
| 4 | 1 | 2 | 2 | 3 | 3 | 4 | 2.5 |
| 5 | 1 | ? | 1 | 2 | 3 | 3 | 2 |

Abbildung 17: User-Item-Matrix mit zugehörigen Ratings

Für das Item-based Collaborative Filtering sind demnach vor allem die Spalten der Matrix wichtig, da sie die Ratings der unterschiedlichen Nutzer enthalten. In der rechten Spalte sind zudem die Durchschnitts-Ratings der Nutzer angegeben. Um der Problematik zu entgehen, dass manche User Items generell besser oder schlechter bewerten, werden „mean-centered ratings“ gebildet, indem das Durchschnitts-Rating von den jeweiligen Ratings in der Zelle abgezogen wird¹²⁸. Nach diesem Schritt sieht die Matrix wie folgt aus:

| Item-Id ⇒ | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------------------------|--------|--------|--------|--------|--------|--------|
| User-Id ↓ | | | | | | |
| 1 | 1.5 | 0.5 | 1.5 | -1.5 | -0.5 | -1.5 |
| 2 | 1.2 | 2.2 | ? | -0.8 | -1.8 | -0.8 |
| 3 | ? | 1 | 1 | -1 | -1 | ? |
| 4 | -1.5 | -0.5 | -0.5 | 0.5 | 0.5 | 1.5 |
| 5 | -1 | ? | -1 | 0 | 1 | 1 |
| Cosine(1, j) (item-item) | 1 | 0.735 | 0.912 | -0.848 | -0.813 | -0.990 |
| Cosine(6, j) (item-item) | -0.990 | -0.622 | -0.912 | 0.829 | 0.730 | 1 |

Abbildung 18: User-Item-Matrix mit zentrierten Durchschnittswerten

Die „mean-centered ratings“ seien in der Berechnung s_{ij} , wobei u für die jeweiligen User und j für die entsprechenden Items steht. Dabei sei U_i der Index aller User, die Ratings für das Item i ¹²⁹ abgegeben haben. Hat ein User also Ratings für das erste, dritte und vierte Item abgegeben, ist $U_i = \{1, 3, 4\}$. Auf dieser Basis lässt sich die Formel zur Berechnung der Kosinus-Ähnlichkeit aufstellen¹³⁰:

¹²⁸ Vgl. Aggarwal: Recommender Systems, S. 36 f.

¹²⁹ Hier beiden Variablen i und j geben hier die beiden zu vergleichenden Items an.

¹³⁰ Vgl. Aggarwal: Recommender Systems, S. 40.

$$\text{CosSim}(i, j) = \frac{\sum_{u \in U_i \cap U_j} s_{ui} \cdot s_{uj}}{\sqrt{\sum_{u \in U_i \cap U_j} s_{ui}^2} \cdot \sqrt{\sum_{u \in U_i \cap U_j} s_{uj}^2}}$$

Anhand der Kosinus-Ähnlichkeit kann nun bestimmt werden, wie ein User u ein bisher nicht bewertetes Item t bewerten würde. Hier ist der erste Schritt, die top- k ähnlichen Items zu t über die Kosinus-Ähnlichkeit zu bestimmen. Diese seien als $Q_t(u)$ bezeichnet. Auf dieser Basis lässt sich das fehlende Rating wie folgt bestimmen¹³¹:

$$\hat{r}_{ut} = \frac{\sum_{j \in Q_t(u)} \text{CosSim}(j, t) \cdot r_{uj}}{\sum_{j \in Q_t(u)} |\text{CosSim}(j, t)|}$$

Mit leichter Abänderung der verwendeten Variablen lässt sich mit dieser Methode auch die Ähnlichkeit zwischen zwei unterschiedlichen Nutzern bestimmen.

In einem Anwendungsbeispiel sollen nun die fehlenden Ratings 1 und 6 des User 3 aus der Abbildung 17 berechnet werden. Zuerst müssen die Kosinus-Ähnlichkeiten zwischen den unterschiedlichen Items berechnet werden. Das geschieht durch das Einsetzen aller Werte in die oben beschriebene Formel:

$$\text{CosSim}(1,3) = \frac{1,5 * 1,5 + (-1,5) * (-0,5) + (-1) * (-1)}{\sqrt{1,5^2 + (-1,5)^2 + (-1)^2} \cdot \sqrt{1,5^2 + (-0,5)^2 + (-1)^2}} = 0,912$$

Die Kosinus-Ähnlichkeiten sind in Abbildung 18 in den unteren beiden Zeilen eingetragen. Weil die Items 2 und 3 am ähnlichsten zu Item 1 sind (Kosinus-Ähnlichkeiten: 0,735 und 0,912 zu 1) werden diese zur Vorhersage des Ratings zu Item 1 genutzt. Gleiches gilt für die Items 4 und 5 zur Vorhersage des Ratings von Item 6 (Kosinus-Ähnlichkeiten: 0,829 und 0,730 zu 1). Diese Werte werden in die oben beschriebene Formel zur Vorhersage der fehlenden Ratings eingesetzt:

$$\hat{r}_{31} = \frac{3 * 0,735 + 3 * 0,912}{0,735 + 0,912} = 3$$

$$\hat{r}_{36} = \frac{1 * 0,829 + 1 * 0,730}{0,829 + 0,730} = 1$$

Die Ergebnisse zeigen, dass aus dem vorhandenen Pool an Items User 3 das Item 1 vergleichsweise besser bewertet hätte, weshalb das Recommender System ihm dieses Item als Empfehlung ausgeben würde.

Die Methode, Ratings mit Hilfe der Kosinus-Ähnlichkeit vorherzusagen, ist nicht die einzige, neben ihr existieren zum Beispiel mit den Regressionsmodellen oder der Pearson-Korrelation noch viele weitere Methoden, mit denen fehlende Ratings berechnet werden können.

¹³¹ Vgl. Aggarwal: Recommender Systems, S. 40.

3.5.3 Die Empfehlungserstellung im Content-based Recommender System

Die Empfehlungserstellung in einem Content-based Recommender System läuft in drei unterschiedlichen Schritten ab¹³²: 1. Preprocessing und Feature-Extraktion: In diesem ersten Schritt werden aus den Produktbeschreibungen der Items textbasierte Vektorräume geschaffen. Dieser Arbeitsschritt läuft aufgrund der hohen Komplexität offline ab und ist sehr domain-spezifisch. 2. Content-basiertes Lernen von Nutzerprofilen: In diesem Schritt wird aus den textbasierten Vektorräumen und Nutzerdaten, die vor allem aus Ratings bestehen, ein nutzerspezifisches Modell erschaffen, das auf den Item-Interessen des Nutzers aufbaut. Dieses Modell wird anhand eines Sets von Trainings- und Test-Daten optimiert. Auch dieser Schritt läuft aufgrund seiner Komplexität offline ab. 3. Ausgabe der Empfehlung: Auf Basis des trainierten User-Modells werden in diesem Schritt Empfehlungen für den spezifischen Nutzer generiert. Dieser Schritt sollte in Echtzeit ausgeführt werden.

Das Preprocessing und das Anlegen der Attribute ist – wie weiter oben beschrieben – sehr domainspezifisch, weil jede Domain unterschiedliche Items enthält. Am Beispiel eines Videos könnten zum Beispiel die Videobeschreibung und weitere Metadaten genutzt werden, um Attribute herauszustellen. Bei der Herausbildung von Attributen aus Beschreibungen kommen Techniken wie Stopwörter-Entfernung, Stemming und Phrasenerkennung¹³³. Die herausgebildeten Attribute der Items können zudem noch nach ihrer Bedeutung gewichtet werden¹³⁴.

Beim Lernen der unterschiedlichen User-Modelle gibt es keinen einheitlichen Ansatz, meist richtet er sich nach den Zielsetzungen des Recommender Systems. Eine bewährte Herangehensweise ist die im Unterkapitel 3.5.2 bereits beschriebene Methode der Kosinus-Ähnlichkeit, über die sich Ähnlichkeiten zwischen den Attributen der Items bestimmen lassen und schließlich auch ein Rating vorhersagen und damit auch eine nutzer-spezifische Empfehlung machen lässt.

3.5.4 Die Empfehlungserstellung im Constraint-based Recommender System

Die Constraint-based Recommender Systems erlauben es dem User, eigene Anforderungen oder Beschränkungen für Items zu formulieren und so den Empfehlungsprozess stark zu

¹³² Vgl. Lops: Content-based Recommender Systems, S. 75 f.

¹³³ Vgl. Aggarwal: Recommender Systems, S. 145.

¹³⁴ Vgl. Ebd., S. 147 ff.

beeinflussen. Problem ist dabei, dass die User ihre Queries¹³⁵ nicht immer so formulieren, wie die Attribute die gesuchten Items beschreiben, weshalb das System vor der Herausforderung steht, die Anfrage des Users mit den vorhandenen Item-Attributen zu verarbeiten. Das funktioniert am besten über Regeln, die dem System zuerst vorgegeben werden müssen¹³⁶. Die Regeln sollten schon vor dem ersten Einsatz des Recommender Systems größtenteils festgelegt sein, auch wenn im Lauf der Benutzung neue hinzugefügt werden können. Grundlage für die Regelerstellung können zum Beispiel die Ergebnisse von Data und Text Minings in den eigenen oder öffentlichen Datenbanken oder Erfahrungen mit den Items der Domain aus der Vergangenheit sein¹³⁷.

Neben den bereits festgelegten Regeln fragen die Constraint-based-Systeme die Nutzeranforderungen ab, um sie zu User-Attributen umzuformen¹³⁸. Am Beispiel Autos könnten die Attribute zum Beispiel wie folgt aussehen: Stadt oder Gelände (binärer Wert), Familiengröße (numerischer Wert), PLZ des Nutzers (Numerischer Wert), Marke (kategorischer Wert), Anzahl der Türen (numerischer Wert), maximaler Preis (numerischer Wert) und PS-Anzahl (numerischer Wert). Die aus dieser Abfrage extrahierten Nutzer-Attribute können also neben den Anforderungen des Nutzers auch demografische Informationen zum Nutzer enthalten.

Die oben beschriebenen Anforderungen und Attribute werden in einem Constraint-based Recommender System in einem interaktiven Prozess vom Nutzer abgefragt und müssen anschließend in Verbindung zu den Attributen der Items gesetzt werden. Das geschieht über die „Knowledge Bases“, die die bereits angesprochenen Regeln enthalten¹³⁹. Diese könnten zum Beispiel so aussehen: Stadt oder Gelände=Gelände → Nutzung im Gelände=[Liste von geländegängigen Autos].

Die Regeln sind folglich Filter-Bedingungen, welche die Produktdatenbank nach den Vorgaben des Nutzers durchsuchen. Gleichzeitig kann das System aus den Anforderungen des Nutzers Beschränkungen extrahieren, zum Beispiel: Familiengröße=6 → min.Sitze≥6. Anhand dieser aus den Anforderungen des Nutzers abgeleiteten Regeln bildet das System nun eine Query, die zur Durchsuchung der Item-Datenbank durchsucht wird. Diese könnte wie folgt aussehen: $(\text{min.Sitze} \geq 6) \wedge (\text{min.Türen} \geq 5) \wedge (\text{Marke} = \text{“Audi“}) \wedge (\text{Preis} < 30.000) \wedge$

¹³⁵ Unter Queries versteht man generell Suchanfragen von Usern, die sich aus verschiedensten Elementen zusammensetzen können.

¹³⁶ Vgl. Aggarwal: Recommender Systems, S. 172.

¹³⁷ Vgl. Ebd., S. 174.

¹³⁸ Vgl. Ebd., S. 172.

¹³⁹ Vgl. Ebd., S. 173.

(PLZ=22303). Die aus dieser Query resultierende Liste an empfohlenen Items kann der Nutzer nun bei Bedarf weiter modifizieren, bis er das für ihn passende Ergebnis gefunden hat. Darüber hinaus ist es sinnvoll, ein Ranking in die ausgegebene Liste einzubetten. Die einfachste Methode ist dabei, die Liste für den User nach Preis, Anzahl der Türen oder Nähe zum eigenen Standort sortierbar zu machen¹⁴⁰. Darüber hinaus ist auch ein Ranking nach Relevanz möglich, was über eine gewichtete Nutzenfunktion umgesetzt werden kann¹⁴¹.

Die Empfehlungserstellung in Constraint-based Recommender Systems funktioniert also in drei unterschiedlichen Schritten, die in folgender Abbildung gut repräsentiert sind:

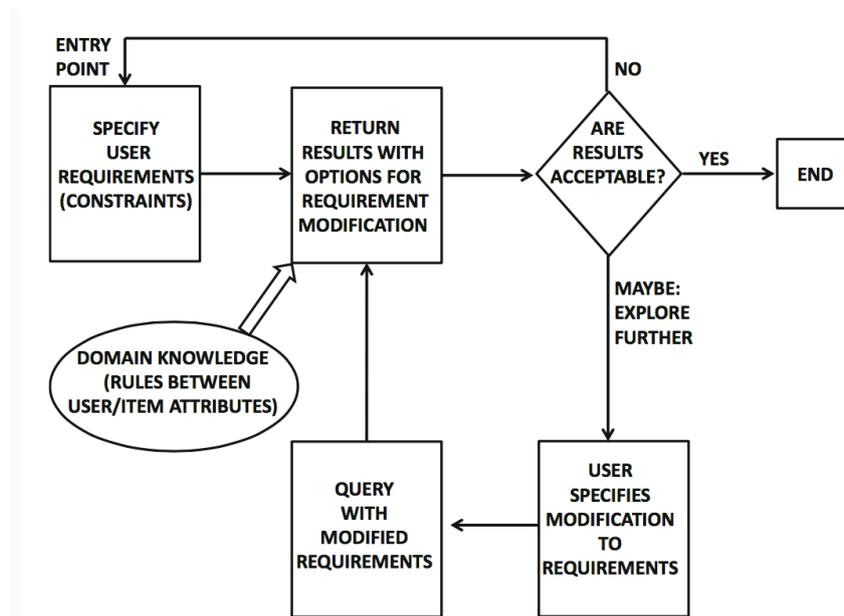


Abbildung 19: Empfehlungserstellung in einem Constraint-based Recommender System

1. Über ein interaktives Interface gibt der User seine Präferenzen und Beschränkungen für die Item-Suche ein, worauf die Datenbank nach den Nutzer-Vorgaben durchsucht wird.
2. Dem User wird eine Liste von relevanten Items ausgegeben. Oft enthält diese Anzeige auch eine Erklärung dafür, wie die Auswahl zustande gekommen ist. Das hilft dem User, die Empfehlung nachvollziehen zu können. Zudem erhält die Anzeige meist Vorschläge, wie der Nutzer seine Suche verfeinern kann.
3. Auf Basis der Empfehlungen verfeinert der Nutzer seine Suche immer weiter, bis er das gesuchte Item gefunden hat.

Der Empfehlungsprozess in einem Constraint-based Recommender System funktioniert folglich über eine iterative Feedbackschleife, die der Nutzer selbst steuern kann¹⁴².

¹⁴⁰ Vgl. Aggarwal: Recommender Systems, S. 178.

¹⁴¹ Vgl. Ebd.

¹⁴² Vgl. Ebd., S. 176.

3.5.5 Zusammenkommen der Empfehlungen im hybriden Recommender System

In den hybriden Recommender Systems laufen die von den unterschiedlichen Algorithmen ermittelten Ratings zusammen und werden zu einem einzigen Wert kombiniert. An dieser Stelle ist es auch möglich, die Ergebnisse der verschiedenen Systeme zu gewichten, um die Empfehlungserstellung noch weiter optimieren zu können. \hat{R} steht dabei für das final ermittelte Rating, $\hat{R}_1 \dots \hat{R}_q$ für die bereits spezifizierten Ratings, vorhersagt von q unterschiedlichen Algorithmen. Mit den Gewichtungen $\alpha_1 \dots \alpha_q$ lässt sich über folgende Formel das gewichtete Rating ermitteln¹⁴³:

$$\hat{R} = \sum_{i=1}^q \alpha_i \hat{r}_{uj}^i$$

Diese Methode ist die simpelste, um die von unterschiedlichen Recommender Systems ermittelten Ratings miteinander zu verrechnen. Darüber hinaus besteht zum Beispiel noch die Möglichkeit, die Gewichtung der Ratings nicht händisch zu ermitteln, sondern diese ebenfalls mathematisch zu bestimmen¹⁴⁴.

3.5.6 Evaluation des Recommender Systems

Ist das finale Rating ermittelt und ein relevantes Item an den User übermittelt, ist der Entstehungsprozess einer Empfehlung eigentlich abgeschlossen. Dennoch nimmt in der vorhandenen Literatur zu den Recommender Systems auch die Evaluation ebendieser, also die Auswertung der für den Nutzer vorgeschlagenen Empfehlungen, eine wichtige Rolle ein. Die große Problematik bei der Evaluation von Recommender Systems ist die bereits angesprochene Diversität bei diesen Systemen. Dementsprechend herrscht auch bei der Evaluation eine große Vielfalt bei den Herangehensweisen, die sich aber meist an den Zielen orientieren, die für das Recommender System vorgegeben wurden.

Weil die Evaluation von Recommender Systems vor dem Hintergrund der Leitfrage der vorliegenden Arbeit nur eine Nebenrolle spielt, soll nur ein interessanter Punkt angesprochen werden, der sich in der Recherche herausstellte¹⁴⁵. Herlocker und seine Kollegen stellen fest, dass gute Recommender Systems, die mit Collaborative Filtering-Algorithmen arbeiten, eine Fehlerrate von 0,73 Punkten auf einer 5-Punkte-Skala

¹⁴³ Vgl. Aggarwal: Recommender Systems, S. 206.

¹⁴⁴ Vgl. hierzu Ebd., S. 206 ff.

¹⁴⁵ Tiefere Einblicke in die Evaluation von Recommender Systems gibt der vielbeachtete Aufsatz von der Forschergruppe um den Computerwissenschaftler Jonathan Herlocker: Herlocker, Jonathan [u.a.]: Evaluating Collaborative Filtering Recommender Systems, In: ACM Transactions on Information Systems (Band 22), New York, 2004, S. 5-53.

aufweisen¹⁴⁶. Sie folgern daraus, dass dieser Wert eine magische Grenze darstellen könnte, die nicht mehr weiter verfeinert werden kann. Als Grund dafür geben Herlocker und seine Kollegen die Inkonsistenzen bei der Rating-Vergabe der User selbst an. So bewerte ein User bei einem zweiten Bewertungsdurchlauf nicht alle Items genauso wie beim ersten Durchlauf, was dafür sorgen könnte, dass Inkonsistenzen entstehen, die sich bis in die Vorhersage der Ratings auswirken könnten¹⁴⁷.

3.6 Netzwerkeffekte und Recommender Systems

Ähnlich wie die Evaluation der Recommender Systems soll auch das Thema „Netzwerkeffekte und Recommender Systems“ nur am Rande behandelt werden, obwohl es eine eigenständige Arbeit wert wäre. Dabei ist der Einfluss von Netzwerkeffekten auf Recommender Systems noch nicht sehr ausführlich untersucht worden, fest steht nur, *dass* Netzwerkeffekte eine Auswirkung auf Empfehlungssysteme haben: „Network effects are relevant to all Recommender Systems that are based on user input or behavior.“¹⁴⁸

Dabei sind die Netzwerkeffekte grundsätzlich in direkte (horizontale) und indirekte (vertikale) Netzwerkeffekte zu unterscheiden.¹⁴⁹ Direkte Netzwerkeffekte können darüber hinaus entweder positiv oder negativ sein.

Die direkten Netzwerkeffekte beschreiben den Umstand, wenn der Nutzen eines Netzwerks mit jedem neuen Teilnehmer, der als Knoten angesehen wird, überproportional ansteigt. Mit der Zahl der Teilnehmer steigt folglich auch der Wert des Netzwerks.¹⁵⁰

Grundlage dafür ist das Metcalf'sche Gesetz zu Kosten-Nutzen-Verhältnissen in Kommunikationssystemen. Ein beliebtes Beispiel für die Anwendung des Gesetzes ist die Vernetzung von Netzwerk-Teilnehmern mit Telefonen: Ist ein Teilnehmer alleine im Netzwerk, hat er keine Vernetzungsmöglichkeiten. Existieren im Netzwerk zwei Teilnehmer, sind zwei Verbindungen möglich, bei drei Teilnehmern sind es sechs Verbindungen, bei vier Teilnehmern sind es 12 Verbindungen und so weiter. Bei den direkten Netzwerkeffekten gibt es positive und negative Effekte. Ein positiver Effekt wäre

¹⁴⁶ Vgl. Herlocker, Jonathan [u.a.]: Evaluating Collaborative Filtering Recommender Systems, In: ACM Transactions on Information Systems (Band 22), New York, 2004, S. 6.

¹⁴⁷ Vgl. Ebd., S. 7.

¹⁴⁸ Vgl. Neumann, Andreas: Recommender Systems for Information Providers – Designing Customer Centric Paths to Information, Heidelberg, 2009, S. 28.

¹⁴⁹ Vgl. Erhardt, Marcus: Netzwerkeffekte, Standartisierung und Wettbewerbsstrategie, Wiesbaden, 2001, S. 28.

¹⁵⁰ Vgl. Katz, Michael; Shapiro, Carl: Product Compatibility Choice in a Market with Technological Process, In: Oxford Economic Papers, London, 1986, S. 146 ff.

zum Beispiel die enorme Nutzensteigerung mit jedem neuen Teilnehmer, ein negativer Effekt die Überlastung des Mobilfunknetzes bei zu vielen Teilnehmern¹⁵¹.

Indirekte Netzwerkeffekte liegen vor, wenn ein Gut bei steigender Nutzerbasis attraktiver wird, ohne dass eine direkte Verbindung zwischen Nutzerbasis und Gut besteht¹⁵². Ein Beispiel dafür bietet die Entwicklung für Apps für mobile Geräte: Hat Apples mobiles Betriebssystem iOS eine gewisse Nutzerbasis, wird es für Entwickler attraktiv, Software für dieses Betriebssystem zu entwickeln. Mit steigender Vielfalt an Apps im Appstore steigen wiederum die Nutzer. Die beiden Gruppen „Nutzer“ und „Entwickler“ hängen dabei nicht direkt zusammen, sondern sind indirekt miteinander in diesem Netzwerk verbunden.

Übertragen auf Recommender Systems lässt sich feststellen, dass auch hier die Nutzer sowohl von direkten als auch von indirekten Netzwerkeffekten profitieren – abhängig von der angewendeten Filter-Methode. Das Collaborative Filtering mit der Neighborhood-Methode kann zum Beispiel als direkter Netzwerkeffekt verstanden werden, da hier die Nutzer des Netzwerks untereinander in Beziehung gesetzt werden und der Wert des Netzwerks mit jeder Nutzereingabe und jedem neuen Nutzer erhöht wird. An dieser Stelle könnte allerdings die Frage gestellt werden, ob unter diesem Blickpunkt die Teilnehmer des Netzwerks genauer kategorisiert werden müssten, da anders als beim Telefon-Beispiel die Interaktion des Nutzers mit einem Recommender System deutlich vielschichtiger ist. So kann es zum Beispiel Nutzer geben, die nur unäre Ratings hinterlassen, indem sie die Items der Seite scannen und andere Nutzer, die für jedes ihrer gekauften Item ein Intervall-basiertes Rating und eine Text-Rezension hinterlassen. Letzterer User würde dem Netzwerk einen höheren Wert hinzufügen als der erstgenannte User.

3.7 Herausforderungen bei der Einführung eines Recommender Systems

Einige Herausforderungen, die bei der Einführung eines Recommender Systems bewältigt werden müssen, sind in dieser Arbeit bereits kurz angesprochen worden. In diesem Kapitel sollen sie noch einmal gesammelt und etwas ausführlicher behandelt werden:

1. Kaltstart-Problem: Das Kaltstart-Problem taucht bei vielen unterschiedlichen Filter-Methoden auf, vor allem sind davon aber die kollaborativen Filter-Methoden betroffen. Grundsätzlich beschreibt das Kaltstart-Problem die Problematik, wenn das Recommender System auf keine oder nur sehr wenige Daten zu Usern und Items zugreifen kann, woraus

¹⁵¹ Vgl. Dewald, Pirmin; Robers, Sebastian: Netzwerkeffekte: Theoretische Grundlagen, Kaiserslautern, 2015, S. 3.

¹⁵² Vgl. Erhardt: Netzwerkeffekte, S. 27.

wenig akkurate Empfehlungen für den Nutzer resultieren¹⁵³. Der Ursprung des Problems liegt in der User-User- oder User-Item-Matrix, die zu viele leere Felder aufweist, als dass ein Recommender System sie mit Ratings füllen könnte, die es aus vorhandenen Ratings berechnet. Eine Lösung für das Kaltstart-Problem ist zum Beispiel die Verwendung von hybriden Systemen, die andere Datenquellen zur Vorhersage nutzen oder wie bei den Knowledge-based Recommender Systems zunächst Anforderungen vom Nutzer abfragen¹⁵⁴.

2. Longtail-Problem: Das Longtail-Problem tritt vor allem beim Neighborhood-based Collaborative Filtering in Erscheinung und ist in folgender Abbildung zu sehen:

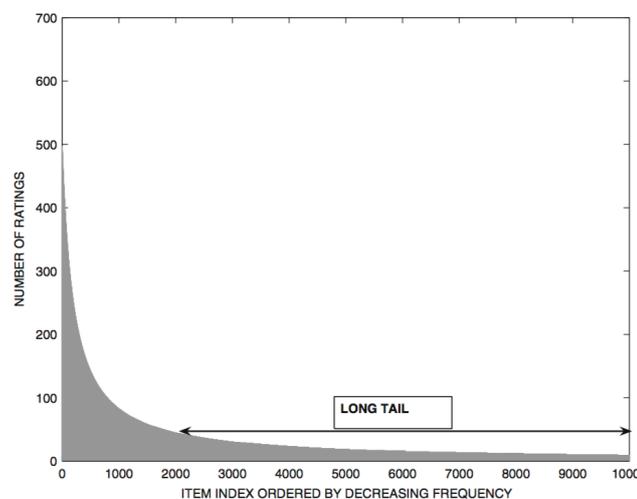


Abbildung 20: Darstellung des Longtail-Problems

Das Longtail-Problem beschreibt den Umstand, dass es nur wenige Items gibt, die mit vielen Ratings bewertet wurden. Das bedeutet im Umkehrschluss, dass es auf einer Plattform sehr viele Items gibt, die keine oder nur sehr wenige Ratings“ erhalten haben.¹⁵⁵ Das führt dazu, dass auf Basis des Neighborhood-Filterns ein Gewichtungsproblem zwischen Items mit vielen Ratings und Items mit wenigen Ratings entsteht. Beachtet man dieses Problem nicht, kann der Recommendation Process zu nicht repräsentativen und missverständlichen Ergebnissen führen¹⁵⁶. Gelöst wird das Problem – ähnlich wie das Kaltstart-Problem – mit hybriden Recommender Systems, in dem zusätzliche Filter-Methoden das Longtail-Problem neutralisieren.

¹⁵³ Vgl. Krohn-Grimberghe, Arthur: Analysis and Evaluation of Multi-relational Recommender Systems, Hildesheim, 2012, S. 33 ff.

¹⁵⁴ Vgl. Aggarwal: Recommender Systems, S. 199.

¹⁵⁵ Vgl. Ebd., S. 32.

¹⁵⁶ Vgl. Ebd.

3. Overspecialization: Mit Overspecialization wird ein Problem benannt, das vor allem in Content-based Recommender Systems entsteht. Aufgrund der Filter-Methode werden dem Nutzer in diesen Systemen in den meisten Fällen sehr naheliegende Empfehlungen gemacht. Auch die oben beschriebenen Zufallsfunde sind durch das Overspecialization-Problem äußerst selten¹⁵⁷. Das Problem entsteht daraus, dass auf Basis der Attribute der Items Ähnlichkeiten berechnet werden, was dazu führt, dass im Fall von Video-Items nur miteinander artverwandte Filme empfohlen werden. Eine Lösung für dieses Problem ist erneut der Einsatz eines hybriden Recommender Systems, das neben der Content-based-Methode auch noch andere Methoden zur Empfehlungserstellung nutzt.

4. Kontext: Der Einfluss von Kontext bei der Empfehlungserstellung kann für ein Recommendation System nicht nur eine Herausforderung, sondern auch eine Chance sein, die Genauigkeit der Empfehlungen deutlich zu verbessern. Die Herausforderung für ein Recommender System besteht darin, dass das Nutzerverhalten zum Beispiel stark von der Tageszeit oder bestimmten Wochentagen abhängig sein kann¹⁵⁸. Die Lösung für das Kontext-Problem ist recht simpel, so kann zum Beispiel die oben beschriebene User-Item-Matrix um eine zusätzliche Dimension wie Zeit oder Wochentag erweitert werden¹⁵⁹.

5. Filter Bubble: Das Filter Bubble-Problem resultiert in gewisser Weise aus dem Problem der Overspecialization. Wie Eli Pariser 2011 erkannte, sorgt die zunehmende Personalisierung von Webseiten dafür, dass eine Filterblase um uns entsteht, aus der wir nur sehr schwer ausbrechen können. In dieser Filterblase werden wir immer wieder mit dem ähnlichem Content konfrontiert, neue Inhalte können uns nicht erreichen¹⁶⁰. Eine Strategie zur Vermeidung der Filter-Bubble-Bildung oder um aus ihnen auszubrechen, sind die „Structural Holes“¹⁶¹, die am Beispiel Recommender Systems in Form von Zufallsfunden auftreten können.

6. Performance: Gerade bei großen Nutzer- und Produktdatenbanken und komplexen Recommender Systems spielt auch die Performance des gesamten Systems eine wichtige Rolle. Im besten Fall sollte die Empfehlungserstellung in Echtzeit funktionieren, was in einem oben genannten Fall eine große Herausforderung sein kann¹⁶². Zur Erhöhung der

¹⁵⁷ Vgl. Aggarwal: Recommender Systems, S. 183.

¹⁵⁸ Vgl. Ricci: Recommender Systems Handbook, S. 6.

¹⁵⁹ Vgl. Aggarwal: Recommender Systems, S. 285.

¹⁶⁰ Vgl. Pariser, Eli: The Filter Bubble – What the Internet is Hiding from You, New York, 2011, S. 17 ff.

¹⁶¹ Vgl. Burt, Ronald: Structural Holes and Good Ideas, In: American Journal of Sociology (Band 110), Chicago, 2004, S. 349 ff.

¹⁶² Vgl. Ricci: Recommender Systems Handbook, S. 26.

Performance kann zum Beispiel die Offline-Berechnung von Teilen des Empfehlungsprozesses beitragen¹⁶³ – dass zusätzlich solch ein System über ausreichend Speicher- und Rechenkapazität verfügt, ist vorausgesetzt.

7. Disruption: Disruptive Veränderungen im Nutzungsverhalten von Webseiten kann für die dahinterliegenden Recommender Systems ebenfalls eine Herausforderung sein¹⁶⁴. So kann eine einschneidende private Veränderung das Kaufverhalten und das Item-Interesse stark verändern. Eine Strategie, um disruptive Veränderungen zu verarbeiten, kann zum Beispiel das Verwenden von Kontext-Informationen wie Tageszeiten oder Demographie des Nutzers in einem Recommender System sein.

8. Vertrauen in die Empfehlungen: Eine weitere Herausforderung für Recommender Systems ist das Vertrauen des Nutzers in die vorgeschlagenen Empfehlungen. So gibt es bei vielen Usern eine Grundskepsis gegenüber maschinell erstellten Empfehlungen. Sie verlassen sich lieber auf von Menschen gemachte Empfehlungen, in der Literatur auch als „word of mouth“ beschrieben¹⁶⁵. Eine Strategie, um dieser Herausforderung zu begegnen, ist relativ neu: Dank der Einbindung von sozialen Netzwerken, kann ein Recommender System ein „social trust network“¹⁶⁶ aufbauen, über das es wichtige Kontext-Informationen sammeln kann, um dem Misstrauen von Usern zu begegnen¹⁶⁷.

9. Sicherheit von Recommender Systems: Da die meisten Recommender Systems an offenen Plattformen ansetzen, auf denen sich jeder User anmelden kann, ist auch die Sicherheit des Systems im Falle eines Angriffs oder bei einem Manipulationsversuch zum Beispiel über Fake-Profile eine große Herausforderung¹⁶⁸. Hier gibt es einige Abwehrmechanismen, meistens in Form von speziell entwickelten Algorithmen, die Attacken erkennen und unterbinden können¹⁶⁹. Eine andere, sehr simple Methode, um solchen Attacken vorzubeugen, ist zum Beispiel der Einsatz von CAPTCHAS bei der User-Registrierung¹⁷⁰.

¹⁶³ Vgl. Ahrens: Recommender Systems, S. 10 ff.

¹⁶⁴ Vgl. Probst, Bastian: Verwendung von Benutzermodellen zur Personalisierung von Inhalten, S. 2 f.

¹⁶⁵ Vgl. Ahrens: Recommender Systems, S. 20 ff.

¹⁶⁶ Vgl. Aggarwal: Recommender Systems, S. 346.

¹⁶⁷ Vgl. Ebd., S. 345 ff.

¹⁶⁸ Vgl. Ebd., S. 385.

¹⁶⁹ Vgl. Ebd., S. 398 ff.

¹⁷⁰ Vgl. Ebd., S. 404.

3.8 Zwischenfazit

Das Kapitel über die Recommender Systems hat einige wichtige Informationen für die vorliegende Arbeit zusammengetragen. Zentrale Erkenntnis des Kapitels ist, dass es nicht *das* Recommender System gibt, das sich eins-zu-eins für alle Vorhaben einsetzen lässt. Vielmehr ist ein Recommender System ein sehr komplexes Konstrukt, das auf viele unterschiedliche Datenquellen zugreifen kann und aus vielen verschiedenen Elementen besteht, die ganz nach dem Einsatzzweck des Systems konfiguriert werden können. Letzterer Punkt gilt vor allem für die unterschiedlichen Filtermethoden, von denen die wichtigsten in den vorhergehenden Kapiteln beschrieben wurden. Auch bei der Strategie der Empfehlungserstellung gibt es viele verschiedene Wege zur Berechnung der fehlenden Ratings in der User-Item- oder User-User-Matrix. Auch hier sollte sich vor dem Hintergrund des Einsatzzwecks für die richtige Methode entschieden werden. Daher sollte im folgenden Kapitel dieser Arbeit die Formulierung eines klaren Einsatzziels für das Recommender System für die NDR Mediathek einen wichtigen Teil ausmachen. Zusätzlich zeigte die Betrachtung der Herausforderungen für Recommender Systems, dass es sowohl schon bei der Konzeptionierung eines solchen Systems einige Punkte gibt, die zu beachten sind, als auch im laufenden Betrieb. Auch hier sollen einige Erkenntnisse mitgenommen werden, um Probleme eines Empfehlungssystems für die NDR Mediathek schon während der Konzeption zu erkennen und erste Lösungsansätze für diese Probleme zu präsentieren.

4. Rechtliche und organisatorische Struktur der öffentlich-rechtlichen Mediatheken

Das folgende Kapitel soll sich näher mit der Struktur des öffentlich-rechtlichen Rundfunks, genauer der öffentlich-rechtlichen Mediatheken, beschäftigen. Ziel soll eine Bestandsaufnahme sein, die zeigt, auf welcher Grundlage ein Recommender System beim NDR aufbauen kann und welche Dinge bei der Einführung eines Empfehlungssystems zu beachten wären.

Dazu beschäftigt sich dieses Kapitel unter anderem mit dem Auftrag des öffentlich-rechtlichen Rundfunks, der sich auf die Mediatheken beziehen lässt, sowie dem Arbeitsprozess in der Mediathek. Darunter fallen auch die Themen Datenstruktur sowie der Arbeitsablauf in der Mediathek. Zudem soll es in diesem Kapitel auch um die datenschutzrechtlichen Voraussetzungen beim NDR gehen.

4.1 Auftrag der öffentlich-rechtlichen Mediatheken

Der Auftrag der öffentlich-rechtlichen Mediatheken ist - wie bereits erwähnt – eng an den Auftrag der öffentlich-rechtlichen Rundfunkanstalten gekoppelt. Das initiale Element des Auftrags des öffentlich-rechtlichen Rundfunks findet sich schon im Grundgesetz. Dort ist in Artikel 5, Absatz 1, Satz 2 die so genannte Rundfunkfreiheit definiert. Zielsetzung der Rundfunkfreiheit ist es, eine umfassende Meinungsvielfalt innerhalb der Gesellschaft zu ermöglichen und zu bewahren. Die Rundfunkfreiheit ist zusätzlich von „Ausgestaltungsgesetzen“¹⁷¹ geprägt, die diesen Terminus noch weiter ausgestalten sollen. Diese Ausgestaltung der Rundfunkfreiheit ist unter anderem im Rundfunkstaatsvertrag (RStV) geregelt, der 1987 zum ersten Mal aufgesetzt und seitdem von einigen Novellen, den so genannten Rundfunkänderungsstaatsverträgen (RÄStV), abgeändert wurde. Im RStV ist unter anderem der Funktionsauftrag der öffentlich-rechtlichen Rundfunkanstalten explizit ausformuliert. Der Auftrag gilt dabei sowohl für die klassischen Rundfunkprogramme als auch für die Online-Angebote, die innerhalb der Rundfunkanstalten und der Gesetzestexte als Telemedien bezeichnet werden¹⁷². Bis im Juni 2009 der 12. RÄStV in Kraft trat, war der Auftrag wie folgt definiert: Der öffentlich-rechtliche Rundfunk habe „durch die Herstellung und Verbreitung von *Rundfunkprogrammen* als Medium und Faktor des Prozesses freier individueller und öffentlicher Meinungsbildung zu wirken.“¹⁷³ Nach Inkrafttreten des 12. RÄStV lautete er folgendermaßen: „Auftrag der öffentlich-rechtlichen Rundfunkanstalten ist, durch die Herstellung und Verbreitung ihrer *Angebote* als Medium und Faktor des Prozesses freier individueller und öffentlicher Meinungsbildung zu wirken und dadurch die demokratischen, sozialen und kulturellen Bedürfnisse der Gesellschaft zu erfüllen.“¹⁷⁴ Geändert wurde in der für die Telemedienangebote des öffentlich-rechtlichen Rundfunks zentralen Novellierung des RStV die Definition des verbreiteten Contents. Zählten in der alten Version ausschließlich Rundfunkprogramme, also Hörfunk und Fernsehen, zum produzierten Content der Anstalten, wurde in der Novellierung mit den „Angeboten“ ein

¹⁷¹ S. Neuhoff, Heiko: Rechtsprobleme der Ausgestaltung des Auftrags des öffentlich-rechtlichen Rundfunks im Online-Bereich, In: Materialien zur rechtswissenschaftlichen Medien- und Informationsforschung (Band 68), hrsg. von Martin Eifert u.a., Hamburg, 2011, S. 182.

¹⁷² Vgl. Todsén, Nele Julie: Grenzen gebührenfinanzierter Telemedien – Eine kritische Bewertung des Drei-Stufen-Tests mit besonderem Blick auf das Wettbewerbsverhältnis zur Presse, In: Deutsches und internationales Wirtschaftsrecht (Band 54), hrsg. von Marian Paschke, Berlin, 2013, S. 119.

¹⁷³ S. Ebd., S. 120 f.

¹⁷⁴ S. Ebd., S. 121.

Sammelbegriff eingeführt. Dieser setzt sich aus drei Angebotstypen zusammen: Zum einen aus den klassischen Rundfunkprogrammen Hörfunk und Fernsehen, zum anderen aus den Telemedien, die damit ganz offiziell die dritte Säule des öffentlich-rechtlichen Funktionsauftrags geworden sind¹⁷⁵.

Des Weiteren enthalten beide Auftragsformulierungen noch zwei zentrale Begriffe: „Medium“ und „Faktor“ der Meinungsbildung. Füllt der Rundfunk die Aufgabe des Mediums aus, übermittelt er in „möglichster Breite und Vollständigkeit Informationen und Meinungen anderer“¹⁷⁶, um den Prozess der Meinungsbildung zu ermöglichen und zu wahren. Als Faktor der Meinungsbildung ist der öffentlich-rechtliche Rundfunk dazu berufen, „eigene Auffassungen zu vertreten sowie den dem Einzelnen und gesellschaftlichen Gruppen Gelegenheit zu meinungsbildendem Wirken zu geben.“¹⁷⁷ Außerdem sei der Funktionsauftrag des Rundfunks technologieneutral, also nicht gebunden an eine bestimmte Plattform oder Technologie¹⁷⁸.

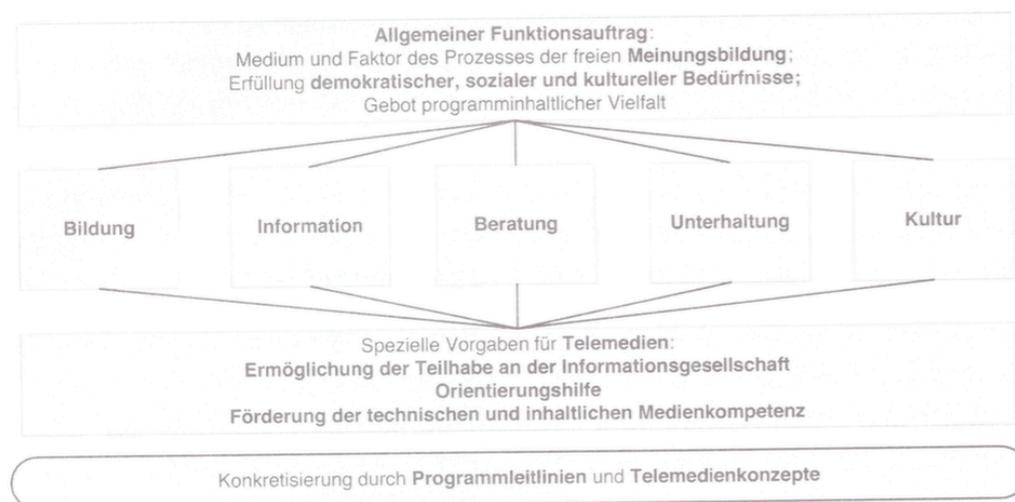


Abbildung 21: Übersicht zum allgemeinen Funktionsauftrag

Abbildung 21 zeigt, wie sich der oben beschriebene Auftrag weiter aufspannt. Der allgemeine Funktionsauftrag legt somit die Grundlage für die verschiedenen Themenkanäle, die der öffentlich-rechtliche Rundfunk zur Verbreitung von Informationen zur Meinungsbildung nutzen kann. Mit den Sparten Bildung, Information, Beratung, Unterhaltung und Kultur steht den Anstalten somit ein sehr breites Spektrum zur Nutzung zur Verfügung.

¹⁷⁵ Vgl. Todsen: Grenzen gebührenfinanzierter Telemedien, S. 121.

¹⁷⁶ S. Ebd., S. 119.

¹⁷⁷ S. Ebd.

¹⁷⁸ Vgl. Ebd.

Die Abbildung greift außerdem die speziellen Bedingungen für Telemedienangebote auf, die der Gesetzgeber aus dem EU-Beihilfekompromiss übernahm, der 2007 unter anderem zwischen der Bundesrepublik und der EU-Kommission geschlossen wurde¹⁷⁹. Zu diesen Bedingungen gehören folgende Punkte: Allen Bevölkerungsgruppen ist die Teilhabe an der Informationsgesellschaft zu ermöglichen, Minderheiten ist Zugang zu den Angeboten zu verschaffen, Bürgern ist der Nutzen der digitalen Dienste zu bieten, Medienkompetenz der Gesellschaft ist zu fördern und die Rundfunksendungen sind dem Wandel der Mediennutzung entsprechend zu begleiten¹⁸⁰. In der amtlichen Begründung zum zwölften RÄStV geben die Rundfunkanstalten sich selbst noch vor, in Zukunft als „verlässlicher Führer durch das Internet („trusted guide“) dienen zu wollen.“¹⁸¹ An dieser Stelle ergänzt der Soziologe und Politologe Peter Badura noch einen interessanten Zusatz: Er sieht in dem Funktionsauftrag speziell für die öffentlich-rechtlichen Telemedien nicht nur ein begleitendes Programmangebot, sondern durchaus die Möglichkeit eines eigenständigen und originären Angebots. Außerdem könne der dem Rundfunk zugewiesene Auftrag angesichts der digitalen Konvergenz der Medien und der fortentwickelten Nutzungsgewohnheiten des Publikums allein durch Live-Streaming und zeitversetzte Darbietung der herkömmlichen Rundfunkprogramme und auf einzelne Programmanwendungen bezogene Begleit-Telemedien nicht sachgerecht erfüllt werden¹⁸². Um diese Bedingungen erfüllen zu können, bedarf es schließlich der Konkretisierung durch Programmleitlinien und Telemedienkonzepte, die den an die öffentlich-rechtlichen Rundfunkanstalten gestellten Funktionsauftrag schließlich in die Praxis umsetzen.

4.1.1 Problematiken bei der Ausgestaltung des Auftrags

Der 12. RÄStV brachte dem öffentlich-rechtlichen Rundfunk nicht nur eine Legitimierung der eigenen Telemedienangebote im Netz, sondern auch neue Regulierungen und Einschnitte, die sowohl die Wettbewerbsfähigkeit der Mitbewerber im Medienmarkt, als auch den Gebührenzahler schützen sollen. Gerade diese Regulierungen wirken sich auf die

¹⁷⁹ Vgl. Held, Thorsten: Online-Angebote öffentlich-rechtlicher Rundfunkanstalten – Eine Untersuchung des verfassungsrechtlich geprägten und einfachgesetzlich ausgestalteten Funktionsauftrags öffentlich-rechtlichen Rundfunks im Hinblick auf Internet-Dienste, Baden-Baden, 2008, S. 359.

¹⁸⁰ Vgl. Todsén: Grenzen gebührenfinanzierter Telemedien, S. 122 f.

¹⁸¹ S. Amtliche Begründung zum 12. RÄStV, S. 18 f.

¹⁸² Vgl. Badura, Peter: Die öffentlich-rechtlichen Rundfunkanstalten bieten Rundfunk und Telemedien an, In: Archiv des öffentlichen Rechts (Band 134), hrsg. von Udo di Fabio u.a., Bonn, 2009, S. 241.

Mediatheken aus, da der dort ausgespielte Content stark von den Einschränkungen betroffen ist.

Als Grund für die Regulierung des öffentlich-rechtlichen Rundfunks gibt der Gesetzgeber nicht mehr nur die Knappheit der Übertragungswege an, die für Telemedienangebote so gut wie keine Gültigkeit mehr hat, sondern die „Sonderheit des Rundfunks“¹⁸³, die sich aus verschiedenen Faktoren zusammensetzt. Dazu gehören die Breitenwirkung des Rundfunks, dessen Aktualität und seine Suggestivkraft. Die Breitenwirkung des Rundfunks ergibt sich aus seiner großen Reichweite innerhalb der Bevölkerung und dem damit sehr großen Beeinflussungspotential für die öffentliche Meinungsbildung. Die Aktualität des öffentlich-rechtlichen Rundfunks folgt aus der zeitgleichen und schnellen Verbreitung seiner Inhalte. Seine besondere Suggestivkraft nimmt der Rundfunk daher, verschiedene Kommunikationsformen wie Text oder Bewegtbild miteinander zu kombinieren und seinem Content damit den Anschein einer hohen Authentizität zu verleihen¹⁸⁴.

Die Verbindung des damals geplanten Telemedienangebots des öffentlich-rechtlichen Rundfunks mit der Sonderheit des Rundfunks und der Gebührenfinanzierung rief zudem die privaten Medienanbieter auf den Plan, die auf eine paradoxe Gefahr hinwiesen: Ein eigenständiges öffentlich-rechtliches Telemedienangebot bedrohe die Meinungsvielfalt in der Medienlandschaft¹⁸⁵. Als Begründung gaben die Anbieter an, dass ein kostenfreies öffentlich-rechtliches Nachrichtenangebot die Nutzer von den teils kostenpflichtigen Angeboten der Privatanbieter wegziehen und so ein Meinungsmonopol entstehen könnte. Außerdem sei durch den ausbleibenden Traffic auf den Seiten privater Anbieter deren Existenz bedroht¹⁸⁶. Mehrere Studien konnten allerdings keine negativen Auswirkungen der öffentlich-rechtlichen Angebote auf die Angebote der privaten Anbieter feststellen¹⁸⁷. Dennoch führte der Gesetzgeber zum Schutz der privaten Anbieter und auch der Gebührenzahler folgende Regulierungen speziell für die öffentlich-rechtlichen Telemedienangebote ein: Bestimmter Content für die Telemedienangebote darf nur mit Verweildauern auf die Online-Plattformen gestellt werden. Dazu gehören zum Beispiel große Events wie die Olympischen Spiele oder Übertragungen der ersten und zweiten

¹⁸³ Vgl. Neuhoff: Rechtsprobleme des Auftrags des öffentlich-rechtlichen Rundfunks im Online-Bereich, S. 64.

¹⁸⁴ Vgl. Ebd.

¹⁸⁵ Vgl., Ebd., S. 75.

¹⁸⁶ Vgl. Ebd.

¹⁸⁷ Vgl. Ebd.

Bundesliga genauso wie tagesaktuelle Nachrichten¹⁸⁸. Erstgenannte Sportereignisse dürfen mit einer Verweildauer von 24 Stunden ausgespielt werden, sendungsbezogene Medien wie Nachrichten dürfen sieben Tage – mit einer Fristverlängerung, die im Einzelfall geprüft werden kann, aber auch länger – online verbleiben¹⁸⁹. Endet die Verweildauerfrist, muss sichergestellt werden, dass die Angebote offline gehen. Mit den Verweildauern sollen private Anbieter, die zum Beispiel Sportübertragen kostenpflichtig anbieten, wettbewerbsfähig bleiben können und außerdem auch der Gebührenzahler vor zukünftigen Erhöhungen des Rundfunkbeitrags geschützt werden, die zum Beispiel aus gestiegenen Lizenzgebühren für die Übertragungsrechte der Bundesliga resultieren könnten¹⁹⁰. Außerdem ist die Ausstrahlung von angekauften Serien und Spielfilmen auf den eigenen Online-Angeboten des öffentlich-rechtlichen Rundfunks nicht erlaubt¹⁹¹. Ausgenommen sind hier allerdings Produktionen, die der öffentlich-rechtliche Rundfunk selbst in Auftrag gegeben hat wie zum Beispiel „Rote Rosen“ oder auch „Der Tatortreiniger“.

Dazu verbietet der Gesetzgeber auch Werbung und Sponsoring auf den Telemedienangeboten des öffentlich-rechtlichen Rundfunks. Anders als im klassischen Rundfunk ist auf den Online-Angeboten jegliche Form von Werbung verboten – ohne Ausnahme oder Schlupflöcher¹⁹².

Zudem gibt der Gesetzgeber im 12. RÄStV eine Negativliste heraus, die bestimmte Formen von Angeboten vom Auftrag des öffentlich-rechtlichen Rundfunks fernhält. Durch die Negativliste werden E-Commerce-Angebote wie Anzeigenportale, Branchenregister, Partner-, Kontakt- und Stellenbörsen, Bewertungsportale für Dienstleistungen, Business-Networks, Wettportale sowie Routenplaner vom Auftrag ausgeschlossen¹⁹³. Mit den beiden letztgenannten Regulierungen schützt der Gesetzgeber vor allem private Anbieter solcher Dienste.

Außerdem führte der Gesetzgeber einen Drei-Stufen-Test für neue und veränderte Telemedienangebote ein, die in Zukunft nach bestimmten Kriterien geprüft werden, um die

¹⁸⁸ Vgl. Neuhoff: Rechtsprobleme des Auftrags des öffentlich-rechtlichen Rundfunks im Online-Bereich, S. 83.

¹⁸⁹ Vgl. Ebd.

¹⁹⁰ Vgl. Ebd.

¹⁹¹ Vgl. Ebd., S. 84.

¹⁹² Vgl. Ebd., S. 106.

¹⁹³ Vgl. Ebd., S. 117 ff.

Zulassung für die Online-Angebote der öffentlich-rechtlichen Rundfunkanstalten zu erhalten¹⁹⁴.

4.2 Der Drei-Stufen-Test als Legitimierung der Online-Aktivität

Mit dem 12. RÄStV fiel die Selbstverpflichtung der öffentlich-rechtlichen Rundfunkanstalten, höchstens 0,75 Prozent des Gesamtbudgets für Online-Aktivitäten auszugeben¹⁹⁵. An die Stelle dieser Verpflichtung trat der Drei-Stufen-Test, mit dem neue und veränderte Telemedienangebote geprüft werden sollen.

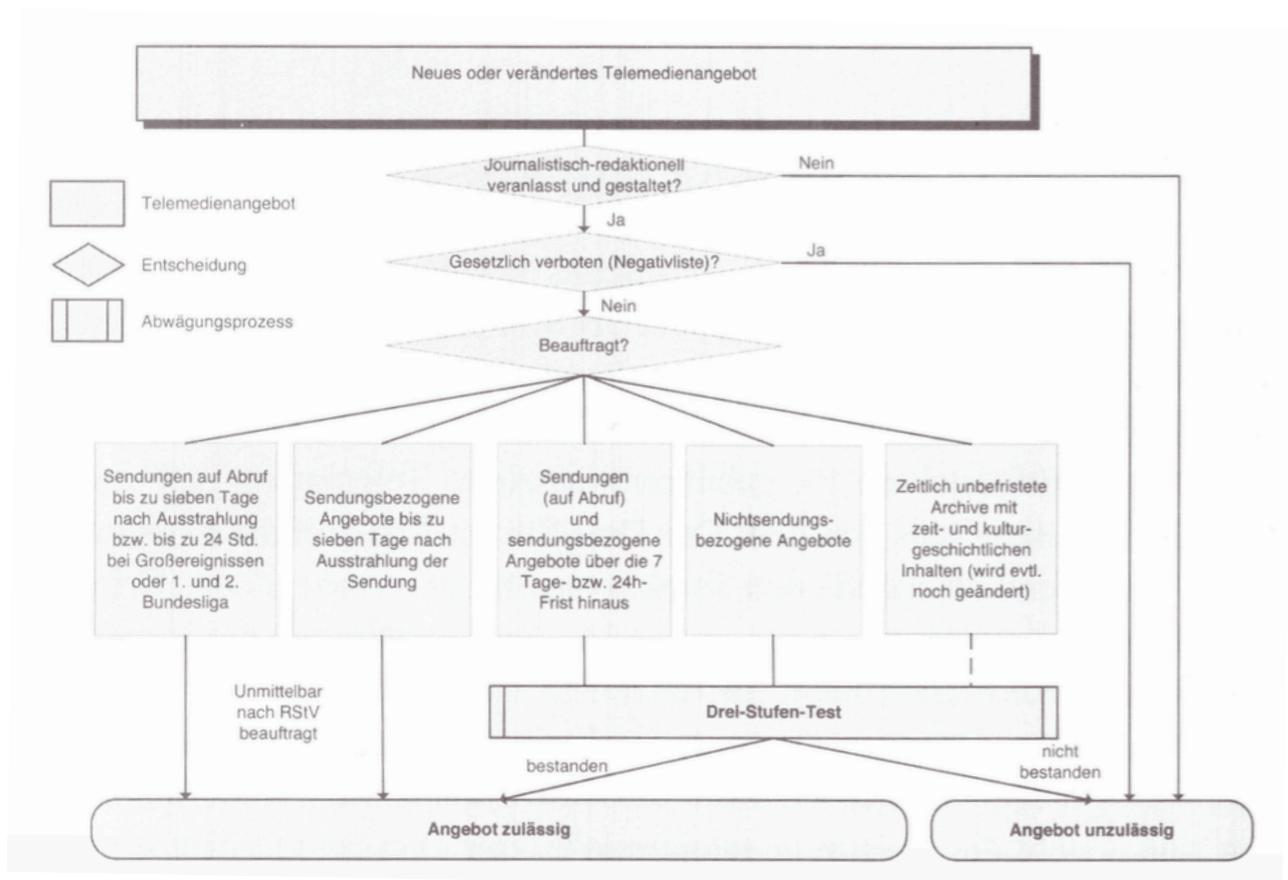


Abbildung 22: Rahmenbedingungen für die Durchführung eines Drei-Stufen-Tests

Grundlage eines jeden neuen Telemedienangebots ist – wie in Abbildung 22 zu sehen – immer das Telemedienkonzept. In diesem Konzept wird das neue Angebot ausführlich beschrieben und mögliche Auswirkungen auf den Auftrag des öffentlich-rechtlichen Rundfunks abgeschätzt¹⁹⁶.

Bevor ein Telemedienkonzept zu einem neuen oder veränderten Angebot für den Drei-Stufen-Test zugelassen wird, muss zunächst einmal geprüft werden, ob ein Drei-Stufen-

¹⁹⁴ Vgl. Neuhoff: Rechtsprobleme des Auftrags des öffentlich-rechtlichen Rundfunks im Online-Bereich, S. 131.

¹⁹⁵ Vgl. Todsen: Grenzen gebührenfinanzierter Telemedien, S. 131.

¹⁹⁶ Vgl. Ebd., S. 137.

Test überhaupt notwendig ist oder ob das Angebot womöglich unzulässig ist. Unzulässig ist es, wenn es kein journalistisch-redaktionell betreutes Angebot ist oder es auf der oben erwähnten Negativliste steht.

Im 12. RÄStV ist außerdem festgelegt, dass ein Angebot, das bereits im Auftrag für die öffentlich-rechtlichen Rundfunkanstalten explizit genannt ist, keinen Test mehr durchlaufen muss¹⁹⁷. Gleiches gilt für das „weniger umfangreiche Telemedienangebot“¹⁹⁸. Ändert sich zum Beispiel nur ein kleiner Teil des Mediatheksangebots, muss für diese Änderung kein kompletter Drei-Stufen-Test durchgeführt werden.

Was aber unter die Begrifflichkeit des „weniger umfangreichen Telemedienkonzepts“ fällt, wird mit der so genannten Aufgreifschwelle weiter definiert. Diese besteht aus Positiv- und Negativkriterien und legt fest, wann es sich um ein neues oder verändertes Telemedienangebot handelt und wann nicht¹⁹⁹. Dabei reicht es, ein Positivkriterium zu erfüllen, damit ein neues oder verändertes Telemedienangebot vorliegt. Werden Negativkriterien erfüllt, ist das Angebot noch nicht so weit neu oder verändert, dass die Aufgreifschwelle für einen Test erreicht wäre. Folgende Positivkriterien sind vorgegeben²⁰⁰:

1. Es muss eine grundlegende Änderung der thematisch-inhaltlichen Ausrichtung des Gesamtangebots vorliegen (zum Beispiel ändert sich das Thema des Gesamtangebots von einem Unterhaltungsangebot zu einem allgemeinen Wissensangebot)
2. Es muss eine substantielle Änderung der Angebotsmischung vorliegen (zum Beispiel der Wechsel von einem unterhaltungsorientierten zu einem informationsorientierten Angebot)
3. Es muss eine Veränderung der Zielgruppe des Angebots vorliegen (zum Beispiel der Wechsel von einem Senioren- zu einem Kinderprogramm)
4. Es muss eine wesentliche Steigerung des Aufwands für die Erstellung eines Angebots vorliegen (zum Beispiel in Form von stark gestiegenen Kosten oder zeitlichem Aufwand)

Folgende Negativkriterien sind vorgegeben²⁰¹:

¹⁹⁷ Vgl. Neuhoff: Rechtsprobleme des Auftrags des öffentlich-rechtlichen Rundfunks im Online-Bereich, S. 239 f.

¹⁹⁸ S. Ebd., S. 239.

¹⁹⁹ Vgl. Ebd., S. 240 f.

²⁰⁰ Vgl. Ebd., S. 240.

²⁰¹ Vgl. Ebd., S. 240 f.

1. Veränderung oder Neueinführung einzelner Elemente, Weiterentwicklung einzelner Formate ohne Auswirkung auf die Grundausrichtung des Angebots
2. Veränderung des Designs ohne direkte Einwirkung auf die Inhalte des Angebots
3. Verbreitung bereits bestehender Telemedien auf neuen technischen Verbreitungsformen
4. Weiterentwicklung im Zuge der technischen Entwicklung auf bereits bestehenden Plattformen
5. Weiterentwicklung aufgrund gesetzlicher Verpflichtungen (zum Beispiel die Barrierefreiheit)
6. Änderungen im Bereich der sendungsbezogenen Telemedienangebote, sobald sich die Ausrichtung des begleiteten Fernsehprogramms grundlegend ändert
7. Vorliegen einer zeitlichen Beschränkung (zum Beispiel in Form von Verweildauern)
8. Vorliegen eines Testbetriebs (das heißt, das Angebot besteht maximal zwölf Monate, steht nur einem eingeschränkten Nutzerkreis zur Verfügung und ist räumlich begrenzt)

Liegt ein Angebot über der Aufgreifschwelle für ein neues oder verändertes Telemedienangebot, muss es zwingend einen Drei-Stufen-Test durchlaufen, um zugelassen zu werden.

4.2.1 Die drei Stufen des Tests

Im Drei-Stufen-Test wird – wie es der Name bereits vermuten lässt – das vorliegende Telemedienangebot auf drei verschiedenen Stufen geprüft.

Die erste Stufe dient beim Drei-Stufen-Test der Feststellung, ob das neue oder veränderte Telemedienangebot dem Auftrag des öffentlich-rechtlichen Rundfunks entspricht oder nicht. Dabei gilt es vor allem zu klären, ob das Angebot gesellschaftliche, also demokratische, soziale und kulturelle Bedürfnisse erfüllt. Dazu gehören zum Beispiel folgende Fragen²⁰²: Wer passt in die Zielgruppe der Mediathek? Welche Funktion erfüllt das Angebot für die Nutzer? Wie wird das Angebot mit Blick auf die Dynamik in diesem Bereich in einigen Jahren aussehen? Wie wirkt sich das Angebot auf die Konkurrenz mit den Privaten aus? Bei der Beantwortung dieser Fragen helfen zum Beispiel empirische Untersuchungen der Marktforschung weiter. Beachtet wird das Angebot aber nicht nur vor dem Hintergrund der Empirie, sondern auch vor der Funktion des öffentlich-rechtlichen

²⁰² Vgl. Todsen: Grenzen gebührenfinanzierter Telemedien, S. 138.

Rundfunks. Wie im Kapitel zum Auftrag des Rundfunks bereits angeführt, soll der öffentlich-rechtliche Rundfunk auch Angebote zur Bildung und zur Wahrung der Meinungsvielfalt bereitstellen und die Neuorientierung der gesellschaftlichen Kommunikation unterstützen²⁰³. Gerade letztgenannte Punkte, die auf ein sehr exploratives Feld führen, werden von den privaten Anbietern nur selten erfüllt, da sie sehr stark an Quoten gebunden sind.

Bestätigt sich das gesellschaftliche Bedürfnis für das neue oder veränderte Telemedienangebot, muss in der ersten Stufe außerdem noch untersucht werden, ob der öffentlich-rechtliche Rundfunk zur Befriedigung des Bedürfnisses überhaupt zuständig ist oder ob es schon von privaten Anbietern erfüllt wird²⁰⁴.

Damit ist der Übergang zur zweiten Stufe fließend: Sie besteht aus mehreren Phasen und soll klären, ob das zu prüfende Angebot in qualitativer Hinsicht zum publizistischen Wettbewerb beiträgt. Bei dieser Untersuchung müssen drei Punkte berücksichtigt werden²⁰⁵: 1. Die Quantität und die Qualität der vorhandenen, frei zugänglichen Angebote. 2. Die marktlichen Auswirkungen des Angebots. 3. Die meinungsbildende Funktion des Angebots unter Berücksichtigung vorhandener Angebote.

Die zweite Stufe unterteilt sich in drei verschiedene Phasen, die Erkenntnisse zu den oben genannten Punkten hervorbringen sollen²⁰⁶:

- Marktanalyse: Hier wird die Situation im betreffenden Marktsegment qualitativ und quantitativ erfasst, um Auswirkungen des öffentlich-rechtlichen Angebots auf private prüfen zu können. So soll ausgeschlossen werden, dass durch das öffentlich-rechtliche Angebot bereits bestehende Angebote aus dem Markt verdrängt werden oder neue, höhere Markteintrittsbarrieren geschaffen werden.
- Publizistischer Mehrwert: Hier soll der publizistische Mehrwert des Angebots in Bezug zu den Qualitätskriterien des öffentlich-rechtlichen Rundfunks gesetzt werden (zum Beispiel Informationsvielfalt, Transparenz, Objektivität, Vollständigkeit, Verständlichkeit, Aktualität)²⁰⁷.

²⁰³ Vgl. Todsen: Grenzen gebührenfinanzierter Telemedien, S. 139.

²⁰⁴ Vgl. Ebd., S. 140.

²⁰⁵ Vgl. Ebd., S. 142.

²⁰⁶ Vgl. Ebd., S. 143 ff.

²⁰⁷ Dieser Punkt birgt großes Konfliktpotential zwischen öffentlich-rechtlichen und privaten Anbietern, da immer im Einzelfall ausgelotet werden muss, ob ein Marktversagen vorliegt oder nicht. Diese Hürde wurde auf Drängen der privaten Anbieter in den RStV aufgenommen.

- Abwägung der gewonnenen Erkenntnisse: Hier steht die Frage im Mittelpunkt, ob die in Phase eins und zwei herausgearbeiteten Erkenntnisse in Konflikt mit der Presse- und Rundfunkfreiheit stehen.

Auf der dritten Stufe des Tests muss das Angebot schließlich unter dem Gesichtspunkt der Finanzierung untersucht werden. Hier steht die Frage im Mittelpunkt, ob der finanzielle Aufwand für das zu prüfende Angebot in einem angemessenen Verhältnis zum öffentlichen Nutzen steht²⁰⁸.

4.2.2 Ablauf des Drei-Stufen-Tests

Der Ablauf des Drei-Stufen-Tests wird in folgender Abbildung sehr gut dargestellt:

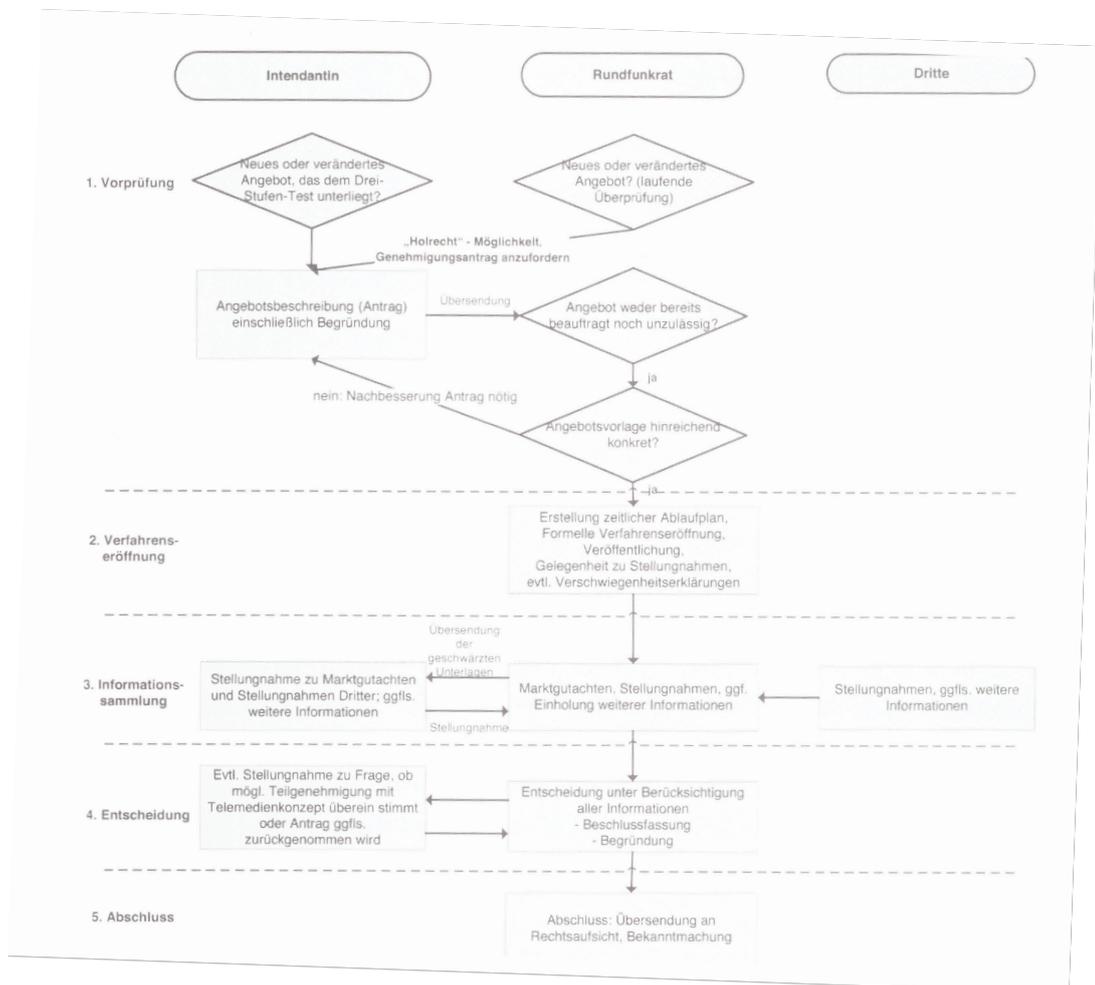


Abbildung 23: Ablauf des Drei-Stufen-Tests

Zu Beginn des Tests muss das Telemedienkonzept für das zu prüfende Angebot erstellt werden. Hier muss nicht nur das Angebot selbst beschrieben werden, sondern auch auf die marktlichen Auswirkungen des Angebots eingegangen werden. Außerdem gehören auch

²⁰⁸ Vgl. Todsen: Grenzen gebührenfinanzierter Telemedien, S. 156.

Aufstellungen zu den Kosten des Angebots in das Konzept²⁰⁹. Ist das Telemedienkonzept fertig gestellt, wird es vom jeweiligen Intendanten an den Rundfunk- oder Verwaltungsrat der entsprechenden Landesrundfunkanstalt geleitet, der für die Prüfung des Angebots zuständig ist²¹⁰.

In der Vorprüfung wird vom Rundfunkrat überprüft, ob das Konzept korrekt erstellt wurde und ob es ein ähnliches Angebot bereits gegeben hat. Nimmt der Rat das Konzept an, wird das Testverfahren eröffnet, indem ein Zeitplan erarbeitet wird und einige formelle Punkte abgearbeitet werden. Nun beginnt der Rundfunkrat damit, Informationen zum vorliegenden Telemedienkonzept in Form von Marktgutachten und Stellungnahmen einzuholen. Hier haben sowohl der Intendant als auch Dritte die Möglichkeit eigene Stellungnahmen einzureichen. Unter Berücksichtigung aller gesammelten Informationen trifft der Rundfunkrat darauf eine Entscheidung für oder gegen das Telemedienangebot. Die Entscheidung des Rates wird schließlich an die Rechtsaufsicht zur rechtlichen Prüfung übermittelt. Zudem wird die Entscheidung im amtlichen Verkündungsblatt öffentlich gemacht.

4.3 Der Drei-Stufen-Test der NDR Mediathek

Noch vor Inkrafttreten des 12. RÄStV beschlossen die Verantwortlichen beim NDR, das Angebot einer Mediathek einem Drei-Stufen-Test zu unterziehen, um Erfahrungen mit dem Ablauf des damals neuen Tests zu sammeln²¹¹. Ausgangspunkte des Tests waren die Angebotsbeschreibung der Mediathek sowie das Telemedienkonzept.

4.3.1 Das Telemedienkonzept der NDR Mediathek

Das Telemedienkonzept für die Mediathek orientiert sich stark an der Struktur des Drei-Stufen-Tests. Deshalb beschreibt das Konzept auf den ersten Seiten den Beitrag der Mediathek zur Erfüllung der demokratischen, sozialen und kulturellen Bedürfnisse der Gesellschaft. Zuerst wird die Zielgruppe genauer definiert: Die Mediathek richte sich an alle Zuschauer und Zuhörer der Rundfunk- und Hörfunkprogramme, die Inhalte zeit- und ortssouverän abrufen möchte. Die Mediathek solle vor allem junge Menschen ansprechen²¹². Darüber hinaus sei die Mediathek ein Schaufenster der publizistischen Leistungen von Hörfunk und erhöhe die Nachhaltigkeit der Programmangebote des

²⁰⁹ Vgl. Todsen: Grenzen gebührenfinanzierter Telemedien, S. 163.

²¹⁰ Vgl. Ebd., S. 164.

²¹¹ Vgl. Vorlage zur 370. Sitzung des NDR Rundfunkrates am 28. November 2008 – Einleitung eines Drei-Stufen-Tests für die NDR Mediathek, Hamburg, 2008, S. 2.

²¹² Vgl. Ebd., S. 3.

NDR²¹³. Über klassische On-Demand-Nutzung, das Abonnieren von Podcasts, Livestreams und eine Suche-Funktion, mit der alle Inhalte der Mediathek durchsucht werden könnten, stelle sie außerdem ein sehr breites Angebotsspektrum für den Nutzer zur Verfügung²¹⁴. Einen weiteren wichtigen Punkt im Telemedienkonzept nimmt das Verweildauerkonzept ein. Die Verweildauern sind – wie weiter oben beschrieben – eine rechtliche Vorgabe, die vor allem dem Schutz privater Mitwettbewerber dienen soll. Im vom NDR erarbeiteten Konzept zu den Verweildauern ist genau festgelegt, welche Inhalte wie lange in der Mediathek stehen dürfen²¹⁵.

- 24 Stunden: Inhalte zu Sport-Großereignissen oder zur 1. und 2. Bundesliga.
- 7 Tage: Inhalte mit tagesaktuellem oder nachrichtlichem Hintergrund, die relativ schnell wieder veraltet sind.
- 3 Monate: Inhalte zu saisonal bedingten Themen. Darunter fallen zum Beispiel Tests von Winterreifen oder Berichte über Musikfestivals. Auch Buch- und Kinotipps fallen unter diese Verweildauerfrist.
- 6 Monate: Inhalte, die Hintergründe zu aktuellen politischen, gesellschaftlichen, wirtschaftlichen und kulturellen Entwicklungen liefern – abseits der täglichen Nachrichten. Darunter fallen zum Beispiel Beiträge der Sendung „Weltbilder“.
- 12 Monate: Ausgewählte Inhalte aus den Magazinsendungen zu bedeutsamen politischen, wirtschaftlichen, kulturellen, norddeutschen oder gesellschaftlichen Ereignissen. Inhalte, die Menschen und Landschaften aus Norddeutschland vorstellen. Inhalte aus Programmschwerpunkten wie den ARD/NDR-Themenwochen. Inhalte aus Wissensformaten und Dokumentationen. Inhalte aus Comedyserien und Unterhaltungsangeboten.
- Zeitlich unbefristet: Inhalte mit nachhaltig dokumentarischem Charakter. Inhalte mit zeit- und kulturgeschichtlicher Bedeutung. Inhalte in niederdeutscher Sprache oder anderen regionalen Dialekten. Inhalte, die in Zusammenhang mit einem besonderen Ereignis (Jahrestag, Todestag) stehen.

Abschließend kommt das Konzept zu dem Schluss, dass die Mediathek einen Beitrag zu den demokratischen, kulturellen und sozialen Bedürfnissen der Gesellschaft leistet, da die Beiträge aus den Rund- und Hörfunkprogrammen aus der Flüchtigkeit des Hörfunks und Rundfunks entzogen und in der Mediathek konserviert würden. Durch die Möglichkeit,

²¹³ Vgl. Vorlage zur Einleitung eines Drei-Stufen-Tests, S. 3.

²¹⁴ Vgl. Ebd.

²¹⁵ Vgl. Ebd., S. 5 f.

diese Inhalte immer wieder abzurufen, leiste die Mediathek einen wichtigen Beitrag zur gesellschaftlichen Kommunikation und zur individuellen Meinungsbildung. Außerdem stärke sie die Nachhaltigkeit der öffentlich-rechtlichen Programme und die Verbreitung von unabhängigen Informationen sowie den transparenten Zugriff auf die Programminhalte. Daher sei das Angebot einer Mediathek im Auftrag des öffentlich-rechtlichen Rundfunks enthalten²¹⁶.

Die zweite Stufe des Tests untersucht die Auswirkungen der Mediathek auf den publizistischen Wettbewerb. Daher wird im Konzept auch dieser Punkt ausgeführt: So biete vor allem die Suchfunktion einen publizistischen Mehrwert für den Nutzer, da er so direkt zum einzelnen Beitrag springen kann und den Beitrag nicht in der ganzen Sendung suchen muss²¹⁷. Zudem erstelle die Mediathek Themen-Dossiers mit Video- und Audioinhalten, die Inhalte multimedial verknüpfen und so einen leichten Einstieg in die nicht-lineare NDR-Welt bieten würden²¹⁸. Außerdem gewährleiste sie als einzige Mediathek im Markt die regionale, hochwertige Berichterstattung mit norddeutschem Bezug²¹⁹.

Zu den marktlichen Auswirkungen beauftrage der NDR mit Hardy Gundlach, der Informations- und Medienökonomie an der HAW Hamburg lehrt, einen externen Gutachter. Gundlach äußert in seinem Gutachten nur wenige Bedenken bezüglich der Marktauswirkungen der Mediathek: Die größten marktlichen Auswirkungen habe die Mediathek auf die werbefinanzierten On-Demand-Dienste, die um so größer würden, je mehr Überschneidungen es mit den Inhalten der NDR Mediathek gebe²²⁰. Dennoch würden aber die negativen Auswirkungen für die „Konsumentenwohlfahrt“²²¹ bei einem Fernbleiben der Mediathek vom Markt die Bedenken überwiegen. Grund dafür sei die Qualitätssteigerung der Inhalte, die eine öffentlich-rechtliche Mediathek für den Markt nach sich zöge²²². Mit diesem Argument zerstreut Gundlach auch ähnliche gerichtete Bedenken zu den Auswirkungen der Mediathek auf entgeltpflichtige Video-On-Demand-Anbieter sowie auf private regionale und lokale Videoplattformen²²³.

²¹⁶ Vgl. Vorlage zur Einleitung eines Drei-Stufen-Tests, S. 7.

²¹⁷ Vgl. Ebd., S. 8.

²¹⁸ Vgl. Ebd.

²¹⁹ Vgl. Ebd. Diesen Service böten zwar auch Zeitungen im Netz an, allerdings nicht so umfassend wie der NDR.

²²⁰ Vgl. Gundlach, Hardy: Medienökonomisches Gutachten zu den marktlichen Auswirkungen der geplanten NDR Mediathek, Hamburg, 2009, S. 179.

²²¹ Vgl. Ebd.

²²² Vgl. Ebd.

²²³ Vgl. Ebd., S. 169 ff.

Zur zweiten Stufe konstatieren das Telemedienkonzept sowie Gundlachs Gutachten also, dass es bei der Einführung der Mediathek zwar marktliche Bedenken gebe, diese aber vom Nutzen für die Rezipienten überwogen würden. Somit werde der Markt nicht negativ beeinträchtigt. Außerdem sei die von anderen Mitbewerbern bisher nicht erreichte, qualitativ hochwertige Berichterstattung ein gesellschaftlicher Mehrwert. Gemäß der dritten Stufe des Tests, die die Finanzierung des Angebots untersucht, führt das Telemedienkonzept zur Mediathek auch die Kosten der Mediathek auf: Bis 2012 seien 1,25 Millionen Euro inklusive Lizenzkosten veranschlagt, die durch Umschichtungen in anderen Bereichen freigemacht werden sollen²²⁴.

4.3.2 Ergebnis des Drei-Stufen-Tests

Ende März 2009 fasste der Rundfunkrat des NDR in seiner 372. Sitzung einen Beschluss zur NDR Mediathek. Auch dieses Schriftstück orientiert sich stark an den drei Stufen des Testverfahrens.

Der Rundfunkrat teilt die im Telemedienkonzept beschriebene Meinung, dass die Mediathek durch ihre thematische Vielfalt und die Möglichkeit der zeit- und ortssouveränen Nutzung die demokratischen, sozialen und kulturellen Bedürfnisse der Gesellschaft bedient und außerdem ein kommunikatives Bedürfnis der Gesellschaft erfüllt, das bisher marktlich noch nicht abgedeckt wird²²⁵. Das Angebot der Mediathek sei erforderlich, um die bisherige Funktion des NDR auch unter dem gewandelten medialen Verhalten der Nutzer erfüllen zu können²²⁶.

Auf der zweiten Stufe des Testverfahrens kommt der Rundfunkrat zu dem Schluss, dass die Mediathek in qualitativer Hinsicht zum Wettbewerb beitrage. Sie halte das Qualitätsniveau auf dem Markt aufrecht, da sie dem Nutzer die Möglichkeit des Qualitätsvergleichs mit anderen Mediatheken oder On-Demand-Angeboten biete²²⁷.

Zudem trage das Verweildauerkonzept dazu bei, Angebote Dritter zu schützen. Dazu sei das Konzept angemessen formuliert²²⁸.

Zudem leiste die Mediathek durch das Angebot der zeit- und ortssouveränen Nutzung, der Such- und Recherchefunktion sowie der Bündelung und Verknüpfung von Audio- und Videobeiträgen große publizistische Mehrwerte.

²²⁴ Vgl. Vorlage zur Einleitung eines Drei-Stufen-Tests, S. 9.

²²⁵ Vgl. Abschließende Beratung und Beschluss über das Angebot der NDR Mediathek, Hamburg, 2009, S. 2.

²²⁶ Vgl. Ebd.

²²⁷ Vgl. Ebd., S. 3.

²²⁸ Vgl. Ebd., S. 5.

Somit erkennt der Rundfunkrat auf der zweiten Stufe des Tests die Bedenken Gundlachs zwar an, folgt aber seiner Einschätzung, dass der Mehrwert der Mediathek für die Konsumentenwohlfahrt überwiegt. Somit stelle die Mediathek auch mit Blick auf den publizistischen Wettbewerb keine Gefährdung dar.

Auf der letzten Stufe prüft der Rundfunkrat die veranschlagte Finanzierung. Er erkennt an, dass der veranschlagte finanzielle Aufwand für die Mediathek „erforderlich und auskömmlich“²²⁹ sei, gibt aber gleichzeitig vor, dass in den Mediathek nur Beiträge gebündelt und aufbereitet werden sollen, jedoch keine exklusiven Beiträge für die Mediathek produziert werden sollen²³⁰.

Der Rundfunkrat kommt so zu dem Schluss, dass die NDR Mediathek vom öffentlich-rechtlichen Auftrag umfasst ist und in ihrer Ausgestaltung im Telemedienkonzept auch den Ansprüchen des 12. RÄStV gerecht wird²³¹.

4.4 Arbeitsprozess in der NDR Mediathek

Dreh- und Angelpunkt des Arbeitsprozesses in der NDR Mediathek ist das von der Hamburger subshell GmbH entwickelte CMS Sophora. Es verarbeitet die Eingaben der Redakteure und verwaltet Inhalte und Daten des NDR-Contents. In diesem Kapitel soll zunächst kurz ausgeführt werden, wozu ein CMS überhaupt genutzt wird, um im nächsten Schritt Sophora selbst anhand des Arbeitsablaufs innerhalb der Mediathek näher zu beschreiben. Dabei soll auch ein Bild der im Arbeitsprozess erzeugten Daten gezeichnet werden, auf die ein Recommender System zurückgreifen könnte.

4.4.1 Was ist ein CMS?

Ein Content-Management-System oder kurz CMS ist im Grunde genommen ein Informationssystem, das der Unterstützung bei der Erstellung und Verwaltung von Informationen dient. Ein CMS bietet dazu verschiedene Funktionalitäten: Zum einen für dispositive Aufgaben wie zum Beispiel der Inhaltsplanung oder der Prozesskontrolle, zum anderen für operative Aktivitäten wie der Erfassung, Bearbeitung und Publikation von Inhalten und auch für unterstützende Tätigkeiten wie zum Beispiel der Systemadministration²³². Mit dem Einsatz eines CMS lassen sich zum einen Kosten senken, indem die Effizienz bei der Contentproduktion gesteigert und die Durchlaufzeit der Produktion verkürzt wird, zum anderen die Qualität der bereitgestellten Inhalte

²²⁹ Vgl. Beschluss über das Angebot der NDR Mediathek, S. 6.

²³⁰ Vgl. Ebd.

²³¹ Vgl. Ebd., S. 7.

²³² Vgl. Bodendorf, Freimut: Daten- und Wissensmanagement, Heidelberg, 2006, S. 100.

steigern, indem vor allem den Layoutprozessen definierte Abläufe zugrunde gelegt werden und der bereitgestellte Content leichter zu pflegen und zu überarbeiten ist²³³.

Wesentliche Aufgabe eines CMS ist die Speicherung aller Inhalte, Metadaten und Steuerungsinformationen. Um diese Aufgabe erfüllen zu können, benötigt das System eine Zugriffsverwaltung, die Inkonsistenzen verhindert, wenn unterschiedliche Nutzer gleichzeitig Inhalte editieren möchten²³⁴. Auf die gespeicherten Inhalte kann mittels einer Suche zugegriffen werden, sodass Dokumente im CMS leicht auffindbar sind.

Außerdem bieten viele CMS eine Schnittstelle für Tools, mit denen zum Beispiel Daten ins CMS importiert werden können oder die spezialisierte Recherchen im Internet ermöglichen²³⁵.

Die Bündelung von Informationselementen erfolgt im CMS über zwei Stufen: Zuerst werden Texte, Audios, Videos und Bilder zu Beiträgen kombiniert, anschließend werden sie der Navigationshierarchie des Online-Angebots untergeordnet²³⁶.

Aus technischer Sicht lässt sich ein CMS in drei Hauptkomponenten unterscheiden²³⁷:

- Das Editorial System: Es umfasst viele Werkzeuge zur Erzeugung, Verarbeitung und Verwaltung von Content. Klassischerweise läuft die Bearbeitung über einen Editor, dessen Oberfläche aus gängigen Textbearbeitungsprogrammen bekannt ist.
- Das Content Repository: Hier werden die im Editorial System erstellten Inhalte und Layouts gespeichert. Dabei sind drei Aspekte wesentlich: Layouts und Inhalte müssen voneinander getrennt gespeichert werden, damit ein Layout für mehrere Inhalte genutzt werden kann und nicht nur mit einem speziellen Inhalt verknüpft ist. Mikro- und Makro-Strukturen müssen im Content Repository abgebildet sein, die es überflüssig machen, Inhalte immer wieder duplizieren und zu neuen Einheiten zusammenfügen zu müssen. Außerdem müssen im Content Repository dynamische Informationsarten wie Video und Audio eingebunden sein, um die Inhalte multimedial aufbereiten zu können.
- Das Publishing System: Dieses System hängt stark vom Zielmedium ab. Bei Online-Angeboten hat es zum Beispiel die Aufgabe, den Content auf einem für den Rezipienten zugänglichen Server bereitzustellen. Dabei sind Umsetzung der

²³³ Vgl. Bodendorf: Daten- und Wissensmanagement, S. 100.

²³⁴ Vgl. Ebd. Über die Verwaltung werden Informationselemente für die Dauer der Bearbeitung für andere Nutzer gesperrt. Erst nach dem Hochladen der veränderten Version ist sie wieder für andere Nutzer bearbeitbar.

²³⁵ Vgl. Ebd., S. 101.

²³⁶ Vgl. Ebd.

²³⁷ Vgl. Ebd., S. 102 ff.

Struktur in Layout und der Zeitpunkt der Erzeugung von Dokumenten zwei wesentliche Aspekte.

Die Grenzen zwischen den Systemen sind durchlässig: So lassen sich Inhalte im Content Repository zum Beispiel immer wieder in das Editorial System zurückholen, um sie zu überarbeiten oder zu aktualisieren und danach immer wieder über das Publishing System dem Rezipienten zur Verfügung stellen.

Das CMS, das neben dem NDR auch fast alle anderen öffentlich-rechtlichen Rundfunkanstalten einsetzen, heißt Sophora und wurde von der Hamburger Firma subshell entwickelt. Es weist im Vergleich zu klassischen CMS einige Besonderheiten auf. So gewährleistet es durch eine intelligente Zugriffsverwaltung, dass auch mehrere hundert Nutzer gleichzeitig mit dem System am Content arbeiten können – ohne dass es zu Datenverlusten kommt²³⁸. Eine weitere Funktion, die Sophora für viele öffentlich-rechtliche Anstalten interessant machte, ist die mächtige Suchfunktion, die es dem Nutzer erspart, sich in komplexen Ordnerstrukturen zurechtfinden zu müssen²³⁹. Zudem erleichtern simple Operatoren die Suche nach Content. Außerdem bietet Sophora die Möglichkeit, zusätzliche Add-Ons an das CMS anzudocken, um dem System zusätzliche Funktionen zu verleihen. So wurde beim NDR kürzlich der YouTube-Connector hinzugefügt, mit dem sich unter anderem Videobeschreibungen und –titel direkt in Sophora editieren lassen²⁴⁰. Ein weiterer Vorzug von Sophora ist eine mächtige Filterfunktion für Inhalte, die sich darüber zum Beispiel in Themenseiten zusammenfassen und online ausspielen lassen²⁴¹.

Vom Aufbau her gleicht Sophora anderen CMS. Hersteller subshell zeigt auf der Produktseite von Sophora eine Übersicht über die Bestandteile des Systems:

²³⁸ Vgl. Gespräch mit A. Svensson, [03:44 Minuten].

²³⁹ Vgl. Ebd., [03:21 Minuten].

²⁴⁰ Vgl. Ebd., [07:03 Minuten].

²⁴¹ Vgl. Ebd., [36:57 Minuten].

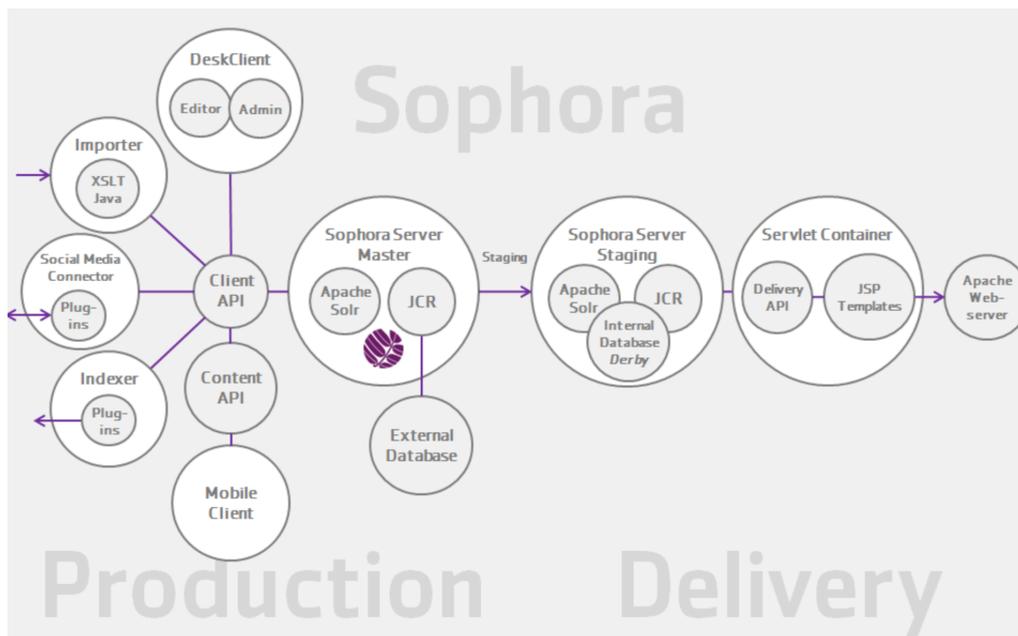


Abbildung 24: Logischer Überblick über das CMS Sophora

Die Abbildung zeigt sehr schön die weiter oben beschriebenen Hauptkomponenten eines CMS: Die aufgefächerten Bubbles ganz links zeigen das Editorial System, das neben Texttools auch die Add-Ons und zusätzliche Datenquellen zeigt, die sich an Sophora anbinden lassen. Die dort erzeugten oder gesammelten Daten werden zusammengeführt und im Content Repository gespeichert. Diese Komponente heißt in der Abbildung „Sophora Server Master“. Die dort gespeicherten Informationen werden nach der Freigabe schließlich über das Publishing System für den Rezipienten zugänglich gemacht. In der Abbildung ist der Schritt in den zwei großen Bubbles rechts vom „Sophora Server Master“ dargestellt. Den unter „Sophora Delivery“ zusammengefassten Schritt beschreibt subshell mit den Worten: „The Sophora Delivery coordinates the delivery of and provides access to the managed content.“²⁴²

4.4.2 Die Datenstruktur in der NDR Mediathek

Da der Begriff „Datenstruktur“ recht weit gefasst sein kann, bedarf es hier einer genaueren Einordnung. Die komplette Datenstruktur des CMS Sophora zu untersuchen böte vermutlich genug Stoff für eine eigenständige Arbeit, weshalb diese Untersuchung den Rahmen der vorliegenden Arbeit sprengen würde. Deshalb sind unter den Begriff „Datenstruktur“ für diese Arbeit zwei Ebenen gefasst: Zum einen beschreibt der Begriff die Struktur der in Sophora eingegebenen und gesammelten Daten und Metadaten für ein zu bearbeitendes Video, zum anderen gibt er an, welche Nutzungsdaten beim Aufruf eines

²⁴² Vgl. <http://www.subshell.com/en/sophora/documentation/delivery/>, zuletzt besucht am 28.05.2016 um 16:53 Uhr.

Videos in der Weboberfläche der Mediathek gemessen werden. Diese Daten würden bei einem Recommender System die Grundlage für die Berechnung der Empfehlung bilden. Dementsprechend lohnt sich eine genauere Betrachtung, welche Daten derzeit beim NDR zu Videos erfasst werden, um am Ende der Untersuchung einen möglichst präzisen Ist-Stand vorweisen zu können.

4.4.2.1 In Sophora erfasste Daten und Metadaten

Die in Sophora erfassten Metadaten zu jedem Video folgen einem vordefinierten Schema.

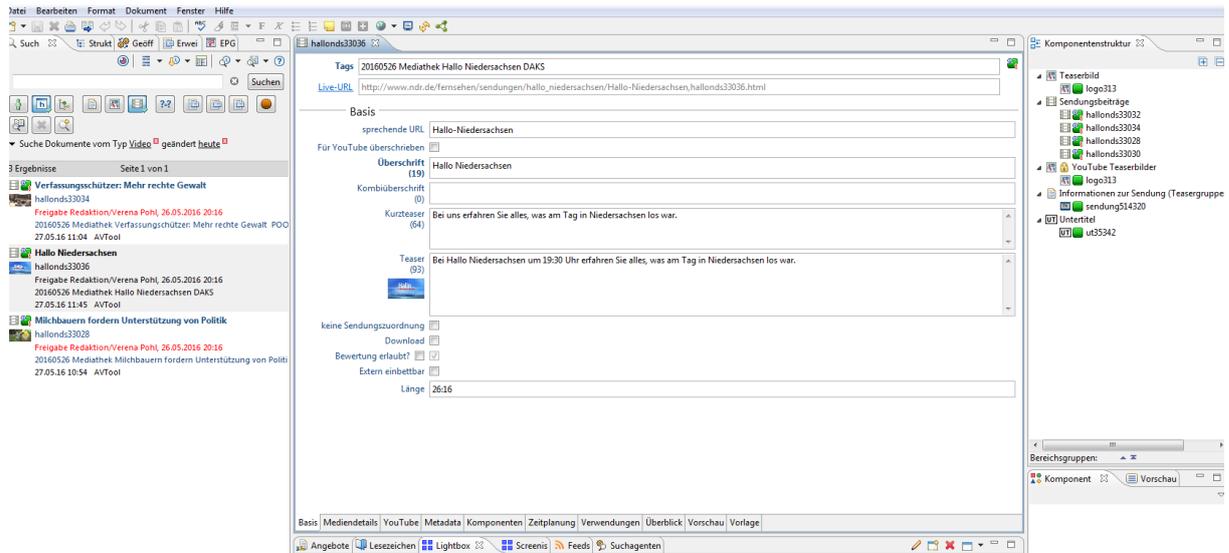


Abbildung 25: Dateneingabe in Sophora, „Basis“-Reiter der Video-Datei

Wird eine Videodatei zu Sophora hinzugefügt, wird der Datei eine einzigartige ID zugeordnet, über die die Datei solange wiederzufinden ist, wie sie im CMS existiert. Diese ID setzt sich aus einer alphanumerischen Zeichenfolge zusammen, die aus dem Stamm der Sendereihe und aus einer immer weiter fortlaufenden Zahl besteht. Der Stamm der Sendereihe ist meist deren Kurzform, das obige Beispiel zeigt die des Regionalmagazins Hallo Niedersachsen, dessen Stamm „hallonds“ ist. Der Stamm ist manuell anpassbar. Dazu werden die Tags des Videos erfasst, die für die interne Sophora-Suche relevant sind. Die übrigen, im „Basis“-Reiter erfassten Daten sind die Überschrift sowie Kurz- und Langteaser des Videos. Diese Daten dienen vor allem der Information des Nutzers, welches Video er in seinem Browser abgerufen hat. Zusätzlich lassen sich an dieser Stelle noch weitere Funktionen festlegen, zum Beispiel, ob das Video zum Download bereit stehen darf²⁴³, ob es bewertet werden darf oder ob es extern eingebettet werden darf. Auch

²⁴³ Diese Option muss bereits beim Export aus dem Videoschnitt-Tool ausgewählt worden sein, lediglich über Sophora funktioniert die Download-Bereitstellung noch nicht.

das Teaserbild wird im Reiter „Komponentenstruktur“ an das Video gehängt. Die Länge des Videos in Minuten übernimmt Sophora aus der Encodinganlage Odeon. Im „Metadaten“-Reiter werden zusätzliche Informationen zum Video erfasst.

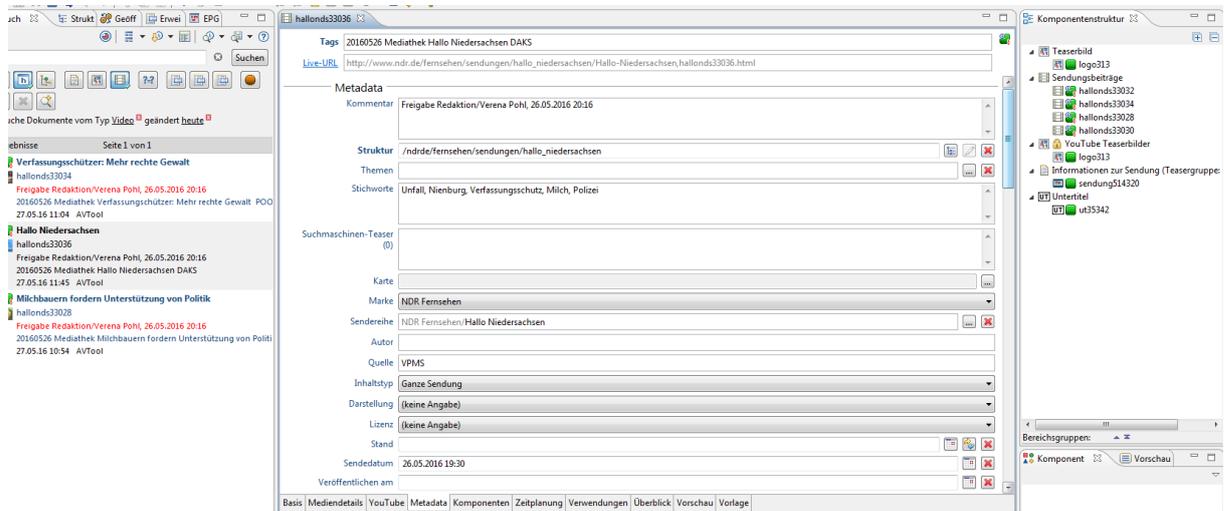


Abbildung 26: Dateneingabe in Sophora, „Metadaten“-Reiter der Datei (1. Abschnitt)

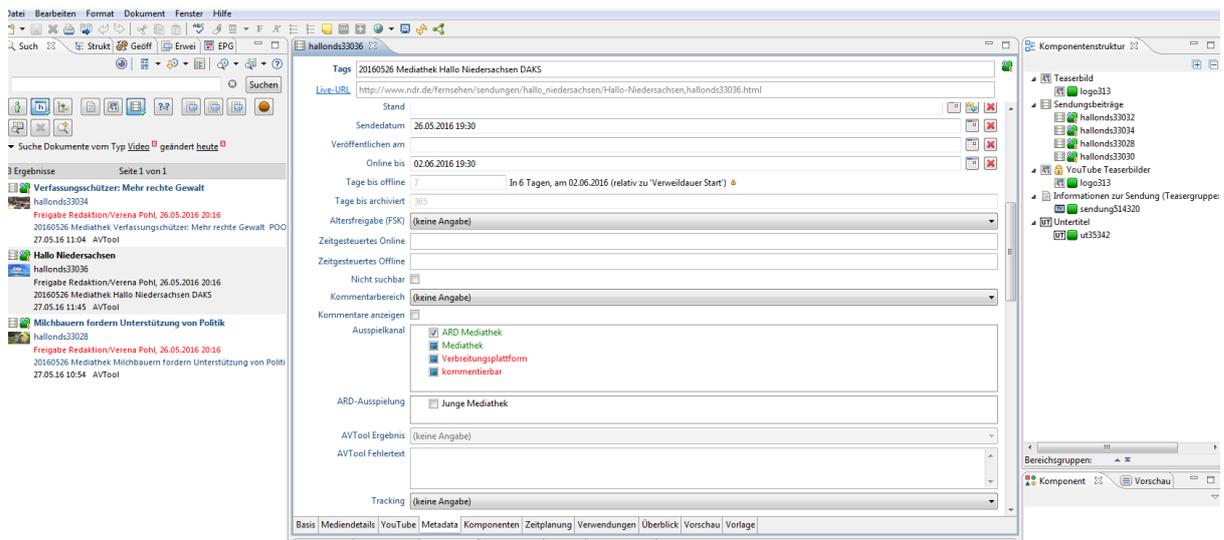


Abbildung 27: Dateneingabe in Sophora, „Metadaten“-Reiter der Datei (2. Abschnitt)

Die Abbildungen 26 und 27 zeigen alle möglichen Dateneingaben, die über den „Metadaten“-Reiter nach derzeitigem Stand erfasst werden können. Ganz oben ist in diesem Reiter das Kommentarfeld, in das bei den Videos immer die Freigaben der Redaktionen eingetragen werden. Das zweite Feld gibt den Speicherort des Videos in der Struktur des Content Repository von Sophora an. Diese Struktur folgt einem logischen Aufbau. Übergeordnet ist im vorliegenden Beispiel nrdde, das Sendungsvideo zum Magazin Hallo Niedersachsen liegt darunter im Unterordner fernsehen/sendungen. Im darauf folgenden Feld lässt sich das vorliegende Video mit fest vorgegebenen Themen verorten. Das Themenspektrum ist sehr fein ausdefiniert und dient weniger der genrespezifischen Einordnung von Texten und Videos, zum Beispiel nach Reportagen oder

Dokumentationen, sondern eher der themenspezifischen Zuteilung der Inhalte. So beinhaltet das Feld unter anderem die Über-Themen Nachrichten, Ratgeber, Gesundheit und Wirtschaft, die sich jeweils noch feiner untergliedern²⁴⁴. Im „Themen“-Feld sind mehrere Themen-Nennungen möglich. Feste Regeln, wie das „Themen“-Feld gefüllt werden soll, gibt es nicht.

Das „Stichworte“-Feld kann mittels freier Texteingabe gefüllt werden. Hier werden hauptsächlich Begriffe aufgeführt, die thematisch zum Video passen – eine feste Regel, welche und wie viele Stichworte vergeben werden müssen, existiert nicht. Das „Stichworte“-Feld ist bisher für die Sophora-Suche auf ndr.de in Verwendung und wird für externe Suchmaschinen ausgespielt²⁴⁵.

Die Felder zum Suchmaschinen-Teaser und zur Karte sind für Videos derzeit nur selten in Benutzung. Das „Marke“-Feld legt fest, welcher Marke das vorliegende Video zugeordnet wird. Neben dem „NDR Fernsehen“ gibt es hier zum Beispiel auch die Auswahlmöglichkeiten „Das Erste“ oder verschiedene NDR-Radiowellens.

Das „Sendereihe“-Feld gibt an, unter welcher Sendereihe das vorliegende Video läuft. Für die meisten Videos, die regelmäßig im NDR laufen, gibt es bestehende Sendereien, es gibt allerdings auch Videos, die keiner Sendereihe zugeordnet werden, da für sie keine Sendereihe existiert. Dazu gehören zum Beispiel Spielfilme, die nur für eine bestimmte Zeit in die Mediathek gestellt werden.

Das „Autor“-Feld wird ebenfalls nur bei manchen, aber nicht bei allen Videos gefüllt. Ob dort der Name des Autors erscheint, geben die Redaktionen in ihren Mails an das Mediatheks-Team vor (vgl. Kapitel 4.4.3).

Das „Quelle“-Feld wird automatisch gefüllt und gibt die Quelle an, aus der Sophora das Material erhalten hat.

Die drei folgenden Dropdown-Fenster werden nur bei wenigen Videos ausgefüllt: Der Inhaltstyp gibt an, ob es sich bei dem vorliegenden Video um die ganze Sendung oder einen Trailer handelt; bei normalen Beitragsvideos bleibt das Feld leer. Das Dropdown-Fenster zur Darstellung ist für Videos nicht relevant; im „Lizenz“-Fenster ist anzugeben, mit welcher Lizenz das vorliegende Video veröffentlicht werden darf. Hier muss zum Beispiel im Fall von „Sturm der Liebe“ die Geoblocking-Option gewählt werden, damit es außerhalb Deutschlands nicht abgerufen werden kann.

²⁴⁴ Vgl. <http://www.ndr.de/home/thema108.html>, zuletzt besucht am 29.05.2016 um 14:11 Uhr.

²⁴⁵ Vgl. Gespräch mit A. Svensson, [36:46 Minuten].

Die folgenden Felder dienen der zeitlichen Einordnung des Videos: Das „Stand“-Feld gibt den aktuellen Änderungsstand an, bleibt aber in den meisten Fällen leer. Das Sendedatum gibt den genauen Ausstrahlungstermin des Videos an – bei Einzelbeiträgen orientiert es sich immer am Sendedatum der zugehörigen ganzen Sendung. Das „Veröffentlichen am“-Feld dient dazu, ein vom Sendedatum unabhängiges Veröffentlichungsdatum einzugeben. Diese Funktion nutzt zum Beispiel extra3 öfter, um Videos vor dem eigentlichen Sendedatum zu veröffentlichen. Mit dem „Online bis“-Feld wird die Verweildauer des Videos in der Mediathek angegeben. Mit den beiden Feldern darunter lässt sich in Tagen festlegen, ob und wenn ja, nach wie vielen Tagen es als Archivinhalt gekennzeichnet wird. Mit dem Feld zur Altersfreigabe lässt sich einstellen, zu welcher Tageszeit Videos abgerufen werden können. Dies betrifft Inhalte, die ab 12, 16 oder 18 Jahren freigegeben sind und demnach nur zu bestimmten Tageszeiten abrufbar sein dürfen.

Die Felder zum zeitgesteuerten Online- und Offline-Gehen bieten die Möglichkeit, das Video automatisch zu einer bestimmten Zeit zu veröffentlichen oder zu depublizieren. Die darauf folgenden Felder bestimmen, wie der Nutzer mit dem Video interagieren kann: Über die „Nicht suchbar“-Funktion kann das Video aus der On-Page-Suche herausgenommen werden, mit den Funktionen zum Kommentarbereich wird festgelegt, ob Nutzer das Video kommentieren dürfen.

Der Ausspielkanal ist eine wichtige Funktion bei der Videoveröffentlichung und gibt an, auf welchen Kanälen das Video ausgespielt wird. An dieser Stelle entscheidet sich zum Beispiel, ob Videos auch in der ARD Mediathek auftauchen wird oder nicht.

Die drei übrigen Felder zum AVTool und zum Tracking sind für Videos nicht weiter relevant.

Neben den Metadaten zu den Videos, die direkt an das Video gehängt werden, laufen auch noch über eine andere Schnittstelle relevante Video-Informationen in Sophora ein. Dieses System nennt sich ProPlan. ProPlan ist das Programmplanungssystem vom NDR und bestückt den Electronic Program Guide – kurz EPG – mit zusätzlichen Informationen zu den Videos, die unter anderem in den Sendungsobjekten zu den Videos ausgespielt werden²⁴⁶. Diese Informationen beinhalten vor allem Details zu Redaktions- und Produktionsteams, die an der Erstellung der entsprechenden Sendungen beteiligt waren.

²⁴⁶ Vgl. Gespräch mit A. Svensson, [18:19 Minuten].

| | |
|------------------------------|--------------------------------|
| Redaktionsleiter/in | Manfred Schröter |
| Produktionsleiter/in | Edgar Rygol |
| Redaktion | Kai-Ove Kessler Dirk Külper |
| Moderation | Julia-Niharika Sen |
| Nachrichtenmoderation | Jens Riewa |

Abbildung 28: Daten aus dem ProPlan-Tool, ausgespielt auf ndr.de

Über ProPlan laufen allerdings auch zusätzliche Metadaten zum Video in Sophora ein. Dazu gehören unter anderem Informationen zu Sendeterminen, zu Reihentiteln und technische Metadaten wie das Format des Videos, ob es in Schwarz-Weiß oder in Farbe ausgegeben wird oder in welchem Tonformat das Video ausgespielt wird²⁴⁷. Ebenso wie in Sophora sind auch die eingegebenen Daten in ProPlan nicht allzu konsistent, sondern die Datenqualität variiert von Sendereihe zu Sendereihe²⁴⁸.

Der Umfang der erfassbaren Metadaten in Sophora ist leicht veränderbar. So lassen sich entweder vom Online-Service-Center des NDR bei Bedarf problemlos neue Felder in den „Metadaten“-Reiter in Sophora einfügen oder – bei tiefergehenden Veränderungen im CMS – eine Anforderung an subshell formulieren, die von den Entwicklern dann umgesetzt wird²⁴⁹. Sollten also für ein Recommender System zusätzliche Metadaten benötigt werden, sind sie mit relativ wenig Aufwand innerhalb kurzer Zeit zu ergänzen. Es bleibt nur die Herausforderung, dass diese Felder schnell und konsistent gefüllt werden müssen.

4.4.2.2 Durchs Tracking gesammelte Nutzungsdaten

Neben den aktiv in Sophora gesammelten und erzeugten Daten und Metadaten, werden beim NDR durch das Web-Tracking auch noch unabhängig von Redakteuren Daten aufgezeichnet. Diese Daten bilden das Nutzungsverhalten auf ndr.de und damit auch in der NDR Mediathek ab und werden über Cookies²⁵⁰ und Zählpixel²⁵¹ erfasst. Die Betreuung

²⁴⁷ Vgl. Gespräch mit A. Svensson, [18:25 Minuten].

²⁴⁸ Vgl. Gespräch mit M. Lahmann [13:45 Minuten].

²⁴⁹ Vgl. Ebd., [28:37 Minuten].

²⁵⁰ Ein Cookie ist eine Textinformation, die entweder über den Webserver an den Browser gesendet wird oder über eine Skript-Datei auf der Webseite erzeugt wird. Ein Cookie enthält nutzerbezogene Daten und hat eine variable Bestandsdauer, die angibt, wann es ausläuft.

²⁵¹ Ein Zählpixel ist eine meist 1x1 Pixel große Grafik, die über eine Skript-Datei auf der Webseite mit Nutzungsdaten gefüllt wird. So lassen sich Logdaten eines Nutzers für spätere Analysen aufzeichnen und auslesen.

des Web Trackings übernimmt für den NDR der Analytics-Dienstleister comScore sowie das branchenweit verbreitete INFOnline-System²⁵².

Die eingesammelten Nutzungsdaten werden schon während der Erhebung anonymisiert, weshalb eine Zuordnung zu einem bestimmten Nutzer nicht möglich ist²⁵³. Trotz der fehlenden Zuordnung können die Nutzungsdaten von ndr.de und speziell die der Videos in der Mediathek aber nützlich für eine Recommender Engine sein, weshalb an dieser Stelle genauer auf die Art und die Struktur der erhobenen Nutzungsdaten eingegangen werden soll.

Der NDR setzt mit den weiter oben beschriebenen Cookies und Zählpixeln zwei verschiedene Tracking-Techniken auf seinen Webseiten ein. Zunächst sollen die in Cookies erfassten Daten genauer untersucht werden: Auf ndr.de und speziell auf den Mediatheks-Seiten kommen nur sehr wenige Cookies zum Einsatz. Das wichtigste Cookie ist hier „JSESSIONID“, das jedem Nutzer eine ID zuordnet, die bis zum Ende seiner Session bestehen bleibt. Die ID ermöglicht es, einen Nutzer auch über mehrere Seitenaufrufe hinweg identifizierbar zu machen.

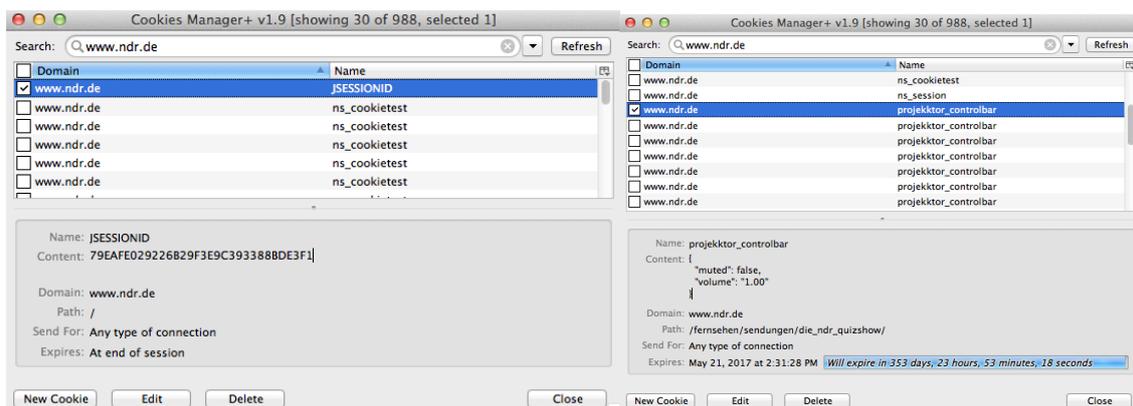


Abbildung 29: Cookie „JSESSIONID“ und „projekktor_controlbar“

Ein weiteres Cookie namens „projekktor_controlbar“ übermittelt die Einstellung des Players in der Mediathek. Projekktor ist dabei der Name des Web Players, der in der NDR Mediathek verwendet wird. Das Cookie enthält Daten dazu, ob der Nutzer den „Lautlos“-Button benutzt hat und mit welcher Lautstärke und Abspielqualität das Video wiedergegeben wurde.

Neben diesen beiden Cookies existieren noch weitere: 1. „ns_session“ gibt an, dass eine neue Session gestartet wurde und läuft dementsprechend am Ende der Sitzung ab. 2.

²⁵² Vgl. Gespräch mit A. Svensson, [45:32 Minuten].

²⁵³ Vgl. Ebd., [44:33 Minuten]. Die Anonymisierung wird durch eine Veränderung der IP-Adresse vorgenommen. Hier werden das letzte Oktett der Adresse entfernt.

„ns_cookies_test“ überprüft, ob bereits ein gültiges Cookie besteht und läuft ebenfalls zum Ende der Sitzung ab.

Auch über die Zählpixel auf den ndr.de-Seiten werden Nutzungsdaten erhoben. Diese lassen sich über ein Sniffer-Tool wie Charles oder httpFox auslesen. Die Untersuchung der Zählpixel zeigte, dass das Tracking auf ndr.de zwischen der normalen und der Videonutzung Unterschiede macht. Ein typischer Server Call²⁵⁴, der zum Beispiel bei der Nutzung der Nachrichtenseite auf ndr.de entsteht, sieht wie folgt aus:

| Parameter | Value |
|----------------------------|--|
| nachrichten.niedersachs... | |
| amp;cid | ndsnachrichtenindex101 |
| amp;pdt | 20160603T1254 |
| amp;pti | Nachrichten_aus_Niedersachsen |
| amp;otp | story |
| amp;bra | - |
| amp;sid | - |
| amp;src | 100 |
| amp;ctp | Index |
| amp;ns_t | 1464951790731 |
| amp;ns_referrer | http://www.ndr.de/nachrichten/index.html |

Abbildung 30: Server Call auf ndr.de, gemessen mit dem Sniffer-Tool httpFox

Der Server Call auf ndr.de liefert vergleichsweise wenige Variablen (in der Abbildung als Parameter angegeben) mit, oft finden sich an dieser Stelle auch noch deutlich mehr Informationen, zum Beispiel zur Technik des anfragenden Rechners wie Bildschirmauflösung oder Betriebssystem. Der NDR und der Dienstleister comScore integrieren in die Variablen des normalen Server Calls trotz seines geringen Umfangs aber die wichtigsten Nutzungsinformationen²⁵⁵:

- „amp:cid“ gibt die Sophora-ID des Objekts an, das die Seite mit Inhalt befüllt.
- „amp:pdt“ hält einen Timecode, der die Uhrzeit und das Datum des Server Calls angibt.
- „amp:pti“ gibt den Namen der Seite an, auf der der vorliegende Server Call gemessen wurde.
- „amp:otp“ gibt den Inhaltstyp der Seite an. Neben dem Typ „story“, der für Textinhalte steht, gibt es zum Beispiel auch den Typ „video“, der für Videocontent steht.

²⁵⁴ Unter einem Server Call versteht man die Browser-Anfrage an den Server, auf dem die abzufragende Datei liegt. Der Server Call enthält Nutzungsdaten und auch technische Daten des anfragenden Rechners.

²⁵⁵ Gespräch mit M. Nellen, [Mail vom 06.06.2016]. Die Tracking-Variablen sind nicht von comScores Tool Digital Analytix vorgegeben, sondern vom NDR selbst erstellt worden.

- „amp:ns_referrer“ gibt die Seite an, über die der Nutzer auf die aktuell betrachtete Seite gelangt ist.

Die übrigen Variablen sind nicht befüllt oder nicht für das Tracking relevant.

Etwas umfangreicher sieht der Server Call aus, der beim Aufruf eines Videos abgeschickt wird:

| Parameter | Value |
|-----------|--|
| name | html5.player |
| ns_type | hidden |
| ns_st_sv | 4.1408.01 |
| ns_st_it | c |
| ns_st_id | 1464702760476_1 |
| ns_st_ec | 2 |
| ns_st_sp | 1 |
| ns_st_cn | 1 |
| ns_st_ev | hb |
| ns_st_po | 20024 |
| ns_st_cl | 140000 |
| ns_st_el | 140000 |
| ns_st_hc | 2 |
| ns_st_mp | Projekktor |
| ns_st_mv | 1.4.00 |
| ns_st_pn | 1 |
| ns_st_tp | 1 |
| ns_st_pt | 20025 |
| ns_st_pa | 20025 |
| ns_st_li | 1 |
| ns_st_ci | extra11426 |
| ns_ts | 1464702785483 |
| ns_st_bt | 0 |
| ns_st_bp | 0 |
| ns_st_pc | 0 |
| ns_st_pp | 0 |
| ns_st_br | 0 |
| ns_st_ub | 0 |
| ns_st_pr | Abgehakt |
| ns_st_ep | Abgehakt |
| c7 | http://www.ndr.de/fernsehen/sendungen/extra_3/extra11426-player_image-1d53ebf5-5f27-4... |
| c8 | Abgehakt |
| c9 | http://www.ndr.de/fernsehen/sendungen/extra_3/Abgehakt,extra11426.html |
| afn | TV-20160525-2208-3700 |
| cid | extra11426 |
| pdt | 20160526T0957 |
| otp | video |
| sid | 9 |
| src | - |
| bra | ndrtv |
| ctp | - |
| pti | Abgehakt |

Abbildung 31: Server Call auf ndr.de, gemessen mit dem Sniffer-Tool httpFox

Das Videotracking funktioniert in comScores Tracking über ein zusätzliches Tool namens Stream Sense. Dazu schreibt comScore auf seiner Webseite: „It [Stream Sense, M. Bahne] can supply powerful, real-time analytics reports on browser behaviour during video or audio streams whether on demand, progressive, download or live, and regardless of where the streams are hosted.“²⁵⁶ Um möglichst genaues Videotracking zu ermöglichen, teilt Stream Sense jedes Video in Abschnitte ein. Wird ein neuer Abschnitt im Video erreicht, wird ein neuer Server Call mit aktualisierten Informationen erzeugt. Wie viele einzelne Wiedergabeabschnitte ein Video hat, ist nicht festgeschrieben. Vielmehr hängt die Anzahl der Abschnitte von der Länge des entsprechenden Videos ab. Allgemein ist die Einteilung

²⁵⁶ Vgl. <https://www.comscore.com/Products/Enterprise-Analytics/Capabilities/Video>, zuletzt besucht am 04.06.2016 um 12:05 Uhr.

in Wiedergabeabschnitte aber so angelegt, dass vermehrt zu Beginn des Videos Server Calls abgesetzt werden. Dies zeigt Abbildung 32 beispielhaft:

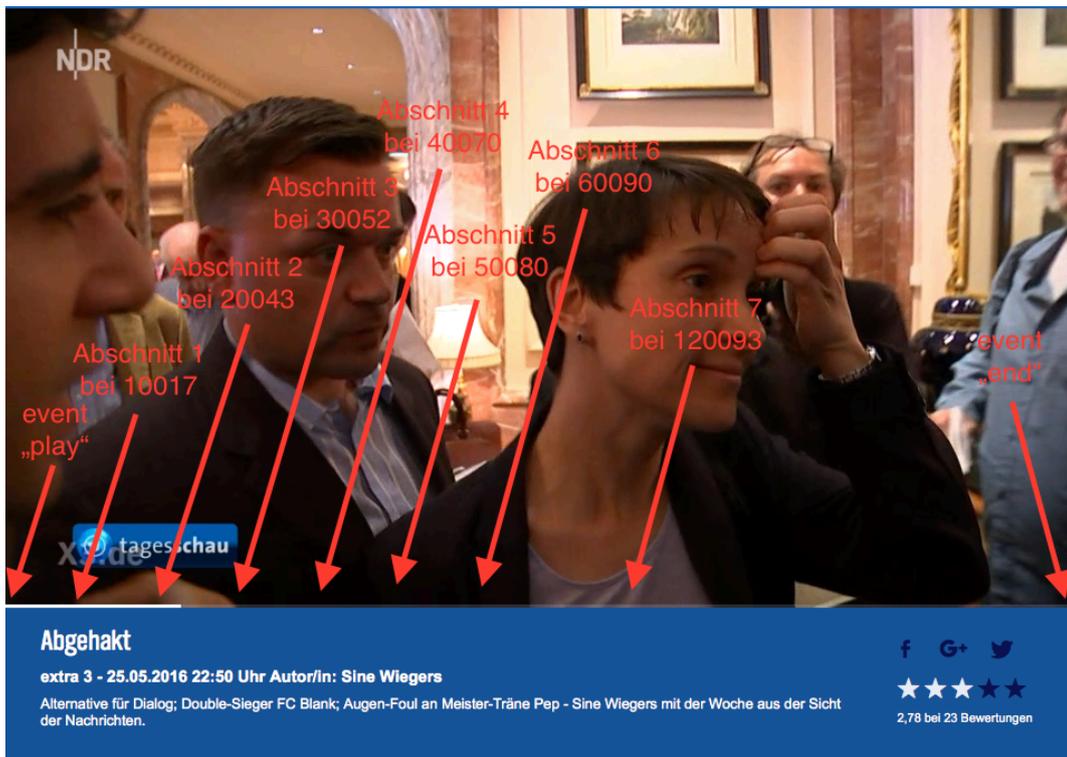


Abbildung 32: Wiedergabeabschnitte eines Videos der NDR Mediathek, Zeitangaben in Millisekunden

Jeder neue Server Call, der bei der Interaktion mit Videocontent auf ndr.de entsteht, enthält folgende für das Tracking relevante Variable²⁵⁷:

- „ns_st_id“ ordnet dem Nutzer eine Playback Session ID zu, die ihn während des gesamten Videoabrufs als einen einzigen Nutzer identifizierbar macht.
- „ns_st_ec“ gibt an, wie viele events²⁵⁸ während des begonnenen Videoabrufs bereits erzeugt wurden.
- „ns_st_ev“ gibt an, in welchem Status sich der Player befindet. Die Variable kann die events „play“, „pause“, „end“ und „hb“ beinhalten. Dabei steht „play“ für den Klick auf den „Play“-Button, „pause“ für das Pausieren des Videos und „end“ für das Erreichen des Videoendes. Das event „hb“ ist im Grunde genommen nur ein

²⁵⁷ Vgl. comScore: Metrics Dimensions Guide Digital Analytix, Amsterdam, 2015, S. 99 ff. Online verfügbar unter: <http://fliphtml5.com/ftym/izdr/basic/51-100>, zuletzt besucht am 04.06.2016 um 13:25 Uhr.

²⁵⁸ „events“ sind Nutzerinteraktionen mit Webinhalten, die unabhängig von den übrigen Trafficdaten gemessen werden können. Ein event kann zum Beispiel der Klick auf den „Play“-Button eines Videos sein oder ein Klick auf den „Absenden“-Button eines bestellten Newsletters.

Platzhalter in der Variable, der ausgegeben wird, wenn ein neuer Wiedergabeabschnitt im Video erreicht wird.

- „ns_st_po“ gibt den Fortschritt auf der Videotimeline im Millisekunden an, den der Nutzer vor dem Überschreiten der Grenze zu einem neuen Abschnitt erreicht hatte.
- „ns_st_cl“ gibt die gesamte Länge des vorliegenden Videos in Millisekunden an.
- „ns_st_hc“ gibt an, welchen Wiedergabeabschnitt des Videos der Nutzer erreicht hat.
- „ns_ts“ gibt die gesamte Wiedergabezeit aller Nutzer für Video in Millisekunden an.
- „ns_st_ep“ gibt den Titel des Videos an.
- „c9“ gibt die URL des abgespielten Videos an.
- „afn“ gibt den Erstellungszeitpunkt des Videos als Timecode an.
- „cid“ gibt die Sophora-ID des abgespielten Videos an.
- „pdt“ gibt den Timecode des aktuellen Abrufs an.
- „otp“ gibt den Inhaltstyp des abgerufenen Contents an.

Die übrigen Variablen, die im Server Call angegeben sind, sind für das Tracking wenig relevant, da sie entweder nicht befüllt sind oder Informationen zum verwendeten Player enthalten. Das trifft zum Beispiel auf die Variablen „ns_st_mp“, „ns_st_mv“, „ns_type“ oder „name“ zu.

4.4.3 Arbeitsablauf bei der Produktion und Veröffentlichung eines Videos

Bis ein Beitrag den Weg von der Produktion bis in die Mediathek gefunden hat, durchläuft er mehrere Bearbeitungsschritte, die für jeden Beitrag leicht unterschiedlich ablaufen. Trotzdem lässt sich der Arbeitsablauf bei der Erstellung eines Videos – von der Produktion bis hin zur Veröffentlichung in der Mediathek – in einen allgemeingültigen Prozess fassen, den folgende Abbildung zunächst bildlich beschreibt:



Abbildung 33: Von der Produktion bis zur Mediathek: Arbeitsablauf bei der Erstellung eines Videos

Die Farbgebung des Prozessdiagramms gibt Aufschluss darüber, welche Abteilungen an der Produktion und der Veröffentlichung eines Videos beteiligt sind. Vor dem Hintergrund der Leitfrage der vorliegenden Arbeit liegt der Fokus natürlich auf dem Bereich des Prozesses, der im Diagramm oben mit der hellblauen Farbe gekennzeichnet ist. Diese Arbeitsschritte finden in der Mediathek statt, während die übrigen Schritte in den anderen Bereichen des NDR und teilweise sogar bei externen Firmen durchgeführt werden. Am Anfang des Prozesses stehen die Produktion und die Aufzeichnung des Videos. Nicht alle Videos, die im NDR ausgestrahlt werden oder in die Mediathek gestellt werden, sind auch vom NDR produziert. Ganz im Gegenteil: Viele Dokumentationen, Reportagen oder Talkshows werden von externen Unternehmen im Auftrag des NDR produziert. So zeichnet sich zum Beispiel die beckground TV GmbH für die Produktion der Unterhaltungsshow „Ina’s Nacht“ verantwortlich²⁵⁹ oder unter anderem die stenner filmproduktion GmbH für „Typisch!“ und „NaturNah“²⁶⁰. Beim NDR werden hingegen viele täglich laufende Sendungen produziert – zum großen Teil mit nachrichtlichem Hintergrund – wie zum Beispiel „NDR//Aktuell“, „NDR//Aktuell extra“ oder die Regionalmagazine²⁶¹, genauso wie Reportagen und Talkshows²⁶². Die Aufzeichnung der Videos läuft – abgesehen von der immer wieder auf den neuesten Stand gebrachten Studioteknik – immer noch so ab, wie zu Beginn des Fernsehens: Kameras zeichnen Inhalte auf, die von Schauspielern oder Moderatoren gesprochen werden. Erst beim Transfer der Inhalte ist die fortschreitende Digitalisierung deutlich zu erkennen: Während vor 2005 noch Redakteure mit Videokassetten durch die Rundfunkanstalten liefen, um die Inhalte zu transferieren, läuft dieser Prozess heute digital und nahezu in Echtzeit ab²⁶³. Noch während der Produktion läuft das aufgezeichnete Material auf einem Ingest-Server ein. So kann jeder Redakteur, der auf den Server zugreifen kann, mit einer Verzögerung von wenigen Sekunden auch das eingespielte Material abrufen. Ein weiterer großer Vorteil der „Ingestierung“ der Videos ist, dass die Produktion nicht mehr standortgebunden ist,

²⁵⁹ Vgl. <http://www.beckground.de/produktionen/unterhaltung.html?p=7>, zuletzt besucht am 05.06.2016 um 13:10 Uhr.

²⁶⁰ Vgl. <http://www.stennerfilm.de/produktionen.html>, zuletzt besucht am 05.06.2016 um 13:11 Uhr.

²⁶¹ Vgl. Geschäftsbericht für den NDR von 2009, Hamburg, 2010, S. 23.

²⁶² 2014 gab es laut A. Svensson im NDR 26,6 Prozent Eigenproduktionen, 6,5 Prozent Auftragsproduktionen, 6,1 Prozent Übernahmen und 59,7 Prozent Wiederholungen.

²⁶³ Vgl. Geschäftsbericht für den NDR von 2009, S. 23.

sondern die Einspielung auf den Server theoretisch von überall auf der Welt möglich ist, was die Videoproduktion bei Großevents wie Olympia erheblich vereinfacht²⁶⁴.

Der zweite wichtige Schritt bei der Verarbeitung eines Videos für Online ist die Rechteklärung. Diese findet in den Redaktionen statt. Sie müssen die Inhalte in das weiter oben beschriebene Verweildauerkonzept einordnen²⁶⁵. Dieser Prozessschritt findet in den Redaktionen lange vor oder während der Produktion des Videos statt.

Ist das Videomaterial schließlich erfolgreich „ingestiert“, können es die Redakteure in der Mediathek abrufen. Das funktioniert über das Tool „VPMS Media Center“, das über Microsoft Silverlight im Internet Explorer ausgespielt wird und auf alle Inhalte der vernetzten Produktion (Ingest, Schnitt, Ton) zugreifen kann. Hauptsächlich greift die Mediathek auf die Suchfunktion des Tools zurück, um die „ingestierten“ Videos zu durchsuchen. Wonach gesucht werden muss, geben die „Bestell-Mails“, die aus den Redaktionen kommen und Anweisungen für die Veröffentlichung des Videos enthalten, oder das Wiki der Mediathek vor, das ebenfalls Informationen zur Befüllung der Video-Metadaten enthält. Die Anweisungen aus der Redaktion und deren Form sind über die Sendungen hinweg nicht konsistent, manche geben nur die Teaser, Videobeschreibungen und Verweildauern an, manche wollen ihre Videos auch noch als Download oder mit freigegebener Kommentarfunktion haben. Das Wiki ist hingegen konsistent für alle Sendungen aufgebaut, allerdings sind die „Bestell-Mails“ meist aktueller.

Die für das Auffinden des Materials relevanten Informationen sind aber sowohl im Wiki, als auch in den „Bestell-Mails“ enthalten: Die Suche funktioniert am besten über die BID. Diese ist ein alphanumerischer Code und existiert stets in zwei Varianten: Zum einen als Such-BID, über die sich das gesuchte Material im „VPMS Media Center“ schnell finden lässt, und zum anderen als Export-BID, die beim Export des Materials aus dem Schnitt-Tool vergeben wird. Jede Sendereihe hat eine Such- und Export-BID, wobei regelmäßig ausgestrahlte Sendungen wiederkehrende BIDs besitzen können.

Ist das gesuchte Material im „VPMS Media Center“ gefunden, kann es direkt in das Schnitt-Tool namens „VPMS Preview-Client“ exportiert werden. In diesem Tool lassen sich lediglich grundlegende Videoschnitt-Aktionen durchführen wie zum Beispiel das Abschneiden und Anheften von Video-Schnipseln oder das Einfügen von Weiß- und Schwarzblenden. Im Schnitt-Tool werden die Beiträge gemäß den Anweisungen der Redaktionen geschnitten. Häufig müssen Inhalte, die keine Online-Freigabe erhalten

²⁶⁴ Vgl. Geschäftsbericht für den NDR von 2009, S. 25.

²⁶⁵ Vgl. Gespräch mit M. Lahmann [2:31 Minuten].

haben, aus den Videos geschnitten werden. Ebenso müssen die Videos oft auf die richtige Länge geschnitten werden, da sie mit Vor- und Nachlauf von mehreren Minuten ingestiert werden. Wenn das Video fertig geschnitten ist, kann es mit ersten Metadaten befüllt werden. In der Export-Maske können Titel, Teaser, Tags, Stichworte und Verweildauern angegeben werden. Zusätzlich müssen dort das Logo des Videos und der weiter oben erwähnte ID-Stamm eingegeben werden. Sind alle Eingaben getätigt, wird das Video ins Encoding-System des NDR überspielt. Im „Odeon“ genannten System werden alle benötigten Versionen des Videos encodiert.

Sobald die Enkodierung abgeschlossen ist, steht das Video zur Bearbeitung in Sophora zur Verfügung. Gemäß den Anweisungen in der Mail aus der Redaktion werden schließlich die zusätzlichen Metadaten wie Ausspielkanal und Verweildauer des Videos ausgefüllt. Hier entscheidet sich auch, ob das Video nur in der NDR Mediathek zu sehen ist, oder ob es auch in die ARD Mediathek überspielt wird, in die neben dem NDR auch alle anderen öffentlich-rechtlichen Rundfunkanstalten in Deutschland Material einfließen lassen.

Ist dieser Prozess in Sophora abgeschlossen, kann das Video in der NDR Mediathek auf ndr.de veröffentlicht werden, wo es für alle Nutzer zum Abruf bereit steht, solange es die Verweildauer des Videos zulässt.

Der letzte Schritt, den das Prozessdiagramm zeigt, gehört im Grunde genommen nicht mehr zum Prozess der Videoproduktion und –verarbeitung, weshalb er im Diagramm auch mit gestrichelten Linien dargestellt ist. Dennoch sollte das Tracking als Erfolgsmessung in dieser Betrachtung noch genannt werden, da es den Prozess der Videoproduktion- und -verarbeitung in gewisser Weise abschließt. Über das Tracking, das – wie weiter oben ausgeführt – der Dienstleister comScore für den NDR übernimmt, lässt sich herausfinden, welche Videos in der Mediathek besonders gut von den Nutzern angenommen wurden. Somit kommt dem Tracking eine Feedback und Controlling-Funktion zu, da die Ergebnisse des Trackings in Form von Reports regelmäßig an die Redaktionen weitergegeben werden. Mit dem Erhalt des Feedbacks ist der Prozess der Videoaufzeichnung und –verarbeitung abgeschlossen.

4.4.4 Inhalte der NDR Mediathek

Dieses Kapitel soll eine Momentaufnahme der Inhalte der Mediathek abbilden, die zum Zeitpunkt der Untersuchung in der Mediathek standen. Es soll herausgearbeitet und geprüft werden, was überhaupt in die Mediathek gestellt wird und ob sich die Inhalte in der Mediathek sinnvoll clustern lassen. Mithilfe dieser Analyse soll ein erster Überblick über

den Fundus der Mediathek geschaffen werden, auf den ein Recommender System zugreifen könnte.

Zunächst ist die Frage relevant, welche produzierten Inhalte überhaupt in der Mediathek ausgespielt werden und welche nicht. Der Sinn der Mediathek ist es, möglichst alle Bewegtbild-Inhalte, die der NDR zur Verfügung hat, auch in der Mediathek für die Nutzer zugänglich zu machen. Dennoch gibt es viele Videos, die nicht in voller Länge oder überhaupt nicht in die Mediathek gestellt werden. Das hat vor allem mit den eingekauften Lizenzen zur Ausstrahlung. So werden zum Beispiel jegliche Inhalte zur Berichterstattung zur 1. und 2. Bundesliga sowie zu anderen Sportevents wie Boxen oder Formel 1 aus den Nachrichtensendungen geschnitten, weil der NDR keine Lizenz zur Online-Veröffentlichung besitzt²⁶⁶. Ein weiterer Grund, der häufig dafür sorgt, dass ein Video nicht in der Mediathek veröffentlicht werden darf, ist ebenfalls rechtlicher Natur: Zeigt ein Nachrichtenmagazin Inhalte, deren Rechte nur für das Fernsehen lizenziert wurden, aber nicht für das Netz, dürfen diese Inhalte ebenfalls nicht in der Mediathek ausgespielt werden, sondern müssen aus den jeweiligen Sendungen geschnitten werden. Das betrifft zum Beispiel Mitschnitte aus Konzerten oder auch historische Aufnahmen.

Bevor nun ein Überblick über die Inhalte der Mediathek gegeben wird, muss kurz angesprochen werden, dass es sich bei diesem Überblick nur um ein temporäres Abbild der Inhalte der Mediathek handelt, da sich die angegebenen Zahlen zum Beispiel durch auslaufende Verweildauern innerhalb weniger Tage verändern können. Um einen Eindruck zu vermitteln, welche Inhalte überhaupt in der Mediathek zu finden sind, ist der Überblick aber zu gebrauchen, da sich trotz dynamischer Contentlage die Verteilung der Inhalte der Videos nicht von einem Tag auf den anderen grundlegend ändern wird.

Zum Zeitpunkt der Untersuchung, dem 6. Juni 2016 um 13:24 Uhr, standen insgesamt 13.029 Videos in der Mediathek zum Abruf bereit. Für eine Inhaltsübersicht wurden die Videos noch weiter geclustert, das erste Clustering erfolgte aber zunächst auf Grundlage der Darstellungsformate.

²⁶⁶ So darf das Hamburg Journal zum Beispiel für 30 Tage in der Mediathek stehen, der darin enthaltene Beitrag zum Hamburger SV nur für 24 Stunden. Das führt dazu, dass der entsprechende Beitrag aus der ganzen Sendung entfernt werden muss.

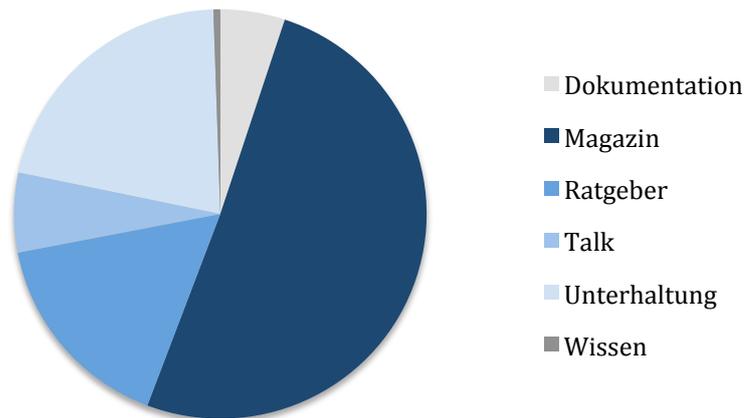


Abbildung 34: Clustering der Inhalte der Mediathek nach Formaten

Dieses Clustering existiert nicht von Seiten des NDR, sondern wurde vom Autor der vorliegenden Arbeit erstellt. Dazu wurden alle Sendereihen mit ihren zugehörigen Videos einem Format untergeordnet²⁶⁷. Unter Dokumentationen fielen zum Beispiel Tierdokumentationen wie „Am Kap der wilden Tiere“, aber auch Reportagereihen wie „7 Tage“ oder „45min“. Das Magazin-Format macht den größten Teil der Mediathek aus, hierunter fallen sämtliche Nachrichtensendungen wie „NDR//Aktuell“ oder die Regionalmagazine. Zusätzliche zählen zu den Magazinen auch Sendungen mit gesellschaftlichen Themen wie Panorama. Zum Ratgeber-Format zählen die Kochsendungen wie „Tim Mälzer kocht!“ sowie die Reiseratgeber wie „Endlich Urlaub!“ oder Verbraucher-Ratgeber wie „Markt“. Das Talk-Format beinhaltet sämtliche Talkshows des NDR wie zum Beispiel die „NDR Talk Show“ oder „Offen gesagt“. Das Segment Unterhaltung macht in der obigen Abbildung fast ein Viertel aller Sendungen aus und beinhaltet vor allem Inhalte wie „extra 3“, „Der Tatortreiniger“ oder „Die NDR Quizshow“. Unter das Wissens-Format fallen Sendungen wie „plictsch.“ oder „X:enius“. Das erste Clustering der Inhalte zeigt sehr deutlich, dass der Schwerpunkt der Videos in der Mediathek auf Magazin-Formaten mit nachrichtlichen Inhalten liegt. Zusätzlich setzt ein großer Teil der Videos in der Mediathek auf unterhaltende und Ratgeber-Formate. Die übrigen Formate Talk, Wissen und Dokumentationen spielen keine zentrale Rolle.

²⁶⁷ Der Begriff „Format“ folgt hier nicht vollständig seiner medienwissenschaftlichen Bedeutung, sondern orientiert sich auch nach den Inhalten der Sendungen. So sind zum Beispiel die Sendungen „Markt“ oder „Visite“ vom Fernsehformat her Magazine, hier aber aufgrund der Ausrichtung der Inhalte den Ratgebern untergegliedert.

Ein etwas feineres Clustering erlaubt die Einteilung der Videos nach Themen. Auch diese sind vom Autor ausgewählt worden und orientieren sich nur in Grundzügen an den bereits existierenden Themen des NDR²⁶⁸.

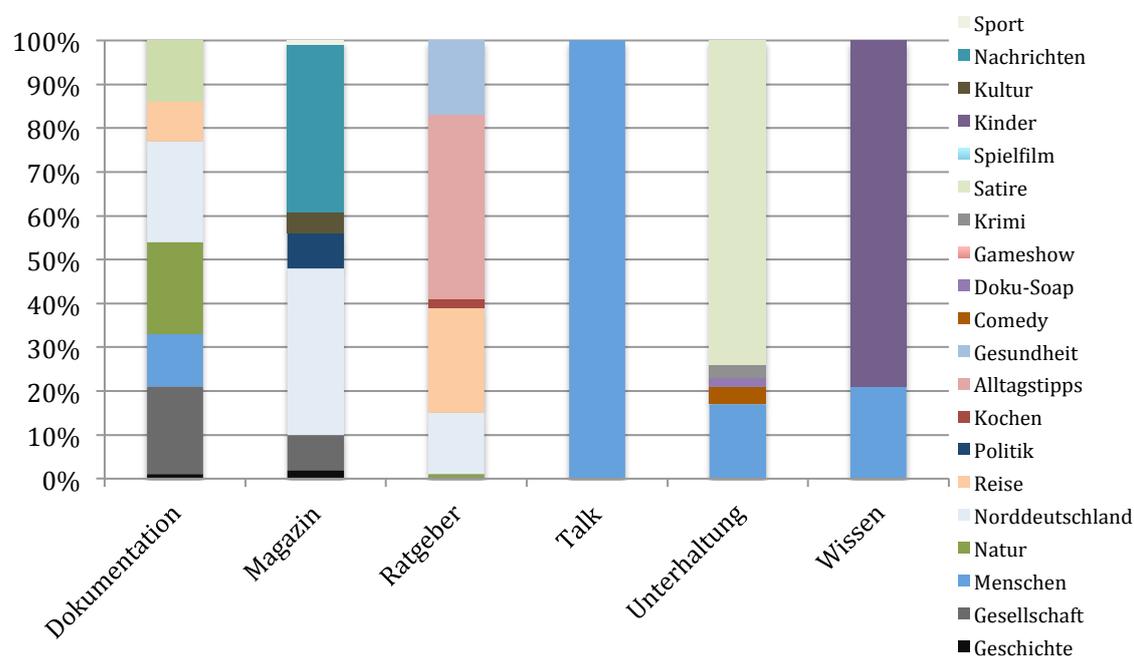


Abbildung 35: Feineres Clustering der Formate nach Themen

Abbildung 35 zeigt die Verteilung der neu erstellten Themen auf die Formate. Es können in der Darstellung mehrere Themen in verschiedenen Formaten enthalten sein, es ist also nicht jedem Format nur eine bestimmte Themenrubrik zugeordnet. Bei der Untersuchung der Abbildung wird vor allem deutlich, dass es nur wenige Themen gibt, die sich durch sämtliche Cluster ziehen. Dazu gehört zum Beispiel das Thema Norddeutschland, das sowohl in den Formaten Dokumentation, Magazin und Ratgeber bedient wird, oder auch das Thema Menschen, das sich in den Formaten Dokumentation, Talk, Unterhaltung und Wissen wiederfindet. Andere Themen wie zum Beispiel Satire, Nachrichten oder Kinder sind zwar anteilig sehr gut in manchen Formaten vertreten, sie beschränken sich aber auch auf das jeweilige Format.

Einen besseren Überblick, wie gut einzelne Themen in der Mediathek tatsächlich vertreten sind, gibt folgende Abbildung:

| | |
|--|--|
| Über 10 Prozent der vorhandenen Mediatheks-Videos | Norddeutschland (26,07%), Nachrichten (22,32%) |
| Unter 10 Prozent der vorhandenen | Satire (9,64%), Alltagstipps (8,25%), |

²⁶⁸ Zu finden ist die gesamte Übersicht über die Themen unter <http://www.ndr.de/home/thema108.html>, zuletzt besucht am 11.06.2016 um 12:17 Uhr.

| | |
|--|--|
| Mediatheks-Videos | Menschen (6,77%), Gesellschaft (5,80%), Politik (5,33%), Reise (5,10%) |
| Unter 5 Prozent der vorhandenen Mediatheks-Videos | Gesundheit (3,34%), Kultur (2,68%), Geschichte (1,37%), Natur (1,03%) |
| Unter 1 Prozent der vorhandenen Mediatheks-Videos | Comedy (0,56%), Kochen (0,45%), Sport (0,37%), Krimi (0,35%), Doku-Soap (0,26%), Kinder (0,26%), Spielfilm (0,02%), Gameshow (0,02%) |

Abbildung 36: Verteilung der Themen in der Mediathek

Abbildung 36 hilft bei der Analyse noch einmal weiter: Die Tabelle verdeutlicht, dass die Videos in der Mediathek zu einem Viertel nachrichtliche Themen enthalten und zu ungefähr einem anderen Viertel das Thema Norddeutschland. Relativ weit hinter den beiden Top-Themen liegen Videos zum Thema Satire und Alltagstipps. Ebenfalls über die Fünf-Prozent-Hürde geschafft haben es die Themen Menschen, Gesellschaft, Politik, Reise und Natur. Die Themen Gesundheit, Kultur und Geschichte weisen noch eine kleine, aber vorhandene Relevanz in der Mediathek auf, während die Themen Natur, Comedy, Kochen, Sport, Krimi, Doku-Soap, Kinder, Spielfilm und Gameshow nur einen verschwindend geringen Anteil der Videos in der Mediathek ausmachen.

Interessantester Punkt der Analyse ist, dass abgesehen vom Thema Satire alle übrigen Themen aus dem Format-Cluster Unterhaltung unter die Ein-Prozent-Hürde fallen und somit die unterhaltenden Elemente in der Mediathek so gut wie nicht vorhanden sind. Stattdessen liegt der Fokus der Mediathek vor allem auf den Themen Nachrichten und Norddeutschland, auch Ratgeber-Themen haben in der Mediathek einen hohen Stellenwert. Diese Informationen sind elementar wichtig für die Konzeption des Recommender Systems, da im Vergleich zu den Engines von Netflix und Co. bei der NDR Mediathek nicht die Unterhaltung, sondern – nach der Analyse der Inhalte gehend – die Information des Nutzers im Mittelpunkt steht.

4.5 Datenschutz beim NDR

Wie alle anderen Portale im Netz hat auch der NDR auf seinen Seiten eine Datenschutzrichtlinie, die dem interessierten Nutzer angibt, welche Standards der NDR auf seinem Portal an den Datenschutz seiner Nutzer anlegt²⁶⁹. Die Datenschutzerklärung des

²⁶⁹ Zu finden ist die Datenschutzrichtlinie des NDR unter folgendem Link: <http://www.ndr.de/service/datenschutz/index.html>, zuletzt besucht am 11.06.2016 um 17:01 Uhr.

NDR liest sich dabei wie viele andere Datenschutzrichtlinien auch. Der NDR gibt an, dass Nutzungsdaten im Rahmen des Trackings in Form von Cookies und Zählpixeln gemessen und gespeichert werden. Daten die über das Tracking hinausgehen, zum Beispiel Mailadressen für Gewinnspiele oder Newsletter, können nur vom Nutzer freiwillig überlassen werden und werden nach dem Ende der Aktion wieder gelöscht. Zusätzlich bietet der NDR Links in der Datenschutzerklärung, über die sich der Nutzer über ein Opt-Out-Verfahren²⁷⁰ vom Tracking ausschließen kann. Außerdem ermöglicht es der NDR jedem Nutzer, sich über jegliche Daten, die dem NDR vorliegen, zu informieren und diese auf Wunsch löschen zu lassen.

Die Basis für die Datenschutzrichtlinien des NDR ist allein das Hamburgische Datenschutzgesetz, Einflüsse aus dem RStV gibt es nicht²⁷¹. Das Hamburgische Datenschutzgesetz lässt dem NDR und auch allen anderen Hamburger Firmen sehr viel Freiraum beim Datensammeln: „Nutzungsdaten dürfen problemlos gesammelt werden, solange der Nutzer darüber informiert ist, welche Nutzungsdaten auf der jeweiligen Internetseite erhoben werden. Bei personenbezogenen Daten wie Mail-Adressen oder Postanschriften muss die Einwilligung des Nutzers vorliegen. Die gibt er in den meisten Fällen automatisch, sobald er sich für ein Gewinnspiel oder einen Newsletter anmeldet²⁷²“, sagt der Datenschutzbeauftragte und stellvertretende Justitiar des NDR, Horst Brendel. Problematischer wird es laut Brendel bei Daten von Kindern und Jugendlichen: Bevor man ihre Daten speichern und nutzen dürfe, sei die Einwilligung der Erziehungsberechtigten notwendig²⁷³.

Die Einführung eines Recommender Systems für die NDR Mediathek hält Brendel trotz letztgenannter Einschränkung für unbedenklich. „Solange dem Nutzer klar ist, für welches Angebot er seine Daten hergibt und welche seiner Daten gesammelt und ausgewertet werden, ist sowohl die Einführung von Mediatheks-Profilen der Nutzer unproblematisch – gesetzt dem Fall, der Nutzer stimmt der Sammlung seiner Daten zu -, als auch die Einführung eines Empfehlungssystems auf Basis freiwilliger Interessensangaben der Nutzer. Gerade letzteres System wäre eine sehr datenschutz- und nutzerfreundliche Lösung.“²⁷⁴

²⁷⁰ Das Opt-Out-Verfahren ist ein einstufiger Prozess, über den sich ein Nutzer zum Beispiel vom Webtracking ausschließen lassen kann. Es reicht in den meisten Fällen ein Klick auf einen Link, um das Verfahren abzuschließen.

²⁷¹ Vgl. Gespräch mit H. Brendel, [8:51 Minuten].

²⁷² S. Ebd., [11:02 Minuten].

²⁷³ Vgl. Ebd., [11:43 Minuten].

²⁷⁴ S. Ebd., [22:37 Minuten].

Trotz des recht großen Spielraums beim Thema Datenschutz würde Brendel bei der Einführung eines Recommender Systems eine zentrale Rolle spielen. Steht zum Beispiel eine Änderung im Tracking der Seite an, müssen die Redaktionen ihm ein Konzept zukommen lassen, das er auf Zulässigkeit überprüft. Dies müsste auch bei der Einführung eines Recommender Systems geschehen. Wichtig für ihn wären dabei allgemeine Geschäftsbedingungen (AGB), die der Nutzer liest und versteht: „Der Nutzer muss auf den ersten Blick erkennen: Das sind die Vorteile des Angebots und diese Daten werden dafür von mir gesammelt und verarbeitet. Dem Nutzer müssen vor dem Beitritt zum Angebot alle relevanten Informationen übersichtlich präsentiert werden. Alle übrigen datenschutzrechtlichen Ausführungen, können sich hinter einem ‚Read more‘-Button verstecken, den Nutzer bei Interesse anklicken und durchlesen können.“²⁷⁵ Brendel ist es wichtig, dass Nutzer nicht mit endlos langen AGB konfrontiert werden, die die meisten höchstens überfliegen, aber nicht lesen und verstehen. Mit seinem Vorschlag erhofft er sich, auf der einen Seite die Menschen für den Datenschutz zu sensibilisieren und auf der anderen Seite die Transparenz der NDR Mediathek zu fördern.

4.6 Der NDR und die Disruption der Medienlandschaft

Das Kapitel 2 der vorliegenden Arbeit beschäftigt sich mit dem veränderten Mediennutzungsverhalten der Rezipienten. Das Zwischenfazit des Kapitels zeigte, dass mit dem Internet ein „Medium“ in Erscheinung trat, dessen Reichweite auf Kosten der übrigen Medien immer weiter zunimmt und das für Medienanbieter eine breite Palette an Werkzeugen bereit hält, um Inhalte zu erstellen und anzubieten. Damit hat das Internet auf viele traditionelle Medienanbieter eine disruptive Wirkung – im Sinne des von Clayton Christensen erstmals eingeführten Begriffs. Bereits in den 1990-er Jahren untersuchte der Wirtschaftsprofessor, warum etablierte Unternehmen neue Technologien unterschätzen und infolgedessen ihre einstmals beherrschende Marktstellung verlieren.

Die disruptive Wirkung des Internets resultiert vor allem aus dem weiter oben beschriebenen Umstand, dass es nicht nur als Ausspielkanal – wie Fernseher oder Radio – dient, sondern parallel zur Content-Erstellung und –Verbreitung genutzt werden sowie gleichzeitig ein Kanal der Diskussion über den erstellten und verbreiteten Content sein kann. Das führt dazu, dass traditionelle Medienanbieter die über viele Jahre bewährten Produktions- und Distributionsprozesse vollständig überdenken müssen, was vielen Anbietern schwerfällt. Dafür gibt es laut Christensen vor allem zwei Gründe: Den

²⁷⁵ S. Gespräch mit H. Brendel, [24:13 Minuten].

Unternehmen, die von einer Disruption betroffen sind, fehle es an der nötigen Weitsicht und der Beweglichkeit, um angemessen und schnell auf die Disruption reagieren zu können²⁷⁶.

Das Kapitel über die Struktur des öffentlich-rechtlichen Rundfunks verdeutlicht, dass dieser von den disruptiven Veränderungen der Medienlandschaft, die das Internet nach sich zieht, besonders betroffen ist. Obwohl die nötige Weitsicht und das Bewusstsein für die Disruption an vielen Stellen in Redaktion und Mediathek vorhanden sind – das zeigten die Experteninterviews, die im Rahmen dieser Arbeit geführt wurden –, sind es bei den öffentlich-rechtlichen Rundfunkanstalten vor allem die starren Strukturen, also die nach Christensens Worten „fehlende Beweglichkeit“, die ein schnelles Reagieren auf die Disruption verhindern. Gerade beim Aspekt Videostreaming können Anbieter wie Netflix oder YouTube, die nicht in dem Maße an Verweildauern, Drei-Stufen-Tests für veränderte Telemedienangebote oder ein Gremium wie den Rundfunkrat gebunden sind, wesentlich agiler sein und dementsprechend schneller auf zukünftige Disruptionen reagieren.

4.7 Zwischenfazit

Das zurückliegende Kapitel beschäftigte sich ausführlich mit der Struktur der öffentlich-rechtlichen Rundfunkanstalten und der öffentlich-rechtlichen Mediatheken sowie der Umsetzbarkeit eines Recommender Systems in dieser Struktur.

Schon die Untersuchung der rechtlichen Grundlagen der Mediatheken warf dabei einige Schwierigkeiten auf, die in den danach folgenden Unterkapiteln noch um einige weitere Punkte ergänzt wurden.

Durch die rechtlichen Vorgaben des 12. RÄStV gibt es für ein Recommender System vor allem zwei wichtige Begrenzungen. Die erste liegt in Form der Verweildauern vor.

Aufgrund der Verweildauern sind die Inhalte der NDR Mediathek teilweise nur für einen sehr kurzen Zeitraum in der Mediathek abrufbar. Gerade nachrichtliche Inhalte in der Mediathek sind davon betroffen. Die Ergebnisse der Inhaltsanalyse, die zeigen, dass gerade die nachrichtlichen Themen prozentual sehr stark in der Mediathek vertreten sind, verstärkt die Herausforderung des extrem dynamischen Contents noch. Auf dieser Basis gute und verlässliche Empfehlungen auszusprechen, dürfte für ein Recommender System kompliziert sein.

²⁷⁶ Vgl. Christensen, Clayton [u.a.]: The Innovator's Dilemma – Warum etablierte Unternehmen den Wettbewerb um bahnbrechende Innovationen verlieren, München, 2013, S. 39.

Hinzu kommen noch die inhaltlichen Begrenzungen für die Mediathek. Hier ist vor allem das Lizenz-Problem zu nennen, das für zugekaufte Serien und Filme gilt, und das deren Upload in die Mediathek verhindert. Dieses Problem ist auch eine Erklärung dafür, dass unterhaltende Themen – abgesehen vom Thema Satire – nur am Rande eine Rolle spielen. Gerade in diesem Punkt sind Streamingplattformen wie Netflix oder Amazon Prime den öffentlich-rechtlichen Mediatheken weit voraus.

Positiv fiel dafür die nähere Untersuchung der Drei-Stufen-Test-Methode aus. Da die NDR Mediathek auch mit einem Recommender System keines der Positivkriterien erfüllt, erreicht sie nicht die Aufreißschwelle für einen neuen Drei-Stufen-Test.

Die Untersuchungen zum CMS und zur Datenstruktur im Online-Bereich des NDR zeigten, dass dem NDR mit dem CMS Sophora ein sehr mächtiges und funktionales Tool zur Seite steht, das bereits sehr viele Daten zu einem Video erfasst. Problematisch sind die vielen unterschiedlichen Nutzer, die die Videos bearbeiten. Besonders durch die Freitext-Felder in Sophora, aber auch bei den übrigen „Metadaten“-Feldern wie den Themen, ist die Inkonsistenz der Daten hoch und könnte bei einem Recommender System Probleme verursachen. Dies gilt auch für die Daten, die über ProPlan eingespielt werden. Ein weiteres Problem ist die Struktur der derzeit existierenden Themen: Sie sind zu fein gegliedert, als dass eine Recommender Engine auf dieser Basis gute und verlässliche Empfehlungen für die Nutzer ausspielen könnte.

Die Untersuchung des Trackings auf ndr.de zeigt, dass bereits einige Daten, die wichtig für die Einführung eines Recommender Systems wären, getrackt werden. Dazu gehören vor allem die Sophora-IDs, damit sich das Video ohne Zweifel den mitgetrackten Daten zuordnen lässt. Ebenso wichtig sind die Wiedergabesegmente, in die jedes Video eingeteilt ist und mit denen sich der Wiedergabestatus auslesen lässt. Praktisch ist hier, dass in jedem Fall ein „end“-Server Call gesendet wird, sobald die Wiedergabe eines Videos beendet wird – egal ob der Nutzer den Browsertab schließt oder nur die Seite verlässt. Zusätzlich müssten noch einige zusätzliche Variablen mit ins Tracking genommen werden, die den Nutzer näher beschreiben, zum Beispiel ob er mit einem mobilen oder Desktop-Rechner die Videos abgerufen hat. Trotzdem ist in diesem Bereich eine gute Grundlage vorhanden, an die eine Recommender Engine ansetzen könnte.

Das Gespräch über den Datenschutz mit dem Datenschutzbeauftragten des NDR brachte ebenfalls einige wichtige Erkenntnisse: So stünde der Einführung eines Recommender Systems aus Sicht des Datenschutzes nichts im Wege, solange der Nutzer ausreichend darüber informiert wird, welche Daten zu welchem Zweck vom NDR erhoben werden.

Hier müsste zusätzlich ein Konzept erarbeitet werden, das dem Datenschutzbeauftragten die Funktionsweise des Recommender Systems näher erläutert.

Der Aspekt der Disruption rundete das Kapitel zur Struktur des öffentlich-rechtlichen Rundfunks ab. Es zeigte, dass es vor diesem Hintergrund noch große Herausforderungen für die Rundfunkanstalten in Form von Verweildauern, Drei-Stufen-Tests und Rundfunkräten gibt, die ein schnelles Reagieren auf disruptive Veränderungen unmöglich machen. Das sorgt im Vergleich zu anderen Anbietern wie Netflix oder YouTube für einen Wettbewerbsnachteil.

5. Konzept eines Recommender Systems für die NDR Mediathek

Im folgenden Kapitel soll ein das Konzept für ein Recommender System erarbeitet werden, das die Erkenntnisse der vorliegenden Arbeit zusammenführt und auf ihnen aufbaut.

Zunächst wird deshalb eine Zielsetzung für das Recommender System der NDR Mediathek formuliert, ehe herausgearbeitet wird, welche Herausforderungen speziell in der Mediathek des öffentlich-rechtlichen Rundfunks zu bewältigen sind und wie ein Empfehlungssystem für die Mediathek schließlich aussehen könnte. Zum Schluss des Kapitels soll ein erstes Grundgerüst für ein Recommender System stehen, an dem der NDR und auch andere öffentlich-rechtliche Rundfunkanstalten weiterarbeiten können.

5.1 Zielsetzung des Recommender Systems für die NDR Mediathek

Das Unterkapitel 3.2 hat gezeigt, dass die Zielsetzung für ein Recommender System stets aus zwei Perspektiven betrachtet werden muss – zum einen aus der Sicht des Anbieters, zum anderen aus der Sicht des Users.

Dabei sollte sich ein Recommender System in der NDR Mediathek zunächst an den übergeordneten Zielen eines Empfehlungssystems orientieren. Ein Großteil dieser Ziele ist bereits in Abbildung 14 aufgeführt. Dazu gehört zum Beispiel: Erhöhen der Glaubwürdigkeit der ndr.de-Seite, Effiziente Informations-Verarbeitung, Nutzergerechtes Anbieten von Inhalten, Unterstützung für den Nutzer bei der Suche nach Informationen und Content sowie das Erhöhen der Nutzerloyalität. Zusätzlich zu den generellen Zielen eines Recommender Systems sollte das Empfehlungssystem des öffentlich-rechtlichen Rundfunks auch noch speziellen Zielen gerecht werden, die sich im Kapitel über die Struktur des öffentlich-rechtlichen Rundfunks herauskristallisierten.

Dazu gehört ein Ziel, das sich aus der Untersuchung des Contents der NDR Mediathek formulieren lässt: Weil die Mediathek im Gegensatz zu Netflix oder anderen großen

Videostreamingportalen eher wenige unterhaltende Elemente enthält, sondern der Fokus vielmehr auf journalistisch und redaktionell hochwertig aufbereiteten Informationen liegt, sollte das Ziel der Mediathek weniger die Unterhaltung, sondern vielmehr die Information des Nutzers mit für ihn relevantem Content sein.

Ein weiteres wichtiges Ziel sollte die Transparenz des Recommender Systems sein. Das bedeutet zweierlei: Zum einen sollte der Nutzer darüber informiert sein, welche Daten der NDR von ihm für die Empfehlungserstellung erfasst und nutzt, zum anderen sollte der Nutzer die „Kontrolle“²⁷⁷ über die erstellten Empfehlungen behalten, indem er – wie im Abschnitt zu den Knowledge-based Recommender Systems beschrieben – jederzeit in einem interaktiven Prozess die ausgespielten Empfehlungen verändern oder sogar über eine „Neustart“-Funktion komplett von vorne beginnen kann. Im Punkt „Transparenz“ mischen sich die Ziele des Anbieters und die der User merklich.

Neben den in der Abbildung 14 beschriebenen Zielen aus Usersicht, gibt vor allem das Kapitel über die Veränderungen im Mediennutzungsverhalten der Rezipienten Aufschluss darüber, was die User zum einen von den öffentlich-rechtlichen Mediatheken und somit auch von einem angebotenen Recommender System erwarten. Hier steht vor allem die Erwartung im Vordergrund, zeit- und ortsunabhängig relevanten Inhalt geboten zu bekommen. Gerade bei den jüngeren Mediennutzern lässt sich dieser Trend erkennen, doch auch die Mediennutzung der älteren Menschen verlagert sich immer in Richtung Internet sowie hin zur Medienkonvergenz.

Setzt man nun voraus, dass sich ein großer Teil der in Abbildung 14 aufgeführten Ziele mit der Implementierung eines gut arbeitenden Recommender Systems erfüllen lässt, bleibt für die Mediatheken der öffentlich-rechtlichen Anstalten folgende Zielsetzung für die Einführung eines Recommender Systems bestehen: *Es sollte ein transparentes Recommender System geschaffen werden, das die „Kontrolle“ über die Daten und Empfehlungen dem Nutzer lässt und ihm eine zeit- und ortssouveräne Nutzung der Inhalte der Mediatheken ermöglicht. Das Recommender System sollte weniger auf die Unterhaltung, sondern vielmehr auf die Information des Nutzers mit für ihn relevanten Content ausgelegt sein. Bei Bedarf gibt das System auch eigenständig Empfehlungen für den Nutzer ab.*

²⁷⁷ In diesem Fall bedeutet „Kontrolle“ lediglich, dass der Nutzer durch die Beschränkungen und Anforderungen, die er zum Beispiel in einem Constraint-Based Recommender vorgeben kann, Inhalte filtern kann, wie in Kapitel 4.5.3 beschrieben.

5.2 Spezielle Herausforderungen für ein Empfehlungssystem in der NDR Mediathek

Auch bei den Herausforderungen für ein Recommender System müssen neben den generellen Herausforderungen, die im Unterkapitel 3.7 näher erläutert werden, auch die NDR-spezifischen Herausforderungen betrachtet werden. Diese bestehen aus zwei großen Komplexen: Hier ist zum einen die extrem dynamische Content-Lage, die durch die Verweildauer-Problematik entsteht, und zum anderen die teilweise vorhandene Inkonsistenz der im CMS Sophora eingegeben Daten zu nennen.

Durch die unterschiedlich langen Verweildauern der Video-Items muss das Recommender System sehr gut und effizient mit Veränderungen in der Produktdatenbank umgehen können. Theoretisch entsteht mit jedem neuen Item in der Produktdatenbank eine neue Kaltstart-Problematik, die es zu lösen gilt. Dabei kann der Einsatz des Content-based Recommender Systems eine effiziente Strategie sein, das die Ratings auf Basis der Ähnlichkeiten zwischen Attributen von Items berechnet.

Hier spielt wiederum die zweite Herausforderung für das Recommender System der NDR Mediathek eine zentrale Rolle: Wie in Kapitel 4.6 beschrieben, sorgen die Freitext-Felder zusammen mit vielen unterschiedlichen Datenverarbeitern dafür, dass die Metadaten zu den Items teilweise sehr inkonsistent sind, was eine Vergleichbarkeit der Attribute zwischen den Items erschweren könnte. Hier sollte also zunächst eine Lösung gefunden werden, die Daten konsistent zu machen, um eine Basis für ein gut funktionierendes Recommender System schaffen zu können. Ein Ansatz wäre hier zum Beispiel, eine einheitliche Linie für die Stichwortvergabe festzulegen oder Cluster vorzugeben, die künftig über ein Drop-Down-Menü auszuwählen sind. Ein weiterer Punkt bei der Datenqualität ist die Vergabe von Themen in Sophora. Sie ist derzeit zu breit aufgestellt, als das sie einen wertvollen Beitrag bei der Attributbildung für Items leisten könnte. Das Clustering sollte vor diesem Hintergrund – möglicherweise speziell für die Verarbeitung der Videos – überarbeitet werden.

5.3 Elemente und Struktur des Recommender Systems für die NDR Mediathek

Im folgenden Unterkapitel soll nun ein Ansatz für das Recommender System in der NDR Mediathek beschrieben werden. Zuerst soll aufgezeigt werden, wie das Empfehlungssystem in die Webseite implementiert werden kann, ehe der Empfehlungsprozess genauer beschrieben wird. Zum Schluss des Unterkapitels sollen die verwendeten Kontext-Informationen näher beschrieben werden.

5.3.1 Implementierung des Recommender Systems in die ndr.de-Seite

Die Implementierung des Recommender Systems in die ndr.de-Seite orientiert sich an der Abbildung 15, die beispielhaft zeigt, wie ein Empfehlungssystem in eine E-Commerce-Webseite eingegliedert werden kann.

Die folgende Abbildung stellt nun einen Ansatz dar, wie ein Recommender System in die ndr.de-Webseite integriert werden könnte:

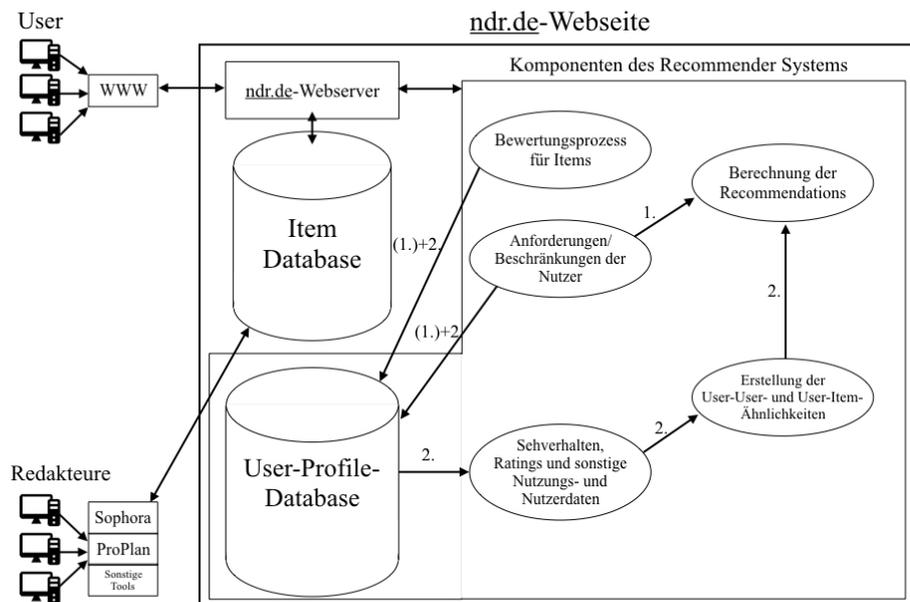


Abbildung 37: Implementierung eines Recommender Systems in die ndr.de-Webseite

Abbildung 37 zeigt zum einen die unterschiedlichen Verbindungen zwischen den Elementen der Webseite auf und versucht außerdem, die Komponenten, die bereits im Hintergrund von ndr.de vorhanden sind, von den Komponenten eines zu implementierenden Recommender Systems abzugrenzen. Der auf der rechten Seite abgebildete Empfehlungsprozess bezieht sich dabei auf die Mediathek und nicht auf die gesamte ndr.de-Webseite.

Zu Beginn fragen viele verschiedene User Inhalte über das Internet und den ndr.de-Webserver an. Dieser greift auf die Item-Database der Webseite zu und gibt die angefragten Inhalte über den Webserver und das Internet an die User aus. Gleichzeitig können die Redakteure des NDR über das CMS Sophora und andere Tools Inhalte zur Item-Database hinzufügen, abrufen und verändern. Diese Infrastruktur ist beim NDR bereits vorhanden.

Neu hinzukommen würden nun die Komponenten des Empfehlungssystems. Hier wäre zum einen die User-Profile-Database, die mit Informationen zum Nutzer sowie zur Nutzung der Webseite der spezifischen Nutzer befüllt wird. Dabei ist das freiwillige

Anlegen eines Nutzerprofils, das im folgenden Kapitel 5.3.2 näher beschrieben werden soll, eine grundlegende Voraussetzung. Die Ellipsen in Abbildung 37 beschreiben die verschiedenen Prozesse, die während des Empfehlungsprozesses im Recommender System ablaufen. Die Zahlen an den Verbindungen zwischen den Komponenten symbolisieren den zweiseitigen Empfehlungsprozess, der ebenfalls in Kapitel 5.3.2 näher beschrieben werden soll. Dabei steht „1.“ für den Constraint-based-Recommendation-Prozess und die „2.“ für den Recommendation-Prozess des hybriden Recommender Systems.

5.3.2 Der Empfehlungsprozess in der NDR Mediathek

Der Empfehlungsprozess des Recommender Systems für die NDR Mediathek sollte in zwei Teilen ablaufen, um die weiter oben formulierte Zielvoraussetzung erfüllen zu können. Wichtig sollte sein: Die Nutzung des Recommender Systems darf für alle Nutzer nicht verpflichtend sein, die Angebote der Mediathek müssen auch abseits des Recommender Systems genutzt werden dürfen. Deshalb sollte das System den Nutzern die Wahl lassen, ob sie die Mediathek nur mit dem vorgelagerten Recommendation-Teil nutzen möchten, mit der vollständigen Empfehlungsgenerierung inklusive Login oder sie komplett vom Recommendation-Prozess ausgelassen werden möchten.

Dem ersten Teil, also dem vorgelagerten Recommender System, liegt die Constraint-based-Methode zugrunde. Diese Methode erlaubt es dem Nutzer, selbst festzulegen, was für ihn relevant ist. In einem interaktiven Prozess führt das Recommender System den Nutzer zu dem für ihn wichtigen Content. Die Methode passt umso besser in das Recommender System der NDR Mediathek, wenn man bedenkt, dass der Fokus der Mediathek weniger auf der Unterhaltung, sondern vor allem auf der Information des Nutzers liegt. Dank der Constraint-based-Methode kann der Nutzer so viel zielgerichteter zu den für ihn relevanten Informationen geführt werden, als es derzeit der Fall ist. Dabei sollte die Menge an Anforderungen und Beschränkungen, die der Nutzer bei dieser Methode angeben muss, eher klein gehalten werden, da er nicht – wie in Kapitel 3.4.3 beschrieben – nach einem sehr komplexen Item wie einem Haus oder einem Auto sucht, sondern nach einem Video. Deshalb würde es sich hier anbieten, den Nutzer zunächst zu fragen, weshalb er die Mediathek besucht. Dort könnte die Inhaltsanalyse der Mediathek in Kapitel 4.4.4 bei der Bildung von Themenclustern helfen. Wählt der Nutzer nun „Nachrichten“ aus, fragt das Recommender System weitere Informationen vom Nutzer ab, zum Beispiel aus welchem Bundesland er kommt und welche Art von Nachrichten ihn interessieren. Zusätzliche Kontext-Informationen wie Tageszeit und Wochentag würden die im Verlauf des Empfehlungsprozesses gebildeten „Knowledge Bases“ automatisch

weiter ergänzen. Diese Form der interaktiven Feedback-Schleife würde den Nutzer bis zu seinem relevanten Content führen. Verlässt der Nutzer die Mediathek, wird auch der aufgezeichnete Recommendation-Prozess gelöscht.

Dem zweiten Teil, also dem nachgelagerten Recommender System, liegt ein hybrides System zugrunde, das einige der in Kapitel 3.4 beschriebenen Methoden miteinander vereint. Grundvoraussetzung für dieses System wäre eine Login-Funktion für die User, damit ihm seine Nutzungsdaten ebenso wie seine Anforderungen und Beschränkungen dauerhaft zugeordnet werden können. Bei der Anmeldung eines Profils sollten die in Kapitel 4.5 herausgestellten Bedingungen des Datenschutzbeauftragten des NDR beachtet werden. So muss dem Nutzer bei der Anmeldung klar erläutert werden, welche Daten von ihm wozu gesammelt werden. Hat der Nutzer sich erfolgreich angemeldet, kann die Constraint-based-Methode genutzt werden, um das Kaltstart-Problem zu überwinden und bereits erste Informationen zum Nutzer zu gewinnen. Zusätzlich setzt das hybride Recommender System Neighborhood-based Collaborative Filtering ein, das sowohl User-based funktioniert. Zusätzlich kommt die Content-based-Methode zum Einsatz. Sie soll der Problematik der kurzen Verweildauern entgegenwirken. Indem konsistente Attribute zu den Items erstellt werden, spielt das Kaltstart-Problem bei neuen Items nahezu keine Rolle mehr. Das hybride Recommender System funktioniert dabei nach dem Ensembled Design, das in Kapitel 3.4.4 näher beschrieben wird, und aus allen Rating-Vorhersagen der einzelnen Recommendation ein gewichtetes finales Rating generiert.

Für die Erstellung der Empfehlungen sollte zunächst eine Rating-Struktur aufgebaut werden, die theoretisch bereits in Form des 5-Sterne-Bewertungssystems in der NDR Mediathek existieren. Diese sollte allerdings deutlich prominenter eingebunden werden als derzeit, da die Ratings für Items das Rückgrat des Recommender Systems bilden. Zusätzlich sollten bei der Empfehlungserstellung noch die übrigen Nutzer- und Nutzungsdaten wie zum Beispiel das Sehverhalten oder demografische Daten berücksichtigt werden – dazu zählen auch die in Kapitel 3.4 beschriebenen unären Ratings, die Interaktionen des Nutzers mit den Items auf der ndr.de-Seite und der Mediathek beschreiben.

In seinem User-Profil hat der Nutzer die Möglichkeit, seine Präferenzen einzusehen, die das Recommender System glaubt zu kennen. Diese kann er in seinem Profil manipulieren oder bei Bedarf sogar komplett löschen, um neu starten zu können. Das soll die Transparenz des Systems und auch das Vertrauen des Nutzers in das System erhöhen.

5.3.3 Im Recommender System verwendete Kontext-Informationen

Um die Recommendations des Systems möglichst akkurat zu halten, sollte es unterschiedliche Kontext-Informationen zu den Nutzern mit in den Prozess der Empfehlungserstellung einfließen lassen. Dazu sollten folgende gehören:

1. Demografische Daten: Diese Daten beinhalten vor allem den Standort und das Alter des Nutzers. Sie sind unter anderem wichtig, um ihm altersgerechten sowie nur für ihn relevanten Content auszuspielen. So interessiert den 56-jährigen Hamburger der Beitrag des Hamburg Journals über die Kunstaussstellung in den Hamburger Messehallen eher als den 21-Jährigen aus Mecklenburg-Vorpommern.
2. Zeitsensitive Daten: Die zeitsensitiven können unabhängig vom Nutzer erhoben werden und zum Beispiel Tageszeit und Wochentag beinhalten. Wie das Kapitel zur Mediennutzung gezeigt hat, hängt sie stark mit der Tageszeit zusammen. So interessieren den Nutzer morgens andere Inhalte als abends. Gleiches lässt sich auch auf den jeweiligen Wochentag übertragen. Hier steht die Mediennutzung am Wochenende im Kontrast zu den Arbeitstagen in der Woche.
3. „Stimmung“ des Nutzers: Dank der Constraint-based-Methode kann der Gemütszustand des Nutzers abgefragt werden – so wie es zum Beispiel Buzzfeed seit einiger Zeit tut. Damit könnten zum Beispiel disruptive Veränderungen im Leben des Nutzers abgefangen werden und die Empfehlung von unpassendem Content weiter unterbunden werden.

5.4 Zwischenfazit

Das letzte Kapitel dieser Arbeit zeigt einen ersten Vorschlag, wie ein Recommender System für den NDR und die übrigen öffentlich-rechtlichen Rundfunkanstalten aussehen könnte.

Es wurde eine Zielsetzung erarbeitet, die ein Recommender System für die öffentlich-rechtlichen Rundfunkanstalten erfüllen sollte. Anhand dieser Zielsetzung wurden aus dem Pool der unzähligen verschiedenen Filter-Methoden einige herausgesucht, die gut zur formulierten Zielsetzung passen und außerdem einen Lösungsansatz für die speziell für den NDR erarbeiteten Herausforderungen bieten. Mit der Abbildung 37 wurde zudem ein Bild gezeichnet, wie das Recommender System in die ndr.de-Webseite implementiert werden könnte.

Damit bietet das zurückliegende Kapitel einen ersten Ansatz für den NDR und die übrigen öffentlich-rechtlichen Rundfunkanstalten, an dem sie sich orientieren und weiterarbeiten können.

6. Zusammenfassung und Ausblick

6.1 Zusammenfassung

Die vorliegende Arbeit setzt sich aus vier großen und thematisch sehr unterschiedlichen Kapiteln zusammen. Das erste Kapitel beschäftigt sich mit den Veränderungen im Mediennutzungsverhalten der deutschen Gesellschaft. Es wurde zunächst die Frage gestellt, wie Gewohnheit bei der Mediennutzung entstehen kann. Dort spielte der Aspekt der positiven Verstärkung bei der Mediennutzung eine wichtige Rolle. Anschließend wurde herausgearbeitet, welche Grundlagen für die Mediennutzung überhaupt geschaffen werden müssen. Die Untersuchung betraf vor allem die technischen Übertragungswege wie Sendefrequenzen oder das Internet. Darauf folgte die Untersuchung der veränderten Mediennutzung selbst. Dort wurden das Mediennutzungsverhalten von 1964 dem von 2015 gegenübergestellt, um Unterschiede möglichst gut erkennen zu können. Das Kapitel lieferte drei wichtige Erkenntnisse: Erstens steht mit dem Internet ein disruptives „Medium“ bereit, das die Reichweiten der übrigen Medien – besonders die der Tageszeitungen – negativ beeinflusst. Zweitens begünstigt das Internet die Medienkonvergenz. Durch das begrenzte Zeitbudget bei der Mediennutzung nutzen immer mehr Rezipienten mehrere Medien zur gleichen Zeit. Drittens geht der Trend bei der Mediennutzung vor allem bei den jüngeren Nutzern immer mehr von der getakteten Nutzung zur enttakteten Nutzung. Die Rezipienten möchten zeit- und ortssouverän entscheiden, welche Medieninhalte sie vor sich haben.

Das folgende Kapitel beschäftigte sich mit den Grundlagen der Recommender Systems. Hier wurde zunächst eine Definition der Systeme gegeben, ehe die Zielsetzung der Recommender Systems sowie deren Implementierung in eine Website abgehandelt wurde. Anschließend wurden die verschiedenen Filter-Methoden der Systeme vorgestellt. Zum Schluss wurde für jede Filter-Methode die Entstehung einer Empfehlung näher erklärt, ehe die Themen Netzwerkeffekte im Recommender System und Herausforderungen für Recommender Systems das Kapitel abschlossen. Im Hinblick auf das Konzept eines Recommender Systems für die NDR Mediathek brachte das Kapitel über die Grundlagen der Empfehlungssysteme vor allem eine wichtige Erkenntnis: *Das Recommender System existiert nicht.* Empfehlungssysteme sind immer sehr domain-spezifisch aufgebaut und orientieren sich stark an der vorgegebenen Zielsetzung sowie den Voraussetzungen, die sie auf der entsprechenden Website vorfinden.

Um solch eine Zielsetzung formulieren zu können, beschäftigte sich das anschließende Kapitel mit der Struktur der öffentlich-rechtlichen Rundfunkanstalten. Es startete mit den rechtlichen Voraussetzungen der öffentlich-rechtlichen Rundfunkanstalten und behandelte in einem zweiten großen Unterkapitel die Struktur und die Arbeitsabläufe in der NDR Mediathek. Abgerundet wurde das Kapitel mit einem kleinen Ausblick auf die Herausforderungen des Datenschutzes sowie die Disruption der Medienlandschaft vor dem Hintergrund der öffentlich-rechtlichen Rundfunkanstalten.

Das letzte Kapitel der vorliegenden Arbeit sammelte die Erkenntnisse aus allen vorhergehenden Kapiteln zusammen und setzte daraus ein Konzept für ein Recommender System für die NDR Mediathek zusammen.

Auch auf die Leitfrage der vorliegenden Arbeit soll zum Schluss noch eine Antwort gegeben werden. Sie fragte nach den Herausforderungen, die die Einführung eines Recommender System für die öffentlich-rechtlichen Rundfunkanstalten mit sich bringt, und welche Lösungsansätze die aktuelle Forschung für diese bereit hält. Die Untersuchung der weiter oben genannten Themenfelder zeigte, dass die Mediatheken des öffentlich-rechtlichen Rundfunks grundlegend einen guten Ansatzpunkt für ein Recommender System bilden und vor dem Hintergrund des weiter oben beschriebenen, sich verändernden Mediennutzungsverhaltens sogar bilden müssen. Gerade für die Nutzer dürfte sich die Funktionalität der Mediatheken durch ein angebundenes Recommender System deutlich erhöhen, zusätzlich dürfte eine so erweiterte Mediathek auch zur Aufwertung der gesamten Webseite der öffentlich-rechtlichen Rundfunkanstalten dienen.

Viele Elemente, die für die Einführung eines Recommender Systems elementar wichtig sind, sind bereits beim NDR vorhanden und müssten lediglich angebunden werden. Jedoch zeigt diese Arbeit auch, dass es noch einige Hindernisse gibt, die überwunden werden müssen. Dazu zählen vor allem die Verweildauern und die Inkonsistenzen bei der Metadaten-Pflege. Für diese Herausforderungen hat diese Arbeit einige Strategien und Lösungswege aufgeführt, die für den NDR hilfreich sein dürften.

Abschließend liefert diese Arbeit also ein Grundgerüst, das der NDR und die übrigen öffentlich-rechtlichen Rundfunkanstalten nutzen können, um bei der Einführung eines Recommender Systems erfolgreich sein zu können.

6.2 Ausblick

Trotz der recht ausführlichen Betrachtung der aufgeführten Themenfelder, konnten manche Bereiche aufgrund des knappen Platzes nur am Rande erwähnt werden, obwohl sich eine genauere Untersuchung lohnen würde. Dazu gehört zum einen die Evaluation der

ausgespielten Empfehlungen, die ein wichtiger Bestandteil für die Optimierung des Recommender Systems ist, und zum anderen die Frage nach dem Vertrauen des Nutzers in ein Recommender System. Dieser Aspekt spielt vor allem im E-Commerce-Bereich eine wichtige Rolle, mit dem vermehrten Einsatz der Recommender Systeme in journalistischen Bereichen dürfte diese Frage aber auch hier aktuell werden. Dabei spielen auch medienethische Fragen eine Rolle, wie jüngst die Vorwürfe gegen Facebook zeigten, die dem Social Network vorwarfen, den Newsfeed der Nutzer bewusst zu manipulieren. Erwähnenswert sind in diesem Zusammenhang auch die möglichen Auswirkungen der Behavioral Analytics, die das Verhalten des Nutzers wie Mausbewegung und Tippgeschwindigkeit auf einer Webseite beobachten, um daraus Schlüsse zu ziehen. Auch diese Informationen wären für ein Recommender System relevant, um die Qualität der herausgegebenen Empfehlungen weiter zu optimieren und zum Beispiel das Kaltstart-Problem noch wirksamer bekämpfen zu können.

Die Aspekte der Evaluation der Recommendations, der Behavioral Analytics und des Vertrauens in Empfehlungssysteme sind für den NDR aber erst in Zukunft relevant. Die nächsten Schritte für die öffentlich-rechtliche Rundfunkanstalt wären, den in dieser Arbeit erarbeiteten Ansatz in ihre Planungen aufzunehmen und zu diskutieren, ob er sich nur für den NDR eignet oder ob der Ansatz auch für andere öffentlich-rechtliche Rundfunkanstalten generalisierbar ist. Idealerweise wäre er auch für die Mediatheken der übrigen Anstalten sowie der übergeordneten ARD Mediathek geeignet. Sollte diese Diskussion ergeben, dass der Ansatz generalisierbar ist, könnte der nächste Schritt sein, den ersten Teil des vorgeschlagenen Recommender Systems mit der Constraint-based-Methode in den Mediatheken zu implementieren und zu testen.

Generell kann das Recommender System für die eigenen Mediatheken für die öffentlich-rechtlichen Rundfunkanstalten ein weiterer wichtiger Schritt in der digitalen Transformation sein, mit dem sie der in dieser Arbeit oftmals angeführten Disruption der Medienlandschaft begegnen können und neben Streaming-Anbietern wie YouTube oder Netflix weiterhin ein attraktives Medienangebot für die Nutzer bereithalten können.

7. Literaturverzeichnis & Quellenverzeichnis

Aggarwal, Charu: Recommender Systems – The Textbook, New York, 2016.

Ahrens, Sophie Charlotte: Recommender Systems – Relevance in the Consumer Purchasing Process, München, 2010.

ARD-Werbung Sales & Services. n.d. Anzahl der durchschnittlich pro Haushalt empfangbaren TV-Sender in Deutschland in den Jahren 1988 bis 2016 (jeweils 1. Januar). Statista. Zugriff am 17. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/160407/umfrage/frei-empfangbare-tv-sender-in-deutschland-seit-1988/>.

Badura, Peter: Die öffentlich-rechtlichen Rundfunkanstalten bieten Rundfunk und Telemedien an, In: Archiv des öffentlichen Rechts (Band 134), hrsg. von Udo di Fabio u.a., Bonn, 2009

Bitkom. n.d. Welche Art von Sendungen oder Videos schauen Sie sich als Stream im Internet an?. Statista. Zugriff am 24. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/316608/umfrage/art-der-nutzung-von-videostreaming-in-deutschland/>.

Bodendorf, Freimut: Daten- und Wissensmanagement, Heidelberg, 2006

Bogner, Alexander [u.a.]: Das Experteninterview, Opladen, 2005.

Breunig, Christian; von Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015.

Burt, Ronald: Structural Holes and Good Ideas, In: American Journal of Sociology (Band 110), Chicago, 2004, S. 349-399.

Christensen, Clayton [u.a.]: The Innovator's Dilemma – Warum etablierte Unternehmen den Wettbewerb um bahnbrechende Innovationen verlieren, München, 2013

Conolly, Dan: A Little History of the World Wide Web, Princeton, 2000.

Dewald, Pirmin; Robers, Sebastian: Netzwerkeffekte: Theoretische Grundlagen, Kaiserslautern, 2015.

Engel, Bernhard; Breunig, Christian: Massenkommunikation 2015 – Mediennutzung im Intermediavergleich, hrsg. von der ARD Werbung Sales & Services, Frankfurt, 2015.

Esser, Hartmut: Die Definition der Situation, In: Kölner Zeitschrift für Soziologie und Sozialpsychologie (Band 48.1), Köln, 1996, S. 1-34.

Felfernig, Alexander [u.a.]: Developing Constraint-based Recommender Systems, In: Recommender Systems Handbook, hrsg. von Francesco Ricci [u.a.], New York, 2011, S. 187-216.

Gantz, John [u.a.]: The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, Framingham, 2014.

Geschäftsbericht für den NDR von 2009, Hamburg, 2010. Online verfügbar unter https://www.ndr.de/der_ndr/unternehmen/bericht123.pdf, zuletzt besucht am 04.06.2016 um 13:45 Uhr.

Goldhammer, Klaus u.a.: Social TV – Aktuelle Nutzung, Prognosen, Konsequenzen, In: Schriftenreihe Medienforschung der Landesanstalt für Medien Nordrhein-Westfalen (Band 76), hrsg. von Jürgen Brautmeier, Düsseldorf, 2015.

Goodnity. n.d. Findest du das mobile Datennetz in Deutschland verbesserungswürdig?. Statista. Zugriff am 17. April 2016. Verfügbar unter

<http://de.statista.com/statistik/daten/studie/533641/umfrage/umfrage-zur-zufriedenheit-mit-dem-mobilen-datennetz-in-deutschland/>.

Gundlach, Hardy: Medienökonomisches Gutachten zu den marktlichen Auswirkungen der geplanten NDR Mediathek, Hamburg, 2009.

Hansen, Hans-Robert [u.a.]: Empfehlungssysteme zur Verkaufsberatung im Internet State-of-the-Art und Konsumentenakzeptanz, In: Wirtschaftsinformatik (Band 49), Wiesbaden, 2007, S. 50-61.

Held, Thorsten: Online-Angebote öffentlich-rechtlicher Rundfunkanstalten – Eine Untersuchung des verfassungsrechtlich geprägten und einfachgesetzlich ausgestalteten Funktionsauftrags öffentlich-rechtlichen Rundfunks im Hinblick auf Internet-Dienste, Baden-Baden, 2008.

Herlocker, Jonathan [u.a.]: Evaluating Collaborative Filtering Recommender Systems, In: ACM Transactions on Information Systems (Band 22), New York, 2004, S. 5-53.

IfD Allensbach. Anzahl der Internetnutzer in Deutschland, die das Internet über das Handy oder Smartphone nutzen, von 2013 bis 2015 (in Millionen).

<http://de.statista.com/statistik/daten/studie/170557/umfrage/internutzung-ueber-das-handy-oder-smartphone/>, zuletzt besucht am 17. April 2016 um 12:34 Uhr.

Katz, Michael; Shapiro, Carl: Product Compatibility Choice in a Market with Technological Process, In: Oxford Economic Papers, London, 1986, S. 146-165.

Krohn-Grimberghe, Arthur: Analysis and Evaluation of Multi-relational Recommender Systems, Hildesheim, 2012.

Lops, Pasquale [u.a.]: Content-based Recommender Systems – State of the Art and Trends, In: Recommender Systems Handbook, hrsg. von Francesco Ricci [u.a.], New York, 2011, S. 75-105.

Marcus, James: Amazonia. Fünf Jahre im Zentrum der E-Commerce-Revolution, Berlin, 2006.

Metrics Dimensions Guide Digital Analytix, Amsterdam, 2015. Online verfügbar unter: <http://fliphtml5.com/ftym/izdr/basic/51-100>, zuletzt besucht am 04.06.2016 um 13:25 Uhr.

Meuser, Michael; Nagel, Ulrike: Das Experteninterview – konzeptionelle Grundlagen und methodische Anlage, In: Methoden der vergleichenden Politik- und Sozialwissenschaft, hrsg. von Detlef Jahn, Wiesbaden, 2009, S. 465-469.

Mock, Thomas: Was ist ein Medium? Eine Unterscheidung kommunikations- und medienwissenschaftlicher Grundverständnisse eines zentralen Begriffs, In: Publizistik (Band 51), Leipzig, 2006, S. 183-200.

Naab, Teresa: Gewohnheiten und Rituale der Fernsehnutzung – Theoretische Konzeption und methodische Perspektiven, In: Reihe Rezeptionsforschung (Band 27), hrsg. von Marco Dohle u.a., Baden-Baden, 2013.

Neuhoff, Heiko: Rechtsprobleme der Ausgestaltung des Auftrags des öffentlich-rechtlichen Rundfunks im Online-Bereich, In: Materialien zur rechtswissenschaftlichen Medien- und Informationsforschung (Band 68), hrsg. von Martin Eifert u.a., Hamburg, 2011.

Neumann, Andreas: Recommender Systems for Information Providers – Designing Customer Centric Paths to Information, Heidelberg, 2009.

Pariser, Eli: The Filter Bubble – What the Internet is Hiding from You, New York, 2011.

Pasquay, Anja: Die deutschen Zeitungen in Zahlen und Daten 2012, hrsg. vom Bund Deutscher Zeitungsverleger, Berlin, 2012.

Pasquay, Anja: Die deutschen Zeitungen in Zahlen und Daten 2016, hrsg. vom Bund Deutscher Zeitungsverleger, Berlin, 2016.

Ricci, Francesco; Rokach, Lior; Shapira, Bracha: Introduction to Recommender Systems Handbook, In: Recommender Systems Handbook, hrsg. von ebendiesen, New York, 2011, S. 1-38.

Todsén, Nele Julie: Grenzen gebührenfinanzierter Telemedien – Eine kritische Bewertung des Drei-Stufen-Tests mit besonderem Blick auf das Wettbewerbsverhältnis zur Presse, In: Deutsches und internationales Wirtschaftsrecht (Band 54), hrsg. von Marian Paschke, Berlin, 2013

Zattoo. n.d. Auf welchen Geräten nutzen Sie Internet-TV Services?. Statista. Zugriff am 24. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/528875/umfrage/nutzung-von-internet-fernsehen-nach-endgeraeten-in-deutschland/>.

8. Abbildungsverzeichnis

Abbildung 1: Pasquay, Anja: Die deutschen Zeitungen in Zahlen und Daten 2016, hrsg. vom Bund Deutscher Zeitungsverleger, Berlin, 2016, S. 5.

Abbildung 2: ARD-Werbung Sales & Services. n.d. Entwicklung der Anzahl der öffentlich-rechtlichen und privaten Radiosender* in Deutschland in den Jahren 1987 bis 2016. Statista. Zugriff am 17. April 2016. Verfügbar unter <http://de.statista.com/statistik/daten/studie/36329/umfrage/radiosender-anzahl-oeffentlich-rechtliche-und-private-seit-1987/>.

Abbildung 3: Breunig, Christian; von Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 511.

Abbildung 4: Breunig, Christian; von Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 511.

Abbildung 5: Breunig, Christian; von Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 510.

Abbildung 6: Breunig, Christian; von Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 516.

Abbildung 7: Engel, Bernhard; Breunig, Christian: Massenkommunikation 2015 – Mediennutzung im Intermediavergleich, hrsg. von der ARD Werbung Sales & Services, Frankfurt, 2015, S. 310.

Abbildung 8: Breunig, Christian; von Eimeren, Birgit: 50 Jahre „Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien, hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 512.

Abbildung 9: Breunig, Christian; von Eimeren, Birgit: 50 Jahre
„Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien,
hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 522.

Abbildung 10: Breunig, Christian; von Eimeren, Birgit: 50 Jahre
„Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien,
hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 506.

Abbildung 11: Breunig, Christian; von Eimeren, Birgit: 50 Jahre
„Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien,
hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 510.

Abbildung 12: ARD, ZDF. n.d. Anteil der Befragten, die folgende Online-
Videoanwendungen nutzen in den Jahren 2007 bis 2015. Statista. Zugriff am
29. April 2016. Verfügbar unter
<http://de.statista.com/statistik/daten/studie/200171/umfrage/nutzung-von-videoanwendungen-im-internet-seit-2007/>.

Abbildung 13: Breunig, Christian; von Eimeren, Birgit: 50 Jahre
„Massenkommunikation“ – Trends in der Nutzung und Bewertung der Medien,
hrsg. von der ARD Service & Sales, Frankfurt, 2015, S. 508/509.

Abbildung 14: Tabelle zu den Anforderungen der Anbieter und Nutzer an ein
Recommender System, erstellt von Max Bahne am 17.06.2016 um 17:43 Uhr.
Vgl. dazu Ahrens, Sophie Charlotte: Recommender Systems – Relevance in the
Consumer Purchasing Process, München, 2010, S. 12 ff.

Abbildung 15: Ahrens, Sophie Charlotte: Recommender Systems – Relevance in
the Consumer Purchasing Process, München, 2010, S. 10.

Abbildung 16: Aggarwal, Charu: Recommender Systems – The Textbook, New
York, 2016, S. 172.

Abbildung 17: Aggarwal, Charu: Recommender Systems – The Textbook, New York, 2016, S. 34.

Abbildung 18: Aggarwal, Charu: Recommender Systems – The Textbook, New York, 2016, S. 34.

Abbildung 19: Aggarwal, Charu: Recommender Systems – The Textbook, New York, 2016, S. 171.

Abbildung 20: Aggarwal, Charu: Recommender Systems – The Textbook, New York, 2016, S. 32.

Abbildung 21: Kops, Manfred u.a.: Rahmenbedingungen für die Durchführung des Drei-Stufen-Tests, Gutachten erstellt für den Rundfunkrat des WDR, S. 31.

Abbildung 22: Kops, Manfred u.a.: Rahmenbedingungen für die Durchführung des Drei-Stufen-Tests, Gutachten erstellt für den Rundfunkrat des WDR, S. 30.

Abbildung 23: Kops, Manfred u.a.: Rahmenbedingungen für die Durchführung des Drei-Stufen-Tests, Gutachten erstellt für den Rundfunkrat des WDR, S. 38.

Abbildung 24: Überblick über den Aufbau von Sophora, erstellt von der subshell GmbH, abrufbar unter <http://www.subshell.com/en/sophora/documentation/>, zuletzt besucht am 28.05.2016 um 16:45 Uhr.

Abbildung 25: Screenshot der Dateneingabe in Sophora, erstellt von Max Bahne am 28.05.2016 um 12:31.

Abbildung 26: Screenshot der Dateneingabe in Sophora, erstellt von Max Bahne am 28.05.2016 um 12:31.

Abbildung 27: Screenshot der Dateneingabe in Sophora, erstellt von Max Bahne am 28.05.2016 um 12:32.

Abbildung 28:

http://www.ndr.de/fernsehen/sendungen/hamburg_journal/Hamburg-Journal,sendung514926.html, zuletzt besucht am 29.05.2016 um 16:54 Uhr.

Abbildung 29: Screenshots aus dem Firefox-Addon „Cookies Manager+“, erstellt von Max Bahne am 03.06.2016 um 13:42 Uhr.

Abbildung 30: Screenshot aus dem Sniffer-Tool „httpFox“, erstellt von Max Bahne am 03.06.2016 um 13:59 Uhr.

Abbildung 31: Screenshot aus dem Sniffer-Tool „httpFox“, erstellt von Max Bahne am 03.06.2016 um 14:28 Uhr.

Abbildung 32: Bearbeiteter Screenshot aus dem Abgehakt-Video der extra3-Sendung vom 25.05.2016, erstellt von Max Bahne am 04.06.2016 um 14:37 Uhr.

Abbildung 33: Diagramm zum Arbeitsablauf bei der Produktion eines Videos, erstellt von Max Bahne am 04.06.2016 um 15:08 Uhr.

Abbildung 34: Diagramm zum Clustering der Sendungen nach Formaten, erstellt von Max Bahne am 09.06.2016 um 13:58 Uhr.

Abbildung 35: Diagramm zum Clustering nach Themen der Formate, erstellt von Max Bahne am 09.06.2016 um 14:02 Uhr.

Abbildung 36: Tabelle der Verteilung der Themen in der Mediatheks-Videos, erstellt von Max Bahne am 11.06.2016 um 13:21 Uhr.

Abbildung 37: Implementierung eines Recommender Systems in die ndr.de-Seite, erstellt von Max Bahne am 03.07.2016 um 12:00 Uhr.

Versicherung über Selbständigkeit

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.

Hamburg, den _____