

Masterthesis
Harald Kirschenmann

Automatisierte Informationsergänzung von Fernsehdokumentationen

FAKULTÄT TECHNIK UND INFORMATIK
Department Informatik

Faculty of Computer Science and Engineering
Department Computer Science

Harald Kirschenmann

Automatisierte Informationsergänzung von Fernsehdokumentationen

Masterarbeit eingereicht im Rahmen der Masterprüfung
im Studiengang *Informatik*
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

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

Eingereicht am: 12. Oktober 2021

Harald Kirschenmann

Thema der Arbeit

Automatisierte Informationsergänzung von Fernsehdokumentationen

Stichworte

Fernsehdokumentation, Informationsergänzung, Second Screen, StopWord Removal, Named Entity Recognition, Named Entity Linking, Geocoding, Wissensbasis, Kartensystem

Kurzzusammenfassung

In dieser Arbeit wurde ein System zur automatisierten Informationsergänzung von Fernsehdokumentationen konzipiert, entwickelt und evaluiert. Die Generierung der Informationsergänzungen erfolgt über die Verarbeitung der Untertitel der Fernsehdokumentationen mit Verfahren des StopWord Removals, Named Entity Recognitions, Named Entity Linkings und Geocodings. Die Darstellung der Informationsergänzungen erfolgt parallel zum Abspielen der Fernsehdokumentation auf einem Second Screen. Das entwickelte System wurde funktional und mit der qualitativen Inhaltsanalyse von durchgeführten Experteninterviews evaluiert.

Harald Kirschenmann

Title of Thesis

Automated information supplementation of documentaries

Keywords

Documentary, information supplementation, second screen, stopword removal, named entity recognition, named entity linking, geocoding, knowledge base, mapping platform

Abstract

In this thesis a system for automated information supplementation of documentaries was designed, developed and evaluated. The information supplements are generated by processing the subtitles of the documentaries using stopword removal, named entity recognition, named entity linking and geocoding methods. The information supplements are displayed on a second screen while watching the documentary. The developed system was evaluated functionally and with the qualitative content analysis of conducted expert interviews.

Inhaltsverzeichnis

1	Einleitung	9
1.1	Zielsetzung	10
1.2	Gliederung	10
2	Analyse	11
2.1	Zielsetzung und Anforderungen	12
2.1.1	Auswahl geeigneter Videos	12
2.1.2	Darstellung der Videos	13
2.1.3	Automatisierte Generierung der Informationsergänzungen	13
2.1.4	Darstellung der Informationsergänzungen	13
2.1.5	Generalisierbarkeit	14
2.2	Anforderungsanalyse	14
2.2.1	Auswahl der Videos	14
2.2.2	Generierung der Informationsergänzungen	14
2.2.3	Darstellung	15
2.3	Second Screen-Phänomen	16
2.4	Datengrundlage	17
2.4.1	Videos	17
2.4.2	Wissensbasen	18
2.4.3	Geodaten	22
2.5	Vorverarbeitung	23
2.5.1	Tokenizer	24
2.5.2	StopWord Removal	25
2.5.3	Named Entity Recognition	26
2.6	Named Entity Linking	29
2.7	Geocoding	31
2.8	Bisherige Arbeiten	32
3	Design und Realisierung	34
3.1	Schichtenmodell	35

3.1.1	Präsentationsschicht	37
3.1.2	Logikschicht.....	38
3.1.3	Datenhaltungsschicht	40
3.1.4	Representational State Transfer (REST)	45
3.1.5	Message Broker.....	47
3.2	Schematischer Programmablauf.....	51
3.2.1	First Screen-Komponente.....	51
3.2.2	Second Screen-Komponente	53
3.3	Ablauf der Generierung der Hintergrundinformationen	54
3.3.1	Vorverarbeitung	56
3.3.2	Entity Linking	57
3.4	Darstellung und Interaktion.....	59
3.4.1	First Screen.....	60
3.4.2	Second Screen	63
4	Evaluation.....	67
4.1	Funktionale Evaluation	67
4.1.1	Precision, Recall und F-Score	68
4.1.2	Videos für die Evaluation.....	69
4.1.3	Vorverarbeitung	70
4.1.4	Geocoding	72
4.1.5	Named Entity Linking.....	75
4.1.6	Geschwindigkeit.....	78
4.1.7	Diskussion.....	81
4.2	Nutzerbefragung.....	85
4.2.1	Forschungsmethode.....	86
4.2.2	Interviewleitfaden.....	90
4.2.3	Interviewpartner und Interviewdurchführung	92
4.2.4	Auswertung	96
4.2.5	Gütekriterien	125
4.3	Gesamtbetrachtung.....	127
5	Zusammenfassung und Fazit	130

6	Ausblick.....	132
7	Literatur.....	134
H	Anhang 1	141
H.1	Messungen der Vorverarbeitung	141
H.1.1	Digitale Zukunft	141
H.1.2	Der große Umbruch.....	141
H.1.3	Wann kommt der Corona-Impfstoff?	142
H.2	Messungen des Geocodings	142
H.2.1	Digitale Zukunft	142
H.2.2	Der große Umbruch.....	143
H.2.3	Wann kommt der Corona-Impfstoff?	143
H.3	Messungen des Named Entity Linking	144
H.3.1	Digitale Zukunft	144
H.3.2	Der große Umbruch.....	144
H.3.3	Wann kommt der Corona-Impfstoff?	145
H.4	Interviewleitfaden	145
H.5	Kurztitel und stichpunktartige Zusammenfassung	148
H.5.1	Interview - B1	148
H.5.2	Interview - B2	149
H.5.3	Interview - B3	150
H.5.4	Interview - B4	151
H.5.5	Interview - B5	152
H.5.6	Interview - B6	153

Abbildungsverzeichnis

Abbildung 1: Bedeutung der Wörter nach der Frequenz der Wörter im Text (vgl. [32])	25
Abbildung 2: Beispielsatz mit identifizierten und klassifizierten Named Entities.....	26
Abbildung 3: Beispiel eines Textes mit Anmerkungen. Zwei Entitäten wurden mit Informationen der Wikipedia verknüpft.....	29
Abbildung 4: Vereinfachte Darstellung der benötigten Komponenten.....	34
Abbildung 5: Schichtenmodell der Software [2].....	36
Abbildung 6: Schematische Darstellung eines Message Brokers	47
Abbildung 7: Schematische Darstellung des Producer/Consumer-Prinzips (vgl. [2])	49
Abbildung 8: Schematische Darstellung des Publish/Subscribe-Prinzips (vgl. [2])	50
Abbildung 9: Ablauf des First Screen als Flussdiagramm [2]	52
Abbildung 10: Ablauf des Second Screen als Flussdiagramm [2].....	53
Abbildung 11: Ablauf der Content-Generierung für die Second Screen - Anwendung als Flussdiagramm [2]	55
Abbildung 12: Darstellung der Videoauswahl des First Screens	60
Abbildung 13: Darstellung der Seite des First Screens zum Abspielen der Videos	62
Abbildung 14: Darstellung der Second Screen-Seite zur Wahl des passenden First Screens.	64
Abbildung 15: Darstellung von zwei automatisiert generierten Nachrichten	65
Abbildung 16: Darstellung einer automatisch generierten Nachrichte mit Kartendarstellung	66
Abbildung 17: Precision-Mittelwerte der Verfahren der Geocoding-Messungen	73
Abbildung 18: Recall-Mittelwerte der Verfahren der Geocoding-Messungen	74
Abbildung 19: F ₁ -Score-Mittelwerte der Verfahren der Geocoding-Messungen	75
Abbildung 20: Precision-Mittelwerte der Verfahren der NEL-Messungen	76
Abbildung 21: Recall-Mittelwerte der Verfahren der NEL-Messungen.....	77
Abbildung 22: F ₁ -Score-Mittelwerte der Verfahren der NEL-Messungen	78
Abbildung 23: Durchschnittliche Berechnungsdauer der Vorverarbeitung in Sekunden	79

Abbildung 24: Durchschnittliche Berechnungsdauer des Geocodings in Sekunden	79
Abbildung 25: Durchschnittliche Berechnungsdauer pro Entität der Geocoding-Verfahren in Millisekunden.....	80
Abbildung 26: Durchschnittliche Berechnungsdauer des Named Entity Linkings in Sekunden	80
Abbildung 27: Durchschnittliche Berechnungsdauer pro Entität der NEL-Verfahren in Millisekunden.....	81
Abbildung 28: Ablaufmodell der qualitativen Inhaltsanalyse.....	88
Abbildung 29: Verteilung der Merkmale der sechs befragten Experten.....	94

1 Einleitung

Digitale Medien sind zu einem ständigen Begleiter im Alltag geworden. Mit dem Fernseher haben erstmals Geräte in jedem Haushalt Einzug erhalten, die Bewegtbilder abspielen. Die Verbreitung von Computern folgte im privaten und beruflichen Feld. Bereits hier hat sich das Verhalten der Menschen nachhaltig verändert. Mit dem Einzug der Smartphones sind digitale Medien für die meisten Menschen ständig verfügbar. Auch weiterhin ist der Medienkonsum einem steten Wandel unterworfen. Seit Anfang des Jahres 2020 hat die Corona-Pandemie zu einem starken Anstieg des Medienkonsums geführt, mit dem auch der Wandel des Konsumverhaltens beschleunigt wurde. Die eingeschränkten Kontaktmöglichkeiten führten dazu, dass die Menschen digital miteinander kommunizierten. Im Privaten haben sich die Menschen digital über die Pandemie informiert und im beruflichen Kontext wurden Konferenzsysteme eingeführt, um das Homeoffice zu ermöglichen. Sowohl im Bereich von Schule und Universität als auch in anderen Bildungsbereichen wurde die Lehre in den digitalen Raum verlegt. Für die digitale Lehre und das digitale Lernen wurden bestehende Angebote genutzt und neue Angebote geschaffen.

In der schulischen Lehre und im Privaten wurden und werden Fernsehdokumentationen zur Vermittlung von Wissen angesehen. Auch für den Zeitvertreib und aus reinem Interesse werden Fernsehdokumentationen konsumiert. Früher erhielten solche Formate die ungeteilte Aufmerksamkeit der Konsumenten. Heutzutage jedoch gehört für viele Menschen die parallele Verwendung eines mobilen Zweitgeräts zum Alltag. Neben der Fernsehdokumentation wird das Zweitgerät zur Ablenkung, Kommunikation und auch zur Suche von weiterführenden Informationen verwendet. All diese Handlungen benötigen Zeit und lenken vom eigentlichen Inhalt der Fernsehdokumentation ab. In der Regel bezieht sich die manuelle Informationssuche auf die Inhalte der Fernsehdokumentationen; doch gerade die manuelle Suche kostet viel Zeit und lenkt vom Hauptbildschirm ab. Würden die gesuchten Informationen bereits zur Verfügung stehen, so würde sich die Situation stark verbessern. Die Konsumenten würden unterstützt werden und könnten sich besser auf den Inhalt der Fernsehdokumentation konzentrieren.

Ein System, das die Informationssuche automatisiert und die gesuchten Informationen dem Anwender in Form von Informationsergänzungen zur Verfügung stellt, kann hier Abhilfe schaffen und zu einer Verbesserung des Verständnisses und des Erlebnisses von Fernsehdokumentationen führen.

1.1 Zielsetzung

Ziel dieser Arbeit ist es, die Konsumenten von Fernsehdokumentationen mit Hilfe von Informationsergänzungen beim Verständnis und dem Erlebnis der Fernsehdokumentationen zu unterstützen. Die Konsumenten sollen Informationen nicht manuell suchen müssen, wenn sie diese zum Verständnis benötigen oder aus Interesse am Inhalt weitere Informationen wünschen. Damit sich die Konsumenten auf die Fernsehdokumentation fokussieren können, soll die Informationssuche und Darbietung der Informationen vom zu entwickelnden System übernommen werden.

Es wird ein System entwickelt, das die Informationsergänzungen anhand von Fernsehdokumentationen automatisch generiert und diese Informationen dem Konsumenten zur Verfügung stellt. Die Fernsehdokumentationen werden durch die Informationen angereichert, indem Inhalte vertieft und inhaltliche Aspekte erläutert werden. Die generierten Informationsergänzungen stehen den Konsumenten parallel zur Betrachtung der Fernsehdokumentationen zur Verfügung. Mit den Ausarbeitungen des Grundprojekts [1] und Hauptprojekts [2] erfolgten bereits Vorarbeiten, auf denen diese Arbeit aufbaut und die Zielsetzung der erfolgten Arbeiten erweitert.

Darüber hinaus soll dieses neu entwickelte System unter verschiedenen Aspekten untersucht werden. Zum einen soll das System funktional evaluiert werden. Zum anderen soll untersucht werden, inwiefern derartige Informationsergänzungen von Fernsehdokumentation und die entsprechende Darstellung das Verständnis und Erlebnis dieser Fernsehdokumentationen fördern.

Das System soll im Rahmen der Untersuchungen zu verwenden sein. Es gehört nicht zum Ziel dieser Arbeit, dass das System Marktreife erlangt. Dies beinhaltet, dass es nicht ausgeschlossen ist, dass es weitere Verbesserungen für Teile des Systems gibt.

1.2 Gliederung

Diese Arbeit wird in mehrere Kapitel gegliedert, die aufeinander aufbauen. Nach dem einleitenden Kapitel 1 erfolgt in Kapitel 2 die Analyse der Zielsetzung dieser Arbeit und die Formulierung der Anforderungen, die die Software erfüllen muss, um die Zielsetzung zu erfüllen. Außerdem werden Daten und Datenquellen vorgestellt, die bei der Erfüllung der Zielsetzung helfen. Und es erfolgt die Diskussion von Methoden, mit denen die Daten im Hinblick auf die Erfüllung der Zielsetzung verarbeitet werden können.

In Kapitel 3 wird ausgeführt, wie die zu entwickelnde Software realisiert wurde. Ebenfalls werden die Design-Entscheidungen, die für die Realisierung der Software getroffen werden, besprochen. Hierzu gehört die Darstellung der Entscheidungen, welche Datenquellen und welche Verfahren zur Datenverarbeitung verwendet werden. Außerdem wird aufgezeigt, wie

die Komponenten des Software-Systems aufgebaut sind und wie diese miteinander interagieren.

In Kapitel 4 wird das entwickelte System evaluiert. Zum einen erfolgt die funktionale Evaluation, bei der die Güte der Verfahren zur Erzeugung der Informationsergänzungen untersucht wird. Und zum anderen wird das entwickelte System mit der qualitativen Inhaltsanalyse von durchgeführten Experteninterviews im Hinblick auf eine formulierte Forschungsthese, bei der der Anwendungsaspekt des Systems im Vordergrund steht, analysiert.

Die wichtigen Aspekte der Arbeit werden in Kapitel 5 zusammengefasst und anschließend wird ein Fazit der Arbeit gezogen. Abschließend wird in Kapitel 6 ein Ausblick auf zukünftige Arbeiten und Entwicklungen gegeben.

2 Analyse

Im Rahmen dieser Arbeit wird ein System entwickelt, mit dem Informationsergänzungen von Fernsehdokumentationen automatisch generiert und den Konsumenten der Fernsehdokumentationen dargeboten werden. In Abschnitt 2.1 wird die Zielsetzung vertieft und die Anforderungen an das zu entwickelnde System ausgearbeitet. In Abschnitt 2.2 werden die ausgearbeiteten Anforderungen analysiert und die Implikationen auf die zu entwickelnde Software dargestellt. Im nachfolgenden Abschnitt 2.3 wird das Second Screen-Phänomen erläutert und in Abschnitt 2.4 wird beschrieben, welche Daten für ein solches System zur Verfügung stehen. In dem Abschnitt 2.5 werden Methoden vorgestellt, mit denen eine erste Verarbeitung der verfügbaren Daten erfolgen kann. Die Methode des *Named Entity Linking* (NEL) zur weiteren Verarbeitung der Daten wird im Abschnitt 2.6 dargelegt. Mit dem NEL können die Daten der Filme mit den Daten von Wissensbasen verknüpft werden. In Abschnitt 2.7 wird die Methode des *Geocodings* vorgestellt. Hierbei werden die Daten der Filme mit Geodaten verknüpft, so dass Ortsinformationen in geographische Koordinaten umgewandelt werden. In dem Abschnitt 2.8 werden Arbeiten vorgestellt, die mit dieser Arbeit oder Teilen dieser Arbeit vergleichbar sind.

2.1 Zielsetzung und Anforderungen

Beim Konsum des Fernsehprogramms ist die parallele Verwendung von mobilen Endgeräten wie Tablets, Smartphones und Laptops weit verbreitet. Diese Geräte werden unabhängig von dem Programm des Fernsehschwerers genutzt, aber auch um sich über die Inhalte des Fernsehprogramms weitergehend zu informieren. Diese Informationen stehen nicht direkt zur Verfügung, sondern müssen von den Anwendern händisch über Suchmaschinen ermittelt werden. Der Fokus der Fernseh-Konsumenten liegt hierbei auf dem mobilen Endgerät und wird dort für die Dauer der Informationssuche gebunden. Die Informationssuche lenkt die Anwender vom Fernsehprogramm ab und kann langwierig sein. Je mehr Informationen gesucht werden, desto größer ist der Anteil der Zeit, der für die Informationssuche verwendet werden muss. Dementsprechend weniger erfassen die Anwender von den Inhalten des Fernsehprogramms mit. Das zu entwickelnde System soll die Anwender bzw. die Konsumenten von Fernsehdokumentationen unterstützen, indem die Suche nach weiteren inhaltsbezogenen Informationen automatisiert wird und die generierten Informationsergänzungen dem Konsumenten, parallel zur Betrachtung der Fernsehdokumentation, verfügbar gemacht werden.

Für die erfolgreiche Umsetzung eines solchen Systems müssen mehrere Anforderungen umgesetzt und erfüllt werden, die nachfolgend dargestellt werden.

2.1.1 Auswahl geeigneter Videos

Die Anwender des zu entwickelnden Systems sollen mit dem Angebot von Informationsergänzungen beim Konsum von Videos unterstützt werden. Die Informationsergänzungen sollen inhaltlich passend zu den Videos generiert werden. Demzufolge bilden die Videos die Basis der Informationsergänzungen. Die Anwender sollen die Möglichkeit erhalten, Videos auszuwählen, die sie mit diesem System betrachten wollen.

Es gibt eine große Menge an Videos und nicht jedes ist für ein solches System geeignet. Die Videos sollen über Untertitel verfügen und Informationen vermitteln, so dass diese vom System vertieft und erweitert werden können. Darüber hinaus sollten die Videos nicht zu kurz sein, da ansonsten die Zeit für die Anwender zu kurz ist, um einerseits das Video und andererseits die Informationsergänzungen nebenläufig zu verarbeiten. Um die Nutzer des Systems nicht zu ermüden und die Konzentration zu beeinträchtigen, sollten die Videos nicht zu lang sein. Das Genre der Videos beeinflusst viele der genannten Faktoren. Fernsehdokumentationen beispielsweise erfüllen die genannten Anforderungen.

2.1.2 Darstellung der Videos

Dem Anwender sollen die Informationsergänzungen parallel zum Betrachten der Videos dargeboten werden. Um dies zu ermöglichen, müssen die Videos auf einem Bildschirm dargestellt werden. Damit die Anwender das Video in Ruhe ansehen können und auch bei der parallelen Betrachtung von Informationsergänzungen viel von den Inhalten des Videos registrieren, bietet sich die Verwendung von großen Bildschirmen wie bei einem Fernseher an.

Der Anwender soll in der Lage sein, die Videos zu starten und zu pausieren. Auch das Überspringen von Teilen der Videos oder das Zurücksetzen auf eine bestimmte Laufzeit der Videos soll dem Anwender ermöglicht werden.

2.1.3 Automatisierte Generierung der Informationsergänzungen

Die Konsumenten der Videos sollen mit Hilfe der Informationsergänzungen unterstützt werden, so dass sie die Informationen nicht explizit suchen müssen, sondern diese direkt vom System zur Verfügung gestellt bekommen. Diese Informationsergänzungen sollen automatisch und inhaltlich zu den Videos passend generiert werden.

Viele Informationsarten sind denkbar, die die Anwender der Software bei der Betrachtung der Videos unterstützen könnten. Im Rahmen dieser Arbeit soll eine möglichst große Bandbreite an Informationen vertieft und ergänzt werden können, so dass die Anwender bei einer Vielfalt von Videos mit den Informationsergänzungen unterstützt werden. Außerdem sollen Inhalte der Videos, die mit bestimmten Orten verknüpft sind, mit weiteren Informationen und einer Kartendarstellung angereichert werden.

2.1.4 Darstellung der Informationsergänzungen

Die Informationsergänzungen müssen den Konsumenten parallel zur Fernsehdokumentation zur Verfügung gestellt werden, damit diese die Informationsergänzungen nebenläufig verarbeiten können. Die Informationen sollen den Anwendern visuell zur Verfügung gestellt werden. Die auditive oder haptische Bereitstellung der Informationen ist nicht Gegenstand dieser Ausarbeitung. Die visuelle Darstellung der Informationsergänzungen soll nicht das Videobild überlagern oder von diesem ablenken.

Die Informationsergänzungen sollen synchronisiert zum Video verfügbar sein. Sobald die Inhalte, für die es Informationsergänzungen gibt, in dem Video besprochen werden, sollen die Infos für den Anwender zur Betrachtung zur Verfügung stehen.

2.1.5 Generalisierbarkeit

Es steht eine Vielzahl an Videos zur Verfügung, die den zuvor formulierten Anforderungen entsprechen. Das zu entwickelnde System soll nicht nur auf einzelne Videos, die die Anforderungen erfüllen, anwendbar sein. So sollen mit den Videos unterschiedliche Themen betrachtet werden können und die Videos auch unterschiedlich lang sein können. Das zu entwickelnde System soll anwendbar auf alle Videos sein, die die genannten Anforderungen erfüllen.

2.2 Anforderungsanalyse

In diesem Abschnitt werden die in Abschnitt 2.1 dargestellten Anforderungen an das zu entwickelnde System analysiert und dargestellt, welche weiteren Aspekte hierdurch impliziert werden bzw. mit welchen Daten und Verfahren ein solches System umgesetzt werden könnte.

In Abschnitt 2.2.1 wird die Auswahl der Videos und in Abschnitt 2.2.2 werden die Anforderungen an die Generierung der Informationsergänzungen näher betrachtet. In Abschnitt 2.2.3 erfolgt die Betrachtung der Darstellung der Informationsergänzung und in Abschnitt 2.3 wird das *Second Screen*-Phänomen vorgestellt. Mit Hilfe eines Second Screens kann die Darstellung der Informationsergänzungen anforderungskonform durchgeführt werden.

2.2.1 Auswahl der Videos

An die Videos werden in Abschnitt 2.1.1 einige Anforderungen gestellt. Darunter fällt die Länge der Videos. Auch sollen Untertitel verfügbar sein und Informationen vermittelt werden, die vertieft werden können. Bei Fernsehdokumentationen wird auf die Länge geachtet, da Wissen vermittelt werden soll. Sie können nicht zu kurz sein, da ansonsten kaum Inhalt erläutert werden kann. Um Ermüdung zu vermeiden, wird auch darauf geachtet, dass die Fernsehdokumentationen nicht zu lang sind. Fernsehdokumentationen erfüllen die Anforderungen und werden daher für die Entwicklung und Evaluation der Software zu Grunde gelegt. Eine tiefergehende Betrachtung von Videos und Fernsehdokumentationen erfolgt in Abschnitt 2.4.1.

2.2.2 Generierung der Informationsergänzungen

In Abschnitt 2.1.3 wurden die Anforderungen an die Generierung der Informationsergänzungen festgelegt. Die Informationsergänzungen sollen automatisch generiert werden und inhaltlich passend zu den Videos sein. Wie in 2.1.1 definiert, stehen Untertitel für jedes Video zur Verfügung. Diese Untertitel müssen von dem zu entwickelnden

System verarbeitet, um aus diesen die Informationsergänzungen zu generieren. Im Rahmen dieser Arbeit wird nicht näher auf die Generierung der Informationsergänzungen auf Basis der Video- und Audiospur eingegangen.

Die Informationsergänzungen sollen eine möglichst große Bandbreite an Informationen abdecken. Um dies zu erreichen, müssen relevante Inhalte aus den Untertiteln extrahiert werden und mit weiterführenden Informationen verknüpft werden. Das Identifizieren bzw. Extrahieren von Kandidaten von relevanten Inhalten kann durch eine Vorverarbeitung der Untertitel erfolgen. Verschiedene Verfahren der Vorverarbeitung der Untertitel werden in Abschnitt 2.5 vorgestellt. Anschließend müssen die Kandidaten mit weiterführenden Informationen verknüpft werden. Das Verknüpfen dieser Kandidaten mit anderen Datenquellen, die weiterführende Informationen enthalten, wird als *Named Entity Linking* bezeichnet und in Abschnitt 2.6 erläutert. Für das Named Entity Linking werden Wissensbasen als Datenquelle benötigt. Diese sollten möglichst umfangreich und thematisch breit aufgestellt sein, um die Anforderungen aus den Abschnitten 2.1.3 und 2.1.5 zu erfüllen.

Eine weitere Anforderung ist es, dass Geodaten identifiziert und ebenfalls mit weiteren Informationen angereichert werden sollen. Für diesen Aspekt kann ebenfalls auf die Vorverarbeitung und das Named Entity Linking zurückgegriffen werden. Darüber hinaus sollen die Geodaten mit einer Kartendarstellung angereichert werden. Hierfür werden Kartensysteme benötigt.

Im Abschnitt 2.4 werden die Daten, die zur Verfügung stehen und für die Umsetzung eines solchen Systems benötigt werden, erläutert. Es wird näher auf die Videodaten eingegangen und einige Wissensbasen, die die Grundlage für das Named Entity Linking bilden und die vertiefenden Informationen enthalten, werden erläutert. Außerdem werden Kartensysteme und Quellen für Geodaten vorgestellt.

2.2.3 Darstellung

Entsprechend der Anforderungen der Abschnitte 2.1.2 und 2.1.4 sollen das Video und die Informationsergänzung für den Anwender zur Betrachtung bereitgestellt werden. Der Anwender soll in der Lage sein, zeitgleich das Video und die Informationsergänzungen zu verarbeiten, ohne dass die Informationsergänzungen das Videobild überlagern oder dauerhaft vom Video ablenken. Hierdurch entfallen Lösungen wie das Einblenden im Video oder ein geteilter Bildschirm.

Das Video soll auf einem großen Bildschirm gezeigt werden. Hierfür bietet sich ein Fernseher oder ein anderer großer Bildschirm an. Mit der Darstellung der Informationsergänzungen auf einem mobilen Zweitgerät können die Anwender diese ihrem Bedarf entsprechend betrachten. Werden keine Informationsergänzungen benötigt oder erwünscht, so ist es einfach, das mobile Zweitgerät nicht zu beachten. Sobald die Informationsergänzungen erwünscht sind, ist das

mobile Zweitgerät schnell verfügbar und intuitiv zu bedienen. Die Anwender können selbst entscheiden, welche Haltung des Zweitgeräts am komfortabelsten ist. Die parallele Verwendung zweier Bildschirme wird von vielen Menschen alltäglich genutzt. Dieses Phänomen wird als Second Screen bezeichnet und in Abschnitt 2.3 erläutert.

2.3 Second Screen-Phänomen

Mit der fortschreitenden Verbreitung mobiler digitaler Endgeräte entwickelte sich ein Phänomen, das das Nutzerverhalten nachhaltig verändert hat. Neben dem Konsum des Fernsehprogramms wurde mit der parallelen Nutzung eines mobilen Zweitgeräts begonnen. Das Fernsehprogramm erhielt den größten Anteil der Aufmerksamkeit der Konsumenten und lief auf dem Hauptbildschirm, häufig ein Fernseher. Das Gerät, auf dem der Hauptfokus liegt, wird als First Screen bezeichnet. Das mobile Endgerät, das nebenläufig zum Fernsehkonsum verwendet wird, wird als Second Screen bezeichnet. Lohmüller & Wolff [3] erweitern eine Definition von Second Screen und schlagen folgende Definition vor:

„The act of second screening involves the use of an additional media screen such as provided by a mobile phone, tablet, or laptop, to consume any kind of content alongside a primary screen such as a TV“ ([3], S. 171)

Die Definition besagt, dass die Handlung des Second Screenings eines zusätzlichen Geräts bedarf, das ein Tablet, Smartphone oder Laptop sein kann, und dieses parallel zum Konsum von Inhalten auf einem Hauptbildschirm, wie einem Fernseher, genutzt wird.

Der Zweck der Verwendung eines Second Screens kann vielfältig sein. So kann das Zweitgerät beispielsweise zur Ablenkung bzw. Unterhaltung, aber auch zur Informationsrecherche, Arbeit und auch zur Kommunikation mit anderen Menschen genutzt werden. Wird der Second Screen zum Austausch mit anderen Konsumenten über das Fernsehprogramm verwendet, so wird auch vom *Social Television* bzw. *Social TV* gesprochen [4]. Werden zwei oder mehr Geräte nebenläufig verwendet, so wird auch der Begriff Multi-Screening genutzt.

In mehreren Studien, die von großen Firmen beauftragt oder durchgeführt wurden, wurde das Nutzungsverhalten der Fernsehkonsumenten im Hinblick auf die Second Screen-Nutzung untersucht. Ein großer Teil der Bevölkerung verwendet bereits einen Second Screen. Eine dieser Studien ergab, dass 86 % der Befragten zwischen 14 und 59 Jahren im Jahr 2013 einen Second Screen verwendeten ([5], S. 12f). In der nachfolgenden Studie erhöhte sich dieser Wert im Jahre 2015 auf 89%. Fast die Hälfte davon nutzte einen Second Screen mindestens täglich und 39% mehr als einmal wöchentlich ([6], S. 18ff). Die Zahlen einer anderen Studie liegen etwas niedriger mit 81% im Jahr 2020, das Alter der Befragten lag zwischen 14 und 69 Jahren. Bei dieser Befragung nutzten 30% häufig und 18% manchmal einen Second Screen. Die

Verbreitung und Häufigkeit der Second Screen-Nutzung ist vom Alter abhängig und geschieht häufig unabhängig vom Inhalt des First Screens ([7], S. 10)[8].

Zumeist wird davon ausgegangen, dass die größeren Bildschirme, wie ein Fernseher, den Hauptbildschirm bzw. First Screen darstellen und die mobilen Endgeräte dem Second Screen entsprechen. Wird die Aufmerksamkeitsverteilung betrachtet, so ist diese Reihenfolge nicht immer stimmig. Häufig erhalten nicht die großen Bildschirme die meiste Aufmerksamkeit, sondern die mobilen Endgeräte. Daher gibt es eine Diskussion über die Verwendung der Termini First- und Second Screen ([4], S. 192) ([6], S. 41). Im Rahmen dieser Ausarbeitung wird der Begriff First Screen für den Hauptbildschirm, auf dem das Video abgespielt wird, und der Begriff Second Screen für das mobile Endgerät, auf dem die Informationsergänzungen dargestellt werden, verwendet.

2.4 Datengrundlage

Um die Zielsetzung der Ausarbeitung und der Anforderungen an die Software umzusetzen, werden Daten benötigt. Die Grundlage der Software bilden die Daten der Videos (siehe 2.4.1), mit denen die weiteren Daten in Bezug gebracht werden. Die Untertitel der Videos sollen mit weiteren Informationen angereichert werden. Für die Anreicherung der Untertitel mit Zusatzinformationen wird eine Datenquelle benötigt, die eben diese Zusatzinformationen enthält und mit der es möglich ist, die Untertitel mit der gewünschten Information zu verknüpfen. Diese Daten sind in Wissensbasen enthalten (siehe 2.4.2). Für die Umwandlung von Ortsbeschreibungen in Geokoordinaten werden Geodaten und für die Darstellung von Kartenausschnitten werden Kartendaten benötigt (siehe 2.4.3).

Die Auswahl dieser Datenquellen erfolgte aufgrund des Umfangs, der Verfügbarkeit und Lizenzierung der Daten. Die aufgeführten Datenquellen sind exemplarisch und können durch andere Daten ersetzt werden. Die aufgelisteten Datensätze enthalten viele Daten und werden kontinuierlich erweitert. Sie sind leicht zu erhalten und mit sehr freien Lizenzen ausgestattet. In der Ausarbeitung des Grundprojekts [1] wurden Teile der nachfolgend erläuterten Daten bereits beschrieben.

2.4.1 Videos

Für diese Software sind die Daten der Videos besonders wichtig, denn die Informationsergänzungen, die für die Anwendung benötigt werden, werden passend zu den Videos generiert. Die Daten der Videos setzen sich aus der eigentlichen Videodatei, den dazugehörigen Metadaten und den Untertiteln zusammen.

Die Videodatei besteht aus einer Audiospur und den bewegten Bildern, der Videospur. Zusätzlich sind Metadaten in der Videodatei gespeichert, die u.a. die Auflösung, Framerate und

Länge des Videos bestimmen. In den Metadaten können weitere vielfältige Informationen enthalten sein. Hier können beispielsweise Daten zu den Urhebern oder *Geo-Tags* gespeichert werden. *Geo-Tags* geben eine nähere Beschreibung über den Aufnahmeort von Fotos und Videos. Sie bestehen aus Daten wie den GPS-Koordinaten, Höhenangaben oder Namen von Ländern, Regionen und Ortschaften.

Viele Videos werden mit Beschreibungstexten präsentiert, die den semantischen Inhalt der Videos kurz zusammenfassen oder beschreiben. Zusätzlich stehen häufig Untertitel für die Videos zur Verfügung. Untertitel sind kurze Texte, die einem bestimmten Zeitabschnitt der Videos zugeordnet werden und parallel zum Abspielen der Videos dargestellt werden. Mit Hilfe der Untertitel werden nicht-gesprochene und gesprochene Inhalte der Videos schriftlich dargestellt. Hierdurch können unverständliche gesprochene Worte angezeigt, Fremdsprachen übersetzt oder höreingeschränkten und gehörlosen Menschen die Möglichkeit gegeben werden, die Audiospur als Text zu lesen. Die Awareness, dass klassische Videos mit Audiospur für viele Menschen eine Barriere darstellen, führt zu einem vermehrten Angebot von Untertiteln. Öffentlich-Rechtliche Medien bauen ihr Angebot aus und achten vermehrt auf die barrierefreie Nutzung ihrer Angebote^{1,2}. Aus der Audiospur können die Untertitel mit Hilfe der Spracherkennung automatisiert generiert werden. Die Qualität der Spracherkennung sollte in weiterer Forschung verbessert werden, so dass unter anderem Menschen mit Behinderung besser von dieser besser profitieren können [9][10].

2.4.2 Wissensbasen

Eine Wissensbasis ist eine Sammlung großer Mengen von Wissen, die in strukturierter und unstrukturierter Form vorliegt. Das Wissen kann in Datenbanken hinterlegt und maschinell verarbeitet werden. Beispiele von Wissensbasen sind Yago [11] und Wolfram|Alpha [12]. Auch Wikipedia bzw. Wikidata und DBPedia werden zu den Wissensbasen gezählt [13]. Die Untersuchung von Färber et al. zeigt auf, dass die Wissensbasen sich für unterschiedliche Zwecke verschieden gut eignen [14]. In der Untersuchung schneiden DBPedia und Wikidata gut ab und sind über ihre Lizenzierung verfügbar. Im Rahmen dieser Ausarbeitung werden mit Wikipedia und DBPedia zwei Wissensbasen vorgestellt, die im weiteren Verlauf zur Anwendung kommen.

¹<https://www.ndr.de/fernsehen/service/Barrierefreie-Angebote-im-NDR,barrierefrei141.html>
(Aufgerufen am 16. Dezember 2020)

² https://www.ndr.de/der_ndr/unternehmen/rundfunkrat/Barrierefreiheit-Untertitelquote-im-NDR-Fernsehen-81-Prozent-Rundfunkrat-begruesst-Fortschritte-pressemeldungndr18164.html (Aufgerufen am 16. Dezember 2020)

Wikipedia

Im Januar 2001 entstand ein schnell wachsendes Projekt der *Wikimedia Foundation*³ mit dem Namen Wikipedia. In vielen Ländern wurden unabhängige Vereine gegründet, die die Wikimedia Foundation unterstützen. Die enzyklopädische Wissensbasis Wikipedia wird in kollaborativer Arbeitsweise entwickelt und umfasste Anfang Oktober 2021 bereits 57,4 Millionen Artikel in über 300 Sprachen⁴. Jede Sprachversion wird unabhängig voneinander entwickelt. Mit 6,2 Millionen Artikeln ist die englische Sprachversion die größte. Mit Hilfe eines Bots, der Artikel automatisiert generiert, folgen die Sprachversionen Cebuano und Schwedisch. Mit 2,5 Millionen Artikeln ist die deutsche Sprachversion die viertgrößte. Die Texte von Wikipedia werden unter der CC-BY-SA – Lizenz veröffentlicht und sind somit unter Namensnennung frei verfügbar.

Wikipedia besteht aus verschiedenen Arten von Seiten. Die Artikel bilden den wichtigsten Baustein von Wikipedia und enthalten den Großteil des Wissens. In den Artikeln gibt es Verlinkungen auf andere Artikel der gleichen Wikipedia Sprachversion. Durch die Verweise sind die Artikel untereinander vernetzt und ermöglichen das Navigieren zwischen verbundenen Artikeln. Das als Drill-Down bezeichnete Navigieren von allgemeineren zu spezifischeren Artikeln kann beim Information Retrieval angewendet werden. Das Information Retrieval bezeichnet den Vorgang des Auffindens von gesuchten Informationen in einer großen Datenmenge, wie beispielsweise Datenbanken.

In vielen Sprachen haben Begriffe mehrere Bedeutungen. Dies wird als Mehrdeutigkeit oder Ambiguität bezeichnet. Die semantische Unterscheidung dieser Begriffe ist eine Herausforderung des Natural Language Processing (NLP) und als Wortsinn-Disambiguierung bekannt [15]. Wikipedia hat die gleiche Herausforderung, da es nicht einen Artikel für mehrere Sachverhalte geben soll. In der Wikipedia wird dieses Problem durch die sogenannten *Disambiguation Pages* (DP) adressiert. Die DPs werden für mehrdeutige Bezeichnungen erstellt, um die jeweiligen Bedeutungen auf einer Seite aufzuführen. Die DPs bestehen aus internen Verlinkungen auf die entsprechenden Artikel. Mitte Dezember 2020 enthält die deutschsprachige Wikipedia über 308.000 DPs⁵. Dies bietet keine vollständige Abdeckung der Ambiguitäten, kann aber für das NLP zur Auflösung von Ambiguitäten als Unterstützung dienen [16][17][18].

Die offene und anonyme Arbeitsweise stellt nur geringe Hürden, um an dem Projekt mitzuwirken und führt zu vielen freiwilligen Unterstützern, wodurch die Anzahl der Artikel

³ <https://wikimediafoundation.org/> (aufgerufen am 29. Dezember 2020)

⁴ https://en.wikipedia.org/wiki/List_of_Wikipedias#Grand_total (aufgerufen am 09. Dezember 2021)

⁵ <https://de.wikipedia.org/wiki/Kategorie:Begriffskl%C3%A4rung> (Aufgerufen am 30. September 2021)

schnell ansteigt⁶. Durch die vielen Mitwirkenden können Fehler leicht erkannt und korrigiert werden. Die Artikel entstehen in kollaborativer Arbeitsweise von meist anonymen freiwilligen Nutzern. Die Anonymität und die große Menge der freiwilligen Unterstützer führten dazu, dass sich Wikipedia eigene Grundsätze und eine Organisationsstruktur gegeben hat, um die Qualität der Artikel zu gewährleisten.

Die offene und anonyme Arbeitsweise führte zu positiven Effekten, gibt aber auch Anlass zur Kritik an dem Projekt. So sei die Qualität der Artikel schwankend. Des Weiteren könnten die Inhalte von Artikeln leicht nach eigenen Bedürfnissen beeinflusst werden. Außerdem können Gruppendynamiken Auswirkungen auf das gesamte Projekt haben.

Wikipedia gibt mit Wikimedia Statistics⁷ einen Einblick in die internen Metriken. Die bereits aufgeführte Zahl von 55,4 Millionen Artikeln und von 21,3 Milliarden Seitenaufrufen im September 2021 verdeutlichen die Bedeutung von Wikipedia. Zur Zeit der ersten Covid-19-Welle zu Beginn des Jahres 2020 stieg die Zahl der monatlichen Aufrufe auf über 25 Milliarden an. Die deutsche Sprachversion von Wikipedia erreichte im Jahr 2020 etwa 1 Milliarde monatliche Aufrufe.

*Wikidata*⁸, *Wiktionary*⁹ und weitere sind Projekte unter der *Wikimedia Foundation*, die an die Grundsätze von Wikipedia angelehnt und zum Teil mit Wikipedia verzahnt sind. Mit *Wiktionary* wird ein frei zugängliches mehrsprachiges Wörterbuch geschaffen. Das *Wikidata*-Projekt startete im Oktober 2012, mit dem Ziel strukturierte Daten zu speichern und unter CC0 1.0 - Lizenz zur freien Verfügung zu stellen. Wie Wikipedia wird auch Wikidata durch die kollaborative Mitarbeit von Freiwilligen erweitert [19]. Bei Wikipedia sind die Artikel hauptsächlich im Fließtext geschrieben. Im Gegensatz hierzu sind die Daten von Wikidata für das maschinelle Lesen konzipiert und ermöglichen komplexere Abfragen. In Wikipedia wurde viel Wissen zusammengetragen, von dem Wikidata profitiert. Doch auch Wikipedia und viele andere Projekte der *Wikimedia Foundation* und darüber hinaus profitieren von Wikidata. Im Dezember 2020 wurden Datenelemente von Wikidata in über 70 % der Wikipedia-Artikel verwendet¹⁰. Während anfangs Wikidata Informationen von Wikipedia einpflegte, bindet Wikipedia nun die Daten von Wikidata sprachübergreifend ein. Änderungen oder Aktualisierungen von Daten müssen somit nur einmalig in Wikidata erfolgen und nicht in jeder einzelnen Sprachversion von Wikipedia. Zugunsten von Wikidata wurde die Freebase-Datenbank eingestellt und die Daten in das Projekt migriert [20].

⁶ <https://stats.wikimedia.org/#/all-projects/content/pages-to-date/normal|line|2-year|~total|monthly> (Aufgerufen am 30. September 2021)

⁷ <https://stats.wikimedia.org/> (Aufgerufen am 09. Oktober 2021)

⁸ <https://www.wikidata.org/> (Aufgerufen am 29. Dezember 2020)

⁹ <https://www.wiktionary.org/> (Aufgerufen am 29. Dezember 2020)

¹⁰ https://wcm.wmflabs.org/WD_percentUsageDashboard/ (Aufgerufen am 29. Dezember 2020)

DBPedia

Das Projekt *DBPedia* wurde 2007 gegründet, um strukturierte Daten aus dem Wikipedia-Projekt zu extrahieren und diese zur Verfügung zu stellen [21]. Wikipedia enthält viel Wissen, das zusammengetragen wird und in Infoboxen und im Fließtext der Öffentlichkeit zur Verfügung gestellt wird. DBPedia erfasst diese Daten mehrerer Sprachversionen von Wikipedia und extrahiert diese in strukturierter Form. Zurzeit gibt es 20 Sprachversionen von DBPedia¹¹. Neben Projekten von Wikimedia erfasst und verknüpft DBPedia auch Daten von weiteren Projekten, wie YAGO2 [22][23] und GeoNames (siehe 2.4.3) [24].

Etwa 5 Jahre nach DBPedia wurde das Wikidata-Projekt gegründet und stellt ebenfalls strukturierte Daten zur Verfügung. Beide Projekte kooperieren miteinander. In der Umsetzung und Arbeitsweise jedoch unterschieden sich beide voneinander, u.a. in den genutzten Datenstrukturen und in der Herangehensweise, die Daten zu sammeln. DBPedia extrahiert die Daten aus anderen Projekten und profitiert so von der Qualität und dem wachsenden Datensatz der genutzten Projekte. Bei Wikidata dagegen werden die Daten von Freiwilligen eingepflegt und mit Verweisen auf die Herkunft der Daten versehen [25].

Monatlich wird ein aktualisierter DBPedia-Datensatz unter der CC BY-SA-Lizenz veröffentlicht. Bereitgestellt werden die Daten im RDF-Format¹², eine SPARQL-Schnittstelle¹³ und über HTTP-URIs zur Verfügung gestellt. Die SPARQL-Schnittstelle ermöglicht komplexe Anfragen. Eine REST-API, die den Zugriff auf die SPARQL-Schnittstelle vereinfachen soll, befindet sich in der Entwicklung¹⁴.

Mit DBPedia Spotlight steht eine Software zur Verfügung, die die strukturierten Daten von DBPedia nutzt, um Informationen eines Textes mit Informationen aus der Wissensbasis DBPedia zu verknüpfen [26].

¹¹ <https://wiki.dbpedia.org/join/chapters> (Aufgerufen am 31. Januar 2021)

¹² <https://www.w3.org/TR/rdf-schema/> (Aufgerufen am 2. Januar 2021)

¹³ <https://www.w3.org/TR/rdf-sparql-query/> (Aufgerufen am 2. Januar 2021)

¹⁴ <https://wiki.dbpedia.org/rest-api> (Aufgerufen am 2. Januar 2020)

2.4.3 Geodaten

Viele Handlungsstränge von Fernsehdokumentationen und anderen Filmgenres werden mit Ortsangaben erzählt und können einen wichtigen Aspekt der Filme ausmachen. Um die Ortsangaben in den Filmen nutzen zu können, sind Geodaten nötig. Die Europäische Union definiert Geodaten als „*alle Daten mit direktem oder indirektem Bezug zu einem bestimmten Standort oder geografischen Gebiet*“ [27]. Die Anwendungsgebiete von Geodaten sind vielfältig. Routenplaner und Kartendienste sind auf Geodaten angewiesen. Auch für Simulationen und Datenvisualisierungen können Geodaten ein wichtiger Baustein sein. Mit der GeoNames-Datenbank und OpenStreetMaps werden nachfolgend zwei Quellen für Geodaten vorgestellt.

GeoNames-Datenbank

Mit *GeoNames* ist eine Datenbank entstanden, die unter der cc-by – Lizenz veröffentlicht wird und eine umfangreiche Sammlung an geographischen Daten bereitstellt. Die umfangreichste Tabelle enthält über 12 Millionen Einträge von Geo-Objekten. Die Geo-Objekte können Länder, Städte und andere besondere Orte repräsentieren. Für die Geo-Objekte lassen sich Informationen, wie die Namen, GPS-Koordinaten, Zeitzone und die Einwohnerzahl finden.

Für viele Geo-Objekte werden verschiedene Bezeichnungen verwendet. Selbst innerhalb einer Sprache kann es mehrere Namen für Geo-Objekte geben, wie „The Big Apple“ für die Stadt New York City. Auch in verschiedenen Sprachen werden viele Geo-Objekte anders benannt. Beispielsweise wird die Hauptstadt Finnlands auf Finnisch als Helsinki und auf Schwedisch als Helsingfors bezeichnet. Diese zahlreichen Namensalternativen werden in der GeoNames-Datenbank abgebildet. Somit wird eine Unterstützung von zahlreichen Sprachen erreicht.

Unter Einhaltung der Lizenz kann die Datenbank auf eigenen Datenbank-Servern betrieben werden.

OpenStreetMap

OpenStreetMap (OSM) ist ein internationales Projekt, das es sich zum Ziel gesetzt hat, eine freie Weltkarte zu erschaffen [29]. Hinter OSM steht eine Datenbank mit Geodaten, die ständig wächst. Die Datenbank wird von vielen Freiwilligen unterstützt, die freie Kartendaten in die Datenbank einpflegen oder Rohdaten sammeln. Das Sammeln der Daten kann mit Hilfe von GPS-Geräten und dem anschließenden Einpflegen in die Datenbank erfolgen. Änderungen in der realen Welt können in der Datenbank schnell abgebildet werden, da die Daten von den Freiwilligen geändert werden können. Dies führt dazu, dass die Daten von OSM ständigen Änderungen unterworfen sein können.

Da das Projekt durch das Einpflegen von Kartendaten durch Freiwillige wächst, ist die Abdeckung der Kartendaten abhängig von der Beteiligung der Freiwilligen. Deutschland ist gut erfasst und die großen Städte sind besonders gut abgedeckt.

Mitte Dezember 2020 enthielt die Datenbank 7,29 Milliarden Objekte in der Datenbank¹⁵. Dies umfasst die drei grundlegenden Datenstrukturen von OpenStreetMaps, der *Node*, *Way* und *Relation*.

Die kleinste Struktur ist eine *Node*, die aus einer ID und einer GPS-Koordinate besteht. Mit 6,55 Milliarden Nodes in der Datenbank umfassen diese den Großteil der Datenbank. Nodes repräsentieren einfache Gegenstände, wie Parkbänke, Müllbehälter und Skulpturen einer Karte oder können die Koordinaten von komplexeren Datenstrukturen sein.

Eine komplexere Datenstruktur ist die *Way*, die mit 725 Millionen Elementen in der Datenbank die zweithäufigste Struktur ist. Eine *Way* besteht aus einer geordneten Liste von Nodes. Mit Hilfe von Ways werden linienartige Strukturen, wie Straßen und Flüsse, aber auch Objekte wie Häuser der realen Welt in der Datenbank abgebildet.

Komplexer als eine *Way* ist die *Relation*, die aus einer geordneten Liste von Nodes, Ways und anderen Relationen bestehen kann. Die Elemente einer Relation stehen in einer logischen oder geografischen Beziehung zueinander. Relationen können in verschiedene Relationsarten unterteilt werden. Dazu gehören Routen, die beispielsweise zur Abbildung von Fahrstrecken von öffentlichen Verkehrsbetrieben dienen, oder auch Multipolygone, mit denen komplexe Flächen, wie Waldstücke mit Lichtungen oder Häuser mit Innenhöfen, gebildet werden können. In der OSM Datenbank sind 8,4 Millionen Relationen enthalten.

Die OpenStreetMap-Datenbank wird unter „Open Database Licence (ODbL) 1.0“ – Lizenz¹⁶ veröffentlicht und kann u.a. zur Kartendarstellung verwendet werden.

2.5 Vorverarbeitung

Im Bereich des *Natural Language Processing* (NLP) ist die Vorverarbeitung der Daten ein wichtiger Schritt, der die Qualität und Performance der gesamten Verarbeitung beeinflussen kann. Die Vereinheitlichung der Groß- und Kleinschreibung reduziert die Komplexität von Texten und kann Fehlerquellen reduzieren. Texte werden initial als Zeichenkette übertragen. Durch einen *Tokenizer* (siehe 2.5.1) wird die Zeichenkette in die einzelnen Wörter aufgeteilt.

¹⁵ https://taginfo.openstreetmap.org/reports/database_statistics (Aufgerufen am 16. Dezember 2020)

¹⁶ <https://opendatacommons.org/licenses/odbl/> (Aufgerufen am 16. Dezember 2020)

Mit Stemming ist es möglich, Wörter auf ihren Wortstamm zurückzuführen und somit die Variationen, die jedes Wort haben kann, zu reduzieren.

Beim Verarbeiten der Untertitel der Videos reduziert das Herausfiltern von Wörtern oder Teilen von Texten den Textkorpus. Dies kann durch das *StopWord Removal* (siehe 2.5.2), bei dem Wörter ohne Mehrwehrt entfernt werden sollen, oder auch dem komplexeren *Named Entity Recognition* (siehe 2.5.3), bei dem wichtige Entitäten des Textes identifiziert werden sollen, geschehen. Das *Named Entity Recognition* ermöglicht eine Einteilung der Entitäten in zuvor festgelegte Kategorien und somit eine spätere gezieltere Weiterverarbeitung. Ein solches System kann z.B. auf Grundlage der Bibliotheken NLTK und SpaCy implementiert werden.

Die aufgeführten Verfahren sind nur ein Ausschnitt der Möglichkeiten zur Vorverarbeitung. Die Vorverarbeitung bietet die Chance, die Resultate der Weiterverarbeitung zu verbessern. Allerdings gibt es keine allgemeingültige Pipeline für alle Anwendungsfälle [30]. Bei den vorgestellten Verfahren gehen Informationen verloren, die für weitergehende Verarbeitungsschritte hinderliche oder positive Auswirkungen haben können.

2.5.1 Tokenizer

Im Rahmen dieser Arbeit sollen die Untertitel von Videos verarbeitet werden. Die Untertitel liegen als zusammenhängende Zeichenketten vor. Um diese Zeichenkette in lexikalische Elemente, wie Wörter und Zahlen, zu unterteilen, wird ein *Tokenizer* verwendet.

Der Tokenizer verarbeitet die Zeichenketten und zerlegt diese in zusammenhängende Einheiten, sogenannte Tokens. Sonderzeichen, wie Punkte und Kommata, werden während des Vorgangs verworfen. In vielen Sprachen zeigen Leerzeichen und Sonderzeichen eine Trennung von Wörtern an. Das Aufteilen von Zeichenketten in Tokens ist nicht trivial, denn die einfache Nutzung von Leerzeichen und Sonderzeichen als Trennung von Tokens kann zu ungewünschten Effekten führen. Hierdurch können zusammengehörige Zeichen getrennt werden. Eigennamen, wie “Los Angeles” bilden eine Einheit und sollen nicht in allen Fällen in die zwei Tokens “Los” und “Angeles” aufgeteilt werden. Ähnlich verhält es sich bei Wörtern, wie das englische “aren’t”. Hier ist fraglich, ob das Wort als ein oder zwei Tokens behandelt werden soll. Es gibt mehrere Optionen zur Auswahl, die die Resultate einer späteren Verarbeitung beeinflussen können. So kann beispielsweise das Matching einer Anfrage mit den ermittelten Tokens zu unterschiedlichen Ergebnissen führen, wenn die Tokens der Anfrage nicht durch das gleiche Verfahren erzeugt wurden. Derartige Schwierigkeiten sind abhängig von der Sprache ([31], S. 22-26).

Neben dem Aufteilen von natürlicher Sprache in Tokens, werden Tokenizer auch zur Vorverarbeitung von formalen Sprachen, wie Programmiersprachen, verwendet und finden in Compilern Anwendung.

2.5.2 StopWord Removal

Die häufigsten Wörter einer Sprache, die ohne gesonderte Bedeutung sind, werden als Stoppwörter bezeichnet. Die ersten Stoppwörter wurden 1958 von Hans Peter Luhn identifiziert und zeichneten die häufigsten vorkommenden Wörter aus, denen keine besondere Bedeutung für den Textinhalt beigemessen wird [32].

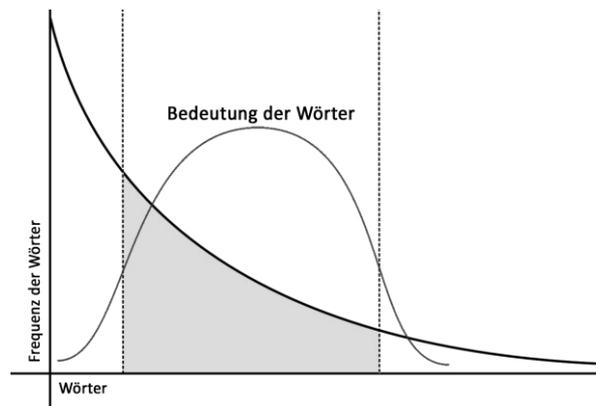


Abbildung 1: Bedeutung der Wörter nach der Frequenz der Wörter im Text (vgl. [32])

H. P. Luhn beschrieb die Bedeutung der Wörter in Abhängigkeit zu ihrer Häufigkeit wie in Abbildung 1. Wörter, die sehr häufig vorkommen, haben für die Bedeutung bzw. Semantik von Texten eine eher niedrige Bedeutung. Ebenso verhält es sich mit Wörtern, die sehr selten vorkommen. Die Semantik von Sätzen wird demnach hauptsächlich von Wörtern getragen, die weder ganz häufig vorkommen noch ganz selten. Werden die häufigsten und die seltensten Begriffe aus einem Textkorpus entfernt, so bleibt demnach der Großteil der Bedeutung erhalten. Wo die Schwelle verläuft, welche bzw. wie viele Begriffe entfernt werden können, so dass ausreichend Semantik erhalten bleibt, ist nicht festgelegt und kann sich je nach Anwendungsfall unterscheiden.

Zu den Stoppwörtern gehören Artikel und Präpositionen. Wie beschrieben, können diese aus dem Textkorpus entfernt werden, ohne dass der semantische Inhalt relevant reduziert wird. Zur Identifizierung gibt es statische und dynamische Ansätze. In klassischen Verfahren werden die Texte mit Hilfe eines Tokenizers in lexikalische Elemente unterteilt und anschließend die Stoppwörter manuell, anhand der Häufigkeiten, in einer Liste zusammengeführt [33]. Derartige Listen existieren für viele Sprachen. Je nach Bedarf können diese unterschiedlich umfangreich ausfallen.

Die Stoppwörter können aus dem Textkorpus entfernt werden, um den Umfang zu reduzieren und so den Umfang nachfolgender Verarbeitungen zu reduzieren. So müssen weniger Wörter verarbeitet werden, was zur Beschleunigung und Verbesserung der Performance der Verfahren führen kann [33][34]. Allerdings kann die Anwendung von Stoppwörtern auch zu einer nicht

signifikanten Verbesserung oder gar Verschlechterung der Performance nachfolgender Verfahren führen [35]. Die Schlagwortsuche wird vom StopWord Removal nicht negativ beeinflusst, allerdings kann dies in der Phrasensuche geschehen, da auch Stoppwörter die Semantik einer Aussage beeinflussen können. So enthält die Phrase „Chancellor of Germany“ ein Stoppwort und präzisiert die Aussage gegenüber den zwei getrennten Tokens „Chancellor“ und „Germany“ ([31], S. 27). Der Einfluss des StopWord Removals auf die spätere Verarbeitung ist abhängig vom Anwendungsfall.

2.5.3 Named Entity Recognition

Beim StopWord Removal (siehe 2.5.2) werden Begriffe ohne besondere Bedeutung entfernt. Im Gegensatz hierzu werden bei der Named Entity Recognition (NER) Begriffe extrahiert und in Kategorien eingeordnet, die für den Inhalt der Texte von Bedeutung sind. Eine spätere Verarbeitung dieser Begriffe kann auf Basis der Klassifizierung gezielt stattfinden.

Der Begriff Named Entity wurde bei der 6. Message Understanding Conference (MUC-6) geprägt [36] und wird im Bereich des Natural Language Processing (NLP) verwendet. Der Begriff beschreibt abstrakte oder physisch existente Objekte der realen Welt mit Eigennamen. Dies kann Personen, Orte, Organisationen, Produkte und weitere Objekte umfassen. Named Entities stehen nicht für sich allein, sondern können in Beziehungen zu anderen Named Entities stehen. Implizite Beziehungen zu Named Entities, also nicht direkt benannten Entitäten, sind für den Menschen leicht zu erkennen und erweitern das Verständnis der Begriffe oder der Semantik von Texten.

Das NER beschreibt den Prozess des Identifizierens und Klassifizierens von Named Entities aus einem Text. Für die Klassifizierung der Named Entities kann es viele Kategorien geben, die nicht festgelegt sind und sich abhängig von der Domäne bzw. vom (Fach-)Bereich unterscheiden können. Mit Person, Ort und Organisation sind drei häufig verwendete Kategorien exemplarisch benannt. Weitere Kategorien können beispielsweise auch Datums- und Zeitangaben sein. Ebenso ist es möglich, dass die Kategorien spezifischer definiert werden, so dass beispielsweise Institutionen, Firmen und Gerichte voneinander unterschieden werden [37][38].



Abbildung 2: Beispielsatz mit identifizierten und klassifizierten Named Entities.

In Abbildung 2 wird ein Beispiel dieses Prozesses dargestellt. In dem dargestellten Satz werden die drei Eigennamen "Peter", "Hamburg" und "Airbus" als Named Entity erkannt und in die entsprechenden Kategorien einsortiert. "Peter" wird als Person (PERS), "Hamburg" als Ort

(LOC) und “Airbus” als Organisation (ORG) erkannt. Der Prozess des NER bildet den Grundstein vieler Anwendungen im Bereich des NLP.

Das Problem des Identifizierens und Klassifizierens von Named Entities ist für einfache Wörterbuchansätze zu komplex. Die Bezeichnungen können sich ändern, in vielen grammatikalischen Formen auftreten, abgekürzt werden, durch Pseudonyme oder andere alternative Namen ersetzt werden und der Umfang eines derartigen Wörterbuches wäre sehr groß. Des Weiteren kann die Klassifizierung von Named Entities abhängig vom Kontext sein, so dass sich mit der Wortsinn-Disambiguierung eine weitere Herausforderung ergibt.

Moderne Ansätze setzen auf das *Machine Learning* und erzielen hiermit sehr gute Ergebnisse. Im Besonderen Deep Learning-Verfahren wurden in den letzten Jahren eingesetzt, da diese aus den Rohdaten selbstständig lernen können und die Entwickler kein domänenspezifisches Wissen über die Texte benötigen. Viele Verfahren wurden für eine spezifische Sprache entwickelt, da sich Sprachen in ihrem Wortschatz und der Struktur unterscheiden. Es gibt mehrere Verfahren, die für mehrere Sprachen geeignet sind. Die multilingualen Ansätze erzielen ebenfalls gute Ergebnisse [39][40].

Für das NER-Problem wurden zahlreiche Algorithmen entwickelt. Mit dem Natural Language Toolkit und SpaCy stehen zwei Open Source - Bibliotheken für das NER-Problem zur Verfügung. Die NLTK- und SpaCy-Bibliothek sind umfangreiche Bibliotheken im Bereich des NLP, leicht zu erhalten und mit freien Lizenzen, wie der MIT¹⁷- und Apache-2.0¹⁸-Lizenz ausgestattet. Beide Bibliotheken sind exemplarisch benannt und können durch andere Bibliotheken für das NER-Problem ausgetauscht werden.

Natural Language Toolkit (NLTK)

Das Natural Language Toolkit (NLTK)¹⁹ [37] ist eine umfangreiche Open Source - Bibliothek für Probleme im Bereich des NLP. Die Bibliothek wurde in Python realisiert und sollte die Lehre der Computerlinguistik bzw. des NLP vereinfachen. Hierbei soll NLTK einfach anzuwenden, erweiterbar und gut dokumentiert sein [41]. Die Bibliothek wird unter der Apache-2.0-Lizenz veröffentlicht.

Das Problem des NER lässt sich mit dem NLTK angehen. Für die englische Sprache steht ein vortrainiertes Modell zur Verfügung, welches mit einem englischen *Automatic Content Extraction (ACE)*-Datensatz des *Linguistic Data Consortium*²⁰ trainiert wurde. Für weitere Kategorien und andere Sprachen müssen eigene Modelle trainiert werden. Im NLTK steht der

¹⁷ <https://opensource.org/licenses/mit-license.php> (Aufgerufen am 13. Januar 2021)

¹⁸ <https://www.apache.org/licenses/LICENSE-2.0.html> (Aufgerufen am 13. Januar 2021)

¹⁹ <https://www.nltk.org/> (Aufgerufen am 4. Januar 2021)

²⁰ <https://www ldc.upenn.edu/> (Aufgerufen am 15. Januar 2021)

CoNLL-2020 Datensatz [42] zur Verfügung. Dieser enthält Trainingsdaten in Spanisch und Niederländisch.

SpaCy

Eine weitere Open Source - Bibliothek im Bereich des NLP ist SpaCy²¹ [43]. Während die Motivation der Entwickler des NLTK in der Lehre von Computerlinguistik bzw. NLP lag, legt SpaCy den Fokus auf die Entwicklung von Software im Bereich des NLP. Die Bibliothek wurde in Python und Cython entwickelt und wird mit der MIT-Lizenz²² zur Verfügung gestellt.

SpaCy enthält Algorithmen für verschiedene Probleme des NLP. Darunter fällt auch das Problem des NER, für das verschiedene vortrainierte Modelle angeboten werden. Für 15 Sprachen stehen jeweils drei vortrainierte Modelle zur Verfügung. Ein weiteres Modell, das NER für mehrere Sprachen unterstützt, wurde mit einem mehrsprachigen Wikipedia-Datensatz trainiert. In den mehrsprachigen Trainingsdaten bilden die englischsprachigen Artikel, gefolgt von deutschsprachigen Artikeln und sieben weiteren Sprachen, die Mehrheit [44]. Die deutschsprachigen Sprachmodelle wurden mit dem TIGER und dem WikiNER-Korpus trainiert. Die Qualität der vortrainierten Modelle unterscheidet sich anhand der Sprachversion und der genutzten Trainingsdaten. Für die deutschsprachigen vortrainierten Modelle sind die Kategorien „Person“, „Ort“, „Organisation“ und „Verschiedenes“ verfügbar.

Für die Unterstützung weiterer Sprachen und Kategorien oder für Modelle spezieller Themenbereiche müssen neue Modelle trainiert werden. Mit Hilfe der Thinc²³-Bibliothek können Modelle anderer *Machine Learning*-Frameworks, wie Tensorflow²⁴ [45] und PyTorch²⁵ [46], mit SpaCy verwendet werden.

Die SpaCy-Bibliothek wird kontinuierlich weiterentwickelt. Projekte des Entwicklerteams und der Community erweitern SpaCy um zusätzliche Funktionalitäten. Mit SpaCy 3.0 ist eine neue Version der Bibliothek in der Entwicklung.

²¹ <https://spacy.io/> (Aufgerufen am 4. Januar 2021)

²² <https://opensource.org/licenses/mit-license.php> (Aufgerufen am 31. Januar 2021)

²³ <https://thinc.ai> (Aufgerufen am 15. Januar 2021)

²⁴ <https://www.tensorflow.org/> (Aufgerufen am 15. Januar 2021)

²⁵ <https://pytorch.org/> (Aufgerufen am 15. Januar 2021)

2.6 Named Entity Linking

Im Abschnitt 2.5.2 wurde eine Methode beschrieben, wie ein Textkorpus verkleinert werden kann. Mit dem Named Entity Recognition wurde in Abschnitt 2.5.3 ein Prozess beschrieben, mit dem Named Entities gezielt erkannt und klassifiziert werden können. Mit beiden Verfahren erfolgt eine Vorverarbeitung des Eingabetextes. Die Named Entities sollen genutzt werden, um diese mit entsprechenden Einträgen einer Wissensbasis (siehe 2.4.2) zu verknüpfen. Wird das StopWord Removal zur Vorverarbeitung genutzt, so sollen aus dem Pool an Wörtern passende Verbindungen zu Wissensbasen gefunden werden. Begriffe, bei denen keine passenden Verbindungen gefunden werden, sollen mit keinerlei Einträgen einer Wissensbasis verknüpft werden. Das Abbilden von Named Entities auf die entsprechenden Einträge einer Wissensbasis wird als *Named Entity Linking* bezeichnet [47]. Die Vorverarbeitung von Eingabetexten kann die Ergebnisse des Named Entity Linking verbessern.

Peter wohnt in **Hamburg** und arbeitet bei **Airbus**



Abbildung 3: Beispiel eines Textes mit Anmerkungen. Zwei Entitäten wurden mit Informationen der Wikipedia verknüpft.

Named Entity Linking findet in verschiedenen Bereichen Anwendung. Einer dieser Bereiche ist die Anreicherung von Texten mit Hintergrundinformationen. In Abbildung 3 ist ein solches Beispiel exemplarisch mit Infoboxen der Wikipedia dargestellt. Aus dem gegebenen Satz werden die Named Entities mit Einträgen einer Wissensbasis verknüpft. Eine solche Verknüpfung erfolgt für die Entitäten “Hamburg” und “Airbus”, für “Peter” wurde kein passender Eintrag gefunden. Weitere Bereiche, in denen das Named Entity Linking genutzt wird, sind Question-Answering-Systeme, automatisierte Inhaltsanalysen und das automatisierte Füllen von Wissensbasen [48][49][50].

Wissensbasen sind oft auf bestimmte Themenfelder spezialisiert und enthalten über diese hinaus kaum Informationen. Auch umfangreiche und themenübergreifende Wissensbasen, wie Wikipedia (siehe 2.4.2), DBPedia (siehe 2.4.2), YAGO oder Wolfram|Alpha, enthalten kein vollständiges Wissen, sondern sind in ihrem Umfang begrenzt. Abhängig vom

Anwendungsfall können sich viele Wissensbasen für das Named Entity Linking anbieten. In dieser Arbeit werden die Untertitel von Videos mit Hintergrundinformationen angereichert. Die Videos sind nicht auf einzelne Themenfelder beschränkt, daher sollten umfangreiche und themenübergreifende Wissensbasen verwendet werden. Wissensbasen, die nur auf bestimmte Themenfelder spezialisiert sind, bieten sich nicht an.

In vielen wissenschaftlichen Veröffentlichungen wird Wikipedia als Wissensbasis für das *Named Entity Linking* - Problem verwendet. Für die Nutzung von Wikipedia als Wissensbasis wurde der Begriff *Wikification* geprägt [51][52]. Die Autoren Mihalcea und Csomai, die den Begriff mit ihrer wissenschaftlichen Veröffentlichung einführten, definieren Wikification als die Aufgabe die wichtigsten Worte und Phrasen eines Dokuments automatisiert zu extrahieren und für jedes dieser Worte und Phrasen eine passende Verknüpfung zu einem Wikipedia-Artikel zu finden (vgl. [51]). Bei der Wikification findet eine Vorverarbeitung der Eingabetexte statt, um die wichtigen Worte und Phrasen zu extrahieren. Anschließend wird das Named Entity Linking auf die extrahierten Worte und Phrasen angewendet. In aktuellen wissenschaftlichen Arbeiten wird das Named Entity Linking und im Speziellen auch das Wikification weiterentwickelt. Während sich viele Arbeiten um die Verbesserung hinsichtlich der Qualität der Ergebnisse kümmern, arbeiten einige aktuelle Arbeiten an der Beschleunigung des Wikification-Prozesses [53][54]. In einer weiteren aktuellen Veröffentlichung wird Wikification zur Verbesserung des Verständnisses von Texten zu Covid-19 eingesetzt [55].

Der Prozess des Named Entity Linking ist nicht trivial. Entitäten, die mit einer Wissensbasis verknüpft werden sollen, sind nicht immer eindeutig. Begriffe und auch ganze Sätze können mehrdeutig sein. Als Beispiel für einen solchen mehrdeutigen Begriff in der deutschen Sprache kann der Begriff "Bienenstich" sein, der zum einen für den Stich einer Biene und zum anderen für das Gebäck stehen kann. Auch der Begriff "Bank" ist mehrdeutig, er kann eine Sitzbank oder auch ein Geldinstitut repräsentieren. Aus diesem Grund sind das Named Entity Linking und die Wortsinn-Disambiguierung eng miteinander verbunden. Mit Hilfe der Wortsinn-Disambiguierung kann ein Named Entity Linking-System bei mehrdeutigen Entitäten entscheiden, mit welchem Eintrag der Wissensbasis die Entität verknüpft werden soll.

Eine weitere Herausforderung stellen Akronyme, Alias-Namen und unterschiedliche Schreibweisen eines Begriffes dar. Mehrere Bezeichnungen, Schreibweisen und Akronyme können die gleiche Entität repräsentieren. Diese sollen mit dem gleichen Eintrag einer Wissensbasis verknüpft werden. Ein Beispiel dieses Phänomens wäre das Vereinigte Königreich, das häufig mit UK abgekürzt wird.

2.7 Geocoding

In der Kartographie werden heutzutage geographische Koordinaten zur eindeutigen Bestimmung eines Standortes verwendet. Über die Erdkugel wird ein Netz aus Längen- und Breitengraden gelegt. Gemeinsam bilden diese Werte ein Koordinatensystem für die Erde. Es findet vielfache Anwendung. Beispielsweise bildet es die Grundlage für Navigationssysteme und Geräte wie Fitnessarmbänder und wird demzufolge von vielen Menschen täglich genutzt.

Auf den Karten der Kartensysteme können die verschiedensten Objekte der Erde dargestellt werden. Zu diesen geographischen Objekten gehören beispielsweise Länder, Kontinent, Städte, Ortschaften, Gebirge, Flüsse, Meere und Straßen. Als Oberbegriff für die Bezeichnungen dieser geographischen Objekte wird der Begriff *geographischer Name* verwendet. Der Begriff Toponym umfasst diese Bezeichnungen ebenfalls. Toponyme jedoch sind weiter gefasst und umfassen ebenfalls Bezeichnungen für Objekte außerhalb der Erde. In dieser Ausarbeitung soll die Kartendarstellung nur für Objekte auf der Erde erfolgen, daher wird der Begriff geographischer Name nachfolgend verwendet.

Bei der Anwendung von Kartendiensten oder Navigationssystemen gibt der Anwender selten geographische Koordinaten in das System ein, sondern üblicherweise einen Text, der den gewünschten Ort bestimmt. Dieser Text kann eine eindeutige Beschreibung des Ortes, die vollständige Adresse oder der konkrete geographische Name sein. Im Hintergrund des Systems wird dieser Text in eine geographische Koordinate umgewandelt. Der Prozess des Umwandeln einer Beschreibung bzw. Bezeichnung in eine geographische Koordinate wird als Geocoding bezeichnet und wird durch umfangreiche Datenbanken ermöglicht [56].

In vielen Kartensystemen kommt auch das Reverse-Geocoding zum Tragen. Hier ist es möglich, das System nach dem geographischen Namen eines beliebigen Punktes der Karte anzufragen. Das System ermittelt daraufhin die Adresse oder den geographischen Namen der gewählten Koordinate.

Geocoding- und Reverse-Geocoding-Systeme finden in vielen Bereichen Anwendung. Die Qualität der Systeme ist in vielen Gebieten sehr gut. In anderen Gebieten jedoch fehlen ausreichende Datenbestände oder die grundlegenden Adress-Systeme unterscheiden sich, so dass das Geocoding noch verbessert werden kann [57][58]. Im Bereich des Geocodings wird an weiteren Verbesserungen gearbeitet [59][60]. Eines der Arbeitsfelder zur Verbesserung ist hierbei die Verarbeitung von räumlichen Beschreibungen, wie beispielsweise der Beschreibung, dass ein gesuchter Ort zwischen zwei anderen Orten liegt.

Im Rahmen dieser Arbeit kann das Geocoding für die Zuordnung von Entitäten zu geographischen Koordinaten genutzt werden. In den Untertiteln von Videos sind häufig geographische Namen enthalten, denen geographische Koordinaten zugeordnet werden sollen,

damit diese auf einem Kartensystem visualisiert werden können. Nicht jede Entität der Untertitel entspricht einem geographischen Namen, so dass eine Vorverarbeitung der Untertitel sinnvoll ist. Mit Hilfe des *Named Entity Recognition* (siehe 2.5.3) kann die Menge der Entitäten gezielt auf mögliche geographische Namen reduziert werden. Mit Hilfe des *Stop Word Removals* (siehe 2.5.2) wird zwar die Menge der Entitäten verkleinert, jedoch nicht gezielt auf geographische Namen reduziert. Mit dem Geocoding wird eine weitere Filterung der Entitäten durchgeführt. Entitäten, denen keine geographischen Koordinaten zugeordnet werden können, werden nicht weiter als geographische Namen betrachtet.

Für das Geocoding werden Geodaten benötigt. Mit OpenStreetMap und der GeoNames-Datenbank wurden zwei Systeme vorgestellt, die zahlreiche Geodaten enthalten und mit denen Geocoding durchgeführt werden kann (siehe 2.4.3).

2.8 Bisherige Arbeiten

In dem Bereich des Second Screens und der Informationsergänzungen wurden bereits vergleichbare Arbeiten durchgeführt. Ivan Demin hat in seiner Masterarbeit eine Second Screen-Plattform entwickelt, bei der dem Anwender Videos auf dem First Screen und zusammengefasste Informationen auf dem Second Screen dargeboten werden. Das Ziel der Arbeit war es, zu untersuchen, ob die Informationsaufnahme der Anwender durch die Verwendung eines Second Screens verbessert wird [61].

Zuvor veröffentlichten Odijk et al. [62] eine Arbeit, bei der Inhalte des Fernsehprogramms anhand der Untertitel in Echtzeit mit Wikipedia verknüpft und auf einem Second Screen dem Nutzer angeboten werden. Angewendet wurde dieses System auf Talkshow-Formate.

In dieser Ausarbeitung werden Fernsehdokumentationen bzw. die Untertitel dieser geocodiert, um den Anwendern Informationsergänzungen in Form von Geoinformationen zu bieten. Im Jahr 2012 veröffentlichten Li et al. ihre Arbeit an einem System, mit dem Videos automatisiert Geocodiert werden [63]. Für die Geocodierung wurden Text-Komponenten, wie die Metadaten der Videos und das bewegte Bild verwendet.

Der Second Screen ist eine wichtige Komponente für diese Arbeit. Die Informationsergänzungen sollen dem Anwender auf einem Second Screen bereitgestellt werden, so dass diese die Informationen parallel zum Video auf dem First Screen betrachten kann. Cauwenberge et al. [64] untersuchten den Effekt eines Second Screens auf die kognitive Belastung und das Lernvermögen der Anwender. Im Testszenario haben die Probanden auf dem First Screen Nachrichten-Formate konsumiert und parallel hierzu auf dem Second Screen Informationen erhalten. Untersucht wurde auch der Effekt von themenbezogenen und themenfremden Informationen auf die Probanden. In einer nachfolgenden Arbeit von Schaap et al. [65] wurde untersucht, welchen Einfluss es auf die Lernfähigkeit hat, ob Informationsergänzungen bereitgestellt werden, aktiv gesucht werden müssen oder es keine

Informationsergänzungen gibt. Die Untersuchungen wurden ebenfalls im Kontext von Nachrichtenformaten auf dem First Screen durchgeführt.

Die nebenläufige Anwendung eines Second Screens zum Fernsehprogramm ist weit verbreitet. Die Konsumenten des Fernsehprogramms informieren sich parallel u.a. über das Programm und tauschen sich mit anderen Konsumenten in sozialen Medien aus. Dieses Phänomen wird als Social TV bezeichnet. Der Fokus liegt nicht ausschließlich auf dem Erkenntnisgewinn oder der Vertiefung des Wissens, sondern kann hauptsächlich auf dem Austausch und der Unterhaltung liegen. In der Schriftenreihe *Social TV in Deutschland* wurden 2015 viele Untersuchungen zum Social TV zusammengeführt [4].

Die Anreicherung von Videos mit Informationen ist eine Idee, die auch von Streaming-Diensten, wie *Amazon Prime* mit *X-Ray*, verfolgt wird. Hier werden die Videos analysiert, die Schauspieler im Bewegtbild erkannt und dem Anwender mit weiteren Informationen dargestellt, welche Schauspieler in den Szenen zu sehen sind.

3 Design und Realisierung

Basierend auf der Zielsetzung und den Anforderungen, die in Kapitel 2 analysiert wurden, wird in diesem Abschnitt das Design und die Realisierung der entwickelten Software besprochen. Es soll eine Software entwickelt werden, die Informationsergänzungen zu einer Fernsehdokumentation anhand der Untertitel automatisiert generiert und diese für die Anwender des Systems zur Verfügung stellt. Die Informationsergänzungen können parallel zum Konsum der Fernsehdokumentation betrachtet werden. Daher werden das Video und die Informationsergänzungen auf zwei unterschiedlichen Geräten dargestellt.

Im Kapitel 2 wurden verschiedene Datenquellen und Verfahren besprochen, die für das Design und die Realisierung der entwickelten Software eingesetzt werden können. Im Besonderen für die Generierung der Informationsergänzungen und die Darstellung sind die Verfahren und Datenquellen wichtig.

Für die Umsetzung der Anforderungen bietet sich ein modularer Aufbau der Software an, bei dem von verschiedenen Komponenten bestimmte Aufgaben übernommen werden.

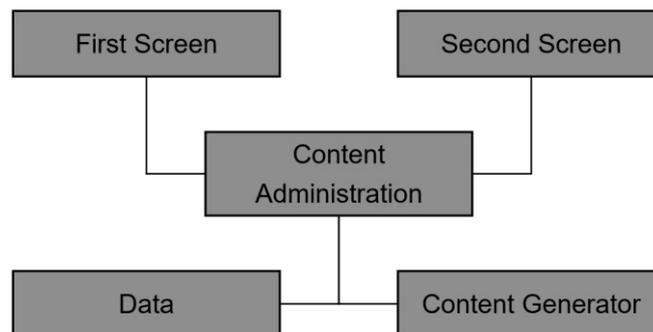


Abbildung 4: Vereinfachte Darstellung der benötigten Komponenten

Aus den Anforderungen können fünf vereinfachte Komponenten extrahiert werden, die im Späteren verfeinert werden. Eine Darstellung dieser Komponenten befindet sich in Abbildung 4. Zu den Komponenten gehören der *First- und Second Screen*, die die Schnittstelle für den Anwender darstellen und sich aus der Analyse der Anforderungen in Abschnitt 2.2.3 ergeben. Der *First Screen* ist für die Auswahl und Darstellung der Videos zuständig und der *Second Screen* für die synchronisierte Darstellung der zugehörigen Hintergrundinformationen. Mit einer weiteren Komponente, dem *Content Generator*, werden die Hintergrundinformationen zu den gegebenen Videos generiert. Die *First- und Second Screens* sowie deren Inhalte, wie die generierten Informationen, werden von einer zusätzlichen Komponente, hier als *Content Administration* bezeichnet, verwaltet. Für die Software müssen Daten zur Verfügung stehen,

die gespeichert und vorgehalten werden oder auf die über externe Schnittstellen zugegriffen werden kann. Dies wird durch die vereinfachte Komponente *Data* repräsentiert.

Eine detaillierte Beschreibung der vereinfachten Darstellung aus Abbildung 4 erfolgt in den nachfolgenden Abschnitten. In Abschnitt 3.1 werden die Komponenten anhand eines Schichtenmodells detaillierter dargestellt und beschrieben. Die Komponenten werden hierbei in drei Schichten unterteilt. In Abschnitt 3.2 werden die Komponenten und ihr schematischer Programmablauf des First- und Second Screens mit Hilfe eines Flussdiagramms aufgezeigt. Ein elementarer Aspekt der Software liegt in der Generierung der Hintergrundinformationen. Der Prozess der Generierung dieser Informationsergänzungen wird in Abschnitt 3.3 aufgeschlüsselt und folgt der Analyse der Anforderungen in Abschnitt 2.2.2. Es werden mehrere Verfahren zur Generierung der Informationsergänzungen umgesetzt, so dass diese in Abschnitt 4.1 untersucht und miteinander verglichen werden können. Nach den technischen Aspekten der Software wird in Abschnitt 3.4 die Darstellung und die Interaktion der Anwender mit der Software ausgeführt.

3.1 Schichtenmodell

Die vereinfachte Darstellung der Komponenten in Abbildung 4 wird in diesem Abschnitt erweitert und detaillierter dargestellt. Anhand des dargestellten Schichtenmodells wird ein Überblick über den Aufbau der Software und seiner Komponenten gegeben. Die Software wird modular aufgebaut, bestehend aus mehreren Komponenten. Jede Komponente erfüllt hierbei einen bestimmten Aufgabenbereich des Gesamtsystems.

Für die Komponenten kommen verschiedene Geräte zur Anwendung. Der Zugriff auf einige Daten erfolgt über APIs externer Anbieter. Die auf verschiedenen Geräten verteilten Komponenten und externen Daten bilden ein verteiltes System.

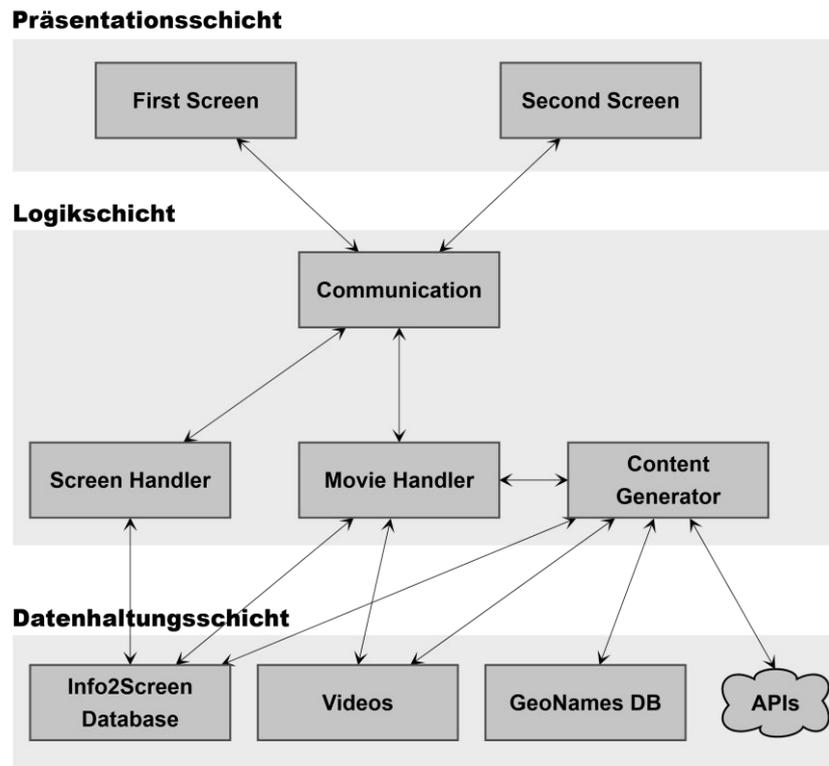


Abbildung 5: Schichtenmodell der Software [2]

Im Sinne der 3-Schichten-Architektur werden die Komponenten der Software, wie in Abbildung 5 gezeigt, entworfen und den drei Schichten zugeordnet. Die Schichten sind logisch voneinander getrennt. In der Präsentationsschicht sind Komponenten enthalten, die für die Darstellung der Inhalte des Systems sorgen und mit denen den Anwendern eine Möglichkeit zur Interaktion mit der Software geboten wird. Dies wird auch als Frontend bezeichnet. In der Logikschicht werden die Daten des Systems und die Eingaben der Anwender verarbeitet. In dieser Schicht wird demzufolge die Logik der Anwendung verortet. In der Datenhaltungsschicht werden die Daten der Anwendung verwaltet. Es wird auf externe Daten zugegriffen und es werden eigene Daten gespeichert und verwaltet.

Die vereinfachte Darstellung von Abbildung 4 wurde mit der Abbildung 5 erweitert und den drei Schichten logisch zugeteilt. Der First- und Second Screen bilden die sichtbare Schnittstelle für den Anwender des Systems und sind der Präsentationsschicht zugeordnet (siehe 3.1.1). Der *Content Generator* wurde der Logikschicht zugeordnet, ebenso wie die *Content Administration*. Die *Content Administration* wurde detaillierter betrachtet und in die Module *Communication*, *Screen Handler* und *Movie Handler* übersetzt (siehe 3.1.2). In der Datenhaltungsschicht befinden sich die Daten der Videos, eine eigene Datenbank, die GeoNames Datenbank und die externen Daten, die über APIs abgerufen werden (siehe 3.1.3).

Die 3-Schichten-Architektur bietet mit der klaren Unterteilung in Datenhaltungs-, Logik- und Präsentationsschicht einige Vorteile. Jede Schicht und ihre Funktionalitäten sind von den anderen Schichten klar getrennt und es werden Schnittstellen zwischen den Schichten definiert. Dies ermöglicht Änderungen einer Schicht, ohne dass Anpassungen in den anderen Schichten erfolgen müssen. Die Logikschicht ist leicht replizierbar, da in dieser üblicherweise keine Zustände gespeichert werden. Die Skalierbarkeit und Fehlertoleranz eines solchen Systems können mit Hilfe der Replikation verbessert werden. Bei Bedarf kann die Logikschicht in weitere Schichten aufgeteilt werden, so dass Teile des Systems unterschiedlich stark repliziert werden können (vgl. [68], S. 23f). Das Konzept folgt einer recht einfachen Struktur, so dass das Entwerfen eines solchen Systems leichtgängig erscheint. Eine nähere Betrachtung der einzelnen Schichten und der dazugehörigen Komponenten erfolgt in den folgenden Abschnitten.

3.1.1 Präsentationsschicht

In der Präsentationsschicht werden die Inhalte der Software dargestellt und Möglichkeiten für Eingaben seitens der Anwender zur Verfügung gestellt. In dieser Schicht befinden sich der First und Second Screen. Mit beiden Komponenten interagieren die Anwender der Software. Über den First Screen wird die Auswahl und das Abspielen der Videos gesteuert. Über den Second Screen wird ein gewünschter First Screen ausgewählt und die dazugehörigen Hintergrundinformationen konsumiert. Der First- und der Second Screen liegen auf einem Webserver und werden als Webseiten über einen Browser verwendet. Die Betrachtung und Interaktion mit diesen Komponenten erfolgen über einen Webbrowser.

Über den First Screen wird ein Video ausgewählt, das im System zur Verfügung steht. Anschließend wird das gewählte Video auf dem First Screen abgespielt. Das Video kann abgespielt, pausiert, fortgesetzt und gestoppt werden. Die Untertitel der Videos stehen zur Verfügung und können ebenfalls angezeigt werden. Der First Screen kommuniziert mit der *Communication*-Komponente der Logikschicht über REST (siehe 3.1.4). Die Metadaten der verfügbaren Videos werden von der *Communication*-Komponente an die Präsentationsschicht übertragen. Nach der Auswahl eines Videos wird die Auswahl an die *Communication*-Komponente übertragen. Während des Abspielens der Videos werden die Laufzeitänderungen ebenfalls an diese Komponente der Logikschicht übermittelt. Die Übermittlung der Laufzeit ermöglicht die Synchronisierung der Second Screens mit der Laufzeit der Videos des First Screens. Für die Betrachtung des First Screen bieten sich Anzeigegeräte, wie Fernseher oder große Bildschirme an, besonders wenn es mehrere Zuschauer gibt. Mehrere Second Screens können die Hintergrundinformationen, parallel zur Laufzeit des Videos im First Screen, erhalten. Der detaillierte Ablauf des First Screens ist in Abschnitt 3.2.1 schematisch aufgezeigt.

Mit dem Second Screen werden die Hintergrundinformationen eines Videos dargestellt. Die Darstellung der Informationen erfolgt synchronisiert zur Laufzeit des Videos eines

ausgewählten First Screens. Vor der Darstellung der Informationen wählt der Anwender einen verfügbaren First Screen aus. Anschließend werden die Hintergrundinformationen auf dem Second Screen dargestellt. Für die Betrachtung des Second Screens bieten sich Anzeigegeräte, wie Smartphones und Tablets, aber auch andere Bildschirme wie beispielsweise ein Notebook, an. Auch der Second Screen kommuniziert über REST mit der *Communication*-Komponente der Logikschicht. Er erhält eine Liste verfügbarer First Screens und kann einen dieser auswählen, um die dazugehörigen Informationen zu erhalten. Initial fragt der Second Screen die *Communication*-Komponente über REST nach bereits versendeten Hintergrundinformationen. So wird ermöglicht, dass ein Second Screen alle Hintergrundinformationen erhält, auch wenn das Video des First Screens bereits abgespielt wird. Für den kontinuierlichen Fluss der Hintergrundinformationen wird ein Message Broker-System (siehe 3.1.5) verwendet. Hierzu registriert sich der Second Screen bei dem Topic des jeweiligen First Screens. Nach der Registrierung erhält wiederum der Second Screen die Nachrichten, die nach dem Publish/Subscribe-Prinzip an das Topic gesendet werden.

Zu jedem First Screen kann es mehrere Second Screens geben. Technisch wird die resultierende Herausforderung der Nachrichtenverteilung an die Second Screens mit dem Message Broker-System gelöst. Nach der Laufzeitänderung im First Screen werden neue Nachrichten von der *Communication*-Komponente der Logikschicht an das dazugehörige Topic gesendet. Über das Topic werden die Nachrichten an alle gewünschten Second Screens verteilt. Die Option mehrere Second Screens zu haben, ermöglicht die gemeinschaftliche Anwendung des Systems mit mehreren Personen, die jeweils ihre eigenen Geräte als Second Screen verwenden können. So werden auch Szenarien beispielsweise im schulischen Umfeld oder im Bereich des E-Learnings ermöglicht. Der Einsatz eines Second Screens im Bereich des E-Learnings wurde beispielsweise von Rohloff et al. untersucht [67].

3.1.2 Logikschicht

Zwischen der Präsentations- und der Datenhaltungsschicht ist die Logikschicht platziert, die sich aus den *Communication*-, *Screen Handler*-, *Movie Handler*- und *Content Generator*-Komponenten zusammensetzt. In dieser Schicht befindet sich die Anwendungslogik der Software. Die Eingaben der Anwender und die Aufrufe durch die Komponenten der Präsentationsschicht, wie die Laufzeitänderungen der First Screens, werden von der Logikschicht verarbeitet. Die Verarbeitung erfolgt mit Hilfe der Datenhaltungsschicht. Auf diese wird durch mehrere Komponenten zugegriffen.

Die *Communication*-Komponente bildet die Kommunikations-Schnittstelle zwischen der Präsentationsschicht und der Logikschicht. Diese Komponente bietet eine eigene REST-Schnittstelle für die Kommunikation mit der Präsentationsschicht an. Der Großteil der Kommunikation erfolgt über REST. Die First Screen-Komponente übermittelt kontinuierlich die Laufzeitänderungen der abgespielten Videos. Die Laufzeitänderungen werden mit Hilfe der

Screen Handler-Komponente verarbeitet. Wenn es Hintergrundinformationen gibt, die verteilt werden sollen, werden diese über einen Message Broker an die Second Screens der Anwender gesendet. Zur Verteilung wird das Publish/Subscribe-Prinzip verwendet, bei der die Nachrichten an alle registrierten Consumer verteilt werden. Als Message Broker wird die Apache ActiveMQ [59] verwendet. Apache ActiveMQ ist ein von mehreren denkbaren Message Brokern, die das Publish/Subscribe-Prinzip ermöglichen. Es werden zahlreiche Protokolle unterstützt, die die Nutzung vieler Programmiersprachen ermöglichen.

Die Screen Handler-Komponente ist für die Verwaltung der First Screens der Anwender zuständig. Es werden die Auswahl eines Videos verarbeitet und die Laufzeitänderungen der Videos verwaltet. Bei einer Änderung der Laufzeit des Videos eines First Screens wird ermittelt, ob es Hintergrundinformationen gibt, die zwischen der alten und der neuen Laufzeit liegen. Gibt es in diesem Zeitraum Informationen, so werden diese über die *Communication*-Komponente an die Second Screens verteilt, die den First Screen ausgewählt haben. Second Screens, die einen First Screen auswählen, können die bereits versandten Hintergrundinformationen abrufen. Diese Anfrage wird u.a. vom Screen Handler verarbeitet. Für die Verwaltung und Verarbeitung wird die *Info2Screen Database*-Komponente der Datenhaltungsschicht aufgerufen.

Die Movie Handler-Komponente verwaltet die Videos, die der Anwendung zur Verfügung stehen. Videos, die neu eingebunden werden, werden mit den Metadaten in die *Info2Screen Database*-Komponente der Datenhaltungsschicht eingetragen. Die dazugehörigen Informationsergänzungen werden mit der *Content Generator*-Komponente aus den Untertiteln generiert. Auf Anfrage werden die in der *Database*-Komponente gespeicherten Metadaten über die *Communication*-Komponente an die Präsentationsschicht gesendet.

Zur Generierung der Hintergrundinformationen werden die wichtigsten Aspekte der Untertitel extrahiert und mit Geodaten und Daten aus einer Wissensbasis verknüpft. Jeder generierten Hintergrundinformation wird ein Zeitstempel zugeordnet. Dieser zeigt an, zu welchem Zeitpunkt des Videos diese Information erscheint. Die Hintergrundinformationen werden in der *Info2Screen Database*-Komponente der Datenhaltungsschicht hinterlegt. Die detaillierte Funktionsweise der *Content Generator* Komponente wird in Abschnitt 3.3 aufgezeigt.

3.1.3 Datenhaltungsschicht

In der Datenhaltungsschicht sind die Komponenten der Anwendung vertreten, die die Daten speichern, verwalten und laden. Die externen Komponenten, auf deren Daten über APIs zugegriffen wird, wurden ebenfalls der Datenhaltungsschicht zugeordnet. Hierin befindet sich eine Datenbank-Komponente, die nachfolgend näher beschrieben wird und in der die Daten gespeichert und verwaltet werden, die durch die Anwendung generiert werden.

Neben der Datenbank sind die Videos ein weiterer wichtiger Baustein der Anwendung. Es werden die Untertitel, Metadaten und Videodaten verwendet. Auf die Videos der Anwendung wird näher eingegangen.

Für die Generierung der Hintergrundinformationen werden die Daten einer Wissensbasis und von Kartendiensten benötigt. Mit einer Wissensbasis werden die Untertitel der Videos mit tiefgehenden Informationen angereichert. Die Wissensbasen, die zur Verwendung kommen, werden im Weiteren erläutert. Mit Hilfe von Kartendiensten können die Untertitel mit Geodaten angereichert werden und in der Präsentationsschicht können die in den Untertiteln benannten Orte mit Kartenausschnitten dargestellt werden. Die verwendeten Kartendienste werden ebenfalls in diesem Abschnitt beschrieben.

Neben den Daten oder Datensätzen, die von der Anwendung erzeugt oder direkt eingebunden werden können, wird auf zahlreiche Daten über externe APIs zugegriffen. Der Kartendienst OpenStreetMap wird über die Nominatim²⁶-API verwendet. Für den Zugriff auf die Daten der Wissensbasen werden die *MediaWiki*²⁷-API für *Wikipedia* und die *DBPedia Spotlight*²⁸-API für DBPedia genutzt. Der Zugriff auf die externen APIs erfolgt über REST (siehe 3.1.4).

Datenbank

In der *Info2Screen Database*-Komponente werden die erzeugten Daten der Anwendung und die Metadaten der Videos gespeichert. Zu den Metadaten der Videos gehören u.a. der Titel und Beschreibungstexte der Videos.

Die Hintergrundinformationen werden ebenfalls in der Datenbank gespeichert und einem spezifischen Video zugeordnet. Zu jedem Datenbankeintrag der Hintergrundinformationen gehört die Nachricht im JSON-Format, ein Label und die Laufzeit, an der die Information im Video erscheint. Mit der Zuordnung der Nachrichten zu einem Label kann eine zusätzliche Filterung der Nachrichten erfolgen. Wenn die Hintergrundinformationen einem

²⁶ <https://nominatim.org/> (Aufgerufen am 11. März 2021)

²⁷ https://www.mediawiki.org/wiki/API:Main_page (Aufgerufen am 11. März 2021)

²⁸ <https://www.dbpedia-spotlight.org/> (Aufgerufen am 11. März 2021)

geographischen Namen entsprechen, sind in dem JSON-Objekt die Geodaten enthalten. Wenn der Datenbankeintrag mit einer Wissensbasis verknüpft werden konnte, dann befindet sich im JSON-Objekt u.a. ein Text mit den gesamten Informationen oder einem einleitenden bzw. zusammenfassenden Inhalt.

Für die Verwaltung der First- und Second Screens werden Daten zum First Screen in der Datenbank gespeichert. Zum einen ist es wichtig zu speichern, welches Video vom First Screen zur Betrachtung ausgewählt wurde und zum anderen spielt die Laufzeit der Videos für jeden First Screen eine weitere wichtige Rolle.

Videos

Die Videos sind ein elementarer Baustein der Software. Das Video wird auf dem First Screen dargestellt und die Untertitel des Videos werden für die Generierung der Zusatzinformationen weiterverarbeitet. Für eine sinnvolle Verwendung der Videos werden im Folgenden einige Rahmenbedingungen erläutert, die für die Videos gesetzt werden, die mit der Software verwendet werden sollen.

Die Videos müssen über Untertitel verfügen, die in maschinenlesbarer Form vorliegen. Es ist möglich, dass aus der Audiospur von Videos die Untertitel generiert werden, was aber nicht Gegenstand dieser Ausarbeitung sein soll [1].

Im Rahmen dieser Arbeit werden Untertitel in der deutschen Sprache verwendet. Es gibt die Möglichkeit, weitere Sprachen zu unterstützen. Der Fokus liegt hier jedoch auf der Verarbeitung deutschsprachiger Untertitel. Für maschinenlesbare Untertitel stehen zahlreiche Formate zur Verfügung. In dieser Ausarbeitung werden die Formate XML und WebVTT verwendet. Das WebVTT-Format bildet für Untertitel den Standard des World Wide Web Consortium (W3C) [69] und basiert auf dem weit verbreiteten SubRip-Format²⁹. In allen Formaten wird einer Zeitspanne ein Textsegment der Untertitel zugeordnet. Diese zeitliche Zuordnung wird für die parallele Anzeige der Untertitel mit dem Video benötigt. Außerdem wird sie genutzt, um die Hintergrundinformationen auf dem Second Screen parallel zum Video darzustellen.

Das System soll den Nutzer beim Betrachten der Videos mit Hintergrundinformationen unterstützen. Diese Informationen sollen vom Anwender wahrgenommen und verarbeitet werden können. Die Auffassung der Zusatzinformationen und das Betrachten der Videos erfolgt parallel. Damit der Nutzer beides verarbeiten kann, dürfen nicht zu viele Informationen in zu kurzer Zeit auf dem Second Screen erscheinen. Das Weglassen von Informationen ist

²⁹ <https://sourceforge.net/projects/subrip/> (aufgerufen am 28. Januar 2021)

keine geeignete Lösung, wenn nicht bekannt ist, über welches Vorwissen der Anwender verfügt und inwieweit die Informationen mit diesem gefiltert werden können. Bei Videoformaten, wie z.B. Nachrichtensendungen, ist mit einer besonders hohen Informationsdichte zu rechnen und könnte dementsprechend überfordernd auf den Anwender wirken.

Um einen Mehrwert erzeugen zu können, müssen die Videos über Informationen verfügen, bei denen es sinnvoll sein kann, weitere Informationen zu erhalten. Im Rahmen dieser Arbeit sollen die Videos daher Wissen vermitteln und aus den Kategorien der Reportage und Dokumentation stammen. Beide Kategorien gehören nach Christian Hißnauer zum Fernsehdokumentarismus. Dokumentationen zeichnen sich durch „inhaltliche und formale Offenheit“ aus. Reportagen dagegen fokussieren den Einzelfall und gelten als „subjektive Darstellungsform“, bei denen Beobachtungen, Beschreibungen und situative Interviews dominieren [70].

Der Themenbereich der Reportagen und Dokumentationen wird nicht eingeschränkt. Es stehen umfangreiche Wissensbasen zur Verfügung, die ein großes Spektrum an Themen umfassen. Die freie Themenwahl innerhalb der Dokumentation bzw. der Reportage ermöglicht eine möglichst umfangreiche Unterstützung der Nutzer.

Die Länge der Videos bzw. Fernsehdokumentationen soll etwa 20 bis 45 Minuten betragen. Viele Fernsehdokumentationen befinden sich in diesem zeitlichen Rahmen. Dieser Rahmen wird festgelegt, damit die Anwender ausreichend Zeit haben, das System zu verwenden und einen Nutzen aus den Ergänzungen zu ziehen. Ein längerer Zeitraum könnte die Konzentration auf die Fernsehdokumentation reduzieren. Mit der Länge eines Videos nimmt die parallele Nutzung anderer Medien zu. Die Studie „Social TV in Deutschland“ aus dem Jahr 2015 hat herausgefunden, dass die Twitter-Aktivität von Zuschauern des Fernsehprogramms abhängig von der Länge des Programms zunimmt (vgl. [4], S. 60). Kurze Videos bieten wenige Gelegenheiten oder wenig Ruhe, in der die Anwender die Informationsergänzungen betrachten und weiteren Nutzen aus ihnen ziehen können.

Die Länge, aber auch die Art der Videos ist relevant. Nachrichten sind in der Regel kurz und die Informationsdichte ist besonders hoch. Van Cauwenberge et al. [64] untersuchten den Effekt eines Second Screens auf das Lernvermögen beim Nachrichtenkonsum. Sie kamen zu dem Erkenntnis, dass sich die Probanden bei der parallelen Nutzung eines Second Screens und dem Nachrichtenkonsum weniger Fakten merken konnten. Werden die Informationen auf dem Second Screen aufbereitet bereitgestellt, so ist dies besser, als wenn die Probanden den Second Screen aktiv nutzen müssen, allerdings können diese sich weiterhin weniger merken, als wenn sie nur die Nachrichten konsumieren. Dies liegt daran, dass die hohe kognitive Belastung beim Nachrichtenkonsum durch die Second Screen-Nutzung gesteigert wurde [65]. Andere Formate, bei denen die Informationsdichte geringer ist, also mehr Zeit zum Verarbeiten der Informationen bleibt, beanspruchen die kognitive Leistung geringer. Formate, wie

beispielsweise Polit-Talkshows eignen sich weniger, da es bei diesen chaotisch zugehen kann und die benannten Themen im Rahmen eines Fact-Checkings gegengeprüft werden sollten.

Wissensbasen

Die Untertitel der Videos sollen mit Hintergrundinformationen angereichert werden. Die Generierung der passenden Hintergrundinformationen erfolgt mit dem *Named Entity Linking* (NEL, siehe 2.6). Hierbei werden die *Named Entities* der Untertitel mit den passenden Einträgen aus einer Wissensbasis verknüpft. In dieser Ausarbeitung wurden mit Wikipedia und DBPedia zwei Wissensbasen für diese Aufgabe ausgewählt (siehe 2.4.2).

Wikipedia ist ein Projekt mit vielen Sprachversionen und einem umfangreichen Datensatz, der von der Community ständig erweitert wird. Da die Videos nicht in ihrem Themenbereich eingegrenzt werden, wird eine umfangreiche und themenübergreifende Wissensbasis wie Wikipedia benötigt. Die Daten von Wikipedia sind leicht zu erhalten und unter der CC-BY-SA – Lizenz unter Namensnennung frei verfügbar. Der Zugriff auf die Daten von Wikipedia erfolgt über die MediaWiki-APIs³⁰. Für die deutsche Sprachversion von Wikipedia wird die entsprechende API verwendet.³¹

Viele der Vorzüge von Wikipedia sind auch in dem DBPedia-Projekt zu finden. DBPedia ist ebenfalls themenübergreifend und enthält umfangreiche Daten. Zurzeit werden 20 Sprachversionen unterstützt, darunter auch die deutsche Sprache. Wie Wikipedia ist auch DBPedia mit der CC-BY-SA – Lizenz unter Namensnennung gut zu verwenden. Es gibt verschiedene Zugriffsmöglichkeiten auf die Daten von DBPedia. Die Daten stehen im RDF-Format, über eine SPARQL-Schnittstelle und über HTTP-URIs zur Verfügung.

Das Projekt *DBPedia Spotlight* stellt eine schnelle und einfache Möglichkeit dar, das *Named Entity Linking* mit den Daten von DBPedia zu anzuwenden. Es werden Texte an das *DBPedia Spotlight*-System übertragen und von diesem automatisiert annotiert. Die *Named Entities* werden extrahiert und mit Einträgen aus DBPedia verknüpft. *DBPedia Spotlight* kann auf einem eigenen Server betrieben oder über eine API³² der Entwickler genutzt werden.

Die aufgeführten Projekte wurden aufgrund ihres Umfangs und ihrer Verfügbarkeit ausgewählt. Das Austauschen dieser und die Auswahl weiterer Wissensbasen ist möglich.

³⁰ https://www.mediawiki.org/wiki/API:Main_page#Endpoint (Aufgerufen am 31. Januar 2021)

³¹ <https://de.wikipedia.org/w/api.php> (Aufgerufen am 31. Januar 2021)

³² <https://api.dbpedia-spotlight.org/de/annotate> - DBPedia Spotlight für die deutsche Sprachversion (Aufgerufen am 31. Januar 2021)

Kartendienste

Die Untertitel der Videos sollen mit Hintergrundinformationen angereichert werden. In den Untertiteln können Orte bzw. geographische Namen enthalten sein. Named Entities, die geographischen Namen entsprechen, können auf Karten dargestellt werden. Mit Hilfe des Geocoding (siehe 2.7) kann den geographischen Namen eine geographische Koordinate zugeordnet werden. Darüber hinaus kann, abhängig von dem geographischen Namen und dem verwendeten System, ein Kartenausschnitt ermittelt werden. Mit den Kartenausschnitten werden große geographische Namen, wie beispielsweise Länder, Kontinente oder Naturreservate, auf einer Karte in passender Größe dargestellt. Mit der Vorverarbeitung des Named Entity Recognition (siehe 2.5.3) können Kandidaten für geographische Namen ermittelt werden. Für das Geocoding und die Darstellung von Karten sind Geodaten erforderlich.

Mit der GeoNames Datenbank und OpenStreetMap wurden zwei Projekte gewählt, die zahlreiche Geodaten enthalten (siehe 2.4.3). Beide Projekte umfassen Geodaten von Geographischen Namen auf der ganzen Erde und eignen sich für das Geocoding.

Mit der GeoNames-Datenbank wurde ein Projekt ausgewählt, das mit einem umfangreichen Datensatz von über 12 Millionen Geo-Objekten für das Geocoding in mehreren Sprachen geeignet ist. Neben mehrsprachigen Bezeichnungen sind auch alternative Bezeichnungen von geographischen Namen in der Datenbank enthalten. Neben den Bezeichnungen und den GPS-Koordinaten von geographischen Namen sind weitere Informationen, wie beispielsweise die Einwohnerzahl, in der Datenbank zu finden. Die Datenbank ist gut zu verwenden, denn sie ist leicht zu erhalten³³ und wird unter der CC-BY – Lizenz veröffentlicht.

Das Projekt OpenStreetMap (OSM) ist eine sehr umfangreiche Datenbank mit Geodaten, die ständig wächst. Die Daten des Projekts können für das Geocoding und weitere Anwendungen hinzugezogen werden. OSM wird unter der „Open Database Licence (ODbL) 1.0“ – Lizenz veröffentlicht und kann unter u.a. mit Namensnennung verwendet werden. Der Datensatz ist ebenfalls leicht zu erhalten und kann für eigene Datenbank-Instanzen genutzt werden. Aufgrund des Datenumfangs von OSM werden im Rahmen dieser Arbeit online auf die APIs von OSM zugegriffen. Für die Kartendarstellung der geographischen Namen auf dem Second Screen wird das Kartenmaterial von OSM verwendet.

³³ <https://download.geonames.org/export/dump/> (Aufgerufen am 31. Januar 2021)

Mit dem Nominatim-Projekt [72] gibt es ein System, das das Geocoding mit den Daten von OpenStreetMap umgesetzt hat. Nominatim wird unter der GPLv2³⁴-Lizenz veröffentlicht und ist über eine API online nutzbar³⁵. Nominatim kann auf eigenen Servern betrieben werden.

Die gewählten Kartendienste eignen sich für die Verwendung zum Geocoding. Es können weitere Kartendienste und Datenbanken mit Geodaten verwendet und die ausgewählten mit diesen ergänzt oder ausgetauscht werden.

3.1.4 Representational State Transfer (REST)

In seiner Dissertation beschrieb Roy Fielding ein Softwarearchitektur-Paradigma für Verteilte Systeme [73]. Mit dem Representational State Transfer (REST) entwickelte Fielding ein Architektur-Paradigma, das auf dem *Hypertext Transfer Protocol* (HTTP) [74] basiert und mit dem Web-Ressourcen über die HTTP-Methoden GET, POST, PUT und DELETE erstellt, angefordert und manipuliert werden kann. HTTP und REST werden üblicherweise gemeinsam verwendet, sind aber nicht zwangsläufig miteinander gekoppelt. REST ermöglicht die Maschine-zu-Maschine Kommunikation in Verteilten Systemen und stellt eine gute Alternative zu anderen Konzepten dar. REST ist ein weit verbreitetes Paradigma für die Maschine-zu-Maschine Kommunikation, für die ein Großteil der benötigten Infrastruktur durch das World Wide Web bereits vorhanden ist.

Viele aktuelle Anwendungen arbeiten als Verteilte Systeme. Beispielsweise benannt seien Anwendungen, die die Microservice-Architektur [76] umsetzen. Hier spielt die Kommunikation über klar definierte Schnittstellen zwischen den Komponenten der Software eine besondere Rolle. Diese Kommunikation kann über REST-Schnittstellen erfolgen. Viele weitere Anwendungen sind mit dem Internet verbunden und greifen auf Schnittstellen wie REST-APIs zurück. Die Kommunikation mit den REST-APIs erfolgt über HTTP und ermöglicht so eine leichte und dennoch mächtige Umsetzung der Schnittstellen.

Auch für die Software, die im Rahmen dieser Ausarbeitung entwickelt wird, werden REST-Schnittstellen verwendet und geschaffen. Es werden externe REST-APIs von Nominatim [72] für OpenStreetMap [29], DBPedia Spotlight [17], DBPedia und Wikipedia (siehe 2.4.2) verwendet. Die Kommunikation zwischen den Komponenten der Präsentationsschicht und der Kommunikations-Komponente der Logikschicht erfolgt über eine eigene REST-Schnittstelle der Kommunikations-Komponente. REST wird in dieser Arbeit für die Kommunikation verwendet, da REST eine schnelle, schlanke und flexible Lösung bietet. Neben der internen Verwendung erfolgt die Kommunikation mit den externen APIs ebenfalls über REST.

³⁴ <https://www.gnu.org/licenses/old-licenses/gpl-2.0.html> (Aufgerufen am 31. Januar 2021)

³⁵ <https://nominatim.org/release-docs/develop/api/Overview/> (Aufgerufen am 31. Januar 2021)

Für REST werden sechs Prinzipien definiert, die eine Software umsetzen muss, um das REST-Paradigma zu erfüllen. Hierbei wird nicht festgelegt, wie diese Prinzipien implementiert werden sollen. Eine Beschreibung dieser Design-Prinzipien von REST erfolgt nachstehend.

Prinzipien von REST

In der Dissertation von Fielding [73] werden Prinzipien für REST definiert, die erfüllt werden müssen. In der Veröffentlichung von Erenkrantz et al. [77] wurden sechs Kern-Prinzipien von REST formuliert, die nachfolgend beschrieben werden. Die Beschreibung von Erenkrantz et al. punktet mit ihrer Prägnanz. Es ist umstritten, welche der beiden Formulierungen die beste sei [75].

Nach Erenkrantz et al. [77] hat REST das Ziel, die Latenz im Netzwerk zu reduzieren und zeitgleich die Implementierung von unabhängigen und skalierbaren Komponenten zu ermöglichen. REST fokussiert sich auf die Kommunikation zwischen den Komponenten und versucht, das ordnungsgemäße Verhalten der Komponenten zu regeln. Hierfür wurden sechs Design-Prinzipien aufgestellt.

1. Ressourcen sind die Hauptabstraktionen von Informationen. Jede Information, die benannt werden kann, kann als Ressource verwendet werden. Ressourcen werden mit Hilfe eines *Uniform Resource Locator* (URL) adressiert.
2. Die Repräsentation einer Ressource ist eine Sequenz von Bytes und den Metadaten der Repräsentation, die diese Bytes beschreiben. Mit REST werden die Ressourcen über die Repräsentationen und nicht die konkrete Ressource aufgerufen und manipuliert. Ressourcen können durch mehrere Repräsentationen ausgegeben werden. Diese Repräsentationen einer Ressource können die Daten in verschiedenen Sprachen oder Formaten, wie JSON und XML, enthalten. Die Form der Repräsentation kann zwischen den REST-Komponenten ausgehandelt werden.
3. Alle Interaktionen mit der REST-Schnittstelle enthalten alle notwendigen Informationen, die für die Verarbeitung der Interaktionen benötigt werden. Die Interaktionen sind unabhängig vom Kontext. Beispielsweise werden keine Informationen vorangegangener Interaktionen mit der Schnittstelle gespeichert, die für die Verarbeitung von Interaktionen benötigt werden.
4. Für REST-Schnittstellen wird eine kleine Menge an gut definierten Methoden angeboten. Mit diesen Methoden wird eine Repräsentation der Ressourcen erstellt, mit der der aktuelle Status oder der beabsichtigte Status nach der Manipulation der Ressource erfasst und anschließend zwischen den REST-Komponenten übertragen

wird. Diese Methoden müssen global von jeder Architekturinstanz von REST unterstützt werden. So müssen alle Ressourcen einer Architekturinstanz, die über HTTP bereitgestellt werden, die gleichen HTTP-Methoden (GET, POST, PUT und DELETE) unterstützen.

5. Es wird empfohlen, für das Caching idempotente Operationen und die Metadaten von Repräsentationen zu verwenden, um die Latenz der Anfragen zu reduzieren. Viele Repräsentationen der Ressourcen können wiederverwendet werden und müssen nicht für jede Anfrage neu berechnet werden. Anhand der Metadaten der Repräsentationen kann entschieden werden, ob eine Neuberechnung einer Repräsentation sinnvoll ist oder diese erst kürzlich erstellt worden ist. Operationen bzw. Methoden, die mehrfach ohne Änderung der Rückgabe ausgeführt werden können, werden als idempotent bezeichnet. Hierzu gehören beispielsweise lesende Operationen. Aufgrund ihrer Eigenschaft können die Repräsentationen dieser Methoden wiederverwendet werden.
6. Es wird empfohlen, mehrere Schichten zu verwenden. Zusätzliche Schichten können die Skalierung des Systems verbessern, wenn Proxies, Firewalls, Gateways oder Load Balancer hinzugefügt werden. Filternde oder weiterleitende Schichten können die Metadaten und die Repräsentationen der Anfragen und Antworten nutzen, um die Anfragen und Antworten zu erweitern, zu reduzieren oder zu modifizieren. Dieses geschieht transparent und ohne Veränderung der Schnittstellen.

3.1.5 Message Broker

In Verteilten Systemen müssen die Komponenten der Anwendung miteinander kommunizieren. Eine Art, die Kommunikation zwischen den Modulen zu ermöglichen, kann über die Verwendung von Message Brokern erfolgen. Eine weitere Möglichkeit stellt das in Abschnitt 3.1.4 bereits besprochene REST dar.



Abbildung 6: Schematische Darstellung eines Message Brokers

Wie in Abbildung 6 schematisch dargestellt, nimmt ein *Message Broker* die Nachrichten von Anwendungen, hier als *Producer* bezeichnet, entgegen, verwaltet diese Nachrichten und leitet diese an die gewünschten Anwendungen, hier als *Consumer* bezeichnet, weiter. Die Nachrichten enthalten die Daten, die vom *Producer* zum *Consumer* gesendet werden sollen.

Die Art und Weise der Verwaltung und Versendung der Nachrichten kann sich unterscheiden. So können Nachrichten an bestimmte *Consumer* gerichtet sein. Zwei verbreitete Ansätze der Nachrichtenverteilung an mehrere *Consumer* sind das Producer/Consumer-Prinzip oder das Publish/Subscribe-Prinzip. Beide Ansätze werden im Späteren erläutert. Ein Message Broker kann diese Ansätze parallel und mehrfach verwirklichen.

In dem Message Broker werden die Nachrichten in Topics bzw. in Message Queues zwischengespeichert. Abstrakt beschrieben erzeugen die *Producer* Nachrichten, die entsprechend ihrer Reihenfolge oder Priorisierung in einer Message Queue zwischengespeichert werden. Wenn der bzw. die *Consumer* verfügbar sind, werden die Nachrichten in der Message Queue sequenziell vom Message Broker verarbeitet und an den bzw. die entsprechenden *Consumer* weitergeleitet.

Der Einsatz von Message Brokern ermöglicht die Kommunikation zwischen Anwendungen oder Komponenten von Anwendungen, ohne dass eine direkte Verbindung zwischen diesen bestehen muss. Das Routing bzw. die Verteilung der Nachrichten übernimmt der Message Broker. Die Kommunikation über die Message Broker erfolgt üblicherweise asynchron. Das bedeutet, dass der Nachrichtenaustausch nicht blockierend ist bzw. auf die Antwort des Empfängers gewartet wird. Der Empfänger kann die Nachricht zu einem späteren Zeitpunkt erhalten.

Es gibt zahlreiche Open Source-Lösungen für Message Broker. Dazu gehören exemplarisch Apache ActiveMQ [78], RabbitMQ [79] und Apache Kafka [80]. Abhängig von dem Anwendungsfall können sich verschiedene Lösungen hervortun. Sie unterscheiden sich in der Performance [81] und auch weiteren Eigenschaften [82]. Für Message Queues und Message Broker-Systeme im Allgemeinen gibt es viele Anwendungsmöglichkeiten, beispielsweise um das Load Balancing und die Skalierung von Anwendungen zu verbessern.

Producer/Consumer-Prinzip

Eine der erwähnten Arten der Nachrichtenverteilung ist das *Producer/Consumer-Prinzip*. Hierbei werden die Nachrichten von einem sogenannten *Producer* generiert und über eine *Message Queue* an die *Consumer* gesendet. Nachdem die Nachrichten in der *Message Queue* verarbeitet worden sind und vorgehalten werden, können diese vom *Consumer* empfangen werden.

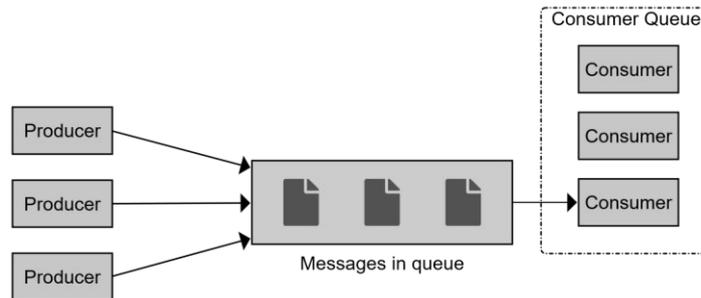


Abbildung 7: Schematische Darstellung des Producer/Consumer-Prinzips (vgl. [2])

In Abbildung 7 ist das *Producer/Consumer-Prinzip* schematisch abgebildet. Eine beliebige Anzahl an *Producern* erzeugt Nachrichten, die an die *Message Queue* versendet werden. Diese Nachrichten werden empfangen und entsprechend ihrer Reihenfolge in der *Message Queue* gespeichert. Einige *Message Broker* unterstützen die Priorisierung von Nachrichten, wodurch die Reihenfolge dieser beeinflusst werden kann. Die Nachrichten werden in der *Message Queue* gespeichert, bis ein *Consumer* zur Verfügung steht und die gespeicherte Nachricht abrufen. Entsprechend der Reihenfolge, also zumeist nach dem „First In – First Out“-Prinzip, werden die Nachrichten an zur Verfügung stehende *Consumer* gesendet und anschließend aus der *Message Queue* entfernt. Für jede *Message Queue* kann es mehrere *Consumer* geben. Jede Nachricht wird jedoch nur an einen *Consumer* weitergeleitet.

Werden Nachrichten über einen definierten Zeitraum nicht abgerufen oder zu viele Nachrichten in der *Message Queue* gespeichert, können Nachrichten verworfen werden.

Beim *Producer/Consumer-Prinzip* wird mit *Message Queues* gearbeitet. Es kann mehrere *Producer* und mehrere *Consumer* geben. Jede generierte Nachricht wird hierbei von maximal einem *Consumer* empfangen. Dieses Prinzip kann beispielsweise für das *Load Balancing* zum Einsatz kommen.

Publish/Subscribe-Prinzip

Wie beim Producer/Consumer-Prinzip gibt es auch beim Publish/Subscribe-Prinzip sendende und empfangende Anwendungen. Es gibt einen oder mehrere Publisher, die Nachrichten erzeugen und an ein sogenanntes Topic senden. In diesem werden die Nachrichten zwischengespeichert und von einem oder mehreren Subscribern abgerufen.

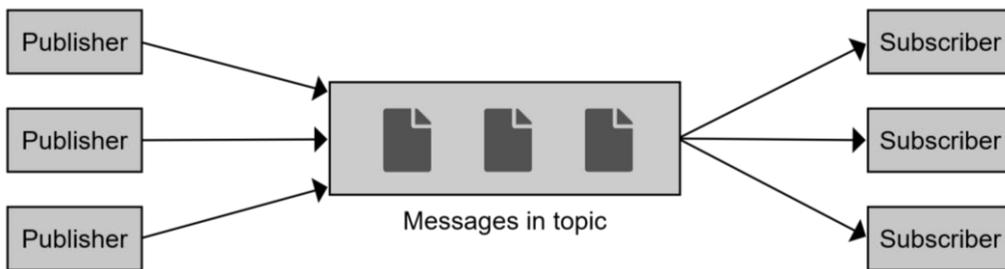


Abbildung 8: Schematische Darstellung des Publish/Subscribe-Prinzips (vgl. [2])

In Abbildung 8 ist das Publish/Subscribe-Prinzip schematisch abgebildet. Hier werden die Anwendungen, die Nachrichten erzeugen, als Publisher bezeichnet. Es kann einen oder mehrere Publisher geben, die die Nachrichten an ein Topic senden, in denen diese verarbeitet und zwischengespeichert werden. Die Topics leiten Kopien der empfangenen Nachrichten an alle Subscriber des Topics weiter. Anwendungen, die die Nachrichten eines Topics empfangen wollen, abonnieren das Topic und werden als Subscriber bezeichnet. Ab dem Zeitpunkt der Anmeldung erhalten die Subscriber alle Nachrichten, die dem Topic zugesendet werden. Einige Systeme speichern die Nachrichten für einen definierten Zeitraum, so dass Subscriber auch ältere Nachrichten erhalten können. Nachrichten, die sich im Topic befinden, werden an alle Subscriber gesendet. Abhängig vom System können Nachrichten von Subscribern zu verschiedenen Zeitpunkten abgerufen werden. Hierbei wird für jeden Subscriber gespeichert, welche Nachrichten bereits empfangen wurden bzw. welche Nachrichten noch nicht an diese gesendet wurden. Die Verteilung der Nachrichten erfolgt für alle Subscriber nach dem „First In – First Out“-Prinzip oder nach der prioritätsbedingten Reihenfolge.

Nachrichten werden aus dem Topic erst entfernt, wenn diese von allen Subscribern erhalten worden sind. Zusätzlich können diese nach einem definierten Zeitraum, oder wenn sich zu viele Nachrichten im Topic befinden, verworfen werden.

Jede Nachricht eines Topics kann von allen Subscribern empfangen werden, die Publisher und Subscriber stehen in einer $n:m$ -Beziehung.

3.2 Schematischer Programmablauf

Mit Hilfe von Flussdiagrammen werden in diesem Abschnitt die schematischen Programmabläufe und die Funktionsweisen der First Screen- und Second Screen-Komponenten aufgezeigt. Wie in Abschnitt 3.1 dargestellt, bieten beide Komponenten die Möglichkeit der Interaktion für die Anwender der Software. Außerdem werden auf dem First Screen die Videos und auf dem Second Screen die Hintergrundinformationen dargestellt.

3.2.1 First Screen-Komponente

Der schematische Programmablauf der *First Screen*-Komponente ist in Abbildung 9 dargestellt. Die First Screen-Komponente ist für die Auswahl und das Anzeigen der ausgewählten Videos zuständig. Es wird mit der *Communication*-Komponente der Logikschicht über REST kommuniziert. Der Ablauf der Komponente wird im Folgenden näher aufgeführt.

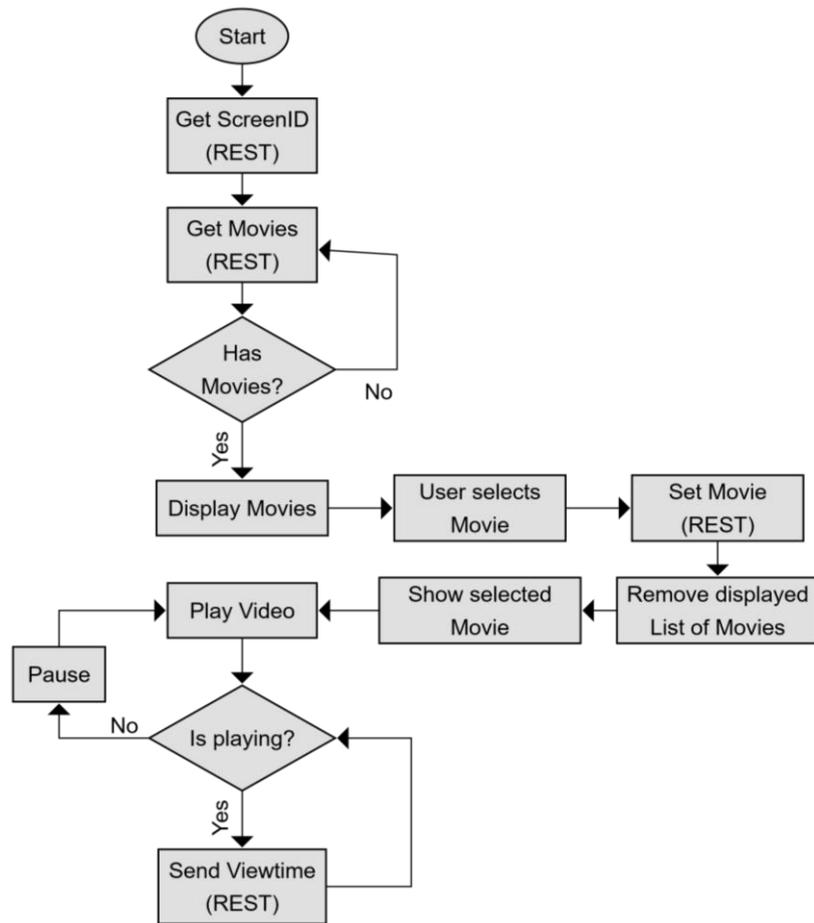


Abbildung 9: Ablauf des First Screen als Flussdiagramm [2]

Mit dem Aufruf der Webseite des First Screens über einen Webbrowser wird der schematische Programmablauf in Abbildung 9 gestartet. Es erfolgen zwei REST-Anfragen an die *Communication*-Komponente der Logikschicht. Mit der ersten Anfrage wird der neuen First Screen-Instanz eine ScreenID zugewiesen, um diese im späteren Ablauf eindeutig identifizieren zu können. Mit der zweiten Anfrage wird eine Liste der verfügbaren Videos angefordert. Wenn keine Videos im System verfügbar sind, so muss die Anfrage zu einem späteren Zeitpunkt wiederholt werden. Wenn es Videos gibt, die im System verfügbar sind, so sollen diese auf der Webseite dargestellt werden. Der Anwender kann eines dieser Videos auswählen. Die Auswahl des Videos wird über REST an die Communication-Komponente in der Logikschicht kommuniziert.

Anschließend wird die Anzeige der verfügbaren Videos entfernt und durch die Anzeige des ausgewählten Videos ersetzt. Die neue Ansicht ermöglicht das Abspielen des Videos über einen eingebundenen Videoplayer. Der Anwender startet das Video zu einem beliebigen Zeitpunkt. Wenn das Video gestartet worden ist, kann es pausiert werden. Ändert sich die

Laufzeit des Videos, so wird die aktualisierte Laufzeit an die Communication-Komponente über REST weitergegeben.

3.2.2 Second Screen-Komponente

Der schematische Programmablauf der Second Screen-Komponente ist in Abb. 10 dargestellt. Die Second Screen-Komponente ist für die Auswahl des First Screens und die Anzeige der Hintergrundinformationen zuständig. Es wird mit der Communication-Komponente der Logikschicht über REST kommuniziert. Des Weiteren werden mit Hilfe eines Message Brokers die Hintergrundinformationen der Videos des First Screens kontinuierlich empfangen. Der Ablauf der Komponente wird im Folgenden näher aufgeführt.

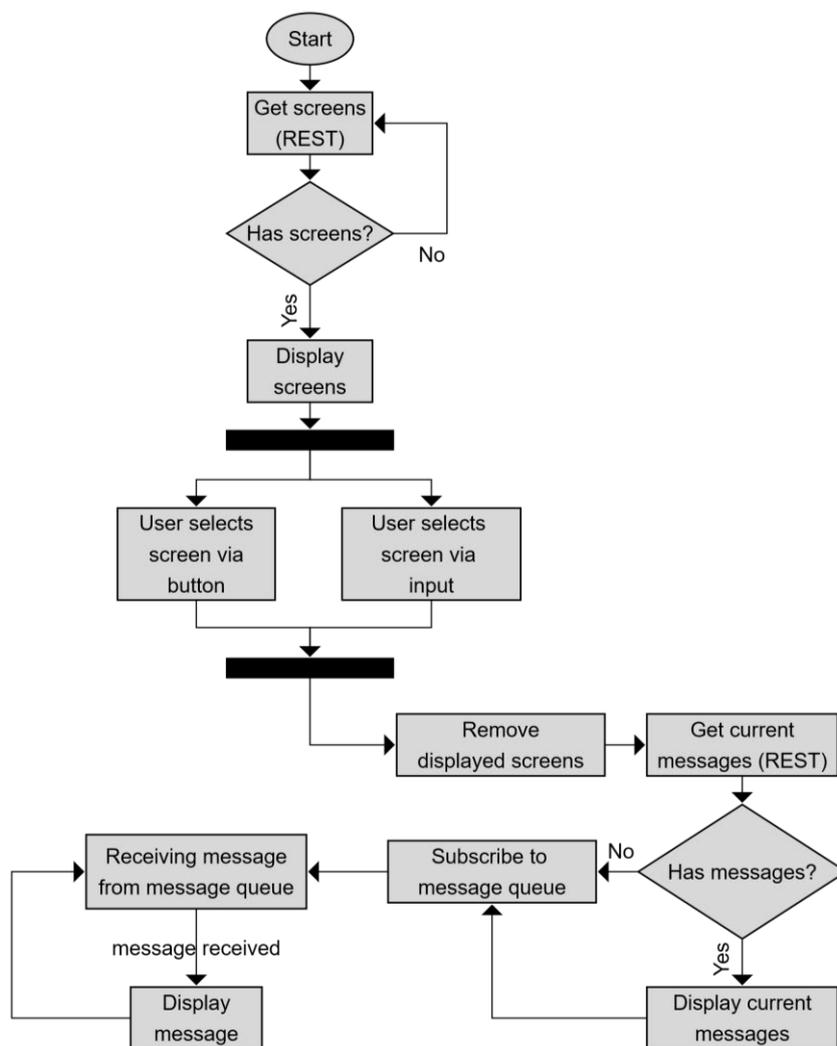


Abbildung 10: Ablauf des Second Screen als Flussdiagramm [2]

Mit dem Aufruf der Webseite der Second Screen-Komponente über einen Webbrowser wird der schematische Programmablauf in Abbildung 10 gestartet. Zu Beginn wird eine REST-Anfrage an die *Communication*-Komponente der Logikschicht gesendet. Hiermit wird eine Liste der verfügbaren *First Screen*-Instanzen angefragt. Wenn es keine verfügbaren First Screens gibt, wird die Anfrage zu einem späteren Zeitpunkt wiederholt. Wenn es First Screens in der Liste gibt, so werden diese auf der Webseite dargestellt. Der Anwender hat zwei Möglichkeiten, den Second Screen mit einem First Screen zu verknüpfen. Er kann den First Screen über einen Button an der dargestellten Liste auswählen oder mit Hilfe der Eingabe der ScreenID des First Screens diesen direkt auswählen. Sobald eine Auswahl getroffen worden ist, wird die Anzeige der dargestellten First Screens entfernt und eine REST-Anfrage an die *Communication*-Komponente der Logikschicht gesendet. Es werden die Hintergrundinformationen abgefragt, die bereits für das Video des ausgewählten First Screens zur Verfügung gestellt wurden. Wenn über diese Anfrage Hintergrundinformationen erhalten werden, so werden diese auf der Webseite der Second Screen-Instanz dargestellt. Im nachfolgenden oder falls es keine empfangenen Informationen gegeben hat, wird das *Topic* des First Screens beim *Message Broker* abonniert.

Entsprechend der Laufzeit des Videos des First Screens werden neue Hintergrundinformationen durch die Logikschicht an das *Topic* gesendet. Diese Hintergrundinformationen werden durch die Second Screen-Instanz asynchron empfangen. Die neuen Informationen werden auf der Webseite der Instanz dargestellt. Aufgrund des asynchronen Versendens kann es notwendig sein, die empfangenen Hintergrundinformationen entsprechend der zugeordneten Laufzeit für die Darstellung zu sortieren.

3.3 Ablauf der Generierung der Hintergrundinformationen

In diesem Abschnitt wird der Ablauf der Generierung der Hintergrundinformationen erläutert. Sie erfolgt automatisiert und kann mit den Mechanismen erfolgen, die in diesem Abschnitt aufgeführt werden. Die Generierung der Hintergrundinformationen lässt sich in die Schritte der Vorverarbeitung (siehe 3.3.1) und des Entity Linking (siehe 3.3.2) unterteilen.

Für die Generierung von Hintergrundinformationen werden Teile der Umsetzung von externen Projekten bzw. Software-Bibliotheken übernommen. Diese kommen in den Arbeitsschritten der Vorverarbeitung, des Named Entity Linking und des Geocodings zur Anwendung. Der Ablauf zur Generierung der Informationen sowie die verwendeten Bibliotheken und Schnittstellen werden nachfolgend ausgeführt.

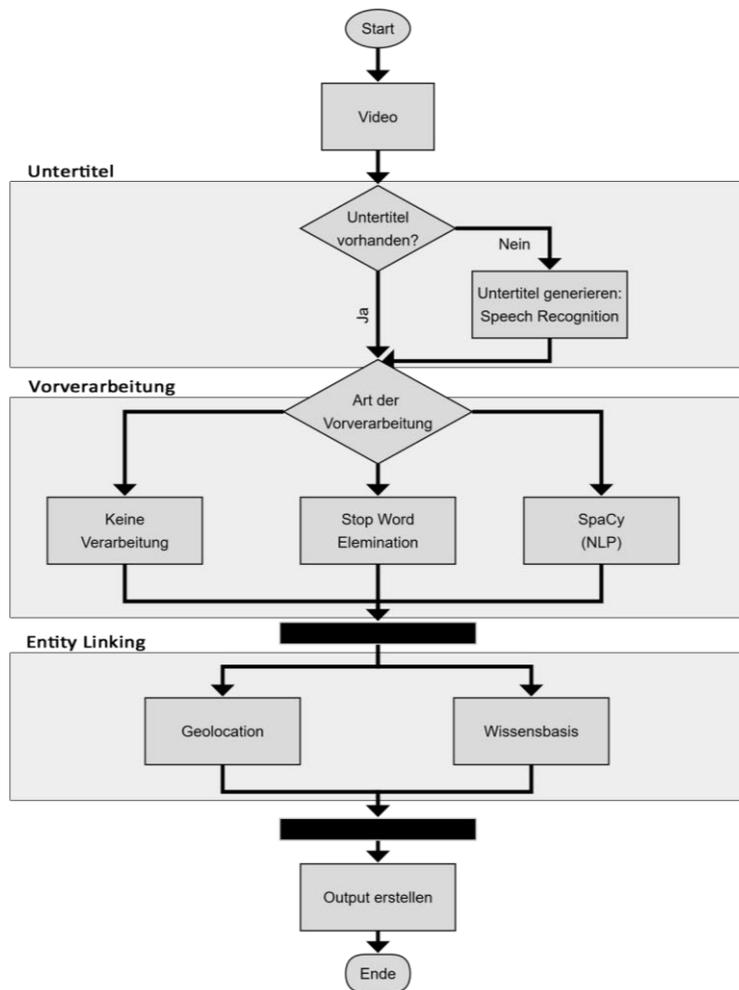


Abbildung 11: Ablauf der Content-Generierung für die Second Screen - Anwendung als Flussdiagramm [2]

In Abbildung 11 ist der Ablauf der Generierung der Hintergrundinformationen dargestellt. Für die Generierung der Informationen müssen Videodaten vorliegen. Die Informationen sollen anhand der Untertitel der Videos generiert werden. Für viele Videos sind Untertitel verfügbar. Gibt es keine Untertitel zu den Videos, so ist die automatisierte Generierung der Untertitel aus der Audiospur der Videos mittels Spracherkennung möglich [2]. In dieser Ausarbeitung wird die automatisierte Generierung der Untertitel nicht näher betrachtet.

Im Folgenden werden die vorhandenen Untertitel durch mehrere optionale Methoden vorverarbeitet. In Abschnitt 3.3.1 wird die Umsetzung des Stopword Removal und des Named Entity Recognition ausgeführt. Die Vorverarbeitung kann optional ausgelassen werden. In der Vorverarbeitung werden die Entitäten, die weiter verarbeitet werden, aus den Untertiteln extrahiert und so der Textkorpus reduziert. Die Art der Vorverarbeitung beeinflusst die Ergebnisse der darauffolgenden Verarbeitungen.

Nach der Vorverarbeitung erfolgt das Entity Linking (siehe 3.3.2). Hier werden Informationen aus den Untertiteln mit Informationen anderer Quellen verknüpft. Die Informationen, die mit den Entitäten der Untertitel verknüpft wurden, werden für die Anreicherung der Fernsehdokumentationen mit Informationsergänzungen verwendet. Beim Geocoding werden die Entitäten der Vorverarbeitung mit Geodaten assoziiert und beim Named Entity Linking werden die Entitäten mit Einträgen einer Wissensbasis verknüpft.

3.3.1 Vorverarbeitung

Die Vorverarbeitung der Daten ist ein wichtiger Schritt. Wie in Abschnitt 2.5 erläutert, beeinflusst dieser Schritt die Qualität und Performance des gesamten Verarbeitungsprozesses. Aus den Untertiteln werden Wörter ohne Mehrwert entfernt und Begriffe können kategorisiert werden. Der Textkorpus bzw. die Anzahl der Entitäten wird so reduziert, was zu einer schnelleren Berechnung und zu einer Verbesserung der Ergebnisse führen kann.

Mit dem Stopword Removal werden Begriffe einer Liste gezielt aus dem Textkorpus entfernt. Mit dem Named Entity Recognition wird ein komplexeres Verfahren eingesetzt, mit dem Begriffe aus dem Textkorpus extrahiert und klassifiziert werden, die als bedeutend angesehen werden.

Stopword Removal

Für den Vorverarbeitungsschritt des StopWord Removal (siehe 2.5.2) ist eine Liste mit Stoppwörtern erforderlich. Diese Liste kann selbst erstellt werden und hinsichtlich ihres Umfangs variieren. Die Stoppwort-Listen unterscheiden sich für jede Sprache, je nach Sprache und definiertem Umfang der Liste.

Das Stopword Removal wird mit Hilfe der NLTK-Bibliothek umgesetzt. Die Bibliothek enthält Stoppwort-Listen in mehreren Sprachen. Im Rahmen dieser Ausarbeitung wird die Stoppwort-Liste für die deutsche Sprache der NLTK-Bibliothek verwendet. Die deutschsprachige Stoppwort-Liste enthält 232 Wörter und die englischsprachige Stoppwort-Liste enthält 179 Wörter.

Die Stoppwort-Listen der NLTK-Bibliothek sind leicht verfügbar und mit der Apache-2.0-Lizenz zu erhalten. Es stehen Stoppwort-Listen für 23 Sprachen zur Verfügung. Diese Liste kommt aufgrund ihres Umfangs und ihrer Verfügbarkeit zur Anwendung. Für das Entfernen der Stoppwörter ist die Anwendung anderer Listen als die gewählte möglich. Die Wahl der Liste kann die Qualität und Performance der Verarbeitung beeinflussen. Das Entfernen der Stoppwörter aus dem Textkorpus der Untertitel erfolgt mit Hilfe der NLTK-Bibliothek.

Named Entity Recognition

Bei dem Prozess des Named Entity Recognition (siehe 2.5.3) werden Named Entities aus einem Textkorpus extrahiert und in verschiedene Kategorien klassifiziert. Hierzu stehen mehrere Bibliotheken, wie die NLTK-Bibliothek und SpaCy, zur Verfügung.

Das Named Entity Recognition wird mit der SpaCy-Bibliothek umgesetzt. SpaCy ist gut dokumentiert und leicht zu verwenden. Bei SpaCy handelt es sich um eine Open Source-Bibliothek, die unter der MIT-Lizenz zur Verfügung steht und vortrainierte Modelle für mehrere Sprachen enthält. Des Weiteren können eigene Modelle trainiert werden. Die Wahl fiel auf SpaCy, doch auch andere Bibliotheken und selbsttrainierte Modelle können verwendet werden.

3.3.2 Entity Linking

Mit Hilfe der Verfahren, die beim Schritt des Entity Linkings zur Anwendung kommen, werden Entitäten mit Informationen assoziiert, die im späteren als Informationsergänzungen die Fernsehdokumentation anreichern. Für die Generierung der Informationsergänzungen ist dieser Schritt elementar.

Mit dem Geocoding und dem Named Entity Linking werden für die Informationsverknüpfungen zwei Gruppen von Verfahren genutzt. Das Geocoding ermöglicht die Zuordnung zwischen einer Entität und geographischen Koordinaten. Zusätzlich erfolgt mit dem Geocoding eine Einordnung, ob die Entität einem geographischen Namen entspricht. Das Named Entity Linking wird für die Zuordnung von den Entitäten und Einträgen von Wissensbasen verwendet. Auch hier wird nicht allen Entitäten ein Eintrag aus der Wissensbasis zugeordnet.

Geocoding

Mit dem Geocoding werden den Entitäten der Untertitel, die Ortsangaben enthalten, die geographischen Koordinaten zugeordnet. Um das Geocoding zu ermöglichen werden Geodaten benötigt. Diese Daten wurden aus der GeoNames-Datenbank und von OpenStreetMap bezogen (siehe 2.4.3).

Die GeoNames-Datenbank ist umfangreich und enthält alternative Namen und Informationen, wie z.B. die Bevölkerungsanzahl von geographischen Namen. Eine eigene Instanz der GeoNames-Datenbank wird erstellt und genutzt. Das Geocoding mit den Geodaten von OpenStreetMap erfolgt über das Nominatim-Projekt [53]. OpenStreetMap ist noch umfangreicher und komplexer als die GeoNames-Datenbank. Nominatim übernimmt den Zugriff auf die Geodaten von OSM und übernimmt den Prozess des Geocodings. Auch das

Reverse-Geocoding wird mit den OpenStreetMap-Daten über Nominatim ermöglicht. Der Zugriff erfolgt online über eine externe API. Beim Geocoding liefert Nominatim die Koordinaten des der gesuchten geographischen Namen. Abhängig von der Art der geographischen Namen wird auch die dazugehörige *Bounding Box* übermittelt. Die *Bounding Box* besteht aus zwei Koordinaten, mit denen die Fläche des geographischen Namens umfasst werden kann.

Neben den Daten der GeoNames-Datenbank und OpenStreetMap können auch andere Kartendienste und Systeme für das Geocoding verwendet werden. Die Wahl fiel auf die GeoNames-Datenbank, OpenStreetMap und das Nominatim-Projekt, da diese leicht verfügbar sind und als Open Source zur Verfügung stehen.

Named Entity Linking

Für das Named Entity Linking (NEL) werden Wissensbasen benötigt, mit denen die Entitäten der Untertitel verknüpft werden. Mit Wikipedia und DBPedia stehen zwei Wissensbasen bereit, die unter freien Lizenzen verfügbar sind und auf die der Zugriff online über externe APIs erfolgen kann (siehe 2.4.2). Der Datensatz von Wikipedia ist besonders groß, umfasst zahlreiche Themen und wächst stetig. Für den Prozess des NEL mit Wikipedia wird die MediaWiki-API genutzt.

Der Datensatz von DBPedia ist ebenfalls groß und wächst stetig an. Die Daten liegen in strukturierter Form vor. Das NEL mit den Daten von DBPedia erfolgt über das *DBPedia Spotlight*-Projekt³⁶. In *DBPedia Spotlight* sind die Prozesse des *Named Entity Recognition* und des NEL integriert und werden automatisiert für gegebene Texte durchgeführt. Die Verwendung des Projekts erfolgt online über die API der Entwickler von *DBPedia Spotlight*.

Neben oder anstelle der Verwendung von Wikipedia und DBPedia ist die Nutzung anderer Wissensbasen möglich. Die Wahl fiel auf Wikipedia und DBPedia aufgrund des Umfangs und der Verfügbarkeit. Beide stehen unter der CC-BY-SA – Lizenz zur Verfügung und sind online über externe APIs zu erreichen.

³⁶ <https://www.dbpedia-spotlight.org/> (Aufgerufen am 25. September 2021)

3.4 Darstellung und Interaktion

In diesem Abschnitt wird die Art der Darstellung und Interaktion mit dem entwickelten System besprochen. Wie bereits zuvor beschrieben, besteht das System aus zwei Geräten, dem First- und dem Second Screen. Wie in Abschnitt 2.2.3 beschrieben, ergibt sich dies aus den Anforderungen. Die Nutzer des Systems sehen eine Fernsehdokumentation und sollen bei Bedarf und Interesse Informationsergänzungen parallel zur Fernsehdokumentation erhalten. Hierfür bietet es sich an, das Video bzw. die Fernsehdokumentation und die Informationsergänzungen auf zwei Geräte aufzuteilen. Wie es beim Konsum vom Fernsehprogramm oder anderen Videos üblich ist, wird auf dem Hauptbildschirm, dem First Screen, das Video gezeigt. Es wird erwartet, dass der Fokus der Anwender überwiegend auf dem First Screen liegt [83]. Es soll keine störenden Elemente geben, so dass das Erlebnis des Videos nicht durch Ablenkung gemindert wird. Die Anwender konzentrieren sich heutzutage häufig auf mehrere Geräte. Besteht das Interesse oder der Bedarf an Informationsergänzungen, so werden diese den Anwendern auf dem Second Screen angeboten. Sobald das Interesse an Informationsergänzungen besteht, sollen diese einfach und intuitiv auf dem Second Screen zur Verfügung stehen.

In der Untersuchung von Lohmüller et al. [83] wurde die Aufmerksamkeit von Anwendern bei der Nutzung von Second Screen-Anwendungen untersucht. Hinsichtlich der Darstellung und Interaktion bietet die Arbeit einige Empfehlungen. Das Design der Informationsergänzungen sollte konsistent sein. Der Second Screen ist als Webseite entworfen und folgt einem einheitlichen Stil. Darüber hinaus sollten die Informationsergänzungen mit Bildern dargestellt oder angereichert werden. Wenn Bilder oder Kartendarstellungen zur Verfügung stehen, dann reichern diese in dieser Anwendung die Texte der Informationsergänzungen an. Die Anwender entscheiden anhand ihres Bedarfs, ob sie den Second Screen zur Unterstützung heranziehen. Daher gibt es keine auditiven oder haptischen Signale, dass weitere Informationsergänzungen hinzugefügt wurden, wie dieses auch von Lohmüller et al. empfohlen wird. Darüber hinaus wird empfohlen, dass vorherige Nachrichten bzw. Informationsergänzungen dem Anwender weiterhin zur Verfügung stehen. Auch dies wird von der entwickelten Software befolgt.

Bei der Interaktion mit dem Second Screen wird auf die üblichen Interaktionskonzepte mit mobilen Endgeräten zurückgegriffen. Handelt es sich beispielsweise um ein Notebook, das als Second Screen verwendet wird, so kann eine Maus und eine Tastatur verwendet werden. Auch Zeigegesten auf dem Touchpad stehen zur Verfügung. Handelt es sich um ein Tablet oder Smartphone, so können die Zeigegesten des Touchscreens verwendet werden. Es handelt sich um die üblichen Interaktionsmöglichkeiten zur Steuerung eines Browsers.

Nachfolgend wird die Darstellung des First- und Second Screens und auch der Informationsergänzungen ausgeführt.

3.4.1 First Screen

Der First Screen umfasst zwei Ansichten, die abhängig von dem gegenwärtigen Zustand der Instanz sind. Die erste ist zuständig für die Auswahl eines verfügbaren Videos zum späteren Abspielen auf dem First Screen. Die zweite Ansicht zeigt das Video in einem Videoplayer und ist für das Abspielen des gewählten Videos zuständig.

Die Darstellung des First Screens wurde als Webseite mit einem responsiven Webdesign umgesetzt. Dementsprechend passt sich die Darstellung dynamisch an die Größe bzw. Auflösung des Anzeigegerätes an. Für die Verwendung des First Screens bietet sich ein großer Bildschirm wie beim Fernseher an, um das Video gut zu sehen. Ein großer Bildschirm ermöglicht die parallele Nutzung mehrerer Second Screens auch durch weitere Benutzer.

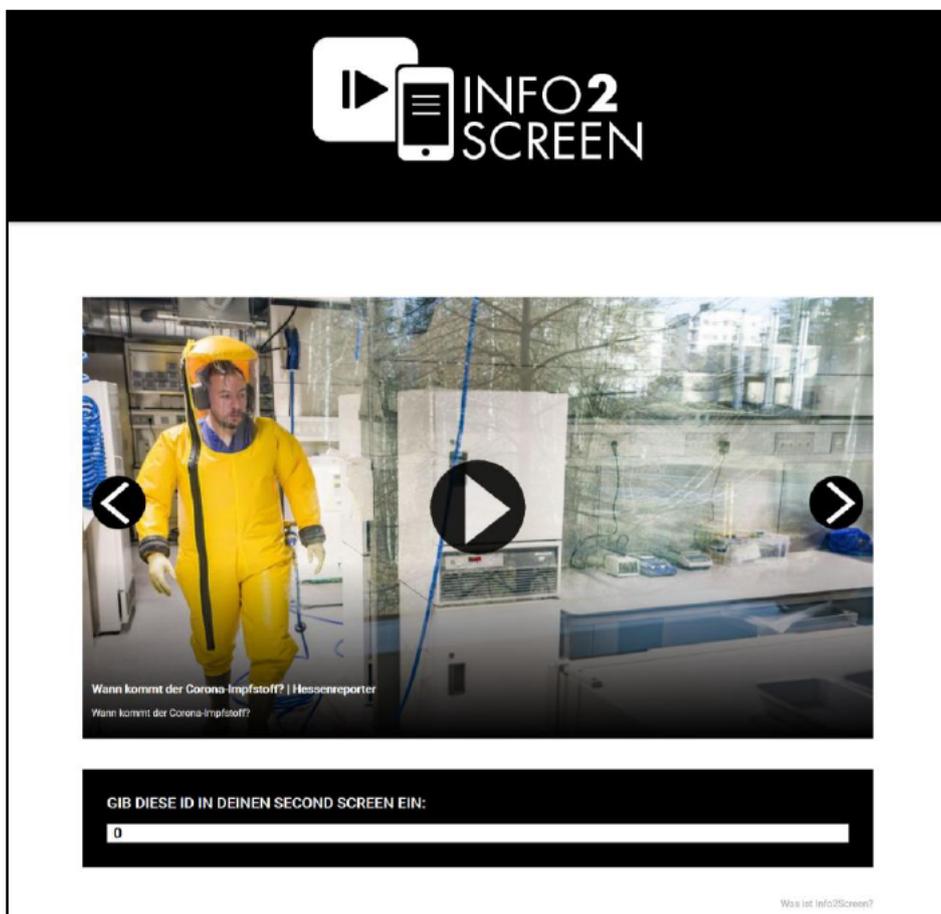


Abbildung 12: Darstellung der Videoauswahl des First Screens

In Abbildung 12 ist die Ansicht zur Auswahl eines verfügbaren Videos für den First Screen zu sehen. Unter dem Header befindet sich eine Galerie, in der die Videos dargestellt werden.

Es wird ein Vorschaubild des Videos gezeigt und im unteren Teil der Galerie befinden sich der Titel und eine Kurzbeschreibung. Je nach Eingabegerät kann die Galerie über die Tastatur, Maus oder Zeigegesten an einem Touchscreen gesteuert werden. Es kann zwischen den verfügbaren Videos gewechselt werden oder durch die Aktivierung des Play-Buttons in der Mitte des Vorschaubildes kann das angezeigte Videos ausgewählt werden. Nach der Auswahl eines Videos ändert sich die Ansicht.

Unterhalb des Galerie-Elements der Webseite befindet sich ein Element, in dem die ScreenID der Instanz des First Screens dargestellt wird. Mit Hilfe dieser Screen-ID können Second Screens mit dem First Screen verbunden werden. Diese Möglichkeit ist optional, die Verbindung zwischen dem First- und Second Screen kann auch auf anderem Wege vollzogen werden.



Abbildung 13: Darstellung der Seite des First Screens zum Abspielen der Videos

Nach der Auswahl eines Videos, ändert sich die Ansicht des First Screens. Wie in Abbildung 13 dargestellt, erscheint nun ein Videoplayer anstelle des vorherigen Galerie-Elements. In dem Videoplayer ist das ausgewählte Video eingebunden und kann abgespielt werden. Die üblichen Funktionen eines Videoplayers stehen dem Nutzer zur Verfügung. So kann das Video gestartet, gestoppt, pausiert und fortgesetzt werden oder auch zu einem bestimmten Zeitpunkt des Videos gesprungen werden. In dem Videoplayer sind die Untertitel des Videos eingebunden und können während des Abspielens des Videos angezeigt werden. Unterhalb des Videoplayers befindet sich weiterhin die individuelle Screen-ID der First Screen-Instanz.

Die Interaktionen des Nutzers mit der Webseite bzw. mit dem Videoplayer erfolgen wie in der vorherigen Ansicht über die verschiedenen Eingabegeräte. Der Videoplayer ermöglicht die

Anzeige des Videos im Vollbild. Während des Abspielens des Videos ist die Verwendung des Vollbildmodus sinnvoll, da auf dem First Screen das Video das relevante Element ist.

3.4.2 Second Screen

Der Second Screen umfasst ebenfalls zwei Ansichten, die abhängig von dem gegenwärtigen Zustand der Second Screen-Instanz sind. Die eine Ansicht ist für das Verbinden bzw. die Auswahl eines First Screens zuständig. Mit der anderen Ansicht werden die Hintergrundinformationen angezeigt, die parallel zur Laufzeit des Videos auf dem First Screen erscheinen.

Ebenso wie der First Screen wurde auch der Second Screen als Webseite mit einem responsiven Webdesign erstellt. Die Darstellung passt sich dynamisch an die Auflösung der Anzeigegeräte an. Für die Verwendung des Second Screens bieten sich mobile Endgeräte an. Tablets oder Smartphones, aber auch Notebooks sind besonders geeignet. Diese Geräte können vom Anwender, parallel zum Betrachten des Videos auf dem First Screen, genutzt werden. Sie sind klein, so dass diese nicht stören. Sie sind mobil, so dass diese beispielsweise spontan am Sitzplatz vor dem Fernseher genutzt werden können, ohne dass die Anwender an einen vorher festgelegten Platz gebunden sind. Diese Geräte können leicht zur Seite gelegt werden, wenn die Aufmerksamkeit vorrangig auf dem First Screen liegt. Sobald die Aufmerksamkeit wieder auf dem Second Screen liegt, ist dieser stets verfügbar.

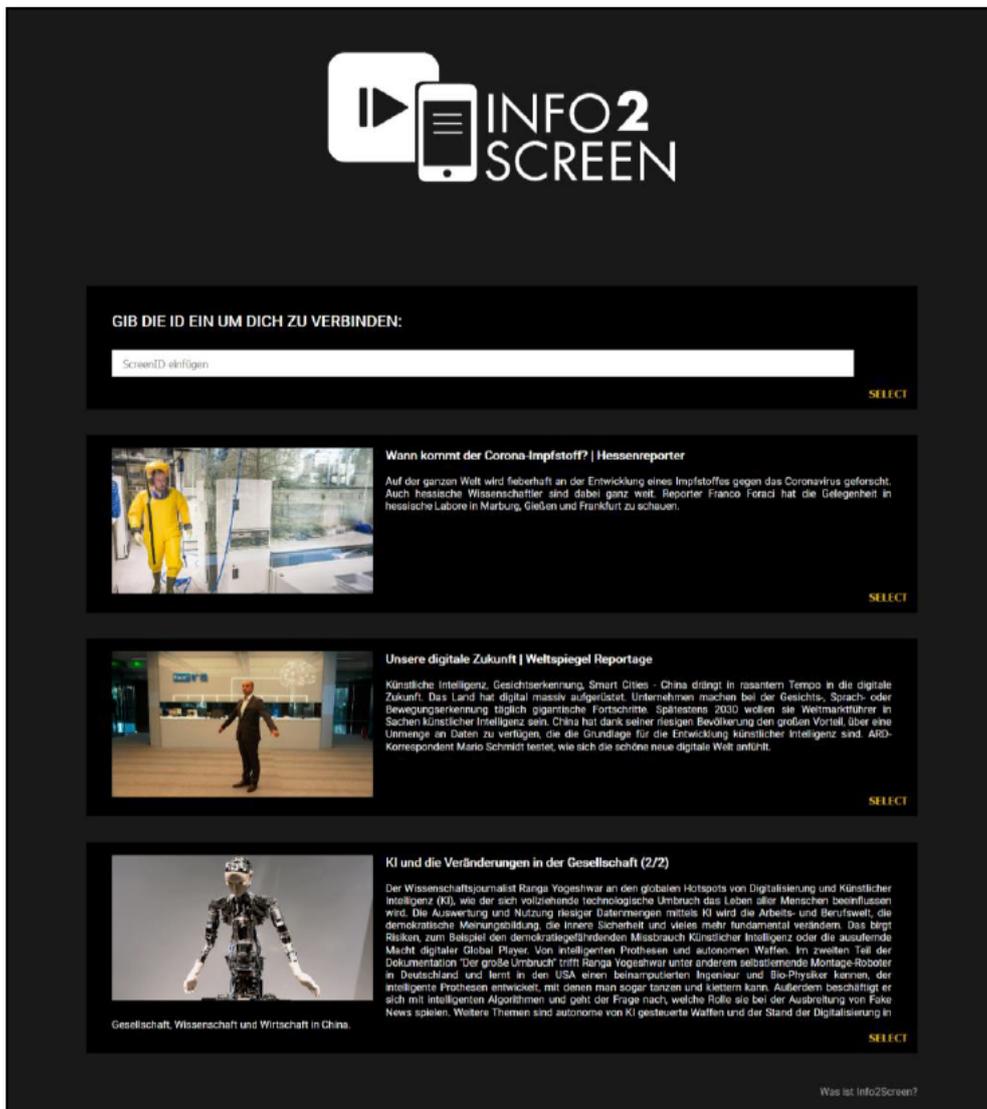


Abbildung 14: Darstellung der Second Screen-Seite zur Wahl des passenden First Screens

In Abbildung 14 ist die Ansicht zur Auswahl der First Screen-Instanz dargestellt. Unter dem Header befindet sich ein Formular, über das eine Verbindung mit dem First Screen hergestellt werden kann. Über den Select-Button kann der gewünschte First Screen ausgewählt werden, so dass sich der Second Screen mit diesem verbindet. Optional kann die ScreenID der gewünschten First Screen-Instanz eingegeben werden.

Unterhalb des Formulars befindet sich eine Auflistung der verfügbaren First Screen-Instanzen. Hier wird das ausgewählte Video mit dem Vorschaubild, dem Titel und der Kurzbeschreibung angezeigt. Über die jeweiligen Select-Buttons kann der gewünschte First Screen ausgewählt

werden. Nach der Auswahl einer First Screen-Instanz ändert sich die Ansicht des Second Screens.

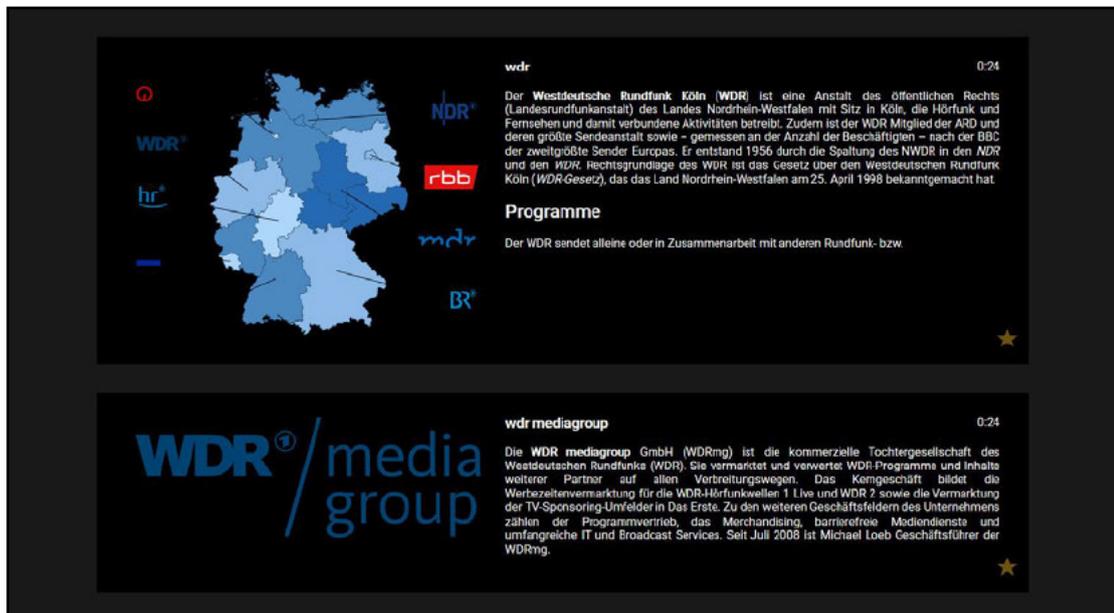


Abbildung 15: Darstellung von zwei automatisch generierten Nachrichten

In der zweiten Ansicht werden die Hintergrundinformationen dargestellt, die für das Video des gewählten First Screens zur Verfügung stehen. Entsprechend der Laufzeit werden die Informationen an die verbundene Second Screen-Instanz gesendet. Wie in Abbildung 15 gezeigt, erscheinen die Informationen untereinander. Sie sind entsprechend der zugeordneten Laufzeiten sortiert. Jede dargestellte Information enthält die zugeordnete Laufzeit, den Titel und einen Kurzttext. Der Kurzttext kann einen längeren Text einleiten oder die gewünschten Informationen bereits enthalten. Wie in Abbildung 15 ersichtlich, können den Informationen Bilder zugeordnet sein.

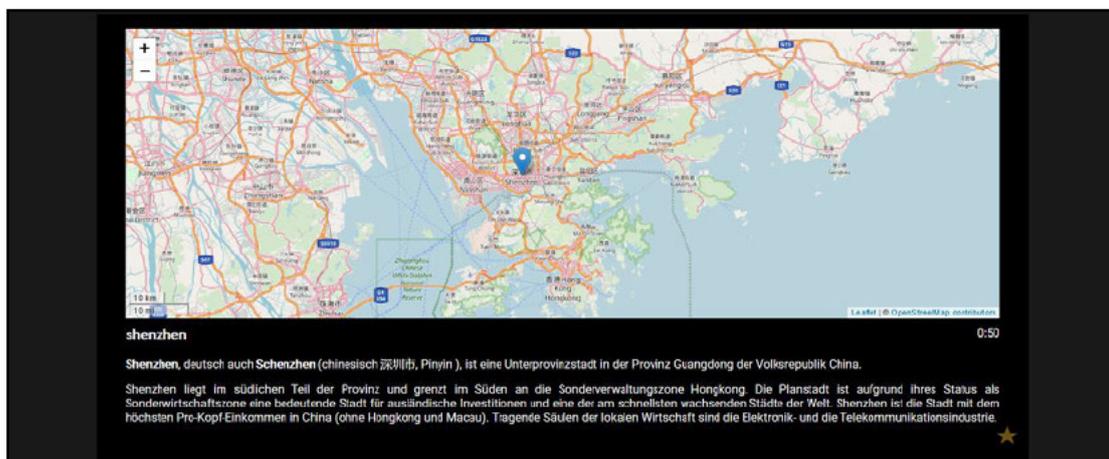


Abbildung 16: Darstellung einer automatisch generierten Nachrichte mit Kartendarstellung

Wenn es sich bei der Hintergrundinformation um einen geographischen Namen handelt, so kann dieser einem Kartenabschnitt zugeordnet sein. In Abbildung 16 ist eine Information zur chinesischen Stadt Shenzhen zu sehen, die zusätzlich einen Kartenausschnitt enthält, auf dem die Stadt Shenzhen markiert wurde. Mit den Karten können auch Regionen, die beispielsweise von Ländergrenzen umgeben sind, hervorgehoben werden.

4 Evaluation

In diesem Kapitel wird das entwickelte System zur automatisierten Informationsergänzung von Fernsehdokumentationen evaluiert. Diese erfolgt in Abschnitt 4.1 über verschiedene Messungen, mit denen die Güte der eingesetzten Verfahren untersucht wird. In Abschnitt 4.2 erfolgt eine Untersuchung des Systems mit der Befragung von Nutzern. Anschließend erfolgt in 4.3 eine Gesamtbetrachtung der Messungen und Befragungen.

4.1 Funktionale Evaluation

In diesem Abschnitt wird die Funktionalität des entwickelten Systems evaluiert. Es werden Messungen zur Evaluation der Funktionalität durchgeführt, die zur Vergleichbarkeit der Verfahren mit den Maßen Precision, Recall und dem F-Score angegeben werden. In Abschnitt 4.1.1 werden die Maßeinheiten erläutert. Die Messungen erfolgen auf Basis von drei Videos, deren Auswahl in Abschnitt 4.1.2 begründet wird. Ein wichtiger Aspekt des Systems ist die automatisierte Generierung von Hintergrundinformationen aus den Untertiteln von Videos. Dieser Abschnitt konzentriert sich auf die Evaluation dieses Teils des Systems. Die drei Videos werden genutzt, um Messungen mit den Verfahren zur Generierung der Hintergrundinformationen (siehe 3.3) durchzuführen. Die Verfahren und Messungen werden, entsprechend der Verarbeitungsschritte, in drei Bereiche eingeordnet. In Abschnitt 4.1.3 werden die Messungen des Vorverarbeitungsschrittes dargestellt. Anschließend erfolgen in Abschnitt 4.1.4 die Messungen für das Geocoding und in Abschnitt 4.1.5 für das Named Entity Linking. Die Messungen werden für jedes Video durchgeführt und untersucht. Die Geschwindigkeit der Verfahren (siehe 4.1.6) fließt ebenfalls in die Diskussion der funktionalen Evaluation (siehe 4.1.7) ein.

4.1.1 Precision, Recall und F-Score

Die einzelnen Verfahren werden miteinander verglichen und die Güte der Verfahren bewertet. Hierzu werden Maße benötigt, die das Vergleichen der Messungen ermöglichen. Für diesen Zweck sind Precision, Recall und der F-Score in dem Bereich des Information Retrievals verbreitet.

		Object is positive	Object is negative
		+	-
Classified as positive	+	True Positive	False Positive
Classified as negative	-	False Negative	True Negative

Tabelle 1: Komponenten einer Konfusionsmatrix zur Unterscheidung richtiger und falscher Klassifikation

Im Kontext des Geocodings ordnen die Verfahren Entitäten eine geographische Koordinate zu. Wird einer Entität eine Koordinate zugeordnet, so wird diese Zuordnung von dem Verfahren als positiv klassifiziert. Wenn einer Entität keine Koordinate zugeordnet wird, so wird diese Zuordnung als negativ klassifiziert. Damit ergibt sich eine binäre Klassifikation, die gemessen werden kann. Durch das Vergleichen von Referenzwerten kann entschieden werden, ob diese Zuordnungen zutreffen. In Tabelle 1 ist dieser Zusammenhang als Konfusionsmatrix dargestellt. Wurde ein Objekt als positiv klassifiziert und mit der Referenz entsprechend bestätigt, so gehört diese Zuordnung zur Menge der True Positives. Wird eine positiv klassifizierte Zuordnung nicht durch die Referenz bestätigt, so kann diese der Menge der False Positives zugeordnet werden. Dies lässt sich auch auf die negativ klassifizierten Zuordnungen übertragen. Eine negativ klassifizierte Zuordnung, die in der Referenz als positiv gilt, wird dementsprechend der Menge der False Negatives zugeordnet. Analog dazu verhält es sich im Kontext des Named Entity Linking.

Für eine Vergleichbarkeit der Verfahren werden die Maße des Precision, Recall und F₁-Score ermittelt.

Die Precision [84][39] gibt die Wahrscheinlichkeit an, mit der ein positiv klassifiziertes Objekt richtig eingeordnet wurde. Es wird auch von der Genauigkeit eines Verfahrens gesprochen. Hierzu wird die Anzahl der richtig positiv klassifizierten Objekte durch die Anzahl aller als positiv klassifizierten Objekte geteilt.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (4.1)$$

Der Wert des Recall [84][39] gibt die Wahrscheinlichkeit an, mit der ein gesuchtes Objekt identifiziert und richtig klassifiziert wird. Es wird auch von der Sensitivität oder Trefferquote

eines Verfahrens gesprochen. Für die Berechnung des Recall werden die richtig positiv klassifizierten Objekte durch die Anzahl aller real positiven Objekte geteilt.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (4.2)$$

Precision und Recall sind zwei Werte, die für die Güte eines Verfahrens betrachtet werden können. Das Vergleichen von Verfahren über mehrere Maße kann problematisch sein. Hier kann der F₁-Score helfen [39]([31], S. 151-157). Mit dem F₁-Score werden die Precision- und Recall-Werte miteinander kombiniert und als harmonisches Mittel dieser beiden Werte berechnet.

$$F_1\text{-Score} = 2 \times \frac{Precision * Recall}{Precision + Recall} \quad (4.3)$$

4.1.2 Videos für die Evaluation

Für die Evaluation der verschiedenen Verfahren zur Generierung der Hintergrundinformationen werden Videos verwendet, die bestimmten Rahmenbedingungen entsprechen. Wie in Abschnitt 3.1.3 erläutert, gehören die Videos zur Kategorie des Fernsehdokumentarismus [70], sie sind zwischen 29 und 44 Minuten lang und es sind deutschsprachige Untertitel in maschinenlesbarer Form verfügbar. Sie erfüllen die Anforderungen aus Abschnitt 2.1.1.

	Unsere Digitale Zukunft	Der große Umbruch	Wann kommt der Corona-Impfstoff?
Länge [Min]	29:07	43:33	29:02
Σ Token in Untertiteln	3328	5417	3976

Tabelle 2: Länge und Σ der Token in den Untertiteln der Videos

Die Wahl traf auf drei Fernsehdokumentationen, die im öffentlich-rechtlichen Rundfunk ausgestrahlt wurden. In der Ausarbeitung des Grundprojekts [1] wurde bereits die Weltspiegel-Reportage „*Unsere digitale Zukunft? – China und die künstliche Intelligenz*“ (UdZ) von Mario Schmidt für die Untersuchungen der Verfahren verwendet. Die Reportage behandelt das Thema der Künstlichen Intelligenz (KI) in China. Auch der Einfluss der KI auf die Gesellschaft wird betrachtet. Wie in Tabelle 2 dargestellt, ist die Reportage 29:07 Minuten lang und die Untertitel umfassen 3328 Token.

In der Ausarbeitung des Hauptprojekts [3] wurden die Verfahren anhand der Dokumentation „*Der große Umbruch – Wie Künstliche Intelligenz unsere Gesellschaft verändert*“ (DgU) von

Tilmann Wolff und Ranga Yogeshwar untersucht. Thematisch behandelt diese Dokumentation ebenfalls die Künstliche Intelligenz und ihren Einfluss auf die Gesellschaft. Neben der KI in China umfasst die Dokumentation geographisch das Thema der KI in den USA und Europa. Wie in Tab. 2 ist die Dokumentaion mit 43:33 Minuten deutlich länger und hat mit 5417 Tokens auch umfangreichere Untertitel als die UdZ-Reportage.

Als drittes Video wurde die Reportage „*Wann kommt der Corona-Impfstoff?*“ von Franco Foraci ausgewählt. Ausgestrahlt wurde die Reportage im Hessischen Rundfunk erstmals im August 2020. Thematisch handelt das Video von der Entwicklung des Corona-Impfstoffs. Mit 29:02 Minuten ist das Video so lang wie das UdZ-Video und hat mit 3976 Tokens in den Untertiteln 19% mehr Tokens.

4.1.3 Vorverarbeitung

In diesem Abschnitt erfolgen die Messungen des Arbeitsschrittes der Vorverarbeitung (siehe 3.3.1) für jedes Video. Die Vorverarbeitung ist der erste Schritt in der Erzeugung der Hintergrundinformationen bzw. der Informationsergänzungen. Die Vorverarbeitung erfolgt mit dem Ziel, den Textkorpus zu verkleinern und die nachfolgenden Verarbeitungsschritte zu unterstützen. Anschließend werden die Arbeitsschritte des Geocoding und des Named Entity Linking durchgeführt.

Für jedes Video werden die gleichen Messungen durchgeführt. Die Untertitel werden händisch verarbeitet, um einen Referenzwert zu erhalten. Aus den jeweiligen Untertiteln werden die relevanten Elemente extrahiert und den Kategorien Person (PER), Organisation (ORG), Location (LOC) und Verschiedenes (MISC) zugeordnet.

Eine weitere Art der Vorverarbeitung ist das StopWord Removal. Hier werden die Stoppwörter der Untertitel aus dem Textkorpus entfernt. In dieser Ausarbeitung wird eine Liste der NLTK-Bibliothek mit 232 unterschiedlichen Stoppwörtern verwendet, anhand derer die Entfernung vorgenommen wird.

Die dritte Art der Vorverarbeitung ist das Named Entity Recognition, bei dem bedeutsame Named Entities aus einem Textkorpus extrahiert werden. Zusätzlich werden die Named Entities klassifiziert. In dieser Ausarbeitung wird die SpaCy-Bibliothek mit dem vortrainierten Modell „de_core_news_md“-Modell in der Version 2.3.0 genutzt, die die Named Entities extrahiert und den oben genannten Kategorien zuordnet.

	PER	ORG	MISC	LOC	Summe	∑ der Token ohne doppelte & Sonderzeichen
Händisch	8,33	10	46,67	17	82	82
SpaCy	21	0	47,33	36,33	104,67	101,33
StopWord	-	-	-	-	2798	1217,67

Tabelle 3: Gemittelte Messergebnisse des Named Entity Recognition und StopWord Removal als Vorverarbeitungsschritte

Die aufgeführten Verfahren wurden für die drei ausgewählten Fernsehdokumentationen (siehe 4.1.2) angewendet. Die Messergebnisse der einzelnen Fernsehdokumentationen befinden sich im Anhang H.1. Da sich die Interpretationen der Messergebnisse der einzelnen Videos ähneln, wurden die Mittelwerte aus den Messungen gebildet und in Tabelle 3 dargestellt. Durchschnittlich enthielten die Untertitel 4240 Token. Nachfolgend werden die Ergebnisse der Vorverarbeitung besprochen.

Die Referenzwerte wurden händisch erstellt. Gemittelt konnten \varnothing 8 Personen, \varnothing 10 Organisationen, \varnothing 17 Orte und etwa \varnothing 47 Entitäten in der Kategorie „Verschiedenes“ identifiziert werden. In der Summe bedeutet dies \varnothing 82 Entitäten mit Bedeutung, die aus den Untertiteln extrahiert wurden.

Mit SpaCy wurden etwas mehr Entitäten extrahiert, als in der Referenz vorgesehen sind. Durchschnittlich wurden mit SpaCy etwa \varnothing 101 einzigartige Entitäten festgestellt. Im Schnitt wurden drei Entitäten mehreren Kategorien zugeordnet oder bestanden aus herausgefilterten Sonderzeichen. In keiner Messreihe wurden durch das vortrainierte Modell Organisationen erkannt. Es wurden \varnothing 21 Personen erkannt und somit etwa 2,5 mal mehr als in der Referenz ermittelt wurden. In der Kategorie „Location“ wurden mit etwa \varnothing 36 Einträgen knapp über dem doppelten der Entitäten der Referenz bestimmt. Knapp die Hälfte aller extrahierten Entitäten wurden in die Kategorie „Verschiedenes“ einsortiert. Mit 47 Entitäten liegt SpaCy in dieser Kategorie sehr nah an dem Wert der Referenz.

Mit Hilfe des StopWord Removal wurde die Anzahl der Token von \varnothing 4240 Token auf \varnothing 2798 Token reduziert und somit bereits um etwa $2/3$ reduziert. Nach der Entfernung aller Dopplungen verblieben \varnothing 1218 Token. Durch das StopWord Removal und dem Entfernen doppelter Entitäten verblieben vom Textkorpus durchschnittlich etwa $2/7$ aller Token. Beim StopWord Removal findet keine Klassifizierung der verbliebenen Entitäten statt. Somit erfolgte keine Einordnung in die Kategorien PER, ORG, LOC oder MISC.

Beim Vergleich zwischen SpaCy und dem StopWord Removal ergeben sich mehrere Unterschiede. Die Vorverarbeitung mit SpaCy reduziert den Textkorpus sehr deutlich. Während das StopWord Removal den Textkorpus deutlich reduzieren konnte, verbleiben hier etwa 12 mal mehr Entitäten als mit SpaCy. SpaCy hat den Vorteil, dass eine Klassifizierung

der Entitäten stattfindet, diese kann für die Weiterverarbeitung genutzt werden. Das StopWord Removal enthält bedeutend mehr Tokens. Dafür wurden nur Stoppwörter aus dem Textkorpus entfernt, so dass weiterhin die bedeutenden Entitäten im Textkorpus vorhanden sind.

4.1.4 Geocoding

In diesem Abschnitt erfolgen die Messungen des Verarbeitungsschrittes des Geocodings (siehe 3.3.2). Beim Geocoding werden geographische Namen, die in Textform vorliegen, in geographische Koordinaten umgewandelt. Die Ortsangaben können beispielsweise Ländern, Städten, Gebiete und vollständige Adressen repräsentieren. Im Rahmen dieser Ausarbeitung findet das Geocoding nach dem Schritt der Vorverarbeitung statt. Die ermittelten Tokens der Verfahren der Vorverarbeitung werden in diesem Arbeitsschritt weiterverwendet. Diese werden in geographische Koordinaten umgewandelt, die im Weiteren für eine Kartendarstellung genutzt werden können. Darüber hinaus können den verwendeten Datenquellen weitere Informationen zu den geographischen Daten entnommen werden.

Die Ergebnisse der Verfahren der Vorverarbeitung (siehe 4.1.3) bilden die Grundlage für die nachfolgenden Messungen. Die Tokens werden mit der Nominatim-API [72] für das OpenStreetMap-Kartensystem [29] und der GeoNames-Datenbank [28] verarbeitet. Zwei Verfahren, mit denen Messungen durchgeführt wurden, nutzen die GeoNames-Datenbank. Das Verfahren „GeoNames“ ist einfach gehalten und wurde in den vorherigen Ausarbeitungen angewendet [1][2]. Das Verfahren „GeoNames-Neu“ ist komplexer und integriert weitere Bestandteile der Datenbank. Es wird erhofft, die Treffer-Qualität der Verfahren mit der GeoNames-Datenbank zu verbessern. Repräsentieren die Tokens einen geographischen Namen bzw. Ortsangaben, so werden diesen die entsprechenden geographischen Koordinaten zugeordnet. Bei der Vorverarbeitung mit dem StopWord Removal-Verfahren werden alle Tokens verarbeitet. Bei der Vorverarbeitung mit dem SpaCy-Verfahren werden die extrahierten Entitäten klassifiziert. Für die Weiterverarbeitung werden die Entitäten verwendet, die in die Kategorie „Location“ eingeordnet wurden.

Der Referenz-Wert der Geocoding-Messungen wurden händisch ermittelt. Die Erstellung des Referenz-Wertes, also der händischen Umwandlung von Entitäten zu geographischen Koordinaten, erfolgte mit Hilfe der Referenz-Entitäten, die im Arbeitsschritt der Vorverarbeitung aus den Untertiteln extrahiert wurden.

Die Messungen erfolgten jeweils mit allen drei vorgestellten Fernsehdokumentationen (siehe 4.1.2). Die Messergebnisse der einzelnen Fernsehdokumentationen befinden sich im Anhang H.2. Die Messergebnisse wurden zusammengefasst und die Mittelwerte der Maße Precision, Recall und F_1 -Score gebildet (siehe 4.1.1). Anhand dieser Werte werden nachfolgend die Messergebnisse dargestellt und besprochen. Die Referenzwerte geben mit 1 den besten zu erreichenden Wert der Maße Precision, Recall und F_1 -Score wieder.

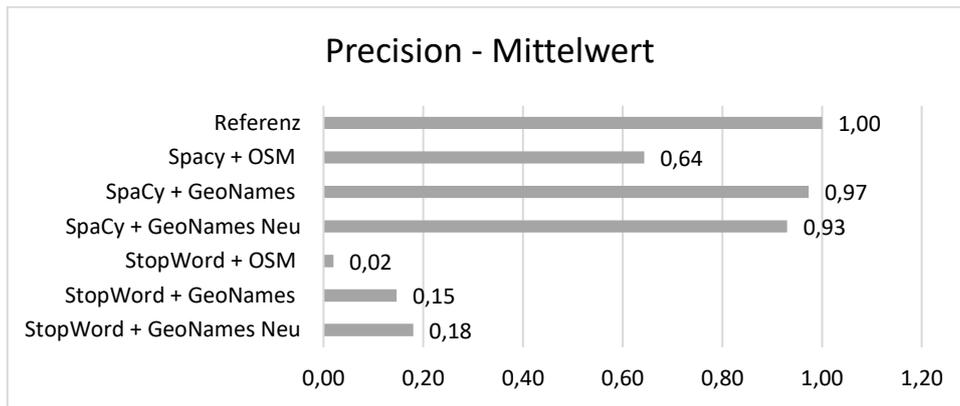


Abbildung 17: Precision-Mittelwerte der Verfahren der Geocoding-Messungen

In Abbildung 17 wird der gemittelte Precision-Wert der gemessenen Verfahren im Bereich des Geocodings abgebildet. In dieser Messreihe gibt der Precision-Wert an, mit welcher Wahrscheinlichkeit eine Entität, der eine geographische Koordinate zugeordnet wurde, mit der richtigen geographischen Koordinate verknüpft wurde und als Entität erwünscht ist. Als Entität können Begriffe enthalten sein, für die es einen gleichnamigen geographischen Namen gibt, im Kontext diesem aber nicht entspricht. Beispielsweise kann es Begriffe, wie die Handlung „essen“ geben, die von den Verfahren mit der Stadt „Essen“ verwechselt werden. Diese Entitäten gelten nicht als korrekte Zuordnungen bzw. False Positive und somit als nicht erwünscht.

Es ist in Abbildung 17 deutlich zu erkennen, dass das SpaCy-Verfahren als Vorverarbeitungsschritt dem StopWord Removal gegenüber deutlich besser abschnitt. Die Kombinationen der Verfahren, bei denen SpaCy beteiligt war, lagen im Bereich von \varnothing 0,64 bis \varnothing 0,97. Die Kombinationen der Verfahren, bei denen das StopWord Removal beteiligt war, lagen nur im Bereich von \varnothing 0,02 und \varnothing 0,18. Es wird deutlich, dass die Art der Vorverarbeitung einen großen Einfluss auf die Precision-Werte haben kann. Beim StopWord Removal werden nur die Stopwörter entfernt, so dass deutlich mehr unerwünschte Wörter, wie beispielsweise Adjektive und Verben, an das Geocoding-Verfahren übergeben werden und von diesem korrekt aussortiert werden müssen.

Ebenfalls lässt sich aus der Abbildung 17 ablesen, dass die einzelnen Verfahren des Geocoding zum Teil deutliche Unterschiede aufweisen. Das Geocoding mit OpenStreetMaps (OSM) erreicht bei der Vorverarbeitung mit SpaCy und dem StopWord Removal die geringsten Precision-Werte. Deutlich besser schneiden die Verfahren mit der GeoNames-Datenbank ab. Das „GeoNames“-Verfahren schneidet bei der Vorverarbeitung mit SpaCy leicht besser ab, als das „GeoNames Neu“-Verfahren. Bei der Vorverarbeitung mit dem StopWord Removal ist dies umgekehrt: Das „GeoNames Neu“-Verfahren erreicht hier einen höheren Precision-Wert als das „GeoNames“-Verfahren.

Die besten Precision-Werte können das „GeoNames“-Verfahren mit \varnothing 0,97 und das „GeoNames Neu“-Verfahren mit \varnothing 0,93 erreichen. Diese Werte liegen relativ nah an dem idealen Wert von 1.

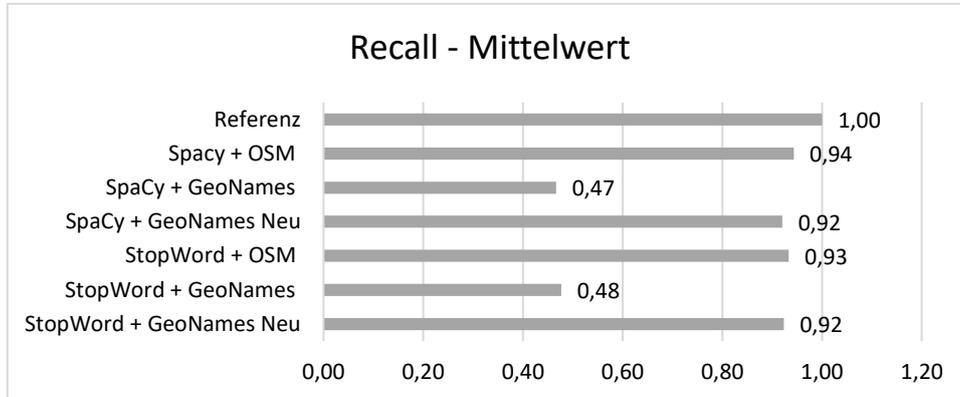


Abbildung 18: Recall-Mittelwerte der Verfahren der Geocoding-Messungen

In Abbildung 18 wird der gemittelte Recall-Wert der gemessenen Verfahren im Bereich des Geocoding abgebildet. In dieser Messreihe gibt der Recall-Wert an, mit welcher Wahrscheinlichkeit die gewünschten Entitäten identifiziert und mit der korrekten geographischen Koordinate verknüpft werden. Je höher der Recall-Wert, desto niedriger ist die Wahrscheinlichkeit, dass erwünschte Entitäten nicht identifiziert oder falsch klassifiziert werden.

Bei jedem der drei verwendeten Geocoding Verfahren haben sich bei den Messungen ähnliche Recall-Werte ergeben, unabhängig vom Verfahren der Vorverarbeitung. Dies ist in den Tabellen im Anhang H.2 gut abzulesen. Die Vorverarbeitung beeinflusst den Recall-Wert nur, wenn gewünschte Entitäten aussortiert bzw. nicht an den weiteren Verarbeitungsschritt übergeben werden. Für die Abbildung 18 wurden die Messergebnisse gemittelt. Hier ist gut zu erkennen, dass die Vorverarbeitungsschritte im Mittel keinen relevanten Einfluss auf die Güte der Verfahrenskombinationen hinsichtlich des Recall-Wertes haben.

Aus der Abbildung 18 ist ersichtlich, dass sich die Verfahren des Geocodings hinsichtlich des Recall-Wertes unterscheiden. Die Verfahren „GeoNames Neu“ und das OpenStreetMap-Verfahren differieren in den ermittelten Werten nur marginal und erreichen Werte zwischen \varnothing 0,92 und \varnothing 0,94. Dies liegt nah bei der Güte der Referenz mit dem idealen Wert von 1. Das „GeoNames“-Verfahren erreicht dagegen nur \varnothing 0,47 oder \varnothing 0,48 Punkte.

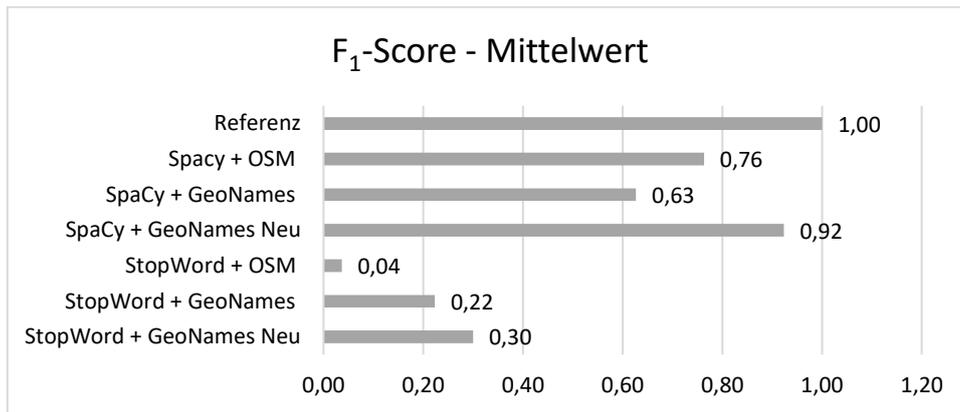


Abbildung 19: F₁-Score-Mittelwerte der Verfahren der Geocoding-Messungen

In Abbildung 19 wird der gemittelte F₁-Score der Messungen mit den Verfahren des Geocodings dargestellt. Der F₁-Score kombiniert die Precision- und Recall-Werte und ermöglicht eine Vergleichbarkeit der Verfahren. Die Güte der Verfahren hinsichtlich des F₁-Scores ist sehr unterschiedlich. Die ermittelten Werte haben eine hohe Bandbreite und liegen zwischen \varnothing 0,04 und \varnothing 0,92. Auch beim F₁-Score liegt der Referenzwert bei idealen 1,00.

Anhand der Abbildung 19 ist ersichtlich, dass die Verfahren mit dem Vorverarbeitungsschritt des StopWord Removals mit \varnothing 0,04 bis \varnothing 0,30 die niedrigsten Werte erreicht. Die Verfahren mit SpaCy als Vorverarbeitungsschritt erreichen mit \varnothing 0,63 bis \varnothing 0,92 deutlich höhere Werte.

Die Kombination aus dem StopWord Removal und dem „GeoNames“-Verfahren erzielt mit \varnothing 0,04 mit Abstand den niedrigsten Wert. Den besten Wert erzielt die Kombination der Verfahren SpaCy und „GeoNames Neu“ mit einem Wert von \varnothing 0,92. Mit diesem Wert liegt diese Kombination recht nah am Wert der Referenz.

4.1.5 Named Entity Linking

Die Ergebnisse der Messungen des Verarbeitungsschritts des Named Entity Linkings (siehe 3.3.2) werden in diesem Abschnitt diskutiert. Beim Named Entity Linking (NEL) werden die Entitäten mit Einträgen einer Wissensbasis verknüpft. So können Informationen den Entitäten zugeordnet und zur Anreicherung von Fernsehdokumentationen verwendet werden. Vor den Verfahren des NEL wurde der Schritt der Vorverarbeitung durchgeführt. Als Basis des NEL werden die ermittelten Tokens bzw. Entitäten des StopWord Removal und SpaCy verwendet.

In diesem Abschnitt werden drei Verfahren der Verarbeitung des NEL untersucht. Für zwei Verfahren wird Wikipedia als Wissensbasis verwendet. Bei beiden erfolgt der Zugriff auf die Wikipedia-Daten über die MediaWiki-API. Bei dem einen Verfahren muss der Eintrag der Wikipedia mit dem Namen der gesuchten Entität übereinstimmen (KeyWord = Link). Bei dem anderen Verfahren wird der Eintrag der Wikipedia priorisiert, der mit dem Namen der

gesuchten Entität übereinstimmt. Gibt es keine Übereinstimmung der Namen, so wird der Eintrag mit dem höchsten Confidence-Wert gewählt. Für beide Verfahren werden SpaCy und das StopWord Removal zur Vorverarbeitung eingesetzt. Das dritte Verfahren DBPedia Spotlight [26] verwendet DBPedia als Wissensbasis und benötigt keine Vorverarbeitung. Die Messungen des DBPedia Spotlight-Verfahrens werden jeweils mit dem Confidence-Parameter von 0.7 und 0.9 gemessen.

Der Referenz-Wert der NEL-Messungen wurde händisch ermittelt. Hierbei wurden die händisch ermittelten Entitäten des Schritts der Vorverarbeitung mit Einträgen der Wikipedia verknüpft.

Die Messungen wurden für die drei vorgestellten Fernsehdokumentationen durchgeführt. Die Ergebnisse der einzelnen Messungen befinden sich im Anhang H.3. Nachfolgend wurden die Messergebnisse des NEL zusammengefasst und für die Maße Precision, Recall und F_1 -Score die Mittelwerte gebildet. Die Referenzwerte entsprechen mit 1,00 dem besten zu erreichenden Wert dieser Maße.

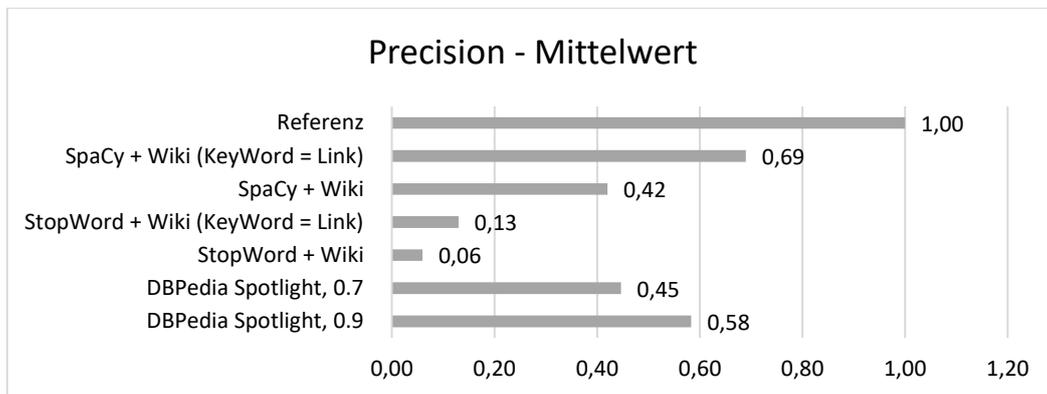


Abbildung 20: Precision-Mittelwerte der Verfahren der NEL-Messungen

In Abbildung 20 wird der Precision-Mittelwert der Messungen im Bereich des NEL abgebildet. Der Precision-Wert gibt hier an, mit welcher Wahrscheinlichkeit eine Entität, die mit einem Eintrag einer Wissensbasis verknüpft worden ist, mit dem richtigen Eintrag einer Wissensbasis verknüpft wurde und als Entität erwünscht ist. Wie beim Geocoding können nach der Vorverarbeitung Entitäten enthalten sein, die nicht erwünscht sind. Diese sollen nicht mit Einträgen einer Wissensbasis verknüpft werden.

Die gemessenen Verfahren erreichen einen Precision-Wert zwischen \varnothing 0,06 und \varnothing 0,69. Die Messungen mit dem StopWord Removal-Verfahren der Vorverarbeitung erreichen mit \varnothing 0,06 und \varnothing 0,13 die geringsten Precision-Werte. Die Messungen mit SpaCy als Vorverarbeitung erreichen dagegen höhere Werte von \varnothing 0,42 und \varnothing 0,69. Die Verfahren von DBPedia Spotlight erreichen mit \varnothing 0,45 und \varnothing 0,58 ähnliche Precision-Werte wie die Verfahren mit SpaCy als Vorverarbeitung. DBPedia Spotlight mit einem Confidence-Parameter von 0.9 schneidet besser

ab als mit einem Parameter von 0.7. Bei den Verfahren, die Wikipedia als Wissensbasis verwenden, erreicht das Verfahren, bei dem die Entität dem Namen des Wikipedia-Eintrags entsprechen muss (KeyWord = Link), bessere Precision-Werte als mit dem zweiten Wikipedia-Verfahren. Am besten schneidet SpaCy mit dem Wikipedia Verfahren (KeyWord = Link) ab.

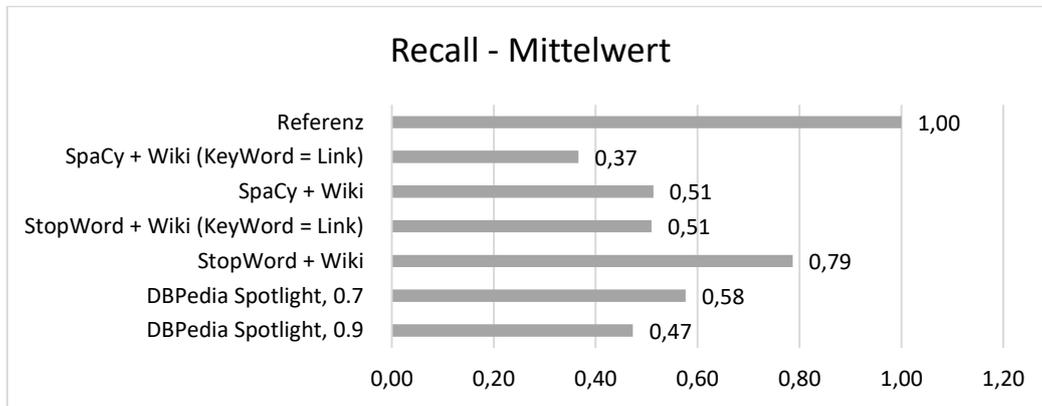


Abbildung 21: Recall-Mittelwerte der Verfahren der NEL-Messungen

Der Recall-Mittelwert der gemessenen Verfahren im Bereich des NEL wird in Abbildung 21 dargestellt. Der Recall-Wert gibt an, mit welcher Wahrscheinlichkeit die gewünschten Entitäten identifiziert und mit den richtigen Einträgen der Wissensbasis verknüpft wurden.

Der Recall-Wert der gemessenen Verfahren liegt zwischen $\approx 0,37$ und $\approx 0,79$. Im Mittelfeld liegen die Messungen für DBPedia Spotlight mit dem Confidence-Parameter 0.7 und 0.9. DBPedia Spotlight 0,7 erreicht mit $\approx 0,58$ einen höheren Wert als DBPedia Spotlight 0.9 mit $\approx 0,47$. Die Art der Vorverarbeitung hat einen deutlichen Einfluss auf die Güte der NEL-Verfahren hinsichtlich des Recall-Wertes. Wird das StopWord Removal zur Vorverarbeitung eingesetzt, so erreichen die NEL-Verfahren mit Wikipedia als Wissensbasis einen höheren Recall-Wert, als wenn SpaCy zur Vorverarbeitung verwendet wird. Dies liegt darin begründet, dass SpaCy die Entitäten extrahiert, die vom Modell für relevant erachtet werden. Hierbei kann es zu Fehlern kommen, so dass relevante Entitäten nicht extrahiert werden. Dementsprechend können diese später nicht mit einem Eintrag einer Wissensbasis verknüpft werden. Beim StopWord Removal dagegen werden nur die Stopwörter entfernt. Somit bleiben alle relevanten Token im Textkorpus erhalten und können mit einem Eintrag einer Wissensbasis verknüpft werden.

Die Kombination aus dem SpaCy zur Vorverarbeitung und der Verknüpfung mit Wikipedia, bei dem der Name der Entität dem Namen des Wikipedia-Eintrags entspricht (Keyword=Link), erreicht mit $\approx 0,37$ den geringsten Recall-Wert. Das StopWord Removal mit der Verknüpfung mit Wikipedia, bei dem der Name der Entität nicht dem Namen des Wikipedia-Eintrags entsprechen muss, erreicht mit $\approx 0,79$ den besten Recall-Wert.

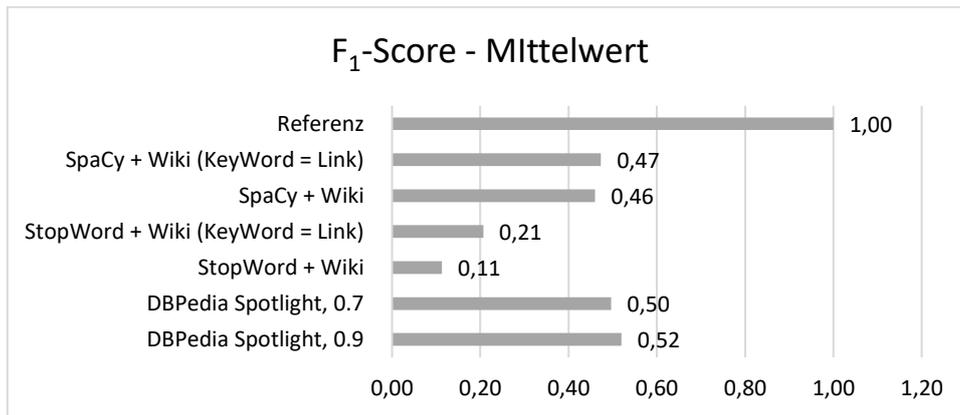


Abbildung 22: F₁-Score-Mittelwerte der Verfahren der NEL-Messungen

Der Mittelwert des F₁-Scores der gemessenen Verfahren im Bereich des NEL wird in Abbildung 22 dargestellt. Es ist ersichtlich, dass die Art der Vorverarbeitung einen signifikanten Einfluss auf die Güte der Verfahren hinsichtlich des F₁-Scores hat. Die Verfahren mit dem StopWord Removal erreichen mit $\bar{\varnothing}$ 0,11 und $\bar{\varnothing}$ 0,21 Punkten den geringsten F₁-Score. Die Verfahren, bei denen SpaCy zur Vorverarbeitung verwendet wurde, erzielen mit $\bar{\varnothing}$ 0,46 und $\bar{\varnothing}$ 0,47 deutlich bessere F₁-Score-Werte. Das NEL-Verfahren mit Wikipedia, bei dem der Name der Entität dem Namen des Wikipedia-Eintrags entspricht (Keyword=Link), schneidet jeweils etwas besser ab als die Verfahrenskombinationen mit dem anderen Wikipedia-Verfahren. Die Art des NEL-Verfahrens beeinflusst den F₁-Score nur in geringem Maße. Die Vorverarbeitung hat hier einen großen Einfluss auf den F₁-Score. Etwas besser aber in einem ähnlichen Bereich liegen beide Messungen mit DBPedia Spotlight. DBPedia Spotlight 0,9 schneidet mit $\bar{\varnothing}$ 0,52 etwas besser ab als DBPedia Spotlight 0,7 mit $\bar{\varnothing}$ 0,50.

4.1.6 Geschwindigkeit

In diesem Abschnitt werden die verwendeten Verfahren für die Generierung der Hintergrundinformationen in Bezug auf ihre Berechnungsdauer untersucht. Sie werden auch hier in die Bereiche der Vorverarbeitung, des Geocodings und des Named Entity Linkings eingeteilt. Die ermittelten Zeiten ergeben sich aus den Messergebnissen mit den Untertiteln der drei Videos aus Abschnitt 4.1.2 mit jeweils vier Messungen. Die Messungen fanden auf einem lokalen Rechner, in dem ein i7-Prozessor der vierten Generation verbaut ist, statt. Über verschiedene APIs wurde über das Internet auf Wikipedia, OpenStreetMap und DBPedia Spotlight zugegriffen. Eine Instanz der GeoNames Datenbank wurde auf dem gleichen Rechner genutzt.

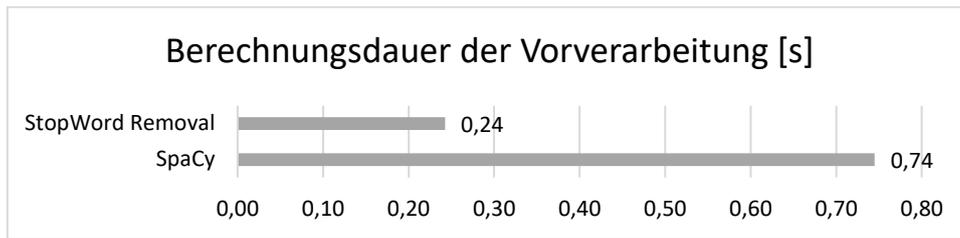


Abbildung 23: Durchschnittliche Berechnungsdauer der Vorverarbeitung in Sekunden

In Abbildung 23 wird die durchschnittliche Berechnungsdauer des Vorverarbeitungsschritts dargestellt. Das Named Entity Recognition-Verfahren von SpaCy und das StopWord Removal mit Hilfe der NLTK-Bibliothek verarbeiten die Untertitel in weniger als einer Sekunde. Das StopWord Removal ist durchschnittlich etwa dreimal schneller als das NER mit SpaCy.

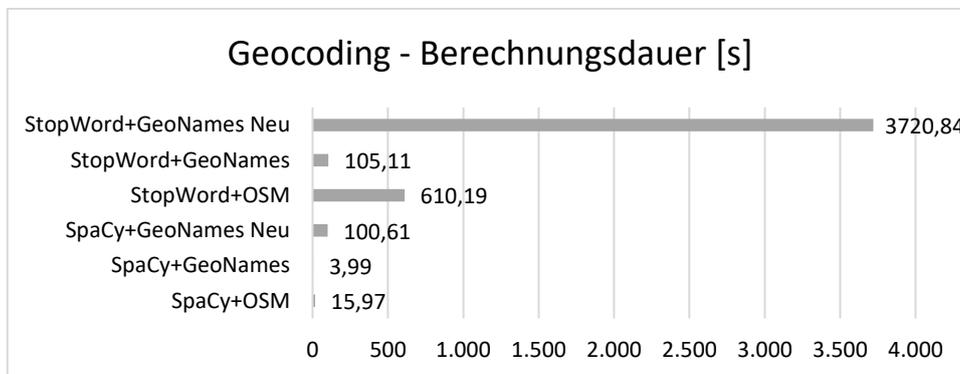


Abbildung 24: Durchschnittliche Berechnungsdauer des Geocodings in Sekunden

Das Diagramm, siehe Abbildung 24, gibt Auskunft über die durchschnittliche Berechnungsdauer der sechs genutzten Verfahren des Geocodings. Das Geocoding erfolgt mit OpenStreetMaps und der GeoNames-Datenbank auf Basis der ermittelten Entitäten des StopWord Removals und der Named Entity Recognition mit SpaCy. Die Berechnungsdauer der Vorverarbeitung fließt nicht in die Berechnung der Werte ein. Für die Verarbeitung mit der GeoNames-Datenbank werden zwei verschiedene Verfahren untersucht. Es ist deutlich zu erkennen, dass die Vorverarbeitung mit SpaCy die Berechnungsdauer der Verfahren deutlich reduziert. Wie in dem Abschnitt 4.1.3 gezeigt wurde, ist in der Vorverarbeitung die Anzahl der berechneten Entitäten mit SpaCy deutlich geringer als mit dem StopWord Removal. Des Weiteren werden mit Hilfe der Named Entity Recognition mit SpaCy die Entitäten in Kategorien, wie Locations, eingeordnet. Für das Geocoding kann die Menge der zu verarbeitenden Entitäten somit weiter reduziert werden, wenn nur Entitäten der Kategorie Location gewählt werden. Die unterschiedliche Anzahl von Entitäten wirkt sich auf die Berechnungsdauer im Geocoding aus. Die Kombination aus SpaCy und der einfachen

Verarbeitung mit der GeoNames - Datenbank ist mit knapp 4 Sekunden das schnellste Verfahren. Die Kombination aus dem StopWord Removal und der komplexeren Verarbeitung mit der GeoNames-Datenbank ist mit durchschnittlich 3720,84 Sekunden, also etwa 62 Minuten, am langsamsten.

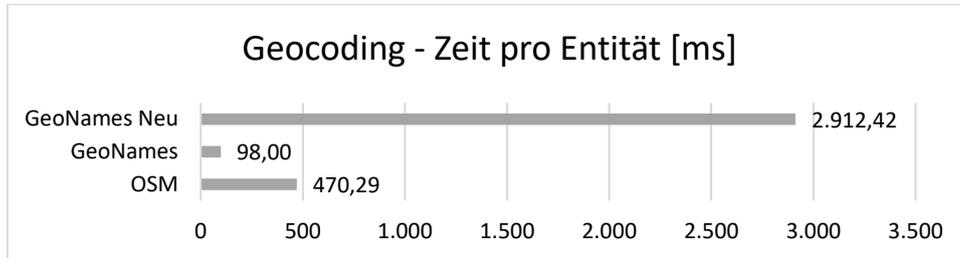


Abbildung 25: Durchschnittliche Berechnungsdauer pro Entität der Geocoding-Verfahren in Millisekunden

Wird die Berechnungszeit pro Entität, wie in Abbildung 25, betrachtet, so können die Verfahren unabhängig von der Vorverarbeitung untersucht werden. Die einfache Verarbeitung mit der GeoNames-Datenbank erzeugt die kürzeste Berechnungsdauer. Die komplexere Verarbeitung mit der GeoNames-Datenbank erhöht die Berechnungsdauer beträchtlich und führt zu dem langsamsten Verfahren. Die Verarbeitung mit OpenStreetMap ist langsamer als die einfache Verarbeitung mit der GeoNames-Datenbank, aber dennoch deutlich schneller als das komplexere Verfahren mit GeoNames. Während der Berechnungen lief die GeoNames-Datenbank Instanz auf dem gleichen Rechner wie die Anwendung. Dies kann zu den deutlich längeren Berechnungszeiten geführt haben. Ein externer und schnellerer Datenbankzugriff kann die Berechnungsdauer der Verfahren mit der GeoNames-Datenbank reduzieren. Dies gilt im Besonderen für das komplexere Verfahren mit der GeoNames-Datenbank, da mit diesem pro Entität mehrere Datenbankzugriffe erfolgen.

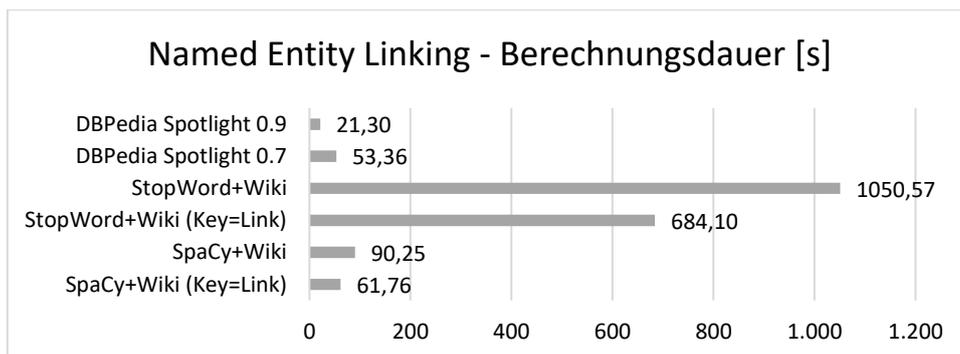


Abbildung 26: Durchschnittliche Berechnungsdauer des Named Entity Linkings in Sekunden

In Abbildung 26 sind die durchschnittlichen Berechnungszeiten der genutzten Verfahren des Named Entity Linkings (NEL) angegeben. Die Berechnungsdauer der Vorverarbeitung fließt nicht in die Berechnung der Werte ein. Das NEL erfolgt mit Wikipedia und DBPedia Spotlight. Die beiden Verfahren mit DBPedia Spotlight unterscheiden sich in der Sensitivität und ergeben in den Messreihen die schnellsten Verfahren.

Beim NEL mit Wikipedia werden zwei Verfahren verwendet. Bei dem einen Verfahren muss der Name der Entität dem Namen des Wikipedia-Eintrags entsprechen (key = link), bei dem anderen Verfahren wird dies bevorzugt, aber auch andere Matches akzeptiert. Auch beim NEL wirkt sich die Wahl des Verfahrens der Vorverarbeitung auf die Gesamtdauer aus. Der Unterschied ist deutlich zu spüren, auch wenn er kleiner als beim Geocoding ausfällt.

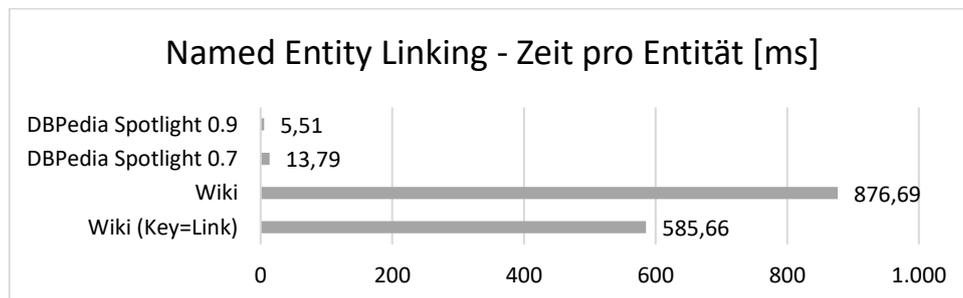


Abbildung 27: Durchschnittliche Berechnungsdauer pro Entität der NEL-Verfahren in Millisekunden

In Abbildung 27 wird die durchschnittliche Berechnungsdauer der NEL-Verfahren pro Entität dargestellt. So können die Verfahren unabhängig von der Vorverarbeitung miteinander verglichen werden. Hier zeigt sich, dass DBPedia Spotlight das schnellste Verfahren darstellt. Die Sensitivität des Verfahrens beeinflusst die durchschnittliche Berechnungsdauer pro Entität. Die untersuchten Verfahren, die auf den Daten von Wikipedia basieren, benötigen am meisten Zeit. Das strengere Verfahren (key=link) ist etwa 33% schneller als das Verfahren, bei dem mehr Matches akzeptiert werden.

4.1.7 Diskussion

In dem Abschnitt 4.1 wurden die Verfahren zur Erzeugung der Informationsergänzungen untersucht. Im Abschnitt 4.1.3 wurden die Verfahren der Vorverarbeitung näher betrachtet. In dem Abschnitt 4.1.4 erfolgte die Untersuchung des Geocodings und im Abschnitt 4.1.5 die Untersuchung des Named Entity Linkings. Die Geschwindigkeit der Verfahren wurde im Abschnitt 4.1.6 als weiterer Aspekt begutachtet. Im Zusammenspiel mit diesen Untersuchungen konnte ein guter Überblick über die Güte der einzelnen Verfahren gewonnen werden. In diesem Abschnitt findet die Diskussion zu den ausgeführten Ergebnissen statt.

Die Güte der Verfahren unterscheidet sich in den Maßen Precision und Recall. Auch wenn es ideal wäre, so ist ein Verfahren, das gut im Precision-Wert ist, in vielen Fällen nicht ebenso gut im Recall-Wert.

Precision

Wird bei einem Verfahren mehr Wert darauf gelegt, dass es möglichst wenige False Positives gibt, dann sollte auf einen hohen Precision-Wert geachtet werden. Im Bereich des Geocodings erzielen zwei Verfahren die besten Ergebnisse. Die Vorverarbeitung mit SpaCy erhöht die Ergebnisse signifikant. Die Kombinationen mit „GeoNames Neu“, mit \varnothing 0,93, und mit „GeoNames“, mit \varnothing 0,97, schneiden in der Precision besonders gut ab. Im Bereich des Named Entity Linking ist die Kombination der Vorverarbeitung mit SpaCy und der Verknüpfung mit Wikipedia, bei dem der Name der Entität dem Namen des Wikipedia-Eintrags entsprechen muss, mit \varnothing 0,69 am besten. DBPedia Spotlight 0,9 folgt mit einem Precision-Wert von \varnothing 0,58.

Recall

Wird bei einem Verfahren mehr Wert darauf gelegt, dass alle erwünschten Entitäten identifiziert und korrekt verknüpft werden, sollte auf einen hohen Recall-Wert geachtet werden. Beim Geocoding konnten die meisten Verfahren einen sehr hohen Recall-Wert im Mittel von über 0,92 erreichen. Nur das Geocoding-Verfahren „Geonames“ schnitt mit beiden Arten der Vorverarbeitung schlechter ab und erreichte etwa \varnothing 0,47. Beim Named Entity Linking konnte die Kombination aus dem StopWord Removal und der Verknüpfung mit Wikipedia, bei dem der Name der Entität nicht dem Namen des Wikipedia-Eintrags entsprechen muss, den höchsten Recall-Wert mit \varnothing 0,79 erzielen. Die anderen Verfahren des NEL lagen zwischen \varnothing 0,37 und \varnothing 0,58 dahinter.

F₁-Score

Wenn der Anspruch an die Verfahren sich nicht auf das Maß des Recalls oder der Precision fokussiert, sondern möglichst gute Werte in beiden Bereichen erfordert, so steht mit dem F₁-Score eine geeignete Metrik zur Verfügung. Beim Geocoding überzeugt die Kombination der Vorverarbeitung mit SpaCy und dem „GeoNames Neu“-Verfahren. Es wird ein F₁-Score von \varnothing 0,92 erreicht. Diese Kombination konnte in der Precision und im Recall sehr gute Werte erreichen. Im Bereich des Named Entity Linking kann kein Verfahren so hohe Werte erzielen, wie beim Geocoding. Vier Verfahren liegen im Durchschnitt um die 0,50. Dies sind beide Messungen mit DBPedia Spotlight und die Verfahrenskombinationen mit SpaCy in der Vorverarbeitung. Diese erreichen die höchsten F₁-Scores beim NEL. Bei diesen Verfahren überwiegt weder der Recall- noch der Precision-Wert signifikant.

Datenverfügbarkeit

Für die Auswahl der Verfahren kann es eine Rolle spielen, welche Daten durch die Verfahren zur Verfügung stehen. Bei den betrachteten Verfahren des NEL spielt dies eine untergeordnete Rolle. DBPedia basiert zu einem Großteil auf den Daten der Wikipedia, so dass bei den verknüpften Entitäten ähnliche Informationen verfügbar sind oder auf die entsprechenden Artikel der Wikipedia verwiesen werden kann. Beim Geocoding stehen mit OpenStreetMap und der GeoNames Datenbank unterschiedliche Informationen zur Verfügung. Bei der GeoNames Datenbank sind Informationen zur Bevölkerungsgröße erhältlich und bei dem OpenStreetMap-System ist eine Bounding Box und ein GeoJSON erhältlich. Die Bounding Box beinhaltet zwei Koordinaten, mit denen ein Bereich definiert wird, in dem sich die gesuchte Entität befindet. Das GeoJSON kann genutzt werden, um beispielsweise Flächen von Ländern auf Kartensystemen darzustellen.

Berechnungszeit

Ein weiterer Faktor für die Auswahl von Verfahren kann die Berechnungszeit sein. Erfolgt die Berechnung der Informationsergänzungen im Vorfeld, so kann dieser Faktor nachrangig sein. Ist es erforderlich, dass die Informationsergänzungen schnell berechnet werden, so ist es wichtig die Berechnungszeit der Verfahren zu beachten.

Bei der Vorverarbeitung ist das StopWord Removal etwa drei Mal so schnell wie die Vorverarbeitung mit SpaCy. Beide Verfahren konnten im Durchschnitt die gesamten Untertitel unterhalb einer Sekunde verarbeiten. Dementsprechend ist die Vorverarbeitung Für die Gesamtdauer der Berechnung der Informationsergänzungen kein großer Faktor.

Beim Arbeitsschritt des Geocodings gibt es gravierende Unterschiede in der Berechnungsdauer der Verfahren. Das schnellste Geocoding-Verfahren ist „GeoNames“ und wird gefolgt von OpenStreetMap mit der Nominatim-API, welches mehr als die 4,5-fache Zeit benötigte. Beide Verfahren verarbeiten eine Entität im Mittel unter einer halben Sekunde. Das Verfahren „GeoNames Neu“ benötigte mit fast 3 Sekunden pro Entitäten am meisten Zeit. Die Verfahrenskombinationen mit der Vorverarbeitung mit SpaCy benötigen deutlich weniger Zeit für die Verarbeitung der Untertitel, da SpaCy gewünschte Entitäten extrahiert und klassifiziert und den Textkorpus so deutlich reduziert. Das StopWord Removal dagegen entfernt nur die Stoppwörter, so dass ein großer Teil des Textkorpus erhalten bleibt. Eine Klassifizierung findet beim StopWord Removal nicht statt. Anschließend müssen die verbliebenen Entitäten verarbeitet werden. Die Anzahl ist beim StopWord Removal deutlich größer, so dass die Gesamtberechnungsdauer dementsprechend zunimmt.

Beim Arbeitsschritt des Named Entity Linking ist die Art der Vorverarbeitung, wie beim Geocoding, ein großer Faktor für die Berechnungsdauer. Am schnellsten ist DBPedia Spotlight mit einem Confidence-Parameter von 0,9.

Keine der untersuchten Kombinationen kann die Untertitel unterhalb einer Sekunde verarbeiten. Die schnellste Kombination an Verfahren für das Geocoding und das Named Entity Linking ergibt sich aus der Verarbeitung mit SpaCy, dem anschließenden Geocoding mit dem „GeoNames“-Verfahren und dem Named Entity Linking mit DBPedia Spotlight 0,9. Gemeinsam werden etwa 26 Sekunden benötigt. Beispielsweise mit einem leistungsfähigeren System könnten die Berechnungen beschleunigt werden.

Generalisierbarkeit

Im Grundprojekt wurde das Konzept der Generierung der Informationsergänzungen erdacht und gezeigt, dass dieses auf die Reportage „*Unsere digitale Zukunft – China und die künstliche Intelligenz*“ anwendbar ist. Mit der Ausarbeitung des Hauptprojekts [2] wurde im begrenzten Rahmen mit der Dokumentation „*Der große Umbruch – Wie Künstliche Intelligenz unsere Gesellschaft verändert*“ gezeigt, dass dieses Konzept auf weitere Videos übertragbar ist. Mit diesen beiden Ausarbeitungen wurde gezeigt, dass das Konzept auf Videos unterschiedlicher Länge und verschiedener Genres innerhalb der Fernsehdokumentationen anwendbar ist. Beide Fernsehdokumentationen thematisierten die Künstliche Intelligenz. In der Verortung beider Fernsehdokumentationen gibt es Überschneidungen, aber auch Unterschiede.

In dieser Arbeit wurden die Messungen für drei Fernsehdokumentationen durchgeführt. Die Messungen für die beiden benannten Fernsehdokumentationen wurden erneut durchgeführt und die Messungen für die Reportage „*Wann kommt der Corona-Impfstoff?*“ erfolgten erstmals. In der dritten Reportage wird die Entwicklung des Corona-Impfstoffs thematisiert. Somit erfolgten Messungen für eine weitere Fernsehdokumentation mit weiteren Eigenschaften, wie einem anderen Thema, das von der Reportage behandelt wird.

Die Ergebnisse der durchgeführten Messungen bestätigen die Ergebnisse des Grundprojekts und des Hauptprojekts. Mit dem dritten Video zeigen die Messergebnisse auf, dass das entwickelte System auch auf Fernsehdokumentationen in andere Themenfelder übertragbar ist und somit die Anforderung aus Abschnitt 2.1.5 erfüllt wird.

4.2 Nutzerbefragung

In Abschnitt 4.1 erfolgte die funktionale Evaluation des entwickelten Systems. In diesem Abschnitt erfolgt die weitere Evaluation über die Befragung von Nutzern des Systems mittels leitfadengestützter Experteninterviews und qualitativer Inhaltsanalyse. Mit Hilfe der Informationsergänzungen sollte ein System geschaffen werden, das die Anwender beim Betrachten der Fernsehdokumentationen unterstützt. Dem Anwender soll zum besseren Verständnis bezüglich der Fernsehdokumentationen geholfen und mögliche Wissenslücken geschlossen oder etwaiges erweitertes Informationsinteresse bedient werden. Zur Evaluation dieser Punkte haben sechs Personen das entwickelte System angewendet und wurden anschließend im Rahmen von Interviews befragt. Im Fokus der Interviews stand folgende Forschungsthese:

„Die Darstellung von Informationsergänzungen auf einem Second Screen fördert das Verständnis und Erlebnis von Fernsehdokumentationen.“

Aus der Forschungsthese ergeben sich mehrere Punkte, mit denen man diese überprüft und bewertet. Die Untersuchungen erfolgen mit dem System, das in dieser Arbeit entwickelt wurde. Die Darstellung des Systems erfolgt, wie in Abschnitt 3 erläutert, über einen First- und Second Screen. Die Fernsehdokumentation wird auf dem First Screen gezeigt und die Informationsergänzungen auf dem Second Screen. Die Darstellung und Handhabung sind Elemente, mit denen mögliche Verbesserungen des Verständnisses und Erlebnisses der Anwender beeinflusst werden können. Dementsprechend ist die Nutzungserfahrung mit dem System ein wichtiger Punkt. Wie wird die Benutzung und Handhabung des Systems mit zwei Geräten beschrieben? Oder wie wird die Darstellung der Informationen beurteilt?

Auf dem Second Screen sollen Informationsergänzungen dargestellt werden. Diese sind ein zentrales Element des entwickelten Systems und spielen auch in der Forschungsthese eine wichtige Rolle. Über die Qualität der Informationen können die befragten Anwender Auskunft geben. Waren die dargebotenen Informationen gut oder gab es überflüssige oder unerwünschte Informationen? Ebenso können Aussagen zur Quantität der Informationsergänzung getroffen werden. Zu viele Informationen könnten überfordernd wirken und zu wenige Informationen erzeugen gegebenenfalls kaum Nutzen. Ein weiterer Aspekt der Informationen ist deren Art. Im entwickelten System werden geographische Daten auf Karten dargestellt und Schlagwörter über das Named Entity Linking (siehe 2.6) mit Daten einer Wissensbasis verknüpft. Hier können Informationen zu Personen, Organisationen und Hintergrundinformationen zu weiteren Schlagwörtern gegeben werden. Gewünschte Informationsarten könnten fehlen oder dargestellte Informationsarten nicht gewollt sein.

Im Rahmen dieser Ausarbeitung ist die Verwendung des entwickelten Systems auf die Informationsergänzung von Fernsehdokumentationen ausgelegt. Die Befragten bringen unterschiedliche Perspektiven ein und können weitere mögliche Anwendungsbereiche eines

solchen Systems überblicken. Die benannten Anwendungsbereiche könnten Rückschlüsse auf die Anwendbarkeit des Systems und die Universalität des Konzepts ermöglichen.

Im Rahmen der Interviews verwenden die Befragten das entwickelte System. Mit möglichen Vorerfahrungen mit einem Second Screen und der Expertise der Befragten können diese das Gesamtsystem betrachten. Sie können bewerten inwiefern die Anwender durch die Nutzung eines Second Screens für die Darstellung der Informationsergänzungen von Fernsehdokumentationen unterstützt werden und inwiefern das Erlebnis und Verständnis gefördert wird.

Die Beschreibung der gewählten Forschungsmethode erfolgt in Abschnitt 4.2.1. Der Interviewleitfaden, mit dem die Interviews strukturiert wurden, wird in Abschnitt 4.2.2 besprochen. Die Beschreibung der Auswahl der Interviewpartner und die Durchführung der Interviews erfolgt in Abschnitt 4.2.3. Die Interviews werden in Abschnitt 4.2.4 ausgewertet und die Ergebnisse präsentiert. Abschließend erfolgt in Abschnitt 4.2.5 eine Überprüfung der durchgeführten Forschungsmethode im Hinblick auf die Gütekriterien.

4.2.1 Forschungsmethode

Als Teil der empirischen Untersuchung des Forschungsgegenstands, der Überprüfung der Forschungsthese, wurde ein leitfadengestütztes Experteninterview als Methode zur Datenerhebung gewählt. Das Experteninterview ist ein Bereich der qualitativen Forschung, bei dem es nicht um Quantifizierung und Repräsentativität geht (vgl. [85], S. 2). Es werden einige Personen befragt, die über Expertise in Bereichen verfügen, die im Zusammenhang mit dem Forschungsgegenstand stehen. Es wird davon ausgegangen, dass das Expertenwissen im gewissen Rahmen verallgemeinert werden kann, es aber durchaus Änderungen unterliegt, es Unstimmigkeiten zwischen Experten geben kann und Experten sich auch subjektiv äußern können (vgl. [86], S. 680f). Als Form des Expertenwissens ist das Deutungswissen, also die subjektive Perspektive der Experten, von Interesse. Das gewonnene Deutungswissen ist perspektivisch und mit der befragten Person verbunden. Allerdings ist es nicht ausschließlich die individuelle Sicht, sondern kann innerhalb des Kreises der Experten geteilt werden (vgl. [85], S. 18f).

In dieser Befragung werden Experten aus mehreren Bereichen befragt, die ihre individuelle Sicht auf den Forschungsgegenstand äußern und somit unterschiedliche Perspektiven abbilden. Wie von Bogner et al. (vgl. [85], S. 87ff) gefordert, wurden Qualitätsmerkmale für Experteninterviews beachtet. So stimmten die Experten dem Interview-Format, der Aufnahme des Interviews und der anschließenden Auswertung im Rahmen dieser Ausarbeitung zu. Die gewonnenen Daten wurden anonymisiert und dementsprechend sind auch die Ergebnisse der Auswertung der Daten anonymisiert.

Für die Befragung gibt es keine vorgegebenen Antwortmöglichkeiten und der Interviewleitfaden (siehe 4.2.2) ist offen gestaltet, so dass die Experten angeregt werden, frei und ausführlich zu sprechen. Der Interviewleitfaden dient der interviewenden Person als Hilfsmittel für die strukturierte Gesprächsführung der Befragung. Gegen Ende der Durchführung der Interviews (siehe 4.2.3) kann die interviewende Person den Interviewleitfaden nutzen, um zu überprüfen, ob Fragen offengeblieben sind. Der Leitfaden muss nicht entsprechend der Reihenfolge abgearbeitet werden. Das Gespräch wird locker geführt und die Experten können eigene Aspekte des Themas einbringen. Durch die Nutzung des entwickelten Systems während des Interviews gewinnen die Experten einen Überblick über die Second Screen-Anwendung. Die Befragung kann daher das entwickelte System als thematischen Gegenstand beinhalten.

Die Datenerhebung in Form des leitfadengestützten Experteninterviews ist der erste Schritt der Untersuchung des Forschungsgegenstands. Anschließend folgt die Auswertung der gesammelten Daten. Die Auswertung der gesammelten Interviewdaten erfolgt in zuvor festgelegter Weise. Zur Fixierung der Schritte wird ein Ablaufmodell definiert, das im Folgenden dargestellt wird und sich an Udo Kuckartz [87], Philipp Mayring [88] und Kuckartz & Rädiker [89] orientiert.

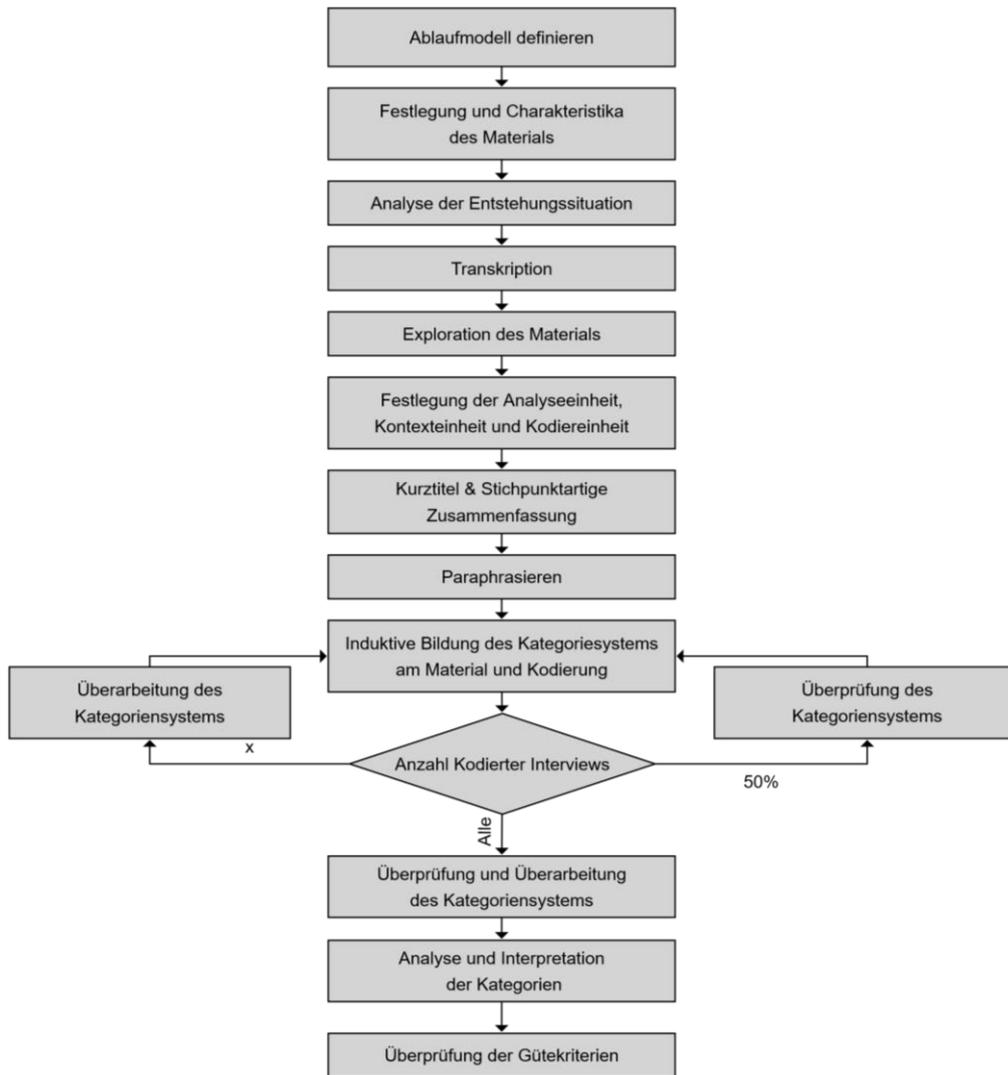


Abbildung 28: Ablaufmodell der qualitativen Inhaltsanalyse

In Abbildung 28 ist das Ablaufmodell dargestellt, anhand dessen die qualitative Inhaltsanalyse erfolgt. Initial wird das Ablaufmodell selbst entwickelt und festgelegt. Anschließend erfolgen die Arbeitsschritte, die sich mit dem Material, also den Daten, die aus den Interviews gewonnen wurden, auseinandersetzen. Zum Material, das im Rahmen dieser Arbeit bearbeitet wird, zählen alle durchgeführten Interviews, die in Rohform als Audioaufnahme vorliegen. Die Interviews wurden vom Autor dieser Ausarbeitung im freiwilligen Dialog mit den Experten durchgeführt. Die Interviews erfolgten in gewohnter Umgebung der Experten (vgl. Abschnitt 4.2.3). Aufgrund der Durchführung der Interviews durch den Autor dieser Arbeit ist ein erster Überblick über das Material gegeben.

Das Material ist zunächst in Rohfassung als Audiodatei verfügbar und wird für die folgende Verarbeitung transkribiert. Die Transkription erfolgt wörtlich und wird für jedes Interview einzeln umgesetzt. Der Abschnitt des Interviews, in dem das entwickelte System durch die Experten angewendet wird, ist nicht in der Transkription enthalten. In diesem Abschnitt ist die Audiospur der gezeigten Fernsehdokumentation dominierend. Beobachtungen des Interviewers werden an dieser Stelle niedergeschrieben.

Anhand der Transkriptionen erfolgt die Exploration der Interviewdaten, bei der sich mit dem Material bekannt gemacht wird und erste markante Aspekte wahrgenommen werden. Für jedes Interview wird eine stichpunktartige Zusammenfassung angefertigt. Auf Basis der Exploration des Materials erfolgt die Festlegung der Analyseeinheit, Kontexteinheit und der Kodiereinheit. Die Analyseeinheit (vgl. [87], S. 30f) entspricht einem ganzen Interview, als Kontexteinheit (vgl. [87], S.44) wird ein Sinnzusammenhang, der sich über mehrere Sätze und Absätze erstrecken kann, betrachtet und die Kodiereinheit (vgl. [87], S.41f), also das kleinste kodierbare Element, ist ein Wort.

Im darauffolgenden Schritt wird jedem Interview ein Kurztitel zugewiesen und die Interviews stichpunktartig zusammengefasst (Siehe Anhänge H.5). Anschließend werden die transkribierten Interviews paraphrasiert. Beim Paraphrasieren werden Textinhalte mit eigenen Worten sinngemäß zusammengefasst. Dies führt zur Komprimierung, Abstraktion und Verdichtung des Materials. Mit der Paraphrasierung der Transkriptionen wird die Vorbereitung des Materials abgeschlossen und die Analyse des Materials beginnt mit der Kategorienbildung.

In dieser Arbeit werden die Kategorien induktiv am Material gebildet (vgl. [87], S. 72ff). Bei der induktiven Bildung der Kategorien werden diese aus dem Material extrahiert. Die Kategorienbildung erfolgt ohne Vorannahmen, wie beispielsweise zuvor erfolgte theoretische Überlegung aus der Forschung oder Voruntersuchungen des Forschungsgegenstands (vgl. [88], S. 85ff). Die Kategorien werden mit Zuhilfenahme der Paraphrasen am Material gebildet. Zur Bildung der Kategorien und Kodierung des Materials wird jedes Interview betrachtet und sequenziell Textstellen bzw. Paraphrasen ermittelt, die kodiert werden sollen. Wenn die entsprechenden Textpassagen in bestehende Kategorien einsortiert werden können, so werden diese jenen subsumiert. Wenn keine passende Kategorie im Kategoriensystem enthalten ist, so wird eine neue Kategorie gebildet.

Die gebildeten Kategorien können miteinander in Beziehung stehen. Das erstellte Kategoriensystem ist hierarchisch aufgebaut, es kann Ober- und Unterkategorien geben. Die Oberkategorien können dazu genutzt werden, die Kategorien der dazugehörigen unteren Hierarchie-Ebenen zu aggregieren. Nach der Kategorienbildung eines Interviews kann das Kategoriensystem überarbeitet werden. Hier können neue Oberkategorien definiert werden, bestehende Kategorien auf eine höhere Hierarchie-Ebene gesetzt werden und bestehende Kategorien den Oberkategorien zugeordnet werden. Ebenfalls findet eine Prüfung statt, ob erstellte Kategorien zusammengeführt werden können.

Nach der Hälfte der Interviews wird das Kategoriensystem gründlich überprüft. Für das Kategoriensystem gibt es verschiedene Kriterien, die eingehalten werden sollen. Es soll plausibel, gut kommunizierbar, erschöpfend und disjunkt bzw. trennscharf sein, die Kategorien sollen sich also deutlich voneinander unterscheiden, so dass sich keine inhaltlichen Überschneidungen der Kategorien ergeben (vgl. [87], S. 85f). Die Ober- und Unterkategorien sollen in einer Beziehung stehen und die Subkategorien Aspekte der Oberkategorie repräsentieren (vgl. [89], S. 69). Mit der fortschreitenden Kodierung kann ein Punkt erreicht werden, an dem sich kaum oder keine Änderungen am Kategoriensystem ergeben. Zu diesem Zeitpunkt ist die gründliche Überprüfung des Kategoriensystems interessant, so dass das Kategoriensystem festgelegt werden kann. Die Kodierung der restlichen Interviews wird mit dem entsprechenden Kategoriensystem durchgeführt, bis das gesamte Material kodiert wurde. Eine spätere Überarbeitung des Kategoriensystems aufgrund neuer Aspekte im weiteren Material wird explizit nicht ausgeschlossen.

Vor der Auswertung und Interpretation des gebildeten Kategoriensystems wird dieses erneut überprüft. Nach der finalen Überprüfung des Kategoriensystems wird dieses in Abschnitt 4.2.4 im Sinne der Forschungsthese interpretiert und ausgewertet. Abschließend erfolgt die Überprüfung der qualitativen Inhaltsanalyse hinsichtlich der Gütekriterien (siehe 4.2.5).

4.2.2 Interviewleitfaden

Die Interviews der Expertinnen und Experten werden im Vorfeld geplant und sollen sich an der Forschungsthese orientieren. Für die strukturierte Durchführung der Interviews dient ein Interviewleitfaden als Hilfsmittel. Mit diesem kann die interviewende Person das Interview thematisch strukturiert leiten, sich an den enthaltenen Fragen orientieren und so dem Forschungsgegenstand, der Forschungsthese, thematisch nähern. Interviewleitfäden erfüllen den Zweck der Unterstützung der interviewenden Personen während der Befragung und der Strukturierung des Themenfelds (vgl. [85], S.27). Es wird ein Leitfaden für alle Interviews entworfen, damit sich die Interviewsituationen ähneln und hierdurch die Interviews vergleichbar werden [86]. In qualitativen Interviews ist es nicht notwendig, dass die Fragen standardisiert werden und allen die identischen Fragen gestellt werden. Das Ziel ist es, die Befragten in den Redefluss zu bringen, so dass Aussagen, die sich um die Fragestellung drehen, getätigt werden (vgl. [85], S. 28).

Mit Hilfe der leitfadengestützten Experteninterviews wird die Forschungsthese untersucht. Aus der Forschungsthese werden Fragen und Themen abgeleitet, mit denen die Befragten zu Aussagen und Bewertungen angeregt werden sollen, die der Untersuchung der These dienen. Nicht jedes für den Forschungsgegenstand relevante Wissen wird von Experten unmittelbar ausgesprochen oder abgerufen. Dies kann an Routinen, Traditionen und Gewohnheiten liegen. Dieses Wissen ist nicht unmittelbar zu erfragen, sondern muss aus den Aussagen rekonstruiert werden (vgl. [90], Seiten 42, 267). Die Befragten sollen angeregt werden, selbstständig in

einem Redefluss über die Themenkomplexe zu sprechen. Um diesen Effekt zu fördern, werden die Fragen möglichst offen gestaltet. Um die Befragten auf die Interviewsituation einzustimmen und deren Redefluss zu fördern, wird, neben der offenen Gestaltung der Fragen, mit der Selbstvorstellung und einer Frage begonnen, bei der die Befragten von ihren Erfahrungen berichten können.

Der Umfang des Interviewleitfadens sollte nicht zu umfangreich sein, aber das Gespräch zwischen der interviewenden und der befragten Person ausreichend unterstützen, so dass ein ausführliches Gespräch entsteht. Für den Umfang eines Interviews von ein bis zwei Stunden empfehlen Bogner et al., dass drei bis maximal acht Themen im Leitfaden enthalten sind und jeweils ein bis drei Fragen zugeordnet werden, die zur Anregung des Dialogs führen. Die Themen können weitere untergeordnete Fragen enthalten (vgl. [85], S. 28).

Die Fragen, die aus der Forschungsthese abgeleitet wurden, wurden Themenblöcken zugeordnet. Fünf Themenblöcke, die nachfolgend näher beschrieben werden, wurden in den Interviewleitfaden aufgenommen.

1. Vorerfahrung und Erwartungshaltung:

Nach einer kurzen Vorstellung wird in diesem Themenblock nach der vorhandenen Erfahrung mit einem Second Screen gefragt. Des Weiteren wird die Erwartungshaltung an ein System, das ergänzende Informationen zu Fernsehdokumentationen anbietet, erfragt. Zum Einstieg werden somit Aussagen zu der Vorerfahrung und Erwartungshaltung gesammelt. Mit Fragen nach der Anwendung der Software kann untersucht werden, ob die getroffenen Aussagen geändert werden oder die Betrachtung beeinflussen.

2. Handhabung eines Second Screens:

In diesem Themenblock wird nach der Erfahrung mit der Handhabung des Second Screens im Rahmen des angewendeten Systems gefragt. Um einen Eindruck zu erlangen, auf welchem Gerät der Fokus lag und ob sich dieser während der Anwendung des Systems änderte, soll die Aufmerksamkeit während der Nutzung des Second Screens bzw. zweier Geräte beschrieben werden. Darüber hinaus wird erfragt, auf welche Weise die Befragten die Informationsergänzungen an die Anwender des Systems übermitteln würden. Nach der Anwendung des Systems mit einem Second Screen könnten hierbei weitere Aspekte benannt werden oder die Übermittlung mittels eines Second Screens besprochen werden.

3. Bewertung der Informationen:

Die ergänzenden Informationen sind ein wichtiges Element des entwickelten Systems. Hier werden die Qualität und Quantität der bereitgestellten Informationen besprochen. Beispielsweise werden Aussagen zu sinnvollen und fehlenden Informationen angeregt, nach dem Empfinden hinsichtlich der Menge an

Informationen sowie nach einer Bewertung des Systems im Hinblick auf die Erwartungen, die vor der Anwendung des Systems geäußert wurden, gefragt.

4. Einsatzszenarien:

Für ein derartiges System sind unterschiedliche Anwendungskontexte denkbar. Die Befragten sollen sich äußern, in welchen Kontexten sie sich den Einsatz derartiger Systeme vorstellen können und welche Änderungen des Systems sie dafür vornehmen würden.

5. Bewertung und Abschluss:

Die Befragten werden nach einer Bewertung des Systems gefragt. Es soll eine Einschätzung bzw. Bewertung gegeben werden, ob das verwendete System fördernd für das Verständnis und Erlebnis von Fernsehdokumentationen ist und wie die Befragten zu dieser Einschätzung gelangen. Abschließend können weitere Punkte ergänzt werden.

Nach dem ersten Themenblock erfolgt die eigenständige Erkundung und Anwendung des entwickelten Systems. Die interviewende Person steht für Nachfragen zur Verfügung. Anschließend wird das Interview mit ausstehenden Themenblöcken des Interviewleitfadens, der sich im Anhang H.4 befindet, fortgeführt. Mit dem Interviewleitfaden wird die Möglichkeit eröffnet, dass Aspekte aus der Perspektive vor und nach der Anwendung des Systems betrachtet werden und sich mögliche Änderungen der Betrachtungsweise ergeben können.

4.2.3 Interviewpartner und Interviewdurchführung

Für die Durchführung der Interviews werden Personen benötigt, die sich als Interviewpartner zur Verfügung stellen und Expertisen in Gebieten einbringen können, die sich im Hinblick auf die Forschungsthese ergeben. In der Literatur gibt es verschiedene Ansichten, wer als Experte betrachtet werden kann (vgl. [90], S. 39ff). So kann der Expertenbegriff sehr weit gefasst werden, bei dem alle Menschen als Experten ihres Lebens gelten oder auch sehr eng gefasst werden (vgl. [86], S. 681). So schreibt M. Pfadenhauer, dass Experten nicht nur über das entsprechende Wissen verfügen müssen, sondern auch Verantwortungsträger für Entscheidungen, basierend auf dieser Expertise, sind (vgl. [90], S. 116f).

Przyborski und Wohlrab-Sahr empfehlen:

„[...] den Begriff der Expertin nur für solche Personen zu verwenden, die – soziologisch gesprochen – über ein spezifisches Rollenwissen verfügen, solches zugeschrieben bekommen und diese besondere Kompetenz für sich selbst in Anspruch nehmen. Das verbindet sich in modernen Gesellschaften häufig mit Berufsrollen, zunehmend aber auch mit Formen eines spezialisierten außerberuflichen Engagements [...]“ ([91], S. 119)

Bogner und Menz benennen mehrere Definitionen des Expertenbegriffs. In der konstruktivistischen Definition beschreiben sie den methodisch-relationalen Ansatz, bei dem der Expertenstatus einer Person von den Forschenden zugesprochen wird (vgl. [90], S. 40f). Es ist abhängig vom Forschungsgegenstand, wer als Experte gelten kann. Der Expertenstatus ist nicht universell, sondern die Expertenrolle einer Person ist auf einige Themenbereiche beschränkt. Die Experten haben ihre Expertise nur in bestimmten Wissensgebieten (vgl. [91], S. 118ff.). Dem methodisch-relationalen Ansatz folgend, kann ein Forscher nach thematischer Einarbeitung einschätzen, ob das Wissen bzw. die Expertise einer Person für den Forschungsgegenstand von Interesse ist. Die Einschätzung der Expertise und der daraus resultierenden Auswahl der Experten erfolgt anhand der zuvor entwickelten Fragen (vgl. [92], S. 72). Die Forscher selbst erkennen den Interviewpartnern die Expertenrolle zu (vgl. [92], S. 39).

Der Forschungsgegenstand dieser Befragung ist die Forschungsthese, die im Hinblick auf das entwickelte System zu betrachten ist. Im Vorfeld der Befragung wurden Überlegungen angestellt, in welchen Bereichen ein derartiges System eingesetzt werden könnte und welche Personen- oder Berufsgruppen hier beeinflusst werden könnten oder profitieren könnten. In den Vorüberlegungen stach im Besonderen der Bildungsbereich hervor. Für den individuellen Nutzer oder ganze Schulklassen und Lerngruppen könnten das System zur Wissenserweiterung und der Förderung des Verständnisses und Erlebnisses genutzt werden. Einen guten Einblick und Expertise in diesen Bereich können Menschen aus der Lehre bieten. Lehre findet beispielsweise in schulischen und auch universitären Kontexten statt. Das Medium, das mit den Informationsergänzungen angereichert wird, sind Fernsehdokumentationen, die gut recherchiert sein sollen und vielfach von Journalisten erstellt werden. Die Expertise eines Journalisten über journalistische Medien und die Veröffentlichung bzw. Verbreitung der journalistischen Recherchen kann Einblicke aus einer weiteren Perspektive bieten und hilfreich für die Überprüfung der Forschungsthese sein. Betrachtet man ein solches System als Möglichkeit des digitalen Lernens, so können sich weitere Überlegungen und Aspekte ergeben. Spaß und Interesse am Lernen fördert den Lerneffekt. Ein System zum Lernen sollte für die Lernenden ansprechend gestaltet sein. Elemente der Gamification können im Bereich des digitalen Lernens das Lernen, beispielsweise durch die Steigerung der intrinsischen Motivation, fördern [93]. Vor diesem Hintergrund kann die Betrachtungsweise eines Experten aus dem Unterhaltungs- oder Marketingbereich die Untersuchung bereichern. Personen mit Expertise im Bereich der Informationstechnologie sind den Umgang mit modernen Medien und Technologien gewohnt. Im besonderen Informatiker haben darüber hinaus Erfahrung mit der Entwicklung und den Möglichkeiten von Softwaresystemen. Experten aus diesem Bereich können die Befragung mit ihrem Wissen und ihren Einblicken ebenfalls bereichern.

Ein qualitatives Experteninterview muss nicht über den untersuchten Sachverhalt hinaus generalisierbar sein und erfordert daher keine repräsentative Stichprobe oder die Befragung

aller Experten. Die Experten wurden anhand der zuvor genannten inhaltlichen Überlegungen ausgewählt (vgl. [92], S. 71).

Die Anzahl der zu befragenden Experten ist abhängig vom Forschungsgegenstand und bestimmt den Umfang der Befragung. Im Hinblick auf die begrenzten Ressourcen, die für Untersuchungen zur Verfügung stehen, sollte die Anzahl der Experten nicht zu groß sein. Erfolgen zu viele Befragungen, so entsteht kein weiterer informationeller Mehrwert und es kam bereits zur „Sättigung“ der generierten Theorie. Die Anzahl der Befragungen muss jedoch dem Forschungsgegenstand angemessen sein (vgl. [92], S.139). In der Literatur gibt es verschiedene Ansichten zur konkreten angemessenen Zahl der Befragungen. Einige Arbeiten nennen konkrete Zahlen, andere empfehlen ein iteratives Vorgehen für jede Untersuchung und die allmähliche Festlegung der Anzahl. Die Anzahl der Befragungen muss sich an dem Forschungsgegenstand und der gewünschten Verallgemeinerung orientieren, aber auch externe Faktoren, wie die verfügbaren Ressourcen beachten (vgl. [94], S. 26f). Nach C. Helfferich werden mindestens sechs Interviews als Stichproben benötigt ([95], S. 175). Im Rahmen dieser Arbeit wurden sechs Experten befragt, die Expertise in allen oben benannten Bereichen haben.

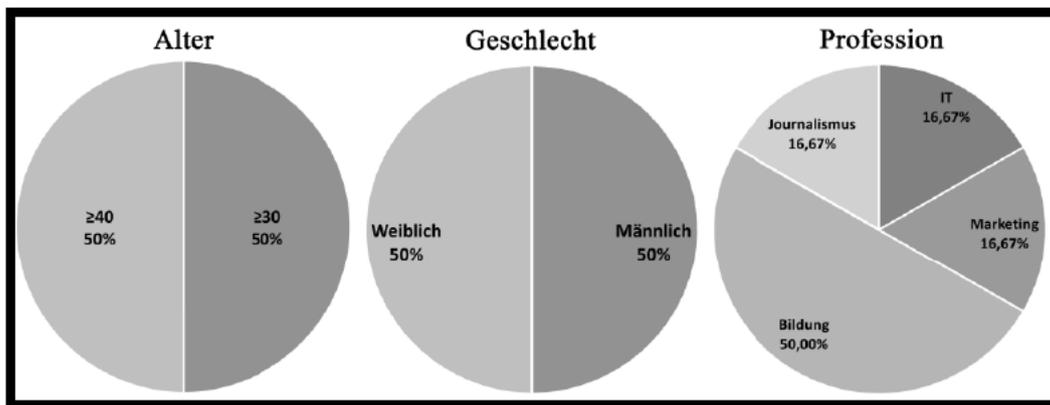


Abbildung 29: Verteilung der Merkmale der sechs befragten Experten

In Abbildung 29 wird die Verteilung der Merkmale der befragten Experten dargestellt. Das Geschlecht der Experten ist gleichermaßen verteilt. Das Alter der Hälfte der Experten ist 30 Jahre oder älter und das Alter der anderen Hälfte ist 40 Jahre oder älter. Die Hälfte der Experten ist im Bildungsbereich tätig und ist deckungsgleich mit der weiblichen Hälfte der Befragten. Zwei dieser Expertinnen sind im Schuldienst als Lehrerinnen am Gymnasium tätig und eine weitere Expertin als Wissenschaftliche Mitarbeiterin an einer Universität, die im Rahmen der Tätigkeit auch lehrt. Die männlichen Befragten arbeiten in unterschiedlichen Berufen. Darunter befindet sich ein Freier Journalist und ein Informatiker, der als IT-Administrator tätig ist. Des Weiteren wurde eine Person befragt, die im Marketing tätig ist und mit einem angefangenen Informatik-Studium ein Grundverständnis der IT mitbringt.

Jedes Interview wurde im Dialog zwischen der interviewenden Person und dem befragten Experten durchgeführt. Es wurde sich jeweils gemeinsam auf den Ort des Interviews geeinigt. Die Experten sollten sich in der Gesprächssituation wohlfühlen, daher fanden die Interviews in gewohnter Umgebung der Experten statt. Dies waren der Garten, das Wohnzimmer oder das Büro. Aufgrund der Covid-19-Pandemie fanden die Treffen corona-konform mit Abstand und frischer Luft statt. Die Interviews fanden in ungezwungener und entspannter Atmosphäre statt. Die Termine der Interviews wurden ebenfalls gemeinsam koordiniert. Diese fanden zwischen Anfang Mai und Anfang Juni 2021 statt.

Die Durchführung der jeweiligen Interviewsituationen dauerte etwa 45-60 Minuten. Der Leitfaden hat die interviewende Person strukturiert durch die Interviews geleitet und wurde flexibel gehandhabt. Der Leitfaden wurde zur Orientierung und Gedächtnisstütze verwendet, aber diesem nicht starr gefolgt. So waren das Nachhaken und weitergehende Fragen möglich, sobald die Interviewsituation dies ergab. Es konnte flexibel auf den Verlauf des Dialogs reagiert werden. Auch der befragten Person war es möglich, weitere Aspekte zu thematisieren. Diese Art der flexiblen Handhabung des Interviewleitfadens wurde u.a. von Kuckartz und Rädiker beschrieben (vgl. [89], S. VI), von Mauser und Nagel (vgl. [90], S. 80) und auch von Kassner und Wasserman (vgl. [90], S. 106)

Nach der Begrüßung und dem technischen Aufbau wurde die Durchführung der Interviews mit thematisch einführenden Worten und der Beschreibung des erstellten Systems begonnen. Anschließend wurde mit dem ersten Themenblock des Interviewleitfadens begonnen. Nach dem Abschluss der Fragen zum ersten Themenblock erfolgte die Erkundung und Anwendung der Software durch die Interviewpartner, für die mobiles Equipment genutzt wurde. Als First Screen wurde der Bildschirm eines Notebooks und als Second Screen ein Tablet verwendet. Die Anwendung erfolgte selbstständig durch die Befragten. Die interviewende Person stand bei Nachfragen unterstützend zur Verfügung. Für die Anwendung und Erkundung des Systems stand ein etwa 10-minütiger Ausschnitt des „Unsere digitale Zukunft“-Videos zur Verfügung (siehe 4.1.2). Mit Hilfe des First Screens sahen die Befragten das Video und über den Second Screen standen die ergänzenden Informationen zur Verfügung. Die Informationsergänzungen wurden durch die Kombination mehrerer Verfahren erzeugt. Mit SpaCy wurde die Vorverarbeitung durchgeführt. Mit der Nominatim-API für OpenStreetMap erfolgte das Geocoding und mit dem Wikipedia-Verfahren, bei dem der Name der Entität nicht dem Namen des Beitrags entsprechen muss, wurde das Named Entity Linking umgesetzt. Nach der Anwendungsphase wurden die weiteren Themenblöcke des Interviewleitfadens besprochen. Abschließend wurde den Befragten Raum für weitere Ergänzungen gegeben und sich für die Teilnahme bedankt.

4.2.4 Auswertung

Wie in den vorherigen Abschnitten des Kapitels 4.2 ausgeführt, wurden die Aufnahmen der durchgeführten Interviews transkribiert, den Interviews Kurztitel zugewiesen und stichpunktartig zusammengefasst. Anschließend erfolgt das Paraphrasieren der transkribierten Interviews. Die stichpunktartige Zusammenfassung gibt eine gute Übersicht über die sechs durchgeführten Interviews, fördert die Exploration des Materials und unterstützt die nachfolgende qualitative Inhaltsanalyse.

B5 - Wissenschaftliche Mitarbeiterin an einer Universität

- B5 hat einen Second Screen im Privaten zur Informationssuche oder Unterhaltung parallel zum Fernsehprogramm verwendet. Im Arbeitskontext wurde ein Second Screen zur Kommunikation parallel zu Videokonferenzen genutzt.
- Im Vorfeld werden Informationen zum Erweitern des Hintergrundwissens und nicht ausreichend vermittelter Inhalte erwartet. Außerdem sollen Informationen zu gezeigten Schauspieler:innen und zur Produktion des Videos zur Verfügung gestellt werden.
- Die Informationen sollen die Themen der Dokumentation aufgreifen, das Gezeigte in historische Kontexte und Debatten einordnen.
- Im Vorfeld wird erwartet, dass die Informationen bei Bedarf abgerufen werden und über Kategorien gefiltert werden können. Die Darstellung soll nicht zu visuell gestaltet werden, damit der Fokus auf dem First Screen liegt.
- Für B5 war die Auswahl und die Menge der Informationen gut. Wenn langsamer gelesen wird oder man an einer Information hängen bleibt, so kann der Dokumentation für einen Moment nicht mehr gefolgt werden.
- Bei der automatisierten Generierung der Inhalte wird befürchtet, dass Menschen tiefer in Filterblasen und Echokammern gelangen könnten. Die Art der Wissensbeschaffung könnte sich ändern.
- Am Anfang war die Aufmerksamkeit darauf gerichtet, sich mit dem System zurechtzufinden. Nach kurzer Zeit erfolgte die Gewöhnung und die Handhabung war gut.
- Die Anwendung kann gut im Bildungsbereich zur Wissensvermittlung genutzt werden. Mit interaktiven Fragen könnte Wissen gefestigt werden.
- Die Verwendung eines Second Screens wird auch im Hinblick auf die Wahrnehmungsweise, mit der Wissen gut aufgenommen werden kann, als gut empfunden. Die Wahrnehmungsweise ist u.a. generationsbedingt unterschiedlich.
- Es gab fehlende, falsch zugeordnete oder uninteressante, aber auch hilfreiche Informationen. Wie sinnvoll Informationen sind, ist vom Anwender abhängig.
- Es wurde zu viel Inhalt geboten, kurze stichpunktartige Informationen wären sinnvoll.

- Die Menge der Informationen war okay, mehr Informationen sollten es nicht sein. Viele Personen können von der Menge der Informationen zu sehr abgelenkt werden.
- Die Informationen haben das Verständnis der Dokumentation gefördert. Aufgrund der doppelten Aufmerksamkeit können die Informationen nachhaltiger gespeichert werden.
- Bei der automatisierten Generierung der Inhalte werden gesellschaftliche Probleme gesehen. Daher wird die redaktionelle Erstellung der Inhalte gewünscht.

Tabelle 4: Kurztitel und stichpunktartige Zusammenfassung des Interviews mit der Befragten B5.

Mit der Tabelle 4 ist die stichpunktartige Zusammenfassung und der zugewiesene Kurztitel des Interviews mit der Befragten B5 beispielhaft abgebildet. Dieses und die weiteren zusammengefassten Interviews befinden sich im Anhang H.5.

Es wurde ein Kategoriensystem anhand des Materials erstellt, das durch die Auswertung der sechs durchgeführten Experteninterviews mittels qualitativer Inhaltsanalyse gewonnen wurde. In dem Kategoriensystem befinden sich sechs Hauptkategorien mit 295 kodierten Textstellen, die teilweise mehreren Kategorien zugewiesen wurden. Die gebildeten Hauptkategorien sind ähnlich der sechs Themenblöcke aus Abschnitt 4.2.2.. Insgesamt wurden die Textstellen 429-mal den Kategorien zugeordnet. Im Nachfolgenden wird das Kategoriensystem interpretiert und analysiert.

Vorerfahrungen mit einem Second Screen

Zu Beginn der Interviews wurden die Experten gefragt, welche vorherigen Erfahrungen diese mit dem Umgang eines Second Screens gemacht haben. Es ergab sich, dass die Befragten über keine oder gelegentliche vorige persönliche Erfahrungen in der Anwendung eines Second Screens verfügen. Die Lehrerin B2 verwendet selten einen Second Screen, wenn ihr etwas auffällt und sie danach suchen möchte: *„Ich benutze einen Second Screen eigentlich nicht, aber es kann sicherlich mal vorkommen, wenn ich gerade am Fernseher sitze und mir da was auffällt, das ich da schon mal nach gegoogelt habe, ja. Also dann auf dem Handy als Second Screen, aber nicht mit `ner Software oder so.“* (B2-Interview, Abschnitt 13)

B1 beschreibt, dass die bisherige Nutzung eines Second Screens unbewusst erfolgte: *„Also ich glaube nicht so, dass es mir so ganz bewusst, dass ich mir, dass ich das so ganz bewusst wahrgenommen habe, dass ich `nen Second Screen gerade bei mir hab.“* (B1-Interview, Absatz 9). Diese unbewusste Handlung der Verwendung eines Second Screens erfolgt bei B1 sehr häufig und nebenläufig zum Medienkonsum: *„Aber, glaube ich, also ich mache das sehr häufig, dass ich irgendwie eine Serie gucke und dabei halt irgendwie auf meinem Handy Nachrichten, also Nachrichten schreibe und lese, oder. Genau. Einfach was auf meinem Handy zu machen, ich glaube am PC oder beim Laptop irgendetwas zu schauen und dann irgendwie am Fernseher und am Handy irgendwelche Sachen zu machen.“* (B1-Interview, Abschnitt 9).

Ein Second Screen wird von den Experten im beruflichen Kontext, aber auch im Privaten verwendet. Die Lehrerin B4 berichtet, dass die Schülerinnen und Schüler (SuS) im Rahmen des Unterrichts einen Second Screen nutzen: *„Ja. Also insofern, als dass mich im Lockdown ja Schüler unterrichte und zum Beispiel Jahrgang 9 auch schon ewig nicht gesehen habe und ich da Online-Unterricht mache. Und zumindest die Schüler entweder nur ein Handy benutzen für den Online-Unterricht, oder auch Handy und Bildschirm benutzen. Teilweise ist das etwas schwierig, wenn sie nur ein mobiles Endgerät haben. Und es funktioniert gut, wenn Sie einen eigenen Computer haben und ein mobiles Endgerät. Wenn Sie zum Beispiel. Wenn sie zum Beispiel „Kahoot!“ spielen im Unterricht oder im Online Unterricht. Und sie denn auf dem Bildschirm Fragen oder Bilder sehen. Und denn die Antwortmöglichkeiten auf ihrem Smartphone halt dann eingeben? Für mich ist das jetzt auch sowas Ähnliches, oder wie, oder ist es das? Fällt das dahin? Und, genau dann würde ich sagen, Ja. nutzen wir das auch. [...] Während ich also mit denen spreche, müssen sie denn auch manchmal ihre Aufgaben rausholen und, oder können auch nebenbei recherchieren, wenn ich zum Beispiel eine Frage stelle und sie ja, wenn sie Glück haben, haben Sie dann auch ein mobiles Endgerät, wo sie dann gleichzeitig, während sie mit mir halt über, über den Bildschirm sprechen, oder ein Dokument über den Bildschirm sehen, dann anderweitig recherchieren, über ihr mobiles Endgerät. Das kommt auch vor. Dass sie dann nach Vokabeln oder Satzstrukturen, dann recherchieren. Ja, also es kommt auf jeden Fall in meinem Lockdown Alltag vor.“* (B4-Interview, Abschnitt 9). Zur Corona-Zeit fand der Unterricht vielfach digital statt und erforderte die Nutzung digitaler Medien. Während des Unterrichts können die SuS ein Zweitgerät zur Recherche nutzen und parallel dazu in einer Videokonferenz im Klassenverband sein. Auch Fragespiele bieten für die SuS einen Anlass, um im Unterricht einen Second Screen zur Informationssuche zu verwenden. Die wissenschaftliche Mitarbeiterin B5 nutzt einen Second Screen im beruflichen Kontext zur begleitenden Kommunikation während langer Videokonferenzen: *„Ansonsten im Arbeitskontext, wenn ich in langen Videokonferenzen rumhänge, nutze ich meine Second Screen zum Chatten. Sei es mit Arbeitskolleg:innen oder mit Freunden.“* (B5-Interview, Abschnitt 11).

Neben der beruflichen Nutzung haben die Experten einen Second Screen auch im privaten Bereich verwendet. So hat B2 einen Second Screen im Zusammenhang mit Spielen genutzt, aber nicht parallel zu Video-Formaten: *„Ich selber habe einen Second Screen nicht benutzt beim Konsumieren von Videos oder Filmen oder sonstwas. Aber ich hab beim Spielen hab ich n Second Screen schon erlebt oder auch benutzt.“* (B2-Interview, Abschnitt 9). Beim Spielen mit einem Second Screen wurden auf diesem andere Informationen gezeigt, als auf dem Hauptbildschirm: *„Das habe ich bei der. Ja, WII war das, glaub ich, ne. Genau bei der WII war das. Dass ich dort halt den Second Screen benutzt habe als Controller und dann hatte ich dort andere Informationen wie auf dem Bildschirm.“* (B6-Interview, Abschnitt 11).

Die meisten befragten Experten haben im Vorfeld einen Second Screen parallel zum Medienkonsum verwendet. So wurden Nachrichten begleitend zur PC-Nutzung auf dem

Mobiltelefon abgerufen, oder neben dem Konsum von Dokumentationen und anderen Video-Formaten Zusatzinformationen mit dem Second Screen abgerufen, wie beispielsweise B6 erzählt: *„Ja, das kommt halt schon vor, dass wenn man halt irgendwie Dokumentationen, Filme, etc. halt irgendwie konsumiert, dass man dann halt, irgendwie parallel dazu halt, irgendwie Zusatzinformationen mit seinem Handy halt einholt, wenn halt irgendwie Fragen aufkommen oder Sachen halt unklar sind.“* (B6-Interview, Abschnitt 7).

Nutzungserfahrung mit dem System

Nach der erstmaligen Anwendung des entwickelten Systems mit einem etwa 10-minütigen Ausschnitt einer Fernsehdokumentation haben die Experten ihre Nutzungserfahrungen mit dem System dargelegt. So wurde in den Interviews beschrieben, wie das System empfunden wurde, ob der First- oder der Second Screen dominierend beachtet wurde und wie die Handhabung des Systems bewertet wird.

Bei der Erstnutzung berichteten vier von sechs Experten von Anfangsschwierigkeiten, die durch die ungewohnte Nutzung und das automatische Anfügen neuer Informationen an das Ende der Seite entstanden. B1 war abgelenkt und konnte sich nicht auf die Fernsehdokumentation konzentrieren. Dies kann daran gelegen haben, dass die Anwender für neue Informationen manuell nach unten scrollen mussten: *„Erst, erst hat's mich. Also, weil ich das, die nicht kannte, hat's mich ziemlich abgelenkt hier so. Dass ich so, konnte ich gar nicht so aufpassen, aber gerade, weil das so. Vielleicht auch, weil mich das mit dem Scrollen ja da irgendwie nicht, also das erstmal ausführen musste, wie das denn genau funktioniert?“* (B1-Interview, Abschnitt 46). Alle Experten, die von Anfangsschwierigkeiten berichteten, sagten ebenfalls, dass sich nach einer Phase der Eingewöhnung die Handhabung verbesserte. B6 vermutet, dass die Nutzung des Systems besser wird, wenn die Anwender das System nutzen und sich an dieses gewöhnen haben: *„Es ist immer schwierig das jetzt halt irgendwie in der ersten Handhabung halt so zu beurteilen. Ich glaube. Das ist natürlich jetzt irgendwie so ein. Ja, was Neues, muss man sich ein bisschen dran gewöhnen, dass es sowas halt gibt. Glaube mit der Handhabung wird auch der Gebrauch davon irgendwie fließender.“* (B6-Interview, Abschnitt 32).

Der Grund für die Anfangsschwierigkeiten liegt unter anderem an der fehlenden Gewöhnung an das System. Vielfach wurde außerdem das manuelle Scrollen als Problemfaktor angeführt: *„Also dieses Scrolling, finde ich, ist halt total hinderlich und das meinte ich auch vorhin mit dieser Anwendung, wo es halt einfach sein, es darf einen nicht bei der Dokumentation halt stören, sondern muss halt einfach so ergänzend wirkend halt sein.“* (B6-Interview, Abschnitt 58). Es wurden einige Vorschläge zur Behebung dieser Schwierigkeit angeführt. Zu den Vorschlägen gehörte, dass die Informationen oben angefügt werden und die älteren Informationen dementsprechend nach unten rutschen: *„[...] dass, halt irgendwie, diese Informationen anders, in der anderen Reihenfolge halt erscheint. Also, dass dieses Scrollen halt aufhört. So, dass einfach die, dass einfach die letzte Information ganz oben steht und dass*

ich einfach entspannt einfach auf das Ding drauf gucken kann und sehe: Okay, es wurd' jetzt halt, was weiß ich, über China halt geredet und ich sehe sofort die Karte von China, ohne dass ich jetzt halt hier nach unten scrollen muss, um dann halt den letzten Beitrag halt irgendwie zu checken so.“ (B6-Interview, Abschnitt 58). Ein weiterer Vorschlag war, dass der Fokus der Seite des Second Screens automatisch auf die neueste Information springt: „Ich hatte, ich weiß jetzt nicht, ob das an mir lag oder ob ich das unterbrochen hab, ich fand es komisch, dass ich scrollen musste quasi und dass er nicht automatisch rollt zu dem Punkt, wo ich war. Dass er immer automatisch zum nächsten scrollt, weil ich sehe okay, es ist was passiert. Es ist eine neue Nachricht gekommen, quasi.“ (B2-Interview, Abschnitt 40). Ein weiterer Vorschlag, der das manuelle Scrollen mit einer Kachel-Darstellung verhindern oder reduzieren könnte, wird im späteren weiter besprochen.

Nicht alle Befragten empfanden Schwierigkeiten bei der Benutzung des Systems. Die Lehrerin befand die Bedienung als sehr leicht, da die Software viel Arbeit abnimmt: „Also ich fand das sehr leicht, weil ich muss ja selbst nichts tun. Ich muss das ja nur in der Hand halten und dann konnte ich direkt sehen, was es da für Hintergrundinformationen liefert. Also ich muss nicht selbst aktiv werden, um das zu suchen. Und da noch groß irgendwie schreiben und eingeben in irgendwas und und auswählen, was ich jetzt für, für Quellen haben will oder so. Sondern das wurde halt einfach mir schon direkt geliefert. Das fand ich ganz praktisch.“ (B4-Interview, Abschnitt 27). Auch das manuelle Scrollen wurde nicht als problematisch betrachtet: „Ja, das war einfach, weil man sah ja vieles schon direkt. Und man musste halt einfach nur nach unten scrollen, um, um dann die aktuellen Schlagwörter, dann zu sehen, was halt gerade gesagt wurde. Das war ja ganz einfach.“ (B4-Interview, Abschnitt 29).

Die Informationen werden fortlaufend und synchronisiert mit der Laufzeit des gezeigten Videos ergänzt. Die beständige Erweiterung der verfügbaren Informationen wurde als positiv bewertet: „Und er ergänzt ja praktisch fortlaufend die Infos. Und das finde ich auch, glaube ich, eine gute Variante, als wenn man jetzt schon ganz am Anfang alle Infos „Buff“ bekäme.“ (B3-Interview, Abschnitt 45). Die Geschwindigkeit bzw. die Zeitverzögerung, mit der die Informationen, nach dem Vorkommen in der Dokumentation, auf dem Second Screen erscheinen, wurde unterschiedlich kritisch betrachtet. Für die Lehrerin B3 war die Zeitverzögerung zu groß. Die Verzögerung ist unerwünscht, es ist jedoch möglich, dass man sich auf diese einlässt: „Ich glaube, es macht auch Sinn, diese Info Stück für Stück sozusagen einzublenden, so wie das hier auch ist. Auch wenn ich manchmal gedacht hab: Ach Mensch, da hätte ich schon gerne was. Da war das noch am, vielleicht am Rechnen oder so, dass ich da schon mal gedacht hab: Och, ich will schon mal scrollen. So ungefähr. Aber da waren die Infos hier gerade noch nicht da, die tauchten dann aber auf. Sozusagen dies Simultane. Einerseits gut. Also, ich wusste nicht genau, wie das aufgebaut ist, einerseits gut, andererseits: Ach Mensch, denn wo ist denn das? Aber das war eben dann sozusagen noch nicht da. Aber wenn man sich darauf einlässt und weiß, das taucht dann auf jeden Fall gleich auf und so, dann ist das gut zu händeln.“ (B3-Interview, Abschnitt 57). Im Gegensatz dazu überraschte das schnelle

Erscheinen der Informationen die Lehrerin B4: „Dass das auch so schnell übermittelt wird, also wenn jetzt da gesagt wird, was zum Beispiel dieser, diese unbemannte, dieses unbemannte Flugobjekt oder so, also da wurde dann auch sehr schnell irgendwie das sehr schnell übermittelt was, was ist das. Also, dass man halt, wenn es lange dauern würde, würde dann keinen Sinn machen? Wenn man, wenn dann in der Doku schon nächstes Thema ist, dann, ja dann liest man nicht darüber, was halt gerade vor 3 Minuten irgendwie passiert war.“ (B4-Interview, Abschnitt 43). B4 merkt an, dass die Verzögerung nicht zu groß sein darf. Der Maßstab bzw. das Empfinden über eine annehmbare Verzögerung ist unterschiedlich.

Mit dem First- und Second Screen umfasst das System zwei Geräte, die die Aufmerksamkeit der Anwender auf sich ziehen. Die Fokussierung auf den First- oder Second Screen war bei den Experten unterschiedlich. B2 hat sich auf den First Screen konzentriert. Nur wenn der Bedarf für mehr Verständnis oder Informationen besteht, würde B2 den Fokus auf den Second Screen verlegen: „Also wenn ich so ein, gerade so eine Doku gucke, ob das Thema jetzt das ist, was ich mir ausgesucht hätte, ist ja mal dahingestellt, dann konzentriere ich mich auf den Bildschirm im Grunde. Und nur wenn ich wirklich speziell irgendwie nicht weiterkomme, oder was wissen will oder es zu langweilig finde, dann würde ich wahrscheinlich auf den Second-Bildschirm zurückgreifen.“ (B2-Interview, Abschnitt 26). Die parallele Nutzung fiel B2 schwer, so dass von den verfügbaren Informationen auf dem Second Screen nur die Überschriften gelesen wurden: „Sobald ich, sobald ich auf den zweiten Bildschirm starre, kann ich gar nicht mehr richtig zuhören quasi, was da ist und wenn ich jetzt anfangen würde, mir jetzt auch wirklich einen Text durchzulesen oder mehrere Sätze durchzulesen, dann bin ich ja komplett raus und muss erstmal wieder reinkommen. Und deswegen habe ich auch, du hast mich auch beobachtet, immer wahrscheinlich nur ganz kurz durchgescrollt und ganz kurz die Überschriften gelesen: Was sind das für Themen? Aber ich habe sie nicht gelesen, weil ich eher zugehört habe, was gerade in der Dokumentation halt erzählt wird.“ (B2-Interview, Abschnitt 30). Auch B3 hatte Schwierigkeiten mit der parallelen Nutzung beider Geräte und hat teilweise das Video auf dem First Screen gestoppt, um den Second Screen zu betrachten: „Also ich bin nicht so multitaskingfähig, ich musste wirklich einmal Stopp machen und mich hier so ein bisschen einfinden.“ (B3-Interview, Abschnitt 45). B2 und B3 bevorzugten die sequenzielle Nutzung der Geräte.

Im Gegensatz dazu lag der Fokus von B6 auf dem Second Screen mit den ergänzenden Informationen. Die parallele Nutzung beider Geräte führte zu keinen Problemen: „Ja, das war jetzt irgendwie schon so, so ein bisschen Fokus halt auf diese Nebeninformationen. Also auf die Informationen halt was, was die App halt zu dieser Doku halt ausgespuckt hat, aber ich, also ich persönlich konnte halt beidem halt folgen, ne.“ (B6-Interview, Abschnitt 34)

Der Fokus der Befragten ändert sich im Laufe der Anwendung und Gewöhnung an das System, wie B3, B4 und B6 erzählten. Nach einer Eingewöhnungsphase, in der das System kennengelernt wurde und der Second Screen vermehrter Aufmerksamkeit bedurfte, konnte sich B4 dem First Screen widmen: „Also am Anfang, ich denke, man muss sich an solche Sachen

dann gewöhnen. Am Anfang dachte ich so, wie funktioniert das jetzt? Wie schnell liefert er jetzt? Welche Fakten und wie viele? Wie viele Infos liefert der? Aber man hat dann schnell gemerkt, dass es nicht zu viel ist. So konnte man der Doku dann noch folgen und ist dann nicht abgelenkt dadurch. Und ich habe mich dann daran gewöhnt. Am Anfang fand ich es etwas verwirrend. Aber dann merkt man, dass das halt eine Hilfe sein kann, ne. Wenn man eine, eine bestimmte Erklärung haben will zu einem Begriff. Oder ein Fremdwort. Dann kann man da kurz nachschauen und muss dem nicht die ganze Zeit folgen, was auf dem Schirm, dann noch auf dem zweiten Gerät ist. Angeboten wird.“ (B4-Interview, Abschnitt 31). Auch B3 sagt, dass sich der Fokus geändert hat und dem First Screen mehr gefolgt wurde, als dem Second Screen: „Ich glaub, ich hab mich dann mehr nochmal tatsächlich dem Video gewidmet, als den Zusatzinfos.“ (B3-Interview, Abschnitt 49).

Der fehlende Hinweis, dass neue Informationen zur Verfügung stehen und die damit verknüpfte Erwartungshaltung, dass neue Informationen erscheinen werden, hat B1 vom Video abgelenkt: „Also ich hab immer das Video angeguckt und dann immer, dann hab ich’s immer nicht gemerkt, dass da was Neues kommt, hab ich dann schon immer auf diesen Screen geschaut und geguckt so: Was ist das denn jetzt, was da sozusagen noch für Informationen kommen? Ja, was mich zum Teil, glaub ich, schon auch mal raus rausgenommen hat aus dem Video. Also so, ne, vom, vom dran denken ja. Also, also, das mich die ja abgelenkt hat auch, auch so.“ (B1-Interview, Abschnitt 46).

Informationen

Die automatisch generierten Informationen sind der elementare Baustein des entwickelten Systems. Die generierten Informationen sollen die Anwender unterstützen und ihnen einen Mehrwert durch sinnvolle Informationen bieten. Die Informationen nahmen in den Interviews den größten Raum ein. Über die Hälfte aller Kodierungen des Kategoriensystems fallen in diese Hauptkategorie. Bevor die befragten Experten die Software sehen und anwenden konnten, wurden ihre allgemeinen Erwartungen an die Informationen, die dieses System anbieten soll, erfragt. Die Befragten erhielten zuvor eine kurze Einleitung, um den Zweck der Software und den Gesprächsgegenstand zu erläutern. Es wurden sechs Eigenschaften genannt, die die Informationen erfüllen sollten. Die hier aufgeführten Punkte wurden jeweils von nur einem Experten benannt.

Es wird erwartet, dass die ergänzenden Informationen kurz und prägnant sind: „Ich erwarte Informationen, die halt irgendwie auf Fragen, halt, kurz und prägnant, halt, antworten.“ (B6-Interview, Abschnitt 15). Dazu passend soll es keine überflüssigen Informationen geben. Die Informationen sollen in dem Moment zur Verfügung stehen, in dem die Anwender diese abrufen möchten: „[...] dass ich immer nur in dem Moment auch, wo ich vielleicht Informationen haben möchte, die auch irgendwie mir nur nehmen kann und nicht die ganze Zeit unnötige Informationen irgendwie auch wird.“ (B1-Interview, Abschnitt 13). Die Entscheidung, welche Informationen sinnvoll und welche überflüssig sind, ist individuell

unterschiedlich. Allerdings kann es Kriterien geben, die einen Rahmen für diese Entscheidung setzen. Im späteren Verlauf der Auswertung wird auf die Qualität der Informationen näher eingegangen.

Im Zusammenhang mit der erwarteten Art der Darstellung der Informationen wurde der Wunsch geäußert, dass es für die angebotenen Informationen mehrere verfügbare Quellen geben soll. Es sollen kurze Teaser der Informationen angeboten werden, aus denen die Anwender die gewünschte Informationsquelle auswählen können, um so die ausführlichen Informationen zu erhalten: *„Ich mein, Google macht das ja beispielsweise halt auch schon so, dass wenn du halt irgendwas suchst, dass sie dir halt einfach so kurz irgendwie so halt Suchergebnisse, 4-5 Antworten halt anzeigt, wo du dann halt einfach relativ schnell nachvollziehen kannst was, sozusagen, die Ergebnisse zu dieser Frage halt sind. Und die sind ja meistens halt mehr Auszüge. Sind meistens ja irgendwie nicht länger als 1-2 Sätze lang und dann kriegt man halt schon, sozusagen, ‘ne Richtung, in was das halt gehen könnte und kann, hat dann sozusagen halt die Möglichkeit, wenn alles halt weiter interessiert, von diesem Link aus dann den kompletten Content-Inhalt einzusehen.“* (B6-Interview, Abschnitt 21).

Zusätzlich zu den bisher geäußerten Erwartungen an die Informationen, wurde der Wunsch geäußert, dass mit den Informationen verschiedene Meinungen und Perspektiven auf ein Thema abgebildet werden: *„[...] dass man eben Meinungsvielfalt abbildet zu bestimmten Themen. Das finde ich interessant. Sowas wie einen historischen Überblick. Dann bestimmte Themen haben ja eben auch, dann sozusagen über, also einen Ursprung schon in der Geschichte. Wissenschaftliche Berichte oder Ausblicke, was sagt ganz aktuell die Wissenschaft dazu? Wie wird das Thema diskutiert?“* (B3-Interview, Abschnitt 23). Nicht nur die Abbildung verschiedener Meinungen und Perspektiven ist von Interesse für die Lehrerin B3. Die Auswahl und Filterung nach einzelnen Perspektiven könnte für den Unterricht nützlich sein. So könnte es mehrere Schülergruppen geben, die das Video aus unterschiedlichen Perspektiven betrachten: *„Oder als Lehrer würde ich mir bei Dokumentationen vielleicht ja auch nochmal einen anderen Fokus wünschen, [...] dann würde ich natürlich nochmal vielleicht fachlich tatsächlich, in Anführungsstrichen vielleicht, als Lehrer schon eine Vorab-Auswahl wählen und sagen so, da guckt sich die eine Schülergruppe den Fokus mit Zusatzinformationen aus Perspektive der Bewohner vor Ort an, eine andere Gruppe kriegt sozusagen Zusatzinformationen schon mal angeboten zu einer Perspektive von einem ausländischen Unternehmen, was sozusagen investieren möchte. Die Regierung könnte dann nochmal eine andere Perspektive darstellen, also dass man da sozusagen vielleicht auch die Möglichkeit hat, diesen Dokumentationsfilm für unterschiedliche Schülergruppen mit einem unterschiedlichen Zusatzinfo-Fokus auszustatten. Dass man vielleicht schonmal vorab Einstellungen machen kann, jetzt nicht die Schüler mit allen Informationen zu bombardieren, sondern schon ‘ne Auswahl zu wählen, können ja auch die Schüler einstellen: Ich möchte jetzt aus der Perspektive der Bewohner Fakten erfahren, oder über Proteste erfahren und eben, wie gesagt, über ein international agierendes Wirtschaftsunternehmen, die Regierung vor Ort usw., also, oder*

Umweltschützer. Also um da eben möglichst multiperspektivisch die Informationen vielleicht schon auch abzuwickeln.“ (B3-Interview, Abschnitt 29).

Neben den verschiedenen Meinungen und Perspektiven werden auch Fachinformationen erwartet. Beispielsweise können diese für Gesprächsrunden verwendet werden, bei denen unterschiedliche Positionen bezogen werden und positionsunabhängige Informationen zum thematischen Hintergrund benötigt werden: *„[...] natürlich auch Fachinformationen, denn das geht häufig eben auch nochmal verloren in solchen Diskussionsrunden. Dass man sich auch noch mal genau anschauen kann, über welchen Inhalt rede ich eigentlich gerade. Und sich damit nochmal auseinandersetzt, oder auseinandersetzen kann, wenn man merkt irgendwie: Ups, da fehlen doch noch weitere Informationen.“ (B3-Interview, Abschnitt 15).*

Die Themen, die in den Videos besprochen werden, sollen mit allgemeinen Eckdaten, wie bei Ländern die Einwohnerzahl und die Währung, angereichert werden: *„Wenn es jetzt eine Dokumentation über Island ist zum Beispiel, dass man sagen kann: Ok, Mensch, was haben die für Währungen, was haben sie vielleicht für 'ne Einwohnerzahl? Und dass man so ein bisschen Informationen einfach zu dem Thema bekommt.“ (B2-Interview, Abschnitt 17)*

Informationsarten

Die Bandbreite an Information, die Themen anreichern könnten, ist nicht zu überblicken. Ebenso gibt es eine Vielzahl an Informationsarten, die eine Bereicherung für die Anwender sein können. Im Rahmen dieser Ausarbeitung werden mehrere Arten von Informationen von dem entwickelten System generiert. Ein Beispiel sind die Ortsinformationen, die zusätzlich mit Kartenmaterial dargestellt werden. Vor der Anwendung der Software haben die befragten Experten ihre Erwartungen an die Informationsarten geäußert. Nach der Nutzung des Systems wurden ebenfalls Informationsarten benannt. Die Erwartungshaltung vor und nach der Anwendung der Software hat sich verändert.

Die verschiedenen Hintergründe der Experten spiegelten sich im Interview in der Vorstellung der Art und Themen der Fernsehdokumentationen wider, zu denen Überlegungen angestellt wurden. Drei Experten stellten fest, dass es abhängig von der Art und Thematik der Dokumentation ist, welche Informationsarten gewünscht werden: *„Dementsprechend angepasst, glaube ich, an die Art der Dokumentation, würde ich mir, glaub ich, unterschiedliche Informationen wünschen.“ (B3-Interview, Abschnitt 23).* Passend zu dieser Feststellung war die Erwartungshaltung der Experten an die Informationsarten unterschiedlich. Es wurde keine Informationsart von allen Experten benannt. Die drei Informationsarten Fakten, Historisches und Personeninformationen wurden jeweils von der Hälfte der Experten vor dem Betrachten der Videos benannt. Viele Informationsarten nur von jeweils einem Experten.

Vielen Experten war es wichtig, dass mit den Informationen Daten und Fakten bereitgestellt werden: *„Daten und Fakten, die gerade aktuell sind vielleicht. Oder die mal aktuell waren.“ (B4-Interview, Abschnitt 17).* Auch nach der Anwendung des Systems werden Fakten erwartet.

B4 hat hierzu ein Beispiel benannt. Das entwickelte System stellte die passenden Fakten und Informationen zu Computertomographen zur Verfügung: „[...] zum Beispiel ich hab auch über Tschernobyl irgendwas gesehen im Fernsehen und hab dann ja tatsächlich nebenbei recherchiert. Und da hab ich mir auch selbst mal angeguckt, wie funktioniert so ein Atomkraftwerk überhaupt noch mal? Und, und das wäre jetzt hier zum Beispiel auch richtig gut gewesen, wäre jetzt so als Thema, weil dann wäre da bestimmt auch gekommen: Das ist jetzt irgendwie Kernspaltung oder Kernfusion oder so. Und hier hat man dann gesehen, dass hier dann auch über den Computertomographen, als das vorkam, da auch was dann so erklärt wurde. Sowas hätte ich dann halt auch gebraucht, ich jetzt als Nichttechnischer oder nicht physikalischer hochgebildeter Mensch und also das kam dann tatsächlich auch.“ (B4-Interview, Abschnitt 53).

Die geschichtlichen Hintergründe eines Themas wecken das Interesse der Experten: „was mich dann interessieren würde, also was ich dann eher recherchiere sind so geschichtliche Hintergründe.“ (B4-Interview, Abschnitt 17). Dokumentationen entstehen eingebettet in einen historischen-Kontext, der die thematische Aufbereitung und den Wissensstand der Produzenten beeinflusst. Daher ist es aufschlussreich, in welchem Jahr bzw. in welchem Kontext die Dokumentation entstanden ist. Darüber hinaus ist es bei Dokumentationen zu einem historischen Thema wissenswert, wie die gesellschaftliche Debatte zu dem Thema war: „[...] wenn du jetzt 'ne Tschernobyl-Dokumentation zum Beispiel anguckst, da ist natürlich interessant: Okay aus welchem Jahr ist die überhaupt? Wie war damals die Debatte? Auch die Debatte zum Thema im Atomkraft etc.. Also auch um, um vielleicht sowas auch mehr zu verstehen, dass ja auch Dokumentationen in so historischen Kontexten entstehen und natürlich sich Blickwinkel da drin auch verändern können.“ (B5-Interview, Abschnitt 23). Nach der Anwendungsphase bestätigte B4 im Interview, dass die gewünschten geschichtlichen Hintergründe in der Anwendungsphase vom System bereitgestellt wurden: „[...] also hier, wie eben schon gesagt, über China kam ja auch. Ja, wurde auch gesagt China, China gehört zu den ältesten Zivilisationen etc etc also da kam ja auch geschichtlicher Hintergrund jetzt über China.“ (B4-Interview, Abschnitt 53).

Bei Filmen und Dokumentationen wirken viele Menschen mit, sei es vor oder hinter der Kamera. Protagonisten und andere Personen vor der Kamera sind für den Konsumenten einer Fernsehdokumentation gut sichtbar. Streaming-Plattformen bieten bereits die Möglichkeit, bei Filmen die Rollennamen und Klarnamen der abgebildeten Schauspieler zu sehen. Auf diese Möglichkeit spielt der Experte B2 an und möchte gerne allgemeine Informationen zu den abgebildeten Personen erhalten: „Dass du einfach stoppen kannst und dann siehst du halt welche Schauspieler spielen da mit? Wo spielen diese Schauspieler gerade mit? In welchen Filmen spielen die sonst noch mit? Was haben die für Sachen gewonnen?“ (B2-Interview, Abschnitt 19). Bei der Produktion von Filmen und Fernsehdokumentationen wirken viele weitere Menschen mit, die für den Zuschauer nicht ersichtlich sind. Die Expertin B5 schlägt vor, dass man diese mit Hilfe der Informationen für die Zuschauer sichtbar machen könnte und

Informationen zur Produktion der Fernsehdokumentation ebenfalls sinnvoll sind: *„Ich denke, es wäre auch sinnvoll, so Informationen zu: Wie wurde meinetwegen der Film oder die Serie produziert? Oder was steckt da so hinter? Dass könnte so ‘ne Sichtbarmachung von sag ich mal Personen, die nicht nur vor der Kamera sind, ermöglichen.“* (B5-Interview, Abschnitt 21). Das Interesse an Informationen zu den Mitwirkenden einer Dokumentation oder eines Filmes ist jedoch individuell. So wären Informationen zu den Personen, die an der Produktion beteiligt waren, für B5 unbedeutend, allerdings gibt es Menschen, die sich dafür interessieren: *„Ich muss aber dazu sagen, mich würde es nicht interessieren. Aber ich denke. Es gibt ja so Menschen, die dann immer wissen, wer der Regisseur ist und was da so an Hintergrundsachen liefen etc. oder wie Filme produziert oder finanziert wurden.“* (B5-Interview, Abschnitt 21). Im späteren Verlauf wurde das Interesse an Informationen zu den Umständen der Produktion der Fernsehdokumentation von der Expertin B5 wiederholt: *„Aber, also ich hab mich schon gefragt: Aha, wie lange waren die denn jetzt da unterwegs? Und sind da jetzt noch nach Dubai geflogen. Also das sind so Sachen, die ich mich dann immer frag, wie lange hat es jetzt wohl gedauert, diese Doku zu drehen? Und in welchem Jahr sind wir jetzt eigentlich?“* (B5-Interview, Abschnitt 60).

Es gibt die mitwirkenden und abgebildeten Personen und Protagonisten. Darüber hinaus können weitere Personen Gegenstand einer Fernsehdokumentation sein. Beispielweise werden in historischen Fernsehdokumentationen oft Personen thematisiert, mit denen nicht unmittelbar interagiert werden kann. So wird Napoleon Bonaparte häufig im Zusammenhang mit der Französischen Revolution benannt. Auch zu solchen Personen werden Informationen erwartet: *„Personenhistorien, die ich dann. Wo ich dann wieder mal nachschauen kann, wer war das denn genau? Wann hat diese Person gelebt?“* (B1-Interview, Abschnitt 15).

Während der Anwendung des entwickelten Systems hat B5 die Informationen zu einer Person registriert, die in dem gezeigten Video thematisiert wurde. Die Zuordnung ist für B5 von Interesse, die erwähnte Person des Beispiels wurde jedoch vom System falsch zugeordnet. So wurden Informationen einer Person mit gleichem oder ähnlichen Namen als Hintergrundinformationen angeboten, die nicht mit der gesuchten Person übereinstimmen: *„Also hier ist ja auch über die Person, die tauchen dann ja auch auf. Einmal der, die. War das der mit den kleinen Mini-Transportdrohnen sozusagen. Sowas finde ich grundsätzlich auch interessant, aber hier taucht dann tatsächlich jemand auf. Oder ist das vertauscht? [...] Auf jeden Fall ist hier ein Name, das ist jemand, der 1914 in China geboren war und das ist natürlich, wenn prominente Leute oder so diesen Namen führen, dann wird natürlich erst mal diese Person verknüpft und nicht irgendjemand, der jetzt seit 2,3 Jahren da irgendein neues Unternehmen hat, ne. Deshalb.“* (B3-Interview, Abschnitt 63).

Eine weitere Informationsart, die erwünscht ist, sind geographische Informationen. Straßen und Orte, die im Video vorkommen, könnten in einem Kartensystem dargestellt werden: *„Interessant wäre da auch noch so, also falls es irgendwie ‘ne Location ist, also so mit irgendwie, der ist gerade hier, die laufen gerade irgendwie hier diese Straße runter übrigens.“*

Guck's dir auf Google Maps an, fertig.“ (B1-Interview, Abschnitt 15). In dem Softwaresystem wurden die thematisierten Städte und Länder auf einem Kartensystem dargestellt. Diese Informationen waren für die Lehrerin B4 von Interesse: *„Also ich für mich persönlich fand ich interessant, wo sind diese Orte in China?“* (B4-Interview, Abschnitt 39).

Darüber hinaus wurden Vorschläge für weitere Fernsehdokumentationen erwartet, die ähnliche Themen besprechen, wie das aktuelle Video: *„oder ähnliche Dokumentationen vielleicht zu dem gleichen Thema.“* (B2-Interview, Abschnitt 19). Allgemein wurden Hintergründe zum aktuellen Video erwartet: *„Hintergründe. [...] Also, dass ich dann halt schnell, dann bestimmte Fragen, ja bestimmte Themen, abrufen kann dann. Oder bestimmte Hintergründe.“* (B4-Interview, Abschnitt 15). Nach der Anwendung der Software wurde die Idee von Inhaltsempfehlungen wiederholt. Die sozialen Medien bieten heutzutage einige Funktionalitäten, wie unter anderem die Empfehlung weiterer zugeschnittener Inhalte. Diesem Gedanken folgend gibt B2 den Vorschlag, dass Informationsinhalte und Videos im gleichen Themenfeld vorgeschlagen werden: *„Ich stell mir das halt auch so ein bisschen vor, wie man heutzutage die ganzen Social Media Dinger nutzt, wenn du jetzt meinetwegen auf Instagram oder sonst was bist. Du klickst auf einen rauf und kriegst aber unten schon gleich so eine Welle an: Okay, das könnte dich auch so interessieren irgendwie. Dass das schon weiß, in welchen Themenumfeld bist du.“* (B2-Interview, Abschnitt 46)

Weitere Informationsarten wurden vor der Anwendung der Software erwartet, im Nachfolgenden jedoch nicht weiter besprochen. Ob es daran liegt, dass diese zuvor erwarteten Informationen nicht weiter von Interesse waren, die erwünschten Informationsarten abhängig von der Thematik des Videos und des Formats sind, die Erwartungen erfüllt wurden oder die vorigen Äußerungen im Verlauf des Videos nicht weiter bedacht wurden, ist nicht zu klären.

Im Kontext von Reisedokumentationen wurden Informationen zu Ausflugszielen in der Umgebung der besprochenen Gegend erwartet: *„[...] wenn ich mir jetzt einfach eine Dokumentation über Schweden angucke, als Reisedokumentation, dann würde ich mir vielleicht noch zusätzlich, wenn ich interessiert bin, Ausflugsmöglichkeiten oder sowas vorstellen können. Wo sind tolle Strände?“* (B3-Interview, Abschnitt 23).

Die wissenschaftliche Mitarbeiterin hat erwartet, dass aktuelle Nachrichten auf dem Second Screen parallel zum Video bereitgestellt werden. Die Nachrichten sollen das Gewissen und den Intellekt der Anwender und Konsumenten unterstützen, wenn diese einfachere Inhalte betrachten: *„Also sei es irgendwie Nachrichten etc., oder ansonsten Inhalte, die ich mir gerne angucke. Ich glaube so, das hat so eine Ebene von schlechtes Gewissen bereinigen, wenn ich mir jetzt, sag ich mal, so RTL-Niveau angucke, dann ist es sinnvoller für mich, wenn so nebenbei jetzt mal aktuelle Nachrichten eingeblendet werden würden, dann könnte ich denken: Okay, schaue ich heute Sendung XY und nebenbei tut der Second Screen, was für meinen emotionalen Intellekt.“* (B5-Interview, Abschnitt 17). Die Nachrichten haben keinen thematischen Bezug mit dem gezeigten Video und sind für das nebenläufige bzw. passive

Betrachten von Videos auf dem First Screen gedacht: „*Dass es mir, wenn ich, sag ich mal, auf meinem First Screen, so Berieselung angucke zu dem auch Informationen bieten kann, die gar nicht Bezug zur Sendung haben*“ (B5-Interview, Abschnitt 17).

Einige weitere Erwartungen an die Art der bereitgestellten Informationen wurden zuvor nicht benannt, im weiteren Verlauf des Interviews jedoch thematisiert. So wünscht sich der It-Administrator allgemeinere Informationen: „*Eher größere Informationen.*“ (B1-Interview, Abschnitt 64). In der Erfahrung von B1 wurden eher einzelne Aspekte erläutert. Sinnvoll wäre dagegen auch, dass komplexere Zusammenhänge erläutert werden. Dies könnte beispielsweise durch eine journalistische Einordnung eines Themenkomplexes geschehen: „*ein bisschen so größer einordnen so 'ne Sachen, so ein bisschen was Journalistisches. So 'n journalistischer Rahmen am ehesten noch.*“ (B1-Interview, Abschnitt 64).

Darstellung der Informationen

Der Aspekt der Darstellung nahm in den Interviews großen Raum ein und wurde von den Experten am häufigsten besprochen. Beinahe die Hälfte der Äußerungen, die sich um die Informationen drehen, wurden zu diesem Aspekt getätigt. In diesem Abschnitt werden die Aussagen und Bewertungen zur Darstellung der Informationen besprochen.

Eine Erwartungshaltung der Experten war, dass die bereitgestellten Informationen den Anwendern stressfrei zu dem Zeitpunkt, zu dem sie diese abrufen wollen, zur Verfügung stehen. In den Momenten, in denen die Nutzer diese nicht beachten wollen, sollen diese nicht stören und ablenken: „*Ich erwarte. Einmal dass es, dass es nicht zu aufdringlich ist, als dass es mich irgendwie von dieser, von der eigentlichen Doku ablenkt oder ich irgendwie dann mich nicht darauf konzentrieren kann, oder es mich stresst. Das wäre. Sondern, dass ich immer nur in dem Moment auch, wo ich vielleicht Informationen haben möchte, die auch irgendwie mir nur nehmen kann und nicht die ganze Zeit unnötige Informationen irgendwie auch wird.*“ (B1-Interview, Abschnitt 13). Auch B5 möchte nicht, dass die Informationen durchgehend angezeigt werden und wünscht sich einen Knopf, mit dem diese angezeigt oder nicht angezeigt werden können: „*Dass es wie so 'n Knopf gibt, den ich andrücken kann. Also, dass sie nicht einfach die ganze Zeit da sind.*“ (B5-Interview, Abschnitt 25).

Für die spätere Betrachtung möchte der Experte B1 die Informationen anklicken können, um spannende und gute Informationen markieren zu können. Im Nachhinein sollen die markierten Informationen wie eine Historie verwendet werden können und für die spätere ausführliche Betrachtung gespeichert werden: „*Vielleicht merke ich dann, wenn ich kurz draufklicke, dass mich, das ist ein super spannendes Thema ist und ich mich im Nachhinein vielleicht auch noch mehr für interessieren möchte. Vielleicht 'ne Sache irgendwie Informationen zu merken, oder sozusagen wie diese Historie einfach nachzuvollziehen, so dass ich dann danach irgendwie denke, mit irgendwie: Ja, dieses Ereignis hatte ich noch gar nicht von gehört, das würde ich mir nochmal - also vielleicht nicht während des Films, weil ich denke: Okay, wenn ich jetzt*

diesen ganzen Wikipedia Artikel lese, dann kommt, krieg ich den Rest nicht mit so. Aber möchte ich mir irgendwann auf jeden Fall nochmal anschauen. Sowas. Merkerinnerungs-Funktion los.“ (B1-Interview, Abschnitt 33).

Im Allgemeinen soll die Anwendung benutzerfreundlich sein, intuitiv zu verwenden und flüssig laufen: *„Ich erwarte, dass es halt benutzerfreundliche Anwendungen halt hat. Also es sollte einfach aufgebaut sein, es sollte selbsterklärend sein und es muss flüssig laufen.“* (B6-Interview, Abschnitt 13). Um diese Punkte zu erfüllen und eine reibungslose Verwendung zu ermöglichen, müssen diese Eigenschaften für die Darstellung der Informationen gelten.

Die Experten haben sich Gedanken darüber gemacht, auf welche Art und Weise die Informationen den Anwendern dargeboten werden können. Hierbei wurden mehrere Punkte formuliert, bevor die Experten das System verwenden konnten.

Die Überlegung, wie die gewünschten Informationen dem Anwender bereitgestellt werden, führte zu der Vorstellung zweier Experten, dass die Informationen dem Nutzer über Suchmasken verfügbar gemacht werden könnten: *„Wie ne, wie Internetseiten, Suchmaschinen aufgebaut sind. Vielleicht ein bisschen einfacher noch, dass man weniger eintippen muss, weil das, man braucht relativ lange.“* (B2-Interview, Abschnitt 21). Ein anderer Experte möchte die Informationen auswählen können, ohne dass eine Eingabe über eine Suchmaske o.ä. erfolgen muss: *„Am besten Fall sehr schnell, intuitiv eher was tippen, äh auswählen als irgendwie den Namen jetzt dann wieder eintragen zu müssen oder so.“* (B1-Interview, Abschnitt 29). Einer Suchmaske wäre B1 auch nicht abgeneigt.

Nach der Benutzung des Systems ist B4 von der Vorstellung einer Eingabemaske zurückgetreten und bewertet die Vorgehensweise des entwickelten Systems als besser. Beim entwickelten System erscheinen die Informationen ohne Eingabe chronologisch sortiert: *„Also am Anfang dachte ich ja, man muss selbst irgendwas ein, eingeben, aber dann, man muss es ja nicht. Von daher fand, ich fand das besser so.“* (B4 - Interview, Abschnitt 33). Aufgrund des Scrollens würde B2 weiterhin eine Suchfunktion haben wollen: *„für mich wäre eine Suchfunktion wichtig, weil mir das einfach auch jetzt zu lange dauert von oben nach unten rum zu scrollen, zu gucken, ok, was habe ich?“* (B2-Interview, Abschnitt 68). Die Lehrerin B4 merkt ebenfalls an, dass die passive Bereitstellung der Informationen gut ist, da die aktive Suche nach den Informationen die Anwender zusätzlich ablenken würde: *„[...] ich würde das, glaube ich, so gut finden, weil man dann, man wäre abgelenkter, wenn man eine Doku verfolgen will, man wäre abgelenkter, wenn man selbst aktiv werden müsste. Dann würde ich dem, glaub ich, nicht so folgen können.“* (B4-Interview, Abschnitt 35).

Fast alle Experten möchten die Möglichkeit haben, die Informationen nach ihren Wünschen zu filtern oder für die spätere Verwendung zu markieren. So möchte B1 Informationen nach eigenem Interesse auswählen: *„Und irgendwie, dass ich vielleicht dann auswählen kann, welche Informationen mich davon interessieren.“* (B1-Interview, Abschnitt 23). Und nach der Filterung möchte B5 einen Teaser der gefilterten Information erhalten: *„Und dass ich dann*

vielleicht so'n paar Kategorien habe, um auszuwählen wie Fun Facts, historischer Hintergrund, zeitgenössischer Hintergrund, weiß, weiß ich auch nicht, irgendwie sowas. Und dann kann ich die anklicken. Und dann kommt so die erste nur und dann will ich so langsam angefütert werden [...]" (B5-Interview, Abschnitt 25).

Zwei Experten möchten, dass die Darstellung der Informationen nicht zu visuell geschieht, da diese eine größere Ablenkung vom First Screen befürchten: *„Auf jeden Fall nicht zu visuell, weil man ja parallel, es ist ja nur der Second Screen, du willst ja noch mitbekommen, was auf dem anderen läuft.“ (B5-Interview, Abschnitt 25).*

Abhängig von der Art der Informationen, möchte B6 ein einfaches Aufführen der Fakten, ohne gesonderte visuelle Elemente: *„Wenn es jetzt aber darum geht, was weiß ich. Was sind die größten Städte Deutschlands? Dann reicht mir sozusagen halt 'ne kurze tabellarische Auflistung mit den Zahlen.“ (B6-Interview, Abschnitt 17).* Im Gegensatz zur einfachen Aufführung der Fakten, möchte B6 bei Personeninformationen ein Bild der Person erhalten: *„Also, wenn ich jetzt zum Beispiel über eine Person halt irgendwie Informationen halt einhole, dann würde ich es gut finden, wenn er, wenn ich, sozusagen, halt visualisiert zu der Person, halt, auch ein Bild angezeigt bekomme.“ (B6-Interview, Abschnitt 17).* Wie bei den Ergebnisanzeigen von Suchmaschinen, sollen die dargestellten Informationen nur wenige Sätze lang sein, bzw. nur kurze Teaser sein, die bei Interesse eine Verlinkung zu der ausführlichen Information zur Verfügung stellen: *„Und die sind ja meistens halt mehr Auszüge. Sind meistens ja irgendwie nicht länger als 1-2 Sätze lang und dann kriegt man halt schon, sozusagen, 'ne Richtung, in was das halt gehen könnte und kann, hat dann sozusagen halt die Möglichkeit, wenn alles halt weiter interessiert, von diesem Link aus dann den kompletten Content-Inhalt einzusehen.“ (B6-Interview, Abschnitt 21).*

Mit dem entwickelten System wurden viele Informationen mit visuellen Elementen, wie Bildern und interaktiven Karten, dargeboten. Nach der Anwendungsphase wurde der Wunsch nach weiteren visuellen Teilen geäußert. So findet B1 interaktive Informationen und Karten besser als Informationen, die nur aus Texten bestehen: *„[...] dann doch gemerkt hat, dass ich diese visuellen Sachen vielleicht doch auch irgendwie, irgendwie doch besser finde, also visuelle Karte oder irgendwie sowas interaktiv darzustellen, als dieser reine Text.“ (B1-Interview, Abschnitt 66).*

Auch B2 möchte viele visuelle Elemente und zugleich bei der Übersicht nur Überschriften dargestellt bekommen. Die Informationen sollen in der Kacheldarstellung, wie bei mobilen Endgeräten häufig genutzt, dargestellt werden. Die einzelnen Kacheln geben die Schlagwörter oder Überschriften der Informationen wieder. Über die Auswahl einer Kachel gelangen die Anwender zu der gesamten Information: *„Ich bin immer ein ganz großer Freund von erstmal visuellen Punkten und nur Überschriften und wenn ich dann haben möchte, dann hätte, mir hätte hier so 'ne Kachel gereicht, quasi. Wo ich diese Überschriften habe und dann sage ich alles klar zu China *Klick* und dann habe ich ein Pop-up, wo ich die Sachen zu China haben*

will.“ (B2-Interview, Abschnitt 36). Die Kacheldarstellung erlaubt mehr Übersichtlichkeit, führt zur Einsparung von Text in der Hauptansicht des Second Screens und reduziert das Scrollen stark: „*Und Kacheln, weil es einfach so typisch ist und verwendet wird halt auch. Ob's ein Windows ist oder, oder Apple. Ich kann irgendwo auf so 'n Button draufklicken und, und hab nicht so viel Text.*“ (B2-Interview, Abschnitt 48).

Diese Idee lässt sich gut mit der Aussage von B1 kombinieren, der sich mehr Interaktionen wünscht, um die Informationen zu erhalten. Zur Übersichtlichkeit sollen nur die Schlagwörter, gemeinsam mit einem Icon, dargestellt werden. Die Icons spiegeln die Kategorien, denen die Informationen zugeordnet werden, wider. Durch die Interaktion mit den Informationen werden die eigentlichen Informationen zur Verfügung gestellt: „*Also wenn gerade meine erste Idee, dass sozusagen erstmal die Sachen, einfach nur die Begriffe aufploppen und ich dann kurz. Und mit einem Icon davor vielleicht, welche Kategorie das ist, dann irgendwie 'nen, 'ne Weltkugel und dann irgendwie China und dann kann ich da. Ja also, das ist irgendwie so ein bisschen, dass ich mehr, vielleicht ein bisschen mehr interagieren muss, wenn ich, wenn ich Informationen haben will.*“ (B1-Interview, Abschnitt 74).

Die von B2 vorgeschlagenen Kacheln, die aus den Schlagworten bestehen, sollen auch hierarchisierbar sein. Dies kann wie eine hierarchische Baumstruktur sein, bei der das erste Element, die erste Kachel, weitere Subelemente, weitere Kacheln, haben kann und man sich immer tiefer in die Hierarchie-Ebenen navigiert, bis man zu den Blättern gelangt, die die vollständigen Informationen repräsentieren: „*[...] ich sehe das wirklich als, als, als Start, diese Schlagworte, die dann hier rein treffen und das sind hier jetzt die Ober-Überschriften. Und ob es jetzt Stimmerkennung ist, dann würden, würde diese Kachel „Stimmerkennung“ kommen, so und dann klicke ich darauf und dann kommt da wahrscheinlich ok, geschichtlich, vielleicht oder, oder, oder welche Software kann ich nutzen? [...] Also dass Themen zu diesem Thema Spracherkennung kommen. [...] Und dann aber auch weitere Subthemen, wo wird das eingesetzt? Und dann klick ich, dann kommen ja andere Schlagworte.*“ (B2-Interview, Abschnitt 50).

Auch zu den Geräten, die für die automatisierte Informationsergänzung von Fernsehdokumentationen genutzt werden können, haben sich die Experten vor der Anwendungsphase geäußert. So hatte sich B3 Gedanken zur Nutzung eines Split Screens gemacht. Hierbei wäre der First Screen geteilt worden. In dem einen Bereich würde das Video laufen und in dem anderen die ergänzenden Informationen gezeigt werden: „*Also vielleicht über so 'ne Art Split Screen, fällt mir das nochmal ein. Dass man sozusagen aufm Vollbildschirm dann die Möglichkeit hätte, den zu teilen und zu sagen so, da werden mir InfoWolken oder sowas angeboten.*“ (B3-Interview, Abschnitt 27). B3 sieht aber auch die Möglichkeit der Nutzung eines Second Screens, auf dem die Informationen dargestellt sind: „*Aber ich hab so, mit so einem Split Screen sonst auch nicht irgendwie gearbeitet, also von daher kann es auch genauso praktisch sein, wenn ich dann den Blick auf mein zweites Gerät senke und da dann eben sehe: Das sind die Infos, die mir angeboten werden*“ (B3-Interview,

Abschnitt 27). Für B6 ist es von der Information abhängig, ob ein Bildschirm oder zwei Geräte verwendet werden: *„Ja, auch das finde ich, ist so ein bisschen von der Situation, also von der Information halt abhängig, die man halt einholen möchte.“* (B6-Interview, Abschnitt 23). Bei sehr kurzen Informationen, wie Schlagwörtern, stellt sich B6 die Darstellung auf dem First Screen vor, bei längeren Informationen, die Nutzung eines Zweitgeräts.

Nach der Anwendungsphase bewertet B3 die Nutzung eines Tablets als Second Screen als gut. Die Größe eines Tablets ist gut und auch besser als die Größe eines Smartphones. Ein Smartphone als Second Screen ist für B3 ebenfalls denkbar: *„Nö, das finde ich schon gut hier mit dem Tablet. Also das ist auch genau die richtige Größe. Tablet ist händelbar, man hat einen relativ großen Überblick hier. [...] Also das finde ich schon, schon sehr gut. Und auf dem Smartphone kann ich mir das auch vorstellen, aber ich glaube, wenn man sich tatsächlich an einen Fernseher setzt oder so dann ist, glaub ich, so ein Tablet schon die richtige Größe auch.“* (B3-Interview, Abschnitt 53).

Die Experten sind sich nicht einig, ob die dargestellten Informationen zu umfangreich sind oder die Länge gut ist. Die Textlänge der einzelnen Informationen wurde von B3 und B4 als gut bewertet: *„so einen kurzen Infotext finde ich gut, mit der Option dann zum Weiterklicken, falls es einen interessiert.“* (B3-Interview, Abschnitt 57). Die anderen Experten hätten kürzere Informations-Teaser besser gefunden. Da man durch die Interaktion weitere Informationen erhalten kann, reichen für den Experten B6 zwei bis drei Sätze für jeden Teaser: *„Teilweise ist mir das ein bisschen zu viel Text, ich würde vielleicht sogar sagen, dass 2-3 Sätze halt, vielleicht sogar maximal 2 Sätze, halt reichen und dass man dann einfach auf „mehr“ klicken kann und, um sich dann halt weiter zu informieren.“* (B6-Interview, Abschnitt 44). Die längeren Informations-Teaser haben B2 vom Video des First Screens abgelenkt: *„Jetzt im Moment bin ich abgelenkt. Jetzt lese, jetzt gucke ich mir das Video an und habe dann so viel Text auf.“* (B2-Interview, Abschnitt 76).

Einige Design-Elemente und Formatierungen der Informationen wurden kritisiert und Verbesserungsvorschläge eingebracht. Beispielsweise war die Formatierung der Überschriften der Informationstexte nicht immer stimmig: *„[...] von der „EH 184“ gibt es einen Beitrag und dann sticht halt visuell erstmal sozusagen das Wort „Geschichte“ halt so ein bisschen raus, weil das Teil des Beitrags ist, ne. Also, das ist so von der Formatierung her“* (B6-Interview, Abschnitt 28).

Die Darstellung der Informationen wird, von den bereits aufgeführten Punkten abgesehen, positiv bewertet. B3 findet gut, dass Informations-Teaser mit Bildern angereichert dargestellt werden und die Informationen erweiterbar sind: *„Ich hatte eben ja gedacht, dass mir oder bevor ich diese Anwendung so einmal teste, dass eben ausschließlich Stichworte geliefert werden, aber hier tauchen ja tatsächlich schon mal kurze Bilder auf, um sozusagen nochmal 'n visuellen Reiz anzusprechen und nicht nur Text, Text, Text, Text. Das finde ich eben sehr gut und so einen kurzen Infotext finde ich gut, mit der Option dann zum Weiterklicken, falls es*

einen interessiert. Genau, das finde ich gut.“ (B3-Interview, Abschnitt 57). Im Besonderen die Visualisierung der Informationen durch die Ergänzung mit Bildern wird von B3 als gut bewertet. Die Visualisierung führt dazu, dass man sich einen relativ schnellen Überblick verschaffen kann: „Ich finde es gut, dass da immer auch Bilder mit dabei sind. Überschriften, Schlagworte und knappe, kurze Infos oder Einleitungen, aber immer mit dem Bild dazu. Das man gleich sehen kann: Ah, dazu die Infos. Dass man da relativ schnell einen Überblick sich verschaffen kann.“ (B3-Interview, Abschnitt 51). Für B1 waren die interaktiven Landkarten, die es zu den geographischen Informationen gab, sinnvoll: „Sinnvoll waren für mich Landkarten.“ (B1-Interview, Abschnitt 58). Über die schlichte Aufmachung der Darstellung der Informationen äußerte sich B6 positiv. Es gab keine ablenkenden Elemente oder eine Reizüberflutung, die ablenken oder anderweitigen Einfluss auf die Anwender nehmen: „[...] was ich aber gut finde ist sozusagen halt die schlichte Aufmachung, lenkt, es gibt da keine Elemente, die halt irgendwie einen ablenkend sind, es gibt keine Werbung da drauf und nichts. So das ist, also es gibt keine weiteren Reizüberflutungen, die dich halt irgendwie noch anderweitig halt beeinflussen.“ (B6-Interview, Abschnitt 58).

Bewertung der Informationen

In den Interviews mit den Experten wurden die bereitgestellten Informationen bezüglich ihrer Quantität und Qualität bewertet. In der Bewertung von B6 gab es für im Verhältnis zur Länge des abgespielten Videos zu viele Informationen: „Teilweise für die 9 Minuten, die wir uns jetzt hier angeguckt haben, bisschen zu viel Informationen.“ (B6-Interview, Abschnitt 54). Für B1 und B5 war die Menge der Informationen okay: „In Ordnung. Sollten auf jeden Fall nicht mehr sein.“ (B5-Interview, Abschnitt 52). Persönlich kam die wissenschaftliche Mitarbeiterin B5 gut mit der Informationsmenge klar: „Und ich kam damit gut klar, weil ich eine unglaublich hohe Auf-, schnelle Auf-, wie heißt das, Auffassungsgabe habe und ganz schnell Texte verarbeiten kann.“ (B5-Interview, Abschnitt 34). B5 bedenkt jedoch auch Personen, die die Informationstexte nicht so schnell lesen und verarbeiten können und geht davon aus, dass es teilweise zu viele bereitgestellte Informationen sind: „So dass ich mir vorstellen könnte, dass es manchmal zu viel Informationen sind, so dass man gar nicht mehr hinkommt. Also, weil das sind alles so Informationen, die man quer liest, wenn man der Doku noch aufmerksam folgen will. Wenn du jetzt aber langsam liest und dann da gedanklich hängen bleibst, dann kriegst du es nicht mehr mit, was da war.“ (B5-Interview, Abschnitt 34). Die Experten B3 und B4 bewerten die Informationsmenge als gut „[...] die Menge an sich fand ich gut. Ja, nicht zu viel, nicht zu wenig.“ (B4-Interview, Abschnitt 45) und genau richtig: „[...] eigentlich finde ich das von der Menge, glaub ich genau richtig.“ (B3-Interview, Abschnitt 65).

Die Qualität der Informationen wurde durch die Experten unterschiedlich bewertet. Ob die Informationen automatisiert generiert oder diese von einer Redaktion ausgearbeitet werden sollen, ist bei den Experten umstritten. Nach der Einschätzung von B6 würde eine redaktionelle Bearbeitung bessere Ergebnisse liefern. Überflüssige Informationen würden vermieden werden

und die Informationen gezielter ausgerichtet sein: *„Ich glaube, wenn man das redaktionell bearbeiten würde, dann würden halt so [...] Beiträge, wie zum Beispiel dieses Wort Englisch, wo über die eigentliche Sprache gesprochen wird, würde da sozusagen gar nicht auftauchen, sondern dann würde sich ‘ne Reaktion halt irgendwie Gedanken machen: Was sind jetzt sozusagen die Keywords, die ich halt für diese Dokumentation halt brauche und dann würden sie das halt einfach selektieren, ne. Und dir dann halt einfach gezielt Informationen halt geben.“* (B6-Interview, Abschnitt 64). Bei der automatisierten Generierung sieht die Expertin B5 die Gefahr, dass die Algorithmen zur Generierung der Informationen die Informationen nicht objektiv und neutral generieren, sondern dass die Informationen derart erzeugt werden, dass die Inhalte der Filterblasen, in denen die Anwender sich befinden, weiter reproduziert wird. Aus diesem Grund befürwortet B5 die redaktionelle Erzeugung der ergänzenden Informationen. Die redaktionelle Bearbeitung kann hierbei eine Gatekeeper-Funktion einnehmen: *„Und dann sehe ich nämlich genau diese Gefahren, die ich angesprochen habe: Dass man in expliziten Filterblasen drinhängt, dass man ganz spezielles Wissen nur bekommt, etc.. Und wir sind ja im journalistischen Bereich und das heißt, ich gehe davon aus, dass journalistisches Wissen hat so ‘ne Gatekeeper-Funktion, nennt sich das. Das heißt auch vorzugenerieren, welches Wissen wird weitergegeben an die Menschen, die den Inhalt weitergeben. Und ich bin da keine Freundin von Algorithmen, sondern sehe es als großen Vorteil an, wenn eben diese Gatekeeper-Funktion auch noch aufrechterhalten wird. Und deswegen fände ich eine redaktionelle Überarbeitung besser.“* (B5-Interview, Abschnitt 66). Der Experte B2 sieht in der redaktionellen sowie der automatisierten Informationserzeugung Gutes. Nach B2 würde die redaktionelle Erstellung der Informationen dazu führen, dass die Informationsergänzungen besser auf die Fernsehdokumentationen und die Bedürfnisse der Anwender zugeschnitten sind: *„Aber redaktionell aufbereitet würde es wahrscheinlich dann wieder so sein, dass man einen Bezug zu dem herstellt, was man gerade gesehen hat, oder was man vermutet, was der Zuschauer beziehungsweise der Benutzer halt gerade erwartet.“* (B2-Interview, Abschnitt 62). Doch auch die automatisierte Generierung liefert gute Schlagwörter, mit denen die Informationsergänzungen verknüpft werden: *„So, wie ich mir das vorstelle halt, wär ‘s wahrscheinlich nötig, so wie ich mir das gerade vorstelle. Wobei auch, auch nicht, ne. Das kann auch automatisiert sein, weil diese, diese Begriffe, die hier hereingekommen sind, die machen ja Sinn.“* (B2-Interview, Abschnitt 62).

Redaktionen bringen eigenes Wissen, Einstellungen und Vorannahmen mit, die bewusst oder unbewusst die Auswahl und Erstellung der Informationen beeinflussen können. In der automatisierten Generierung sieht die Expertin B3 den Vorteil, dass diese die Informationen neutral und unabhängig generieren kann: *„Grundsätzlich nehme ich an, dass sowas automatisch Generiertes, Generiertes neutraler ist, von der Meinung her. Und jetzt eine Redaktion. Natürlich auch neutral aus neutraler Position das, die Informationen aufbereiten könnte. Aber könnte natürlich auch unter einem gewissen Fokus das Ganze aufbereitet sein. Also, das müsste dann sozusagen für mich transparent gemacht werden. Wer diese Infos dann bearbeitet hat und da gehe ich jetzt davon aus, dass es rein auf Schlagworten basiert und*

sozusagen unbearbeitet im Moment ist, ne.“ (B3-Interview, Abschnitt 77). Die Expertin B4 bewertet die redaktionelle Erstellung der Informationen zwar als besser, im Hinblick auf die Menge der Fernsehdokumentationen und Videos überwiegt jedoch der Vorteil der automatisierten Generierung der Informationsergänzungen: „Also von daher ist das, es muss schon ein Format geben, was automatisch das liefert. Sonst würde das einfach nicht so gut funktionieren in der heutigen Welt, ne. Also von daher. Ja redaktionell wäre das natürlich, wäre das natürlich besser auf-, gehe ich von aus, dass das besser aufbereitet ist. Aber es ist schon okay so, weil. Ja, anders kann ich mir das schlecht vorstellen, also über den ganzen Mengen an Clips oder Dokus, oder Ausschnitten aus irgendwelchen Nachrichten oder so, oder da muss das automatisch passieren.“ (B4-Interview, Abschnitt 65).

Ein Großteil der Experten merkt an, dass die Bewertung der Informationsqualität individuell unterschiedlich ausfällt. Die gewünschten und benötigten Informationsergänzungen sind abhängig von dem aktuellen Interesse der Anwender: „Also das liegt ja immer in der Betrachtung des Anwenders, was. Also. Und vor allem was ihn halt interessiert.“ (B6-Interview, Abschnitt 48). Darüber hinaus ist es ebenfalls vom Vorwissen der Anwender abhängig und davon, wie detailliert die Infos betrachtet werden: „Also es ist vielleicht auch nochmal ein Unterschied, wer sich diese Dokumentation anguckt, auf welchem Level man Infos haben möchte.“ (B3-Interview, Abschnitt 55).

Von den Experten wurden vier negative Aspekte für die Qualität der Informationen ausgemacht. So haben fast allen Experten Informationsergänzungen gefehlt und allen Experten sind Informationen aufgefallen, die als überflüssig eingestuft wurden: „Paar Dinge waren mir ja aufgefallen, die mir persönlich noch fehlen würden oder die ich überflüssig finde.“ (B3-Interview, Abschnitt 89). Außerdem gab es Schlagwörter, die mit falschen Einträgen der Wissensdatenbank verknüpft wurden: „[...] wobei hier jetzt ja, da, wie eben ja auch schon angesprochen, praktisch die, das Stichwort richtig ist, aber die Info dahinter nicht ganz das widerspiegelt, was ich erwartet hätte.“ (B3-Interview, Abschnitt 55). Auch ist der Bezug der Informationsergänzung zur Fernsehdokumentation nicht immer ausreichend gegeben, wie B2 anmerkt: „Ich erfahre zwar was zu Dubai, aber ich erfahre jetzt nicht: Okay, wann wurde dieses Gerät in Dubai benutzt? Wer hat das da vorgestellt? Wieviele Zuschauer waren da? Welche Strecke sind sie geflogen? Wie lang war die Strecke? Also der Bezug zu diesem Artikel oder im Besonderen zu diesen, dieser Dokumentation, der fehlt mir.“ (B2-Interview, Abschnitt 64).

Die Experten haben einige Aspekte zur Qualität der Informationen herausgestellt, mit denen diese positiv bewertet werden. Für die Lehrerin B4 gab es kaum Informationen, die ihr gefehlt hätten: „[...] weil ich nicht viel über China an sich weiß, würde mich dann, würde mich interessieren, dass ich dann noch mehr über China an sich, die Geschichte Chinas, erfahren möchte. Aber das wäre natürlich jetzt. Ja, das würde natürlich wieder den Rahmen sprengen für dieses Format jetzt. Ich muss überlegen. Nee, also sonst könnte ich nicht sagen, dass mir irgendetwas Besonderes gefehlt hat.“ (B4-Interview, Abschnitt 49). Für den Experten B1 war

die Erkennung ganz in Ordnung: *„Die Erkennung war schon so ganz okay.“* (B1-Interview, Abschnitt 70). Und die Expertin B3 bewertet die Breite der Informationen als gut: *„Aber ich fand sonst den Mix aus, das ist ja eigentlich ein breiter Mix an Informationen. Das fand ich gut.“* (B3-Interview, Abschnitt 63). Fast alle befragten Experten bewerten die Auswahl der Schlagwörter und Informationen als gut, die Expertin B5 merkt zusätzlich an, dass sie es als gut empfand, dass die Auswahl der Information durch das entwickelte System geschah: *„Ansonsten war’s gut, dass ich das nicht selber auswählen musste. Sondern es mir direkt zur Verfügung gestellt wurde. Und ich die Auswahl der Inhalte jetzt auch gut fand [...]“* (B5-Interview, Abschnitt 34).

Erweiterungen des Systemdesigns

Die befragten Experten haben einige Anregungen für Erweiterungen des Systemdesigns eingebracht. B1 und B2 schlugen vor, dass der Second Screen auch als Controller für den First Screen dienen könnte. Die ergänzenden Informationen sind den Laufzeiten zugewiesen, in denen diese im Video erscheinen. Mit Hilfe des Second Screens könnte das Video des First Screens auf die Laufzeit einer gewählten ergänzenden Information gestellt werden: *„Da wär’s dann für irgendwann später ganz geil, glaub ich, dann auch vielleicht sogar zu der Stelle vom Video zu springen. Da stehen hier zumindest immer noch die Timestamps drunter. Oder? Also das wär‘ irgendwie theoretisch möglich.“* (B1-Interview, Abschnitt 44). Auch B2 möchte den Second Screen als Fernbedienung nutzen können und die Laufzeit des First Screens verändern: *„Ich finde ja auch interessant, wenn, wenn das interaktiv wird. Wenn ich einen Second Screen habe. und ich sage Stopp. Auf dem Second Screen könnte ich Stopp sagen, der Film hält an und sage ich, spring mal hier weiter und dann startet da ‘n Film quasi. Dass ich, dass ich das als Fernbedienung benutze, das finde ich schon sinnvoll beim Second Screen.“* (B2-Interview, Abschnitt 74).

Der befragte B6 schlägt die Erweiterung um eine Teilen-Funktionalität vor, um im Bildungsbereich das kooperative Arbeiten zur erleichtern: *„Und wenn’s jetzt irgendwie so in den schulischen Bereich halt geht und so. [...] Heute ist das alles so generell halt alles im Wandel, dass man vielleicht auch irgendwie, wenn man auch die Möglichkeit halt hat irgendwie die Informationen auch mit anderen halt zu teilen oder so. Dass wenn wir jetzt, nehmen wir jetzt mal diesen, dieses Beispiel. Es gucken mehrere, halt ‘ne Dokumentation und sollten sich halt im Nachhinein halt darüber austauschen, dass man dann anhand dieser Beiträge, die halt die App generiert hat, dass man die vielleicht in so’n Pool zusammenstellen kann oder die halt austauschen kann, oder so.“* (B6-Interview, Abschnitt 72).

Des Weiteren wurde vorgeschlagen, dass das System das Filtern und Gruppieren von Informationen ermöglicht und dem Verhalten bzw. den Bedürfnissen der Anwender gegenüber sensitiv sein könnte. Die Art der Informationen, die von den Anwendern gewünscht werden, können sich unterscheiden. Diese Wünsche könnte das System registrieren und bei der Bereitstellung der Informationen beachten: *„Und vielleicht irgendwie einfach dahin geleitet.“*

Wenn es mehrere sind, also da ja wahrscheinlich viele Informationen möglich sind anzuzeigen, dass ich irgendwie, dass die irgendwie gruppiert werden. Dass ich ungefähr weiß: Okay jetzt, ich interessiere mich gerade für 'nen Land und dass ich das dann vielleicht so gruppiert habe. Dass ich sage: Okay, hier, ich möchte hier diese Länder und dann sie waren vielleicht gerade da, sowas. Vielleicht, dass er das selber merkt, dass ich, irgendwie, mich für die Länder informiere, so. Und nicht an die Schauspieler oder Personen mich eh nicht interessieren, weil ich mir die Namen nicht merken kann oder so.“ (B1-Interview, Abschnitt 31).

Ein weiterer Vorschlag ist die Möglichkeit, dass das System die Historie der betrachteten Videos und Informationen speichert und dem Anwender zu einem späteren Zeitpunkt zur Nachbetrachtung zur Verfügung stellt. Dies könnte durch Nutzerprofile ermöglicht werden: *„Und dann möchte ich schon eigentlich eine Auswahlmöglichkeit haben: okay, das sind das und das hast du geguckt, hier sind die Informationen dazu. Also wirklich so eine Art Protokoll führen. Was habe ich geguckt und welche Informationen gibt es dazu? Weil dann erinnere ich mich vielleicht morgen daran: Ah ich hab ja gestern die Doku über tolles Essen gesehen und ich möchte das jetzt nachkochen zum Beispiel. also so praktische Sachen. Wie ist denn das Rezept dazu? Und dann könnte ich dann ja gucken, okay, was habe ich gestern geguckt? Ahja, da ist ja die Doku und da sind Rezepte. Zum Beispiel. Deswegen eigentlich so 'ne. Ob das jetzt wie gesagt, ja, auf dem, auf dem Tablet oder iPhone, das ist ja egal, aber es ist ja quasi im Hintergrund, es ist eine Datensammlung natürlich.“ (B2-Interview, Abschnitt 32).* Die Sammlung dieser Daten empfindet B2, der im Marketing-Bereich tätig ist, als unbedenklich, wenn der Anwender der Datensammlung zustimmt und sie nur diesem zugänglich ist: *„Ich bin da, solchen Sachen immer relativ offen, ich hab da keine Angst, irgendwie, dass meine Daten irgendwie genutzt werden und weitergegeben werden. Und wenn ich damit einverstanden bin, dass eben gesammelt wird, was habe ich geguckt und dann solche Informationen gesammelt gespeichert werden und die nur für mich zugänglich sind, finde ich das vollkommen in Ordnung.“ (B2-Interview, Abschnitt 32).*

Anwendungsbereiche des Systems

Das System wurde im Rahmen dieser Arbeit für die automatisierte Informationsergänzung von Fernsehdokumentationen entwickelt. Entsprechend der angedachten Handlungsweise kann das System im privaten Bereich verwendet werden, wie es auch B3 aufzählt: *„Aber eben auch so in der Heimanwendung, wenn man interessiert ist, sich verschiedene Dinge, Dokumentation, so wie hier jetzt, anschaut und einfach nochmal ein bisschen in die Tiefe gehen möchte.“ (B3-Interview, Abschnitt 81).*

Weitere Einsatzmöglichkeiten eines solchen Systems sind denkbar. Einige dieser Einsatzmöglichkeiten wurden von den befragten Experten vorgeschlagen und werden im Folgenden erläutert. Für die Anwendung in anderen Bereichen können Änderungen nötig sein.

Eine naheliegende Anwendungsmöglichkeit ist die Verwendung des Systems nicht nur für Videos, die ins System hinzugefügt wurden, sondern als Informationsergänzungs-System für Dokumentationen im linearen Fernsehen: *„Ansonsten glaube ich, könnte sowas gut funktionieren für so Dokumentation, wenn man die im Fernsehen guckt, dass man halt da die Möglichkeit hat, sei es über ‘ne App oder eben, weiß nicht wirklich, gibt ja auch so Knöpfe jetzt bei so Fernbedienungen mit „Zusatzwissen abrufen“ oder so. Dass automatisch das dann einem zur Verfügung gestellt wird, passend zur jeweiligen Dokumentation.“* (B5-Interview, Abschnitt 38).

Ebenso naheliegend und einfach umzusetzen ist es, die erzeugten ergänzenden Informationen nachträglich zur Verfügung zu stellen, so dass die Anwender diese als eine Art Protokoll oder Skript der Fernsehdokumentation lesen können. Nahezu alle befragten Experten artikulieren dies Möglichkeit der Nutzung des Systems. So auch der befragte B2, der das Tracken der gezeigten Inhalte vorschlägt, um die gesehenen Informationen auch nach einem längeren Zeitraum nachschlagen zu können: *„Ja, wie ich gesagt habe, wenn, wenn das einfach mittrackt, um zu zeigen halt, als, als Notizblock quasi, wo ich rein stöbern kann, wenn ich sage, ich habe das und das gestern geguckt und ich habe das und das letzte Woche geguckt, was, dann weiß ich ja ungefähr: Mensch, ich hab hab ja diese, diese Doku vor 2 Wochen geguckt. Wie hieß nochmal der Regisseur zum Beispiel. Da weiß ich, ich habe hier einfach diese Doku, würde einfach gucken: Okay, was hab ich hier nochmal für Themen? Da hat mich interessiert, dieser, dieser Hersteller von dem, oder von dem Bus oder von diesem Jeep, der dann vorgestellt worden ist. Was ist das für ein Jeep? Also wirklich diese Sachen, dass ich die nachschlagen kann.“* (B2-Interview, Abschnitt 66). Das spätere Betrachten der Informationen unterstützt die Anwender bei der Erinnerung und bietet Ansätze zur weitergehenden Recherche: *„Und gerade auch für, um, damit ich irgendwie nicht wieder Notizen machen muss. Was ich nicht mache, so, aber trotzdem irgendwie Sachen denke, find ich interessant, so. Will ich mir merken, merke ich mir dann aber vielleicht nicht. Und ich kann es mir sozusagen danach dann, sozusagen, also gar nicht nur für die Zeit, wo ich mir dieses Ding anschau, sondern auch sozusagen nicht, also so, dass es mir noch einen tieferen Mehrwert bringt, weil ich einfache Ansatzpunkte habe, um mich weiter zu beschäftigen mit diesem Thema.“* (B1-Interview, Abschnitt 88). Generell wird es als sinnvoll erachtet, die Möglichkeit zu eröffnen, die Informationen nachträglich zu betrachten. Beispielsweise hatte B3 Probleme mit der simultanen Nutzung beider Geräte und musste das Video des First Screens stoppen, um den Second Screen zu beachten. Die nachträgliche Nutzung der Informationen würde durch B3 bevorzugt werden: *„Und die Zusatzinfos, glaube ich, würde ich mir jetzt hintendran nochmal eher angucken als ganz genau nebenbei. Ich glaube vielleicht ist der 10 Minuten Ausschnitt auch dann zu kurz, oder ich müsste wirklich immer Stopp machen. Also ich glaube so ‘n simultanes Info checken, glaube ich, kommt für mich eher nicht infrage. Sondern ich brauch dann. Muss mich dann nur dem Second Screen widmen und sozusagen die Hauptinfoquelle erstmal auf Pause stellen. Und dann hier gucken, was wird mir angeboten?“* (B3-Interview, Abschnitt 49).

Ein weiterer häufig geäußelter Vorschlag ist die Nutzung eines solchen Systems im Bildungsbereich. Dies gilt zum einen für die Schule mit SuS, aber auch bei Fortbildungen und der Erwachsenenbildung, wenn Fernsehdokumentationen gezeigt werden. Jede Person bringt unterschiedliche Vorkenntnisse mit, so dass sie durch ein solches System individueller unterstützt werden kann. Die Anwender können selbstständig entscheiden, welche Information sie überspringen oder näher betrachten: *„Weil ich denke, also gerade im Bildungsbereich, wie ich das eben schon angesprochen hatte, würde ich es gut finden, aber auch sozusagen im, sagen wir mal bei Fortbildungen oder Erwachsenenbildung usw.. In, wenn man jetzt ‘n digitalen Lehrgang oder sowas hat und da dann sozusagen die zusätzlichen Infos immer nochmal abrufen kann. Weil da ja. Jetzt gerade beim Fortbildungsangebot, eben Leute mit unterschiedlichen Vorerfahrungen usw. kommen und die kriegen praktisch Informationen angeboten. Also ein Angebot und können dann selektiv auswählen: Das interessiert mich und da möchte ich noch mal genau etwas zu wissen. Wenn jemand anderes mit einem anderen Hintergrund sagt: Da kenne ich mich aus, da brauche ich keine Infos, sondern in einem anderen Bereich. Als sozusagen Bildungsbereich kann ich mir das sehr gut vorstellen. Digitalisierte Fortbildung, wie auch immer.“* (B3-Interview, Abschnitt 81). Im Bildungsbereich könnten die Lehrkräfte die Informationsergänzungen mit Zwischenfragen erweitern, um die Aufmerksamkeit der lernenden Personen zu erhöhen. Zusätzlich könnte ein Aufgabenblatt integriert werden: *„Besonders cool fänd ich, wenn man dann noch so Zusatzfunktion hätte, wie ich selbstständig noch so Zwischenfragen mit einblenden kann oder so. Wo man einfach nur so, ja nein beantworten kann, um so ein bisschen die Aufmerksamkeit nochmal zu reaktivieren etc.. Und am Ende dann am besten, falls noch so Aufgabenblatt 1, wo man einfach drauf drückt und dann kommen so die Aufgaben, die ich in Bezug zur Dokumentation an meine Zielgruppe richte.“* (B5-Interview, Abschnitt 74). B6 gibt jedoch zu bedenken, dass die automatisierte Bereitstellung der Informationen auch Nachteile für die SuS haben könnte. Da bereits tiefgehende Informationen zur Verfügung stehen, könnte sich die selbstständige Recherche der SuS reduzieren: *„Das hat auch immer alles Vor- und Nachteile. Hindert natürlich vielleicht auch daran, dass man dann vielleicht selbst recherchiert. So. Gerade wenn ‘s jetzt um diesen schulischen Aspekt halt geht.“* (B6-Interview, Abschnitt 68).

Im Kontext des Fremdsprachenunterrichts werden Filme, Dokumentationen und sonstige Videos in der Fremdsprache eingesetzt. B4 legt nahe, dass man beispielsweise anstelle der bisherigen Informationen die verwendeten Vokabeln erhält. Dies kann zu einem besseren Verständnis des Inhalts führen: *„[...] also jetzt zum Beispiel im Sprachunterricht macht das natürlich auch viel Sinn, ne. Also wenn ich jetzt, da ich ja Französisch als Fach habe, kann man natürlich auch Vokabeln, ne. Also dann schnell erfassen, ja oder schnell. Oder wenn man jetzt fremde Filme oder Dokumentationen oder irgendwo Fremdsprachen schaut, kann man die Sprache ja auch fördern. Und wenn man jetzt komplett einen Film auf einer Fremdsprache schaut, da besser den Film auch verstehen, weil man denn ja Vokabeln geliefert bekommt.“* (B4-Interview, Abschnitt 73). Auch der Unterrichtsablauf könnte sich flüssiger gestalten. Es kommt vor, dass die gezeigten Videos mehrfach pausiert werden müssen, damit die Lehrkräfte

Inhalte oder Vokabeln der Fremdsprache erläutern. Ein solches System könnte das Pausieren vermeiden: *„Oder wenn du dann bestimmte Begriffe in einem Film hast, die immer wieder auftauchen und du kannst, du verstehst es deswegen nicht, dann kannst du es halt schnell nochmal, ohne dass du im Wörterbuch nachschaust. Ja, da weißt du das halt schnell und ohne, dass ich jetzt, als Lehrer, alle 3 Minuten den Film stoppe, um irgendwas zu erklären. Weil die Schüler haben ja ein individuelles Vorwissen und einige können das halt einfach und kennen das und verstehen das, aber andere können, ja, sind weiter abgeschlagen oder können zu dem Themengebiet die Vokabeln nicht und so ist es individueller.“* (B4-Interview, Abschnitt 77).

Auch Menschen mit Einschränkungen, wie z.B. Seh- oder Hörschädigungen, könnten von einem System, das weiterführende Informationen auf einem Zweitgerät anbietet, profitieren: *„Oder vielleicht auch für Menschen mit Sehbehinderungen oder Hörgeschädigte, weiß ich nicht, das führt jetzt sehr weit, aber vielleicht, vielleicht ist das für die dann auch einfacher, ne. Das man dann, wenn man ein Gerät hat, was man überall mitnehmen kann das. Ja, oder auch im Fernsehen, dass man dann Dinge besser erkennen kann, ne. Oder ein Sehbehinderter, der natürlich auf seinem mobilen Gerät, was auf ihn zugeschnitten ist, sind auch Dinge besser, nachverfolgen kann. Oder ein Hörgeschädigter natürlich, der sowieso denn vieles auf, auf einem großen Bildschirm dann nicht so, ja nicht hören kann und dann aber seine Hilfestellung über ein Zweitgerät bekommt.“* (B4-Interview, Abschnitt 73).

Im Bereich des Marketings könnte ein solches System genutzt werden, um passend zum Inhalt der Fernsehdokumentationen Produkte anzupreisen. Dies könnten Reisebuchungen für gezeigte Reiseziele oder auch Angebote von gezeigten Produkten sein: *„Oder man könnte natürlich auch das im Marketing benutzen, dass man vielleicht Reisen buchen kann, oder Angebote bekommt, oder dass man sogar shoppen kann. Man kann ja natürlich dann sofort, du siehst im Fernsehen eine Dokumentation, siehst du irgendwelche tollen Sachen, die da geschnitzt wurden oder sonst was. Und wenn du die natürlich gleich online bestellen könntest, das könnte ich mir vorstellen.“* (B2-Interview, Abschnitt 17).

Nicht alle Vorträge und Vorlesungen arbeiten mit visuellen Elementen und erläutern alle Aspekte des Vorgetragenen ausführlich. Auch haben alle Zuhörer unterschiedliche Vorkenntnisse. Ein solches System könnte das Vorgetragene mit Hintergrundinformationen und visuellen Elementen anreichern: *„Also ich könnt mir sowas auch sehr gut, auch für Vorträge oder sowas vorstellen. Also so die, also so das geht, ne. Live ist alles ‘n bisschen technisch, ‘n bisschen schwieriger, aber einfach so für Sachen, die weniger. Also bei ‘ner Doku ist es ja manchmal schon so, dass dann ne Karte angezeigt wird, oder sowas. Also ohne, dass sie sozusagen dann schon noch mehr selber mit, mit visuellen Sachen arbeiten, oder mit Hintergrundinformationen, oder sowas. Bei ‘nem, bei einem Vortrag ist es eher noch weniger, oder eben, es kann auch ‘n aufgezeichneter Vortrag sein, zum Beispiel so, ne. Wäre dann technisch einfacher. Oder eine Vorlesung, oder sowas, die aufgezeichnet wurde, irgendwas, wo es wirklich um Informationen geht, so, die aber selber nicht visuell großartig ist, so ne. Da könnte ich mir das sehr gut vorstellen, dass jemand erzählt einfach nur was, hat selber*

eigentlich nicht groß 'ne Darstellung. Und ich kann dann. Die Person spricht über eine Person, und ich sehe dann das Foto von der, so über diese Person, die sie spricht, auf diesem Bildschirm oder sowas daneben her. Das könnte ich mir sehr gut vorstellen.“ (B1-Interview, Abschnitt 82)

Die wissenschaftliche Mitarbeiterin B5 weist darauf hin, dass die Wahl der Zielgruppe eines solchen Systems eine wichtige Überlegung darstellt, da ein solches System nicht für alle gleichermaßen geeignet sei: *„Ich glaube, das ist ein System, was nicht generell umsetzbar, einsetzbar ist für alle Zielgruppen. Sondern man schon überlegen muss, in welchen Kontexten das sinnvoll ist.“ (B5-Interview, Abschnitt 74).* Beispielsweise unterscheidet sich die Wahrnehmungsweise der Generationen untereinander. Die unterschiedliche Wahrnehmung liegt an unterschiedlicher Prägung, durch den Umgang mit digitalen Medien, der sich zwischen den Generationen deutlich unterscheidet. Dieser Aspekt sollte in die Überlegungen hinsichtlich der Zielgruppe einfließen und bei der entsprechenden Ausgestaltung des Systems bedacht werden: *„Es ist aber so, dass in der Wahrnehmungsweise, ich weiß nicht, ob du dich damit beschäftigt hattest. Es ja unsere Generation anders wahrnimmt als Generationen vor uns. Das heißt, wir haben so ein kreisrundes Wahr-, Aufmerksamkeitserfassungsvermögen, oder wie das heißt, in diesen ganzen digitalen Bereichen. Deswegen sind wir schneller mit Bildverarbeitung, mit Instagram, mit all diesen ganzen Sachen und können quasi, ich kann dahin gucken und wenn ich das Tablet hier liegen hab, wie in so'nem Kreis sehen. Personen, die nicht der Generation Y, also Y ist die Generation ab 87, die als erste Digital Natives mit gelten, also noch später, als also noch vor, jetzt sind ja die Generation Z, die Millennials, aber die Generation davor, alle davor nehmen wirklich auch alles visuelle im digitalen Bereich wahr, wie 'n Buch. Das heißt, die wären gar nicht in der Lage, also meine Mutter wäre gar nicht gleichermaßen in der Lage, die Inhalte so schnell hier zu verarbeiten und mit dem, weil sie von rechts nach links hier gucken würde und eben auch genau das. Das heißt, da wäre die Handhabung anders und dann könnte es sein, dass es sinnvoll wäre, ein anderes Format zu nehmen, also eben nicht auf zwei Screens das zu verteilen, sondern da müsste man, glaub ich, gucken mit: Ab welcher Altersstufe macht unterschiedliche Präsentationen Sinn? Und dieses runde Lesen, das kann erlernt, also ist halt durch die Sozialisation bedingt. Genau, das gab's in vielen Gaming Studien, dass zum Beispiel Leute, die ganz viel Games gespielt haben, ganz anders visuell Kontexte erschließen können, wahrnehmen können, viel mehr auf Kleinigkeiten achten etc..“ (B5-Interview, Abschnitt 40).*

Bewertung des Systems

In der Bewertung des Systems ziehen alle befragten Experten ein positives Fazit. In der Bewertung am kritischsten waren B1 und B3. Die Erwartungen von B1 wurden zum Teil gut erfüllt, allerdings erschienen die Informationen zu textlastig und zu klein. Nach B1 hätte der Fokus der Darstellung der Informationen eher auf visuellen Elementen liegen sollen: *„Also Teile davon schon gut erfüllt, mit dem, dass ich das, also, dass ich das speichern konnte, dass*

ich das das gut schnell lesen konnte. Zum Teil halt zu, zu klein, textlastig so, was mich dann doch eher dann rausgezogen hat so. Ich das gar nicht immer bräuchte, sondern halt eher, der so dann doch gemerkt hat, dass ich diese visuellen Sachen vielleicht doch auch irgendwie, irgendwie doch besser finde [...].“ (B1-Interview, Abschnitt 66). Für B3 waren die Zusatzinformationen gut und die gezeigte Fernsehdokumentation verständlich. Allerdings hätte sich B3 mehr Zeit für die Erkundung und Einarbeitung der Informationsergänzungen gewünscht und auch gerne die verknüpften Informationsquellen betrachtet: „Grundsätzlich fand ich die Dokumentation schon allgemein verständlich. Ich finde hier diese Zusatzinfos gut, hätte mich aber, glaub ich, sonst nochmal, wie gesagt mit einem Link nochmal weiter rein gearbeitet und hätte, glaub ich, noch ein bisschen mehr Zeit gebraucht, um da tatsächlich das so anzuwenden.“ (B3-Interview, Abschnitt 71). Im Hinblick auf die geäußerten Erwartungen wurden diese für B3 gut erfüllt. Abstriche wurden jedoch für überflüssige und fehlende Informationen gemacht: „Ja, eigentlich schon soweit ganz gut. Ich hatte ja angesprochen, dass da so ein paar Dinge für mich sozusagen zu lapidar, überflüssig, Englisch, Deutsch usw. tauchte hier auf, es tauchten aber die Uiguren auf, um eben sozusagen so ein Konflikt anzusprechen, den es ja gibt, der mich interessieren würde. Dieser Datenschutz fehlte mir hier so ein bisschen, das als Konfliktfeld.“ (B3-Interview, Abschnitt 69). In der Gesamtbetrachtung fällt das Fazit bei B3 gut aus: „[...] eigentlich hab ich ja, glaub ich, schon so die Punkte, wo ich sage: Mensch, da ist das super. Vielfalt. Paar Dinge waren mir ja aufgefallen, die mir persönlich noch fehlen würden oder die ich überflüssig finde. Also das genau finde ich ne tolle, rundum gelungene Sache.“ (B3-Interview, Abschnitt 89).

Für B6 ist es überraschend, dass ein solches System möglich ist. Damit wurden die Erwartungen übertroffen: „Ich würde sagen: Krass, dass sowas funktioniert. Also dass sowas halt überhaupt technisch möglich ist, dass man so mit, anscheinend irgendwie auf so einer Datenbank oder so, oder irgendwie automatisiert halt Sachen aus so Pools halt herausziehen kann [...].“ (B6-Interview, Abschnitt 62). Auch für B4 wurden die eigenen Erwartungen übertroffen: „Ja, also wie schon gesagt würde ich, würde ich sagen, dass ich es besser so finde, als das, was ich gedacht habe.“ (B4-Interview, Abschnitt 51). Die Erwartung, dass die Anwender die Informationen selbstständig suchen müssen und eine Art Suchmaschine zum Einsatz kommt, wurde revidiert. Dass die Informationsergänzungen automatisiert dargeboten werden, bewertet B4 als besser: „Ich persönlich würde mich ablenken, wenn ich noch selbst recherchieren muss, oder wenn ich selbst noch in eine Suchmaschine was eingeben müsste. Und wenn ich mich dann noch vertippen würde dabei. Und dann noch erstmal mir was angeboten wird, was ich an Informationen jetzt bekommen könnte. Da ist mir jetzt natürlich lieber, dass mir das sofort geliefert wird, wovon ich. [...] Und von der finde ich das so dann besser.“ (B4-Interview, Abschnitt 51).

Für den freien Journalisten B6 ist das entwickelte System faszinierend und eine Zeitersparnis. In vielen Bereichen wird das System als hilfreich bewertet: „Also ich find's faszinierend, dass sowas funktioniert. Ich glaube, es kann uns halt in vielen Bereichen einfach hilf, hilfreich halt

sein. Dass man, ja, das ist einfach eine Zeitersparnis.“ (B6-Interview, Abschnitt 62). Die Lehrerin B4 kann sich beispielsweise die Verwendung eines solchen Systems im Beruf vorstellen: „Und natürlich denke ich auch daran, wie ich das in meinem Beruf machen kann und dann würde es. Ja, würde es auch, ja könnte ich mir viele Möglichkeiten vorstellen, wo es dann Sinn macht.“ (B4-Interview, Abschnitt 79).

Im Hinblick auf die Forschungsthese kommen alle Experten zu einem positiven Fazit. Wie sehr das Verständnis und Erlebnis von Fernsehdokumentationen durch die Darstellung von Informationsergänzungen auf einem Second Screen gefördert wird, ist nach Aussagen von B4 und B6 abhängig von dem Kenntnisstand und Interesse der Anwender: „Also es kommt drauf an, denke ich. Denn es gibt natürlich Dokumentationen über Themengebiete, die ich ja schon irgendwie kenne. Und wo ich das dann einfach nur nochmal wiederholend, vielleicht mir angucken möchte. [...] Aber wenn es jetzt um Themengebiete geht, die irgendwie sehr komplex sind oder die ich einfach nicht so kenne oder auch von, ja. Dann, dann halte ich das für sinnvoll.“ (B4-Interview, Abschnitt 79).

Der Experte B1 schwankt in seiner Bewertung, in welchem Umfang das System das Verständnis von Fernsehdokumentationen fördert. Er stellt fest, dass das System das Verständnis nur leicht fördert, da einige Aspekte, wie die Karten fördernd sind und andere Informationen nicht schnell genug erfasst wurden oder nicht präzise genug waren. In der späteren Betrachtung der Informationen sieht B1 eine weitere Möglichkeit, mit der das Verständnis der Fernsehdokumentationen gesteigert werden kann: „Ich glaube, in dem konkreten Fall jetzt so direkt noch nicht, weil ich nicht so. Also leicht, weil ich dann also, wie gesagt, diese Karte oder sowas, das waren Sachen, die ich jetzt so direkt schon wie schnell aufnehmen konnte, während ich das geschaut hab. Für die anderen Sachen waren das die, auch die Informationen, die da waren, eher so Hintergrund also eher so zu, nicht prägnant genug, um sie während des Filmschauens schnell mehr Informationen dazu zu kriegen, so. Sondern das wären dann eher so ein. Um mehr Hintergrund, für mich den Punkt wichtige Hintergrundinformationen, wahrscheinlich kriegen zu können, hätte ich mehr als sozusagen nur die ersten 10 Wörter von der Beschreibung lesen müssen. Und genau, von daher sozusagen jetzt so direkt beim Film schauen noch nicht, wäre vielleicht eher was für danach nochmal halt irgendwie sich zu vertiefen.“ (B1-Interview, Pos. 68). Im späteren Verlauf des Interviews bestärkt B1, dass das System fördernd für das Verständnis von Fernsehdokumentationen ist. Die Anwender können die Informationen betrachten, die interessant für sie sind. Außerdem kann das System genutzt werden, um eine Art Notiz oder Merkhilfe durch die Informationen zu haben und diese später verwenden zu können: „Finde ich schon fördernd. Ich find es interessant, gerade auch, um sich auch. Dann. Also sich so die Aspekte herausgreifen zu können, die einen gerade besonders interessieren, die man gerade besonders spannend findet, oder man hingengeblieben ist, so. [...] Und gerade auch für, um, damit ich irgendwie nicht wieder Notizen machen muss. Was ich nicht mache, so, aber trotzdem irgendwie Sachen denke, find ich interessant, so. Will ich mir merken, merke ich mir dann aber vielleicht nicht. Und ich

kann es mir sozusagen danach dann, sozusagen, also gar nicht nur für die Zeit, wo ich mir dieses Ding anschau, sondern auch sozusagen nicht, also so, dass es mir noch einen tieferen Mehrwert bringt, weil ich einfache Ansatzpunkte habe, um mich weiter zu beschäftigen mit diesem Thema.“ (B1-Interview, Abschnitt 88).

Die Expertin B3 wertet die Anwendung als fördernd bzw. sehr förderlich. Die Möglichkeit, die Informationen während des Konsums der Fernsehdokumentation zu erfassen, aber auch die Option die Informationsergänzungen im Anschluss zu betrachten, führen zur positiven Bewertung hinsichtlich der Forschungsthese: *„Ja also, ich find das schon fördernd, wenn man eben, wie gesagt, nicht nicht nur sich nur berieseln lassen will, sondern dann eben auch noch mal ein paar Sachen überprüfen möchte oder sich hier neue Kontexte aufgeworfen worden sind, wo man sagt, das möchte ich mir noch mal genauer angucken. Dann hat man hier einen leichten, schnellen Zugriff und finde das, genau, sehr förderlich. Wobei, genau. Ich finde im, im Nachgang einer solchen Dokumentation und dann auch noch rückblickend sozusagen sich diese Zusatzinformationen angucken kann. Das man also nicht immer mittendrin, mal Stopp macht. Aber dass man eben auch diese Möglichkeit hat, entweder mittendrin mal zu stoppen oder sich erstmal die gesamte Dokumentation anzugucken und dann im Nachgang nochmal sich die Infos anzugucken. Das finde ich beides gut.“ (B3-Interview, Abschnitt 87)*

Der Experte B6 bewertet das vorgestellte System als fördernd im Sinne der Forschungsthese. Dass das System bei Sachverhalten unterstützt, die unklar sind oder bei denen Interesse besteht, sorgt bereits dafür, dass es die Anwender in ihrem Verständnis diesbezüglich stärkt und sie unterstützt. B6 merkt zusätzlich an, dass die Einschätzung, ob etwas hilfreich oder fördernd ist, individuell entschieden wird: *„Also fördernd ist ja schon alleine der Teil, dass du halt einfach zusätzliche Informationen halt bekommst, du die halt unklar sind, oder die dich halt interessieren. Sind ja sozusagen so zwei Aspekte und dementsprechend ist das fördernd. Ja. Ja. ja oder ja? Mal darüber nachdenken. Ich finde, das ist auch immer alles so individuell.“ (B6-Interview, Abschnitt 74).* Auf die gezeigte Fernsehdokumentation bezogen, konnte das System B6 unterstützen, indem er einen schnellen Überblick gewinnen konnte. Dies galt beispielsweise für die geographischen Informationen, die mit interaktiven Karten angereichert wurden: *„Ja, speziell jetzt auf diese Dokumentation. Ja, dass man einfach, nehmen wir jetzt mal das Beispiel mit dieser geographischen, diesem geographischen Aspekt, dass man halt einfach schnell halt weiß so, in welche, welche Regionen, über welche Länder wurde halt angesprochen? Dass man da sozusagen schnellen Überblick gewinnt.“ (B6-Interview, Abschnitt 60).*

Die Expertin B5 erläutert, dass durch die Darstellung der Informationen auf dem Second Screen ein weiterer Aspekt im Lernprozess hinzugefügt wird. Neben der auditiven und visuellen Vermittlung des Wissens über die Fernsehdokumentation auf dem First Screen kommen die dargestellten Informationen auf dem Second Screen hinzu. Die Anwender interagieren mit dem Second Screen und den dargebotenen Informationen. Die Informationen bestehen aus Texten und Bildern, die visuell vom Anwender erfasst werden. Der Second Screen fördert den Lernprozess als zusätzliches Element: *„Das Wissen, was ich visuell erlangt hab, aus der*

Dokumentation, wurde hier dann schon angereichert durch das, durch, durch die Schriftform. Was dazu führt, dass es dann nochmal doppelt im Kopf quasi 'ne Erwähnung findet, mit unterschiedlichen Arten. Also einmal rein visuell und natürlich ist das auch visuell, aber 'ne schriftliche Wissensaufnahme. Und dadurch können Informationen nachhaltiger abgespeichert werden im Kopf [...]' (B5-Interview, Abschnitt 62).

4.2.5 Gütekriterien

Für diese Ausarbeitung wurden Interviews mit sechs Experten durchgeführt. Die anschließend durchgeführte qualitative Inhaltsanalyse wird im Hinblick auf Gütekriterien betrachtet. Sowohl für die quantitative Forschung als auch für die qualitative Forschung gelten die Kriterien der Validität, Reliabilität und Objektivität in der Literatur als die „klassischen“ Gütekriterien“ ([88], S. 123) ([91], S. 21) ([94], S. 249). Diese Gütekriterien werden in der Forschung seit einigen Jahrzehnten für die qualitative Forschung diskutiert. Es werden verschiedene Positionen in der Sache eingenommen. Eine Position ist es, dass die Gütekriterien in der empirischen Forschung universell gelten sollen, also für die quantitative und die qualitative Forschung gleichermaßen. Eine andere Position ist die Ablehnung oder Neu-Formulierung von Gütekriterien für die qualitative Forschung ([87], S. 202f). Übereinstimmend mit Mayring kommt ein Review von qualitativen Arbeiten in der naturwissenschaftsdidaktischen Forschung zu der Erkenntnis, dass die drei klassischen Gütekriterien einen hohen Stellenwert besitzen. In dem Review ergibt es sich, dass die Kriterien Validität, Reliabilität und Objektivität am häufigsten verwendet werden [96]. Nachfolgend werden die klassischen Gütekriterien näher beschrieben.

Reliabilität

Damit das Gütekriterium der Reliabilität zu erfüllen ist, müssen die Ergebnisse der Untersuchung reproduzierbar sein, wenn die gleichen Bedingungen vorliegen. Ein zentraler Faktor für die Reliabilität ist die genaue Dokumentation der durchgeführten Interviews und die Offenlegung der Forschungsmethode. Es wurde offengelegt, wie und mit wem die Interviews vorbereitet und durchgeführt wurden. Dies geschah mit Besprechung des Interviewleitfadens (siehe 4.2.2) und der Beschreibung der Interviewpartner und der Interviewdurchführung (siehe 4.2.3). Die durchgeführten Interviews wurden transkribiert, so dokumentiert und für die weitere Analyse bereitgestellt. Darüber hinaus wurde die Methodik der Auswertung des Materials dargestellt. Hierzu gehört die Offenlegung der Methode zur Erstellung des Kategoriensystems mittels induktiver Kategorienbildung am Material.

Validität

Die Validität ist das Gütekriterium, das bemisst, wie gut die Forschungsmethode den gewünschten Sachverhalt untersucht. Im vorliegenden Fall, wie gut die Forschungsmethode geeignet ist, die Forschungsthese zu untersuchen. Damit das Validitäts-Kriterium erfüllt werden kann, muss auch die Reliabilität erfüllt sein (vgl [88], S. 125). Das heißt, dass die durchgeführte qualitative Forschung nur valide sein kann, wenn die Ergebnisse reproduzierbar sind. Der Punkt der Reliabilität wurde bereits besprochen. Für den Punkt der Validität wurde zusätzlich sichergestellt, dass die Forschungsthese in der Konzeption der Forschungsmethode berücksichtigt wurde. Im Besonderen geschah dies in der Entwicklung des Interviewleitfadens und der Auswahl der Experten, die Expertisen in Bereichen haben, die den Forschungsgegenstand tangieren. Auch die Auswertung der erhobenen Daten erfolgte vor dem Hintergrund der Forschungsthese.

Objektivität

Um das Qualitätskriterium der Objektivität zu erfüllen, müssen die Resultate der Forschungsmethode unabhängig von der durchführenden Person sein. Eine andere forschende Person muss zu den gleichen oder ähnlichen Ergebnissen gelangen. Für die Datenerhebung kann dieses Gütekriterium für die qualitative Forschung nicht in diesem Sinne erfüllt werden, da in der Kommunikation, also den Interviews mit den Experten, subjektive Aspekte vorhanden sind ([86], S. 683f). Auch bei einer Standardisierung der Interviewsituation mit offenen Fragen, kann es bei Wiederholung eine Variation des Antwortverhaltens geben. Die Subjektivität bleibt in gewissem Rahmen erhalten.

Es wird darauf geachtet, dass die Einflüsse der einzelnen Forschenden minimiert bzw. eliminiert werden, so dass die einzelnen Forschenden keinen relevanten Faktor darstellen. Das Kriterium der Objektivität wird eher für die Auswertung und nicht die Erhebung der Daten angewendet. Um den Faktor der individuellen Forschenden für die Analyse der erhobenen Daten zu minimieren, wurden die Forschungsmethoden (siehe 4.2.1) vor der Durchführung festgelegt, befolgt und transparent gemacht.

4.3 Gesamtbetrachtung

In Abschnitt 4.1 erfolgte mit der funktionalen Evaluation die Untersuchung und Besprechung der Verfahren zur Generierung der Informationsergänzungen. Es wurden mehrere Verfahren für die Vorverarbeitung, das Geocoding und das Named Entity Linking näher betrachtet und die Unterschiede aufgezeigt. In jedem dieser Bereiche gibt es untersuchte Verfahren, die gute oder sehr gute Ergebnisse erzielen. Abhängig von der Anforderung an diese bieten sich verschiedene Verfahren an.

Die Verfahren der Vorverarbeitung beeinflussen die Güte der nachfolgenden Verarbeitungsschritte. Die Vorverarbeitung mit dem StopWord Removal entfernt nur die Stoppwörter aus dem Textkorpus, wodurch alle relevanten Wörter erhalten bleiben, aber der Textkorpus weniger reduziert wird als bei der Vorverarbeitung mit SpaCy. Bei der Vorverarbeitung mit SpaCy werden die Wörter aus dem Textkorpus extrahiert, die für relevant erachtet werden. Hierdurch reduziert sich die Zahl der Wörter, die weiterverarbeitet werden müssen, deutlich. Es kann passieren, dass gesuchte Begriffe nicht extrahiert wurden. Es erfolgt eine Klassifikation dieser Wörter, die nachfolgende Verfahren ausnutzen können.

Für das Identifizieren von geographischen Namen und das Geocoding ist die Kombination aus SpaCy und dem „GeoNames Neu“-Verfahren am besten, unabhängig davon, ob möglichst alle geographischen Namen gefunden werden sollen oder False Positives vermieden werden sollen. Die Verarbeitung mit dem „GeoNames Neu“-Verfahren ist sehr zeitaufwendig. Trotz der Vorverarbeitung mit SpaCy, bei der die Anzahl der zu verarbeitenden Tokens deutlich reduziert wird, ist diese Kombination, mit einer Berechnungsdauer von etwa 100 Sekunden, nicht für höhere Anforderungen an die Berechnungszeit geeignet. Werden die Informationsergänzungen, wie beim entwickelten System, vor der Anwendung generiert, so ist die Berechnungszeit ausreichend.

Beim Verknüpfen der Entitäten mit einer Wissensbasis, also dem Named Entity Linking, ist die Auswahl von den Anforderungen abhängig. Es gibt kein Verfahren, das im Bereich des Precision und Recall besonders hohe Werte erreicht. Wenn es bei der Anforderung, ob es möglichst wenige False Positive-Zuordnungen geben darf oder möglichst alle gefunden werden sollen, keine Priorisierung gibt, dann bieten sich vier Verfahren an. Alle vier Verfahren haben unter 100 Sekunden für die Berechnung benötigt. Von diesen Verfahren war DBPedia Spotlight 0,9 mit etwa 21 Sekunden am schnellsten.

Die Verfahren sind in der Lage, für die automatisierte Generierung der Informationsergänzungen die gewünschten Entitäten aus den Untertiteln zu extrahieren. Diese können als geographische Namen klassifiziert, den geographischen Namen eine Koordinate zugeordnet und die Entitäten mit Einträgen einer Wissensbasis verknüpft werden. Die besseren Verfahren des Named Entity Linking erreichen in den Messungen einen F₁-Score von etwa

0,50. Dies zeigt, dass es in diesem Bereich Potenzial für Verbesserungen gibt. Es könnten weitere Verfahren untersucht oder die bestehenden verbessert werden.

Im Abschnitt 4.2 erfolgte die Evaluation des entwickelten Systems mit Hilfe einer Nutzerbefragung in Form von Experteninterviews und anschließender qualitativer Inhaltsanalyse. Hinsichtlich des entwickelten Systems zogen die befragten Experten ein positives Fazit. Es gab Kritiken und Verbesserungsvorschläge. Für die Experten überwogen jedoch die positiven Aspekte. Das System führt zu einer Zeitersparnis, fördert den Lernprozess sowie das Verständnis der Fernsehdokumentationen und kann auch im beruflichen Kontext unterstützen. Neben der Verwendung im beruflichen Kontext und dem angedachten Zweck der Anwendung sahen die Experten auch für weitere Bereiche Potenzial für den Einsatz eines solchen Systems. Nach dem Konsum der Videos können die generierten Informationen weiterhin als Notizen, Skript oder als Einstiegspunkt für weitergehende Recherchen genutzt werden. Ein solches System könnte nicht nur zur Generierung von inhaltlichen Informationsergänzungen dienen, sondern auch als Vokabel-Unterstützung bei Videos in Fremdsprachen. Erfüllt ein solches System die Anforderungen, die ein Echtzeitsystem hinsichtlich der Berechnungszeit mit sich bringt, so könnte dies parallel zum linearen Fernsehen oder bei Vorträgen und Vorlesungen eingesetzt werden.

Die Experten hatten nicht viel Vorerfahrung mit der bewussten Anwendung eines Second Screens. Die parallele Nutzung des First- und Second Screens funktionierte unterschiedlich gut. Während einige Experten beidem folgen konnten, mussten andere das Video auf dem First Screen pausieren, um die dargestellten Informationen des Second Screens zu verarbeiten. Viele der Anfangsschwierigkeiten spielten im Laufe der Nutzung, die nur etwa 10 Minuten andauerte, für viele Experten keine relevante Rolle. Es wurde eine gewisse Eingewöhnungszeit an das System benötigt. Dies spiegelte sich im Besonderen darin wider, dass sich viele Experten nach der Eingewöhnung mehr auf den First Screen konzentrierten und der Second Screen weniger als ablenkend empfunden wurde.

Zu einigen Aspekten der dargebotenen Anwendung haben die Experten Verbesserungsvorschläge gemacht. Beispielsweise wurde die Reihenfolge der neuen Informationen kritisiert. Es wurde gewünscht, dass die neuesten Informationen oben angefügt werden, so dass das manuelle Scrollen reduziert wird. Das manuelle Scrollen war ein Faktor, der vielfach als störend empfunden wurde. Eine veränderte Darstellungsweise, wie z.B. mit Kacheln, Filterung und hierarchisierten Informationen, könnte diesen Faktor minimieren.

Die Experten waren sich nicht einig darüber, ob es zu viele Informationsergänzungen gab, oder die Menge der Informationen gut war. Für einen Experten war die Menge zu umfangreich, ein Experte tendierte dazu, dass die Quantität „okay“ war, legte sich aber nicht fest. Drei Experten bewerteten die Menge der Informationsergänzungen als gut bzw. adäquat.

Die Informationsergänzungen lassen sich in verschiedene Arten einteilen. Im Vorfeld der Nutzung des Systems äußerten die Experten ihre Vorstellungen über die Informationsarten. Im

Anschluss an die Anwendungsphase änderten die Experten teilweise ihre Vorstellungen. Viele der gewünschten Informationsarten wurden von dem System angeboten.

Zur Qualität der Informationen wurden negative und positive Aspekte benannt. Es gab fehlende, überflüssige und falsch zugeordnete Informationsergänzungen. Dies lässt sich anhand der in Abschnitt 4.1.5 durchgeführten Messungen nachvollziehen. Für die geographischen Namen, also den Kartendarstellungen, dürfte dies weniger gelten. Während der parallelen Anwendung des Systems sind die fehlerhaften Informationen nicht allen Experten aufgefallen. Die Spannweite der Informationen wurde als gut befunden. Ebenso kamen die meisten Experten zu dem Fazit, dass die Auswahl der Schlagwörter und Qualität der Informationsergänzungen angemessen ist.

Die untersuchte Forschungsthese, dass die Darstellung der Informationsergänzungen auf einem Second Screen das Verständnis und Erlebnis von Fernsehdokumentationen fördert, wurde von den Experten anhand des entwickelten Systems bestätigt. Die Experten führen mehrere Punkte auf, durch die das Verständnis und Erlebnis verbessert werden. So gibt es die Möglichkeit, sich, mit Hilfe der Informationsergänzungen, während und nach dem Konsum der Fernsehdokumentation inhaltlich damit auseinanderzusetzen. Es werden Informationen bereitgestellt, die für die Anwender interessant oder unbekannt sind. Uninteressante oder fehlerhafte Informationsergänzungen können schnell übersprungen werden, reduzieren jedoch das Erlebnis. Eine Expertin erläuterte, dass die Verwendung des Second Screens und der dortigen Darstellung der Informationsergänzungen den Lernprozess um einen weiteren Aspekt bereichert und somit das Lernen fördert. Herausgestellt wurden die interaktiven Karten, die die geographischen Informationsergänzungen anreicherten. Diese sorgten dafür, dass ein schneller Überblick gewonnen werden konnte. Dies verbessert das Erlebnis, aber auch das inhaltliche Verständnis der Fernsehdokumentationen.

5 Zusammenfassung und Fazit

Mit dem Kapitel 1 wurde in das Thema eingeleitet und die Ziele dieser Arbeit wurden festgelegt. Im nachfolgenden Kapitel 2 wurden Anforderungen definiert, die von der zu entwickelnden Software erfüllt werden müssen. Hierzu gehören Anforderungen an die Auswahl der Videos und die Generierung der Informationsergänzungen. Ebenfalls umfasst dies die Anforderungen an die Darstellung des Videos und der Informationsergänzungen. Es wurde analysiert, dass die Informationsergänzungen auf einem Zweitgerät dargestellt werden sollen, während das Video auf dem Hauptbildschirm gezeigt wird. Die Art der Handhabung mit beiden Bildschirmen entspricht dem Second Screen-Phänomen. Ebenfalls in Kapitel 2 wurden Verfahren und Datenquellen vorgestellt, die zur Erzeugung der Informationsergänzungen genutzt werden können und für ihre spätere Darstellung benötigt werden.

In Kapitel 3 wurde dargestellt, mit welchen Verfahren, Methoden und Design-Entscheidungen die Software realisiert wurde, so dass diese die Anforderungen aus Kapitel 2 erfüllt. Es wurde begründet herausgearbeitet, welche Verfahren zur Erzeugung der Informationsergänzungen verwendet werden und welche Verfahren zur Umsetzung der Architektur des Software-Systems zur Anwendung kommen. Für die späteren Untersuchungen wurden mehrere Verfahren zur Erzeugung der Informationsergänzungen umgesetzt. Der Grundaufbau der Verfahren besteht aus der Vorverarbeitung der Untertitel der Fernsehdokumentationen und das anschließende Named Entity Linking sowie das Geocoding. Als Wissensbasen für das Named Entity Linking wurden Wikipedia und DBpedia verwendet. Für das Geocoding wurden die GeoNames-Datenbank und die Daten von OpenStreetMap genutzt. Die Darstellung der erzeugten Informationsergänzungen wurde anhand von Screenshots der entwickelten Software erläutert.

Kapitel 4 diente der Evaluation des entwickelten Systems. Es wurde funktional und qualitativ-empirisch evaluiert. Bei der funktionalen Evaluation wurde die Generierung der Informationsergänzungen untersucht. Es wurden Messungen für alle Verfahren der Arbeitsschritte zur Erzeugung der Informationsergänzungen und auch für die Berechnungsdauer dieser Verfahren durchgeführt, die Messergebnisse untersucht und die Unterschiede der jeweiligen Verfahren herausgearbeitet. Entsprechend der Anforderungen an das zu entwickelnde System müssen die passenden Verfahren ausgewählt werden.

Ein Verfahren, das besonders gut beim Precision-Wert abschneidet, hat zumeist einen niedrigeren Recall-Wert und umgekehrt. Eine Ausnahme bildet ein Verfahren des Geocodings. Die Kombination aus SpaCy und dem „GeoNames Neu“-Verfahren schneidet bei beiden Werten besonders gut ab und erreicht somit auch einen hohen F₁-Score. Diese Verfahrenskombination hat jedoch eine erhöhte Berechnungsdauer. Die Berechnungsdauer könnte beispielsweise mit Hilfe eines leistungsfähigeren Systems verbessert werden. Da die

Generierung der Informationsergänzungen in dieser Arbeit vor dem Abspielen der Videos geschieht, ist die erhöhte Berechnungsdauer ein untergeordneter Faktor. Die erreichten Werte des Precision, Recall und F_1 -Scores sind für diese Arbeit die entscheidenden Werte.

Die Verfahren des Named Entity Linkings können die Entitäten der Untertitel gut mit den Wissensbasen verknüpfen. Jedoch gibt es kein Verfahren, das sowohl im Hinblick auf die Precision- als auch die Recall-Werte besonders gut abschneidet. In der Berechnungsdauer gibt es Unterschiede zwischen den Verfahren. Je nach Anforderungen an das zu entwickelnde System muss das entsprechende Verfahren gewählt werden.

Die Messungen wurden anhand mehrerer Fernsehdokumentationen mit unterschiedlichen Eigenschaften und Thematiken durchgeführt. Es wurde gezeigt, dass die Generierung der Informationsergänzungen funktioniert und die implementierten Verfahren in den Messungen gute Ergebnisse erzielen. Die Untersuchungen ermöglichen die begründete Auswahl der besten Verfahren hinsichtlich der Anforderungen an das System. Darüber hinaus zeigen die Untersuchungen, dass das Konzept und die Verfahren zur Generierung der Informationsergänzungen auf die unterschiedlichsten Fernsehdokumentationen angewendet werden kann. Die Verfahren sind unabhängig von der Länge und der Thematik der Fernsehdokumentationen. Ebenso sind die Verfahren auf verschiedene Genres innerhalb der Fernsehdokumentationen anwendbar. Dies wurde anhand von Beispielen für Reportagen und Dokumentationen deutlich.

Neben der funktionalen Evaluation erfolgte in Kapitel 4 die qualitativ-empirische Untersuchung des entwickelten Systems. Dazu wurden sechs Experteninterviews durchgeführt. Die Expertinnen und Experten haben ihre Expertise aus verschiedenen Bereichen eingebracht und verschiedene Aspekte des Systems beleuchtet. Mit der qualitativen Inhaltsanalyse wurden die Experteninterviews ausgewertet. Die Experten haben vielfältige Ansichten und Punkte geäußert, die sich für die Bewertung des Systems eignen und sich auch auf zukünftige Verbesserungen und Anwendungsfelder beziehen.

Die Erwartungen an ein System, das den Anwendern Informationsergänzungen zu Fernsehdokumentationen liefert, wiesen vor und nach der praktischen Anwendung des Systems Unterschiede auf. Es gab Aspekte des entwickelten Systems, die von den Experten und Expertinnen unterschiedlich bewertet wurden. Ebenso verhielt es sich z.B. beim Aspekt der Qualität der Informationsergänzungen, zu denen positive und negative Punkte benannt wurden. Von den Expertinnen und Experten wurde die Verwendung eines Second Screens zur Darstellung der Informationsergänzungen positiv aufgenommen. In der Gesamtsicht waren sie sich einig, dass die Darstellung von Informationsergänzungen auf einem Second Screen das Verständnis und Erlebnis von Fernsehdokumentationen fördert. Somit wurde die untersuchte Forschungsthese von den Expertinnen und Experten bestätigt.

6 Ausblick

Im Rahmen dieser Arbeit wurde ein System zur Anreicherung von Fernsehdokumentationen mit Informationsergänzungen entwickelt. Anhand der Untertitel der Fernsehdokumentationen wurden Informationsergänzungen automatisiert generiert. Diese werden den Anwendern des Systems auf einem Second Screen dargeboten, während parallel dazu die Fernsehdokumentation auf dem First Screen angesehen werden kann. Die Quellen der Informationsergänzungen sind Geodaten und Wissensbasen.

Die durchgeführten Untersuchungen haben gezeigt, dass ein solches System machbar ist und die Umsetzung funktioniert. Das System wird von den Anwendern insgesamt positiv bewertet. Die Befragung der Expertinnen und Experten hat die Forschungsthese bestätigt und ergeben, dass das entwickelte System fördernd für das Verständnis und Erlebnis von Fernsehdokumentationen ist.

In dieser Arbeit war das entwickelte System auf die Verwendung von Fernsehdokumentationen beschränkt. In weiterführenden Arbeiten kann das System auf weitere Genres und andere Medienformate übertragen werden. Es könnten getroffene Annahmen überprüft werden. Beispielsweise gehören hierzu die Annahmen, dass Nachrichtenformate zu kurz sind und ihre Informationsdichte zu hoch ist. Außerdem wird angenommen, dass Spielfilme selten informativen Nutzen erzeugen und häufig fiktiv sind.

Es wurde gezeigt, dass die verwendeten Algorithmen ihren Zweck gut erfüllen. Es gibt Potenzial für Verbesserungen der verwendeten Algorithmen und es können weitere untersucht oder entwickelt werden. So könnten beispielsweise Machine Learning-Verfahren mit Datensätzen aus redaktionell erzeugten Informationsergänzungen zu den Videos überwacht lernen. Außerdem könnten Daten, die von Anwendern des Systems erzeugt werden, gesammelt und für weitere Verbesserungen genutzt werden.

Die Zielgruppen für ein derartiges Konzept können einen weiteren Aspekt zukünftiger Untersuchungen darstellen. Die Fragestellung, für wen sich die Verwendung eines Second Screens mit Informationsergänzungen eignet und wie das Konzept angepasst werden könnte, um den Nutzen für weitere Zielgruppen zu verbessern, kann in weiteren Arbeiten ausgeführt werden. Die Betrachtung von beispielsweise unterschiedlichen Alters- und Bildungsgruppen kann weitere Erkenntnisse erbringen.

Ein System, das Informationsergänzungen von Fernsehdokumentationen bereitstellt, ist für verschiedene Anwendungsbereiche denkbar. In der Befragung der Expertinnen und Experten wurden viele Bereiche benannt, in denen derartige Systeme eingesetzt werden können. Zu den genannten Bereichen gehört der Bildungsbereich, in dem die Hälfte der Befragten tätig ist. Beispielsweise das digitale Lernen kann durch die Anreicherung von Fernsehdokumentationen

und anderen medialen Inhalten unterstützt werden. In weiterführenden Untersuchungen kann die Übertragbarkeit dieses und ähnlicher Konzepte auf andere Bereiche untersucht werden.

Menschen mit Einschränkungen, wie Seh- und Hörschädigungen, könnten von einem System, das Informationsergänzungen anbietet, profitieren und im Kontext von fremdsprachigen Medien können Menschen unterstützt werden. Es kann untersucht werden, wie ein System gestaltet sein muss, um den Menschen möglichst gut zu assistieren.

Mit der Corona-Pandemie haben Falschinformationen und verschwörungsideologische Erzählungen einen besonderen Schub erfahren. Die Sachverhalte sind komplex und waren bzw. sind für viele Menschen unbekannt oder unverständlich. Hier setzten und setzen Produzenten und Multiplikatoren von Falschnachrichten und Verschwörungsideologien an, um diese in der Öffentlichkeit zu verbreiten. Spätestens seitdem sind diese in der Gesellschaft sehr präsent. Ein System, das dem Grundkonzept dieser Arbeit folgt, könnte ein Baustein sein, um dieser Entwicklung entgegenzuwirken. Es könnten komplexe Sachverhalte, die im Fernsehprogramm oder in Videoformaten vermittelt werden, in einfacherer Sprache erläutert und aufgeschlüsselt oder mit Hilfe von Faktenchecks überprüft werden.

Seit vielen Jahren sind verschiedene Formen des digitalen Lernens zunehmend in der Gesellschaft verbreitet. Mit der Corona-Pandemie jedoch hat sich die Verbreitung des digitalen Lernens stark beschleunigt und die mediale Nutzung ist auf einem hohen Niveau. In der schulischen und universitären Lehre hat das digitale Lernen im breiten Maßstab Einzug erhalten. Es ist zu erwarten, dass die weite Verbreitung des digitalen Lernens und die gesteigerte mediale Nutzung auch nach der Corona-Pandemie auf einem hohen Niveau bleiben wird.

Der Wandel der Mediennutzung ist ein laufender Prozess und bietet immer neue Möglichkeiten. Die parallele Nutzung mehrerer Endgeräte, wie das Second Screen-Phänomen, ist bereits weit verbreitet. Die Potenziale des Second Screen-Phänomens wurden beispielsweise im Bereich des Marketings für die Werbung erkannt.

Es bleibt abzuwarten, wie sich der Bildungsbereich und die anderen Anwendungsfelder eines solchen Systems entwickeln. Das Second Screen-Phänomen kann für viele positive und wünschenswerte Zwecke genutzt werden. Mit der kontinuierlichen und parallelen Nutzung multipler Endgeräte kann jedoch die Konzentration auf einzelne Geräte abnehmen. Die Toleranz von Langeweile könnte sich verringern, da häufig die Ablenkung von uninteressanten oder spannungsarmen Nutzungsmomenten gewählt wird. In der Zukunft wird sich zeigen, wie sich dies auswirkt und sich die Mediennutzung entwickelt.

7 Literatur

- [1] KIRSCHENMANN, HARALD.: Automatisierte Generierung von Content für Second Screen Anwendungen. In: Grundprojekt Ausarbeitung an HAW-Hamburg (2020) – URL: <https://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2020-proj/kirschenmann.pdf> (Aufgerufen am 10. Januar 2021)
- [2] KIRSCHENMANN, HARALD.: Untersuchung einer Second Screen Anwendung mit automatisierter Generierung von Content. In: Hauptprojekt Ausarbeitung an HAW-Hamburg (2020) – URL https://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2020-proj/kirschenmann_hp.pdf (Aufgerufen am 10. Januar 2021)
- [3] V. LOHMÜLLER AND C. WOLFF: Towards a Comprehensive Definition of Second Screen. In: *Association for Computing Machinery. : Proceedings of Mensch Und Computer 2019.*, 2019, S. 167-177
- [4] BUSCHOW, C., SCHNEIDER, B.: Niedersächsische Landesmedienanstalt (Hrsg.): *Social TV in Deutschland*. Band 30. Leipzig: VISTAS Verlag (2015)
- [5] InteractiveMedia CCSP GmbH (Hrsg.); United Internet Media AG (Hrsg.): *Catch Me If You Can! – Grundlagenstudie zur Multi-Screen-Nutzung*. 2013
- [6] United Internet Media AG (Hrsg.): *Catch Me If You Can! 2.0 – Update Multi-Screen-Grundlagenstudie*. 2015
- [7] Ericsson Consumerlab: Ericsson (Hrsg.): *TV and Media 2016 – The evolving role of TV and media in consumers' everyday lives*. 2016
- [8] JOHNEN, M. & STARK, B.: Wenn der Fernseher nicht mehr ausreicht: Eine empirische Analyse der Second Screen-Nutzung When watching television becomes insufficient: An empirical analysis of second screen usage. In: *Studies in Communication Media* 4 (2015), S. 364-405
- [9] AZENKOT, S. & LEE, N. B.: Exploring the Use of Speech Input by Blind People on Mobile Devices. In: *Association for Computing Machinery. : Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility.*, 2013
- [10] BUTLER, J.; TRAGER, B. & BEHM, B.: Exploration of Automatic Speech Recognition for Deaf and Hard of Hearing Students in Higher Education Classes. In: *Association for Computing Machinery. : The 21st International ACM SIGACCESS Conference on Computers and Accessibility.*, 2019, S. 32-42

- [11] SUCHANEK, F. M.; KASNECI, G. & WEIKUM, G.: Yago: A Core of Semantic Knowledge. In: *Association for Computing Machinery. : Proceedings of the 16th International Conference on World Wide Web.*, 2007, S. 697-706
- [12] WOLFRAM ALPHA. Wolfram|Alpha. <http://www.wolframalpha.com/>. (Aufgerufen am 16. Dezember 2020)
- [13] LEHMANN, J.; ISELE, R.; JAKOB, M.; JENTZSCH, A.; KONTOKOSTAS, D.; MENDES, P. N.; HELLMANN, S.; MORSEY, M.; VAN KLEEF, P.; AUER, S. & BIZER, C.: DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. In: *Semantic Web 6 (2015)*, S. 167-195
- [14] FÄRBER, M.; BARTSCHERER, F.; MENNE, C. & RETTINGER, A.: Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. In: *Semantic Web 9 (2017)*, S. 1-53
- [15] HIRST, G.: *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press, Cambridge, U.K., 1987
- [16] BUNESCU, R. & PASCA, M.: Using Encyclopedic Knowledge for Named entity Disambiguation. In: *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, April 2006
- [17] MIHALCEA, R. & CSOMAI, A.: Wikify! Linking Documents to Encyclopedic Knowledge. In: *Association for Computing Machinery: Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management.*, 2007, S. 233–242.
- [18] MILNE, D. & WITTEN, I. H.: Learning to Link with Wikipedia. In: *Association for Computing Machinery: Proceedings of the 17th ACM Conference on Information and Knowledge Management.*, 2008
- [19] VRANDECIĆ, D. & KRÖTZSCH, M.: Wikidata: A Free Collaborative Knowledgebase. In: *Commun. ACM 57 (2014)*, Nr. 10, S. 78-85
- [20] PELLISSIER TANON, T.; VRANDECIĆ, D.; SCHAFFERT, S.; STEINER, T. & PINTSCHER, L.: From Freebase to Wikidata: The Great Migration. In: *International World Wide Web Conferences Steering Committee : Proceedings of the 25th International Conference on World Wide Web.*, 2016, S. 1419-1428
- [21] AUER, S.; BIZER, C.; KOBILAROV, G.: DBpedia: A Nucleus for a Web of Open Data. In: Springer Berlin Heidelberg (Hrsg.) : *The Semantic Web.*, 2007, S. 722-735
- [22] HOFFART, J.; SUCHANEK, F. M.; BERBERICH, K.; LEWIS-KELHAM, E.; DE MELO, G. & WEIKUM, G.: YAGO2: Exploring and Querying World Knowledge in Time, Space, Context, and Many Languages. In: *Association for Computing Machinery.: Proceedings of the 20th International Conference Companion on World Wide Web.*, 2011, S. 229-232
- [23] HOFFART, J.; SUCHANEK, F. M.; BERBERICH, K. & WEIKUM, G.: YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia. In: *Artif. Intell.* 194 (2013), S. 28-61
- [24] LEHMANN, J.; ISELE, R.; JAKOB, M.; JENTZSCH, A.; KONTOKOSTAS, D.; MENDES, P. N.; HELLMANN, S.; MORSEY, M.; VAN KLEEF, P.; AUER, S. & BIZER, C.: DBpedia - A large-

- scale, multilingual knowledge base extracted from Wikipedia. In: *Semantic Web 6* (2015), Nr. 2, S. 167-195
- [25] ISMAYILOV, A.; KONTOKOSTAS, D.; AUER, S.; LEHMANN, J. & HELLMANN, S.: Wikidata through the Eyes of DBpedia. In: *Semantic Web 9* (2015)
- [26] MENDES, P. N.; JAKOB, M.; GARCA-SILVA, A. & BIZER, C.: DBpedia Spotlight: Shedding Light on the Web of Documents. In: *Association for Computing Machinery. : Proceedings of the 7th International Conference on Semantic Systems.*, 2011, S. 1-8
- [27] EUROPÄISCHE UNION: Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates., 2007. Artikel 3, Nr 2. - https://www.bmu.de/fileadmin/bmu-import/files/pdfs/allgemein/application/pdf/richtlinie_inspire.pdf (Aufgerufen am 16. Dezember 2020)
- [28] GEONAMES. GeoNames. <http://geonames.org/>. (Aufgerufen am 16. Dezember 2020)
- [29] OPENSTREETMAP. OpenStreetMap. <http://www.openstreetmap.org>. (Aufgerufen am 16. Dezember 2020)
- [30] Pre-processing Boosting Twitter Sentiment Analysis?. In: *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*., 2015, S. 748-753
- [31] MANNING, C. D.; RAGHAVAN, P. & SCHÜTZE, H.: Introduction to Information Retrieval, USA: Cambridge University Press., 2008, S. 22-26
- [32] LUHN, H. P.: The Automatic Creation of Literature Abstracts. In: *IBM Journal of Research and Development 2* (1958), Nr. 2, S. 159-165
- [33] LADANI, D. J. & DESAI, N. P.: Stopword Identification and Removal Techniques on TC and IR applications: A Survey. In: *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*., 2020, S. 466-472
- [34] DOLAMIC, L. & SAVOY, J.: When Stopword Lists Make the Difference. In: *Journal of the American Society for Information Science and Technology*, 61 (2010), Nr. 1, S. 200-203
- [35] GHOSH, K. & BHATTACHARYA, A.: Stopword Removal: Why Bother? A Case Study on Verbose Queries. In: *Association for Computing Machinery. : Proceedings of the 10th Annual ACM India Compute Conference.*, 2017, S. 99-102
- [36] GRISHMAN, R. & SUNDHEIM, B.: Design of the MUC-6 Evaluation. In: *Association for Computational Linguistics: Proceedings of the 6th Conference on Message Understanding.*, 1995, S. 1-11
- [37] BIRD, S.; LOPER, E.; KLEIN, E.: Natural Language Processing with Python. In: *O'Reilly Media Inc.*, 2009
- [38] E. LEITNER; G. REHM AND J. SCHNEIDER: Fine-Grained Named Entity Recognition in Legal Documents. In: *Semantic Systems. The Power of AI and Knowledge Graphs.* 2019, S. 272-287
- [39] LI, J.; SUN, A.; HAN, J. & LI, C.: A Survey on Deep Learning for Named Entity Recognition. In: *IEEE Transactions on Knowledge and Data Engineering* (2020)
- [40] Y. LIN; S. YANG; V. STOYANOV AND H. JI: A Multi-lingual Multi-task Architecture for Low-resource Sequence Labeling. In: *Association for Computational Linguistics.* :

- Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018, S. 799-809
- [41]LOPER, E. & BIRD, S.: NLTK: The Natural Language Toolkit. In: *Association for Computational Linguistics. : Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics - Volume 1.*, 2002, S. 63-70
- [42]TJONG KIM SANG, E. F.: Introduction to the CoNLL-2002 Shared Task: Language-Independent Named Entity Recognition. In: *Association for Computational Linguistics. : Proceedings of the 6th Conference on Natural Language Learning - Volume 20.*, 2002, S. 1-4
- [43]HONNIBAL, M. & MONTANI, I.: SpaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing., To appear (2017)
- [44]NOTHMAN, J.; RINGLAND, N.; RADFORD, W.; MURPHY, T. & CURRAN, J.: Learning multilingual named entity recognition from Wikipedia. In: *Artif. Intell.* 194 (2013), S. 151-175
- [45]ABADI, M.; AGARWAL, A.; BARHAM, P.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015
- [46]PASZKE, A.; GROSS, S.; MASSA, F.; LERER, A.; BRADBURY, J.; (Hrsg.): *Curran Associates, Inc.* : PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: *Advances in Neural Information Processing Systems 32.*, 2019, S. 8024-8035
- [47]BALOG, K.: Entity Linking, Cham: *Springer International Publishing.* : *Entity-Oriented Search*, 2018, S. 147-188
- [48]JI, H. & GRISHMAN, R.: Knowledge Base Population: Successful Approaches and Challenges. In: *Association for Computational Linguistics. : Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1.*, 2011, S. 1148 – 1158
- [49]HE, J.; DE RIJKE, M.; SEVENSTER, M.; VAN OMMERING, R. & QIAN, Y.: Generating Links to Background Knowledge: A Case Study Using Narrative Radiology Reports. In: *Association for Computing Machinery. : Proceedings of the 20th ACM International Conference on Information and Knowledge Management.*, 2011, S. 1867 – 1876
- [50]SHEN, W.; WANG, J. & HAN, J.: Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions. In: *IEEE Transactions on Knowledge and Data Engineering* 27 (2015), Nr. 2, S. 443-460
- [51]MIHALCEA, R. & CSOMAI, A.: Wikify! Linking Documents to Encyclopedic Knowledge. In: *Association for Computing Machinery: Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management.*, 2007, S. 233–242.
- [52]MILNE, D. & WITTEN, I. H.: Learning to Link with Wikipedia. In: *Association for Computing Machinery: Proceedings of the 17th ACM Conference on Information and Knowledge Management.*, 2008

- [53] SHNAYDERMAN, I.; EIN-DOR, L.; MASS, Y.; HALFON, A.; SZNAJDER, B.; SPECTOR, A.; KATZ, Y.; SHEINWALD, D.; AHARONOV, R. & SLONIM, N.: Fast End-to-End Wikification. In: *CoRR* (2019)
- [54] LI, B. Z.; MIN, S.; IYER, S.; MEHDAD, Y. & YIH, W.-T.: Efficient One-Pass End-to-End Entity Linking for Questions. In: *Association for Computational Linguistics. : Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*., 2020, S. 6433-6441
- [55] LYMPEROPOULOS, P.; QIU, H. & MIN, B.: Concept Wikification for COVID-19. In: *Association for Computational Linguistics. : Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*., 2020
- [56] OWUSU, C.; LAN, Y.; ZHENG, M.; TANG, W. & DELMELLE, E.: Geocoding Fundamentals and Associated Challenges, 2017, S. 41-62
- [57] MANORUANG, D. & ASAVASUTHIRAKUL, D.: A Tax-Map-Based Address Point Data Model for Geocoding Thai addresses. In: *2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON)*., 2019, S. 225-229
- [58] SRIVASTAVA, V.; TEJASWIN, P.; DHAKAD, L.; KUMAR, M. & DANI, A.: A Geocoding Framework Powered by Delivery Data, New York, NY, USA: *Association for Computing Machinery. : Proceedings of the 28th International Conference on Advances in Geographic Information Systems*., 2020, S. 568-577
- [59] AL-OLIMAT, H. S.; SHALIN, V. L.; THIRUNARAYAN, K. & SAIN, J. P.: Towards Geocoding Spatial Expressions (Vision Paper). In: *Association for Computing Machinery. : Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*., 2019, S. 75-78
- [60] ALEXIS, K.; KAFFES, V.; VARKAS, I.; SYNGROS, A.; TSAKONAS, N. & GIANNOPOULOS, G.: Improving Geocoding Quality via Learning to Integrate Multiple Geocoders. In: *Association for Computing Machinery. : 32nd International Conference on Scientific and Statistical Database Management*., 2020
- [61] DEMIN, IVAN.: Entwicklung einer Plattform für Second Screen Experimente. In: Masterarbeit an HAW-Hamburg (2015) – URL <https://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/master/demin.pdf> (Aufgerufen am 29. August 2021)
- [62] D. ODIJK; E. MEIJ AND M. DE RIJKE: Feeding the Second Screen: Semantic Linking Based on Subtitles. In: *LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE. : Proceedings of the 10th Conference on Open Research Areas in Information Retrieval*., 2013, S. 9-16
- [63] L. T. LI; D. C. G. A. PEDRONETTE; J. ALMEIDA; O. A. B. PENATTI; R. T. CALUMBY AND R. DA S. TORRES: Multimedia Multimodal Geocoding. In: *Association for Computing Machinery. : Proceedings of the 20th International Conference on Advances in Geographic Information Systems*., 2012, S. 474-477

- [64] CAUWENBERGE, A. V.; SCHAAP, G. & VAN ROY, R.: "TV no longer commands our full attention": Effects of second-screen viewing and task relevance on cognitive load and learning from news. In: *Computers in Human Behavior*. 38 (2014), S. 100-109
- [65] SCHAAP, G.; KLEEMANS, M. & VAN CAUWENBERGE, A.: *Second Screening for news: Effects of presentation on information processing and program liking*. In: *Computers in Human Behavior* 84, 2018, Seite 76-85
- [66] SevenOne Media (Hrsg.): *Media Activity Guide 2020 – Trends in der Mediennutzung*. 2020
- [67] ROHLOFF, T.; RENZ, J.; BOTHE, M. & MEINEL, C.: Supporting Multi-Device E-Learning Patterns with Second Screen Mobile Applications. In: *Association for Computing Machinery : Proceedings of the 16th World Conference on Mobile and Contextual Learning*., 2017
- [68] SCHILL, A. & SPRINGER, T.: Verteilte Systeme: Grundlagen und Basistechnologien. In: *eXamen.press* : Springer-Verlag Berlin Heidelberg, 2012, S. 22-26
- [69] WEBVTT. WebVTT. <https://w3c.github.io/webvtt>. (Aufgerufen am 28. Januar 2021)
- [70] HIBSNAUER, C.: Fernsehdokumentarismus. In: GEIMER, A.; HEINZE, C. & WINTER, R. (Hrsg.), Wiesbaden: *Springer Fachmedien Wiesbaden : Handbuch Filmsoziologie*., 2019, S. 1-20
- [71] SCHAAP, G.; KLEEMANS, M. & VAN CAUWENBERGE, A.: *Second Screening for news: Effects of presentation on information processing and program liking*. In: *Computers in Human Behavior* 84, 2018, Seite 76-85
- [72] NOMINATIM. Nominatim. <https://nominatim.org/>. (Aufgerufen am 29. Januar 2021)
- [73] FIELDING, R. T. & TAYLOR, R. N.: Architectural Styles and the Design of Network-Based Software Architectures: *University of California, Irvine*., 2000
- [74] BELSHE, M.; PEON, R. & THOMSON, M.: Hypertext Transfer Protocol Version 2 (HTTP/2), RFC 7540: *RFC Editor*., Nr. 7540, 2015
- [75] FIELDING, R. T.; TAYLOR, R. N.; ERENKRANTZ, J. R.; GORLICK, M. M.; WHITEHEAD, J.; KHARE, R. & OREIZY, P.: Reflections on the REST Architectural Style and "Principled Design of the Modern Web Architecture" (Impact Paper Award). In: *Association for Computing Machinery : Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*., 2017, S. 4-14
- [76] NEWMAN, S.: *Building Microservices: O'Reilly Media, Inc.*, 2015
- [77] ERENKRANTZ, J. R.; GORLICK, M.; SURYANARAYANA, G. & TAYLOR, R. N.: From Representations to Computations: The Evolution of Web Architectures. In: *Association for Computing Machinery : Proceedings of the the 6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on The Foundations of Software Engineering*., 2007, S. 255-264
- [78] APACHE ACTICEMQ. ApacheActiveMQ. <http://activemq.apache.org/>. (Aufgerufen am 25. Februar 2021)
- [79] RABBITMQ. RabbitMQ. <https://www.rabbitmq.com/>. (Aufgerufen am 27. Februar 2021)

- [80] APACHE KAFKA. Apache Kafka. <https://kafka.apache.org/>. (Aufgerufen am 27. Februar 2021)
- [81] IONESCU, V. M.: The analysis of the performance of RabbitMQ and ActiveMQ. In: *2015 14th RoEduNet International Conference - Networking in Education and Research (RoEduNet NER)*., 2015, S. 132-137
- [82] FU, G.; ZHANG, Y. & YU, G.: A Fair Comparison of Message Queuing Systems. In: *IEEE Access* 9 (2021), S. 421-432
- [83] V. LOHMÜLLER; P. EIERMANN; P. ZEITLHOFLER AND C. WOLFF: Attention Guidance in Second Screen Applications. In: *Association for Computing Machinery. : Proceedings of Mensch Und Computer 2019.*, 2019, S. 179-187
- [84] BAEZA-YATES, R. A. & RIBEIRA-NETO, B.: Modern Information Retrieval, USA: Addison-Wesley Longman Publishing Co., Inc., 1999, S. 74-82
- [85] BOGNER, Alexander ; LITTIG, Beate ; MENZ, Wolfgang: *Interviews mit Experten. Eine praxisorientierte Einführung*. Wiesbaden : SPRINGER VS, 2014
- [86] C. HELFFERICH: Leitfaden- und Experteninterviews. In: N. BAUR AND J. BLASIUS (Hrsg.), Wiesbaden: *Springer Fachmedien Wiesbaden. : Handbuch Methoden der empirischen Sozialforschung.*, 2019, S. 669-686
- [87] U. KUCKARTZ: Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. 4. Auflage., Weinheim; Basel: *Beltz Juventa.*, 2018
- [88] MAYRING, PHILIPP: Qualitative Inhaltsanalyse : Grundlagen und Techniken. 12. Auflage., Weinheim; Basel: *Beltz Juventa.*, 2015
- [89] U. KUCKARTZ AND S. RÄDIKER: Fokussierte Interviewanalyse mit MAXQDA: Schritt für Schritt, 2020
- [90] A. BOGNER; B. LITTIG AND W. MENZ: Das Experteninterview. Theorie, Methode, Anwendung, 2002
- [91] A. PRZYBORSKI AND M. WOHLRAB-SAHR: Qualitative Sozialforschung: Ein Arbeitsbuch: *Oldenbourg Wissenschaftsverlag.*, 2014
- [92] R. KAISER: Qualitative Experteninterviews. Konzeptionelle Grundlagen und praktische Durchführung., Wiesbaden: *Springer VS.*, 2014
- [93] S. SCHÖBEL: Exploring Gamification in Digital Learning Environments: *kassel university press.*, 2020
- [94] G. MEY AND K. MRUCK: Handbuch Qualitative Forschung in der Psychologie Band 2: Designs und Verfahren: Band 2: Designs und Verfahren, 2020
- [95] HELFFERICH, Cornelia: *Die Qualität qualitativer Daten: Manual für die Durchführung qualitativer Interviews*. 4. Aufl.. Wiesbaden : VS-Verlag, 2011.
- [96] M. GÖHNER AND M. KRELL: Qualitative Inhaltsanalyse in naturwissenschafts-didaktischer Forschung unter Berücksichtigung von Gütekriterien: Ein Review. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 2020.

H Anhang 1

H.1 Messungen der Vorverarbeitung

H.1.1 Digitale Zukunft

	PER	ORG	MISC	LOC	Summe	Summe ohne doppelte & Sonderzeichen
Händisch	8	8	28	12	56	56
SpaCy	12	0	31	42	85	83
StopWord	-	-	-	-	2311	1112

Tabelle 5: Named Entity Recognition und StopWord Removal als Vorverarbeitungsschritt des DZ-Videos mit 3328 Token

H.1.2 Der große Umbruch

	PER	ORG	MISC	LOC	Summe	Summe ohne doppelte & Sonderzeichen
Händisch	12	13	74	22	121	121
SpaCy	27	0	73	34	134	127
StopWord	-	-	-	-	3526	1489

Tabelle 6: Named Entity Recognition und StopWord Removal als Vorverarbeitungsschritte des DgU-Videos mit 5417 Token

H.1.3 Wann kommt der Corona-Impfstoff?

	PER	ORG	MISC	LOC	Summe	Summe ohne doppelte & Sonderzeichen
Händisch	5	9	38	17	69	69
SpaCy	24	0	38	33	95	94
StopWord	-	-	-	-	2557	1052

Tabelle 7: Named Entity Recognition und StopWord Removal als Vorverarbeitungsschritt des CI-Videos mit 3968 Token

H.2 Messungen des Geocodings

H.2.1 Digitale Zukunft

	Kandidaten	True Positive	False Negative	False Positive	Precision	Recall	F ₁ -score
Referenz	12	12	0	0	1,00	1,00	1,00
Spacy + OSM	20	11	1	9	0,55	0,92	0,69
SpaCy + GeoNames	6	6	6	0	1,00	0,50	0,67
SpaCy + GeoNames Neu	12	11	1	1	0,92	0,92	0,92
StopWord + OSM	760	12	0	748	0,02	1,00	0,03
StopWord + GeoNames	40	7	5	33	0,18	0,58	0,27
StopWord + GeoNames Neu	73	12	0	61	0,16	1,00	0,28

Tabelle 8: Geocoding des DZ-Videos mit verschiedenen Verfahren und Vorverarbeitungsschritten

H.2.2 Der große Umbruch

	Kandidaten	True Positive	False Negative	False Positive	Precision	Recall	F ₁ -score
Referenz	22	22	0	0	1,00	1,00	1,00
Spacy + OSM	30	20	2	10	0,67	0,91	0,77
SpaCy + GeoNames	13	12	10	1	0,92	0,55	0,69
SpaCy + Geonames Neu	20	18,5	3,5	1,5	0,93	0,84	0,88
StopWord + OSM	1053	19	3	1034	0,02	0,86	0,04
StopWord + GeoNames	77	11	11	66	0,14	0,50	0,22
StopWord + GeoNames Neu	104	17	5	87	0,16	0,77	0,27

Tabelle 9: Geocoding des DgU-Videos mit verschiedenen Verfahren und Vorverarbeitungsschritten

H.2.3 Wann kommt der Corona-Impfstoff?

	Kandidaten	True Positive	False Negative	False Positive	Precision	Recall	F ₁ -score
Referenz	17	17	0	0	1,00	1,00	1,00
Spacy + OSM	24	17	0	7	0,71	1,00	0,83
SpaCy + GeoNames	6	6	11	0	1,00	0,35	0,52
SpaCy + GeoNames Neu	18	17	0	1	0,94	1,00	0,97
StopWord + OSM	698	16	1	682	0,02	0,94	0,04
StopWord + GeoNames	49	6	11	43	0,12	0,35	0,18
StopWord + GeoNames Neu	79	17	0	62	0,22	1,00	0,35

Tabelle 10: Geocoding des CI-Videos mit verschiedenen Verfahren und Vorverarbeitungsschritten

H.3 Messungen des Named Entity Linking

H.3.1 Digitale Zukunft

	Anzahl Entities	Found Matches	True Positives	False Negative	False Positive	Precision	Recall	F ₁ -score
Referenz	57	46	46	0	0	1,00	1,00	1,00
SpaCy + Wiki (KeyWord=Link)	83	33	21	25	12	0,64	0,46	0,53
SpaCy + Wiki	83	74	25	21	49	0,34	0,54	0,42
StopWord + Wiki (KeyWord=Link)	1112	255	24	22	231	0,09	0,52	0,16
StopWord + Wiki	1112	828	35	11	793	0,04	0,76	0,08
DBPedia Spotlight, 0.7	-	82	29	17	53	0,35	0,63	0,45
DBPedia Spotlight, 0.9	-	48	24	22	24	0,50	0,52	0,51

Tabelle 11: Named Entity Linking des DZ-Videos mit verschiedenen Verfahren und Vorverarbeitungsschritten

H.3.2 Der große Umbruch

	Anzahl Entities	Found Matches	True Positives	False Negative	False Positive	Precision	Recall	F ₁ -score
Referenz	121	102	102	0	0	1,00	1,00	1,00
SpaCy + Wiki - (KeyWord=Link)	127	44	31	71	13	0,70	0,30	0,42
SpaCy + Wiki	127	106	49	53	57	0,46	0,48	0,47
StopWord + Wiki (KeyWord=Link)	1489	332	41	61	291	0,12	0,40	0,19
StopWord + Wiki	1489	1085	70	32	1015	0,06	0,69	0,12
DBPedia Spotlight, 0.7	-	117	52	50	65	0,44	0,51	0,47
DBPedia Spotlight, 0.9	-	74	41	61	33	0,55	0,40	0,47

Tabelle 12: Named Entity Linking des DgU-Videos mit verschiedenen Verfahren und Vorverarbeitungsschritten

H.3.3 Wann kommt der Corona-Impfstoff?

	Anzahl Entities	Found Matches	True Positives	False Negative	False Positive	Precision	Recall	F ₁ -score
Referenz	69	64	64	0	0	1,00	1,00	1,00
SpaCy + Wiki - (KeyWord=Link)	94	30	22	42	8	0,73	0,34	0,47
SpaCy + Wiki	94	72	33	31	39	0,46	0,52	0,49
StopWord + Wiki (KeyWord=Link)	1052	222	39	25	183	0,18	0,61	0,27
StopWord + Wiki	1052	769	58	6	711	0,08	0,91	0,14
DBPedia Spotlight, 0.7	-	69	38	26	31	0,55	0,59	0,57
DBPedia Spotlight, 0.9	-	46	32	32	14	0,70	0,50	0,58

Tabelle 13: Named Entity Linking des CI-Videos mit verschiedenen Verfahren und Vorverarbeitungsschritten

H.4 Interviewleitfaden

Art des Interviews: Leitfadengestütztes Experteninterview

Dauer: maximal 1 Stunde, Einzelinterviews

Experten: Informatiker, Journalist, Lehrerinnen, Dozentin

- B1 (Informatiker, männlich, 30+ Jahre)
- B2 (Person aus dem Marketing-Bereich mit Informatik-Hintergrund, männlich, 40+ Jahre)
- B3 (Lehrerin, weiblich, 40+ Jahre)
- B4 (Lehrerin, weiblich, 40+ Jahre)
- B5 (Dozentin, weiblich, 30 Jahre)
- B6 (Freier Journalist, männlich, 30+ Jahre)

Forschungsthese:

Die Darstellung von Informationsergänzungen auf einem Second Screen fördert das Verständnis und Erlebnis von Fernsehdokumentationen.

Informationsphase:

- Dank für die Bereitschaft mitzumachen

-
- Ziel und Ablauf der Befragung sowie Erklärung der vertraulichen Handhabung der Daten
 - Kurzvorstellung der entwickelten Software

Themenblöcke:

1. Vorerfahrung und Erwartungshaltung
2. Handhabung eines Second Screens
3. Bewertung der Informationen
4. Einsatzszenarien

Einstiegsfragen:

- Können Sie sich kurz vorstellen?

1. Vorerfahrung und Erwartungshaltung:

- Welche Erfahrungen haben Sie bereits mit einem Second Screen gemacht?
 - Nutzen Sie einen Second Screen in ihrem Alltag, während Sie fernsehen oder einer ähnlichen Tätigkeit nachgehen?
 - In welchen Situationen nutzen Sie einen Second Screen?
- Was erwarten Sie von einem System, das ergänzende Informationen zu Fernsehdokumentationen anbietet?
 - Was für ergänzende Informationen erwarten Sie?
 - Wie stellen Sie sich die Wiedergabe bzw. Darstellung der Informationen vor?

Erkundung der entwickelten Software:

- Die Befragten wenden die entwickelte Software selbstständig an. Der Interviewende steht für Nachfragen zur Verfügung.

2. Handhabung eines Second Screens

- Beschreiben Sie die Handhabung eines zweiten Bildschirms während der Betrachtung des Videos.
- Wie beschreiben Sie ihre Aufmerksamkeit während der parallelen Nutzung der Geräte?
- Auf welche Weise würden Sie die Informationsergänzungen an die Anwender übermitteln?

3. **Bewertung der Informationen**

- Welche ergänzenden Informationen waren sinnvoll?
- Welche Aspekte des Systems waren für Sie überraschend?
- Wie empfanden Sie die Menge der Informationen?
- Welche Informationen haben gefehlt oder wären eine gute Ergänzung gewesen?
- Wie bewerten sie das System im Hinblick auf Ihre zuvor geäußerten Erwartungen?
- Wie förderten die Informationen Ihr Verständnis der Fernsehdokumentation?
- Wie bewerten Sie die automatisierte Generierung der ergänzenden Informationen?

4. **Einsatzszenarien**

- In welchen Anwendungskontexten können Sie sich den Einsatz eines Systems zur Informationsergänzung von Fernsehdokumentationen vorstellen?
- Welche Aspekte des Systems würden Sie dafür ändern?

Abschlussfrage:

- Halten Sie das vorgestellte System fördernd für das Verständnis und das Erlebnis von Fernsehdokumentationen? Was bewegt Sie zu dieser Einschätzung?
- Möchten Sie noch einige Punkte ergänzen?

H.5 Kurztitel und stichpunktartige Zusammenfassung

H.5.1 Interview - B1

B1 - Studierter Informatiker (Ba) und IT-Administrator

- Keine Nutzung von Second Screen-Anwendungen im Vorfeld. Nur die parallele thematisch unabhängige Nutzung mehrerer Geräte.
- Erwartet im Vorfeld, dass die Informationen bei Bedarf zur Verfügung stehen, filterbar sind und nicht aufdringlich von der Dokumentation ablenken.
- Erwartet im Vorfeld, dass die Informationen in Textform folgende Aspekte enthalten: Orte, Namen, Personenhistorien, Fakten und weitere wichtige Dinge.
- Zu Beginn der Anwendung der Software war es verwirrend und die Bedienung hat abgelenkt, später hat alles funktioniert.
- Einige Informationen waren überflüssig und einige haben gefehlt. Die interaktiven Karten und weitere Informationen waren sinnvoll, es gab keine falschen Informationen. Viele Informationen waren zu allgemein, eine journalistische Perspektive dieser wäre gut gewesen.
- Die Menge und die Erkennung der Informationen waren okay.
- Die Informationen hätten kompakter, prägnanter und visueller in Kategorien, mit extra Icons zu den Kategorien, dargestellt werden sollen. So könnten die Schlagworte mit den Icons dargestellt werden und nach Auswahl zur Anzeige der Informationen expandiert werden.
- Die Informationen könnten als Schlagworte zum Markieren dargestellt werden, um die markierten Informationen nachträglich zu betrachten.
- Das Verständnis der Dokumentation wurde leicht gefördert, das nachträgliche Vertiefen könnte das Verständnis weiter verbessern.
- Das System könnte ebenfalls im Kontext von Vorlesungen oder Vorträgen verwendet werden, um diese visuell und mit weiterem Hintergrundwissen anzureichern.
- Es ist wichtig, dass die Informationen für die spätere Verwendung gespeichert werden. Die Nachbereitung ist ein wichtiger Faktor.

H.5.2 Interview - B2

B2 - Angestellter im Marketing-Bereich mit IT-Background

- Hat einen Second Screen für Spiele und parallel zum Fernsehkonsum genutzt, ist aber kein regelmäßiger Second Screen-Nutzer.
- Die erwarteten Informationen sind abhängig von der Dokumentation. Im Vorfeld erwartet wurden: Basisinformationen und Hintergrundwissen zum Thema, Geschichtliches, weitere themenbezogene Dokumentationen, Informationen zu Schauspielern, für Marketingzwecke Möglichkeiten zum Buchen von Reisen, ...
- Im Vorfeld wird erwartet, dass die Informationen mit großen Buttons zur Verfügung stehen. Durch die Auswahl dieser kann tiefer in die Information navigiert werden. Des Weiteren soll es eine Suchmaske geben.
- Empfand die Dokumentation als spannend genug, so dass die Zeit für die Nutzung des Second Screens fehlte. Nur die Überschriften wurden betrachtet, diese entsprachen gewünschten Informationen. Fällt die Aufmerksamkeit auf den Second Screen, so konnte B2 dem Inhalt auf dem First Screen nicht folgen.
- Die parallele Nutzung anderer Medien und Ablenkung vom Konsum von Dokumentationen wird von B2 als störend empfunden. Eine solche Anwendung als Nachschlagewerk im Anschluss wird gewünscht.
- Das manuelle Scrollen für neue Informationen lenkt ab, Training könnte dieses verbessern. Wenn es möglich sein soll zu scrollen, so sollte automatisiert zu den neuesten Informationen gesprungen werden.
- Es wurde zu viel Text gezeigt: es ist nicht möglich, parallel zur Dokumentation alles zu lesen. Es sollten kurze Texte verwendet werden, die erweitert werden können.
- Anstelle des Textes sollten die Schlagwörter in Kacheln dargestellt werden. Die Kacheln können erweitert werden und wiederum Kacheln oder Informationen enthalten, so dass die Anwender hierarchisch durch die Kacheln bzw. Informationen navigieren können.
- Es gibt viele uninteressante Informationen. Die Auswahl der Schlagwörter ist gut, aber die dargestellten Informationen zu den Schlagwörtern entsprachen nicht den interessantesten Aspekten.
- Das Genre der Videos, die Art der Nutzung und die Zielgruppe entscheiden darüber, welche Informationen sinnvoll sind.
- Die angebotenen Informationen haben das Verständnis der Fernsehdokumentation gefördert; diese war jedoch in sich schlüssig genug.
- Die redaktionelle Bearbeitung der Informationen ist sinnvoll, um einen besseren thematischen Zusammenhang mit der Dokumentation zu bieten. Die automatisierte Generierung der Informationen ist ebenfalls möglich.

-
- Ein solches System könnte als Historie bzw. eine Art digitaler Notizblock verwendet werden, in dem die Informationen nachgeschlagen werden können. Hierzu sollte es eine Suchfunktion geben.
 - Die Nutzung eines Second Screens parallel vom Konsum weiterer Medien ist interessant.
 - Ein Second Screen könnte genutzt werden, um Dokumentationen oder andere Videos interaktiv zu gestalten.

H.5.3 Interview - B3

B3 - Gymnasiallehrerin, die vor Corona tätig war

- Hat einen Second Screen nur im privaten Bereich und selten genutzt.
- Die Anwendung soll ein neutrales Bild geben durch die Darstellung fachbezogener Informationen und verschiedener Meinungen.
- Die Art der gewünschten Zusatzinformationen ist abhängig von der Art der Dokumentation.
- Darstellung der Informationen auf einem Second Screen; Darstellung über einen Split Screen kann ebenfalls sinnvoll sein.
- Es sollen mehrere Perspektiven auf die Dokumentation geboten werden. SuS können sich eine Perspektive auswählen.
- Die Benutzung eines Second Screens war ungewohnt. Da es Schwierigkeiten mit dem Multitasking gab, wurden Pausen benötigt.
- B3 geht davon aus, dass es sich einspielen würde, dass die Informationen nachträglich statt parallel betrachtet werden.
- Es gab Informationen, die falsch oder nur punktuell gut zugeordnet waren.
- Es ist gut, dass die Informationen als Schlagworte, als Überschrift und kurze knappe Texte dargestellt und mit Bildern angereichert werden.
- Die Menge der Informationen ist richtig
- Die Auswahl der Schlagwörter ist gut und diese entsprechen den gewünschten Informationen. Es gibt aber nicht immer die erwarteten Informationen.
- Die synchronisierte Darstellung der Schlagwörter ist gut, wenngleich es eine zeitliche Verzögerung gegeben hat.
- Es fehlten einige gewünschte Informationen und es wurden nicht mehrere unterschiedliche Meinungen abgebildet.
- Bei mehr Informationen sollte nur die Liste der Schlagwörter angezeigt werden, die sich für weitere Informationen erweitern lassen.
- Im Hinblick auf die Erwartungen war das System gut. Es fördert das Verständnis und erweitert das Wissen zur Dokumentation.

- Die automatisierte Generierung der Informationen funktioniert sehr gut. Es gibt jedoch einige überflüssige oder fehlerhafte Informationen.
- Eine redaktionelle Bearbeitung der Informationen müsste transparent gemacht werden; die automatisierte Generierung wird als neutraler betrachtet.
- Im Bildungsbereich oder in der Erwachsenenbildung kann die Anwendung genutzt werden. Die Anwender können die angebotenen Informationen entsprechend ihrer Vorerfahrung nutzen oder überspringen.
- Im schulischen Bereich könnten SuS Dokumentationsthemen multiperspektivisch betrachten. Eine Bearbeitungsmöglichkeit der Informationen für Lehrer wäre sinnvoll.
- Das System ist gut, um sich vertiefend mit Dokumentationen auseinanderzusetzen, auch wenn einige Dinge fehlten oder überflüssig sind.

H.5.4 Interview - B4

B4 - Gymnasiallehrerin, die während Corona tätig war

- Im Privaten wurde ein Second Screen zum Informieren über das Fernsehprogramm genutzt. Beruflich wurden Second Screens von SuS im Rahmen des Unterrichts genutzt.
- Erwartet im Vorfeld, dass (geschichtliche) Hintergründe, Daten, Fakten und Erläuterungen zu Themen schnell geliefert werden.
- Erwartet im Vorfeld, dass die Informationen über eine Suchmaske zur Verfügung gestellt werden.
- Die Handhabung und Bedienung der Anwendung war sehr leicht. Es wurde jedoch Zeit zur Eingewöhnung benötigt.
- Es ist gut, dass die Informationen ohne aktive Eingaben angeboten werden.
- Die ergänzenden Informationen waren sinnvoll. Es gab auch überflüssige und falsch zugeordnete Informationen.
- Es überraschte, dass die Informationen parallel zur Dokumentation auf dem Second Screen erschienen.
- Die Menge und Länge der Informationen waren gut, auch dass diese mit Bildern und Karten angereichert wurden.
- Es fehlten keine besonderen Informationen, die Auswahl der Schlagwörter war gut. Es hätte Interesse für weitere Informationen bestanden, dies hätte jedoch den Rahmen gesprengt.
- Die Anwendung war besser als zuvor erwartet, da die Informationen ohne aktive Eingaben angeboten wurden. Auch für die Aufmerksamkeit ist dies besser.

- Die Informationen förderten das Verständnis der Dokumentation, da Sachverhalte dazugelernt wurden, auch solche, nach denen, ohne die Anwendung, nicht gesucht worden wäre.
- Die automatisierte Generierung der Informationen funktioniert und die Auswahl der Schlagwörter ist gut.
- Die redaktionelle Bearbeitung der Informationen wäre besser. Es gibt jedoch zu viele Dokumentationen, als dass es praktikabel sei.
- Das System könnte im privaten oder im beruflichen Kontext angewendet werden. Seit Corona wird im schulischen Kontext vermehrt mit digitalen Medien und Videos gearbeitet, die die SuS zu Hause betrachten.
- Ein solches System kann den Sprachunterricht unterstützen, indem benötigte Vokabeln angeboten werden oder bei Übersetzungen unterstützt wird. Auch Menschen mit Sehbehinderung oder Hörgeschädigte könnten unterstützt werden.
- Die Informationen sollten gespeichert werden, so dass diese später nachgelesen werden können.

H.5.5 Interview - B5

B5 - Wissenschaftliche Mitarbeiterin an einer Universität

- B5 hat einen Second Screen im Privaten zur Informationssuche oder Unterhaltung parallel zum Fernsehprogramm verwendet. Im Arbeitskontext wurde ein Second Screen zur Kommunikation parallel zu Videokonferenzen genutzt.
- Im Vorfeld werden Informationen zum Erweitern des Hintergrundwissens und nicht ausreichend vermittelter Inhalte erwartet. Außerdem sollen Informationen zu gezeigten Schauspieler:innen und zur Produktion des Videos zur Verfügung gestellt werden.
- Die Informationen sollen die Themen der Dokumentation aufgreifen, das Gezeigte in historische Kontexte und Debatten einordnen.
- Im Vorfeld wird erwartet, dass die Informationen bei Bedarf abgerufen werden und über Kategorien gefiltert werden können. Die Darstellung soll nicht zu visuell gestaltet werden, damit der Fokus auf dem First Screen liegt.
- Für B5 war die Auswahl und die Menge der Informationen gut. Wenn langsamer gelesen wird oder eine Information viel Aufmerksamkeit beansprucht, so kann der Dokumentation für einen Moment nicht mehr gefolgt werden.
- Bei der automatisierten Generierung der Inhalte wird befürchtet, dass Menschen tiefer in Filterblasen und Echokammern gelangen könnten. Die Art der Wissensbeschaffung könnte sich ändern.
- Anfangs war die Aufmerksamkeit darauf ausgerichtet, sich mit dem System zurechtzufinden. Nach kurzer Zeit erfolgte die Gewöhnung und die Handhabung war gut.

- Die Anwendung kann gut im Bildungsbereich zur Wissensvermittlung genutzt werden. Mit interaktiven Fragen könnte Wissen gefestigt werden.
- Die Verwendung eines Second Screens auch im Hinblick auf die Wahrnehmungsweise, mit der Wissen gut aufgenommen werden kann, gut. Die Wahrnehmungsweise ist generationsbedingt unterschiedlich.
- Es gab fehlende, falsch zugeordnete oder uninteressante, aber auch hilfreiche Informationen. Wie sinnvoll Informationen sind, ist vom Anwender abhängig.
- Es wurde zu viel Inhalt geboten, kurze stichpunktartige Informationen wären sinnvoll.
- Die Menge der Informationen war okay; mehr Informationen sollten es nicht sein. Viele Personen können von der Menge an Informationen zu sehr abgelenkt werden.
- Die Informationen haben das Verständnis der Dokumentation gefördert. Aufgrund der doppelten Aufmerksamkeit können die Informationen nachhaltiger gespeichert werden.
- Bei der automatisierten Generierung der Inhalte werden gesellschaftliche Probleme gesehen. Daher wird die redaktionelle Erstellung der Inhalte gewünscht.

H.5.6 Interview - B6

B6 - Freier Journalist

- B6 nutzte einen Second Screen zur Informationssuche während des Konsums von Dokumentationen, Filmen etc., wenn Fragen oder Unklarheiten aufkamen. Dies geschieht nicht regelmäßig.
- Das System soll benutzerfreundlich, einfach, intuitiv und flüssig sein. Während der Nutzung darf es keine technischen Schwierigkeiten geben.
- Mit den Informationen sollen offene Fragen kurz und prägnant beantwortet werden.
- Abhängig von der Information sollten diese mit einem Bild visualisiert werden.
- Die Art der gewünschten Informationen ist abhängig von dem Suchbegriff bzw. der Thematik. Für die Informationen könnten mehrere Quellen angeboten werden.
- Bei kurzen Informationen reicht die Anzeige der Fakten. Längere und komplexere Informationen könnten zur späteren Betrachtung markiert werden.
- Das vorgestellte System kam sehr gut an. Es gibt Punkte, die verbessert werden könnten. Die Formatierung und das Design könnten angepasst werden und das manuelle Scrollen führt zu Schwierigkeiten.
- Einige Schlagwörter sind irrelevant für den Beitrag.
- Die Aufmerksamkeit ändert sich während der Erstbenutzung, die Handhabung wird flüssiger. Bei der Erstbenutzung ist der Anwender abgelenkt; nach einer Gewöhnungsphase sollte die Software eine gute Alternative zur Informationsgewinnung sein.

-
- Maximal zwei Sätze sollten zu jeder Information angezeigt werden. Zur parallelen Erfassung müssen die Informationen kurz und prägnant sein.
 - Es ist vom Anwender abhängig, welche Informationen sinnvoll sind.
 - Die geographische Einordnung und Kartendarstellung waren äußerst hilfreich. Es war gut, dass diese mit einem Text ergänzt wurden.
 - Es gab zu viele Informationen, einige Informationen waren überflüssig, andere fehlten.
 - Die Anwendung erfüllt den angekündigten Zweck, für eine Marktreife müssen Verbesserungen erfolgen.
 - Das Verständnis der Dokumentation wurde gefördert, da zusätzliche Informationen angeboten und mit den geographischen Informationen ein schneller Überblick gewonnen werden konnte.
 - Das Eigeninteresse beeinflusst den Erkenntnisgewinn. Die Informationen sollten kurz und prägnant sein, um Informationen schnell zu erfassen und Uninteressantes zu überspringen.
 - Es ist faszinierend, dass eine solche Anwendung möglich ist. Die Anwendung kann hilfreich sein und eine Zeitersparnis bieten.
 - Die redaktionelle Bearbeitung kann zu besseren und gezielteren Informationen führen, das Bedürfnis der Individuen ist jedoch unbekannt.
 - Die Anwendung kann im schulischen Kontext genutzt werden und interessant zur Nachrecherche und Referatsvorbereitung von SuS sein.

Erklärung zur selbstständigen Bearbeitung einer Abschlussarbeit

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich gemacht.

Ort

Datum

Unterschrift im Original