Abschlussbericht – Teil II – Eingehende Darstellung

Abteilung Computational Humanities, Institut für Informatik, Universität Leipzig
Förderkennzeichen: **16KIS1517**

Vorhabenbezeichnung: **Verbundprojekt: Narrative der Desinformation verstehen: ein Vergleich zwischen öffentlich-rechtlichen und alternativen Nachrichtenvideos – FakeNarratives –; Teilvorhaben: Digital Humanities-Ansätze für ein Scalable Viewing**

Laufzeit des Vorhabens: **01.10.2021 – 31.12.2024**

Inhaltsverzeichnis

1. Projektkontext und Ziele 2
2. Stand der Wissenschaft: Videoanalyse in den Digital Humanities und Scalable Viewing 2
3. Verwendung der Zuwendung 3
4. Erzielte Ergebnisse 4
   AB 1 – Korpusauswahl und Formalisierung von Narrativitätsstrategien (Lead: ULpz) 4
      AP 1.1 Aufbau der Datenkorpora 4
      AP 1.2 Formalisierung von Narrativitätsstrategien 5
   AB 2 – Multimodale Informationsextraktion (Lead: LS3) 7
   AB 3 – Erkennung von Narrativitätsstrategien (Lead: UBre) 8
      AP 3.1 Ableiten von Narrativitätsstrategien und konkreten Merkmalen aus der Multimodalitätstheorie für den Bereich „Nachrichtenvideos“ 8
      AP 3.2 Semi-automatische Erkennung von Narrativitätsstrategien 9
      AP 3.3 Überprüfung der intermodalen Konsistenz von Informationen 10
   AB 4 – Explorative Datenanalyse mittels Scalable Viewing (Lead: ULpz) 10
      AP 4.1 Close Viewing 10
      AP 4.2 Distant Viewing 12
      AP 4.3 Scalable Viewing 15
   AB 5 – Kontrastive Analyse audiovisueller narrativer Strategien (Lead: UBre) 17
      AP 5.1 Kontrastive Analyse audiovisueller narrativer Strategien 17
      AP 5.2 Identifikation von narrativen Lücken 18
      AP 5.3 Visualisierung und Ausformulierung von Gegenbotschaften zu Desinformation 18
6. Ergebnisverwertung 18
7. Literaturverzeichnis 19

1. Projektkontext und Ziele

Irreführende und manipulative Informationen verbreiten sich über soziale Netzwerke und virale Mechanismen mit einer enormen Geschwindigkeit und erreichen so ein großes Publikum. Eine zentrale Erkenntnis der neueren Forschung (Allen et al., 2020) ist jedoch, dass nicht nur die alternative Nachrichtenberichterstattung zur öffentlichen Desinformation beiträgt, sondern vor allem auch seriöse, öffentlich-rechtliche Nachrichtenformate Erzählmuster enthalten können, die desinformierend wirken und damit die öffentliche Meinung maßgeblich negativ beeinflussen. Wie die Studie von Gräf und Hennig (2020) zeigt, kann bereits die selektive Auswahl und Präsentation von Informationen in eigentlich seriösen Formaten zu unbeabsichtigter Desinformation führen. Dies geschieht unter anderem durch die Praxis des sogenannten 'false balancing' in den öffentlich-rechtlichen Medien (Brüggemann & Engesser, 2017), bei dem aus Gründen der ‚Objektivität' und ‚Ausgewogenheit' allen Akteuren, egal ob sie seriöse oder unseriöse Informationen verbreiten, die gleiche Redezeit eingeräumt wird. Narrative Strukturen der Desinformation in öffentlich-rechtlichen Nachrichten sind hochgradig gefährlich, weil sie bei Teilen des Publikums ein grundsätzliches Misstrauen schüren und sie damit in die Arme alternativer, bewusst irreführender Informationskanäle treiben. Es stellen sich daher kritische Fragen nach Ähnlichkeiten und Unterschieden zwischen öffentlich-rechtlichen Nachrichtenquellen und alternativen Nachrichtenkanälen der gezielten Desinformation, die in diesem Projekt untersucht werden sollen.

Ziele des Teilprojekts der Universität Leipzig (ULpz): Ein zentrales Ziel des Teilprojekts ULpz liegt bei diesem interdisziplinären Unterfangen in der Vermittlung zwischen Multimodalitätstheorie (Teilprojekt Universität Bremen – UBre) und Informatik (Teilprojekt Forschungszentrum L3S), ganz im Sinne der Digital Humanities. Insbesondere werden dabei Ansätze der automatisierten Erkennung und Analyse von Text, Bild, Audio und Video auf die diskursive Ebene des *distant viewing* übersetzt, wo sie dann hinsichtlich narrativer Strategien untersucht werden können. Eine weitere Zielsetzung besteht in der Entwicklung explorativer Analysetools und insbesondere skalierbarer Visualisierungen, um die kontrastive Interpretation und allgemein die Deutung der automatisch identifizierten Narrativitätsmuster im Sinne eines *scalable viewing*-Ansatzes überhaupt erst zu ermöglichen.

2. Stand der Wissenschaft: Videoanalyse in den Digital Humanities und Scalable Viewing

Derzeit erleben wir in den Digital Humanities ein verstärktes Interesse für nicht-textuelle Medientypen, zuletzt für Film und Video (Ross et al., 2009; Sayers, 2018). Traditionell haben empirische Filmstudien mit "Einstellungen" (shots) als kleinster messbarer Einheit gearbeitet (Salt, 2006), d.h. sie haben die Längen und Häufigkeiten verschiedener Einstellungen in Filmen gemessen und berechnet (Salt, 1974; Tsivian, 2009). In jüngerer Zeit wurde auch über andere berechenbare Merkmale von Filmen nachgedacht (Rasheed et al., 2005), z. B. die Verteilung von Farben, Bewegungsmuster, Aspekte der Beleuchtung oder die Sprache von Filmen mittels Untertitel (Burghardt et al., 2016; Everingham et al., 2006; Cour et al., 2008).

Angelehnt an Designprinzipien aus dem Feld der Visual Analytics (Keim et al., 2008), geht man in der jüngeren Forschung zur computergestützten Film- und Videoanalyse davon aus, dass Visualisierungen dabei helfen können, bestimmte Aspekte von Filmen in ihre elementaren Bestandteile zu zerlegen und somit interessante Muster und Trends in einem größeren Maßstab aufzudecken (Manovich, 2013). Dieser größere Maßstab wurde in Analogie zum Konzept des *distant reading* (vs. *close reading*) zuletzt als *distant viewing* bezeichnet (Arnold & Tilton, 2019). In den letzten Jahren wurden zahlreiche Tools für die Analyse und Visualisierung von Filmdaten vorgeschlagen (für eine Übersicht siehe Schoeffmann et al., 2015). Während die meisten der existierenden Werkzeuge und Visualisierungen von der Informatik-Community entwickelt wurden, gibt es auch eine wachsende Anzahl von Visualisierungen mit einem dezidierten Fokus auf Digital Humanities. Lev Manovich (2013) etwa stellt eine Reihe von Werkzeugen zur Verfügung, die sich stark auf die Visualisierung als explorative Schnittstelle zu Dziga Vertovs Filmen stützt. *Script-Threads* (Hoyt et al., 2014) ist ein weiteres Tool, das zur Visualisierung von Beziehungen zwischen den Figuren eines Films verwendet werden kann. Andere Tools berücksichtigen die Beziehung von dominanten Farben und Sentiment-Informationen, die aus Filmuntertiteln abgeleitet werden (Burghardt et al., 2016; Hohman et al., 2017).

Der in diesem Projekt vorgeschlagene Ansatz des *scalable viewing* (Burghardt et al., 2018; Burghardt et al., 2019) erweitert das Paradima des *distant viewing* wesentlich, indem er ein dynamisches Wechseln zwischen Detail- und Distanzperspektive erlaubt. Der Ansatz orientiert sich an bestehenden Techniken der *visual analytics*, inbesondere der *visual movie analytics* (Kurzhals et al., 2016). Gleichzeitig werden ganz neue Methoden und Visualisierungen für die Darstellung von Narrativitätsstrategien entwickelt und evaluiert, die wesentlich zum aktuellen Forschungsstand beitragen sollen.

3. Verwendung der Zuwendung

Die Zuwendung wurde hauptsächlich zur Finanzierung einer wissenschaftlichen Mitarbeiterstelle in der Abteilung Computational Humanities am Standort Leipzig verwendet. Die eingesetzte Person war über die gesamte Laufzeit hinweg zentral verantwortlich für den Aufbau des Videokorpus. Die Hauptaufgabe des wissenschaftlichen Mitarbeiters lag in der Entwicklung von Visualisierungs- und Analyseansätzen im Sinne eines *scalable viewing*-Ansatzes. Weitere Mittel wurden für Reisekosten eingesetzt. Diese umfassten zum einen die regelmäßige Teilnahme an standortübergreifenden Projekttreffen im Verbund, zum anderen die Präsentation von Projektergebnissen auf nationalen und internationalen Konferenzen, bei denen einschlägige Fachöffentlichkeiten adressiert wurden. Alle Mittel wurden sachgerecht, zweckgebunden und entsprechend der im Zuwendungsantrag geplanten Verwendung eingesetzt (vgl. Abb. 1). Die zentralen Zielsetzungen des Teilprojekts konnten mit den verfügbaren Ressourcen planmäßig und erfolgreich erreicht werden. Eine detaillierte Darstellung der Arbeitsergebnisse anhand der Arbeitspaketstruktur folgt im nächsten Abschnitt des Berichts.

	2021			2022													2023											2024										
Arbeitsbereich (AB) und Lead (unten); **Projektmonat** (rechts)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
AB1: Korpusauswahl und Formalisierung von Narrativitätsstrategien (Lead: ULpz)	1	1	2	2	2																																		
AB2: Multimodale Informationsextraktion (Lead: L3S)					2	3																																	
AB3: Erkennung von Narrativitätsstrategien (Lead: UBre)							1	1	1		1	1	1	2			2	2			2	2	2		2														
AB4: Explorative Datenanalyse mittels Scalable Viewing (Lead: ULpz)							1	2		1	1	1	1		2	1		2	2	2	2		3	2				3	3	3	3	3	3	3	3	3	3	3	3
AB5: Kontrastive Analyse audiovisueller narrativer Strategien (Lead: UBre)																1	1						2	1	1	1	2	2		2	2	3		3	3	3	3	3	3

Abbildung 1: Arbeitsplan des Teilprojekts ULpz – Die Zahlen in den einzelnen Monatsscheiben repräsentieren das jeweilige Arbeitspaket (AP) im übergeordneten Arbeitsbereich (AB).

4. Erzielte Ergebnisse

Das Arbeitsprogramm umfasst insgesamt fünf große Arbeitsbereiche (AB), die jeweils mehrere Arbeitspakete (AP) beinhalten und nachfolgend dargestellt sind. Jeder AB wird federführend von jeweils einem Teilvorhaben gelenkt und in Kooperation mit den anderen Verbundpartnern bearbeitet. Die Ergebnisse der ULpz je Arbeitspaket sind nachfolgend beschrieben.

AB 1 – Korpusauswahl und Formalisierung von Narrativitätsstrategien (Lead: ULpz)

AP 1.1 Aufbau der Datenkorpora

In diesem AP war die Hauptaufgabe des Teilvorhabens ULpz der Aufbau zweier Subkorpora mit jeweils repräsentativen Video-Samples für die Bereiche „etablierte Fernsehnachrichten“ und „alternative Online-Nachrichtenvideos“. Für das Subkorpus der etablierten Nachrichten wurden Videos aus “Tagesschau”[1] und “ZDF Heute Nachrichten” verwendet. Das Subkorpus „alternativer Nachrichtenvideos“ beinhaltet Videos der Kanäle “Welt”, “BildTV” und “CompactTV”. Abweichend vom ursprünglichen Zeitplan wurde der Task der Korpuserstellung nicht vollständig zu Beginn des Projekts durchgeführt, sondern vielmehr iterativ während der Projektlaufzeit aufgebaut, um auch aktuellen politischen Entwicklungen, wie bspw. dem Russland-Ukraine-Konflikt, Rechnung tragen zu können. Das finale Korpus zu Projektende umfasst 1.065 Videos für den Zeitraum 03/2022 bis 03/2024. Tabelle 1 gibt einen Überblick zur Anzahl an Videos je Nachrichtensendung sowie auch grundlegende Angaben zur Gesamtdauer aller Videos sowie zur durchschnittlichen Videolänge und zum genauen Zeitraum, der durch die Videos abgedeckt wird.

[1] Die Tagesschau-Videos wurden vom Verbundpartner L3S zur Verfügung gestellt.

channel	videos	total_duration	avg_duration	min_time	max_time
BILD TV	223	41:54:22	11:16	2022-01-03	2024-03-27
CompactTV	292	76:06:50	15:38	2022-01-03	2024-03-27
Tagesschau	235	60:52:12	15:32	2022-01-03	2024-03-27
WELT	133	16:35:14	07:28	2022-01-03	2024-03-13
ZDFheute Nachrichten	182	52:57:04	17:27	2022-01-05	2024-03-27

Tabelle 1: Übersicht zum Videokorpus, bestehend aus insgesamt 1.065 Videos aus fünf unterschiedlichen Nachrichtenkanälen.

Abb. 2 zeigt überblicksartig, welche Themen in den Videos abgedeckt werden. Die Themen wurden aus den Hashtags in der Beschreibung von auf YouTube bereitgestellten Videos ermittelt[2]. Darin enthalten sind bspw. der Russland-Ukraine-Konflikt oder die Corona-Politik der Bundesregierung.

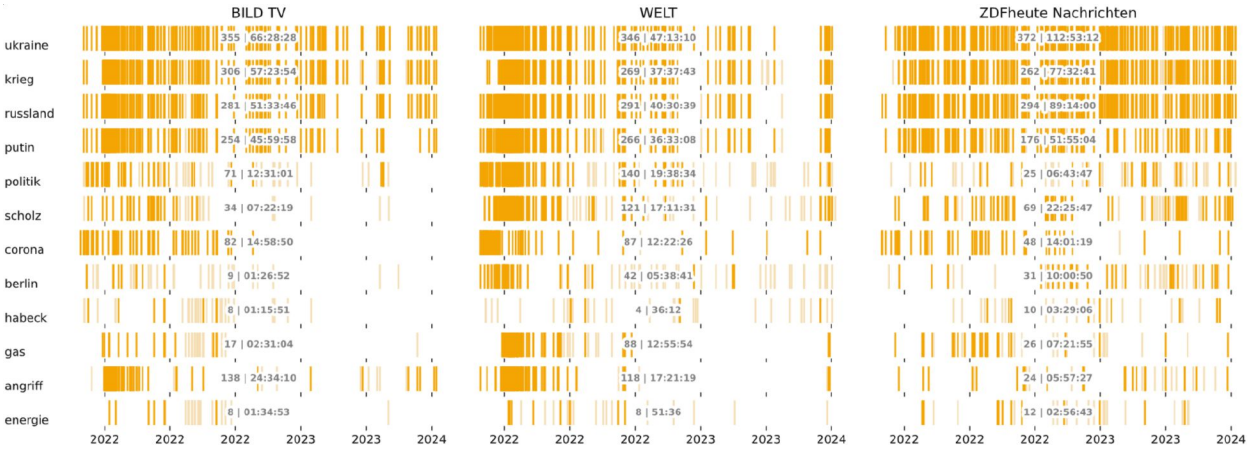

Abbildung 2: Diachrone Übersicht zur Verteilung von YouTube-Hashtags für die im Korpus enthaltenen Videos der Nachrichtensender BildTV, Welt und ZDF Heute.

AP 1.2 Formalisierung von Narrativitätsstrategien

Die Aufgabe des Teilvorhabens ULpz lag bei diesem AP insbesondere in der Identifikation von Merkmalen zu übergeordneten Narrativitätsstrategien, die sich objektiv bestimmen und später entsprechend des *scalable viewing*-Ansatzes visualisieren lassen. In der Anfangsphase des Projekts wurde zur Unterstützung dieses APs in Leipzig das Tool *Zoetrope*[3] entwickelt, welches verschiedene formale Parameter für einzelne Videos in einer Zeitleistendarstellung visualisieren kann und dadurch die explorative Identifikation von narrativen Mustern und Strategien unterstützt (vgl. Abb. 3). Die in *Zoetrope* dargestellten Video-Informationen enthalten:

[2] Sind nur für BildTV, Welt und ZDF Heute Nachrichten verfügbar.
[3] Der Quellcode von *Zoetrope* ist über ein frei zugängliches GitHub-Repositorium verfügbar: https://github.com/poke1024/zoetrope

- **Vorschaubilder** der einzelnen Frames
- Ergebnisse der ***optical character recognition*** (OCR), d.h. im Video textuell dargestellte Sprache (bspw. Bauchbinden, Schriften auf Plakaten, etc.)
- Ergebnisse der ***automatic speech recognition*** (ASR), d.h. gesprochene Sprache aus den Videos
- Ein **Clustering** zu im Video erkannten **Gesichtern**
- Quantitative Informationen zur **Größe** der in einem Frame dargestellten **Gesichter** (bspw. close-up vs. Totale)
- Quantitative Informationen zur **Anzahl** der in einem Frame dargestellten **Gesichter**, um bspw. Massenszenen von Szenen mit nur einer oder zwei Personen unterscheiden zu können
- Quantitative Informationen zum Audio-Signal in Form von **Amplitudeninformationen** und einer **spektrographischen Analyse**
- Darstellung von dominanten **Farbwerten** als Movie Barcode
- Quantitative Informationen zur **visuellen Dynamik** als Maß wie stark sich das Video visuell über einen kurzen Zeitraum hinsichtlich des Gezeigten verändert (ermittelt anhand von CLIP embeddings, vgl. Radford et al., 2021)
- **Aggregation** von ganzen Video-Abschnitten in einzelne Bilder im Sinne des *distant viewing* mit Hilfe von Methoden des *volumetric cinema* (Ferguson, 2019)

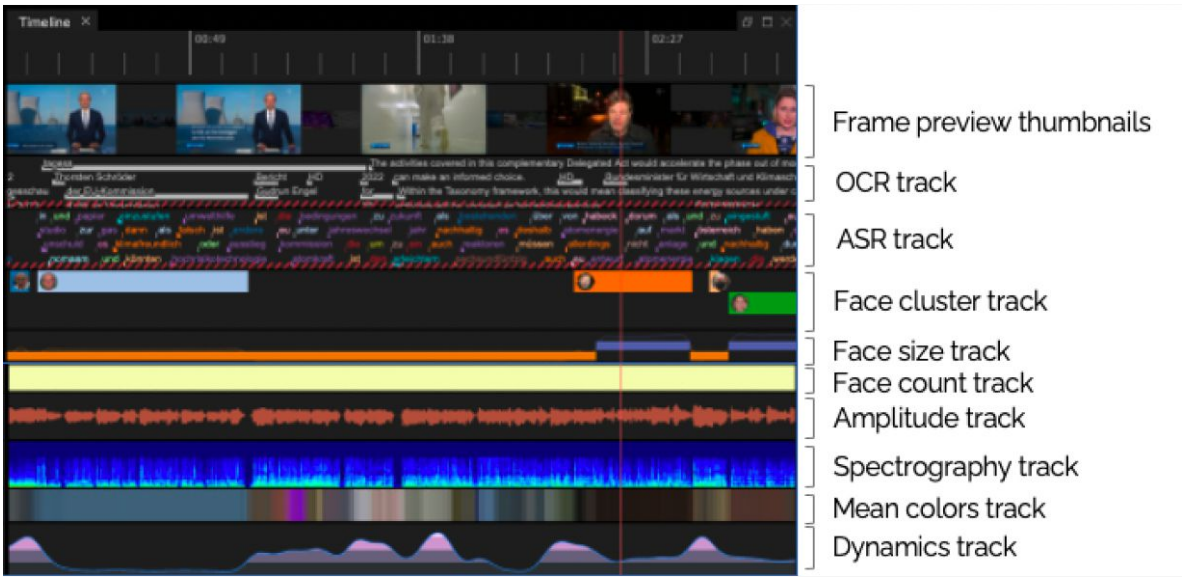

Abbildung 3: Übersicht über die unterschiedlichen Informationen die in *Zoetrope* quantitativ in einer Zeitleiste dargestellt werden.

Das *Zoetrope*-Tool wurde von den Multimodalitätsexpert:innen des Verbundpartners UBre genutzt um in Nachrichtenvideos zur Schau gestellte Narrativitätsstrategien wie bspw. „Personali-

sierung“ und „Dramatisierung“ in Videos anhand der vielfältigen quantitativen Parameter zu formalisieren. Eine Darstellung des *Zoetrope*-Tools sowie auch erste Ergebnisse dieser Formalisierung finden sich in Tseng et al. (2023) sowie in Liebl & Burghardt (2023a/b).

AB 2 – Multimodale Informationsextraktion (Lead: LS3)

Die Aufgabe des Teilprojekts ULpz lag bei diesem AB in der Formalisierung und Klassifizierung von Merkmalen, um diese in den nachfolgenden Arbeitsschritten für Verfahren des *scalable reading* nutzbar zu machen. Dieser Task wurde unter Zuhilfenahme verschiedener Frameworks und Modelle des bereits erwähnten *Zoetrope*-Tools umgesetzt.

AP 2.1 Analyse der Audiodaten und Erzeugung von Sprachtranskripten

Vom Verbundpartner ULpz war keine Beteiligung an diesem AP geplant.

AP 2.2 Informationsextraktion aus Einzelbildern

Für das in AP 1.2 bereits vorgestellte Tool *Zoetrope* wurden verschiedene Informationen aus den einzelnen Frames der verschiedenen Nachrichtenvideos extrahiert. Dazu kamen unterschiedliche Frameworks und Modelle zum Einsatz, die zum Zeitpunkt des Projekts den state-of-the-art gut abbildeten:

- OCR (*optical character recognition*), für die Extraktion von geschriebenem Text (vgl. Chen et al., 2021)
- Semantische Beschreibung von Bildinhalten mittels automatisch erstellter *captions* (vgl. das BLIP-Modell von Li et al., 2022)
- Movie Barcodes wurden mithilfe von numpy aus den RGB-Farbdaten der Video-Frames extrahiert

AP 2.3 Multimodale Informationsextraktion aus Nachrichtenvideos

In diesem Arbeitspaket wurden analog zu AP 2.2 weitere Informationen extrahiert, dieses Mal allerdings nicht für statische Einzelbilder, sondern für multimodale Informationen entlang der Zeitachse eines Nachrichtenbeitrags. Auch hier kamen wiederum verschiedene Frameworks und Modelle zum Einsatz, die zum Zeitpunkt des Projekts den state-of-the-art gut abbildeten:

- ASR (*automatic speech recognition*), für die Extraktion von Informationen zu gesprochener Sprache mithilfe des Mozilla Deep Speech-Modells[4]
- Audio-Features wurden mithilfe des Librosa-Frameworks extrahiert[5]
- Diverse Informationen (Größe, Anzahl, Ähnlichkeitsclustering) zu in den Nachrichtenvideos dargestellten Gesichtern wurden extrahiert (vgl. die Ansätze von Deng et al. 2019 sowie auch Deng et al. 2020)

[4] Mozilla Deep Speech: https://github.com/mozilla/DeepSpeech
[5] Librosa: Librosa: https://librosa.org/doc/latest/index.html

AB 3 – Erkennung von Narrativitätsstrategien (Lead: UBre)

AP 3.1 Ableiten von Narrativitätsstrategien und konkreten Merkmalen aus der Multimodalitätstheorie für den Bereich „Nachrichtenvideos“

Gemeinsam mit den Expert:innen aus der Multimodalitätstheorie wurden hier konkrete Narrativitätsstrategien aus der bestehenden Literatur abgeleitet. Das Teilprojekt ULpz war hier insbesondere dafür zuständig, die Erstellung und formale Repräsentation eines Annotationsschemas für die besagten Narrativitätsstrategien durch die Bereitstellung eines interaktiven Annotationstools zu unterstützen. Zu diesem Zwecke wurde das bestehende Tool *Zoetrope* um zusätzliche Annotationsfunktionen ergänzt. Die Annotation erfolgt dabei im Wesentlichen über die Möglichkeit nach spezifischen Parametern, wie bspw. bestimmten Keywords (vgl. Abb. 4) oder einer bestimmten Anzahl von Personen in einem Video zu suchen. Auch die in AP 2.2 erstellten BLIP *captions* zu den einzelnen Videoframes sind über *word embeddings* (FastText[6]) flexibel durchsuchbar (vgl. Abb. 5). *Word embeddings* bilden formal die Bedeutungsähnlichkeiten zwischen Wörtern ab, sodass z. B. eine Suche nach „Fahrzeug“ auch Frames anzeigen würde, bei denen das Wort „Auto“ oder „LKW“ in der *caption* vorkommt,

Zoetrope liefert jeweils individuelle Annotationsebenen für das Video zurück, welche alle Fundstellen für den gesuchten Parameter beinhaltet. Zur weiteren Detailannotation der einzelnen Annotationsebenen ist schließlich ein Export der Annotationen im ELAN-Format möglich. ELAN[7] ist ein weitverbreitetes Tool zur händischen Annotation von Videodaten. Der vollständige Annotationsworkflow sieht hier also zunächst eine automatisierte Annotation innerhalb *Zoetrope* auf Basis ausgewählter Such- und Filterparameter vor, mit der Möglichkeit diese Daten dann für eine manuelle Detailannotation in das bestehende, frei verfügbare ELAN-Tool zu exportieren.

Abbildung 4: Beispiel für zwei automatisch in *Zoetrope* erstellte Keyword-Annotationsebenen, die entlang der Zeitachse unterhalb des Videos dargestellt werden. In der ersten Ebene wurde dabei nach Vorkommen des Keywords “Lindner” im geschriebenen Text (OCR) des Videos gesucht, in der zweiten Ebene wurde die gesprochene Sprache (ASR) nach dem Keyword “Minister” durchsucht.

[6] FastText Repository: https://github.com/facebookresearch/fastText
[7] ELAN Annotationstool: https://archive.mpi.nl/tla/elan

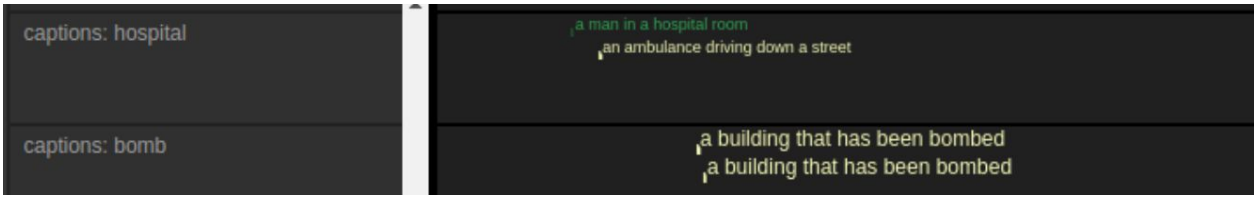

Abbildung 5: Beispiel für zwei automatisch in *Zoetrope* erstellte *caption*-Annotationsebenen, die entlang der Zeitachse unterhalb des Videos dargestellt werden. In der ersten Ebene wurde dabei nach Vorkommen des Keywords “hospital” in den Video *captions* gesucht, in der zweiten Ebene wurde nach “bomb” gesucht. Da das verwendete Modell englische *captions* erstellt, müssen die Suchbegriffe hier entsprechend englischsprachig sein. Durch die *word embeddings* werden auch *captions* gefunden, die nicht explizit das Keyword enthalten, wie etwa beim Beispiel “an ambulance driving down a street”, was semantisch eng verwandt mit der Suchanfrage “hospital” sein dürfte.

AP 3.2 Semi-automatische Erkennung von Narrativitätsstrategien

Grundlegende Arbeiten zur Auswertbarkeit und Evaluation der Ausgabedaten der automatischen Verfahren wurden durch den Verbundpartner ULpz in drei Schritten durchgeführt:

1. Erstellung und Formalisierung eines einheitlichen, hierarchisch strukturierten und stark typisierten **Datenschemas** (vgl. Abb. 6), das die domänenspezifischen narrativen Konzepte auf einem angemessenen Abstraktionsniveau modelliert und damit die Grundlage für Retrieval und Evaluation bildet.

2. **Transformation** der technischen Ausgabedaten, die vom Verbundpartner L3S extrahiert werden in oben genanntes formales Schema durch verschiedene Schritte der Aggregation und Kombination

3. Design und Implementierung einer flexiblen ***Query Language***, die komplexe Abfragen und Evaluationen auf dem nun verfügbaren homogenen Datensatz ermöglicht

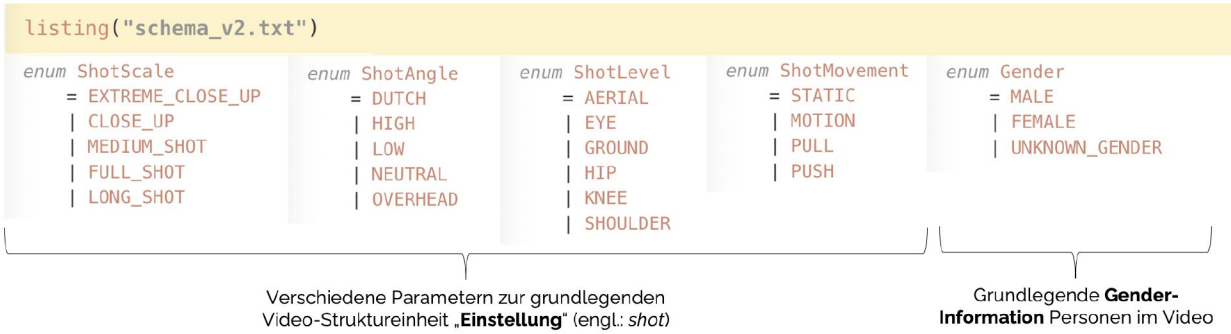

Abbildung 6: Auszug aus dem umfangreichen Datenschema für die formale Abbildung narrativer Muster.

Alle drei Schritte sind die wesentliche Grundlage für das in Leipzig entwickelte Tool *Narrascope*[8], welches im AB4 als zentrales Ergebnis für den Bereich des *scalable viewing* im Detail beschrieben wird. Großer Wert wurde bei der Entwicklung neben Aspekten der Visualisierung auch auf die Anpassbarkeit der Queries und deren Ausführbarkeit in Echtzeit gelegt. Durch seine in hohem Maße interaktiven und flexiblen Retrieval- und Visualisierungsfähigkeiten konnte *Narrascope* im Projekt erfolgreich zur Kontrolle der Qualität der automatischen Verfahren eingesetzt werden. Darüber hinaus ist – wie schon bei *Zoetrope* – ein Export in die Software ELAN zur Korrektur von Annotationen möglich.

AP 3.3 Überprüfung der intermodalen Konsistenz von Informationen

Vom Verbundpartner ULpz war keine Beteiligung an diesem AP geplant.

AB 4 – Explorative Datenanalyse mittels Scalable Viewing (Lead: ULpz)

In diesem Arbeitsbereich finden sich aus wissenschaftlicher Sicht die wichtigsten Ergebnisse des Verbundpartners ULpz, welche im Bereich des *scalable viewing* liegen. *Scalable viewing* bedeutet dabei eine dynamische Informationsvisualisierung für Videoinhalte, auf unterschiedlichen Detail- bzw. Abstraktionsgraden (= *scales*). Die Ansätze folgen dabei im Kern dem *Visual Information Seeking Mantra* von Shneiderman (1996): “Overview first, zoom and filter, then details-on-demand”. Shneiderman fasst mit diesem Mantra prägnant einen gestuften Zugang zu komplexen Informationen zusammen: Zunächst soll eine Übersicht präsentiert werden, dann ermöglicht man das gezielte Ein- oder Ausblenden relevanter Inhalte, und schließlich kann man bei Interesse weitere Details abrufen. Für das vorliegende Szenario von Nachrichtenvideos wurden unterschiedliche Ansätze der Skalierbarkeit in insgesamt drei Softwaretools entwickelt (*Zoetrope*, *ClusterGraph*, *Narrascope*). *Zoetrope* und *Narrascope* wurden bereits in vorherigen Arbeitspaketen erwähnt, sollen in diesem Abschnitt aber erneut aus Perspektive der Visualisierungsmöglichkeiten vorgestellt werden. Dabei werden zunächst Aspekte des *close viewing* (also der Detailanalyse) in AP 4.1, dann Möglichkeiten des *distant viewing* (also der Analyse von großflächigen Mustern und Relationen) in AP 4.2 und zuletzt die Synthese aus beiden Ansätzen im Sinne eines *scalable reading* in AP 4.3 beschrieben. Eine grundsätzliche Methodologie zur Analyse von Filmen und Videos mithilfe empirischer Ansätze und Tools aus den Digital Humanities wurde von den Verbundpartnern des Projekts FakeNarratives in einer gemeinsamen Methodenpublikation mit dem Titel “Computergestützte Werkzeuge und Methoden für die Film- und Videoanalyse beschrieben” (vgl. Burghardt et al., 2024).

AP 4.1 Close Viewing

In der Einzeldarstellung von Videos greift *Zoetrope* auf eine Multi-Track-Darstellung zurück, die - wie bereits in AP 1.2 ausgeführt - aus mehreren spezialisierten *data tracks* besteht. Hier liegt eine

[8] Zu *Narrascope* liegen zum Zeitpunkt des Berichts noch keine veröffentlichten Publikationen vor. Es wird aber an entsprechenden Beiträgen gearbeitet, die zeitnah in einschlägigen Publikationsorganen der Öffentlichkeit zugänglich gemacht werden sollen. Für einen ersten Überblick zu den Möglichkeiten des *Narrascope*-Tools wurde ein Demo-Video erstellt: https://tinyurl.com/4ahtzwyz

typische Visualisierung von Videos vor, die als Standardverfahren für das *close viewing* von Bewegtbildern gelten mag: entlang einer Zeitachse können beliebige Stellen des Videos inspiziert und dabei die automatisch extrahierten Informationen näher analysiert werden (vgl. Abb. 7).

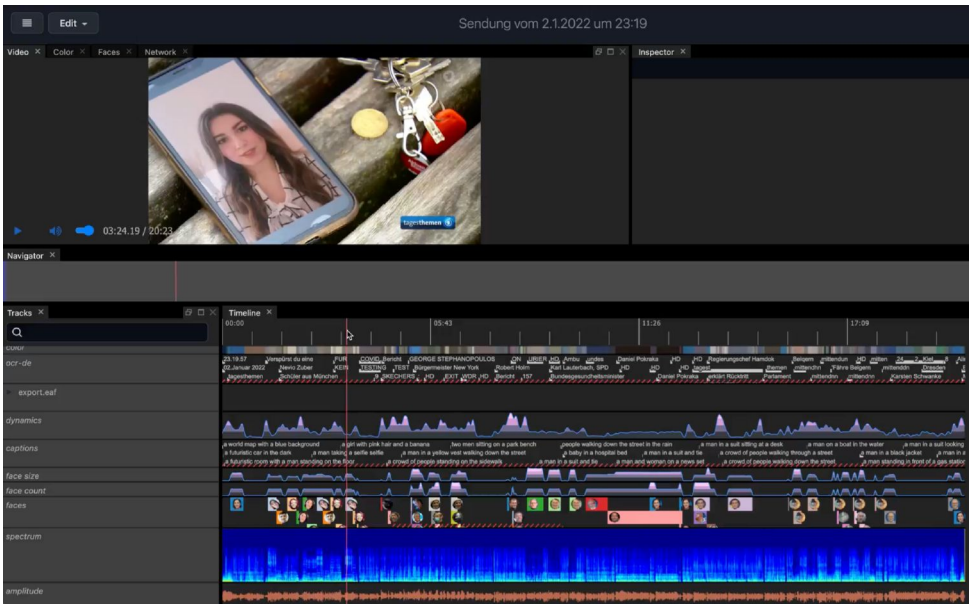

Abbildung 7: Darstellung des *Zoetrope*-Tools für Zwecke des *close reading*. Das Video kann in einem video player (links oben) dargestellt werden. Parallel dazu werden verschiedene, automatisch extrahierte Informationen in mehreren Ebenen (*data tracks*) dargestellt (untere Hälfte der Abbildung).

Während *Zoetrope* in der ersten Hälfte der Projektlaufzeit entwickelt wurde, ist *Narrascope* primär das Ergebnis der zweiten Projekthälfte. *Narrascope* fungiert dabei hauptsächlich als Information Retrieval-System und ermöglicht die Suche von Videos in einem größeren Korpus, anhand bestimmter Parameter, die in einer spezifischen Suchanfrage (*query*) abgefragt werden können. In Abb. 8 ist eine *query* zur Abfrage eines spezifischen Patterns der narrativen „Intensivierung“ über drei Einstellungen *(shots*) hinweg gezeigt. Abgefragt wird hier die kontinuierlich größer-werdende Darstellung von Personen im Video. Die *query language* und allgemein die starke Formalisierung narrativer Strategien in Nachrichtenvideos ist inspiriert durch Arbeiten von Wu et al. (2018), die filmische Patterns formal beschreiben.

```
(): Tuple

hits <| batch corpus.videos:
annotate [
    s1, s2, s3
] if (
    some [a, b, c] in actors(s1, 0.1) + actors(s2, 0.2) + actors(s3, 0.3) gathered by id:
            b.size / a.size > 1.25 and c.size / b.size > 1.25
) as FEP_INTENSIFY_V1
from shots
```

Abbildung 8: Beispielhafte Darstellung einer formalisierten Suchanfrage über die eigens entwickelte *query language* von *Narrascope*. Häufig verwendete Suchanfragen können auch über *shortcuts* gespeichert werden, und müssen nicht jedes Mal von Neuem getippt werden.

Die Visualisierung der Ergebnisse dieser konkreten Suchanfrage sind in Abb. 9 zu sehen. Hier wird ein kleiner Ausschnitt der vollständigen Ergebnisliste dargestellt. Jede Zeile repräsentiert ein Video, in dem das gesuchte narrative Pattern vorkommt. Die vertikalen grauen Linien markieren die Schnittgrenzen einzelner Einstellungen innerhalb jedes Videos. Grün hervorgehobene Bereiche zeigen diejenigen Sequenzen von Einstellungen, die mit der angegebenen Abfrage übereinstimmen. Die fokussierte Zeile in der Mitte ermöglicht eine genauere Betrachtung im Detail, was u.a. die *close viewing*-Möglichkeiten von *Narrascope* demonstriert (von einer gefilterten Überblicksdarstellung aller passenden Videos in die Ergebnisliste von *shots* kommend). Die Darstellung zeigt mittels grüner Markierungen zwei Übereinstimmungen, den Namen des Musters auf der linken Seite sowie die genauen Zeitstempel am unteren Rand. Das durch die Zeitleiste dargestellte Video wird darunter angezeigt. Durch Bewegen des roten Abspielkopfs können Nutzer:innen interaktiv zu verschiedenen Positionen springen und das Video mit allen Details betrachten, um so das narrative Muster mittels *close viewing* im Kontext konkreter Videos nachvollziehen zu können.

Abbildung 9: Beispielhafte Darstellung der Ergebnisse zu einer Suchanfrage in *Narrascope*. Jede Zeile stellt ein Video dar, welches das gesuchte narrative Muster enthält. Die grünen Markierungen zeigen an, an welcher Stelle des Videos das Muster gefunden wurde. Jedes Video kann direkt in der Ergebnisliste an beliebiger Stelle abgespielt werden.

AP 4.2 Distant Viewing

Während *close viewing* die qualitative Detailanalyse einzelner Videos beschreibt, so bedeutet *distant viewing* die Analyse von Mustern und Beziehungen mehrerer Einstellungen in einem Video oder gar die Analyse von größeren Videokorpora. Bereits das zu Projektbeginn entwickelte Tool *Zoetrope* enthält *distant viewing*-Funktionen. So zeigt etwa der Video-Browser (vgl. Abb. 10) von *Zoetrope* ein vollständiges Video-Korpus durch die Darstellung von Vorschaubildern und sogenannten Movie-Barcodes (vgl. Burghardt et al., 2018) in einer Überblicksdarstellung an (*distant*

view). Der Movie-Barcode stellt für sich eine Überblicksdarstellung für jedes einzelne Video dar und bietet dabei eine visuelle Zusammenfassung, die auf einen Blick eine einfache Übersicht über dominante Farben entlang der Zeitachse eines Videos ermöglicht. Wenn der Mauszeiger über das Video bewegt wird, aktualisiert sich das Vorschaubild entsprechend der horizontalen Cursorposition und zeigt das zugehörige Bild des Videos an. Dies erleichtert eine gezielte Vorschau und Navigation. Videos können weiterhin mithilfe einer Volltextsuche gefiltert werden, die sich nicht nur auf Metadaten erstreckt, sondern auch auf automatisch extrahierte multimodale Informationen (vgl. APs 2.2 + 2.3) der Track-Daten. Wie in Abb. 10 dargestellt, wird mit der Abfrage „Regierung" jedes Video angezeigt, das mindestens Teile des gesuchten Begriffs entweder in der gesprochenen Audiospur (ermittelt durch ASR) oder als eingeblendeten Text (ermittelt durch OCR) enthält. Der Video-Browser ermöglicht es Nutzern, sich – im Sinne eines *distant viewing* – einen schnellen Überblick über ein größeres Korpus zu verschaffen und nach bestimmten Kriterien Subkorpora zu filtern. Die detaillierte *close viewing*-Ansicht eines Videos in *Zoetrope* (vgl. AP 4.1) öffnet sich durch Klicken auf das jeweilige Vorschaubild. Durch diese Möglichkeit zwischen *close* und *distant viewing* hin und her zuschalten finden sich bereits erste Ansätze für das final angestrebte *scalable viewing*, welches in AP 4.3 noch weiter elaboriert wird.

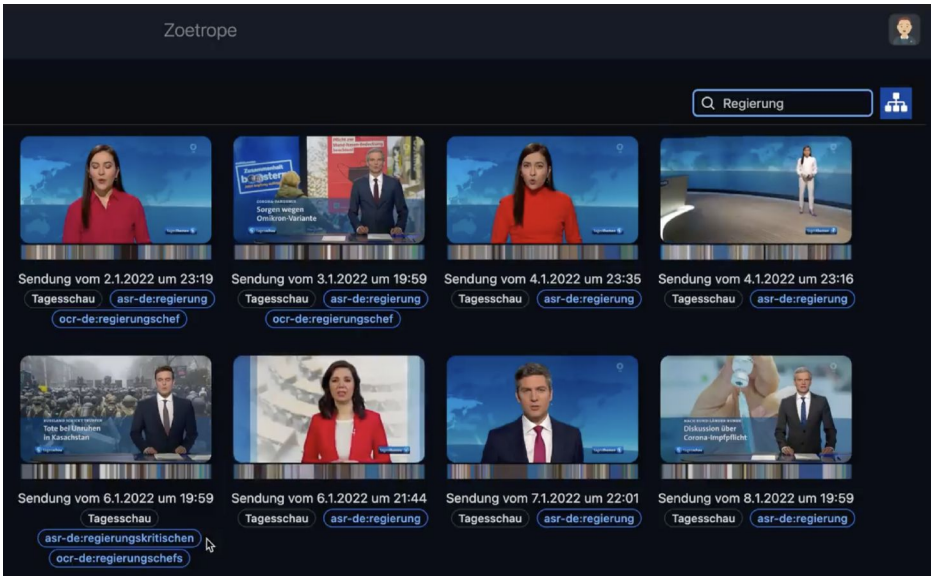

Abbildung 10: Der Video-Browser als grundlegende *distant viewing*-Komponente (Überblick / Filtermöglichkeiten) in *Zoetrope*.

Auch in der Einzelvideodarstellung von *Zoetrope* finden sich *distant viewing*-Komponenten im Sinne aggregierter, zusammenfassender Visualisierungen für einzelne Videos. So fasst etwa das *face network* (vgl. Abb. 11) in *Zoetrope* Informationen über das gemeinsame Auftreten von Gesichtern auf der Bildebene zu einem Netzwerk zusammen. Derlei Netzwerkvisualisierung in *Zoetrope* sind dabei nicht auf Gesichter beschränkt, sondern lassen sich auf andere Parameter wie etwa gesprochene und geschriebene Texte verallgemeinern, die ebenfalls als Ereignisse in das Kookkurrenz-Netzwerk aufgenommen werden können, und deren gemeinsames oder zeitgleiches Auftreten dann mit entsprechenden Kanten kodiert wird.

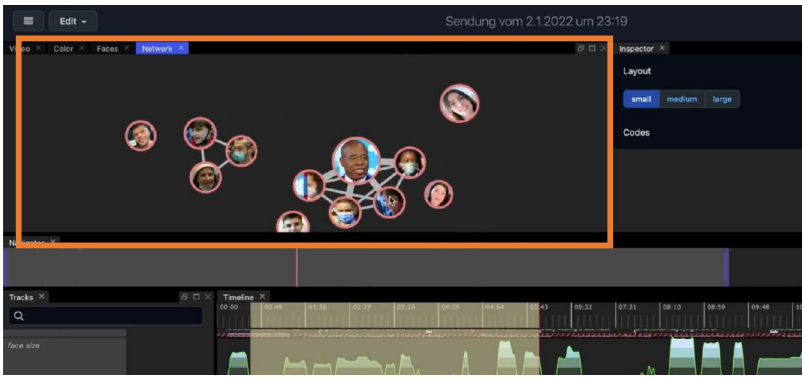

Abbildung 11: *Face network*-Visualisierung in *Zoetrope* als *distant viewing* für ein ganzes Video.

Inspiriert durch die Kookkurrenz-Netzwerkdarstellung verschiedener Parameter für Einzelvideos, wurde der Ansatz mit *ClusterGraphs* als experimentelle Visualisierung für ein Sample von 117 Tagesschausendungen weiterentwickelt (Ruth, Liebl & Burghardt, 2023a/b). Da schon bei diesem überschaubar großen Video-Korpus und einer relativ niedrigen Auflösung von einem Frame je 0.5 Sek. bereits mehr als 300.000 Einzelframes anfallen, wurde zunächst ein Clustering inhaltlich ähnlicher Frames (anhand der bestehenden *embeddings*, vgl. AP 2.2) vorgenommen, um die Netzwerkvisualisierung nicht zu überfrachten. Im Ergebnis entstehen so 231 Cluster von Frames aus 117 Einzelvideos, die ähnliche Inhalte zeigen. In Abb. 12 ist links eine Detailanansicht mit Einzelframes für ein Cluster mit Abbildungen von Putin und weiteren einschlägigen Personen zu sehen. Jeder Cluster wird in der Netzwerkdarstellung (vgl. Abb. 12, rechts) als Knoten dargestellt. Die Kanten-Verbindungen zwischen den einzelnen Frame-Clustern ergeben sich aus dem gemeinsamen Auftreten von Frames innerhalb eines Videos. Auf diese Weise ist es möglich, thematische Cluster und deren Kookkurrenz in Nachrichtenvideos zu visualisieren und im Falle von *ClusterGraphs*[9] interaktiv als 3D-Visualisierung explorierbar zu machen.

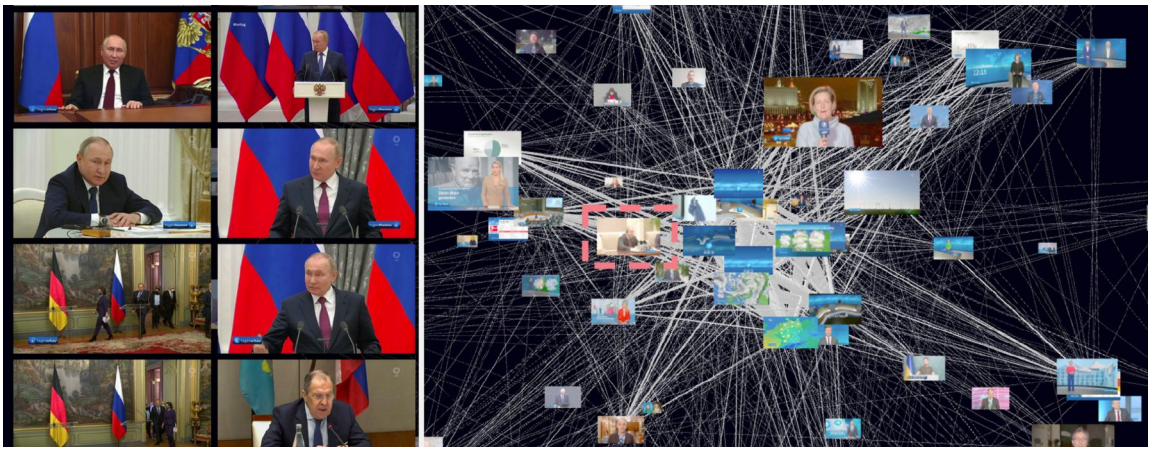

Abbildung 12: Ausschnitt aus ClusterGraphs – links zu sehen ist eine Detailsicht auf eines der Cluster, welches einen Überblick zu allen enthaltenen Einzelframes liefert. Rechts zu sehen ist die Gesamtdarstellung aller Cluster in einem Kookkurrenznetzwerk.

[9] Der Quellcode sowie auch ein Demovideo sind über ein entsprechendes GitHub-Repositorium verfügbar: https://github.com/Nicolas-le/from-clusters-to-graphs

Durch die Netzwerkmodellierung der Frame-Kookkurrenzen in verschiedenen Videos ist es weiterhin möglich, *communities* innerhalb des Netzwerks zu identifizieren, d.h. Gruppen von Cluster-Knoten, die besonders häufig miteinander kookkurieren und dementsprechend einen bestimmten thematischen Bereich abdecken. Abb. 13 etwa zeigt eine solche automatisch berechnete *community* im Gesamtgraphen, bei dem alle Cluster im weiteren Sinne dem Themenbereich “Covid-19 Pandemie” zuzuordnen sind. Bei näherer Betrachtung der Cluster finden sich hier bspw. Frames die medizinische Handlungen zeigen oder auch Szenen der Bundespressekonferenz enthalten. Mit diesen verschiedenen Detail- und Abstraktionsstufen beinhaltet die *ClusterGraphs*-Visualisierung bereits erste Merkmale des scalable viewing, die in AP 4.3 noch weiter ausgeführt werden.

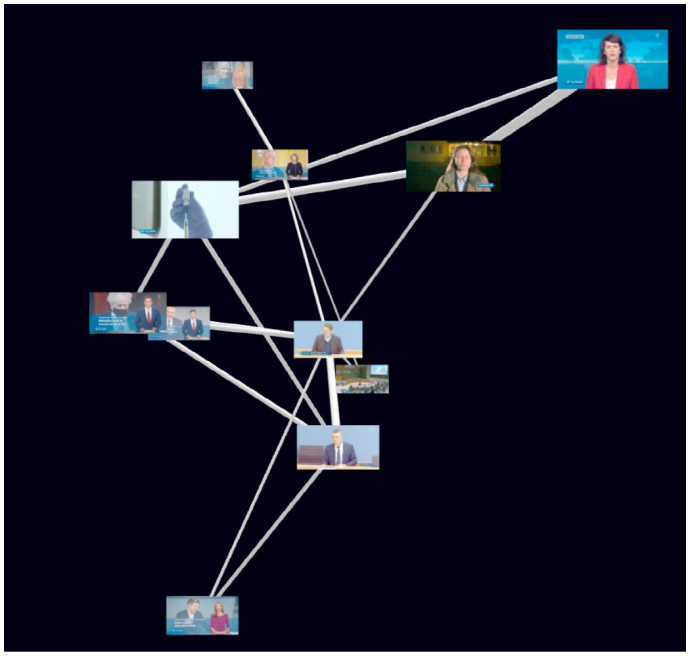

Abbildung 13: Teilgraph aus dem Tagesschau-Graphen, der die Thematik “Covid-19 Pandemie mit insgesamt 11 tendenziell kookkurierenden Clustern wiedergibt.

AP 4.3 Scalable Viewing

Neben grundlegenden Skalierungsmöglichkeiten in *Zoetrope* (vgl. AP 4.1+4.2) liefert insbesondere auch *Narrascope* innovative Visualisierungen für ein *scalable viewing*, dass sowohl Möglichkeiten des Überblicks, des Filterns und der Detailanalyse (vgl. AP 4.1) bereitstellt. Da all diese Perspektiven im selben Tool bereitstehen, ist ein mehr oder weniger nahtloses *scalable viewing* von *Nachrichtenvideos* möglich. Hervorgehoben werden soll an dieser Stelle insbesondere eine innovative Überblicksvisualisierung in *Narrascope*, die in Analogie zur klassischen *n-gram* Analyse in textbasierten Korpora konzipiert wurde. Ein *n-gram* steht dabei für eine Analyseeinheit, die aus n Elementen (bspw. Zeichen oder Wörtern) besteht. Die visuellen n-grams in *Narrascope* bestehen dann aus definierbaren Abfolgen von n Einstellungen. Abb. 14 zeigt beispielsweise die Auswertung der Einstellungsgrößen (*shot scales*) in allen Videosequenzen die aus drei Einstellungen bestehen. Die Visualisierung kann hier Kontraste zwischen unterschiedlichen Nachrichten-Kanälen aufzeigen.

Je nach Position, Farbe und Größe des Kreises im Diagramm, lässt sich ablesen, welcher Kanal eine bestimmte Dreier-Folge von *shot scales* (z.B. die Abfolge *close up*, *full shot*, *full shot*) häufiger benutzt als die anderen Kanäle. Genaue Frequenzen werden beim *hovering* über einen Kreis eingeblendet. Gleichzeitig lassen sich durch diese Art der Darstellung auf einen Blick komplexere Muster erkennen. Liegt ein Kreis zum Beispiel auf der Linie zwischen zwei *shot scales*, so kodiert der Kreis eine Abfolge, die nur aus diesen beiden *shot scales* besteht. So zeigen z.B. die größeren türkisen Kreise oben, dass die Nachrichtensendung WELT relativ häufig ganz bestimmte Dreierfolgen verwendet, die nur aus *close up* und *full shot* bestehen. Gleichzeitig nutzt die Tagesschau typischerweise Dreierfolgen zwischen *medium shot* und *full shot*, während bei CompactTV Kombinationen aus *close up* und *medium shot* typisch erscheinen. Die Visualisierung lässt sich für verschiedene n-gram-Längen und verschiedene Eigenschaften (wie *shot angle*, *shot level*, und viele andere mehr) anpassen.

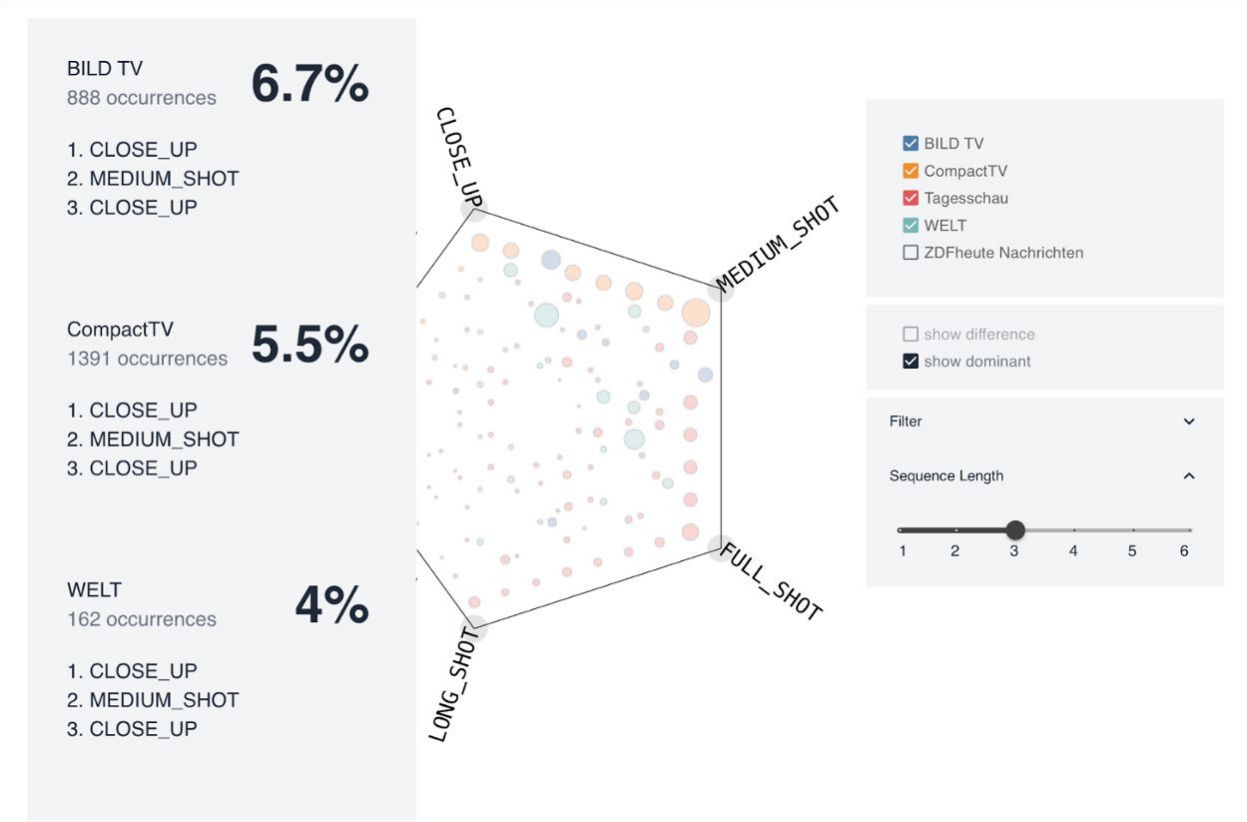

Abbildung 14: Kontrastive Überblicksvisualisierung von *shot* n-grams in *Narrascope*.

Diese kontrastive Überblicksvisualisierung (*distant viewing*) kann durch Klick auf einen der Kreise wiederum in eine Detaildarstellung skalieren, indem eine Liste von relevanten Videos analog zur Ergebnisanzeige im *query* Modus angezeigt wird. Die entsprechenden Videos, in denen bestimmte *shot* n-grams enthalten sind, können sodann im Sinne des *close viewing* mit allen filmischen Kontextinformationen analysiert werden (vgl. Abb. 15).

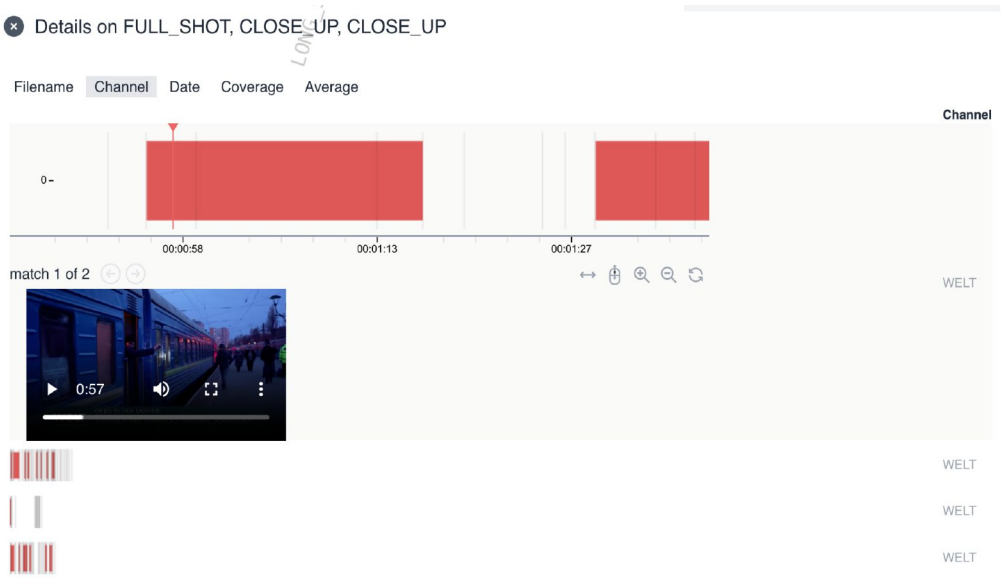

Abbildung 15: Ergebnisliste und Detaildarstellung von Videos in *Narrascope*.

AB 5 – Kontrastive Analyse audiovisueller narrativer Strategien (Lead: UBre)

AP 5.1 Kontrastive Analyse audiovisueller narrativer Strategien

Dieses AP wurde schwerpunktmäßig vom Verbundpartner UBre bearbeitet. Der Partner ULpz hat an dieser Stelle im Wesentlichen die in AB4 ausführlich beschriebenen Visualisierungs- und Analysetools beigesteuert. Kontrastive Analysen unter Zuhilfenahme von *Zoetrope* werden etwa in Tseng et al. (2023) dokumentiert. Für weitere kontrastive Analysen, zu denen zum Zeitpunkt dieses Berichts noch keine veröffentlichten Publikationen vorliegen, wurde in der zweiten Hälfte schwerpunktmäßig das *Narrascope*-Tool eingesetzt (vgl. AP 4.3). Dabei wurden entsprechende Queries formuliert, um gezielt nach narrativen Strategien in Nachrichtenvideos zu suchen (vgl. Abb. 16).

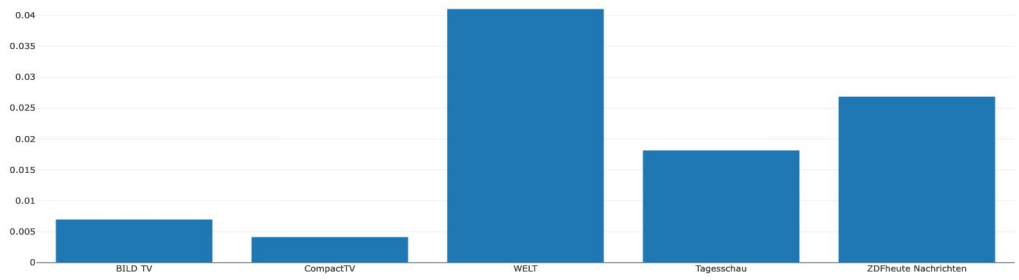

Abbildung 16: Kontrastive Analyse aller fünf Nachrichtenformate hinsichtlich der Verwendung des narrativen Musters der “Individualisierung von Reporter:innen” in den jeweiligen Beiträgen (durchgeführt vom Verbundpartner UBre).

AP 5.2 Identifikation von narrativen Lücken

Auch dieses AP wurde schwerpunktmäßig vom Verbundpartner UBre bearbeitet und von Leipzig durch die Bereitstellung und fortwährende Anpassung der Tools *Zoetrope* und insbesondere *Narrascope* unterstützt. Eine exemplarische Analyse findet sich etwa in Bateman, Burghardt & Tseng (2023).

AP 5.3 Visualisierung und Ausformulierung von Gegenbotschaften zu Desinformation

Dieses AP wurde ebenfalls schwerpunktmäßig vom Verbundpartner UBre bearbeitet und von Leipzig insbesondere durch die Bereitstellung und fortwährende Anpassung der Tools *Narrascope* unterstützt. Mit den diversen *close* und *distant viewing*-Visualisierungen (vgl. AP 4.3) liefert *Narrascope* vielfältige Möglichkeiten zur Unterstützung dieses Arbeitspakets

6. Ergebnisverwertung

Die entwickelten Systeme *Zoetrope* und *ClusterGraphs* sind als Open-Source-Software verfügbar und wurden im Rahmen wissenschaftlicher Beiträge veröffentlicht. Geplant ist eine Weiterentwicklung im Rahmen zukünftiger Drittmittelprojekte. Die technischen Ansätze sowie auch ganz allgemein die Methodik des *scalable viewing* werden zudem in anderen Forschungsprojekten der Computational Humanities Group in Leipzig aufgegriffen und fortgeführt. So etwa in einem laufenden Promotionsprojekt zur computergestützten Untersuchung von ideologischen Mustern in NS-Vorbehaltsfilmen (Ruth & Burghardt, 2025) sowie auch in einem Drittmittelprojekt zur Analyse von Illustrationen in einem großen Korpus von Kinder- und Jugendliteratur des langen 19. Jahrhunderts (Helm et al, 2025).

Neben zahlreichen wissenschaftlichen Veröffentlichungen und weiteren geplanten Papers, insbesondere zum *Narrascope*-Tool, finden die im Projekt entwickelten Tools und Methoden auch regelmäßig Eingang in die Lehre der Bachelor- und Masterstudiengänge Digital Humanities, an der Universität Leipzig (auch für die Zukunft geplant):

- **Wintersemester 2022/2023**: "Desinformation in Nachrichtenvideos erkennen und verstehen – Grundlagen der Extraktion und Visualisierung von multimodalen Informationen“, Mastermodul (Seminar + Übung), Manuel Burghardt
- **Wintersemester 2024/2025**: „Analyse und Visualisierung von Videoinhalten“, Modul Aktuelle Trends in den Digital Humanities (Seminar + Übung), Bernhard Liebl
- **Sommersemester 2025**: „Cultural Analytics“, Mastermodul (Vorlesung + Seminar), Bernhard Liebl
- **Abschlussarbeit** im B. Sc. Digital Humanities (abgegeben 10/2023) zum Thema "Ein computergestützter Ansatz zur Erkennung von narrativen Strategien in Nachrichten", bei der hauptsächlich ein Beitrag zur Evaluation des Zoetrope-Tools geleistet wurde (Betreuer: Manuel Burghardt)

7. Literaturverzeichnis

Hinweis: Alle fett markierten Quellen sind Publikationen, die im Rahmen des Projekts FakeNarratives veröffentlicht wurden.

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the Fake News Problem at the Scale of the Information Ecosystem. *Science Advances, 6*(14).

Arnold, T., & Tilton, L. (2019). Distant Viewing: Analyzing Large Visual Corpora. Digital Scholarship in the Humanities, 34(Supplement_1), i3–i16.

Bateman, J. A., Burghardt, M., & Tseng, C.-I. (2023). Narrative Strukturen Als Mittel Der Indirekten Kritik in Der Audiovisuellen Nachrichtenberichterstattung. Bestimmte Unbestimmtheit: Offene Struktur Und Funktionale Lenkung in Audiovisuellen Medien, 12, 279.

Brüggemann, M., & Engesser, S. (2017). Beyond False Balance: How Interpretive Journalism Shapes Media Coverage of Climate Change. *Global Environmental Change, 42*, 58–67.

Burghardt, M., Kao, M., & Walkowski, N.-O. (2018). Scalable Movie Barcodes – An Exploratory Interface for the Analysis of Movies. In 3rd IEEE VIS Workshop on Visualization for the Digital Humanities, Berlin.

Burghardt, M., Kao, M., & Wolff, C. (2016). Beyond Shot Lengths – Using Language Data and Color Information as Additional Parameters for Quantitative Movie Analysis. In Book of Abstracts of the International Digital Humanities Conference (DH).

Burghardt, M., Müller-Budack, E., Bateman, J. A., & Ewerth, R. (2024). Computergestützte Werkzeuge Und Methoden Für Die Film- Und Videoanalyse. In C. A. Nunn & F. van Oorschot (Eds.), Kompendium Computational Theology, Bd. 1: Forschungspraktiken in den Digital Humanities (pp. 161–189). heiBOOKS.

Burghardt, M., Pause, J., & Walkowski, N.-O. (2019). Scalable Viewing in Den Filmwissenschaften. In Book of Abstracts, DHd 2019, Frankfurt/Mainz.

Chen, X., Jin, L., Zhu, Y., Luo, C., & Wang, T. (2021). Text Recognition in the Wild: A Survey. ACM Computing Surveys (CSUR), 54(2), 1–35.

Cour, T., Jordan, C., Miltsakaki, E., & Taskar, B. (2008). Movie/Script: Alignment and Parsing of Video and Text Transcription. In European Conference on Computer Vision (pp. 158–171). Springer.

Deng, J., Guo, J., Ververas, E., Kotsia, I., & Zafeiriou, S. (2020). RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5203–5212).

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive Angular Margin Loss for Deep Face Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4690–4699).

Everingham, M., Sivic, J., & Zisserman, A. (2006). Hello! My Name Is... Buffy – Automatic Naming of Characters in TV Video. In BMVC (Vol. 2, p. 6).

Ferguson, K. L. (2019). Volumetric Cinema. In M. K. Gold & L. F. Klein (Eds.), Debates in the Digital Humanities 2019 (pp. 335–349). University of Minnesota Press.

Gräf, D., & Hennig, M. (2020). Die Verengung der Welt zur medialen Konstruktion Deutschlands unter Covid-19 anhand der Formate ARD Extra – Die Coronalage und ZDF Spezial. Preprint, verfügbar über ResearchGate. https://www.researchgate.net/publication/343736403

Helm, W., Borst, J., Burghardt, M. & Schmideler S. (2025). Buchkindheiten digital – Werkstattbericht zur computergestützten Analyse von Illustrationen in historischen Kinder- und Jugendbüchern. In Book of Abstracts, Dhd-Konferenz 2025, Bielefeld.

Hohman, F., Soni, S., Stewart, I., & Stasko, J. (2017). A Viz of Ice and Fire: Exploring Entertainment Video Using Color and Dialogue. In Proceedings of the 2nd Workshop on Visualization for the Digital Humanities (VIS4DH), Phoenix.

Hoyt, E., Ponto, K., & Roy, C. (2014). Visualizing and Analyzing the Hollywood Screenplay With ScripThreads. DHQ: Digital Humanities Quarterly, 8(4).

Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. In Information Visualization (pp. 154–175). Springer.

Kurzhals, K., John, M., Heimerl, F., Kuznecov, P., & Weiskopf, D. (2016). Visual Movie Analytics. IEEE Transactions on Multimedia, 18(11), 2149–2160.

Li, J., Li, D., Xiong, C., & Hoi, S. (2022). BLIP: Bootstrapping Language-Image Pre-Training for Unified Vision-Language Understanding and Generation. In International Conference on Machine Learning (pp. 12888–12900). PMLR.

Liebl, B., & Burghardt, M. (2023a). Zoetrope – Interactive Feature Exploration in News Videos. In A. Baillot, W. Scholger, T. Tasovac, et al. (Eds.), Digital Humanities 2023: Book of Abstracts. Graz.

Liebl, B., & Burghardt, M. (2023b). Designing a Prototype for Visual Exploration of Narrative Patterns in News Videos. In M. Burghardt & C. Weiß (Eds.), 4th Workshop InfDH, Proceedings of the INFORMATIK 2023 Conference, Lecture Notes in Informatics (LNI) (pp. 831–840). Gesellschaft für Informatik.

Manovich, L. (2013). Visualizing Vertov. Russian Journal of Communication, 5(1), 44–55.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. In Proceedings of the 38th International Conference on Machine Learning (Vol. 139, pp. 8748–8763). PMLR.

Rasheed, Z., Sheikh, Y., & Shah, M. (2005). On the Use of Computable Features for Film Classification. IEEE Transactions on Circuits and Systems for Video Technology, 15(1), 52–64.

Ross, M., Grauer, M., & Freisleben, B. (2009). Digital Tools in Media Studies: Analysis and Research. An Overview.

Ruth, N., Burghardt, M., & Liebl, B. (2023a). From Clusters to Graphs – Toward a Scalable Viewing of News Videos. In A. Šeļa, F. Jannidis, & I. Romanowska (Eds.), Proceedings of the Computational Humanities Research Conference 2023, Paris.

Ruth, N., Liebl, B., & Burghardt, M. (2023b). A Scalable View on the Visual Narrative: Exploring Relationships in News Videos. Poster Presented at Digital Humanities Day Leipzig (DHDL). https://doi.org/10.36730/2024.1.dhdl.1

Ruth, N. & Burghardt, M. (2025). Echoes of Ideology – Toward an Audio Analysis Pipeline to Unveil Character Traits in Historical Nazi Propaganda Films. In Book of Abstracts, ADHO DH-Conference 2025, Lissabon.

Salt, B. (1974). Statistical Style Analysis of Motion Pictures. Film Quarterly, 28(1), 13–22.

Salt, B. (2006). Moving Into Pictures: More on Film History, Style, and Analysis. Starword.

Sayers, J. (Ed.). (2018). The Routledge Companion to Media Studies and Digital Humanities. Routledge.

Schoeffmann, K., Hudelist, M. A., & Huber, J. (2015). Video Interaction Tools: A Survey of Recent Work. ACM Computing Surveys (CSUR), 48(1), 1–34.

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proceedings of the 1996 IEEE Symposium on Visual Languages (pp. 336). IEEE.

Tseng, C.-I., Liebl, B., Burghardt, M., & Bateman, J. (2023). FakeNarratives – First Forays in Understanding Narratives of Disinformation in Public and Alternative News Videos. In Proceedings of the DHd Conference 2023, Luxemburg/Trier.

Tsivian, Y. (2009). Cinemetrics, Part of the Humanities' Cyberinfrastructure.

Wu, H. Y., Palù, F., Ranon, R., & Christie, M. (2018). Thinking like a director: Film editing patterns for virtual cinematographic storytelling. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *14*(4), 1-22.