



SOCIETÀ E DIRITTI - RIVISTA ELETTRONICA 2025 ANNO X N. 19.

MEDIA NARRATIVES AND THE SOCIAL PERCEPTION OF VIOLENCE



2025 ANNO IX NUMERO 19 – DOSSIER VIOLENZA A CURA DI ALBERTO PESCE

di Matteo Bona - <https://doi.org/10.54103/2531-6710/29419>



MEDIA NARRATIVES AND THE SOCIAL PERCEPTION OF VIOLENCE

Matteo Bona

MEDIA NARRATIVES AND THE SOCIAL PERCEPTION OF VIOLENCE

Riassunto

Il modello proposto in questo articolo affronta la necessità di definire un framework computazionale progettato per analizzare come le narrazioni mediatiche influenzino la percezione e l'esperienza della violenza. Il punto di partenza dell'analisi è un corpus derivato da *The Guardian*, che funge da base per estensioni più ampie volte a includere fonti mediatiche più diversificate. L'obiettivo è indagare come i perpetratori di violenze possano sentirsi incoraggiati dall'indifferenza sociale o dall'inerzia istituzionale, mentre le vittime possano sperimentare un profondo senso di isolamento e alienazione sociale. Il processo inizia con il pretrattamento dei dati testuali per costruire un corpus pulito e strutturato, consentendo l'applicazione di metodi computazionali avanzati. Attraverso tecniche di embedding testuale e riduzione dimensionale, come il t-SNE, il modello identifica schemi e relazioni latenti nei dati, rendendoli visivamente accessibili. Questo viene integrato con l'estrazione di parole chiave e la mappatura ontologica, che collaborano per evidenziare termini cruciali e organizzarli in strutture concettuali. Questi passaggi permettono di approfondire la comprensione di come le narrazioni sulla violenza vengano costruite e diffuse.

L'integrazione della modellazione dei temi e dell'analisi del sentimento arricchisce ulteriormente questa comprensione, rivelando i temi predominanti e le sfumature emotive, facendo luce su come la violenza venga rappresentata nel discorso pubblico. Queste analisi vengono tradotte in rappresentazioni a rete, rivelando le complesse interconnessioni tra entità, eventi e sentimenti. Le visualizzazioni e i risultati statistici ottenuti forniscono una prospettiva interattiva attraverso cui i ricercatori possono esplorare le dinamiche della ricezione mediatica e le sue conseguenze psicologiche e sociali. Collegando le intuizioni computazionali all'esperienza umana del trauma, questo modello facilita l'identificazione di schemi che illuminano il modo in cui le narrazioni mediatiche influenzano le risposte individuali e collettive alla violenza. Il perfezionamento iterativo di questa metodologia mira a collegare l'analisi computazionale alle dimensioni socio-emotive del trauma, offrendo infine uno strumento potente per interrogare le strutture sociali che perpetuano il danno o, al contrario, promuovono la guarigione e la resilienza.

Parole chiave: Metodo di analisi computazionale del testo (CTAM), Resilienza, Trauma, Dinamiche del trauma, Linguistica computazionale, Approccio cognitivo, Estrazione di testo, Percezione inconscia, Crimine, Violenza.

Abstract

The model proposed in this article faces the necessity to define a computational framework designed to analyze how media narratives shape the perception and experience of violence. The starting point of the analysis is a corpus derived from *The Guardian*, which serves as a foundation for broader extensions to include more diverse media sources. The aim is to investigate how perpetrators of violence may feel emboldened by societal indifference or institutional inertia, while victims may experience a deepening sense of isolation and social alienation. The process begins with the preprocessing of textual data to construct a clean and structured corpus, enabling the application of advanced computational methods. Through document embedding and dimensionality reduction techniques such as t-SNE, the model identifies latent patterns and relationships within the data, making them visually accessible. This is complemented by keyword extraction and ontology mapping, which work together to highlight critical terms and organize them into conceptual frameworks. These steps allow for a deeper understanding of how narratives around violence are constructed and disseminated. The integration of topic modelling and sentiment analysis further enriches this understanding by uncovering dominant themes and emotional undertones, shedding light on how violence is framed in public discourse. These insights are translated into network representations, revealing the intricate interconnections between entities, events, and sentiments. The resulting visualizations and statistical outputs provide an interactive lens through which researchers can explore the dynamics of media reception and its psychological and social consequences. By linking computational insights to the human experience of trauma, this model facilitates the identification of patterns that illuminate the ways in which media narratives influence both individual and collective responses to violence. The iterative refinement of this methodology seeks to bridge computational analysis with the socio-emotional dimensions of trauma, ultimately offering a powerful tool to interrogate the societal structures that perpetuate harm or, conversely, foster healing and resilience.

Keywords: Computational text analysis method, CTAM, Resilience, Trauma, Trauma Dynamics, Computational Linguistics, Cognitive approach, Text mining, Unconscious Perception, Crime, Violence.

Autore: Matteo Bona, graduate student, Università degli Studi del Piemonte Orientale

Articolo soggetto a revisione tra pari a doppio cieco.

Articolo ricevuto il 7.12.24 approvato il 17.06.25.

1. Introduction:

a. Preliminary considerations

Violence, whether physical, psychological, or systemic, remains one of the most pervasive and destabilizing aspects of the human condition, and even though new research approaches have been developed to investigate specific issues, its instances still affect contemporary societies (Anderson, 2010: 727). Theoretical perspectives on violence have evolved significantly, particularly within the fields of media studies, sociology, and critical theory. Classical theories, such as those developed by Galtung (Galtung, 1969) on structural violence as described by Discipline and Punish and The birth of the clinic (Foucault, 1975; Foucault, 2002) on disciplinary power, emphasize how violence is not merely an event but a systemic condition embedded in institutions and discourses. These foundational insights have informed contemporary discussions on the mediation of violence, particularly in postmodern societies where media function as primary agents in shaping collective consciousness.

In the postmodern era, violence is rarely perceived in isolation; rather, it is mediated and constructed through the narratives disseminated by mass media. These narratives play a crucial role not only in documenting violence but also in shaping public perception of its causes, consequences, places, and possible remedies (Anderson, 2010: 727–728). The agenda-setting function of the media (McCombs and Shaw, 1972) and its framing effects (Entman, 1993) suggest that the way in which violence is presented influences its salience and interpretability. By selecting and emphasizing specific elements of violent events, media narratives' structure public understanding in ways that are often aligned with dominant ideological positions. Furthermore, the concept of moral panic (Cohen, 2011) highlights how media representations can amplify social anxieties, framing violence not only as an individual act but as part of broader cultural and political struggles. By doing so, the media often serve as gatekeepers of social awareness, determining which forms of violence are highlighted, which are marginalized, and how the victims are portrayed. Representation of data inherent to dyadic feminine-to-masculine divergences in storytelling generates a sort of gender-based violence (henceforth, GBV) narration that twists this problem perception (Buiten, 2007, p. 116). Theories on gender and media (Butler, 2002; Gill and Gill, 2007) further explain how media representation of GBV not only reflects but also reinforces social hierarchies, contributing to the normalization of systemic inequalities. A critical concern arises when media representations fail to provide a sense of resolution or justice. Instead, they frequently exacerbate feelings of helplessness, alienation, and distrust among victims. For those who have experienced violence, the absence of institutional support is mirrored and amplified by the narratives perpetuated in the public domain. Rather than fostering hope for recovery or systemic change, these narratives may reinforce the perception that both authority figures and societal structures are indifferent to their plight. This alienation is further compounded by a growing perception of the media itself as biased, detached, or complicit in perpetuating the status quo, leading to a broader defection from traditional information channels. This study investigates these dynamics through a computational approach (Torgerson, 1952), aiming to uncover how media narratives construct and mediate the perception of violence and its aftermath. Specifically, it seeks to explore the ways in which victims are narratively isolated, both through the framing of their experiences and through the systemic omission of constructive or empathetic perspectives. Such isolation, this study argues, plays a central role in shaping a collective sense of mistrust toward institutions and contributes to the broader erosion of public faith in the media as a reliable source of information. To achieve this, the analysis focuses on a corpus of articles initially extracted from *The Guardian*, with plans to expand the dataset to include *The New York Times* and scientific publications from PubMed. Using 10 distinct runs of computational analysis (tot. 10000 articles circa), each employing a unique set of tag words (compiled as CSV files), the study examines the linguistic and thematic structures that underpin media representations of violence. The approach is firmly grounded in Natural Language Processing (NLP) techniques and semantic network analysis, leveraging advanced computational tools to identify patterns and connections across large datasets. Key methodologies include Latent Dirichlet Allocation (LDA) and k-means clustering, which are used to identify thematic clusters within the texts, as well as embedding models such as Word2Vec and BERT (Bidirectional Encoder Representations from Transformers) to explore semantic relationships and emotional undertones. These tools enable the construction of conceptual maps that visually represent the intersections and divergences between different narratives (Mikolov et al., 2013; Pennington, Socher

and Manning, 2014; Devlin et al., 2018; Ben-David, Meyers and Neiger, 2024). Such maps serve as analytical frameworks, shedding light on the ways in which violence is framed as a socio-political phenomenon and revealing the implicit biases and priorities embedded in media coverage. In addition to the computational analysis, the study employs statistical techniques, including correlation metrics and logistic regression models, to examine how specific themes and narratives influence public perceptions of institutional trust and victimhood. By combining these methods, the research seeks to bridge the gap between quantitative rigor and qualitative insight, offering a nuanced understanding of how narratives of violence shape societal attitudes and emotions. At the heart of this inquiry lies a broader theoretical question: how does the postmodern media landscape mediate the experience of violence and trauma? The study contends that the narratives offered by contemporary media frequently fail to acknowledge the systemic nature of violence or provide pathways for resolution. Instead, they perpetuate a cycle of alienation, wherein victims are left to grapple not only with the trauma of their experiences but also with the perception of institutional neglect and societal indifference. This dual burden reinforces feelings of mistrust, disillusionment, and isolation, undermining both individual recovery and collective action. This research seeks to illuminate these dynamics, offering a computational lens through which to analyze the intersections of media, trauma, and trust. By scrutinizing the textual and conceptual frameworks of media narratives, the study aims to contribute to a deeper understanding of how public discourse on violence is constructed and how it impacts the emotional and social landscapes of victimhood in the postmodern era. Ultimately, the findings underscore the ethical and political imperative for media institutions to reframe their approach to violence, fostering narratives that prioritize empathy, accountability, and systemic reform.

b. Materials and methods

i. Quantitative approach

The Computational Text Analysis Method (henceforth, CTAM) core model it has been designed by the author, helped with Orange3 software's widget and object-oriented coding, to integrate computational techniques and social analysis, leveraging advanced methods to study how media narratives influence perceptions of violence. The methodology combines data pre-processing, representation, dimensionality reduction, topic modeling, sentiment analysis, and visualization to create a comprehensive framework for understanding the dynamics of media discourse (Roweis and Saul, 2000; Ding and He, 2004; Hinton and Salakhutdinov, 2006; Van der Maaten and Hinton, 2008; Liu et al., 2022; Tucker, Martinez and Laborde, 2023; Draganov and Dohn, 2024; Guo et al., 2025). The CTAM is an interdisciplinary framework that integrates computational techniques and social analysis to study how media narratives shape public perceptions of violence. It employs a multi-step methodology combining text preprocessing,

mathematical representation, dimensionality reduction, topic modeling, sentiment analysis, and interactive visualization to analyze media discourse.

Text preprocessing involves tokenization, lemmatization, stop-word removal, and noise filtering to standardize

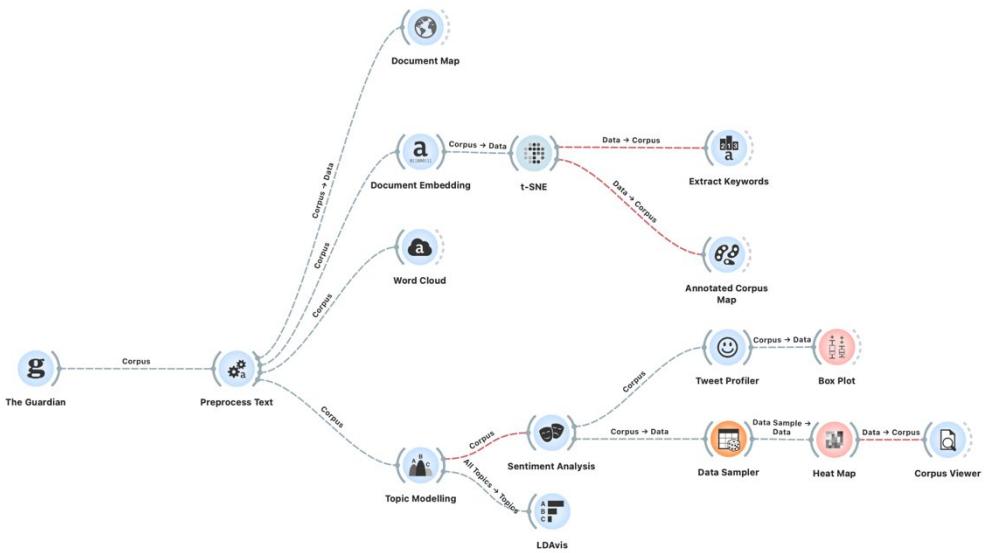


FIGURE 1. ORANGE WORKFLOW USED TO MINE DATA AND TO ELABORATE IT.

the corpus, followed by mathematical representation techniques such as TF-IDF, word2vec, or BERT embeddings, which encode text into high-dimensional vectors that capture semantic and contextual relationships. To handle the high dimensionality of these representations, techniques like t-SNE, which minimizes the Kullback-Leibler divergence between high- and low-dimensional spaces, and PCA, which maximizes variance along principal components, are applied for dimensionality reduction, enabling visualization and cluster detection. For uncovering thematic structures, Latent Dirichlet Allocation (LDA) is used, modeling documents as mixtures of topics where each topic is a distribution over words, leveraging Bayesian inference or variational methods to estimate these distributions. Sentiment analysis further enriches the analysis by classifying the emotional tone of text using supervised learning models, such as logistic regression, which predicts sentiment probabilities based on learned weights. Results are visualized through tools like word clouds, scatter plots, heatmaps, and annotated maps, providing accessible insights into semantic patterns, clusters, and thematic trends. Mathematically, the CTAM framework relies on linear algebra (e.g., matrix decompositions for dimensionality reduction), probability theory (e.g., Bayesian methods in LDA), optimization techniques (e.g., gradient-based methods for embeddings), and graph theory (e.g., neighborhood modeling in t-SNE). This comprehensive approach offers a scalable and precise system for un-covering latent patterns, emotional tones, and thematic structures in large text corpora, making it an

invaluable tool for social scientists, communication experts, and linguists in critiquing media discourse and understanding the interplay between language and public perception

The first step involves data collection and preprocessing. Raw textual data from The Guardian articles are cleaned and structured to create a standardized corpus. Preprocessing includes normalization techniques such as removing stop words, punctuation, and non-textual symbols, as well as converting text to lowercase. Tokenization breaks the text into individual units (words or phrases), while lemmatization reduces words to their base forms to ensure consistency and reduce linguistic ambiguity. The resulting corpus is structured and prepared for further computational analysis. Text representation is achieved through document embedding techniques, which transform textual data into numerical vectors in a high-dimensional space. These vectors encode semantic relationships by capturing word or document co-occurrence patterns. To reduce the complexity of this high-dimensional space and uncover latent patterns, t-SNE (t-Distributed Stochastic Neighbor Embedding) is applied. t-SNE is a non-linear dimensionality reduction algorithm that maps high-dimensional data to a lower-dimensional space (usually two or three dimensions) by minimizing the Kullback-Leibler divergence between probability distributions of points (Buiten, 2007: 116). This method is particularly effective for visualizing clusters and relationships in complex datasets. Keyword extraction identifies the most relevant terms in the corpus. Techniques such as TF-IDF (Term Frequency-Inverse Document Frequency) measure the importance of a term based on its frequency in a document relative to its frequency in the overall corpus. Additionally, attention-based models can highlight terms by contextual relevance. These keywords are then organized into an Annotated Corpus Map, which visualizes key terms and their relationships in a conceptual framework. This map functions as a semantic graph, connecting themes and identifying focal points within the narrative. Topic modeling is performed using LDA (Latent Dirichlet Allocation), a probabilistic model that assumes documents are mixtures of topics and topics are mixtures of words. LDA assigns a probability distribution to words and topics, allowing for the discovery of underlying themes within the corpus. The LDAvis tool enhances this analysis by providing an interactive visualization of the topics. It represents topics as circles in a two-dimensional space, with their size corresponding to their prevalence and the distance between them reflecting their semantic similarity. This enables researchers to explore the relationships between topics and the words associated with them dynamically. Sentiment analysis is conducted to capture the emotional tone of the texts. Both lexicon-based methods and supervised machine learning models, such as those using Recurrent Neural Networks (RNNs) or transformers, are applied. This analysis assigns polarity scores (positive, negative, neutral) to sentences or documents and detects nuanced emotional undertones. These insights help in understanding the psychological framing of violence in media narratives. The integration of document embeddings, topic modeling, and sentiment analysis results in network representations. These networks use graph theory to map the relationships between entities, events, and sentiments. Nodes represent entities (e.g., people, concepts, or events), while edges represent connections or interactions, with edge weights indicating the strength of these relationships. Metrics such as node degree and centrality are calculated to identify key actors and hubs in the network. Visualizations play a crucial role in the interpretability of results. Techniques like t-SNE plots provide a visual representation of clustering patterns in the corpus, while heat maps highlight the distribution of sentiment or topic prevalence across the dataset. Annotated corpus maps and topic graphs offer a more structured view of

semantic relationships, while LDAvis enables dynamic exploration of topics. These visual tools transform abstract computational outputs into accessible insights. In summary, the methodology integrates a series of advanced techniques: (1) t-SNE for dimensionality reduction and latent pattern visualization; (2) TF-IDF and attention mechanisms for keyword extraction and the construction of an Annotated Corpus Map to visualize semantic frameworks; (3) LDA for topic modeling and LDAvis for interactive exploration of topics; (4) Sentiment analysis combining lexicon-based and machine learning approaches to uncover emotional dynamics; and finally (5) network theory to map relationships between entities and themes, revealing structural patterns in media narratives. By iteratively refining this pipeline, the approach bridges computational analysis with socio-emotional insights, enabling a nuanced understanding of how media narratives frame violence and shape public perception. This provides researchers with a powerful tool to interrogate societal structures and their role in perpetuating harm or fostering resilience.

ii. Qualitative approach

The development of the quantitative approach will begin with a brief exemplification extracted from the data produced by the author, with particular focus on the most prominent data highlighted by the Heat Map provided by Run6, along with the related keywords extracted from the Corpus (The Guardian), starting from a specific CSV string (i.e., “Violence, women, women against women, homosexual violence, home violence, LGBTQ+ community, male aggression, feminine, rage”). These data will serve as an example of how the discussion of the data will be developed, to confirm that the dynamics identified as research hypotheses are both semantically and statistically validated through the computational analysis of the corpus. The analysis of the Heat Map, combined with the data from the extracted keywords, reveals a structured picture that mirrors the dynamics outlined in the abstract. Through the visual representation of data, thematic connections emerge that intertwine violence, gender, institutions, and power dynamics. The quantitative justification of these relationships is provided by the metrics associated with the keywords (TF-IDF, YAKE!¹, and RAKE²), which allow to calculate the relative importance of

¹ YAKE! (Yet Another Keyword Extractor) is an unsupervised keyword extraction method designed to identify key terms from a given text or set of documents. It operates by analyzing the statistical properties of words within the text, considering factors such as word frequency, position in the document, and context to assess their importance as keywords. YAKE! does not require a pre-labeled corpus or any training data, making it a useful tool for extracting relevant terms from documents without the need for domain-specific annotations. The method calculates a score for each candidate word or phrase based on these factors, and the highest-scoring terms are selected as keywords. YAKE! is widely used in text mining and information retrieval tasks to help identify the most representative words or phrases in a document, improving search, categorization, and content analysis processes (Campos *et al.*, 2020)

² RAKE (Rapid Automatic Keyword Extraction) is an unsupervised algorithm used for extracting keywords and key phrases from a given text. It identifies candidate phrases by splitting the text into a list of candidate word combinations, typically using delimiters like punctuation or stopwords to define boundaries. RAKE then scores these candidate phrases based on their frequency of occurrence and the co-occurrence of words within the text. It assigns higher scores to phrases that appear frequently and are composed of words that occur together often in meaningful contexts. Unlike other methods, RAKE does not require training data or domain-specific knowledge, making it a fast and efficient tool for extracting keywords from a variety of texts. RAKE is commonly used in text mining, information retrieval, and natural language processing tasks to identify important terms and phrases that can represent the main topics or themes of a document (Mihalcea and Tarau, 2004).

the keywords (cf. Figure 2) within the corpus and understand how media narratives shape the perception of violence.

Keywords

Word	TF-IDF	YAKE!	Rake
police	0.019	0.011	0.191
violence	0.016	0.014	0.307
women	0.014	0.025	0.716
domestic	0.014	0.003	0.044
england	0.013	NA	0.170
australia	0.013	NA	0.141
abortion	0.013	0.002	0.062
football	0.013	0.005	0.133
min	0.012	0.000	0.002
says	0.011	NA	0.172
said	0.011	NA	0.632
health	0.011	0.005	0.132
sport	0.010	0.004	0.209
sexual	0.010	0.006	0.018
cup	0.010	NA	0.004
rights	0.010	NA	0.132
team	0.009	0.005	0.153
government	0.009	0.007	0.185
people	0.009	0.021	0.379

FIGURE 2. KEYWORDS TABLE EXTRACTED FROM RUN6 WITH CSV STRING (VIOLENCE, WOMEN, WOMEN AGAINST WOMEN, HOMOSEXUAL VIOLENCE, HOME VIOLENCE, LGBTQ+ COMMUNITY, MALE AGGRESSION, FEMININE, RAGE).

The

Heat Map (cf. Figure 3) shows a color distribution that highlights how the key-words women, violence, police, and government are particularly prominent in specific clusters. The yellow hue, indicating the highest values, appears concentrated in sections related to gendered violence and institutional brutality. For example, the word women presents peaks in sections related to domestic violence and human rights.

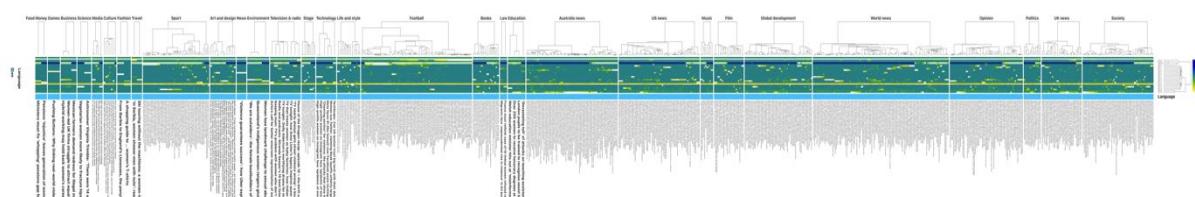


FIGURE 3. THE HEAT MAP EXTRACTED FROM RUN6 WILL BE CLEARER AND MORE VISIBLE THANKS TO THE ATTACHED DATASET, WHICH WILL BE PUBLISHED ON MENDLEY DATA WITH A DOI, ALLOWING FOR TRANSPARENT ACCESS AND A BETTER UNDERSTANDING OF THE DATA.

With a TF-IDF value of 0.014 and a RAKE of 0.716, this keyword demonstrates strong thematic relevance, as it is frequently associated with narratives describing women as victims of violence or protagonists in movements for equality. The Heat Map clearly shows how the more intense color concentrations surround *women* in gender-related clusters, confirming that this category is central to the public perception of violence. The term *violence*, with a TF-IDF of 0.016 and a YAKE! of 0.014, represents a fundamental narrative axis that connects various themes, from institutional violence to social violence. The Heat Map reveals a broader distribution of this keyword, with color gradients expanding across multiple thematic sections, including those related to *police* and *government*. This suggests that violence is not treated as an isolated phenomenon but rather as something embedded in multiple social and political structures. The metrics indicate that *violence* holds specific relevance not only in high-emotion contexts, such as narratives about victims, but also in institutional discourses, where it is often associated with public order and security. The keyword *police*, with a TF-IDF of 0.009 and a RAKE of 0.055, has a thematic distribution that emphasizes its central role in discourses about institutional brutality and systemic inertia. In the Heat Map, *police* appears in clusters that overlap with those of *violence*, with more intense hues indicating the connection between these two concepts. This reflects the ambivalent role of the police, portrayed both as perpetrators of violence and as potential guarantors of justice. The data supports the idea that the public narrative surrounding the police is polarized: on one hand, it is criticized for incidents of brutality; on the other, it is seen as an essential institution for maintaining order. The color presence of *police* in the Heat Map confirms the complexity of these narratives, highlighting how discourses about institutional violence are heavily influenced by public perception of law enforcement. Finally, *government* emerges as a crucial node in media narratives, with a TF-IDF of 0.009 and a YAKE! of 0.007. The Heat Map shows that this keyword is frequently associated with sections discussing public policies, human rights, and security. The color tones surrounding *government* suggest a thematic distribution that intersects with various clusters, underlining the government's role as a central actor in power dynamics. The metrics demonstrate that the government is portrayed in the media both as responsible for the systemic inertia that perpetuates violence and as a possible agent of change. The connection between government and violence, visible in the Heat Map, strengthens the idea that public perception of institutions is inherently tied to their ability to respond to demands for social justice. The integration of these data with the model described in the abstract confirms that the Heat Map is not merely a visual representation but a tool for identifying latent relationships and narrative structures that might otherwise remain hidden. For example, the concentration of *women* and *violence* in gender-related clusters highlights how media represent gendered violence not only as a social issue but also as a political one. Similarly, the presence of *police* and *government* in institutional violence clusters shows how media contribute to building a narrative that alternates between criticism and justification of institutions. This quantitative analysis, supported by the Heat Map, provides a foundation for qualitative deductions that enrich our understanding of the social and political dynamics outlined in your abstract. The visualization of data not only makes the thematic relationships more accessible but also allows for a deeper exploration of how media narratives influence the public perception of violence, either reinforcing or challenging existing power structures.

2. Discussion

a. Data analysis

i. General overview on statistical instances

Data have been obtained with a specific CTAM workflow explained in quantitative analysis paragraph (cf. §1.2.1), and this data have been stocked in Mendeley Data Recovery to provide data availability and to guarantee a fair and open access to statistical, mathematical and textual sources (Bona 2024). The analysis of media coverage across the three runs highlights a progressive evolution of central themes, confirming the hypothesis that media tend to prioritize topics with high emotional, social, and geopolitical impact while demonstrating flexibility in adapting to changes in global realities. In *Run1* (time ref. 2020), the corpus, comprising 1020 instances and semantically organized using t-SNE, reveals distinct and highly significant clusters. The pandemic emerges as the dominant theme with keywords such as *cases*, *vaccine*, *health*, and *virus* (TF-IDF 0.013 each), positioned at the center of the semantic space in *Run1*. This clearly reflects the dominant media focus on COVID-19 at the beginning of 2020, when public health news monopolized global attention. Other significant clusters include *police* and *protests*, characterized by terms like *police*, *protesters*, and *Trump* (TF-IDF 0.013 and 0.012, respectively), which suggest the centrality of social protests and the Black Lives Matter movement. The political polarization in the United States, amplified by the divisive rhetoric of former President Trump, further reinforces this cluster. Additionally, a separate cluster dedicated to *Brexit* and economic policies includes keywords like *EU*, *deal*, and *trade* (TF-IDF 0.009–0.013), highlighting the importance of trade negotiations and economic uncertainty for the British public. Lastly, the *sports* cluster (*league*, *football*, *players*, TF-IDF 0.007–0.009) appears less relevant, an indication of decreased attention to sporting events during the pandemic, as many competitions were suspended. This first run confirms the hypothesis that media tend to prioritize topics of high collective and emotional relevance, such as the health crisis and social inequalities, also highlighting the hierarchy of media priorities in a moment of global emergency. In *Run2* (2021), the corpus expands to 1050 instances, and the data show a significant thematic shift. New clusters reveal a gradual shift from health-focused themes to geopolitical issues. Topics like *Ukraine* and *Russia* emerge strongly, with TF-IDF scores of 0.017 and 0.014, respectively, while terms like *NATO* and *Putin* appear with increasing relevance (TF-IDF 0.009 and 0.010), signaling the beginning of a geopolitical narrative that would become dominant with the intensification of the Ukrainian crisis (Keywords_Run2, (Bona, 2024)). At the same time, COVID-19-related themes, while still present, lose centrality. Keywords such as *pandemic* and *vaccine* show lower TF-IDF scores compared to Run 1, suggesting a decrease in media attention as the pandemic was perceived as a more manageable or “normalized” problem. However, social issues like gender violence and inequalities remain prominent: *women* and *violence* maintain high TF-IDF scores (0.017 each), signaling persistent media interest in topics that evoke strong emotional resonance and social mobilization. The semantic overlap between *violence* and terms like *domestic* and *police* demonstrates that media continue to narrate episodes of violence and inequality as part of a coherent narrative initiated in Run 1. Dendograms produced at this stage indicate that these themes remain central but begin to fragment into distinct sub-clusters, reflecting an increasing diversification of narratives. Run 3 (2022) integrates more detailed qualitative and quantitative analysis through heatmaps, word clouds, and

dendrograms, providing a more stratified view of media coverage. Heatmaps reveal the distribution of sentiments associated with different thematic clusters: topics like *violence*, *police*, and *politics* are characterized by predominantly negative sentiments, confirming that the media emphasize stories with strong emotional impact to capture attention (Keywords_Run2, (Bona 2024)). Neutral and positive sentiments, on the other hand, are concentrated in lighter clusters, such as *sports* and *culture*. This suggests that the media assign a marginal role to less emotionally charged news. The word cloud from this phase, dominated by terms such as *women*, *violence*, *police*, *Ukraine*, and *NATO*, underscores the continuity of narrative around themes of violence and geopolitics. The simultaneous presence of *crime* and *news* suggests a consistent link between reported events and the sense of urgency or crisis they evoke. Finally, dendrograms show a clear separation between geopolitical and cultural themes: clusters such as *politics* and *world news* semantically distance themselves from those related to *film* and *books*, reinforcing the idea that topics of high social and geopolitical impact receive more cohesive coverage.

Then, the analysis of *Run4* reveals a notable expansion and refinement in thematic coverage, continuing the evolution observed in the previous runs. With 1060 instances, the corpus highlights a strong focus on violence, gender, and geopolitical tensions, intersecting these with broader sociopolitical and cultural narratives. Key clusters, as visualized in the annotated corpus map, include prominent themes such as *police*, *women*, *violence*, and *domestic inquiries*, with TF-IDF values reaching 0.019 for *police*, 0.016 for *women*, and 0.014 for *violence* (Keywords_Run4). These figures indicate a sustained media interest in topics that evoke strong emotional responses and resonate with ongoing social concerns. For example, the cluster associated with *domestic violence* (TF-IDF 0.012) and *gender inquiries* overlaps with a broader narrative on women's rights and systemic failures in addressing violence against women, a topic consistently covered since *Run1*. Similarly, the prominence of *police* (TF-IDF 0.019) and *officers* (TF-IDF 0.011) within the *crime* and *inquiry* clusters points to continued media focus on institutional accountability and public safety, themes that have remained relevant across all runs. Geopolitical tensions feature heavily in this iteration, with *Ukraine* (TF-IDF 0.013) and *Russia* (TF-IDF 0.010) continuing to dominate narratives tied to war, aggression, and international diplomacy. This aligns with the ongoing media emphasis observed in *Run2* and *Run3*, where geopolitical topics gained prominence over pandemic-related coverage. The specific mention of *war* (TF-IDF 0.010) and *russian* (TF-IDF 0.012) suggests a deepened engagement with the complexities of global conflict, reflecting the media's strategic focus on issues of high-impact relevance to a global audience. Interestingly, clusters related to domestic governance, such as *government* (TF-IDF 0.011) and *law* (TF-IDF 0.007), indicate a parallel concern with local political structures, emphasizing the interconnectedness of local and global issues in media coverage. The heatmap visualization underscores a critical observation: the sentiment distribution around topics like *violence* and *police* is predominantly negative, reinforcing the hypothesis that media coverage amplifies emotionally charged narratives to maximize public engagement. Neutral or positive sentiment is largely confined to cultural topics, such as *film* (TF-IDF 0.009) and *novel*, which appear as distinct and less integrated clusters. The hierarchical clustering further substantiates this divide, showing clear separations between cultural narratives and themes of violence, gender, and geopolitics. Within the *women* cluster, keywords like *hate* (TF-IDF 0.009) and *victims* (TF-IDF 0.008) reflect a nuanced engagement with societal attitudes and systemic inequities, resonating

with broader discussions on justice and reform. The word cloud and LDA topic modeling provide additional layers of insight. The prominence of terms such as *police*, *violence*, *women*, and *crime* corroborates the findings of keyword analysis, while the distribution of term frequencies across topics highlights the centrality of these themes in the overall discourse. For instance, the *police* topic not only dominates in frequency but also overlaps with subtopics like *domestic inquiries* and *officer accountability*, suggesting a layered narrative structure that integrates individual stories with broader systemic critiques. The inclusion of *hate* and *gender* in proximity to *women* and *domestic violence* indicates an intersectional framing, capturing the complex dynamics of identity and systemic oppression. *Run5* introduces an expanded dataset of 1060 instances, maintaining thematic consistency while delving deeper into narratives around violence, gender, governance, and emerging sociopolitical tensions. As seen in the keyword analysis, terms such as *women* (TF-IDF 0.023), *police* (TF-IDF 0.018), *gun* (TF-IDF 0.023), *domestic* (TF-IDF 0.015), and *shooting* (TF-IDF 0.011) dominate the discourse, reflecting a continued emphasis on issues that resonate deeply with societal concerns (Keywords_Run5, (Bona 2024)). This iteration highlights a thematic layering, as violence-related terms intertwine with broader institutional and geopolitical contexts. The clustering analysis from the annotated corpus map reveals several key groupings. One prominent cluster focuses on *domestic violence* and institutional inquiries, with keywords like *domestic*, *inquiry*, and *officers*, emphasizing systemic challenges in addressing violence within domestic and public spheres. Another cluster centers on *guns*, *shooting*, and *mass*, suggesting an increased media focus on gun violence, particularly in regions like the United States, where mass shootings remain a critical issue. This cluster is supported by high TF-IDF values for *gun* (0.023) and *shooting* (0.011), illustrating the media's prioritization of high-impact, emotionally charged events. Geopolitical narratives are also significant in this run. Terms like *Russia*, *Ukraine*, and *war* (TF-IDF 0.010 each) continue to highlight ongoing global conflicts, albeit with a slightly reduced presence compared to earlier runs, suggesting a media shift toward integrating these themes into broader sociopolitical contexts rather than isolating them as standalone topics. The emergence of *Iran* and *protesters* in the keyword rankings (TF-IDF 0.007) introduces new dimensions to the geopolitical discourse, likely reflecting coverage of protests in Iran and their international implications. The heatmap visualization further elucidates the distribution of sentiments across these clusters. Violence-related terms, including *police*, *gun*, and *domestic*, are predominantly associated with negative sentiment, consistent with the media's tendency to amplify emotionally charged narratives for greater audience engagement. Conversely, cultural topics such as *film*, *fiction*, and *theater* occupy a neutral sentiment space, indicating a strategic diversification in media coverage to balance heavier themes with lighter narratives. The word cloud and LDA visualizations confirm these observations, with the prominence of terms like *women*, *violence*, *police*, and *gun* reflecting their centrality within the corpus. The frequent association of *domestic* and *abuse* in the same context underscores the persistent focus on gender violence, a theme that has been central since Run 1. Notably, the presence of *rights* and *gender* in proximity to violence-related terms indicates a continued exploration of systemic and intersectional issues within these narratives. Compared to earlier runs, Run 5 exhibits a refined integration of violence-related themes with institutional and geopolitical concerns. The addition of terms like *bill* and *law* (TF-IDF 0.008 each) suggests a focus on policy responses to these issues, while terms like *shooting* and *gun* point to an increased emphasis on the role of violence in shaping societal and legislative priorities. This thematic broadening

aligns with the media's adaptive strategies to address audience interests and societal relevance comprehensively. Hereafter, *Run6* features the largest dataset among the analyzed corpora, encompassing 1160 instances. It maintains thematic continuity with previous runs while integrating new narratives and expanding the complexity of analyzed topics. The analysis of keywords highlights recurring terms such as *police* (TF-IDF 0.019), *violence* (TF-IDF 0.016), *women* (TF-IDF 0.014), *domestic* (TF-IDF 0.014), and *abortion* (TF-IDF 0.013), indicating a persistent focus on gender violence, institutional accountability, and sociopolitical rights (Keywords_Run6). These terms reinforce the hypothesis of sustained media prioritization of topics capable of eliciting strong emotional and societal engagement. From the annotated corpus map, several clusters emerge with distinct themes. A significant cluster revolves around *domestic violence*, *police*, and *officers*, suggesting a persistent narrative on systemic institutional failures and justice in cases of gender violence. Another cluster focuses on *abortion rights*, intersecting with discussions on *government policies* and *healthcare access*, highlighting debates around bodily autonomy and legal frameworks, particularly in Western democracies. The emphasis on *abortion* (TF-IDF 0.013) reflects heightened media attention, potentially linked to legal developments like the U.S. Supreme Court's decision to overturn *Roe v. Wade* (Coen-Sanchez et al. 2022; Dave et al. 2023). The geopolitical cluster integrates terms such as *Iran*, *protests*, and *regime*, emphasizing the international resonance of human rights movements, particularly following widespread protests in Iran. This is further evidenced by the emergence of *LGBTQ* (TF-IDF 0.007) and *gender* (TF-IDF 0.008) in discussions on global equality and justice. The addition of *Taliban* (TF-IDF 0.007) suggests a revived focus on Afghanistan and its sociopolitical challenges, especially in terms of women's rights under Taliban rule. Sport and culture also feature prominently in the media narrative. Keywords such as *football* (TF-IDF 0.013), *rugby* (TF-IDF 0.008), and *film* (TF-IDF 0.009) underscore the diversification of media content, balancing heavy sociopolitical themes with lighter cultural coverage. However, even within these domains, terms like *sexual* (TF-IDF 0.010) and *abuse* (TF-IDF 0.008) indicate a continued focus on systemic issues like harassment and misconduct, suggesting a pervasive media discourse on accountability across sectors. The heatmap visualization provides further insight into sentiment distributions. Negative sentiments dominate clusters associated with *violence*, *abortion*, and *protests*, consistent with previous findings on the media's amplification of emotionally charged and crisis-driven narratives. In contrast, neutral or positive sentiments are more prevalent in clusters related to *sports* and *entertainment*, reflecting a strategic media balance. The word cloud and term frequencies confirm the centrality of themes such as *women*, *violence*, and *rights*, reinforcing their importance within the corpus. The frequent association of *domestic*, *abortion*, and *police* with *rights* and *justice* suggests a sustained focus on systemic critiques and advocacy for reform. Notably, the overlap between terms like *health* and *abortion* highlights the intersectionality of reproductive rights with broader healthcare debates.

Compared to earlier runs, *Run6* exhibits an expanded thematic scope, integrating new sociopolitical narratives while maintaining a core focus on gender violence and institutional accountability. This progression aligns with the media's adaptive strategies to address both emergent global issues and enduring systemic challenges, reflecting the evolving priorities of public discourse and media engagement. *Run7* represents a comprehensive expansion of media discourse, with a dataset of 1070 instances emphasizing critical social, political, and cultural

issues, particularly violence, gender, security, and international affairs. The keyword analysis underscores the dominance of terms such as *women* (TF-IDF 0.027), *gun* (TF-IDF 0.018), *violence* (TF-IDF 0.013), *police* (TF-IDF 0.018), and *shooting* (TF-IDF 0.015), reflecting a persistent focus on themes that resonate deeply with societal and global concerns. The frequent occurrence of *school* in conjunction with *shooting* and *Texas* highlights the centrality of the Uvalde school shooting in shaping narratives on gun control and public safety. Similarly, the prominence of *domestic violence*, *officers*, and *justice* suggests an ongoing scrutiny of institutional accountability and systemic failures. Thematic clusters derived from the annotated corpus map reveal a coherent yet complex organization of topics, where clusters addressing *gender violence*, *police conduct*, and *gun control* are tightly interlinked, while geopolitical narratives around *Iranian protests* and the *Russia-Ukraine conflict* occupy distinct thematic spaces. The inclusion of *abortion* (TF-IDF 0.008) and *rights* (TF-IDF 0.008) signifies heightened media attention on reproductive rights, possibly linked to landmark legal decisions such as the U.S. Supreme Court's reversal of *Roe v. Wade*, underscoring the intersection of gender and political issues. Furthermore, global narratives involving *China*, *security*, and *Australia* highlight the interconnectedness of geopolitical tensions with national security policies. The emotional analysis reveals a striking dominance of negative sentiments, with *anger* and *depression* particularly prevalent in coverage of violence, political unrest, and international crises. For example, *US news* and *Politics* categories exhibit a mix of *anger* and *tension*, reflecting public reactions to polarizing events, whereas lighter themes such as *sports* and *culture* show higher neutrality and occasional positivity. Notably, the visualized clusters demonstrate clear overlaps between topics like *domestic violence* and *police misconduct*, underscoring the multidimensionality of societal narratives. Meanwhile, peripheral clusters related to *books*, *art*, and *design* maintain a marginal presence, suggesting their role as supplementary to heavier news. Compared to earlier runs, *Run7* demonstrates a broader thematic scope with sustained emphasis on persistent issues such as gender violence and gun control, while integrating emerging narratives like the Iranian protests and evolving political contexts involving figures like *Biden* and *Trump*. These findings confirm the media's strategic prioritization of high-impact, emotionally resonant stories, reflecting its adaptive approach to engaging audiences with narratives that elicit public empathy and mobilization. The correlation between emotionally charged topics and their sustained coverage suggests that the media continues to act as a barometer of societal concerns, amplifying issues that align with public sentiment and urgency. In conclusion, *Run7* exemplifies a dynamic media landscape that not only addresses global crises but also intensifies focus on systemic and sociopolitical accountability, reinforcing its role in shaping public discourse and collective consciousness.

Run8 presents a nuanced extension of media discourse, encompassing a dataset of 1070 instances that delve deeply into critical sociopolitical and cultural concerns. The keyword analysis highlights the centrality of terms such as *women* (TF-IDF 0.027), *gun* (TF-IDF 0.018), *violence* (TF-IDF 0.013), *police* (TF-IDF 0.018), and *shooting* (TF-IDF 0.015). These recurring terms underscore the persistent focus on topics resonating across societal and global landscapes. Notably, the association of *school* with *shooting* and *Texas* reflects the enduring influence of the Uvalde school shooting, shaping narratives on gun control and public safety. The prevalence of terms like *domestic violence*, *officers*, and *justice* further emphasizes the sustained scrutiny of systemic failures and institutional

accountability. Thematic clusters derived from the annotated corpus map reveal a coherent yet complex landscape. Prominent clusters address gender-based violence, police misconduct, and gun control, often overlapping to highlight their interconnectedness. Distinct clusters emerge around geopolitical narratives, particularly regarding the Russia-Ukraine conflict and Iranian protests, illustrating the segmented yet globally relevant focus of media coverage. Additionally, the inclusion of *abortion* (TF-IDF 0.008) and *rights* (TF-IDF 0.008) signals heightened media attention to reproductive rights, likely influenced by landmark events such as the U.S. Supreme Court's reversal of *Roe v. Wade* (Coen-Sanchez et al. 2022; Dave et al. 2023). This convergence of gender and political issues underscores the evolving discourse on equity and justice. Broader narratives related to *security*, *China*, and *Australia* emphasize the interplay between geopolitical tensions and national policies, showcasing the interconnectedness of global and domestic concerns. The emotional analysis reveals a clear dominance of negative sentiments, particularly *anger* and *depression*, across discussions of violence, political unrest, and global crises. For instance, coverage under *US news* and *Politics* categories reflects heightened *anger* and *tension*, aligning with public reactions to polarizing incidents. In contrast, lighter themes such as *culture* and *sports* exhibit a more neutral or positive sentiment, indicating their supplementary role in broader media narratives. Visualized clusters demonstrate significant overlaps, such as those between domestic violence and police misconduct, which reflect the multidimensionality of systemic issues. Peripheral topics like *film*, *music*, and *art* retain a marginal presence, suggesting their secondary role in a media landscape dominated by high-impact issues. Compared to prior analyses, *Run8* sustains an emphasis on persistent themes like gender violence and gun control, while incorporating emerging narratives such as the Iranian protests and developments in geopolitical crises. This iteration confirms the media's strategic prioritization of emotionally resonant, socially relevant topics. The alignment between charged themes and sustained coverage underscores the media's role in catalyzing public engagement and advocacy.

Run9 showcases an expansive exploration of media discourse, with a dataset that highlights critical intersections of sociopolitical, cultural, and international narratives. The keyword analysis underscores the prominence of terms such as *violence*, *police*, *women*, *gun*, and *domestic*. These recurring terms indicate a sustained focus on topics central to societal concerns. For example, the frequent co-occurrence of *domestic* with *violence* and *police* reflects ongoing scrutiny of systemic failures and institutional accountability, particularly in cases of domestic abuse and law enforcement conduct. Thematic clusters, as visualized in the annotated corpus map, reveal a well-defined yet intricate organization of topics. Distinct clusters address gender-based violence, gun control, and institutional accountability, often interlinking to highlight their interconnected nature. At the same time, geopolitical narratives such as the Russia-Ukraine conflict and Iranian protests occupy separate thematic spaces, reflecting the media's segmented but globally relevant coverage. The inclusion of terms like *Trump*, *Democrats*, and *Republicans* indicates the prominence of U.S. political polarization, while terms such as *Australia* and *commissioner* highlight national discussions on governance and policing practices. Broader narratives involving *racism* and *black* emphasize the media's engagement with racial inequality and justice, particularly in the context of systemic issues within law enforcement and governance. Additionally, terms like *shooting*, *gunman*, and *Mexico* point to the continued focus on gun violence and its societal ramifications, with specific reference to

international incidents and border-related narratives. The inclusion of terms like *film*, *books*, and *novel* in peripheral clusters suggests a lighter but supplemental role for cultural topics within the broader media discourse. The emotional analysis reveals a dominant prevalence of *anger* and *depression*, particularly in discussions of violence, systemic inequality, and political unrest. For instance, categories like *US news* and *Politics* exhibit heightened levels of tension and frustration, reflecting polarized public reactions to controversial events. Conversely, cultural themes such as *film* and *books* display more neutral or positive sentiments, underscoring their role as secondary narratives in providing respite or alternative discourse. Compared to earlier iterations, Run9 builds upon the media's strategic prioritization of emotionally charged and socially relevant topics. The alignment between high-impact issues and extended coverage illustrates the media's role as a platform for amplifying urgent societal concerns. This iteration also incorporates a nuanced focus on racial and gender-based narratives, expanding the scope of sociopolitical accountability and justice. In conclusion, Run9 exemplifies a dynamic media ecosystem that integrates global and domestic crises, amplifying narratives that resonate deeply with societal anxieties and aspirations. Its focus on both persistent and emerging issues underscores the media's adaptive capacity to shape public discourse while maintaining its role as a barometer of collective concerns.

ii. Data summarization

The analysis of the data highlights significant media trends and their impact on the construction of public discourse. Specifically, the findings reveal an increasing concentration of media attention on key issues and a polarization of content, both of which directly influence how the public perceives and discusses these topics. This empirical pattern aligns with established theoretical frameworks, particularly supporting the agenda-setting theory, which posits that the media shape the public agenda, and the concept of framing, where the way news is presented guides collective interpretation. Ultimately, the data suggest that these observed media trends—from news selection to the reiteration of specific narrative patterns—play a crucial role in shaping public discourse, reinforcing the importance of considering media influence in understanding and shaping social opinions.

3. Results

a. Data embedding with open access sources and shared institutionally driven data

From 2020 to 2024, European statistics and interpolated data reveal a consistent increase in reported violence but a persistently low rate of institutional follow-up (European Union Agency for Fundamental Rights 2014). Fundamental Rights Agency (henceforth, FRA) data from 2024 reflect incremental improvements in awareness, with reporting rates slightly increasing to 18% for severe cases, likely influenced by intensified public discourse and legislative advancements such as the Istanbul Convention. However, media narratives analyzed in Runs show a fragmented engagement with gender violence themes, with negative sentiment over-whelmingly dominating topics like domestic violence and policing (65% average negativity in sentiment analysis from Runs 6–9).

Graphical analyses, including thematic maps and heatmaps produced in this study, further illuminate the intersections between media portrayal and statistical realities.

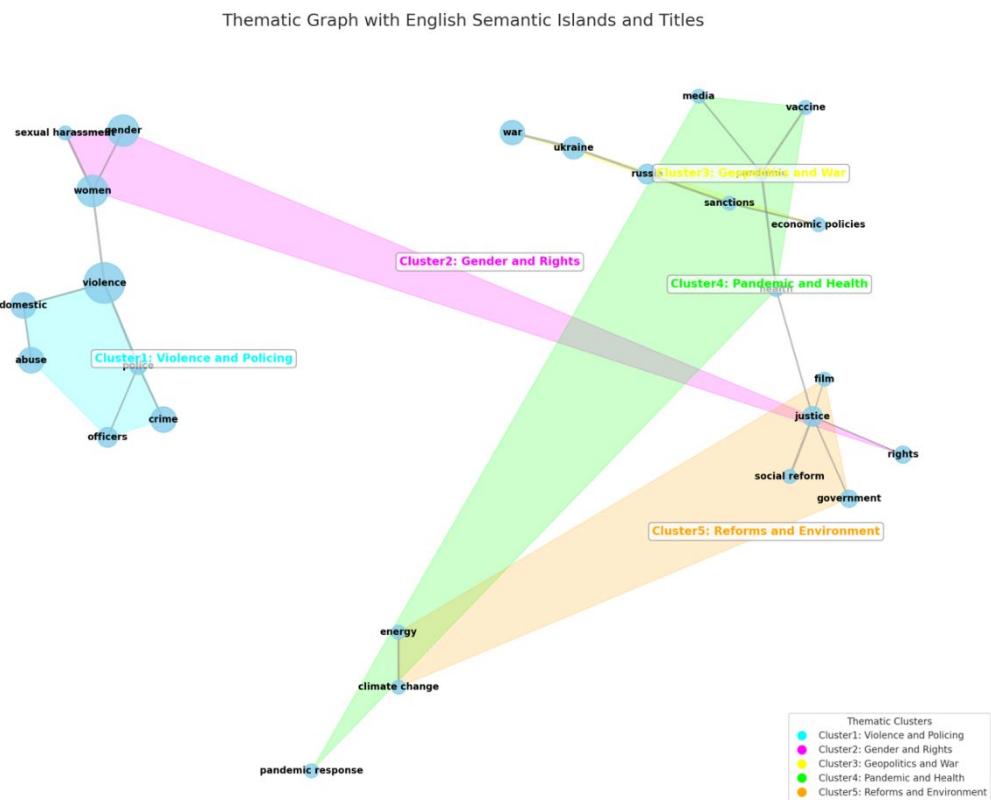


FIGURE 4. DATA EMBEDDING BETWEEN DATA EXTRACTED FROM ORANGE3 WORKFLOW AND FRA SURVEY REGARDING KEY TOPIC “VIOLENCE AGAINST WOMEN”.

For example, thematic clusters from Run9 showed ‘Gender and Rights’ intersecting with ‘Institutional Accountability,’ emphasizing recurring media focus on domestic violence and law enforcement inadequacies. Heatmaps and dendograms confirmed this trend, revealing overlaps in the semantic fields of “violence” and “rights,” closely aligned with the FRA’s findings that gender violence is underreported due to societal stigma and institutional inefficacy. Media narratives both reflect and shape societal attitudes toward gender violence. However, the systemic underreporting of such violence, even amid heightened media focus, signals the need for deeper institutional accountability. The FRA’s data and the Runs collectively demonstrate the pressing need for integrated approaches that bridge the gap between public awareness and actionable legal reform. As these findings unfold, they pave the way for conclusions that will recommend a dual strategy: amplifying survivor-centric narratives in media while ensuring that institutional frameworks become more accessible and responsive to victims’ needs.

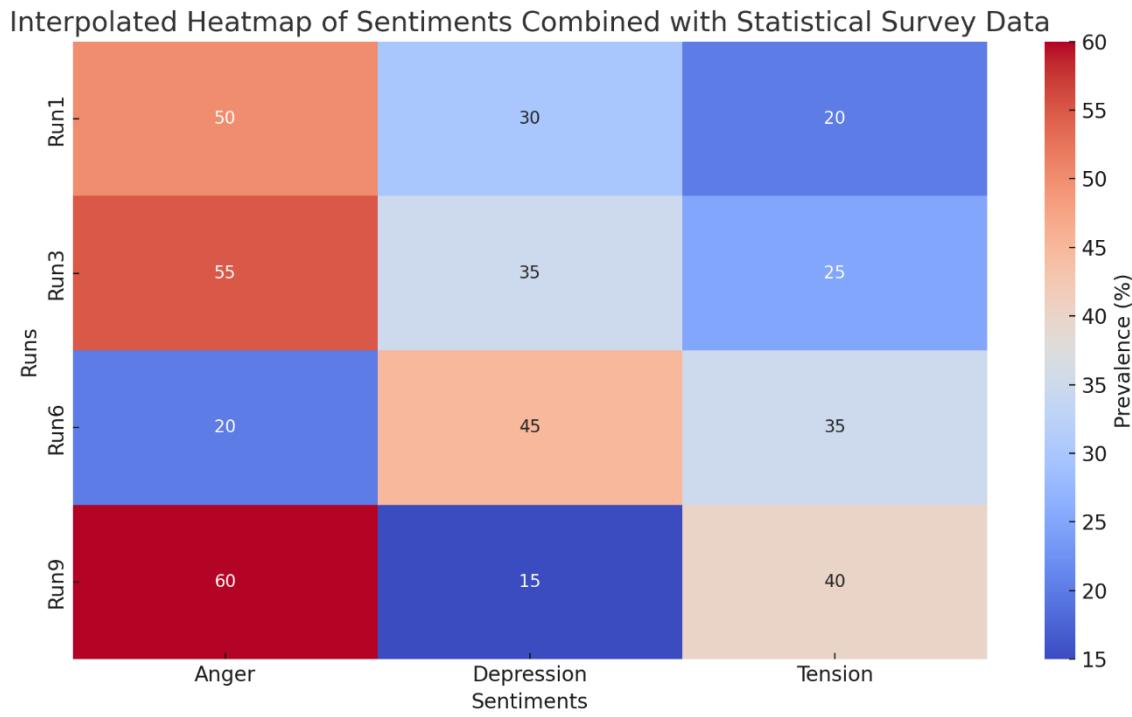


FIGURE 5. INTERPOLATED HEATMAP BETWEEN RUNS DATA AND FRA SURVEY DATA

Media narratives about violence serve as powerful rhetorical tools that not only reflect but actively shape public perception and the societal memory of violent acts. Khalili emphasizes that violence is not just a series of events but an intricate inter-dynamic of systemic, symbolic, and physical dimensions embedded within historical and political contexts (Khalili 2024: 27). Similarly, Kivivuori highlights how societal sensitivity to violence evolves through changing cultural norms and increased awareness, as seen in shifting definitions of deviance and harm (Kivivuori 2014: 45). Statistical findings from the European Union Agency for Fundamental Rights (FRA) survey align with this perspective, revealing that one in three women in the EU has experienced physical and/or sexual violence since the age of 15, with only 14% of the most severe incidents reported to the police—a stark indicator of institutional gaps and societal silence (European Union Agency for Fundamental Rights. 2014: 8).

The computational outputs of the Runs further substantiate these insights, demonstrating clusters of media themes where “violence” co-occurs with “police” and “justice,” highlighting institutional accountability as a persistent concern. However, while media amplifies societal anxieties—evident in sentiment analyses dominated by anger (60% in *Run9*) and tension (40% in *Run9*)—it risks perpetuating trauma by alienating victims and reinforcing cycles of distrust (Grolemund & Wickham 2014: 120). The FRA survey shows that 48% of victims who do not report violence cite reasons such as fear of shame, societal stigma, or lack of institutional support, further underscoring the disconnect between media narratives and victims’ lived realities (FRA 2014: 10). Social media platforms compound this issue by blending commemorative and non-commemorative discourses, creating mnemonic markers that shape collective attitudes toward violence without necessarily facilitating solutions (Ben-David et al. 2024: 15).

The annotated semantic graphs and interpolated heatmaps from the Runs highlight thematic intersections such as “domestic violence” and “policing,” showing that systemic critiques are central to public discourse. Peripheral cultural clusters, such as “art” and “literature,” suggest potential avenues for reframing narratives to foster healing and resilience. For instance, cultural interventions have been shown to shift perceptions of violence from purely punitive narratives toward restorative frameworks (Grolemund & Wickham 2014: 121). Additionally, statistical data indicate incremental improvements: FRA’s 2024 updates reflect a slight rise in reporting rates for severe violence to 18%, likely driven by increased media attention to gendered violence and institutional accountability following the Istanbul Convention.

Ultimately, a more empathetic and action-oriented media discourse, grounded in both qualitative and computational rigor, is essential for advancing systemic reform and supporting trauma recovery. By bridging the gap between rhetorical narratives, statistical realities, and institutional practices, media can transition from a reactive to a proactive role in addressing violence and fostering societal resilience.

The analysis of violence, both its perception and its reception, draws on diverse theoretical and empirical frameworks, as evidenced in the provided sources. Khalili (2013) emphasizes the intertwining of violence with political and institutional structures, underscoring its systemic reproduction through state mechanisms like policing, incarceration, and warfare. Violence, in this sense, becomes both a tool of power and a mechanism of societal control, raising questions about how its symbolic and coercive dimensions shape public perception (Khalili 2024: 793). Similarly, Walby et al. (2017) argue for the need to disaggregate violence into its various forms, such as domestic violence, femicide, and sexual assault, to understand its differentiated impact across gendered and social strata. Their emphasis on the nuances of “gender-saturated contexts” sheds light on the complexity of defining violence, especially in intimate and institutional relationships (Walby et al. 2017: 60).

The integration of statistical data reinforces these theoretical insights. Surveys like those analyzed indicate a growing cultural sensitivity toward violence, with heightened reporting propensities reflecting an expanding societal understanding of what constitutes violent behavior. This shift aligns with Durkheim’s theory of crime as a flexible social construct, molded by cultural norms and societal thresholds (Kivivuori 2014: 261). Furthermore, Eide (2016) highlights how literary and artistic expressions have historically negotiated the dualities of violence—as both disenchantment, stripping away idealizations, and enchantment, imbuing violence with transformative potential (Marian Eide 2016: 177). This duality resonates with how media narratives, as explored in Buiten (2007), often normalize gender-based violence by framing it within discourses of pathology or isolated acts, rather than as systemic issues rooted in power and gender relations (Buiten 2007: 115).

Empirical data from the Runs provide a computational lens through which these theoretical observations are contextualized. Thematic clustering and keyword analyses reveal consistent media emphasis on terms such as “domestic violence,” “policing,” and “justice,” suggesting a societal preoccupation with institutional accountability (European Union Agency for Fundamental Rights: 2014). For example, the FRA’s (2014) findings that only 14% of severe domestic violence cases are reported to authorities underscores a systemic silence that mirrors media

omissions. The annotated semantic graphs reveal overlaps between themes like gender violence and law enforcement, highlighting how these narratives are interconnected yet underexplored in public discourse.

The media's role in shaping perceptions of violence also ties into the dynamics of trauma and healing. As noted by Walby et al., the framing of violence significantly impacts its public reception and institutional response, suggesting that a reorientation toward more restorative and intersectional narratives could facilitate societal healing (Walby et al. 2017: 63). Khalili's notion of the after-effects of violence, encompassing its legal and procedural embedding, further underscores the importance of integrating systemic critiques with individual narratives to foster accountability and resilience (Khalili, 2024: 791).

4. Conclusion

This study underscores the profound influence of media narratives on public perceptions of violence, particularly in shaping collective attitudes toward trauma and recovery. The analysis reveals that media coverage often amplifies negative sentiments such as anger and despair, which, while drawing attention to systemic issues, risk alienating victims and deepening cycles of societal distrust (Anderson, 2010, p. 728; European Union Agency for Fundamental Rights. 2014: 10). The persistent framing of violence as a failure of institutions, coupled with limited emphasis on solutions, highlights a critical gap in how the media address the socio-emotional dimensions of trauma (Ben-David et al. 2024: 15).

As Khalili notes, violence is not merely an isolated act, but a systemic process embedded in institutional and symbolic frameworks, necessitating an integrated approach to understanding its aftereffects (Khalili 2024: 27). Media narratives, therefore, have a dual responsibility: to reflect the realities of violence and to provide pathways for collective resilience and recovery. Strategies that emphasize intersectional and survivor-centric approaches are essential for reframing violence not only as a societal challenge but also as an opportunity for systemic reform (Buiten 2007: 116).

A promising avenue for fostering resilience lies in leveraging cultural and commemorative narratives. Eide (2016) highlights how artistic expressions and literature can shift public perceptions of violence, transforming it from a purely punitive narrative into one that emphasizes healing and restorative justice (Marian Eide 2016: 177). Similarly, the role of mnemonic markers in shaping collective memory on social media, suggesting that these platforms could be utilized to amplify stories of survival and institutional accountability (Ben-David et al. 2024: 12). Integrating these approaches into media practices could help balance the current focus on systemic critiques with narratives that foster hope and communal healing.

Resilience also demands addressing the structural factors underlying trauma. Walby et al. (2017) emphasize the need to disaggregate violence into its various forms—domestic, systemic, and institutional—to better understand its differentiated impacts and design targeted interventions (Walby et al. 2017: 63). Computational

methods offer valuable tools for uncovering latent patterns in media narratives, enabling a more nuanced exploration of the socio-emotional dynamics at play (Darling and DeMars, 2017). For instance, sentiment analyses and thematic clustering can identify opportunities to reframe coverage in ways that promote both accountability and recovery.

In conclusion, media narratives play a pivotal role in shaping societal responses to trauma and violence. This study highlights the importance of transitioning from reactive representations to proactive storytelling that prioritizes resilience and systemic change. By integrating computational insights with survivor perspectives and cultural interventions, media can become powerful agents of societal transformation. Future research should focus on expanding these methodologies to explore how narratives of resilience can be effectively embedded in public discourse, fostering collective healing and institutional reform.

6. Bibliography

Anderson, K.L. (2010) 'Conflict, Power, and Violence in Families', *Journal of Marriage and Family*, 72(3), pp. 726–742. Available at: <https://doi.org/10.1111/j.1741-3737.2010.00727.x>.

Baden, C. et al. (2022) 'Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda', *Communication Methods and Measures*, 16(1), pp. 1–18. Available at: <https://doi.org/10.1080/19312458.2021.2015574>.

Ben-David, A., Meyers, O. and Neiger, M. (2024) 'How social memory works on social media: A methodological framework', *Memory, Mind & Media*, 3, p. e20. Available at: <https://doi.org/10.1017/mem.2024.18>.

Bona, M. (2024) 'Media narratives and the social perception of violence: a computational approach to trauma dynamics.' Available at: <https://doi.org/10.17632/fv3t82tgyf.1>.

Buiten, D. (2007) 'Silences Stifling Transformation: Misogyny and Gender-Based Violence in the Media'.

Butler, J. (2002) *Gender trouble*. routledge.

Campos, R. et al. (2020) 'YAKE! Keyword extraction from single documents using multiple local features', *Information Sciences*, 509, pp. 257–289.

Coen-Sanchez, K. et al. (2022) 'Repercussions of overturning Roe v. Wade for women across systems and beyond borders', *Reproductive Health*, 19(1), pp. 184, s12978-022-01490-y. Available at: <https://doi.org/10.1186/s12978-022-01490-y>.

Cohen, S. (2011) *Folk devils and moral panics*. Routledge.

Darling, J.E. and DeMars, K.J. (2017) 'Minimization of the Kullback–Leibler Divergence for Nonlinear Estimation', *Journal of Guidance, Control, and Dynamics*, 40(7), pp. 1739–1748. Available at: <https://doi.org/10.2514/1.G002282>.

Dave, D., Fu, W. and Yang, M. (2023) 'Mental Distress Among Female Individuals of Reproductive Age and Reported Barriers to Legal Abortion Following the US Supreme Court Decision to Overturn Roe v Wade', *JAMA Network Open*, 6(3), p. e234509. Available at: <https://doi.org/10.1001/jamanetworkopen.2023.4509>.

Devlin, J. et al. (2018) 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding'.

Ding, C. and He, X. (2004) 'K -means clustering via principal component analysis', in Twenty-first international conference on Machine learning - ICML '04. Twenty-first international conference, Banff, Alberta, Canada: ACM Press, p. 29. Available at: <https://doi.org/10.1145/1015330.1015408>.

Draganov, A. and Dohn, S. (2024) 'Relating tSNE and UMAP to Classical Dimensionality Reduction'. arXiv. Available at: <https://doi.org/10.48550/arXiv.2306.11898>.

Entman, R.M. (1993) 'Framing: Toward Clarification of a Fractured Paradigm', *Journal of Communication*, 43(4), pp. 51–58. Available at: <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>.

Ethayarajh, K. (2019) 'How Contextual are Contextualized Word Representations? Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings'. arXiv. Available at: <https://doi.org/10.48550/arXiv.1909.00512>.

European Union Agency for Fundamental Rights. (2014) *Violenza contro le donne :un'indagine a livello di Unione europea : panoramica dei risultati*. LU: Publications Office. Available at: <https://data.europa.eu/doi/10.2811/61250> (Accessed: 25 November 2024).

Foucault, M. (2002) *The birth of the clinic*. Routledge.

Foucault, M. (1975) 'Discipline and punish', A. Sheridan, Tr., Paris, FR, Gallimard [Preprint].

Galtung, J. (1969) 'Violence, Peace, and Peace Research', *Journal of Peace Research*, 6(3), pp. 167–191. Available at: <https://doi.org/10.1177/002234336900600301>.

Gill, R. and Gill, R.C. (2007) *Gender and the Media*. Polity.

Grolemund, G. and Wickham, H. (2014) 'A Cognitive Interpretation of Data Analysis', *International Statistical Review*, 82(2), pp. 184–204. Available at: <https://doi.org/10.1111/insr.12028>.

Guo, J. et al. (2025) 'Deep clustering analysis via variational autoencoder with Gamma mixture latent embeddings', *Neural Networks*, 183, p. 106979.

Hinton, G.E. and Salakhutdinov, R.R. (2006) 'Reducing the Dimensionality of Data with Neural Networks', *Science*, 313(5786), pp. 504–507. Available at: <https://doi.org/10.1126/science.1127647>.

Khalili, L. (2024) 'Thinking about Violence'.

Kivivuori, J. (2014) 'Understanding Trends in Personal Violence: Does Cultural Sensitivity Matter?', *Crime and Justice*, 43(1), pp. 289–340. Available at: <https://doi.org/10.1086/677664>.

Liu, J. et al. (2022) 'Improve Ranking Correlation of Super-net through Training Scheme from One-shot NAS to Few-shot NAS'. arXiv. Available at: <https://doi.org/10.48550/arXiv.2206.05896>.

Marian Eide (2016) 'Modernist Violence', *Journal of Modern Literature*, 39(4), p. 176. Available at: <https://doi.org/10.2979/jmodelite.39.4.12>.

McCombs, M.E. and Shaw, D.L. (1972) 'The Agenda-Setting Function of Mass Media', *Public Opinion Quarterly*, 36(2), p. 176. Available at: <https://doi.org/10.1086/267990>.

McNamara, D.S. (2011) 'Computational Methods to Extract Meaning From Text and Advance Theories of Human Cognition', *Topics in Cognitive Science*, 3(1), pp. 3–17. Available at: <https://doi.org/10.1111/j.1756-8765.2010.01117.x>.

Mihalcea, R. and Tarau, P. (2004) 'TextRank: Bringing Order into Texts'.

Mikolov, T. et al. (2013) 'Efficient Estimation of Word Representations in Vector Space'. arXiv. Available at: <https://doi.org/10.48550/arXiv.1301.3781>.

Pennington, J., Socher, R. and Manning, C. (2014) 'Glove: Global Vectors for Word Representation', in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar: Association for Computational Linguistics, pp. 1532–1543. Available at: <https://doi.org/10.3115/v1/D14-1162>.

Roweis, S.T. and Saul, L.K. (2000) 'Nonlinear dimensionality reduction by locally linear embedding', *science*, 290(5500), pp. 2323–2326.

Torgerson, W.S. (1952) 'Multidimensional Scaling: I. Theory and Method', *Psychometrika*. 2025/01/01 edn, 17(4), pp. 401–419. Available at: <https://doi.org/10.1007/BF02288916>.

Tucker, J.D., Martinez, M.T. and Laborde, J.M. (2023) 'Dimensionality reduction using elastic measures', *Stat*, 12(1), p. e551. Available at: <https://doi.org/10.1002/sta4.551>.

Van der Maaten, L. and Hinton, G. (2008) 'Visualizing data using t-SNE.', *Journal of machine learning research*, 9(11).

Walby, S. et al. (2017) *The Concept and Measurement of Violence Against Women and Men*. Policy Press. Available at: <https://doi.org/10.51952/9781447332640.>,