

# Humanitarian Crisis Coverage Report



Marius Dragomir

Benjamin Marks

August 2025

# About the project

The **Humanitarian Crisis Coverage Report** is part of MJRC's **Media Content Analysis Series**, which focuses on systematically examining media output to uncover patterns in coverage, bias, framing, and editorial choices. This portfolio includes both thematic studies—such as crisis coverage, disinformation, and political framing—and cross-national comparisons of media narratives. In recent years, MJRC has integrated AI-driven tools and machine learning models into its methodology, enabling large-scale analysis of news texts, sentiment, and visibility trends across multiple languages and platforms. The work as part of this portfolio combines computational analysis with media research to expose trends in global media attention.

## Authors

### Marius Dragomir

Marius Dragomir is an award-winning media scholar and journalist, now serving as the director of the Media and Journalism Research Center (MJRC), a global think tank dedicated to the study of media, journalism, politics, and technology. He is also a researcher at the University of Santiago de Compostela in Spain and has served as a visiting professor at Central European University (CEU) in Vienna. Dragomir has received the 2025 Media and Democracy Karol Jakubowicz Award for his original and rigorous research on media ownership and regulatory frameworks.

### Benjamin Marks

Benjamin Marks is a campaign strategist. His initiatives have raised hundreds of millions of dollars and reached billions of people worldwide. He's obsessed with tackling the global information crisis and helping individuals and communities build resilience against misinformation, disinformation, and media manipulation. Previously, he was CNN's Global Director of Innovation.

# Table of Contents

<b>Introduction</b>	<b>Page 1</b>
<b>Executive Summary</b>	<b>Page 2</b>
<b>Methodology</b>	<b>Page 4</b>
Define the Scope	Page 4
<b>Findings</b>	<b>Page 6</b>
The Big Picture	Page 6
Comparative Framing: Gaza and Ukraine as Media Priorities	Page 11
Outlet-Specific Disparities	Page 16
<b>Conclusions</b>	<b>Page 21</b>
<b>References</b>	<b>Page 23</b>
<b>Appendix 1: Data Curation and Processing Framework</b>	<b>Page 25</b>
<b>Appendix 2. List of humanitarian crises studied</b>	<b>Page 35</b>

# Introduction

---

In an era where information flows ceaselessly and headlines shape global perceptions, understanding the relationship between media coverage and humanitarian crises has never been more vital. **The Humanitarian Crisis Coverage Report** was born out of a pressing need to uncover how media narratives influence—or sometimes distort—public awareness and policy response, and, implicitly, funding allocation during the world’s most severe humanitarian crises.

The motivation for conducting this research was driven by a critical question: to what extent does media representation align with the actual severity of crises, and how does this alignment—or lack thereof—impact the mobilization of resources and international response?

Despite the proliferation of news, disparities persist in how crises are portrayed, often amplifying access and visibility for some emergencies while marginalizing others. These disparities have far-reaching implications, affecting where humanitarian agencies focus their efforts and how donors prioritize funding.

By analyzing thousands of news articles from diverse outlets, we aimed to shed light on the dynamics between media visibility, perceived urgency, and real-world needs. The goal behind this research is to encourage meaningful dialog and collaboration between the humanitarian and media sectors to ensure that coverage accurately reflects crisis severity and catalyzes effective action.

The **Humanitarian Crisis Coverage Report** serves a dual purpose—bridging academic inquiry with tangible real-world impact.

On the academic front, while this report offers only a distilled snapshot of our broader findings, the full dataset generated through our research is openly available to scholars, practitioners, and journalists seeking to build upon, challenge, or expand this work. We welcome its use in further studies, articles, or additional research.

Equally, this project was conceived as a practical instrument for advocacy—a research-driven resource aimed at reshaping the narratives we construct and the voices we elevate. It is only through collective reckoning with the forces that shape media coverage that we can begin to re-center the stories of those most in need. Ensuring that the world’s most vulnerable are not just acknowledged, but urgently seen, heard, and supported, demands nothing less.

# Executive Summary

The **Humanitarian Crisis Coverage Report** offers a data-driven analysis of how media attention is distributed across 10 major humanitarian emergencies. Through a combination of quantitative and qualitative methods, the study reveals deep imbalances in coverage, critically interrogates how media frames these crises, and assesses how journalistic trends may influence public perception, humanitarian policy, and funding patterns.

The central research question — to what extent media representation aligns with crisis severity — underpins an analysis of 78,667 news articles spanning 10 of the world's most urgent humanitarian crises, from Sudan and Ethiopia to Ukraine and Gaza. Drawing from English-language media in eight countries, the study is based mostly on Western sources and three non-Western outlets to offer a cross-sectional view of editorial priorities in the Anglophone information ecosystem.

Key findings demonstrate that global humanitarian coverage is highly uneven. Crises in Gaza and Ukraine dominate international media attention, averaging 58.5 and 19.4 articles per day, respectively. Conversely, some of the world's most devastating, yet protracted crises—including those in Chad (0.06 articles/day) and the Democratic Republic of Congo (1.5 articles/day)—receive little sustained attention. This disproportionate focus does not correlate with humanitarian severity or affected populations. For example, the DRC, where over 21 million people need aid, garners a fraction of the visibility of Gaza, which affects 3.3 million people.

These disparities are symptomatic of a broader hierarchy in global journalism. Crises with geopolitical resonance—such as those involving Western allies or adversaries—are amplified, while those centered in the Global South are often sidelined. The study identifies longstanding logics of newsworthiness at play, prioritizing conflict proximity, elite involvement, dramatic visuals, and ideological utility over humanitarian need.

Event-driven spikes further distort coverage. Media interest surges during acute developments—such as the Taliban's return in Afghanistan—but quickly subsides, perpetuating a feeling of “media fatigue.” This inconsistency undermines humanitarian advocacy that depends on sustained public and political engagement.

The framing analysis of the Gaza and Ukraine crises conducted as part of this project reveals that the humanitarian lens is the most dominant across media coverage. Articles prioritize imagery of suffering civilians, refugee stories, and basic needs, registering higher average mentions (0.316 per article) than military (0.281) or geopolitical (0.267) frames. However, while framed empathetically, these narratives risk oversimplifying multidimensional conflicts and depoliticizing structural injustices.

Notably, Ukraine coverage exhibits a relatively balanced emotional register: 49.6% of sentiment-coded content is negative, but 27.7% is positive, showcasing themes of solidarity and resistance. In contrast, Gaza coverage is overwhelmingly negative (66.1%), with just 2.5% positive sentiment. This divergence points to geopolitical and editorial influences in how crises are emotionally framed—and consequently understood.

Media ownership and geopolitical alignment also matter. Outlets such as the public service broadcaster BBC, and the privately held The New York Times often aim for perceived balance, while others like Al Jazeera and RT, both state-owned, tend to reflect their respective governments' foreign policy agendas. These editorial differences influence which crises are covered, how they are framed, and whom audiences are encouraged to empathize with.

Ultimately, the study concludes that media power not only informs public understanding but shapes global humanitarian priorities. Crises that disappear from headlines risk disappearing from the policymaking agenda. Rectifying this will require collaborative efforts across journalism, humanitarian organizations, academia, and civil society to ensure fairer, more accountable crisis coverage.

# Methodology

---

Given the interdisciplinary nature of the project, the methodology was designed to meet both computational and analytical standards. It blends automation (for scale) with expert oversight (for rigor), balancing efficiency with interpretive accuracy.

While the methodology involved several automated components—particularly in article retrieval and relevance filtering—it also incorporated expert-driven design choices at critical stages: the selection of media outlets, the construction of crisis-specific keyword sets, and the validation of AI-assisted outputs.

The methodology for this study was designed both for robustness and scalability as well as with a view toward ethical data practices and interdisciplinary applicability. It can serve as a reusable model for similar studies investigating media representations of complex, high-stakes global issues.

## Define the Scope

### Humanitarian crisis selection

A total of 10 humanitarian crises were selected by overall impact, such as the number of people requiring humanitarian aid, as follows: Sudan, Syria, Gaza, Ukraine, Democratic Republic of the Congo (DRC), Afghanistan, Yemen, Ethiopia, Myanmar and Chad. To determine this, we used the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) 2025 Global Humanitarian Overview. (United Nations OCHA, 2025).

### Media outlet selection

This study focuses on English-language media outlets from five key countries—the United States, the United Kingdom, Canada, Australia, and Ireland—chosen for their outsized role in shaping global information flows and their relevance in international reporting on humanitarian crises. These countries were selected based on linguistic accessibility, regional diversity, and their media systems' influence in the Anglophone world.



To complement this core country selection and ensure a broader geopolitical and editorial representation, the study also included three internationally prominent outlets that are not headquartered in these five countries but are frequently cited in global or region-focused crisis coverage and widely indexed in digital platforms: **RT (Russia Today)**, a Russian state-funded outlet known for its alternative framing of Western narratives; **Al Jazeera English**, based in Qatar and widely recognized for its focus on the Global South and underreported regions; **IRNA (Islamic Republic News Agency)**, the official state news agency of Iran, offering a perspective aligned with Iranian state policy.

These additions were made to capture coverage from non-Western and state-aligned perspectives, enriching the study's ideological and geopolitical scope.

Media outlets were selected using the following criteria:

- **Inclusion in Google News indexing**, ensuring visibility in algorithmically curated environments.
- **Editorial and ideological diversity**, covering a wide range of political alignments and ownership structures—from public broadcasters and independent dailies to state-sponsored international media.
- **Technical accessibility**, prioritizing outlets that allow for efficient automated or semi-automated content collection.

During the full-text extraction stage (see *Appendix 1: Data Curation and Processing Framework*), some outlets were excluded from qualitative analysis due to legal or technical constraints, such as strict rules or aggressive anti-bot measures. These exclusions are noted later in the document.

## Data curation

The data curation process for this study combined expert-driven keyword construction, structured metadata collection via the Google News API, and advanced filtering techniques to ensure both breadth and depth in the coverage of 10 humanitarian crises.

Keywords were manually curated through a hybrid approach: a general template (e.g., {Crisis Name} + humanitarian terms) and crisis-specific terms informed by expert input, resulting in 198 unique keyword phrases.

Articles were collected using SerpAPI, ensuring compliance with ethical standards and enabling precise, outlet-specific queries. Preprocessing included deduplication, outlet normalization, and rule-based and AI-based relevance classification using GPT-4o, validated through stratified human sampling with over 93% F1-score.



Two crises (Gaza and Ukraine) were selected for full-text analysis, extracted using Newspaper3k and custom scrapers, with excluded outlets documented. Each article was enriched with humanitarian metadata (e.g., people in need, funding required), facilitating normalization across time and impact.

Quantitative analysis used these metrics to identify over- and under-represented crises, while qualitative analysis applied an AI-assisted framing schema to examine how narratives were constructed. This layered, ethically sound methodology enabled both scale and depth in analyzing global media coverage.

*(For more details, see Appendix 1: Data Curation and Processing Framework).*

## Findings

---

### The Big Picture

#### Unequal Spotlight: Disproportionate Media Attention to Global Crises

The analysis of global humanitarian crisis coverage reveals deep and persistent imbalances in media attention, with some crises commanding the limelight while others remain consigned to the margins. Data gathered from the media outlets analyzed in this project show that journalistic attention is often drawn less by humanitarian severity than by geopolitical salience.

Since October 7, 2023, the day when Hamas launched a large-scale attack on Israel, involving rocket fire and ground incursions from the Gaza Strip, the crisis in Gaza has dominated global headlines, averaging 58.5 articles per day—a figure that far outpaces coverage of other humanitarian emergencies. Coverage spikes coincided with Israeli military offensives, diplomatic flashpoints, and moments of widespread civilian suffering, pointing to an editorial logic driven by spectacle and strategic stakes.

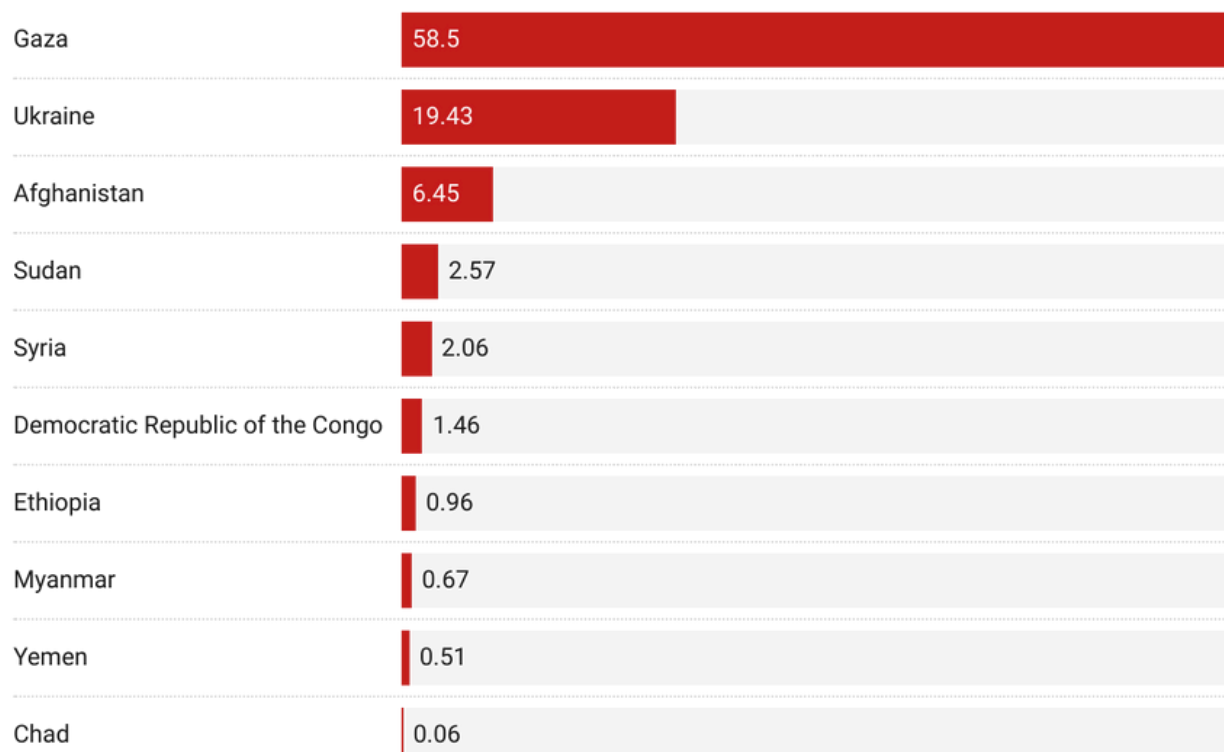
This pattern reflects a well-established tendency in media logic: crises involving international confrontation, Western allies or adversaries, and high political stakes are more likely to be elevated to global importance (Galtung & Ruge, 1965; Wolfsfeld, 2004).

Ukraine, while still receiving significant attention (averaging 19.4 articles per day), has experienced a gradual decline in media coverage since the outbreak of the Gaza war. Although the 2022 full-scale invasion initially sparked a wave of media mobilization—bolstered by Europe’s geographic proximity and the enduring Cold War-era East–West dichotomy—coverage has since plateaued. This trend reflects a classic case of “media fatigue,” where protracted conflicts struggle to maintain visibility in the absence of new, dramatic developments (Robinson, 2018; Scott, 2014). It also suggests the presence of multiple, uneven waves of attention. Nonetheless, a clear downward trend in coverage over time is evident in our analysis.

More troubling, however, is the chronic neglect of crises in the Global South. Despite enormous humanitarian tolls, including mass displacement, food insecurity, and protracted violence, many of these crises have remained virtually invisible in mainstream news coverage. Chad, for example, received just 0.06 articles per day, despite being identified by humanitarian agencies as facing one of the world’s most severe emergencies (United Nations OCHA, 2024). This disparity is not merely an oversight; it signals entrenched structural biases within international journalism, what scholars have described as “compassion fatigue” (Moeller, 1999) and a “hierarchy of human life” (Chouliaraki, 2006), where suffering in distant or non-Western locales is rendered morally remote and journalistically marginal. Structural factors such as the lack of foreign correspondents, commercial pressures, and editorial gatekeeping are believed to further exacerbate this imbalance (Franks, 2013).

## Average daily coverage by crisis

Number of articles per day



Source: MJRC • Created with Datawrapper

Quantitative data demonstrate that the crises attracting the most media attention are not necessarily those affecting the most people or demanding the largest humanitarian responses. The Gaza crisis, for example, generated the highest average media coverage (58.5 articles per day) while affecting approximately 3.3 million people. By contrast, the Democratic Republic of the Congo (DRC), where 21.2 million people are in urgent need, received just 1.5 articles per day.

Similar discrepancies are evident across other conflicts. Ethiopia, Afghanistan, Sudan, and Chad—each facing extreme food insecurity and displacement—received far less coverage than conflicts with greater geopolitical symbolism, such as Gaza and Ukraine. These disparities reaffirm longstanding critiques of “news values” that prioritize drama, elite nation involvement, and visual impact over actual humanitarian need (Harcup & O’Neill, 2001).

### Coverage disparity, comparison of people affected

	People affected (in million)	Articles/day	Coverage per million affected
Gaza	3.3	58.5	17.7
Ukraine	14.6	19.4	1.3
DRC	21.2	1.5	0.07
Ethiopia	20.1	0.97	0.05
Chad	7.7	0.06	0.01

The misalignment becomes even more pronounced when we examine the ratio of coverage per dollar of humanitarian funding required. Gaza emerges as an outlier, with media attention orders of magnitude higher relative to its aid needs.

Such disparities reflect deep-rooted asymmetries in global information flows, where dominant media ecosystems are centered in the Global North (Thussu, 2018). This leads to what Cottle and Nolan (2007) termed “selective empathy,” where suffering becomes newsworthy only when it fits established editorial and ideological frameworks.

These distortions have high-stakes implications. When media visibility drives donor interest and policy agendas, crises that lack coverage are effectively left without a lifeline. The absence of the so-called “CNN effect”[1] in low-visibility crises erodes public solidarity and hampers global humanitarian response.

[1] The “CNN effect” refers to the theory that real-time, global media coverage—especially by 24-hour news networks like CNN—can influence government policy decisions and humanitarian intervention. It suggests that emotionally powerful images and stories broadcast during crises (e.g., famine, war, genocide) can pressure policymakers to respond quickly to public outcry, bypass traditional diplomatic channels by putting issues directly on the public agenda or accelerate humanitarian interventions, even in strategically marginal regions.

Moreover, the skewed attention reinforces hierarchies of visibility and legitimacy. As Chouliaraki (2006) observes, the media often curate a "theatre of suffering," in which some lives are made visible and grievable, while others remain in the shadows. The result is a global narrative that elevates some crises to moral emergencies while allowing others to quietly persist.

## Monthly Trends Reflect Event-Driven Spikes

The analysis of monthly media coverage trends across humanitarian crises reveals a pattern where news attention is highly event-driven, with dramatic spikes in coverage coinciding with moments of geopolitical rupture or heightened violence. These surges are followed by sharp declines, indicating that sustained media engagement is rare and often dependent on sudden developments that align with the values of "newsworthiness."

The project's longitudinal data shows that coverage volume fluctuates significantly month by month, with increases tightly linked to key geopolitical events. For instance, Afghanistan experienced a major spike in August 2021, when the Taliban seized control of Kabul following the withdrawal of U.S. and NATO forces. International media attention surged in response to dramatic images of the evacuation from Hamid Karzai International Airport and the collapse of the Western-backed government. Gaza coverage surged in October 2023 after the Hamas-led attack on Israel and the subsequent large-scale Israeli military response in Gaza. This escalation brought significant attention to the region, dominating front pages and broadcast segments for weeks.

Afghanistan's visibility plummeted after September 2021, while the humanitarian crisis deepened, with food shortages and economic collapse.

At the same time, coverage decisions are shaped by the degree to which a crisis intersects with national interests, ideological polarization, and elite political agendas. The Taliban's rise implicated U.S. foreign policy directly, while the Gaza conflict galvanized domestic audiences across the political spectrum in Western democracies.

The reliance on event-driven surges creates a temporal imbalance in coverage. Moments of visibility are intense but fleeting, often failing to convey the chronic nature of humanitarian needs. This poses challenges for humanitarian advocacy and policy formation, which require sustained public attention and donor engagement.

The monthly trends in crisis coverage captured by our research demonstrate a strong correlation between geopolitical events and media attention, underscoring the event-driven nature of global journalism. While these spikes enhance awareness during critical moments, they also reinforce selective visibility and narrative volatility.

# Monthly coverage of selected crises

Crisis	Year-month	Coverage count
Afghanistan	Aug, 2021	3,306
Afghanistan	Sep, 2021	981
Afghanistan	Oct, 2021	305
Afghanistan	Nov, 2021	212
Afghanistan	Dec, 2021	156
Afghanistan	Jan, 2022	100
Gaza	Oct, 2023	7,299
Gaza	Nov, 2023	3,692
Gaza	Dec, 2023	2,292
Gaza	Jan, 2024	1,049
Gaza	Feb, 2024	682
Gaza	Mar, 2024	638
Sudan	Apr, 2023	498
Sudan	May, 2023	130
Sudan	Jun, 2023	55
Sudan	Jul, 2023	48
Ukraine	Feb, 2022	3,576
Ukraine	Mar, 2022	3,346
Ukraine	Apr, 2022	1,318
Ukraine	May, 2022	799
Ukraine	Jun, 2022	635
Ukraine	Jul, 2022	467
Ukraine	Aug, 2022	443

Source: MJRC

# Comparative Framing: Gaza and Ukraine as Media Priorities

The research also reveals a deep imbalance in media prioritization. Despite the presence of multiple large-scale humanitarian emergencies worldwide, two conflicts—Gaza and Ukraine—have overwhelmingly dominated international media discourse. This disproportionate focus is not merely a function of the events' scale or impact, but rather of how media outlets frame certain crises as globally salient while marginalizing others.

As a result, we chose the two conflicts for our qualitative analysis of their coverage, which is being summarized in this section.

Gaza and Ukraine received dramatically more media attention than other crises, even those with larger affected populations or longer durations. On average, Gaza coverage was nine times greater than Afghanistan, and 60 times greater than Ethiopia. Ukraine coverage was three times higher than Afghanistan, and 20 times higher than Ethiopia. This disparity cannot be explained by humanitarian impact alone. Afghanistan, Ethiopia, and the Democratic Republic of the Congo have all experienced mass displacement, famine, and civilian casualties at comparable or greater levels. Yet these crises are consistently underreported.

These findings echo the core premises of news values theory (Galtung & Ruge, 1965; Harcup & O'Neill, 2001), which explains that journalists and editors are more likely to select stories that involve proximity to Western powers, elite nation involvement, conflict, and visual drama. Both Gaza and Ukraine meet these criteria.

Gaza is at the epicenter of a decades-long geopolitical conflict involving Israel, the U.S., and Arab nations. It is framed through a lens of terrorism, retaliation, and regional instability, offering a morally polarized and emotionally charged narrative.

Ukraine represents a proxy conflict between Russia and the West, framed as a battle between democracy and authoritarianism. The war has implications for NATO, European security, and global energy markets. In contrast, crises like those in Ethiopia or Chad are framed as complex, distant, and often indecipherable, which discourages sustained engagement by mainstream media and their audiences (Franks, 2010; Cottle, 2008).

This is further reinforced by editorial resource allocation. Outlets with limited foreign correspondents prioritize crises where they can ensure audience interest and advertiser value (Hamilton, 2004). This explains why crises in sub-Saharan Africa or Central Asia receive only episodic coverage, even during acute humanitarian disasters.

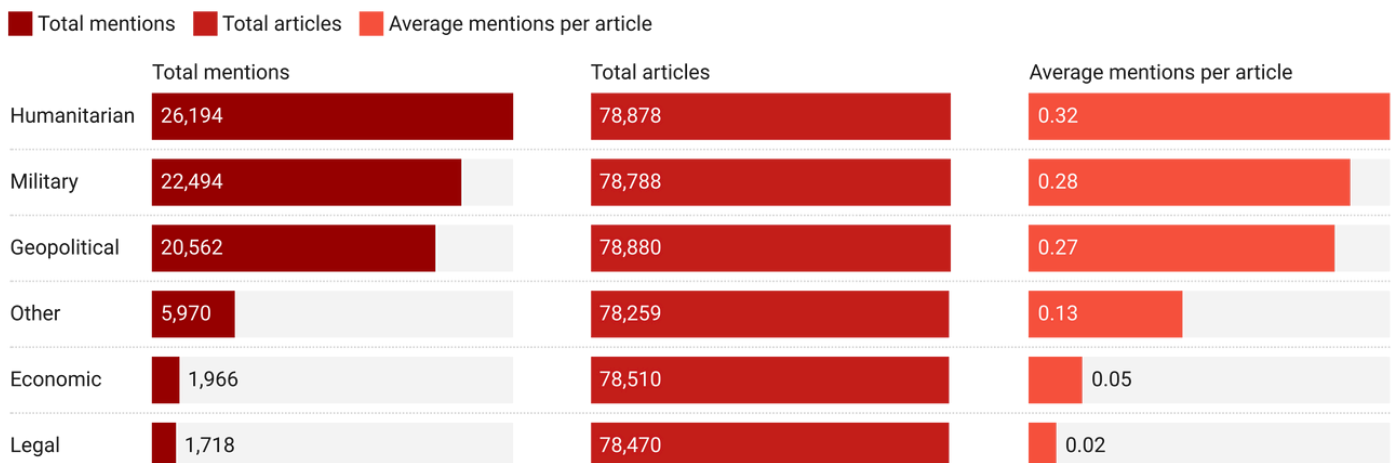
Media prioritization has tangible effects. Studies show that coverage volume strongly influences international donor behavior, political advocacy, and humanitarian response (Moeller, 1999). When Ethiopia's Tigray conflict or Chad's displacement crises are virtually absent from news cycles, they are also likely to be deprioritized in diplomatic, financial, and public policy agendas.

Furthermore, this skewed focus distorts global understanding of humanitarian needs. It fosters a myopic view of crisis geography, wherein only conflicts involving Western allies or adversaries seem to merit attention, while the structural causes of violence and suffering in the Global South are rendered invisible.

## Dominant Frames in Conflict Coverage

An analysis of conflict coverage across all outlets in our dataset reveals a clear predominance of the humanitarian frame, especially in reporting on the two most prominently covered crises: Gaza and Ukraine. Averaged across all articles and outlets, humanitarian framing registers the highest frequency, with 0.316 mentions per article, surpassing both military (0.281) and geopolitical (0.267) frames. This hierarchy suggests a journalistic inclination toward emphasizing human suffering, displacement, and the immediate moral imperatives of crisis response.

### Framing overview: Gaza and Ukraine crises coverage



Source: MJRC • Created with Datawrapper

This finding aligns with a broader media tendency noted in the literature: the rising salience of human-interest reporting in crisis contexts, particularly where visual imagery and emotive narratives can mobilize public empathy (Höjjer, 2004). In the cases of Gaza and Ukraine, coverage often centers on wounded civilians, refugees, or destroyed infrastructure. This lens is effective in generating affective resonance, yet it may also depoliticize or oversimplify the structural roots of conflict.



While humanitarian framing dominates on a per-article basis, military framing remains a significant secondary mode of representation, appearing on average 0.281 times per article. Its persistence reflects the durable legacy of war reporting, in which narratives of combat, weaponry, and territorial dynamics are central. Particularly in the Ukrainian context, where battlefield developments evolve rapidly, military frames provide strategic updates and signal turning points. However, reliance on official or embedded sources can reinforce nationalistic perspectives or narrow understandings of combat (Bennett et al., 2008).

Geopolitical framing, though slightly less frequent per article (0.267), is notable for its cumulative volume—over 20,000 mentions across the dataset—signalling its entrenched role in defining the macro-narrative of each conflict. In both Gaza and Ukraine, media often frame local violence within broader strategic rivalries, such as NATO vs. Russia, or Israel’s security vis-à-vis Iran-backed actors. This reflects what Entman (2004) calls “cascading activation,” where elite political discourse dictates the characteristics of public debate. Yet, geopolitical frames can marginalize civilian experiences by privileging state actors and global power dynamics (Chouliaraki, 2010).

Other framing types—economic and legal, for instance—appear far less frequently, with average mentions per article at 0.047 and 0.134, respectively. This relative absence may indicate editorial prioritization of emotional and strategic narratives over structural or accountability-focused perspectives. It also shows that coverage of conflict economics (e.g., sanctions, trade disruption) or international law (e.g., war crimes, occupation legality) is sporadic or secondary, despite their critical relevance.

Taken together, the framing landscape in coverage of Gaza and Ukraine reveals a three-layer hierarchy: humanitarian urgency, military spectacle, and geopolitical strategy. Each plays a role in shaping public understanding, yet their combined dominance also risks narrowing the lens through which conflict is viewed. By foregrounding immediacy and spectacle, media may inadvertently sideline deeper causal explanations or alternative perspectives.

As some media scholars have long argued, framing is never neutral (Entman, 1993). The humanitarian frame, while normatively appealing, can serve political ends when selectively deployed. In both Gaza and Ukraine, the prevalence of empathy-inducing narratives correlates with broader Western support for affected populations, raising questions about how humanitarianism intersects with ideology and foreign policy agendas.

## Sentiment Landscape: Balanced Tone or Strategic Ambiguity?

A disaggregated analysis of sentiment in media coverage reveals marked contrasts between the Gaza and Ukraine crises. While previous studies have pointed to a broadly balanced sentiment profile across conflict reporting, this focused comparison tells a more complex story.

In coverage of Gaza, sentiment is overwhelmingly negative, accounting for 66.1% of all sentiment-tagged mentions. Neutral sentiment follows at 31.4%, while positive sentiment constitutes only 2.5%. This pattern suggests that Gaza coverage is framed predominantly through the lens of violence, destruction, and humanitarian crisis. Such a skew toward negativity aligns with established news values privileging conflict, casualty, and dramatic imagery (Galtung & Ruge, 1965). However, it also raises questions about the potential reinforcement of despairing or fatalistic narratives, which may contribute to audience desensitization or the “compassion fatigue” already described above.

By contrast, coverage of Ukraine reflects a more diversified emotional register. Negative sentiment is still dominant at 49.6%, but positive sentiment reaches a substantial 27.7%, with neutral sentiment at 22.7%. This more even distribution suggests that Ukraine coverage accommodates not only reports of violence and hardship, but also narratives of resistance, resilience, and diplomatic progress. Stories of Ukrainian defiance, Western military support, or international solidarity likely feed into this more hopeful tonal balance.

The divergence in sentiment patterns across the two crises may stem from both geopolitical positioning and editorial calculus. Western media have generally aligned with Ukraine as a sovereign nation resisting invasion, which allows for constructive or valorizing narratives that support both policy agendas and audience sympathies.

In the case of Gaza, however, the media often operate within a more constrained discursive terrain—balancing complex geopolitical alliances, accusations of bias, and deeply polarized audience expectations. This can result in a narrative space where negative sentiment predominates, while positivity is rare and politically fraught.

In sum, the sentiment landscape of Gaza and Ukraine coverage reveals not only differing editorial tones but also divergent discursive opportunities. Where Ukraine coverage blends tragedy with optimism, Gaza is framed almost exclusively in terms of catastrophe, which may be to a large extent influenced by Hamas’ editorial strategy, focused on depicting the conflict in such terms. These contrasts reflect deeper ideological, geopolitical, and cultural filters through which conflict is made intelligible to global audiences. They also call for critical reflection on how sentiment is distributed, not just as an emotional register, but as a narrative tool that shapes public understanding, empathy, and policy resonance.

## Attribution of Victimhood and Responsibility: Patterns of Blame and Sympathy

A core finding from this analysis is the asymmetric attribution of victimhood and causality in media representations of the Ukraine and Gaza conflicts. Drawing on data across major outlets, the results show a consistent pattern: Ukraine and Gaza/Palestine are predominantly framed as victims, while Russia and Israel are more frequently portrayed as aggressors.

In quantitative terms, Ukraine registers the highest average mentions as a victim across the dataset (0.481 mentions per article), with Gaza/Palestine following closely at 0.388. On the aggressor side, Russia stands out with 0.693 average mentions per article, followed by Israel at 0.270. Meanwhile, Gaza/Palestine and Ukraine receive negligible aggressor attribution (0.036 and 0.025, respectively), underscoring the sharp division in how moral roles are assigned.

### Victim and aggressor attribution by crisis

Crisis	Group	Framing type	Total mentions	Average mentions per article
Gaza	Israel	aggressor	15,256	0.270
Gaza	Israel	victim	19,578	0.388
Gaza	Palestine	aggressor	1,780	0.036
Gaza	Palestine	victim	25,698	0.481
Ukraine	Russia	aggressor	25,874	0.693
Ukraine	Russia	victim	930	0.034
Ukraine	Ukraine	aggressor	806	0.025
Ukraine	Ukraine	victim	23,544	0.637

Source: MJRC • Created with Datawrapper

This framing dichotomy echoes Entman's (1991) theory of differential framing, where foreign policy alignment influences media depictions. In Western press environments, Ukraine's victimhood fits squarely within NATO and EU narratives that emphasize Russian aggression. Similarly, the depiction of Palestinian suffering—particularly under siege or bombardment—is often contrasted with portrayals of Israeli military dominance, reinforcing a victim-aggressor binary rooted in perceptions of power asymmetry.

Yet these narratives are not simply editorial constructs; they are reinforced by the discursive ecosystems of international advocacy, diplomacy, and humanitarian work. NGOs and global institutions routinely produce content that frames Palestinians and Ukrainians as “deserving victims,” which the media then circulate.

At the same time, such dichotomous framing can flatten complexity. The positioning of Israel as the aggressor, for example, minimizes the role of Hamas and its financiers in causing the conflict. Similarly, the focus on Israeli state actions may underrepresent internal dissent within Israel or the contested nature of Palestinian political authority.

These dynamics reflect episodic framing (Iyengar, 1991), where stories are presented through immediate, emotional snapshots rather than through broader structural or historical lenses. In this context, victim framing elicits empathy and solidarity, while aggressor framing mobilizes blame and calls for accountability—both central to journalistic storytelling, yet both susceptible to ideological bias.

The absence of shared or overlapping victim-aggressor roles reflects a binary moral architecture in media discourse. These narratives rarely accommodate ambiguity or shared responsibility. This may benefit clarity and advocacy, but it also simplifies conflict realities and may obscure deeper structural causes and political nuances.

## Outlet-Specific Disparities

### Country-Based Media Patterns in Humanitarian Crisis Coverage

The project’s dataset—drawn almost exclusively from English-language media in the Global North—is not sufficient to provide a comprehensive global media analysis. Yet, it reveals important structural dynamics within Anglophone international news production.

U.S. and UK outlets often shape the tone and framing of crisis narratives consumed by English-speaking audiences worldwide. Canadian, Australian and Irish outlets were represented in the data, but to a much lesser extent. Their coverage tended to mirror patterns established by larger UK and U.S. counterparts, often relying on the same international wires or reproducing similar geopolitical angles.

Our dataset suggests that editorial attention correlates strongly with political and cultural proximity. U.S. outlets devoted considerable space to Afghanistan and Ukraine—conflicts with clear strategic relevance to American foreign policy—while offering less consistent coverage of crises in places like Chad or the Democratic Republic of Congo.

The same applies to UK media, whose reporting on Gaza, Sudan, and Syria often reflected government positions and domestic political debates. Across all countries in the dataset, crisis salience appeared to be less a function of humanitarian severity than of perceived national relevance and audience familiarity.

Another factor influencing these patterns is media infrastructure. Large outlets like BBC, CNN, The New York Times, and The Guardian maintain international news desks and deploy correspondents to conflict zones. Their capacity to generate original reporting—not just syndicate wire content—gives them narrative control over how crises are framed.

Smaller outlets in the dataset, particularly regional or digital-native platforms from Canada, Ireland, and Australia, were more likely to follow the lead of dominant newsrooms. This replication reinforces editorial hierarchies within the English-language media sphere, where a relatively small number of institutions effectively shape the global crisis narrative.

While our dataset was heavily weighted toward Anglophone media from the Global North, a handful of non-Western or geopolitically distinct outlets—such as Russia Today (RT), Iran’s IRNA, and Al Jazeera—offered valuable counterpoints to prevailing Western narratives.

RT, a Russian state-backed international broadcaster, emerged as a notable outlier due to both its volume and editorial framing. While RT’s crisis coverage often mirrored the topical focus of major U.S. and UK outlets (e.g., Ukraine, Gaza), its framing sharply diverged, frequently emphasizing Western double standards, NATO aggression, or humanitarian impacts of Western-aligned military action. The inclusion of RT provides a window into how Russian state propaganda attempts to frame the conflict versus Western narratives.

Islamic Republic News Agency (IRNA), the official news agency of Iran, appeared much less frequently in the dataset, but its inclusion is nonetheless significant. Its crisis reporting is largely aligned with Iran’s regional foreign policy interests, highlighting, for instance, the humanitarian crisis in Gaza through a lens critical of Israel and U.S. policy. While its overall influence in the dataset was marginal, IRNA’s presence serves as a reminder that state-aligned news actors outside the Western sphere do participate—albeit selectively—in the global information environment.

Al Jazeera represents a unique case. Though headquartered in Qatar, it operates with a global editorial model and significant reach in English-language media. Unlike RT or IRNA, Al Jazeera is more broadly recognized as a professional international news outlet, albeit one that also reflects regional perspectives often missing from Western reporting. Al Jazeera played a major role in Gaza coverage, offering extensive reporting and commentary that foregrounded Palestinian civilian experiences and regional political dynamics, which may be entirely expected as the broadcaster is owned by the state, hence it is known as supporting Qatar’s geopolitical position, in the same way at RT aligns with the Russian government propaganda.

Al Jazeera's editorial focus often centered on Israeli security concerns or diplomatic frameworks. Its presence in the dataset thus added depth and regional nuance, particularly in Middle East-related crises.

These outliers injected complexity into the narrative homogeneity of Anglophone Global North media. Their inclusion exposed the presence of alternative geopolitical frames in an otherwise Western-leaning dataset, highlighted the selective visibility of certain crises based on national or regional foreign policy priorities, and emphasized the need for critical engagement with source agendas, particularly when analyzing crisis framing and international solidarity discourses.

While RT, IRNA, and Al Jazeera represent only a small fraction of the total dataset, they underscore the fact that media power is not monolithic, even within a heavily filtered sample. Their divergent framings offer opportunities for comparative analysis and reveal the contested nature of humanitarian storytelling on the global stage. In future research, a more comprehensive inclusion of such actors—across linguistic, political, and regional divides—would help build a genuinely multipolar understanding of crisis coverage.

## Framing as a Function of Editorial Identity

The data in our project shows that certain outlets consistently produced more coverage across multiple crises, while others exhibited selective engagement, often shaped by geographic, political, or ideological proximity. For instance, BBC and Al Jazeera both maintained high volumes of coverage on crises like Afghanistan and Gaza, but likely differed in tone and narrative framing. The BBC traditionally adheres to a “balanced” and official source-aligned journalistic approach (Wahl-Jorgensen et al., 2016) while Qatar government-owned Al Jazeera—especially its Arabic service—has been noted for amplifying voices from the Global South and focusing on human rights narratives (Figenschou, 2014).

Even without explicit tone categories, this divergence in editorial stance allows us to infer that identical crises may be portrayed with different emotional valence, agency attribution, and moral framing, depending on the outlet.

Dispositional variation across outlets has real-world effects. Research has shown that framing influences whether a crisis is perceived as solvable or inevitable, political or humanitarian, urgent or peripheral (Cottle & Nolan, 2007). For example, crisis reporting that emphasizes government failure or international injustice may spur political advocacy, whereas a focus on suffering without context can depoliticize and dehistoricize the issue. Moreover, when coverage tone varies by outlet, so too does its resonance with different audiences. For example, CNN's framing of Ukraine as a fight for democracy against authoritarianism mobilized U.S. public support, while RT offered a counter-narrative of NATO aggression.



Outlets' reputations and ownership structures appear closely linked to the way they frame crises. BBC, a publicly funded but editorially independent broadcaster, tends to offer more neutral and humanitarian-driven coverage. Its consistent attention to Ukraine correlates with the UK government's geopolitical interests, but is balanced by strong editorial oversight structures. Al Jazeera, funded by the state of Qatar, demonstrates a clear editorial line favoring Palestinian narratives, visible in the keyword dominance of "refugees," "siege," and "occupation" in Gaza reporting. Fox News, although not among the highest-volume contributors, displays a more security and geopolitically driven frame when it does cover international crises, particularly Israel-Palestine, consistent with its right-leaning ownership and audience orientation.

Public perception and trust dynamics also play a key role in the media's coverage, as identified by this research. BBC enjoys high trust in many countries but has seen erosion domestically due to perceived political interference and culture war issues. Its still-strong international reputation lends authority to its crisis framing. Al Jazeera is perceived as credible in much of the Global South and Arab world but less so in Western contexts due to perceptions of state influence. CNN and The New York Times, both widely cited in global crisis coverage, face polarized perceptions, often seen as liberal leaning in the U.S. but highly reputable internationally.

Trust ratings often correlate with frame acceptance: outlets with higher public trust see wider resonance of their narrative frames. The coverage patterns and keyword choices in our research datasets (e.g., "humanitarian crisis" in Ukraine, "siege" in Gaza) demonstrate how framing is both shaped by and contributes to public sentiment, particularly when propagated by high-trust outlets.

## Overview of key correlations between public perception, ownership and framing

Outlet	Ownership type	Frame dominance	Public perception	Observed correlation in data
BBC	Public (UK license fee)	Humanitarian, balanced	High international, mixed UK	High volume, diverse frame usage
Al Jazeera	State-funded (Qatar)	Pro-Palestinian, anti-occupation	Trusted in Global South, mixed West	Gaza-focused, keywords around civilians
AP	Cooperative, non-profit	Factual, wire-style, event-driven	High trust globally	Constant crisis output, tone-neutral
Fox News	Private (News Corp)	Geopolitical, state security	Trusted by right-wing audiences	Less frequent, more aggressor-focused
CNN	Private (Warner Bros.)	Liberal-leaning, emotion-heavy	Politically polarized	High Ukraine presence, narrative-driven

Source: MJRC • Created with Datawrapper



Overall, framing data from the geopolitics, humanitarian, and political accountability datasets shows some thematic consistencies. Publicly- or state-funded broadcasters like BBC and Al Jazeera show greater use of humanitarian and victim-oriented framing (e.g., keywords like “refugees,” “ceasefire,” “humanitarian crisis”), though with nuanced differences reflecting their respective state’s foreign policy interests. Commercial and U.S.-based outlets (e.g., CNN, Fox News) lean more toward geopolitics and causality, particularly in coverage of Russia and China. These outlets frequently highlight themes of strategic conflict and Western-aligned accountability framing. On the other hand, public service outlets also show broader thematic diversity but often lack the volume and virality of commercial players.

# Conclusions

The **Humanitarian Crisis Coverage Report** sheds light on the powerful, often unseen mechanisms that govern how the world learns about suffering. It reveals that the global media landscape does not report on humanitarian crises in line with their severity or scale. Instead, editorial choices tend to follow the geopolitical interests, visual drama, and alignment with prevailing ideologies in dominant media ecosystems. As a result, while some crises are thrust into the spotlight, commanding headlines and stirring public outcry, others—often those unfolding in regions of the Global South—are left in the shadows, regardless of the number of lives affected.

What takes shape is a media hierarchy, in which suffering is not only unequally covered but unequally valued. Some tragedies are made visible, and therefore urgent, while others are allowed to fade into background noise, if covered at all. This selective visibility carries weighty consequences. Public awareness, donor generosity, and political responses are all, to a significant extent, shaped by the stories the media chooses to tell. Crises that fail to capture journalistic attention risk being overlooked not only by the public but also by the very policymakers and humanitarians tasked with responding and allocating funds. These overlooked emergencies thus suffer a double blow: first, from the crisis itself, and second, from being effectively erased from the global agenda.

Beyond uncovering these troubling asymmetries, the **Humanitarian Crisis Coverage Report** also opens up a number of critical avenues for future research—each of which could deepen our understanding of the connections between media representation and real-world impact.

One such avenue is the exploration of geopolitical bias and foreign policy influence: to what degree do media narratives mirror national strategic interests, and where—if at all—do they push back against them? A related strand concerns algorithmic gatekeeping. In a news era increasingly dictated by automated feeds and search engine results, it is essential to investigate how algorithms shape access to information about crises. Which stories are amplified, and which are filtered out before they can reach the public eye?

Another important direction involves analyzing how framing choices land with audiences. While this study analyzed how stories are told, the next step is understanding how they are received: how do different cultural or political audiences interpret stories framed around victimhood, military action, or humanitarian appeal?

Equally important is the expansion of comparative work beyond Anglophone or Global North-heavy media samples. This study made a start by including outlets like Al Jazeera, RT, and IRNA, but further engagement with non-Western, multilingual, and regional media ecosystems is vital. Doing so will shed light on whether patterns of neglect and bias are universal or culturally specific.

Lastly, there's a pressing question of impact: precisely how and in what ways does a lack of media visibility contribute to disparities in humanitarian response, from the volume of aid allocated to the speed and scope of international interventions?

These threads of inquiry would bring welcome complexity and granularity to the field of humanitarian communication research. By taking this study further, scholars, journalists, and aid actors alike could help reimagine a more equitable and principled approach to crisis coverage—an approach that sees beyond geopolitical blinders and that is aimed to tell the stories that matter most, whether they're trending or not.

# References

- Bennett, W. L., Lawrence, R. G., & Livingston, S. (2008). *When the press fails: Political power and the news media from Iraq to Katrina*. University of Chicago Press.
- Chouliaraki, L. (2010). Post-humanitarianism: Humanitarian communication beyond a politics of pity. *International Journal of Cultural Studies*, 13(2), 107–126.
- Chouliaraki, L. (2006). *The spectatorship of suffering*. Sage.
- Cottle, S. (2008). *Global crisis reporting*. McGraw-Hill Education (UK).
- Cottle, S., & Nolan, D. (2007). Global Humanitarianism and the Changing Aid-Media Field. “Everyone was dying for footage.” *Journalism Studies*, 8(6), 862–878. <https://doi.org/10.1080/14616700701556104>.
- Entman, R. M. (2004). *Projections of power: Framing news, public opinion, and U.S. foreign policy*. University of Chicago Press.
- Entman, R. M. (1993). Framing: Towards clarification of a fractured paradigm. *McQuail's reader in mass communication theory*, 390, 397.
- Entman, R.M. (1991). Framing U.S. Coverage of International News: Contrasts in Narratives of the KAL and Iran Air Incidents. *Journal of Communication*, 41(4), 6–27, <https://doi.org/10.1111/j.1460-2466.1991.tb02328.x>
- Figenschou, T. U. (2014). *Al Jazeera and the global media landscape: The South is talking back*. Routledge.
- Franks, S. (2013). *Reporting disasters: Famine, aid, politics and the media*. Hurst.
- Franks, S. (2010). *The neglect of Africa: How Western media covers famine and war*. Hurst & Co.
- Galtung, J., & Ruge, M. H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of peace research*, 2(1), 64-90.

Hamilton, J. T. (2004). *All the news that's fit to sell: How the market transforms information into news*. Princeton University Press.

Harcup, T., & O'Neill, D. (2001). What Is News? Galtung and Ruge revisited. *Journalism Studies*, 2(2), 261–280. <https://doi.org/10.1080/14616700118449>

Höijer, B. (2004). The discourse of global compassion: The audience and media reporting of human suffering. *Media, Culture & Society*, 26(4), 513–531.

Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. University of Chicago Press.

Moeller, S. D. (1999). *Compassion fatigue: How the media sell disease, famine, war and death*. Routledge.

Robinson, P. (2018). *Media, propaganda and the politics of intervention*. Routledge.

Robinson, P. (2002). *The CNN effect: The myth of news, foreign policy and intervention*. Routledge.

Scott, M. (2014). *Media and development*. Bloomsbury Publishing.

Thussu, D. K. (2018). *International communication: Continuity and change*. Bloomsbury Publishing.

United Nations Office for the Coordination of Humanitarian Affairs. (2025). *Global Humanitarian Overview 2025 [Report]*. Retrieved July 22, 2025, from <https://www.unocha.org/publications/report/world/global-humanitarian-overview-2025-enarfres>

United Nations Office for the Coordination of Humanitarian Affairs. (2023). *Global humanitarian overview 2024 [Report]*. United Nations. <https://www.unocha.org/publications/report/world/global-humanitarian-overview-2024-enarfres>

Wahl-Jorgensen, K., Berry, M., Garcia-Blanco, I., Bennett, L., & Cable, J. (2016). Rethinking balance and impartiality in journalism? How the BBC attempted and failed to change the paradigm. *Journalism*, 18(7), 781-800. <https://doi.org/10.1177/1464884916648094> (Original work published 2017)

Wolfsfeld, G. (2004). *Media and the path to peace*. Cambridge University Press.

# Appendix 1: Data Curation and Processing Framework

## 1. Keyword Framework Construction

The retrieval of articles was guided by a structured and scalable keyword framework that balanced **generalizability** with **crisis-specific sensitivity**. The keyword strategy followed a two-pronged approach:

- a). **General Template**: for every crisis, a set of general terms was generated using a consistent template: {Crisis Name} + [general humanitarian terms]
- b). **Crisis-Specific Keywords**: in addition to the general template, each crisis was associated with a curated list of specific terms that reflected the unique features, actors, and discourse surrounding that situation. These keyword sets were developed through expert consultation and informed by sources such as. This hybrid keyword approach aimed to ensure both breadth (capturing generic mentions) and depth (capturing nuanced and locally relevant discourse). The result was a total of **198 unique keyword phrases** across the ten selected crises. The keyword selection was carried out entirely manually, based on contextual understanding of each crisis without the use of automated lexicographic tools.

## 2. Article collection and enrichment

The primary mechanism for article collection in this project was the **Google News API**, selected for its broad index coverage, structured metadata, and compliance with ethical and legal data access practices. The rationale for this choice was both **practical** and **methodological**.

Google News is one of the most comprehensive aggregators of news content globally. It indexes articles from thousands of media outlets across diverse regions, languages, and political spectrums. Compared to alternatives such as Media Cloud, GDELT, or custom scraping pipelines, Google News offers key advantages:

- **High Index Coverage**: Aggregates content from both major and minor outlets, improving representativeness;
- **Reliable Metadata**: Each article comes with clean fields including title, snippet, publication date, and outlet source—crucial for downstream processing.
- **Outlet-Scoped Querying**: The API supports site: parameters, allowing precise targeting of specific outlets, which aligned well with our expert-defined media source list.
- **Language and Regional Filtering**: Though primarily English-focused for this project, Google News supports multilingual retrieval.
- **Search Algorithm Maturity**: Google's backend ensures deduplication and relevance prioritization, which complements our own filtering logic.

All article data was retrieved using SerpAPI, a third-party service that interfaces with Google News, ensuring compliance with Google's terms of service and minimizing ethical concerns typically associated with direct web scraping. No direct crawling of publisher websites occurred during this phase, and article previews were restricted to metadata already publicly available via Google's platform.

This design decision reflects a deliberate commitment to responsible research practices:

- **No bypassing of paywalls or unauthorized access**
- **All data collected was publicly indexed and accessible via query.**

By using a well-maintained and legally permissible API, we also reduced maintenance overhead, API instability, and the risk of data contamination caused by inconsistent site structures.

The query strategy combined expert-defined **outlets** and **crisis-specific keywords**, producing search strings of the format: {site}: outlet\_website keyword {after}: crisis\_start\_date

Each (crisis, keyword, outlet) combination was queried, resulting in a highly structured retrieval matrix. For each result, the following metadata was collected:

**Title; Snippet (preview text); Article URL; Source (media outlet); Publication date**

This structured metadata allowed for precise matching, deduplication, relevance evaluation, and temporal aggregation—all necessary for rigorous quantitative analysis.

### 3. Full Article Retrieval (Selected Crises)

To support the qualitative analysis pillar, full-text retrieval was performed for the **two most-covered crises** in the dataset (based on article count). The selection was made after preprocessing and relevance filtering (see Processing: Cleaning, Filtering and Classification), ensuring focus on the most data-rich contexts.

Two primary tools were used:

- **Newspaper3k**: An open-source article extractor optimized for clean content extraction (title, body, publication date).
- **Custom Fallback Scrapers**: Deployed in cases where Newspaper3k failed, typically due to page structure irregularities or dynamic content.

Articles were matched using the previously collected metadata (URL) and processed for clean text extraction. Some outlets explicitly blocked scraping through anti-crawling policies. These were excluded from the qualitative phase to maintain ethical and legal compliance. The details of such failures are provided in Table 1.



## Appendix 1. Table 1: Article exclusion overview

Outlet name	Number of articles failed to retrieve	Percentage of failed to retrieve
Reuters	2,731	99
Newsweek	2,532	100
MSN	2,320	100
Wall Street Journal	1,084	99
News.com.au	495	70
Washington Post	342	19

Source: MJRC • Created with Datawrapper

### 4. Crisis Metadata Enrichment

To contextualize media coverage relative to the real-world scale of each crisis, we enriched each article record with structured humanitarian metadata. These statistics were sourced from the United Nations Office for the Coordination of Humanitarian Affairs and other reputable sources, which were also cited in the main research design.

For each crisis, the following attributes were added: **Crisis start date**, **Number of people in need**, **Required humanitarian funding (USD)**, **Duration of crisis (days)**, computed from start date to the last article date.

These variables were later used to normalize article volume and assess alignment between **media attention** and **humanitarian severity**.

The integration of this external data allows for analytical comparisons that account for real-world impact, rather than treating all crises as equal media units. This is essential for identifying under- or over-represented crises in subsequent analysis.

### 5. Processing: Cleaning, Filtering and Classification

Once collected via the Google News API, the articles' data underwent a multi-stage processing pipeline to ensure consistency, remove noise, and retain only the records relevant to the humanitarian crises under investigation.

## Preprocessing

The initial cleaning stage was designed to sanitize the raw article set, remove redundancies, and standardize fields for downstream use. While the Google News API delivers relatively structured data, inconsistencies and duplication still arise due to overlapping keyword matches or syndicated content.

### Key Preprocessing Steps:

#### Duplicate Removal

- Articles were deduplicated based on a combination of title, URL, and snippet similarity.
- This step helped address duplicates caused by multi-keyword hits or slight title variants.

#### Malformed Record Filtering

Entries missing either title or snippet were discarded.

URLs were validated for proper structure (e.g., HTTP/HTTPS and domain parsing).

#### Outlet Name Normalization

- Some outlets appeared under varying aliases (e.g., “NYT”, “The New York Times”).
- A standard naming schema was applied to unify source tags.

## Relevance Classification

Given that keyword-based retrieval yields many false positives—e.g., generic geopolitical news or indirect references—a robust filtering method was required to isolate articles that truly engaged with the humanitarian aspects of each crisis.

To achieve this, we implemented a **two-phase relevance classification pipeline, followed by a human validation step**:

#### Phase 1 – Keyword Match (Rule-Based Auto-Relevance)

Articles that explicitly mentioned the **crisis name and search keywords** in either the title or snippet were automatically labeled as **relevant**. This rule-based pass was designed to maximize recall early on.

Example: A title like “Sudan conflict displaces thousands” would qualify immediately.

This method captured clear-cut cases with minimal computational cost.

## Phase 2 –AI-Based Contextual Relevance Filtering

Articles that did not meet the initial keyword-matching criteria were subjected to a second relevance check using OpenAI’s **GPT-4o** language model. This model was deployed in a **few-shot prompting** setup, where each prompt included a series of labeled examples to help guide the model’s understanding of relevance beyond surface-level term matching. Each input consisted of the article’s **title and snippet**, reflecting the constraints of the metadata collected via the Google News API.

The examples included in the prompts were deliberately varied in their degree of ambiguity. Some were clearly relevant or irrelevant, while others represented borderline or nuanced cases. Additionally, the prompts were customized with **crisis-specific instructions**, which provided context about the humanitarian dimensions associated with each conflict (e.g., displacement, aid access, violations). This tailoring enabled the model to evaluate whether an article’s focus aligned meaningfully with the crisis in question, even in cases where key terms were implied or indirectly phrased.

The model produced a **binary classification**: “Relevant” or “Irrelevant.” Its role was to serve as a semantic filter, complementing the rigid keyword logic of Phase 1 by identifying content that referred to the crisis in substantive ways, despite potential variations in terminology or framing.

## Evaluation – Human Validation via Stratified Sampling

To ensure the reliability of AI-generated relevance labels, a **manual validation phase** was conducted using a **stratified sampling approach**. The dataset was divided into strata along three primary dimensions: crisis name, media outlet, and **publication period**. This structure ensured that the sample captured variability in source, geography, and temporal context.

From each stratum, a balanced number of articles were randomly selected for review. Human experts independently assessed whether the articles had been correctly labeled by the model, with attention given to both **true positives** (correctly identified relevant articles) and **true negatives** (correctly excluded irrelevant instances). Particular scrutiny was applied to ambiguous or borderline cases, such as articles referencing multiple crises or using metaphorical language.

The validation revealed a high degree of alignment between the model’s output and expert judgment, with an overall agreement rate of 93%, a precision of 96%, a recall of 90%, and an F1-score of 93%. Discrepancies tended to occur in marginal cases, such as brief mentions of a crisis in unrelated contexts or headlines that referred to humanitarian issues without direct reference to the specific crisis being evaluated. These findings supported the robustness of the AI-assisted filtering approach while also informing minor refinements to the structure and classification thresholds.

To provide a clear overview of how the dataset evolved through each phase of the processing pipeline, see the article counts at key stages in Table 2. This includes the initial volume retrieved from the Google News API, the number of records retained after cleaning and filtering, and the final distribution of articles across the two analytical pillars. The reduction in volume reflects successive refinements in data quality, relevance, and availability of full-text content.

## Appendix 1. Table 2: Dataset Volume by processing stage

Stage	Description	Article count
Raw Collection	All articles retrieved before any processing	240,134
After Preprocessing	Cleaned for duplicates, malformed entries, etc.	233,180
Assigned to Pillar 1	Metadata-based analysis (title/snippet level)	78,667
Assigned to Pillar 1	Full-text + framing schema applied	39,564

Source: MJRC • Created with Datawrapper

## 6. Quantitative Structuring & Normalization

With a filtered and validated dataset of relevant articles, the next step in the curation pipeline was to structure the data in a way that allowed for **fair and comparative quantitative analysis** across the ten humanitarian crises. Because each crisis varied significantly in both **duration** and **humanitarian severity**, raw article counts would not have yielded meaningful insights. To address this, we applied three different normalization strategies that adjusted coverage metrics based on time, population impact, and financial need.

Comparing the number of articles per crisis without normalization risks skewing the interpretation toward crises that are either recent, politically salient, or more frequently covered by the media. To counteract this, we incorporated external crisis metadata into our analysis, enabling us to calculate **coverage intensity relative to need**.

This approach aligns with the core research question: *Does the amount of media attention a crisis receives correspond to its humanitarian impact?* Normalization helps answer this by controlling for factors that inherently influence visibility (e.g., how long a crisis has been ongoing), and by contextualizing media salience against measurable indicators of human suffering and funding requirements.

The core normalization metric was based on normalizing the total count of articles per crisis by crisis days, the number of people affected (in scale of millions) and amount of funds required (in scale of billion USD).

The normalized coverage scores were used throughout the quantitative pillar of the project to:

- Highlight **under-reported crises**, where humanitarian need outweighed media attention
- Identify **over-represented crises**, which received disproportionate coverage relative to impact
- Support **cross-crisis comparisons** free from temporal or population-size distortion
- Inform deeper analyses correlating **media salience with funding outcomes** or geopolitical events

This step laid the foundation for identifying systemic gaps in public attention and potential agenda-setting biases within international media coverage.

## 7. Qualitative Analysis Process

In addition to quantitative measures of media salience, the research also examined the **qualitative framing** of news coverage for two of the most extensively covered humanitarian crises: **Gaza** and **Ukraine**. For this purpose, full-text retrieval (see *Full Article Retrieval (Selected Crises)*) was performed for the two most-covered crises in the dataset.

The goal was to explore not only *how much* coverage each crisis received but also *how that coverage was shaped*, specifically in terms of the narratives, moral framing, political positioning, and emotional cues embedded in media discourse.

### Framing Schema Definition

The qualitative framework was developed by subject-matter experts and drew upon established theories in media studies—most notably **agenda-setting theory** and **framing theory**. Four high-level frames were defined to guide the interpretation of news articles:

- **Humanitarian Crisis**
- **Political Accountability**
- **Geopolitics**
- **Historical Legacy & Future Projections**

Each frame was broken down into multiple attributes, each tied to specific research questions. For example, within the “Humanitarian Crisis” frame, analysts looked for how victims were characterized, how causes were discussed, and how moral language was invoked. The schema was designed to be both interpretable by human reviewers and implementable by a language model in an automated setting.

## Trial Setup Using O3 Model

For the pilot run, a **sample of full-text articles** was selected from the datasets of Gaza and Ukraine. These texts had been successfully extracted using the methods described in *Full Article Retrieval*.

We used OpenAI's **O3** reasoning-aware AI model to analyze each article in relation to the defined framing schema. Each article was processed with a tailored prompt that:

- Provided the framing definitions and example outputs
- Asked the model to classify each attribute under the relevant frame
- Encouraged structured and explainable outputs (e.g., short descriptions or yes/no tags)

Unlike prior phases that worked only with headlines or snippets, this analysis leveraged the **entire body text** of the article. This allowed the model to interpret narrative structure, tone, and embedded viewpoints more holistically.

## Expert Validation and Alignment Assessment

To ensure that the model's frame interpretations aligned with human judgment, a subset of the model-labeled articles was manually reviewed by the qualitative research team. This process mirrored the stratified sampling structure used in the relevance validation phase (see *Processing: Cleaning, Filtering and Classification*), with samples drawn from both crises and multiple outlets.

Each attribute within each frame was assessed for:

- **Correctness** of the classification
- **Completeness** of the model's interpretation
- **Justifiability** of the provided rationale or examples

The results showed a **>96% agreement** rate between the AI-generated frame assessments and those produced by expert reviewers. In most cases, disagreements stemmed from subtle interpretive nuances rather than outright errors. This high alignment supported the viability of using language models as a first-pass framing tool, especially when working with large volumes of data where full manual coding would be infeasible.



## Key Learnings and Final Implementation

The trial phase provided an opportunity to refine the prompt design, framing alignment, and output formatting based on expert feedback. Once the approach was validated, with **over 96% alignment** between model-generated outputs and human-coded evaluations, it was no longer treated as a mere proof of concept. Instead, it became the foundation for the **full-scale implementation** of the qualitative analysis pillar across the entire corpus of relevant full-text articles for Gaza and Ukraine.

Informed by the expert revisions during the trial, the final model prompts were adapted to ensure greater consistency and alignment with the framing schema. The analysis was then **applied systematically** to the complete dataset designated for qualitative study, enabling the extraction of large-scale, thematically coded insights that directly contributed to the research findings on narrative framing and comparative media treatment of the two crises.

This integration of human-in-the-loop AI enabled the project to **scale qualitative analysis without sacrificing interpretive integrity**, making it possible to draw structured comparisons across outlets, time periods, and framing categories in a way that would have been prohibitively labor-intensive through manual analysis alone.

## 7. Methodological Evaluation: Strengths and Limitations

### Strengths

One of the most significant strengths of this methodology lies in its combination of **expert-driven design** and **scalable automation**. From the outset, the project grounded its data sourcing in curated lists of media outlets and keyword sets provided by experts. This ensured that the data pipeline was aligned with real-world media dynamics and crisis-specific discourse patterns.

The use of the **Google News API** for article collection provided a structured, legally compliant, and globally representative index of news content. Unlike ad hoc web scraping or social media mining, this approach ensured that data was drawn from high-authority sources while remaining within ethical and legal boundaries.

A major methodological contribution was the **two-phase relevance filtering system**, which paired simple rule-based heuristics with contextual evaluation via GPT-4o. This architecture provided a strong balance between precision and recall, and its reliability was confirmed through **stratified human validation**, which showed over **90% agreement** with expert assessments.



The framework also distinguished itself through its normalization logic, which accounted for the varying intensity and duration of each crisis. By adjusting coverage metrics based on people affected, required funding, and crisis duration, the analysis allowed for meaningful cross-crisis comparisons that did not simply reflect media visibility alone.

Finally, the project demonstrated a scalable model for qualitative analysis using AI, validated through expert comparison. The integration of the O3 model into the framing schema process allowed researchers to move beyond volume metrics and into the domain of meaning, tone, and narrative structure, while retaining interpretive integrity.

## Limitations

Despite its methodological rigor, the pipeline was not without limitations. First, the reliance on **titles and snippets** for relevance filtering (in cases where full-text was not available) constrained the model's ability to assess nuance or indirect references to crises. While GPT-4o performed well in many cases, brief or ambiguous summaries still posed a challenge, particularly when the article's full context was critical to understanding its relevance.

Second, **access to full article content** was only possible for a subset of outlets and crises. Technical and legal constraints meant that some articles were excluded from the qualitative phase, potentially introducing bias if those outlets had distinct framing styles or regional emphasis.

The use of AI models also raises concerns about **interpretive bias**. While prompts were tailored to each crisis and underwent extensive testing, GPT models reflect patterns present in their training data, which may or may not carry subtle ideological or cultural leanings. Although mitigated by human validation, this remains an area requiring caution.

## Appendix 2. List of humanitarian crises studied

### Sudan:

- People needing aid: 30.4 million
- Funding required: \$10.28 billion
- Overview: Ongoing violence following a coup and subsequent conflicts have led to widespread humanitarian needs, with many people displaced and in desperate need of assistance.
- **Start date: April 2023 (Source)**

### Syria:

- People needing aid: 16.7 million
- Funding required: \$8.58 billion
- Overview: Over a decade of conflict has resulted in the world's largest refugee crisis, with a vast majority of the population living in poverty and extensive infrastructure destruction
- **Start date: March 2011 (Source)**

### Gaza:

- People needing aid: 3.3 million
- Funding required: \$4 billion
- Overview: A prolonged state of violence and displacement has left the majority of Gazans in need of humanitarian assistance, exacerbated by a dire humanitarian situation.
- **Start date: October 2023 (Source)**

### Ukraine:

- People needing aid: 12.7 million
- Funding required: \$3.32 billion
- Overview: Since the escalation of violence in 2022, millions face ongoing conflict-related challenges, including significant energy and resource shortages.
- **Start date: February 2022 (Source)**

## Democratic Republic of the Congo (DRC):

- People needing aid: 21.2 million
- Funding required: \$3.23 billion
- Overview: Ever-increasing violence related to land and resource disputes has compounded a longstanding humanitarian crisis, affecting millions in need.
- **Start date: March 2022 (Source)**

## Afghanistan:

- People needing aid: 22.9 million
- Funding required: \$3.04 billion
- Overview: Decades of conflict and political instability have left millions grappling with severe poverty and ongoing humanitarian needs.
- **Start date: August 2021 (Source)**

## Yemen:

- People needing aid: 19.5 million
- Funding required: \$2.5 billion
- Overview: The protracted civil war has devastated the economy and food systems, leaving a significant portion of the population in need of urgent assistance.
- **Start date: Sept 2014 (Source)**

## Ethiopia:

- People needing aid: 10 million
- Funding required: \$2 billion
- Overview: Ongoing conflict, disease outbreaks, and climate change have created overlapping crises, severely impacting humanitarian conditions.
- **Start date: November 2020 (Source)**

## Myanmar:

- People needing aid: 19.9 million
- Funding required: 1.1 billion
- Overview: Widespread violence and human rights abuses in Myanmar have created a humanitarian emergency, affecting millions amid ongoing conflict.
- **Start date: August 2017 (Source)**

## Chad:

- People needing aid: 7.8 million
- Funding required: \$1.5 billion
- Overview: A growing humanitarian crisis fueled by regional conflicts, hunger, and climate challenges has significantly impacted the population's well-being.
- **Start date: July 2009 (Source)**

## Appendix 3. List of media outlets included in the analysis

Outlet name	Website	Country
NYT	nytimes.com	United States
CNN	cnn.com	United States
Fox News	foxnews.com	United States
MSN	msn.com	United States
USA Today	usatoday.com	United States
New York Post	nypost.com	United States
Forbes	forbes.com	United States
AP News	apnews.com	United States
CNBC	cnbc.com	United States
yahoo	news.yahoo.com	United States
Washington Post	washingtonpost.com	United States
Newsweek	newsweek.com	United States
NBC News	nbcnews.com	United States
CBS News	cbsnews.com	United States
Wall Street Journal	wsj.com	United States
Buzzfeed	buzzfeed.com	United States
ABC News	abcnews.go.com	United States
Business Insider	businessinsider.com	United States
Drudge Report	drudgereport.com	United States
Reuters	reuters.com	United States
BBC	bbc.com/news	United Kingdom
Mail Online	dailymail.co.uk	United Kingdom
Mirror	mirror.co.uk	United Kingdom
The Sun	thesun.co.uk	United Kingdom
The Independent	independent.co.uk	United Kingdom
The Guardian	theguardian.com/international	United Kingdom
Sky News	news.sky.com	United Kingdom
Daily Express	express.co.uk	United Kingdom
The Telegraph	telegraph.co.uk	United Kingdom
ITV	itv.com/news	United Kingdom
Metro	metro.co.uk	United Kingdom
Manchester Evening News	manchestereveningnews.co.uk	United Kingdom
Times	thetimes.co.uk	United Kingdom
The Evening Standard	standard.co.uk	United Kingdom
Birmingham Live	birminghammail.co.uk	United Kingdom
GB News	gbnews.com	United Kingdom
Radio Times	radiotimes.com	United Kingdom
Channel 4	channel4.com/news	United Kingdom
Daily Record	dailyrecord.co.uk	United Kingdom
CBC (Canadian Broadcasting Corporation)	cbc.ca/news	Canada
CTV News	ctvnews.ca	Canada
The Globe and Mail	theglobeandmail.com	Canada
ABC News	abc.net.au/news	Australia
Sky News Australia	skynews.com.au	Australia
Sydney Morning Herald	smh.com.au	Australia
News.com.au	news.com.au	Australia
RT	rt.com	Russia
Al Jazeera	aljazeera.com	Qatar
IRNA (Islamic Iranian News Agency)	en.irna.ir	Iran



## Media and Journalism Research Center

### Legal address

Tartu mnt 67/1-13b, 10115,  
Tallinn, Harju Maakond, Estonia

### Postal address

6 South Molton St, London,  
W1K 5QF, United Kingdom

### Academic affiliation

Universidade de Santiago de Compostela (USC)  
Colexio de San Xerome, Praza do Obradoiro s/n,  
CP 15782 de Santiago de Compostela.

### Contact

[www.journalismresearch.org](http://www.journalismresearch.org)  
[mjrc@journalismresearch.org](mailto:mjrc@journalismresearch.org)

### Artificial Intelligence (AI) Disclosure Statement

AI tools were used in this research project for data collection and preprocessing purposes, as outlined in the Methodology section and detailed further in Appendix 1: Data Curation and Processing Framework. Grammarly was used to correct grammar during the editing phase. No AI tool was involved in drafting or shaping the analytical content of this study.

