WHEN TRANSPARENCY FAILS:

META'S POLITICAL AD POLICY DURING ISRAEL'S WAR ON GAZA

TABLE OF CONTENT

List of Abbreviations

AIPAC: American Israel Public Affairs Committee

API: Application Programming Interface

BSR: Business for Social Responsibility

CSV: Comma-Separated Values

DOI: Dangerous Organizations and Individuals

ICJ: International Court of Justice

IDF: Israel Defense Forces

MSF: Médecins Sans Frontières (Doctors Without Borders)

NGO: Non-Governmental Organization

OSB: Oversight Board

SIEP: Social Issues, Elections or Politics

UN: United Nations

WFP: World Food Programme

Executive Summary

This research examines Meta's enforcement of its Social Issues, Elections, or Politics (SIEP) advertising policy in the context of Israel's war on Gaza. It aims to evaluate whether moderation and enforcement practices on Meta's platforms demonstrate systemic bias in the treatment of pro-Palestinian and pro-Israeli political ads. In doing so, the study also critically assesses the reliability and limitations of Meta's Ad Library as a research and transparency tool, given its central role in facilitating public access to political advertising data.

Using a combination of quantitative and qualitative methods, the research analyzed more than 4,500 ads extracted from Meta's Ad Library Comma-Separated Values (CSV) datasets. Ads were selected based on three key timeframes: October 16–21, 2023; May 7–21, 2024; and August 5–10, 2024. We chose these periods due to their proximity to the events of October 7, the International Court of Justice's (ICJ) advisory opinion on Israel's occupation, and policy changes within Meta. The sample was limited to English-language ads published in the United States with large audience reach. Each ad was categorized according to political alignment (pro-Palestine or pro-Israel), removal status, factual accuracy, and presence of hate speech or violent language.

Findings reveal that while pro-Palestinian ads are often published by internationally recognized non-governmental organizations (NGOs) (such as Doctors Without Borders or UN bodies) and call for humanitarian aid or ceasefire. Pro-Israeli ads are usually run by media companies (private companies funded by Israeli lobbies such as Facts for Peace) and in support of the war. When an ad violated Meta's ads or content moderation policies, pro-Palestinian ads were removed at a faster rate than pro-Israeli ads. As a result, pro-Israeli ads remained visible for longer periods despite presenting a violation of Meta's policies. The disparity in enforcement contributed to unequal reach and engagement.

The study also found significant shortcomings in the functionality and transparency of the Ad Library itself. Ads visible in the Comma-Separated Values (CSV) export were often inaccessible via public search. Metadata related to removal justifications, targeting criteria, or moderation

timestamps was either missing or inconsistent. Furthermore, the lack of filtering and analysis tools within the interface made comprehensive analysis time-consuming and technically disconnected. These issues were further compounded by the shutdown of CrowdTangle[1] in August 2024, which eliminated a key mechanism for real-time monitoring of ad performance and content dissemination.

The research demonstrates that Meta's enforcement of political ad policies is marked by **procedural inconsistency, limited accountability, and structural opacity**. These flaws not only silence Palestinian perspectives but also hinder civil society's ability to assess whether content moderation practices are applied fairly. The report recommends greater policy clarity, improved transparency tools, and the reinstatement of real-time monitoring systems to ensure equitable and accountable digital governance.

Methodology

We relied on Meta's Ads Library to collect and analyze data related to political ads. The library is sorted by country and allows data to be downloaded in a CSV format using the following parameters.

Meta's Ads Library can be accessed through two main channels: the public web interface and the developer Application Programming Interface (API). Each access modality offers distinct features that directly impact the scope and depth of research.

Public access to the Ads Library is available through Meta's website. Any user can visit the page for a specific country and manually search for advertisements based on keywords, page names, topics, or advertisers. Through this interface, users can view active and inactive advertisements, along with associated metadata such as impressions range, spending range, and information about the sponsoring entity for political or issue-based ads. However, this method does not support automated querying or large-scale data extraction; searches must be conducted manually, and results are limited to what is rendered through the user interface.

[1]CrowdTangle was a public insights tool developed by Meta (formerly Facebook) to explore public content on social media platforms like Facebook and Instagram. It allowed researchers, journalists, and watchdog organizations to monitor and analyze the spread of content, including misinformation and disinformation, by providing access to data on public posts and engagement metrics. As of August 14, 2024, Meta discontinued CrowdTangle, replacing it with the Meta Content Library and Content Library API, which have more restricted access and functionality. More info: https://transparency.meta.com/researchtools/other-datasets/crowdtangle

In contrast, **developer access** is provided via the Ads Library API, which allows registered developers and approved researchers to conduct programmatic queries. The API offers structured metadata outputs, facilitating broader, more flexible, and reproducible data collection. It provides access to detailed attributes of advertisements, including demographic distribution of viewers, platform placements, and additional targeting information where available. Access to the API requires registration through "Meta for Developers," submitting an application, and complying with Meta's data use policies, particularly for accessing data related to political or social issue ads.

For the purposes of this research, **three filters** were applied when accessing the Meta Ads Library for the United States. First, only advertisements published in English were selected. Second, both active and inactive advertisements were included. Third, the research focused on advertisements with an estimated audience size exceeding one million users, thereby prioritizing ads with significant reach and potential impact.

Following the application of these filters, the resulting data was exported in the form of Comma-Separated Values (CSV) files. A CSV file is a widely used text-based data format that organizes information into a structured table, where each line represents a separate data entry and each value within a line is separated by a comma. This format is particularly suited for handling large datasets, as it enables efficient storage, analysis, and interoperability across various software tools such as spreadsheet applications and data analysis programs. Using spreadsheet software, we applied various filters to categorize and assess the ads based on key factors:

- **Political Alignment**
  Advertisements were categorized based on their alignment with either a pro-Palestinian or pro-Israeli stance. Ads calling for a ceasefire were generally considered to reflect a pro-Palestinian stance,[2] while those advocating for continued military action tended to align with a pro-Israeli position. Ads that called for humanitarian aid for both Palestinians and Israelis were excluded from this categorization. However, ads that requested aid exclusively for Palestinians or exclusively for Israelis were classified accordingly.

---

[2] These ads asked for a ceasefire in order to address the humanitarian crisis evolving in Gaza. We did not include the very small number of ads that asked for a ceasefire in a neutral tone because they don't represent a substantial dataset.

- **Factual Accuracy**

   We assessed the veracity of advertisement content to identify instances of misinformation, relying on fact-checks conducted by established news organizations such as *Le Monde*, *Reuters*, and *Forbes*. In addition, we drew on analyses published by recognized human rights and digital rights institutions, including *Human Rights Watch* and *Access Now*. Particular attention was given to claims presented as factual but which were demonstrably false or misleading.

   For example, according to analysis from the BBC, the claim that "The ICJ wants to destroy Israel" is entirely baseless.[3] The International Court of Justice has no mandate or authority to "destroy" any state. Similarly, ads that definitively stated that either Israel or Hamas was responsible for bombing the Al-Ahli Hospital were treated as misinformation, as no conclusive evidence had been presented by independent investigations at the time those ads were circulating[4]. Today, investigations confirm that Israel was responsible for the attack.

- **Disclosure and Removal**

   Instances where advertisements were marked with a "Paid by" disclaimer and the subsequent impact on their removal was recorded.

- **Online Violence and Hate Speech**

   Ads were examined to assess whether they included content that violated Meta's policies on violence, incitement to violence[5] or hate speech.

Additionally, we analyzed content based on a comprehensive list of keywords:

- *Palestine, Israel, From the river to the sea, War, Genocide, Genocidal, Conflict, Hostages, Palestinians, Israeli, Terrorism, Terrorist, Stand with Israel, Slaughtering, Hamas, IDF, Barbaric, Jewish state, Gaza, Ceasefire, Bombing, Holocaust, Dead,*

---

[3] Dominic Casciani, *"Israel-Gaza: What Did the ICJ Ruling Really Say?"* BBC News, May 17, 2024. Available at: https://www.bbc.com/news/articles/c3g9g63jl17o

[4] Human Rights Watch. *Gaza: Findings on October 17 al-Ahli Hospital Explosion.* November 26, 2023. https://www.hrw.org/news/2023/11/26/gaza-findings-october-17-al-ahli-hospital-explosion.

[5] Meta. *Violence and Incitement.* Transparency Center. https://transparency.meta.com/en-us/policies/community-standards/violence-incitement/.

Attack, Civilians, Occupation, Resistance, Casualties, Humanitarian aid, War crimes, Anti-Semitism.

It is important to note that many advertisements contained multiple keywords from the selected list, resulting in the same advertisement appearing across different keyword searches. Consequently, when duplicates were removed from the dataset to ensure the accuracy and integrity of the analysis, certain keywords no longer had distinct entries associated with them. This phenomenon reflects the overlap in content rather than a lack of presence of particular keywords within the ads.

We limited our review to advertisements published in English and targeted toward U.S. audiences. This scope allowed us to analyze a broader sample of content, while ensuring that Meta's content moderation policies were applied where their enforcement is most consistent. According to Meta's own policies, the U.S. English version of the Community Standards serves as the most up-to-date and authoritative reference for enforcement.[6] Meta's "Dangerous Organizations and Individuals" (DOI) list[7] also relies heavily on U.S. law and policy designations, which reinforces the centrality of U.S. standards in determining enforcement thresholds. Additionally, while whistleblower disclosures[8] and leaked internal documents[9] reported by *The Intercept* have revealed significant internal disagreements related to the Israeli war on Gaza, *The Guardian* reported operational struggles and shortcomings within Meta regarding the content moderation in the same context.[10] By focusing on the environment where Meta's policy framework and moderation capacity are supposed to be most developed, we aimed to assess whether any observable bias in enforcement reflects systemic patterns rather than technical limitations or isolated errors.

---

[6] Meta. *Community Standards*, Introduction, Transparency Center.
https://transparency.meta.com/policies/community-standards
[7] Meta. *Dangerous Individuals and Organizations*, Community Standards, Transparency Center.
https://transparency.meta.com/policies/community-standards/dangerous-individuals-organizations/.
[8] Meta Employees. *Stop Censoring Palestine: An Open Letter from Meta Employees*.
https://metastopcensoringpalestine.com.
[9] Sam Biddle, "Meta's Israel Policy Chief Tried to Suppress Pro-Palestinian Instagram Posts," *The Intercept*, October 21, 2024.
https://theintercept.com/2024/10/21/instagram-israel-palestine-censorship-sjp/.
[10] Kari Paul, "Meta Struggles with Moderation in Hebrew, According to Ex-Employee and Internal Documents," *The Guardian*, August 15, 2024.
https://www.theguardian.com/technology/article/2024/aug/15/meta-content-moderation-hebrew

Timeframe

As for the time period, we initially chose to analyze ads from two time periods: October 16–21, 2023, and August 5–10, 2024. An additional and special time period was later added: May 7, 2024 and May 21, 2024 during heightened "debate"[11] regarding the phrase "from the river to the sea."

The **first period** falls nine days after October 7, 2023, a time when content moderation policies were under heightened scrutiny.[12] This period represents a time of crisis for Meta's content moderation systems, which were under intense pressure. We aimed to examine how Meta enforced its Social Issues, Elections, and Politics (SIEP) policy during this critical moment. Meta had previously acknowledged that its content moderation system can be overly aggressive and prone to errors during times of crisis.[13] However, these errors have been reported to disproportionately affect pro-Palestinian content.[14]

The **second period** in August 2024, was chosen for several reasons. Meta had introduced updates to some of its policies, including the SIEP policy and related enforcement practices, in early summer 2024, with additional changes made in late August.[15]

For the purposes of this research, Meta added a special exception where it excluded hate and violent speech against entities designated as "terrorist" in August 2024 under the DOI policy: This includes "aspirational or conditional threats of violence, including expressions of hope that violence will be committed, directed at terrorists and other violent actors." This explains why violent or hateful towards Hamas may have been allowed to proliferate on the platform.

By selecting a window situated between these policy shifts, we aimed to avoid collecting data too close to any single update, thereby reducing the risk of capturing short-term anomalies or transitional enforcement inconsistencies. This period also followed the ICJ's advisory opinion on

---

[11] Bryan Metzger, 'From the River to the Sea': What to Know About the Pro-Palestinian Slogan That's Roiling American Politics," *Business Insider*, November 6, 2023. https://www.businessinsider.com/what-from-river-to-sea-means-israel-palestine-rashida-tlaib-2023-11

[12] Meta. "Meta's Ongoing Efforts Regarding the Israel-Hamas War." *Meta Newsroom*, October 13, 2023. https://about.fb.com/news/2023/10/metas-efforts-regarding-israel-hamas-war/.

[13] Alex Heath, *"Meta Says It's Mistakenly Moderating Too Much,"* The Verge, December 3, 2024. Available at: https://www.theverge.com/2024/12/3/24311513/meta-content-moderation-mistakes-nick-clegg

[14] Marwa Fatafta, *"How Meta Censors Palestinian Voices,"* Access Now, February 19, 2024. Available at: https://www.accessnow.org/publication/how-meta-censors-palestinian-voices/ .

[15] Meta. *Transparency Center: Policies*. https://transparency.meta.com/policies/.

Israel's presence in the Occupied Palestinian Territory,[16][17] a major legal development that sparked renewed political discourse across platforms.

More importantly, this second window allowed for the passage of time since the events of October 2023, when Meta's content moderation practices were operating under higher pressure. This helped us examine whether any enforcement irregularities observed in the first period were specific to crisis management or reflected broader, systemic issues in SIEP policy enforcement. By comparing these two periods, our research aims to capture both **short-term crisis enforcement** and **longer-term trends** in Meta's application of the SIEP policy.

The **third**, special timeframe falling between May 7 and May 21, 2024 was selected after Meta's Oversight Board (OSB) took on the case to examine whether the phrase "from the river to the sea" constitutes a call for genocide and whether it should be banned on Meta's platforms. As part of its process, the OSB invited public comments, with the comment period lasting until **May 21, 2024**.

The OSB is an independent body that reviews Meta's content moderation decisions.[18] It takes on cases where users believe content was wrongly removed, or that certain content should be banned in the future. OSB decisions on individual cases are binding; Meta must implement them. We selected this timeframe to capture data during a period when the discourse around the case was active.

In parallel with this case-specific analysis, we also examined how Meta enforces its misinformation policies. As outlined in its July 2024 policy,[19] Meta differentiates between types of misinformation and applies varying levels of enforcement, ranging from removal to reduced visibility. For the purposes of this research, we focused exclusively on the criteria and instances that led to content removal.

According to its Community Standards, Meta removes misinformation **only** in two cases:

---

[16] *The ICJ, Israel's Occupation, and the Realization of Palestinian Rights*, Democracy for the Arab World Now (DAWN), July 2024. Available at:
https://dawnmena.org/the-icj-israels-occupation-and-the-realization-of-palestinian-rights/

[17] International Court of Justice, *Summary of the Advisory Opinion of 19 July 2024: Legal Consequences Arising from the Policies and Practices of Israel in the Occupied Palestinian Territory, Including East Jerusalem*, July 19, 2024. Available at: https://www.icj-cij.org/node/204176

[18] https://www.oversightboard.com

[19] Meta. *Misinformation*. Transparency Center. https://transparency.meta.com/policies/community-standards/misinformation.

1. **"Misinformation that is likely to directly contribute to the risk of imminent physical harm."**

    ○ Example: False information about a violent conflict that could escalate tensions.

2. **"Misinformation that is likely to directly contribute to interference with the functioning of political processes."**

    ○ Example: False claims about voting procedures that could suppress voter participation.

As such, we assessed the presence of misinformation in political ads. Here, we did not expect Meta to remove all ads that we qualified as misinformation. Instead, we looked to achieve an objective analysis of what percentage of the ads from each side disseminated misinformation. We classified content as misinformation based on our own objective definitions. In our analysis, if an ad met any of the following conditions, we qualified them as misinformation:

1. **Opinions Stated as Facts**: Personal opinions or perspectives that are presented as objective truths.

2. **Debunked Theories and Misinformation**: Claims that have been fact-checked and proven false by credible sources.

3. **Statements and Facts Taken Out of Context**: Selectively presenting factual information in a misleading way, altering its intended meaning.

4. **Omission of Key Information to Support False Claims**: Deliberately leaving out crucial details to create a misleading or false narrative.

Some examples of what we categorize as misinformation according to this criteria:

1. **Claims that Hamas beheaded or burned babies on October 7**
   These allegations have been investigated and debunked by multiple independent sources, including a detailed report by *Le Monde*, which found no evidence supporting

these claims.[20]

2. **Equating Hamas with ISIS**

Statements claiming Hamas is equivalent to ISIS are inaccurate and misleading. Despite similarities in tactics or ideological references, the two organizations differ significantly in their histories, ideologies, objectives, and political contexts.[21]

3. **Attributing the bombing of Al-Ahli Hospital conclusively to either Hamas or Israel**

Claims definitively assigning responsibility for the bombing of Al-Ahli Hospital to either Hamas or Israel lack conclusive evidence. At the time these ads were published, Independent analyses had not reached a consensus, and investigations had yet to provide definitive proof.[22] Recent analysis shows that the attack was in fact conducted by the IDF.

4. **Labeling the phrase "from the river to the sea" as inherently antisemitic and as calling for the removal of Israel or Jews**

Stating categorically that the phrase "from the river to the sea" is antisemitic or that it inherently advocates for violence or removal of Jews or Israel is incorrect. The meaning of this phrase has been debated extensively,[23] including by Meta's own Oversight Board,[24] and there is no conclusive evidence that it universally implies genocidal intent or antisemitism.

We did not classify opinions as misinformation unless they are presented as factual statements or used in a way that supports false or misleading claims. This approach ensures that our

---

[20] Assma Maad, William Audureau, and Samuel Forey, *"'40 Beheaded Babies': Deconstructing the Rumor at the Heart of the Information Battle Between Israel and Hamas," Le Monde*, April 3, 2024. Available at: https://www.lemonde.fr/en/les-decodeurs/article/2024/04/03/40-beheaded-babies-the-itinerary-of-a-rumor-at-the-heart-of-the-information-battle-between-israel-and-hamas_6667274_8.html

[21] Colin P. Clarke and Michael Kenney. "Hamas Is Not ISIS — and the Comparison Itself Is Counterproductive." *Politico*, November 21, 2023. https://www.politico.com/news/magazine/2023/11/21/hamas-isis-are-not-the-same-00128107.

[22] Human Rights Watch, *"Gaza: Findings on October 17 al-Ahli Hospital Explosion: Evidence Points to Misfired Rocket but Full Investigation Needed,"* November 26, 2023. Available at: https://www.hrw.org/news/2023/11/26/gaza-findings-october-17-al-ahli-hospital-explosion

[23] SMEX. *Submission for OSB regarding "from the river to the sea."* May 2024. https://www.oversightboard.com/wp-content/uploads/gravity_forms/37-2d41e975b04d088bd93c8135cc47cfe6/2024/05/PAO-SMEX.pdf.

[24] Oversight Board. *Decision on Posts That Include "From the River to the Sea."* Decision no. BUN-86TJ0RK5, September 4, 2024. https://www.oversightboard.com/decision/bun-86tj0rk5/.

misinformation assessment is grounded in verifiable standards rather than subjective interpretations.

Our analysis is based strictly on Meta's policies in place before the changes to the content moderation policies and third-party fact-checking program that took effect in January 2025 (in the United States).[25] We relied on Meta's own descriptions of these policies and their definitions of terms related to political ads, misinformation, hate speech, and violence.

Challenges

We engaged in a comparative analysis of thousands of political ads about Palestine and Israel collected from Meta's public Ad Library, aiming to identify patterns in enforcement, removal timelines, misinformation treatment, and visibility. The research initially sought to explore whether flaws in Meta's detection systems, such as false positives and false negatives, might disproportionately favor one side over the other. However, during data collection and analysis, it became clear that the research's main questions and objectives could not be separated from the tools and mechanisms Meta provides for transparency. This is where the limitations of Meta's Ad Library, the opacity of enforcement criteria, and the lack of accessible data on content moderation decisions became prominent challenges to be addressed.

Throughout the process, we encountered significant obstacles related to data access, inconsistency, and traceability, all of which demonstrate the potential unreliability of Meta's transparency systems themselves. As a result, the research also sheds light on the **structural deficiencies of Meta's transparency center**, which hinder external oversight and allow critical patterns of policy enforcement to go unmonitored or obscured.

Systemic issues within Meta's ads transparency infrastructure that we identified included inconsistent data availability, limited traceability of ads, and opacity in how content moderation decisions are made. This translates to built-in lack of transparency within Meta's ad systems that prevents researchers and the public from understanding how moderation decisions are made, undermining public oversight/scrutiny. Without accurate and necessary transparency tools, civil society may not be able to hold Meta accountable and assess whether its policies are applied fairly.

---

[25] Liv McMahon, Zoe Kleinman, and Courtney Subramanian, *"Facebook and Instagram Get Rid of Fact Checkers,"* BBC News, January 7, 2025. Available at: https://www.bbc.com/news/articles/cly74mpy8klo

During this research, several other challenges arose that influenced both the scope and depth of our analysis.

Distinction Between Reach and Impressions: One of the challenges in analyzing the data was understanding and interpreting the key metrics used by Meta to measure ad exposure, reach, and impressions. Reach quantifies the total number of unique accounts that have seen an ad at least once, while impressions represent the total number of times an ad has been displayed, including repeated views by the same account. This distinction is crucial because while reach provides an idea of how many distinct users encountered an ad, impressions reflect how many times the ad was shown. The difference between these metrics complicates the ability to assess the exact level of exposure of an ad, especially since both are influenced not only by the content itself but also by the advertiser's bidding strategy, budget, and targeting parameters.

Limited Targeting Information: Meta's disclosure of targeting transparency remains highly limited. Meta provides only basic demographic details about ad audiences, age, gender, and location. The ad library offers no insight into the specific targeting criteria used by advertisers. This lack of transparency makes it difficult to fully understand the dynamics of how ads are delivered, leaving significant gaps in our ability to assess the exact factors influencing the reach and exposure of ads related to the Israeli-Palestinian conflict.

Lack of Detailed Information Regarding the Violation of Meta's policies: Meta does not provide information on which rule or policy an ad violated when it is removed. Instead, the platform simply states that the ad violated its Ads Standards[26] or Community Standards,[27] without detailing whether the removal was due to hate speech, misinformation, sexually explicit content, incitement to violence or other violations (*Figure 1*) . This lack of specificity prevents a deeper understanding of why certain ads were taken down, making it harder to assess whether there was any bias in enforcement of its ads policies.

---

[26] Meta. *Meta Advertising Standards*. Transparency Center. https://transparency.meta.com/policies/ad-standards/.
[27] Meta. "Community Standards." *Transparency Center*. https://transparency.meta.com/policies/community-standards/.
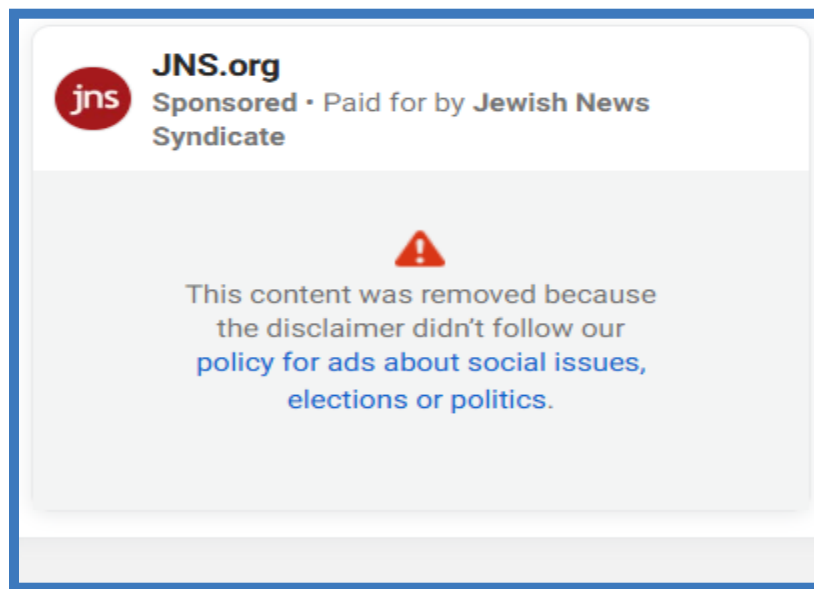
Figure 1. A screenshot of an ad removed for the violation of the SIEP policy.

Limited Functionality of Meta's Ad Library: The Meta Ads Library tool is a useful resource but lacks several essential features for understanding the dynamics of ad moderation, visibility, and policy enforcement over time. It does not offer enough filters or other tools to allow for deeper analysis or more precise segmentation of ads. Advanced research is also not possible as quotes are not allowed in the search box in the library. These limitations hindered our ability to efficiently categorize and assess the ads in relation to our research objectives, especially when dealing with large amounts of data. (As an example, the tool does not filter whether an ad was removed or not.)

Data Access and Ad Identification: Another significant challenge was the inability to link specific ads with identifiable data. The Ad Library did not allow for ad searches based on ad ID numbers, nor did it provide a way to retrieve data that connects ads with their corresponding identifiers.[28]

Archiving Issues: Meta's pages, including individual ad URLs and policy documents, are notoriously difficult to archive through standard web archiving tools such as the Wayback

---

[28] At the time of research, from October 2024 to January 2025, Meta's Ads Library tool did not allow search by ID. As of our last attempt, May 13, 2025 , Meta seems to have addressed the issue.

Machine.[29] In many cases, ad URLs only function for the user who originally accessed them and become inaccessible to others, even when shared directly. This severely limits the ability to preserve evidence, ensure transparency, and maintain verifiable records of removed or altered content over time. As a result, researchers face significant obstacles when attempting to document and audit Meta's content moderation practices.

Introduction

Political advertising on social media has reshaped how political messages are disseminated, allowing advertisers to micro-target audiences at an unprecedented scale. Among these platforms, Meta has emerged as a key player in the political advertising ecosystem. However, civil society is still concerned over the company's enforcement of its Ads about Social Issues, Elections or Politics[30] (SIEP) policy, particularly in sensitive geopolitical conflicts such as the Israeli war on Gaza.

Under public pressure after the Cambridge Analytica scandal in 2018, Meta (at the time still Facebook) implemented a set of policies governing ads about elections, social issues, and politics.[31] Advertisers running such ads must complete an authorization process, disclose their identity, and label their ads with a "Paid for by" disclaimer. Meta was already regulating election-related ads as required by U.S. law,[32] with some limited transparency governing campaign advertisements. However, in the aftermath of the Cambridge Analytica scandal, which exposed the misuse of personal data for political targeting,[33] Meta expanded its ad regulations to include not only election ads but also those related to broader social issues and politics. A key change happened in 2024, which is the AI update, where a disclaimer is added to posts made with AI. However, the rest of the SIEP policy has remained the same until today.

---

[29] The Wayback Machine is a digital archive of the internet created by the **Internet Archive**, a nonprofit organization. It allows users to view how websites looked at different points in time by capturing and storing snapshots (called *crawls*) of web pages. https://web.archive.org/.
[30] Meta, *"Ads about Social Issues, Elections or Politics,"* Meta Transparency Center. Available at: https://transparency.meta.com/policies/ad-standards/siep-advertising/siep
[31] Issie Lapowsky, *"Facebook Imposes New Restrictions on Ads and Popular Pages,"* Wired, April 6, 2018. Available at: https://www.wired.com/story/facebook-issue-based-ads-pages-restrictions
[32] United States Congress, *Bipartisan Campaign Reform Act of 2002*, H.R. 2356, 107th Cong. (2002). Available at: https://www.congress.gov/bill/107th-congress/house-bill/2356
[33] Carole Cadwalladr and Emma Graham-Harrison, *"Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach," The Guardian*, March 17, 2018. Available at: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

According to Meta's official policy,[34] ads that fall under the SIEP policy category include "any paid content, that is made by or on behalf of a candidate for public office, a political party, or advocates for the outcome of an election or referendum; or/and any paid content that references social issues that are debated or discussed across society, especially those that could influence public opinion, government policy, or behavior."

Meta defines social issues broadly and includes topics such as civil and social rights, health and medical care, immigration, security and foreign policy, environmental policies, crime and justice and political values and identity. In addition, Meta considers the context and intent of the ad. For example, an ad that merely mentions a social issue may not be considered SIEP, but if it seeks to influence opinion or action on the issue, it likely will be. However, despite these measures, multiple studies have highlighted enforcement inconsistencies.

For example, research[35] by Paul Bouchaud titled *"On Meta's Political Ad Policy Enforcement: An Analysis of Coordinated Campaigns & Pro-Russian Propaganda"* (2024), found that undeclared political ads frequently evade detection, while legitimate political ads sometimes face unnecessary restrictions. Using data from Meta's Ad Library across 16 EU countries, the study highlights that Meta's content moderation is often imprecise, with a false-positive rate of 60%, where non-political ads are incorrectly flagged as political and a false-negative rate of 95.2%, meaning political ads often go undetected.

The report points at a potential risk of bias in Meta's SIEP policy enforcement, particularly in times of heightened geopolitical tension. If Meta struggles to accurately distinguish between political and non-political content in a structured and regulated environment like the European Union, it becomes crucial to investigate how these challenges manifest in more polarizing and fast-moving contexts. This prompted us to turn our attention to the Israeli war on Gaza, a context where content is not only politically charged but also deeply entangled with issues of identity, human rights, and international law.

---

[34] Facebook Business, "About Ads with Political Content," *Facebook Business Help Center*,, https://www.facebook.com/business/help/167836590566506.
[35] Paul Bouchaud, *On Meta's Political Ad Policy Enforcement: An Analysis of Coordinated Campaigns & Pro-Russian Propaganda*, HAL Open Science, April 2024, p. 2. Available at: https://hal.science/hal-04541571v1/file/PoliticalAds.pdf .

We intended to look at whether the enforcement of Meta's ads policies is biased or prone to error in relation to the Israel-Palestine war, as after October 7, 2023,[36] a wave of ads from both sides engulfed social media platforms, especially on Meta. A report from CalMatters[37] shows data from Meta's ad transparency reports indicating a sharp increase in spending on ads mentioning Israel after October 7, 2023. In the months following the attack and Israel launching a war on Gaza, organizations and individuals spent between USD 14.8 and 22.1 million on Israel-related ads, generating between 1.5 billion and 1.7 billion views on Meta's platforms. This marks a 450% increase in ad spending compared to the same period in the previous year, when Israel-related ads cost between USD 2.4 and 4 million and reached between 373 million and 445 million views.
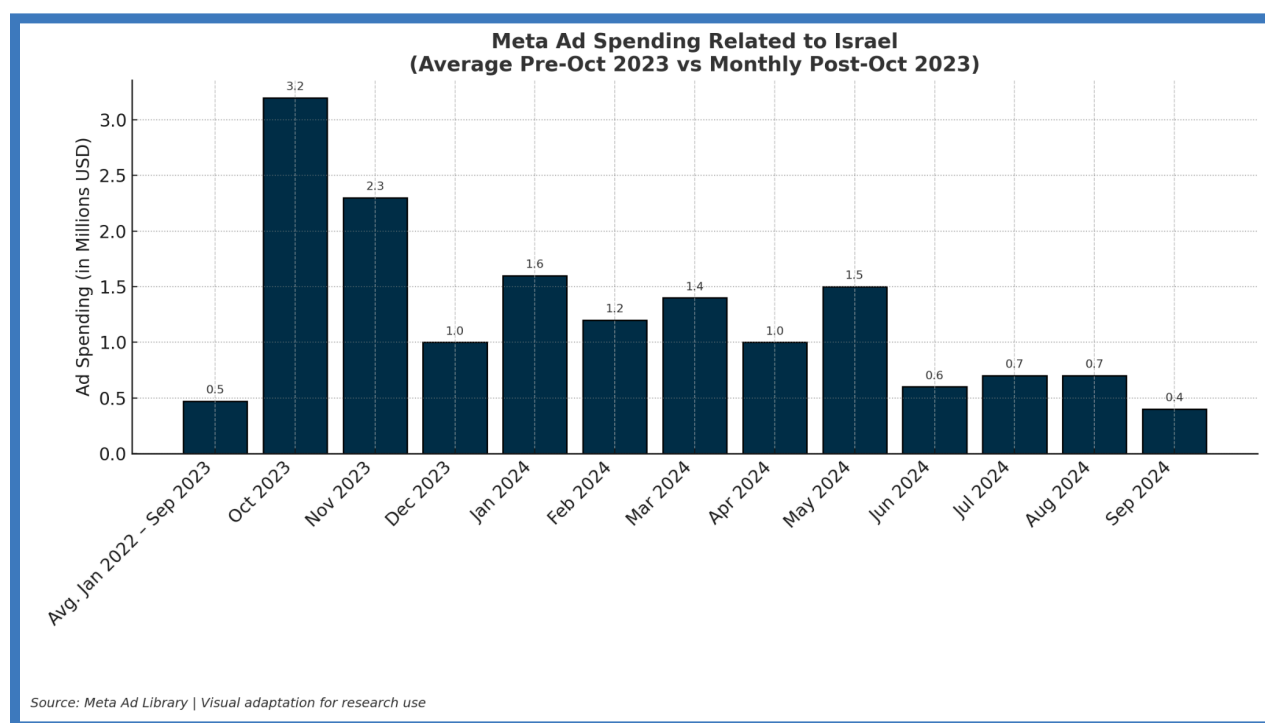
Figure 2. Source: Calmatters.

[36] France24, *"Hamas's 7 October Attacks: The Deadliest Day in Israel's History, One Year On,"* October 7, 2024. Available at:
https://www.france24.com/en/middle-east/20241007-hamas-terrorist-attacks-7-october-deadliest-day-israel-history-anniversary
[37] Colin Lecher and Tomas Apodaca, "How Meta Brings in Millions Off Political Violence," *CalMatters*, October 4, 2024,
https://calmatters.org/economy/technology/2024/10/how-meta-brings-in-millions-off-political-violence/.

According to CalMatters, among the top spenders was the American Israel Public Affairs Committee (AIPAC), a lobbying group that promotes Israeli policies in the United States. AIPAC's ad spending on Meta increased more than 300% in the six months after October 2023, totaling between USD 1.8 and 2.7 million (*Figure 2*). The organization used Facebook and Instagram ads to defend Israel's actions in Gaza and pressure U.S. politicians to support Israel.

Given Meta's deficiencies in moderating political ads, it is important to examine whether these shortcomings, intentional or not, have led to a bias related to the war.

Meta's over-moderation of Palestinian content, as outlined in a 2021 BSR (Business for Social Responsibility) report,[38] has since further intensified through the policies that the company adopted after October 7, 2023.[39]

In fact, Human Rights Watch documented[40] over 1,050 cases of censorship on Instagram and Facebook between October and November 2023. These cases primarily included content in support of Palestine that isn't normally expected to be removed in accordance with Meta's policies, with the overwhelming majority being suppressed, while only one case involved content supporting Israel. Meta's moderation practices were inconsistent and relied heavily on automated tools, which contributed to a significant increase in censorship of Palestinian voices, including those discussing human rights abuses. This point was also highlighted in a letter[41] from the U.S. Senators Elizabeth Warren and Bernie Sanders, who criticized Meta for its secrecy around content moderation and the censorship of Palestinian-related posts on Instagram and Facebook.

The Wall Street Journal reported that, following Israel's war on Gaza, Meta significantly lowered the sensitivity of the automation used to detect violent content for users in Palestinian territories

---

[38] Business for Social Responsibility (BSR), *Human Rights Due Diligence of Meta's Impacts in Israel and Palestine in May 2021: Insights and Recommendations*, SIEPtember 2022. Available at: https://www.bsr.org/reports/BSR_Meta_Human_Rights_Israel_Palestine_English.pdf

[39] Meta, *"Meta's Efforts Regarding the Israel-Hamas War,"* October 13, 2023. Available at: https://about.fb.com/news/2023/10/metas-efforts-regarding-israel-hamas-war/

[40] Human Rights Watch, *Meta's Broken Promises: Systemic Censorship of Palestine Content on Instagram and Facebook*, December 21, 2023. Available at: https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content-instagram-and

[41] U.S. Senators Elizabeth Warren and Bernie Sanders, *"Warren, Sanders Slam Zuckerberg, Meta for Suppression of Palestinian-Related Content, Urge Greater Transparency,"* March 26, 2024. Available at: https://www.warren.senate.gov/newsroom/press-releases/warren-sanders-slam-zuckerberg-meta-for-suppression-of-palestinian-related-content-urge-greater-transparency

to just 25%, far below the 80% threshold that *The Wall Street Journal* claims to be the usual threshold.[42] This alleged threshold change implies that Meta's system started removing or restricting content with far less certainty. Meta's moderation threshold refers to the confidence level its automated systems must reach before taking an action. According to *The Wall Street Journal*'s report, Meta usually requires its system to be at least 80% confident that a piece of content violates a rule before acting automatically, such as hiding the comment or post behind a graphic content warning label before a human moderator reviews it.

The lowering of the threshold meant that if the system was only 25% confident that a post might violate rules (particularly policies related to violence), it could still automatically take action, like hiding or limiting the content's visibility. This appears to have resulted in over-censorship of Palestine-related content, raising serious questions about whether Meta's moderation practices have resulted in a disproportionate take down of pro-Palestinian ads.

Noticibly, during the war in Gaza, many civil society organizations[43] and media workers began classifying Meta's moderation decisions as censorship and double standards since Palestinian voices were being stifled compared to Israeli content.[44] Some media outlets such as *The Intercept* are reporting that Meta is approving violent political ads in Arabic and Hebrew calling for the assassination of Palestinian activists[45] without implementing any mechanisms to control or oversee this content.

Meta's policy for political ads is ambiguous and limited. Overall, Meta focuses on elections when it comes to political ads and overlooks the situations of armed conflict such as the current context in Palestine.[46] Based on our literature review, there is no policy that regulates war time propaganda on Meta which can have major real life consequences during active conflict. On the

---

[42] Sam Schechner, Jeff Horwitz, and Newley Purnell, *"Inside Meta, Debate Over What's Fair in Suppressing Speech in the Palestinian Territories,"* *The Wall Street Journal*, October 20, 2023. Available at:
https://www.wsj.com/tech/inside-meta-debate-over-whats-fair-in-suppressing-speech-in-the-palestinian-territories-6212aa58
[43] Access Now, It's not a glitch: how Meta systematically censors Palestinian voices,
https://www.accessnow.org/publication/how-meta-censors-palestinian-voices/
[44] Al Jazeera, "Meta 'stifling' pro-Palestine voices",
https://www.aljazeera.com/news/2023/12/21/meta-stifling-pro-palestine-voices-on-social-media-hrw
[45] The Intercept, Facebook Approved an Israeli Ad Calling for Assassination of Pro-Palestine Activist
https://theintercept.com/2023/11/21/facebook-ad-israel-palestine-violence/
[46] Meta's Website: https://www.facebook.com/business/help/253606115684173

other hand, it is important to point out the fact that a crisis response control seems to exist as it was deployed[47] right after Russia's invasion of Ukraine in 2022.

As noted above, what Meta does have is a set of rules to publish political ads, including a needed authorization process and a disclaimer for political ads, based on information shared on their website.[48] Meta's claimed policy on political advertising is based on the principle of transparency, aiming to maintain an open space for users to engage with political messages and hear from political leaders, as long as such content complies with the platform's Community Standards. They also share this data about the political ads, their sources, the amounts of funding in their Ad Library including the ads that have been removed for violating Meta's policy.

Meta's longstanding issues[49] with content moderation and the substantial investment by Israeli lobbying groups in shaping online discourse have raised concerns about the neutrality and fairness in the enforcement of its political ads policy. A report from Forbes from November 2023 shows a rate of 91% compliance from Meta platforms to Israeli content removal requests.[50]

Aware of the reports of bias in Meta's content moderation policies and their enforcement, we decided to pursue this research after encountering a striking example: Mossad's official Facebook page had begun publishing ads targeting users in Beirut, asking Lebanese citizens to "provide information" in exchange for protection.[51] This sparked apprehension about a political entity's ability to exploit Meta's platforms with the purpose of distributing war propaganda within the territory of another sovereign state. It also prompted a broader question: Are Meta's ad policies, specifically the SIEP policy, adequately adapted to the conditions and sensitivities of wartime environments, especially in West Asia and North Africa? We wanted to explore whether flaws in Meta's detection systems, such as false positives and false negatives, might

---

[47] Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine: https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/
[48] Meta Website: https://about.fb.com/news/2020/01/political-ads/
[49] Deborah Brown and Rasha Younes, *Meta's Broken Promises: Systemic Censorship of Palestine Content on Instagram and Facebook* (New York: Human Rights Watch, 2023), https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content-instagram-and.
[50] Thomas Brewster, *"Israel Has Asked Meta And TikTok To Remove 8,000 Posts Related To Hamas War,"* Forbes, November 13, 2023. Available at: https://www.forbes.com/sites/thomasbrewster/2023/11/13/meta-and-tiktok-told-to-remove-8000-pro-hamas-posts-by-israel
[51] Souad Lazkani, *"Israel's Mossad Is Reportedly Contacting Lebanese People on Facebook,"* The961, December 15, 2020. Available at: https://www.the961.com/mossad-facebook-requests-lebanon/

disproportionately favor one side of the conflict over the other. While the final focus of this research centered on potential bias between pro-Palestinian and pro-Israeli political ads, the initial case served as a catalyst for a broader investigation into how political advertising is moderated and enforced on Meta platforms, particularly in times of conflict. In the course of our work, we also encountered significant challenges in using Meta's Ad Library tools, revealing serious limitations in transparency and accessibility, an important finding we explain thoroughly in this research.

Overview of Meta's Ad Policies

Definitions

Meta's advertising system operates under several policies. All content on the platform, including ads, must comply with the platform's Community Standards.These policies prohibit content involving hate speech, incitement to violence, misinformation, and more, and they apply universally across all content, including paid advertisements.

In addition, Meta maintains a separate Ads Policy, which includes further content and targeted restrictions specific to advertising. For example, ads promoting adult products, alcohol, or misleading business practices are restricted or banned under this policy. However, political and advocacy ads fall under a more specific category known as **Social Issues, Elections, or Politics (SIEP)**. This policy applies to ads that touch on topics deemed sensitive or capable of influencing public opinion on governance, human rights, or national identity. Under SIEP rules, advertisers must complete an authorization process, disclose their identity, and attach a disclaimer indicating who paid for the ad. Ads in this category are also subject to additional scrutiny and may be stored in Meta's Ad Library for increased transparency. This layered structure means that while all ads must respect the Community Standards, only certain content is flagged for enhanced regulation under SIEP.

Meta defines[52] advertisements about social issues, elections or politics as: "Made by, on behalf of, or about a candidate for public office, a political figure, a political party, a political action

---

[52] Meta. "Ads about Social Issues, Elections or Politics." *Transparency Center*. https://transparency.meta.com/policies/ad-standards/siep-advertising/siep.

committee or advocates for the outcome of an election to public office; or about any election, referendum, or ballot initiative, including go out and vote' or election campaigns; or about social issues in any place where the ad is being placed; or regulated as political advertising." Meta considers social issues "heavily debated and highly politicized sensitive topics that can influence many people and may impact the outcome of an election or result in legislation." The company further adds that "there are a variety of methods that people use to address social issues, in addition to elections or politics:

● *Leaders build their agendas around issues to inspire voters*

● *People vote on issues or support organizations and brands that align with their ideals*

● *Brands, advocacy groups and organizations may try to use paid advertising to broaden their voice or influence the sentiment of people to shape culture and inspire change.*"

To be able to place ads about social issues, elections or politics, advertisers need to verify their accounts by providing proof of their identity. Advertisers can only place such ads in the country in which they are authorized. Advertisers themselves need to label their ads as social issues, elections or political, and need to provide a disclaimer about who paid for the ad. Such ads then enter the Ad Library and remain there for seven years.

Banned Content in Ads

- **Unsafe or Discriminatory Practices:** Ads and advertisers must comply with applicable laws in their jurisdiction. Ads that promote or engage in illegal activities, fraud, or scams are banned. This includes misleading products, services, or schemes, such as counterfeit goods and Ponzi schemes.

- **Content That Compromises User Experience:** Ads that are excessively violent, sexualized, or not family-friendly are not permitted. Content that detracts from the overall user experience may be rejected.

- **Prohibited Content such as:** Weapons or Ammunition, Immoral or Unethical Practices (Ads that encourage cheating, manipulation, or other immoral behaviors), deceptive

Financial Services, (Ads for misleading or false financial products)

- **Unacceptable or Offensive Content:**

    - **Illicit Products or Services:** Ads that promote illegal products or services are strictly prohibited.
    - **Hate Speech and Discriminatory Content:** Ads that target or discriminate against people based on race, ethnicity, gender, sexual orientation, or other attributes are not allowed. Advertisers also cannot exclude specific groups from ad targeting.
    - **Incorrect Information:** Ads that fail third-party fact-checking, such as spreading false health claims, misleading news, or anti-vaccination content, will be removed.
    - **Hateful Content:** Ads that portray specific groups as inferior or a threat to public safety will be rejected.
    - **Endorsements of Extremist Individuals or Groups:** Ads that support or praise extremist groups or individuals are prohibited.

Key Findings: Meta's Opacity and Inconsistency in Policy Enforcement
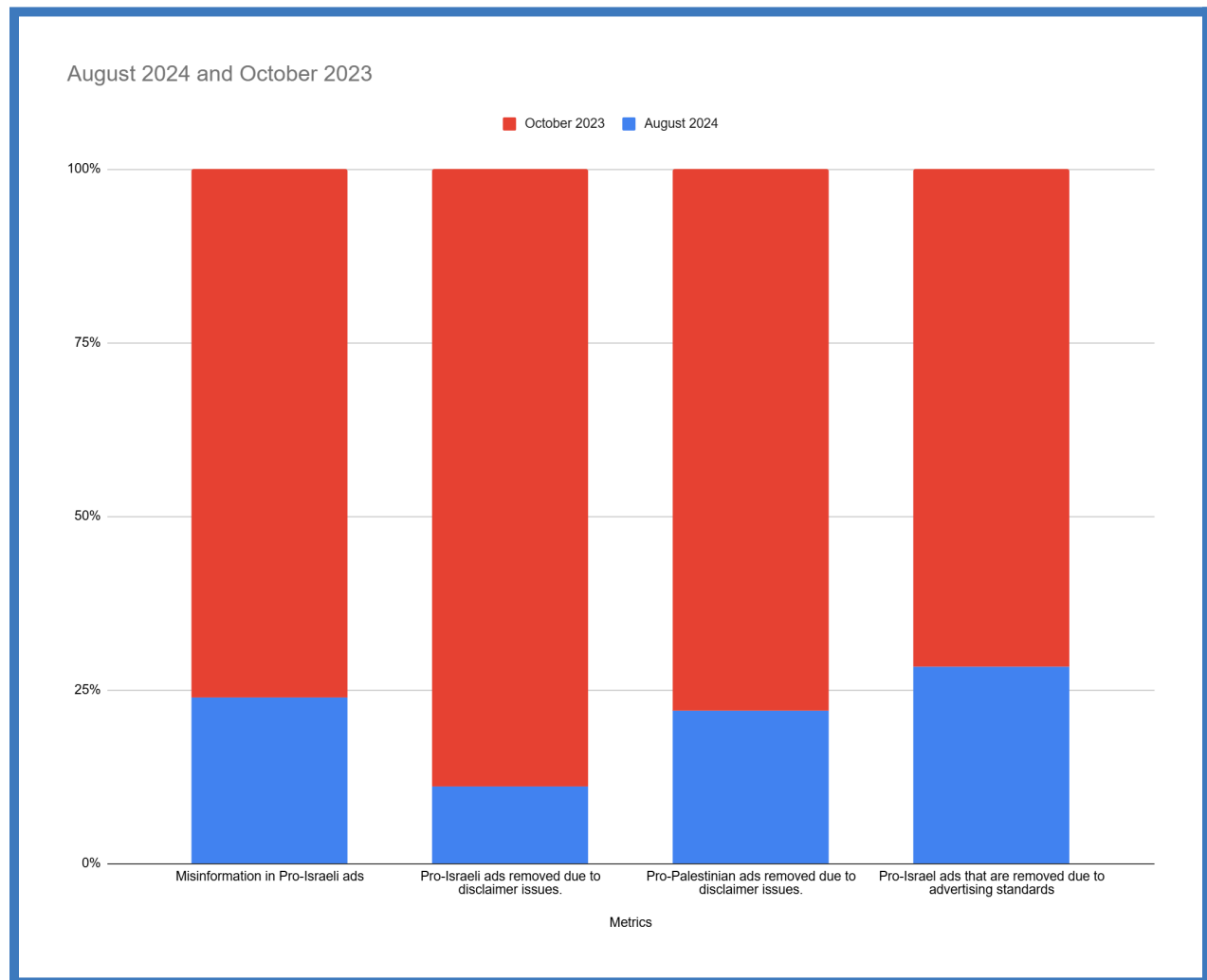
Figure 3: Shifts in Meta's Political Ad Enforcement on Israel-Gaza Content (Oct 2023 vs. Aug 2024).

One of the key findings of our research is that the majority of removed ads were taken down due to a **lack of a proper disclaimer, rather than violations of content policies**. This suggests that enforcement efforts are more focused on **procedural compliance rather than the substance of the ads themselves**.

1. A significant trend we identified is the **prevalence of fundraising ads related to the conflict**. These ads originate from both pro-Palestinian and pro-Israeli sources, either

solicit donations or sell merchandise under the pretense of raising funds for affected communities. While some of these fundraisers are from official NGOs or institutions, for other solicitors, there is no clear verification or proof that these funds actually reach their intended recipients. This highlights the risk of potential exploitation of the crisis for profit or misleading financial practices that go unchecked under Meta's current ad policies. The failure to verify the origin of these fundraisers allowed for the proliferation of harmful content and caused damage before. As demonstrated by a report from the American Sunlight Project,[53] large-scale cross-national investment scam campaigns exploited this loophole to reach a massive audience of over 100 million accounts in early 2024.
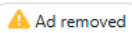
2. A critical issue we encountered during this research is the **lack of transparency around ad removal decisions**. Meta does not specify the exact reason an ad has been taken down, often citing only a vague reference to a violation of "advertising standards." In one example from our dataset (Figure 4), a pro-Palestinian ad was removed under this justification, yet the content of the ad contained no language, imagery, or claims that would reasonably warrant enforcement under Meta's own policies. There was no incitement, misinformation, or policy-violating material. Without any further explanation or access to a detailed review log, it is impossible to understand the rationale behind such enforcement actions. It appears that removals citing violations of Meta's advertising standards are not applied uniformly across all advertisers, but instead seem to disproportionately affect certain pages. In our dataset, multiple ads from both **Democracy Now** (a pro-Palestinian outlet) and the **Jewish News Syndicate** (a pro-Israeli outlet) were taken down under the same vague justifications such as violation of Ads Standards or violation of SIEP policy, suggesting that specific publishers may be subject to heightened scrutiny. The recurrence of removals from these particular sources raises questions about whether enforcement is being guided by consistent policy application or by undisclosed internal criteria affecting certain pages more than others.

---

[53] France24, *"'Sleeper Agent' Bots on X Fuel US Election Misinformation, Study Says,"* October 9, 2024. https://www.france24.com/en/live-news/20241009-sleeper-agent-bots-on-x-fuel-us-election-misinformation-study-says

Figure 4. *Two* *Screenshots showing an ad removed due to violation of Ads standards*

Figure 5. Screenshot showing an ad removed due to violation of SIEP.

3. Another key finding of our research is that **Meta rarely addresses misinformation within political ads**. As noted above, Meta takes down misinformation only in limited situations, and is more likely to address it through fact-checking or limited reach. However, these remed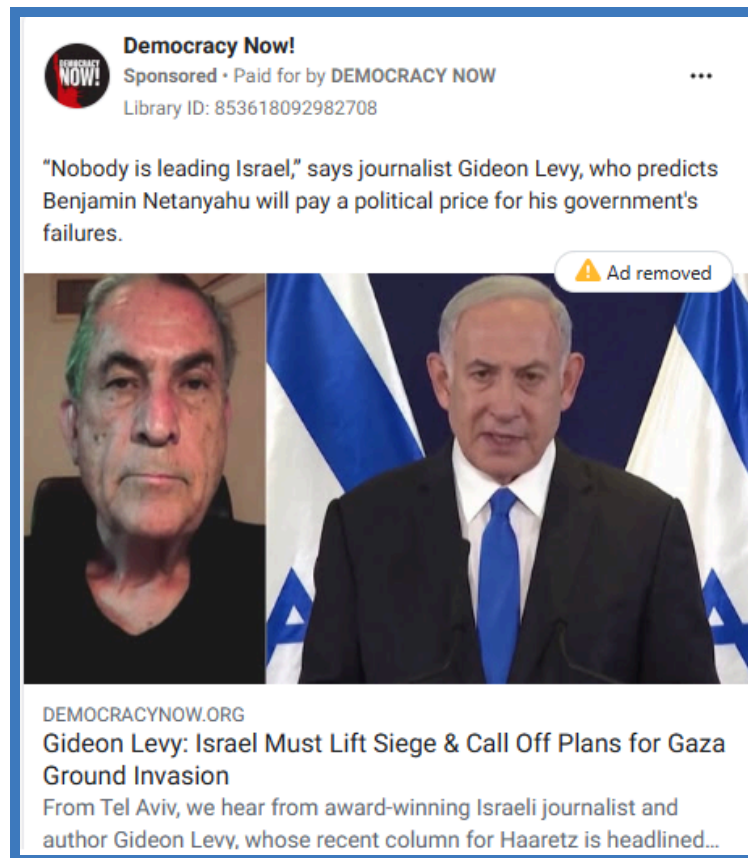ies do not appear to apply to ads. We observed a significant amount of misinformation in pro-Israeli ads, particularly in video format, which tends to evade detection more easily. For example, a media company called Facts for Peace, ran multiple ads containing misleading or unverified claims.[54] Despite Meta's policy of reducing the spread of misleading content, these ads were kept online.

4. **Some ads cannot be traced back to the Ad Library** by searching for their content, even while they are still active. The only way to locate these ads is through their unique ad ID, which requires a developer account and access to Meta's API. However, Meta's

---

[54] Al Jazeera. "Billionaires Are Teaming Up for Pro-Israel, Anti-Hamas Media Drive: Report." *Al Jazeera*, November 12, 2023.
https://www.aljazeera.com/news/2023/11/12/billionaires-are-teaming-up-for-pro-israel-anti-hamas-media-drive-report.

transparency tools do not allow searching by ad ID, making it impossible to verify or analyze certain ads unless they are discovered through alternative means. This made us question the effectiveness of Meta's ad transparency measures and the ability to track the full scope of political advertising on the platform.

5. Some ads were removed for violating Meta's advertising standards, while identical ads with the same content remained active, the only difference being the publisher (*Figure 6*).
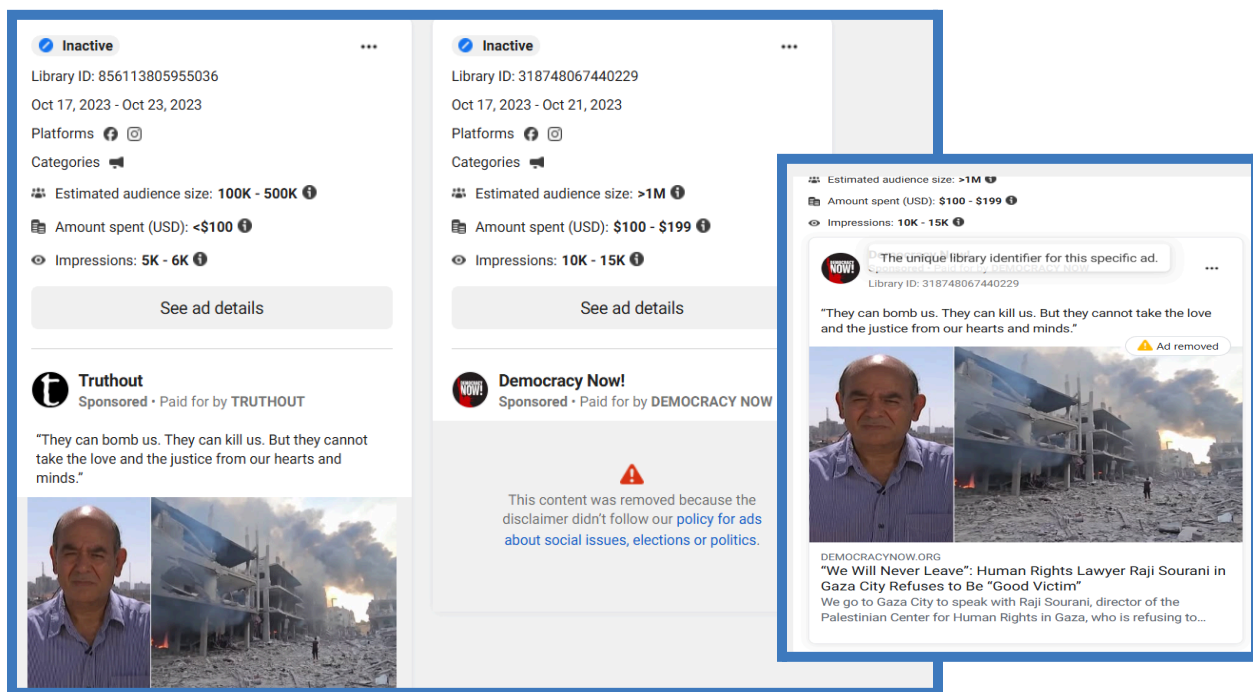
Figure 6: Three screenshots showing the same ad removed due to violation of SIEP in one page but not the other.

6. Another key observation from our analysis is the **discrepancy in removal timelines between pro-Israeli and pro-Palestinian ads** *(Figure 7)*. When Israeli ads were removed, regardless of the reason, whether for misinformation, violation of Meta's SIEP policy, or lack of disclaimer, the ads remained active for a longer period of time compared to pro-Palestine ads. As a result, these ads were able to accumulate significantly more views before enforcement action was taken. In contrast,

pro-Palestinian ads were removed more swiftly, often within the same day of publication, limiting their reach even when the violations are similar or less severe. This inconsistency in enforcement timing contributes to unequal visibility and influence, effectively giving more exposure to one side of the discourse while suppressing the other.

Average Days before Removal (October 2023)
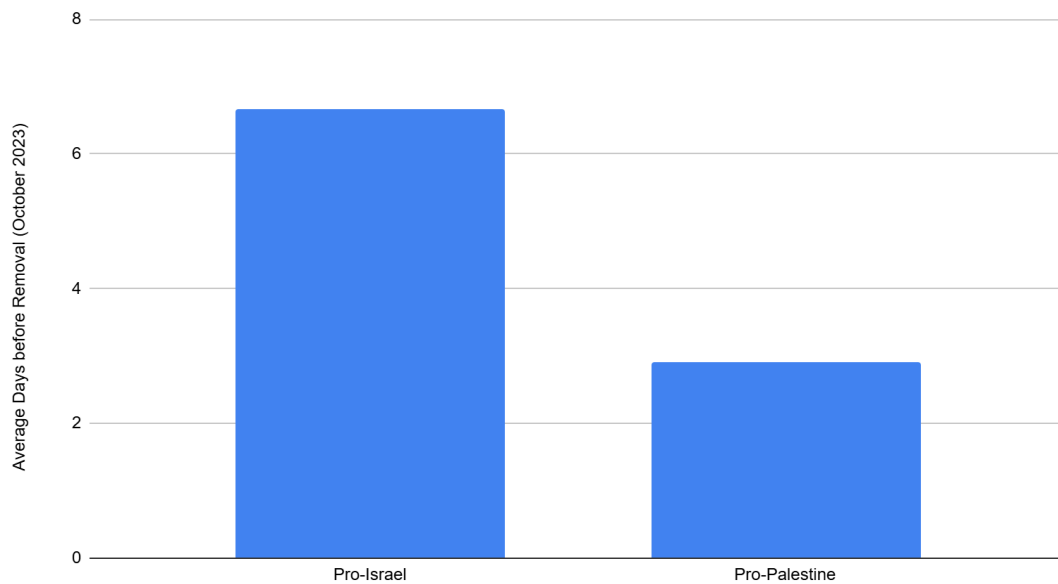
Average Days before Removal (August 2024)
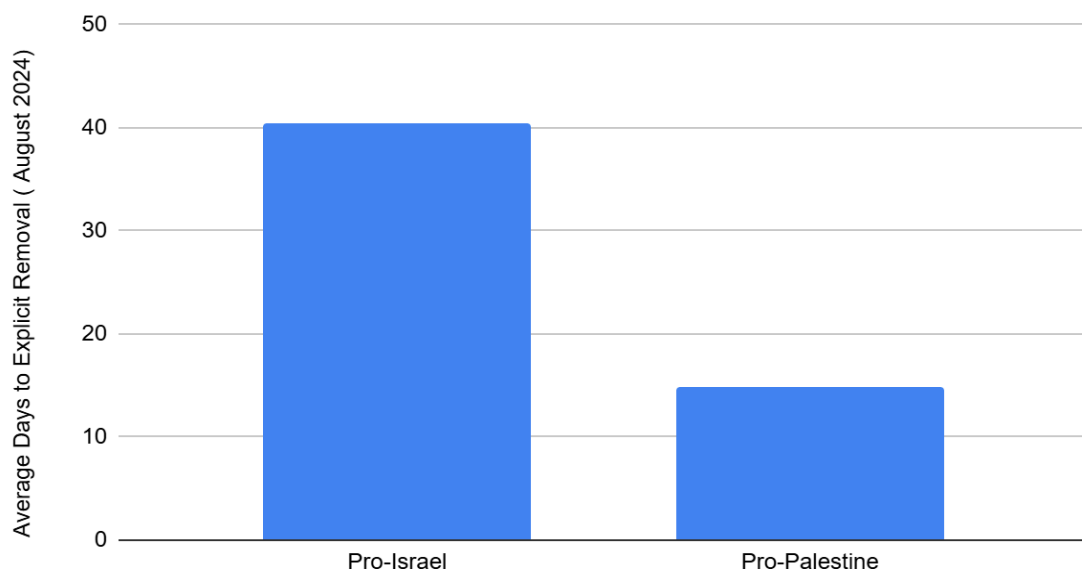
7. We observed an interesting **contrast between the sources and messaging of pro-Israeli and pro-Palestinian ads**. The majority of pro-Israeli ads originated from media companies and advocacy groups, often promoting narratives that supported the continuation of military operations or framing the conflict in absolute terms. These ads regularly employed strong political messaging aimed at justifying or escalating the conflict. In contrast, the pro-Palestinian ads largely came from reputable international NGOs, including humanitarian organizations, whose focus was predominantly on calls for humanitarian aid, ceasefires, and relief for civilians affected by the conflict.

Research Challenges

One of the major obstacles encountered in this research was the **limited usability and transparency of Meta's Ad Library**, which significantly shaped both the scope and execution of the study. While the Ad Library is promoted as a public transparency tool, its practical limitations undermine its potential to support independent research, particularly in politically sensitive and high-conflict contexts such as the Israeli-Palestinian issue.

1. Fragmented Access and Search Limitations

Meta's Ad Library interface provides only a narrow set of filters such as the language, location, platform, and search keywords. It lacks the ability to perform advanced searches like multi-filtering ads by multiple criteria (for example: date range + keywords + removal status + demographics), or query by advertiser ID. As a result, researchers cannot conduct precise or targeted searches, which increases the likelihood of missing relevant ads and makes trend mapping inconsistent.

2. Inconsistent and Incomplete Public Records

While Meta provides a CSV (Comma-Separated Values) export option through its Ad Library API, this format is essentially a raw spreadsheet file that lacks any integrated filtering or analytical capability. Each row in the CSV file corresponds to one ad and includes details like ad ID, start and stop time, the number of impressions, and the funding range. However, crucial pieces of information, such as whether the ad was removed or not, the reason for ad removal, details of targeting, or time and views before removal are not provided.

Some ads in the CSV datasets cannot be located through the Ad Library's search using the exact terms or content from the ad itself. This discrepancy suggests either a technical failure in the platform or might raise concerns regarding selective withholding of certain ads.

3. Missing or Delayed Metadata

Many ads were removed, but Meta does not specify why. The only indicator is a generic label stating "violates political ad policies," **without stating which clause or standard was breached**. This obscures whether an ad was removed due to lack of a disclaimer, hate speech, incitement, misinformation, or something else entirely. This absence of justification prevents meaningful interpretation of enforcement patterns.

→ We were unable to analyze engagement levels (views, shares) with full accuracy, as reach and impression data were often incomplete or inconsistent across records.

→ The inability to verify the reasons for removals led us to focus only on visible removal patterns and publicly accessible claims, which may have provided an inaccurate calculation of Meta's under- or over-enforcement. It also hindered the research results and detailing as we could not assume by ourselves the reasons for the removal.

Implications for Future Research

If Meta continues to limit access to granular data while maintaining vague policy enforcement disclosures:

- Researchers will face many hardships in auditing the effectiveness and fairness of Meta's moderation policies, particularly during crises.
- Without the possibility of using the ad ID numbers or saving specific ads, researchers risk loss of data.
- These constraints create a knowledge asymmetry. Meta and third-party contractors retain full datasets and AI moderation tools, while researchers and the public are limited to partial, delayed, or decontextualized data.

While Meta's Ad Library is presented as a tool for transparency, our experience in conducting this research demonstrates that it **falls short of supporting serious investigative or academic work.** The absence of clear removal justifications, inconsistent visibility of ad content, and the technical limitations of CSV data all contribute to an environment where platform accountability is severely restricted.

For example, despite the greater scalability offered by the developer access compared to the public access to the Ad Library, both access modalities present challenges. Public access is constrained by the inability to automate data collection and the limited scope of search results. Developer access, while enabling structured and automated data retrieval, is subject to rate limits, quota restrictions, and an approval process that can delay or constrain research timelines. Furthermore, certain data points may be redacted or generalized to protect user privacy, impacting the granularity of available information.

These limitations were further compounded by Meta's decision to shut down CrowdTangle in August 2024.[55] CrowdTangle had previously allowed journalists, researchers, and civil society actors to monitor real-time engagement data across Facebook and Instagram. It was one of the few tools offering reliable insight into how content spreads on Meta platforms, especially during periods of crisis or electoral activity. Its removal not only narrowed the public's ability to track viral disinformation and political amplification, but also significantly weakened the ecosystem of independent oversight over Meta's enforcement practices.

For researchers, the lack of a functional alternative to CrowdTangle, combined with the gaps and inconsistencies in the Ad Library, creates serious blind spots. It is now harder than ever to

---

[55] Meta, *"CrowdTangle,"* Meta Transparency Center. Available at:
https://transparency.meta.com/researchtools/other-datasets/crowdtangle.

verify how certain narratives are promoted, how quickly problematic content is removed, and whether moderation policies are being applied equitably across different geopolitical contexts.

Meta's claim to uphold transparency cannot be sustained when critical monitoring infrastructure is dismantled and when the data that is made available is fragmented, stripped of context, and technically inaccessible to most civil society organizations. If Meta genuinely aims to preserve a fair and open civic space, especially during conflicts or elections, it must not only restore research tools like CrowdTangle but also rebuild its transparency systems to support independent analysis, public scrutiny, and human rights monitoring.

Some Examples of "Coded Hate Speech"

In this section, we present examples of ads that, based on Meta's own policies on hate speech and online violence, should have been removed, but were not. These ads often escape moderation because they are published in video format or include external links, formats that appear to be less effectively monitored by Meta's enforcement systems. As a result, they remain online long enough to reach wide audiences, despite containing dehumanizing language, coded hate speech, or outright misinformation. This highlights a significant gap in Meta's content moderation, particularly in how its systems handle harmful content embedded in formats that are harder to flag automatically.

1. In this ad, the speaker in the video, Mark Levin, says *"Jews aren't even safe in their homeland,"* followed by, *"I would encourage everyone to open their Bible, I don't care if you look at the Old Testament or New Testament, show me where the Palestinians are, they don't exist."* The video accompanying this statement continues for several minutes, repeatedly asserting that Palestinians have no historical presence in the region. The speaker refers to them dismissively as *"so-called Palestinians"* and falsely claims that Palestinians "offered soldiers to the Nazis," invoking a discredited narrative often used to portray Palestinians as inherently violent or aligned with fascism. While the content avoids explicit calls to violence, it functions as coded hate speech by erasing Palestinian identity, linking them to historical atrocities, and delegitimizing their existence. The religious framing adds another layer of danger by suggesting that such erasure is divinely justified. According to our research criteria, this type of content constitutes online incitement through historical denial, demonization, and identity erasure.

The ad was not removed.

2. [Another ad](#) from a pro-Israeli media group called "Facts for Peace" has an ad title which contains both **coded hate speech** and **misinformation** that fall under multiple criteria:

First, the ad makes a false factual claim, asserting that Hamas decapitated babies during the October 7 attacks, a claim that has been debunked by credible media outlets, including *[Le Monde](#)*, and lacks verified evidence. Presenting such a claim as fact constitutes misinformation.

Second, the ad uses dehumanizing language and collective blame. By referring to *"Hamas's allies here at home"* and *"Palestinian sympathizers and progressive apologists"* who supposedly justify atrocities like rape, burning families alive, and executing the elderly, the ad blurs the distinction between Hamas and all pro-Palestinian individuals or movements. This tactic stigmatizes not only Palestinians but also anyone who criticizes Israeli policies.

The ad is linked to another website where the article contains multiple forms of coded and explicit hate speech targeting Palestinians. It repeatedly engages in historical erasure by claiming that "there never was such a thing as Palestinian land" and using "so-called," to describe Palestinians, effectively denying their identity and existence. It dismisses internationally recognized political grievances, such as occupation, apartheid, and land theft as "tired slanders" and "baseless lies," framing any Palestinian narrative as inherently deceitful. The text also draws dangerous associations between Palestinians and Nazism or Islamist extremism, portraying them as participants in a global religious war against Jews and Christians. By collapsing all pro-Palestinian voices into the category of "apologists" for terrorism, the distinction between Hamas and the broader Palestinian population or their supporters is blurred. Meta does not have any statement regarding links to other websites, nor an external links policy like Google[56] which brings the question of the possibility to circumvent Meta's policies through the use of external links just like in this example.

---

[56] YouTube Help. "External Links Policy." *YouTube Help Center*.
https://support.google.com/youtube/answer/9054257?hl=en.

Conclusion and Recommendations

This research set out to examine the enforcement of Meta's Social Issues, Elections or Politics (SIEP) advertising policy in the context of the Israeli-Palestinian conflict. While the initial objective was to assess whether moderation practices revealed bias between pro-Israeli and pro-Palestinian political ads, the investigation expanded to address a broader concern: the lack of transparency and structural limitations within Meta's ad infrastructure.

As the research progressed, it became increasingly clear that the issue was not only one of unequal moderation outcomes but also one of systemic opacity that obstructs meaningful public scrutiny.

The findings show that pro-Palestinian ads, often originating from humanitarian organizations and focused on calls for ceasefires or humanitarian aid, were more frequently and more rapidly removed than pro-Israeli ads regardless of the violation. Many of the latter, which came from media outlets or politically motivated actors, remained online longer and accumulated significantly higher reach, even when they contained misinformation, coded hate speech, or incitement. These patterns suggest a clear asymmetry in enforcement, raising serious questions about the consistency and neutrality of Meta's moderation processes.

In addition to the unequal treatment of content, the research highlights the severe limitations of Meta's ad transparency tools. The Ads Library, which is meant to enable public oversight, fails to provide critical metadata such as clear reasons for ad removals, enforcement timestamps, or targeting details. The CSV files lack coherence with the public interface, and many removed ads are impossible to trace or verify. Basic analytical features like filtering by removal status, tracking ad IDs, or reviewing engagement metrics are entirely absent. These deficiencies make it extremely difficult for civil society and researchers to determine how, why, and when ads are removed or allowed to circulate.

The situation worsened following the discontinuation of CrowdTangle in 2024. This tool was essential for real-time monitoring of how content spreads on Meta platforms. Its removal further limited independent researchers' ability to assess Meta's content governance, particularly during moments of crisis when oversight is most needed.

Taken together, these issues reveal more than individual enforcement flaws. They expose a structural opacity that shields Meta's decision-making processes from meaningful accountability. This undermines Meta's public claims to transparency, fairness, and human rights responsibility. For any political ad policy to be credible, it must be enforceable in a consistent and transparent manner. Meta's current system does not meet that standard.

Moving forward, it is essential for Meta to reform its political ads enforcement and transparency mechanisms. This includes restoring effective monitoring tools, providing detailed and accessible metadata for each ad, applying content policies evenly, and committing to genuine engagement with civil society. Only then can Meta begin to ensure that its platforms support a fair, safe, and rights-respecting digital public sphere.

Recommendations to Meta

1. **Reinstate and Maintain Access to Monitoring Tools such as CrowdTangle**
   Meta should immediately reverse its decision to sunset CrowdTangle and restore a functionally equivalent tool for civil society, journalists, and researchers. Real-time access to engagement and dissemination data is critical for monitoring the spread of political narratives, disinformation, and content moderation practices, especially during armed conflict or electoral periods.

2. **Improve the Functionality and Transparency of the Ad Library**

   - Include detailed metadata for each ad, such as specific reasons for removal and the enforcement timestamp.
   - Reinstate access to archived or removed ads, including visual content, for a defined period post-removal.
   - Enable search by ad ID and adding advanced filters (e.g. removal status, advertiser type, funding range, and date range).
   - Ensure consistency between data available through public interfaces and CSV exports.

3. **Clarify and Document Policy Enforcement Rationale**

Meta should provide precise, policy-based justifications when political ads are removed or restricted. General references to "violation of ad standards" are insufficient and prevent any meaningful review. This should apply to all enforcement actions under the SIEP policy and Community Standards.

4. **Establish Independent Auditing Mechanisms for Ad Enforcement**

Meta should commission or allow independent audits of its political ad moderation practices, especially during periods of crisis or heightened conflict. These audits should be periodic, publicly accessible, and include an assessment of enforcement disparities along political, linguistic, and geographic lines.

5. **Protect Humanitarian and Rights-Based Content from Political Misclassification**

Meta must ensure that ads advocating for humanitarian aid, ceasefires, or human rights, particularly by established NGOs, are not erroneously classified as political ads subject to heavy regulation or disclaimer requirements. Clear guidelines must be issued to distinguish advocacy from political influence.

6. **Address Evident Misinformation and Reinstate Fact-Checking Partnerships**

Meta must strengthen enforcement mechanisms against misinformation in political ads, especially when such content is paid and amplified. The company should **reinstate its independent fact-checking program** across all relevant regions and ensure that fact-checkers are authorized to review content flagged under the SIEP policy. This is essential to limit the spread of harmful, false, or misleading narratives, particularly during periods of conflict and crisis.

7. **Monitor Scam Campaigns and Review External Links in SIEP Ads**

Meta should introduce additional safeguards to detect and prevent **scam or fraudulent ad campaigns**, especially those requesting donations for humanitarian causes without transparency or verification. The company must also establish a robust process to **review external links embedded in SIEP ads**, which can serve as a loophole to evade Meta's advertising standards. These links may direct users to malicious, deceptive, or policy-violating content that escapes moderation simply by being hosted off-platform. At

a minimum, Meta should verify that external destinations in political ads comply with its policies before approval and throughout the duration of the campaign.

8. **Engage Civil Society in Policy Development**

Meta must create structured and regular channels of consultation with digital rights organizations, particularly those working in high-risk regions. This engagement must include transparent reporting on how feedback is incorporated into policy revisions.