



Holocaust education in networked publics: audience responses to a facebook campaign against holocaust distortion

Priscilla Dionesalvi¹ · Stefania Manca³ · Giorgia Minello²

Received: 26 June 2026 / Revised: 17 August 2026 / Accepted: 4 September 2026
© The Author(s) 2026

Abstract

This study examines audience responses to “Proteggi I Fatti”, an Italian Facebook campaign developed within the European project *Deconstruct: Deconstructing and Countering Holocaust Distortion via Campaign, Education and Training*. The campaign aimed to raise awareness of Holocaust distortion and denial by disseminating historically accurate information and supporting critical engagement with online content. The analysis is based on 10,160 public comments collected from the comment sections of nine Facebook posts published between December 2024 and May 2025. Methodologically, the study combines lexical and bigram analysis, sentiment analysis, word embeddings, and LLM-assisted classification through the Gemini API, followed by manual validation. Findings show that many users reframed the campaign’s educational messages through the lens of the Israeli-Palestinian conflict, especially the war in Gaza. Recurring patterns include Holocaust distortion, Nazi analogies, victim-perpetrator reversal, and accusations that Holocaust memory is instrumentalized for political purposes. The paper contributes to research on digital Holocaust education by showing how social media campaigns may become sites where remembrance, political conflict, antisemitism, and distortion intersect. It also highlights the value and limits of AI-assisted methods for identifying implicit and context-dependent forms of Holocaust distortion in platformed publics.

Keywords Holocaust distortion · Social media campaign · Facebook · Public engagement · NLP · LLMs



1 Introduction

Holocaust memory is increasingly produced, circulated, and contested within digital and platformed environments (Walden, 2021). Social media have expanded the publics that encounter Holocaust-related content, enabling educational campaigns, memory activism, and new forms of participation (Manca et al., 2025; Neiger, 2026). At the same time, these environments expose Holocaust memory to processes of trivialization, politicization, appropriation, and distortion (González-Aguilar & Makhortykh, 2022; Parnell & Stuckey, 2023). As several authors (Neiger et al., 2023; Novis-Deutsch et al., 2023) argue, the Holocaust is not only a historical event bounded in space and time but also an ongoing discursive event repeatedly mobilized in contemporary communication.

This study examines how Holocaust distortion emerges in audience responses to digital Holocaust education on Facebook. It focuses on “Proteggi I Fatti”, an Italian campaign designed to disseminate historically accurate information, raise awareness of Holocaust distortion and denial, and provide users with critical thinking tools for evaluating Holocaust-related content online. While existing research has examined Holocaust memory, distortion, and antisemitism in digital environments (Becker et al., 2024; Manca et al., 2025; Neiger et al., 2023; Walden, 2021, 2022), less is known about how audiences respond to campaigns explicitly designed to counter Holocaust distortion, and how such educational messages are reframed in comment spaces.

The Italian context is particularly relevant because Holocaust distortion intersects with unresolved memories of Fascism, Shoah, Italian responsibility for antisemitic persecution and contemporary antisemitism (Fogu, 2006; Gordon, 2012; Sarfatti, 2017). Italian public memory has often privileged narratives of anti-Nazi resistance, rescue, and German guilt, while engaging only unevenly with the role of Fascist institutions, racial laws, and local participation in the persecution and deportation of Jews (Foot, 2009; Sarfatti, 2006; Sierp, 2012). In this context, distortion frequently takes indirect forms, including minimization, relativization, symbolic appropriation, false analogy, and the equation of Jews, Israel, or Zionism with Nazism (CDEC, 2026; Chapelan & Becker, 2024; Gerstenfeld, 2007; Giesel, 2021).

This study is situated at the intersection of digital Holocaust education, networked publics, and computational analysis of online discourse. Digital media have become central to Holocaust remembrance and education, but their use raises important ethical, pedagogical, and methodological questions (Manca et al., 2025; Rich & Dack, 2022; Ulloa et al., 2026; Walden, 2022). Social media campaigns can extend educational outreach beyond schools, museums, and commemorative institutions; however, they also expose educational messages to publics who may negotiate, reject, redirect, or distort their intended meaning (Manca, *in press*; Neiger et al., 2023).

The present article analyses public comments posted in response to nine posts published by the “Proteggi I Fatti” Facebook page between December 2024 and May 2025. Combining Natural Language Processing techniques and Large Language Model-assisted interpretation, the study maps the narratives, emotional tones, semantic associations, and forms of distortion that emerged in the comment corpus.

The research is guided by the following question: How do Facebook users engage with and reframe Holocaust education content in the context of a campaign against

Holocaust distortion? More specifically, it investigates how Holocaust-related educational messages are received in a platformed environment, what narratives emerge around the Shoah, and how audience responses reproduce forms of Holocaust distortion, antisemitic discourse, Holocaust inversion, or victim-perpetrator reversal.

2 Background

Holocaust distortion refers to the misrepresentation, relativization, minimization, trivialization, or instrumentalization of the Holocaust (Gerstenfeld, 2007; Lipstadt, 2019; Weitzman, 2023). Unlike explicit denial, which rejects the historical reality of the genocide of European Jews (Kahn, 2011), distortion often acknowledges the Holocaust while weakening its historical specificity or mobilizing its symbolic authority for present political purposes (Weitzman, 2023). It may appear through inappropriate comparisons, moral inversions, accusations that Jews or Israel exploit Holocaust memory, or claims that contemporary actors are equivalent to Nazis (Becker et al., 2024; Chapelan & Becker, 2024; Giesel, 2021; Klaff, 2023).

Approaching Holocaust distortion as memory politics shifts the focus from historical inaccuracy alone to the social uses of Holocaust memory in the present (Assmann, 2010; Levy & Sznajder, 2002; Rothberg, 2009). The key issue is not only what people know about the Shoah, but how Holocaust memory is mobilized to claim victimhood, assign guilt, legitimize outrage, construct analogies, or reverse the positions of victims and perpetrators (Gerstenfeld, 2007). Because Holocaust memory has acquired a powerful moral status in European and transnational public culture, it can be invoked to educate, commemorate, and warn, but also to accuse, relativize, appropriate, or compete over suffering (Chapelan & Becker, 2024; Levy & Sznajder, 2002; Novis-Deutsch et al., 2023; Rothberg, 2009; Sierp, 2014; Subotić, 2023). This distinction between historically meaningful comparison and trivializing analogy is especially important for Holocaust education: while academic comparisons may support historical understanding, superficial Holocaust analogies used for political purposes constitute Holocaust trivialization and require explicit pedagogical attention (Eargle & Rubin, 2025). In this sense, distortion should be understood not simply as a failure of knowledge, but as a practice through which the past is made usable for contemporary moral and political conflicts.

One recurrent form of distortion is the use of Nazi comparisons and Holocaust inversion. Giesel (2021) shows that comparisons between Israel and Nazi Germany operate as communicative strategies of defamation and outrage, relativizing National Socialism while recasting the victims of the past as perpetrators in the present. Klaff (2023) similarly conceptualizes Holocaust inversion as an inversion of reality and morality, in which Jews or Israel are represented as “Nazis” and Palestinians or other groups as the “new Jews.” In this logic, Holocaust memory is not denied but appropriated and turned against Jews, Israelis, or Zionists. Bolton (2024) further shows that, in online debates, accusations of Israeli genocide are often accompanied by analogies between Israel and Nazi Germany and by attempts to equate Auschwitz or the Warsaw Ghetto with Gaza. In these cases, the concept of genocide functions

not only as a moral accusation against Israel, but also as a gateway to Holocaust distortion.

Recent studies on digital Holocaust memory show that social media intensify these dynamics by shaping how Holocaust references become visible, repeatable, affective, and contestable (Neiger et al., 2023; Walden, 2021, 2022). Holocaust discourse on social media ranges from commemorative and reverential uses to exploitative political uses and playful or inappropriate representations (Neiger et al., 2023). User-generated content is central to this transformation: it can support memory activism and education, but also enable trivialization, ridicule, or distortion of the past (González-Aguilar & Makhortykh, 2022; Manca et al., 2025; Neiger, 2026). For education, this means that correction and factual accuracy remain essential but insufficient. If distortion operates through the social and digital circulation of Holocaust memory, countering it also requires attention to how historical references are selected, framed, circulated, and interpreted in public and digital cultures (Becker et al., 2026).

Holocaust distortion also operates through broader processes of politicized memory and digital knowledge production. In their study of Holocaust-related debates on Twitter in Poland, Trzoss et al. (2023) show how historical memory can become a field of political conflict, identity construction, and selective narratives of victimhood, innocence, or moral superiority. Similarly, Grabowski and Klein (2023) show how digital knowledge environments can become sites of intentional distortion when ideological actors reshape public-facing historical narratives. These studies highlight that digital platforms do not simply transmit memory; rather, they structure the conditions under which memory is contested, reframed, and sometimes distorted. Building on this perspective, Jumle and Makhortykh (2026) conceptualise the evolving character of digital Holocaust memory through the notion of ‘protean memorialization’, showing how the discursive composition of Holocaust remembrance on social media changes over time through the interaction of users, platform affordances, and broader contextual factors.

Italy provides a revealing context for examining these processes. Holocaust remembrance is supported by a structured educational and commemorative infrastructure, including the “Giorno della Memoria”, school initiatives, Jewish organizations, memorial institutions, and national guidelines for teaching the Shoah (Manca & Guetta, 2023). Yet this infrastructure does not automatically prevent distortion. When remembrance becomes ritualized or detached from critical historical understanding, it may fail to equip learners to recognize false analogy, symbolic appropriation, or victim-perpetrator reversal (Sarfatti, 2017; Schwarz, 2021). This makes Italy a relevant context for examining how an educational campaign on Holocaust distortion is received in platformed publics.

The post-7 October context has further intensified the dynamics through which Holocaust memory is contested and distorted. Public and online debates surrounding Israel, Gaza, antisemitism, and Holocaust memory have made the Shoah one of the symbolic languages through which contemporary conflict is articulated (CDEC, 2026; Chapelan & Becker, 2024). They have also amplified pre-existing online repertoires in which accusations of Israeli genocide intersect with Nazi analogies and Holocaust distortion (Bolton, 2024). In such debates, Holocaust memory may be

mobilized through claims that Israel is carrying out a “new Holocaust,” accusations that Jews use the Shoah as a political shield, or attacks on Holocaust survivors and Jewish institutions for alleged silence, complicity, or instrumentalization of memory (CDEC, 2026). Recent research has also emphasised the need for a more differentiated understanding of such comparative uses of Holocaust memory. Examining user responses to the Auschwitz-Birkenau State Museum on X after 7 October 2023, Boyd and Merrill (2026) show that the comparative instrumentalization of Holocaust memory can take both commemorative and contestatory forms, ranging from empathetic engagement to polarizing equations between the Holocaust and contemporary violence. Their findings caution against treating all forms of comparison or instrumentalization as inherently antisemitic or distortive. A related process has been identified on Instagram, where representations of the Gaza war have become entangled with visual references to the Second World War and the Holocaust, including Nazi imagery and symbolic juxtapositions that collapse the historical distance between past and present conflicts (Grabowski, 2026).

Relatedly, Topor (2025) identifies terrorism denial as another post-7 October discursive pattern circulating online, characterized by the minimization, denial, or justification of the Hamas attacks, blame-shifting, and conspiracy narratives. Although the present study focuses on Holocaust distortion rather than terrorism denial, this broader discursive environment is relevant because it shows how denial, blame-shifting, and historical reframing circulated simultaneously in post-7 October online debates. These patterns connect the Italian case to broader international research on politicized memory, competitive victimhood, Nazi comparisons, Holocaust inversion, and the platformed circulation of antisemitic narratives (Giesel, 2021; Klaff, 2023; Trzoss et al., 2023).

Against this background, the present study approaches audience comments not as secondary reactions to educational content, but as traces of public meaning-making (Livingstone, 2004). The comments posted in response to the “Proteggi I Fatti” campaign allow us to observe how users receive, support, challenge, redirect, or distort Holocaust education in a networked public environment. In particular, the study examines whether the campaign’s educational messages remain anchored to Holocaust remembrance and historical accuracy, or whether they are reframed through contemporary political conflicts, especially the war in Gaza and debates on Israel, Zionism, and antisemitism.

3 Case study and data

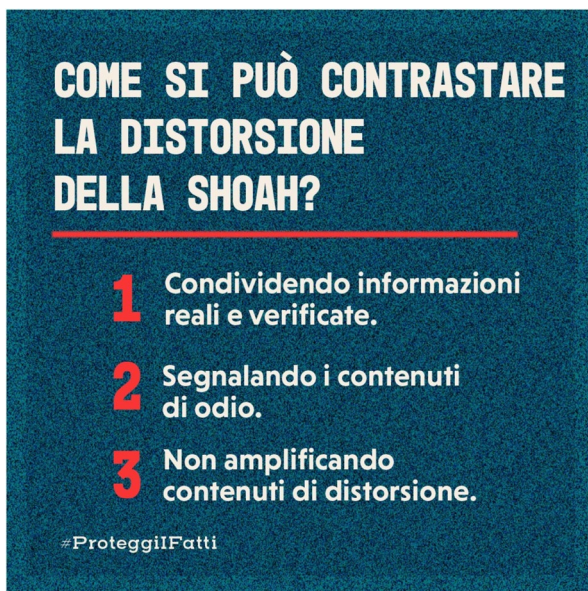
This study focuses on the “Proteggi I Fatti” Facebook page, established as part of the European project *Deconstruct: Deconstructing and Countering Holocaust Distortion via Campaign, Education and Training*. Through online communication campaigns, educational materials, and learning activities, the campaign aimed to disseminate historically accurate information and provide critical thinking tools to raise awareness of Holocaust distortion and denial (USC Shoah Foundation, n.d.). Facebook was selected within the DECONSTRUCT project primarily because the campaign

was intended to reach a broad, non-specialist adult public. The platform was considered particularly suitable for this purpose because of its potential to reach audiences beyond formal educational and commemorative settings. The campaign was implemented through a Facebook page that published nine posts between December 2024 and May 2025 and combined organic dissemination with paid advertising in order to increase visibility.

As shown in Fig. 1, the nine posts had a consistent structure: each included a brief description introducing the content displayed in one or more images, which followed the same visual style, with a neutral background and short text designed to capture the audience's attention. In some cases, the educational messages directly addressed Holocaust denial and distortion, while others focused on the role of educators and the importance of checking sources before posting or sharing content on social media. The content of the posts is listed in Appendix A.

The corpus under study was constructed by extracting 10,160 public comments via Meta's Facebook Graph. This number does not represent all comments published on the Facebook Page, but only those collected from the comment sections of the nine posts. For each post, comments were collected from the date of their publication until the data collection date (September 2025). The dataset included both comments and replies. This corpus makes it possible to examine how users engaged with an educational page dedicated to Holocaust distortion and how they received, reinterpreted, challenged, supported, or rejected the page's educational messages. During the campaign, the organisers did not directly engage with users by replying to comments or otherwise intervening in the discussion. Very light moderation was applied to hide a limited number of comments containing particularly offensive language, including personal insults and profanity. No systematic moderation protocol

Fig. 1 Example of a visual post published on Facebook page



was applied. Importantly, hidden comments remained accessible through the Meta Graph API and were therefore retained in the analytical dataset.

The study complied with relevant legal and ethical frameworks for internet-mediated research, including the EU General Data Protection Regulation (GDPR), the platform's terms of service and privacy policy, the ESRC Framework for Research Ethics, and guidelines on the use of social media data in research (European Commission, Directorate General for Research and Innovation, 2021; Stevens et al., 2015). Formal ethics committee approval was not sought, as the study used only publicly accessible comments posted on a public Facebook page and did not involve private communication or restricted-access data.

Data minimization and anonymization procedures were adopted throughout the research process. Usernames, profile information, account links, and other potentially identifying details were excluded from the analytical dataset. The analysis focuses on aggregate discursive patterns rather than individual users. Comments presented as examples were anonymized and, where necessary, lightly paraphrased or translated to reduce traceability while preserving analytical meaning. The dataset was analysed through the mixed computational and interpretive approach described in the following section.

4 Methodology

The methodology integrates traditional Natural Language Processing (NLP) techniques with Large Language Model (LLM)-assisted qualitative analysis. The analysis was divided into two phases. The first phase examined the corpus at the lexical and semantic levels, while the second phase used Gemini 2.5 Pro to identify and classify potential narratives of Holocaust distortion. Traditional NLP techniques were used during the initial exploratory phase to establish an empirical basis for subsequent analyses. This phase helps ensure that the LLM's results are grounded in the actual semantic distribution of the data. The LLM was introduced primarily to capture implicit forms of distortion that traditional NLP methods were unable to detect, given their limitations.

Each method was selected to capture a different dimension of audience engagement. Frequency analysis was used to identify the dominant lexical themes in the corpus, while bigram analysis highlighted recurrent associations between concepts. Sentiment analysis provided a broad indication of the affective orientation of the comments, whereas toxicity detection was used as an indicator of overt verbal aggression. Word embeddings were employed to explore the semantic environment surrounding the term 'Shoah'. Finally, the LLM-assisted analysis was used to identify broader narrative patterns and context-dependent forms of Holocaust distortion that could not be reliably captured through lexical or classification-based measures alone.

In the first phase, the comments were tokenized using the Italian spaCy model (Honnibal et al., 2020) and filtered to remove terms that were irrelevant to the analysis. The twenty most frequent terms were then identified for each post. The twenty most frequent bigrams were also extracted in order to identify recurring associa-

tions between pairs of terms. The threshold of twenty was adopted as a pragmatic exploratory choice, balancing breadth and interpretability: a smaller set would have provided a more restricted representation of recurrent lexical patterns, whereas a substantially larger set would have made the exploratory comparison across posts less manageable. In addition, sentiment analysis was used to examine the emotional tone of users' comments. It was conducted with the Italian model MilaNLP/feel-it-italian-sentiment (Bianchi et al., 2021), accessed through the HuggingFace Transformers library (Wolf et al., 2020), and used to classify each comment as positive or negative. A methodological limitation is that no independent manual validation of the sentiment analysis was performed on a subset. Furthermore, the presence of terms that directly refer to the traumatic event of the Holocaust may have influenced the model toward the negative pole. This issue made it necessary to use LLM-assisted analysis to interpret the nuances in the content of the comments.

Subsequently, toxicity detection was conducted to measure harmful language. This step used the multilingual model provided by the Detoxify library, based on the XLMRoBERTa architecture (Conneau et al., 2020). The model assigned each comment a probability score between 0 and 1 for toxicity. However, in this study, toxicity detection was not a sufficient measure of antisemitism or Holocaust distortion because the model tends to identify direct insults or swear words more effectively than implicit hostility. To make the results easier to interpret, the numerical values were grouped into five categories:

- “Very low”: scores below 0.20;
- “Low”: scores from 0.20 to 0.39;
- “Moderate”: scores from 0.40 to 0.59;
- “High”: scores from 0.60 to 0.79;
- “Very High”: scores of 0.80 or above.

Semantic analysis was performed through word embeddings, which transform words into vectors in a multidimensional space, where proximity between terms indicates that they tend to appear in similar contexts. This method was used to explore the semantic field surrounding the target word “Shoah”. The analysis was inspired by Levis Sullam et al. (2022), who expanded the work of Tripodi et al. (2019) on semantic shift in historical corpora. Following this approach, Gensim's Word2Vec model was trained on the corpus (Řehůřek & Sojka, 2010), using the Skip-Gram architecture with a five-word context window and a vector size of 100, given the relatively small size of the dataset. The twenty words closest to “Shoah” were identified through cosine similarity. The vectors were normalized to prevent terms with similar directions from appearing artificially distant (Milvus, 2025). The identified terms and the target word were then projected through Principal Component Analysis (PCA) to visualize their two-dimensional semantic positioning. PCA was used as an exploratory step to represent the terms identified by the model as semantically closest to “Shoah”. The visualization presented in the Findings section illustrates this analysis for the first post and supports the transition from the quantitative analysis to the subsequent LLM-assisted qualitative analysis.

The second phase used Gemini 2.5 Pro via the Gemini API (Google, 2025) and prompt engineering techniques. All prompts are listed in Appendix B. The prompts were defined prior to the final analysis and were not iteratively modified on the basis of the classification results. The refinement prompt constituted a separate verification step rather than a modification of the original classification prompt.

In the first step, the model was asked to conduct an inductive analysis of the comments using a zero-shot technique, since the prompt did not provide predefined categories or examples. The model was run with a temperature of 0.0 to reduce output variability (DAIR.AI, n.d.-a). This step was conceived as an exploratory phase to identify users' narratives surrounding the Holocaust before moving to a structured classification. Rather than providing the model with predefined categories, the prompt asked it to identify the dominant representations of the Shoah emerging from the comments. The prompt designed for this analysis considered only those comments in which the term "Shoah" was directly mentioned. The model also grouped and classified these representations.

The second step focused on detecting potential antisemitic narratives and Holocaust distortion. Detailed instructions and examples were provided to the model in order to guide the analysis, following a few-shot prompting approach (DAIR.AI, n.d.-b). This was necessary because recent reports, such as the ADL Center for Tech & Society's AI Index (ADL, 2025), suggest that LLMs can identify general hate speech more easily than implicit or context-dependent antisemitic narratives. For this reason, the classification prompt was based on eight categories from the Decoding Antisemitism lexicon (Becker et al., 2024), selected because they were directly connected to the Shoah and consistent with the IHRA framework for recognizing and countering Holocaust distortion (IHRA, 2021): (1) "demands for a clean break", (2) "Holocaust distortion", (3) "instrumentalization", (4) "Nazi-Jewish collaboration", (5) "Jews have not learned from the past", (6) "admonishers", (7) "affirmation of Nazis/Hitler", and (8) "Nazi analogy". The model was guided by this taxonomy and run with a temperature of 0.0 in order to reduce variability in the outputs and improve consistency across classifications (IBM, n.d.-a). A second refinement prompt was then used to review the initial classification and identify possible inconsistencies. Although the model was run with a temperature of 0.0, this setting does not guarantee the complete absence of stochasticity (Vziatysheva & Makhortykh, 2024). To assess the stability of the classification, the same Decoding Antisemitism-based classification prompt was therefore rerun on a subset of 50 comments while maintaining identical hyperparameters.

To evaluate the reliability of Gemini's final classification, a manual validation procedure was conducted using a Ground Truth approach (IBM, n.d.-b). A sample of 200 comments was manually annotated according to the Decoding Antisemitism guidelines (Becker et al., 2024) by a single human annotator. Group A consisted of 100 comments that Gemini had classified as antisemitic or as containing Holocaust distortion. This subset was used to assess the precision of the model-labelled cases. Group B consisted of 100 randomly selected comments from the whole dataset and was used to identify possible false negatives and evaluate whether the model had missed problematic content in the broader corpus. The manual annotation helped

assess the reliability of the LLM-assisted classification while acknowledging the need for expert interpretation in identifying implicit, context-dependent, or ambiguous forms of Holocaust distortion.

The validation produced quantitative performance metrics, reported in the findings section: true positives, true negatives, false positives, false negatives, precision, and recall. Since the validation was performed on a small subgroup, the metrics obtained should be interpreted with caution, as the results do not provide definitive validation for the entire corpus.

5 Findings

The analyses described in the previous section yielded both quantitative and qualitative results. The first phase focused on the lexical structure of the corpus and on the identification of sentiment, toxicity, and semantic proximity. These analyses provided a transparent observation of the corpus, but they also showed limitations in capturing implicit hostility. For this reason, the second phase used LLM-assisted qualitative analysis to study how users referred to the Shoah and to classify antisemitic narratives. In particular, it focused on identifying implicit forms that could not be captured using quantitative methods alone.

5.1 Computational findings on the comment corpus

The frequency analysis shows that the most frequent terms are primarily related to Israel, Palestine, Gaza, genocide, and the Holocaust. The ten most frequent terms are *israele* (Israel, 1,385), *palestinesi* (Palestinians, 1,222), *gaza* (Gaza, 1,099), *genocidio* (genocide, 938), *hamas* (Hamas, 911), *ebrei* (Jews, 907), *shoah* (Holocaust, 770), *palrestina* (Palestine, 531), *israeliani* (Israelis, 527), and *bambini* (children, 417).

The bigram analysis further specifies these lexical patterns. The ten most frequent bigrams identified are: *genocidio gaza* (genocide Gaza, 113), *pulizia etnica* (ethnic cleansing, 73), *donne bambini* (women children, 67), *genocidio palestinesi* (genocide Palestinians, 54), *genocidio palestinese* (Palestinian genocide, 51), *gaza cisiordania* (Gaza West Bank, 46), *guerra mondiale* (world war, 36), *distorsione shoah* (distortion Shoah, 34), *striscia gaza* (Gaza Strip, 29), and *medio oriente* (Middle East, 28).

Sentiment analysis was used as an exploratory measure to provide a broad description of the sentiment distribution in the corpus. The model processed 9,619 comments. The 541 excluded comments were not classified due to a lack of content to analyse. The sentiment model categorized comments only as negative or positive. There was no neutral category. Among the analysed comments, 8,638 (89.8%) were classified as negative, while 981 (10.2%) were classified as positive. Given the binary nature of the model and the absence of a neutral category, these results should be interpreted with caution. In a corpus dealing with the Holocaust, war, genocide, and violence, negative sentiment may reflect the traumatic subject matter itself rather than hostility, antisemitism, Holocaust distortion, or rejection of the campaign.

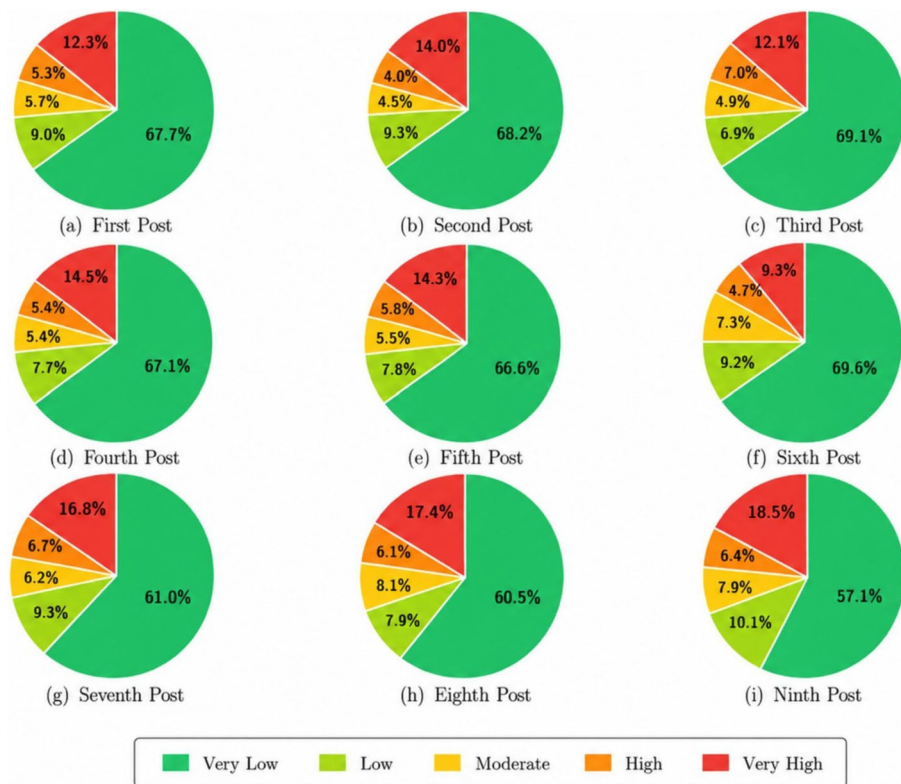


Fig. 2 Toxicity distribution per post

The next step measured the toxicity (Fig. 2). The model divided the comments into five categories. Across the nine posts, the largest category was “very low toxicity,” which ranged from 57.1% (Post 9) to 69.6% (Post 6). The second-largest category was “very high toxicity,” ranging from 9.3% (Post 6) to 18.5% (Post 9). The intermediate categories remained lower: “low toxicity” ranged from 6.9% to 10.1%, “moderate toxicity” from 4.5% to 8.1%, and “high toxicity” from 4.0% to 7.0%. Since Detoxify is more sensitive to explicit profanity, insults, and trigger words, toxicity scores were used as an indicator of verbal aggression. Implicit forms of antisemitism were therefore examined in the LLM-assisted qualitative phase.

The PCA results indicate that the terms associated with “Shoah” include *memoria* (memory), *sterminati* (exterminated), *superstiti* (survivors), *tragedia* (tragedy), *gaza*, *sionisti* (Zionists), *coloni* (settlers), *alibi*, *scudo* (shield), *giustificazione* (justification), and *genocidio* (genocide). Overall, “Shoah” appears within a semantic field that combines the lexicon of Holocaust memory with references to the Israeli-Palestinian conflict and to the political instrumentalization of the past.

Figure 3 illustrates this pattern for the first post. In this visualization, “Shoah” is positioned near memory-related terms such as *memoria* (memory), *sterminati* (exterminated), *subita* (suffered), *superstiti* (survivors), and *ebraici* (Jewish), but also

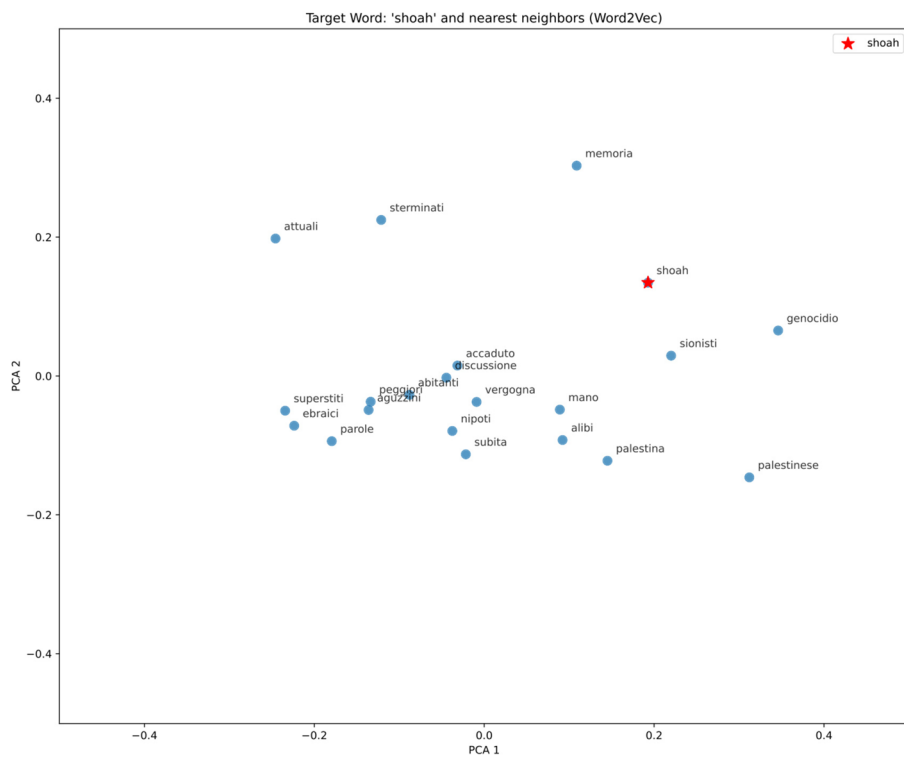


Fig. 3 PCA projection of the Word2Vec nearest neighbours for the target word “shoah” in the first post

near terms such as *palestina* (Palestine), *palestinese* (Palestinian), *sionisti* (Zionists), *genocidio* (genocide), *alibi*, *vergogna* (shame), and *aguzzini* (tormentors).

The computational analysis therefore suggests that the campaign’s educational focus was reframed through a semantic field dominated by Israel, Gaza, genocide, and the Israeli-Palestinian conflict. At the same time, the toxicity results indicate that Holocaust distortion does not necessarily rely on overtly abusive language, but often emerges through comparison, insinuation, moral inversion, and political reframing. These limits of lexical, sentiment, and toxicity-based measures motivated the subsequent LLM-assisted qualitative analysis.

5.2 Qualitative findings: Shoah-related and antisemitic narratives

The first step of the qualitative analysis focused on identifying users’ narratives surrounding the Shoah. The analysis yielded five categories: (1) political instrumentalization, (2) comparison, (3) minimization/dismissal, (4) denial/hate, and (5) historical reference. The distribution of categories calculated for each post is shown in Fig. 4.

Political instrumentalization was the most frequent category across the posts, ranging from 25% to 50%. Comparison was the second most consistent category, ranging from 12% to 43%. Minimization/dismissal ranged from 10% to 25%. Denial/hate

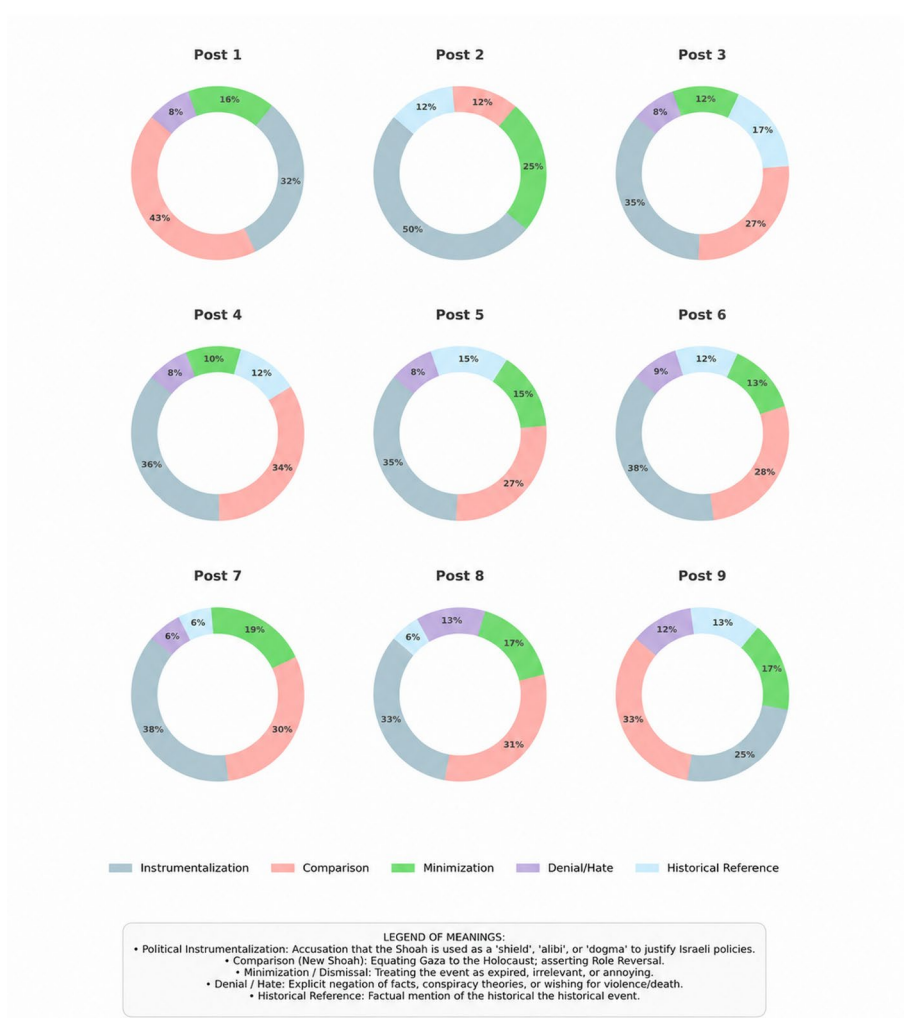


Fig. 4 Distribution of Shoah-related narratives across the nine Facebook posts

ranged from 0% to 13%, while historical reference ranged from 0% to 17%. These percentages refer to the distribution of Shoah-related narrative categories within each post and should therefore be read as per-post distributions rather than as percentages of the entire corpus.

The second step of the qualitative analysis used the eight categories introduced in the methodology section. Some comments fell into more than one category, so Gemini calculated the number of occurrences rather than assigning a single classification to each comment. As shown in Table 1, the most frequent category is Holocaust distortion, with 1,194 occurrences. This was followed by “Jews have not learned from the past”, Nazi analogy and instrumentalization. The other categories appear less frequently.

Table 1 LLM-assisted classification of antisemitic narratives and representative examples

Category	Occurrences	Narrative pattern	Example
Holocaust distortion	1,194	Equating the Holocaust with scenarios of violence/conflict for political aims	"In Gaza, Israelis are deliberately starving Palestinians. Like Auschwitz."
Jews have not learned from the past	695	Suggestion that the lesson has NOT been learned	"It happened, but it was forgotten by those who suffered it."
Nazi analogy	606	Comparing Israeli political processes/actions with Nazi actions	"People looking for food, not soldiers! Kids used as targets by the Nazis with the Star of David."
Instrumentalization	533	Accusation that Jews/Israel exploit the Holocaust for financial or political terms	"You are erasing an entire defenseless people, and it is shameful to hide behind the Shoah."
Demands for a clean break	311	Demands to end dealing with the Holocaust	"After what Israel is doing in Gaza, don't come and talk to me about the Shoah anymore."
Affirmation of Nazis/Hitler	159	Attempts to endorse/celebrate the Holocaust	"You have done more than what you suffered. Unfortunately, they did not eliminate you 100%..."
Admonishers	70	Characterises Jews/Israelis as 'troublemakers of remembrance' for excessively keeping the memory alive	"Arrogant and historical whiners, millions of other people have died and have been strictly forgotten. Disappear."
Nazi-Jewish collaboration	60	Claiming the Holocaust was co-organised/instigated by Jews/Zionists	"In my opinion, given how things are going in the world, I am starting to think that it was not all from 'Adolf's' own bag..."

Within the dominant category, Holocaust distortion, there are comments in which the war in Gaza and the Holocaust are compared. For example, some comments refer to Gaza as the "new Holocaust," while others use Holocaust terminology to describe the situation of the Palestinians. Within the same category, there are comments that directly downplay the number of Jewish victims at Auschwitz.

The category "Jews have not learned from the past" contains comments that describe the Holocaust as a lesson for Jews and Israelis, who are expected to have learned from their suffering. For example, one comment states that calling Jews "hypocrites" is a "compliment", given what has been happening in Palestine since 1949. In many cases, this category overlaps with that of Nazi analogy. One example is when comparisons are made between Israel's actions and those of Nazi Germany.

The Nazi analogy category includes comments that compare Israel, Israeli politicians, or the Israeli military to Nazi Germany, Hitler, or other Nazis. Some comments claim that Israelis are "Nazis 2.0". Another comment states that Netanyahu had an excellent teacher, adding a reference to "the little mustache." Others use the terms "Nazis," "Hitler," or "Third Reich" to describe Israeli military actions in Gaza.

The instrumentalization category includes comments asserting that the Holocaust is used as a political tool or justification. For example, one comment describes the memory of the Holocaust as a tool to defend Israeli policy. The most explicit com-

Table 2 Manual validation and performance metrics of Gemini 2.5 Pro

Metric	Description	Value
True positives	Correctly identified positive cases	84
True negatives	Correctly identified negative cases	91
False positives	Negative cases incorrectly classified as positive	16
False negatives	Positive cases incorrectly classified as negative	9
Precision	Share of predicted positives that were correct	84.0%
Recall	Share of actual positives that were correctly identified	90.3%

ments fall under the affirmation of Nazis/Hitler category, as they directly express approval of Hitler or Nazism. One example states that the Nazis were stopped too soon. Other examples are listed in Table 1.

Manual annotation of the subset of 200 comments yielded the results shown in Table 2. Overall, manual annotation identified 93 comments as positive for antisemitism and/or Holocaust distortion and 107 as negative. From the reported metrics, accuracy and F1-score can be derived. In this case, precision and recall are outlined here because they provide a more direct interpretation of the balance between false positives and false negatives. Repeating the classification on a subset of 50 comments produced an exact-output consistency rate of 84% (42/50). In the remaining eight cases, the primary distortion category remained unchanged, while variation concerned comment extraction and/or the assignment of an additional secondary category.

The model had more difficulty with the “Holocaust distortion” category, even after the second refinement prompt. This difficulty reflected the ambiguity of some comments, which either lacked sufficient evidence of hostility or remained on the borderline between legitimate political criticism and Holocaust distortion.

6 Discussion

The findings show that the reception of Holocaust education on social media cannot be understood only in terms of visibility, reach, or engagement. Although the “Proteggi I Fatti” campaign generated a substantial volume of commenting activity, many comments did not remain within the educational frame proposed by the campaign. Users frequently redirected the discussion toward the Israeli-Palestinian conflict, especially the war in Gaza. This is consistent with the view that, in networked publics, Holocaust-related messages do not circulate as fixed content. Once they enter platformed environments, they become part of a broader ecology of meaning-making in which users reinterpret, contest, appropriate, and sometimes transform the intended message (Neiger, 2026; Neiger et al., 2023; Novis-Deutsch et al., 2023).

This ambivalence is central to digital Holocaust education. Social media campaigns can extend educational outreach beyond schools, museums, memorial organizations, and commemorative initiatives, but the same participatory affordances that

enable wider engagement also expose educational content to political reframing, contested interpretation, and distortion (González-Aguilar & Makhortykh, 2022; Manca et al., 2025; Walden, 2022). The campaign's comment space therefore became not only a site of reception, but also a site where Holocaust memory was actively negotiated and resignified, consistent with recent accounts of the mutable and context-dependent character of digital Holocaust memory (Jumle & Makhortykh, 2026). For this reason, engagement metrics should be interpreted cautiously: high visibility may indicate successful dissemination, but it does not necessarily correspond to educational uptake or historically responsible engagement.

The lexical, semantic, and LLM-assisted analyses point to a broader process of conflictual reframing. The recurrence of terms related to Israel, Gaza, Palestine, genocide, Hamas, and war indicates that users often interpreted the campaign's educational messages through the language of contemporary conflict. This does not mean that all references to current events or all comparative uses of Holocaust memory are necessarily distortive (Boyd & Merrill, 2026; Novis-Deutsch et al., 2023). However, the recurrent redirection of the discussion from Holocaust distortion to Gaza and Israel suggests that the campaign's intended focus was frequently redirected toward a highly polarized interpretive frame. This finding resonates with recent research on the online circulation of genocide accusations, competitive martyrology, and Holocaust-related analogies in debates on Israel and Gaza (Bolton, 2024; CDEC, 2026; Chapelan & Becker, 2024).

The semantic associations surrounding the target word "Shoah" further clarify this process. The proximity of "Shoah" to terms such as "Gaza", "genocide", "Zionists", "shield", "alibi", and "justification" points to a narrative in which Holocaust memory is represented not only as remembrance or historical knowledge, but also as a political resource allegedly used to legitimize present actions. This is significant because Holocaust distortion does not always take the form of explicit denial of the genocide of European Jews. More often, it emerges through analogy, equivalence, relativization, instrumentalization, and victim-perpetrator reversal (Becker et al., 2024; Gerstenfeld, 2007; Weitzman, 2023). In such cases, the Shoah is acknowledged, but its historical specificity is weakened or redirected toward contemporary political accusation.

The classification based on the Decoding Antisemitism categories provides further evidence of the forms through which distortion emerged. The prominence of categories such as Holocaust distortion, Nazi analogy, instrumentalization, and "Jews have not learned from the past" indicates that many comments drew on recognizable antisemitic and distortive repertoires (Becker et al., 2024). These findings are consistent with research showing that accusations of Israeli genocide, Nazi comparisons, and Holocaust-related analogies frequently intersect in online representations and debates about Israel and Gaza (Bolton, 2024; Giesel, 2021; Grabowski, 2026; Klaff, 2023). In the present corpus, these patterns were not external to the campaign's reception; they became part of the way users engaged with, redirected, or challenged its educational messages.

These findings have implications for how Holocaust distortion should be addressed educationally. Countering distortion requires more than correcting false claims or providing historically accurate information. Factual accuracy remains essential, but

it is not sufficient when distortion operates through moral comparison, affective framing, symbolic inversion, and the political appropriation of memory. Digital Holocaust education should therefore help users understand not only what the Holocaust was, but also how Holocaust memory can be misused, instrumentalized, or weaponized in public debate (Becker et al., 2026; Walden, 2022). This includes developing the ability to distinguish historically responsible comparison from analogies that collapse historical specificity, trivialize the genocide of European Jews, or reverse the positions of victim and perpetrator (Eargle & Rubin, 2025).

The study also highlights the importance of treating comments as part of the educational environment rather than as external or secondary reactions. In platformed spaces, the meaning of a campaign is shaped not only by the original posts, but also by the responses, disputes, and interpretations that accumulate around them. For educators, institutions, and campaign designers, this implies that digital interventions should anticipate contested reception and include strategies for contextualization, moderation, and follow-up engagement. Campaigns against Holocaust distortion may need not only to disseminate accurate content, but also to prepare for the ways in which such content can be reframed in highly polarized comment spaces.

Methodologically, the study illustrates the potential value of combining computational and interpretive approaches. Lexical analysis, bigrams, sentiment analysis, and word embeddings made it possible to identify broad patterns across a large comment corpus, while LLM-assisted analysis helped reconstruct narrative patterns and classify more implicit forms of distortion. At the same time, the analysis confirms that AI tools cannot replace expert interpretation. Categories such as Holocaust distortion, Nazi analogy, instrumentalization, and victim-perpetrator reversal are context-dependent and require historical, linguistic, and cultural knowledge (Becker et al., 2024, 2026). Manual validation is therefore a necessary component of responsible AI-assisted analysis, especially when dealing with implicit antisemitism and ambiguous or contested uses of Holocaust memory.

Overall, the case of “Proteggi I Fatti” illustrates how digital Holocaust education can take place within a contested communicative environment. Social media can amplify educational messages, but they can also expose them to reinterpretation, politicization, and distortion. The challenge is therefore not only to make Holocaust education visible online, but to understand how its meanings change once it enters platformed publics.

7 Conclusion

This study contributes to research on digital Holocaust education and networked publics by showing that the reception of educational campaigns is shaped not only by their original messages, but also by the meanings produced in audience response spaces. In the case of “Proteggi I Fatti”, Facebook comments became a site where Holocaust memory was supported, challenged, politically reframed, and sometimes distorted. This suggests that campaigns against Holocaust distortion should be assessed not only through visibility or engagement metrics, but also through the interpretive dynamics that emerge around them.

The study has several limitations. It focuses on a single campaign, one platform, and a specific national context. The corpus consists only of public Facebook comments and replies, and therefore does not capture silent audiences, private sharing practices, or users' motivations. In addition, while LLM-assisted analysis helped identify recurring narratives and implicit forms of distortion, the interpretation of antisemitism and Holocaust distortion remains dependent on contextual and expert judgment. Moreover, the manual validation was conducted by a single human annotator, meaning that inter-annotator agreement could not be assessed. The validation results should therefore be interpreted with appropriate caution. A further limitation concerns the reproducibility of the LLM-assisted classification. Although the model was run with a temperature of 0.0 and a stability check was conducted, LLM outputs may remain stochastic. Moreover, the robustness check was limited to a subset of 50 comments and therefore cannot establish the stability of the classification across the entire corpus. A further limitation concerns the sentiment analysis, which was not independently validated on a manually annotated subset. The binary positive/negative classification may be particularly problematic for Holocaust-related discourse, where references to persecution, violence, and genocide can produce negative sentiment independently of the speaker's stance toward the campaign or the Holocaust.

Future research could compare audience responses across platforms, countries, languages, and campaign formats, and examine how moderation practices, message framing, visual design, and educator involvement affect the reception of Holocaust-related content online. Further work is also needed to refine AI-assisted methods for detecting implicit antisemitism and Holocaust distortion, especially in multilingual and politically polarized environments. This should include systematic investigation of possible LLM biases in recognizing antisemitism, particularly where antisemitic meanings are implicit, coded, or dependent on political and historical context.

From an educational perspective, these findings suggest that countering Holocaust distortion online may require attention not only to historical accuracy, but also to civic, digital, and media literacy, particularly in helping users recognize how Holocaust memory can be reframed, appropriated, or weaponized in platformed publics.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42803-026-00136-w>.

Acknowledgements This article draws on research conducted as part of the Master's thesis in the Master Degree in Digital and Public Humanities at Ca' Foscari University of Venice. The "Proteggi I Fatti" campaign analysed in this study was developed within the project "DECONSTRUCT - Deconstructing Distortion and Disinformation via Campaign and Digital Education" (CERV-2023-EQUAL, anti-racism and non-discrimination).

Author contributions P. D. conducted the study, carried out all analyses, prepared the figures and tables, and wrote the Methodology and Results sections. S. M. provided the conceptual and theoretical framework and was responsible for drafting the introductory sections and the Discussion. G. M. contributed substantially to the design of the research, provided methodological supervision, and guided the development of the analytical framework. All authors reviewed and approved the final version of the manuscript.

Funding Open access funding provided by Consiglio Nazionale Delle Ricerche (CNR) within the CRUI-CARE Agreement.

Data availability This study consist of public Facebook comments and replies collected from the comment sections of the “Proteggi I Fatti” campaign posts. The full dataset is not publicly available due to ethical and privacy considerations, as individual comments may remain traceable even after anonymization. Aggregated results, coding categories, and anonymized examples are reported in the article. Further information on the analytical procedures may be made available by the first author upon reasonable request, where compatible with ethical and legal requirements.

Declarations

Competing interests The authors declare no competing interests

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- ADL. (2025). *ADL AI index*. ADL Center for Tech & Society. <https://www.adl.org/adl-ai-index/full-report/GK>
- Assmann, A. (2010). The Holocaust—a global memory? Extensions and limits of a new memory community. In A. Assmann & S. Conrad (Eds.), *Memory in a global age: Discourses, practices and trajectories* (pp. 97–117). Palgrave Macmillan. https://doi.org/10.1057/9780230283367_6
- Becker, M. J., Troschke, H., Bolton, M., & Chapelan, A. (2024). *Decoding antisemitism: A guide to identifying antisemitism online*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-49238-9>
- Becker, M. J., Blatter, J., & Stanevich, O. (2026). Decoding antisemitism online: Linguistic and multi-modal challenges in the age of AI. *Frontiers in Communication*, 10, 1729279. <https://doi.org/10.3389/fcomm.2025.1729279>
- Bianchi, F., Nozza, D., & Hovy, D. (2021). FEEL-IT: Emotion and sentiment classification for the Italian language. In *Proceedings of the eleventh workshop on computational approaches to subjectivity, sentiment and social media analysis* (pp. 76–83). <https://aclanthology.org/2021.wassa-1.8>
- Bolton, M. (2024). More like genocide: The use of the concept of genocide in UK online debates about Israel. In M. J. Becker, L. Ascone, K. Placzynka, & C. Vincent (Eds.), *Antisemitism in online communication: Transdisciplinary approaches to hate speech in the twenty-first century* (pp. 107–135). Open Book. <https://doi.org/10.11647/OBP.0406.04>
- Boyd, E., & Merrill, S. (2026). The digital multidirectionality of anniversaries: 7 October 1944/2023, Catch-22 commemoration and the comparative instrumentalization of Holocaust memory on X. *Memory Mind & Media*, 5(e9), 1–25. <https://doi.org/10.1017/mem.2026.10032>
- Centro di Documentazione Ebraica Contemporanea (2026). *Annual report on antisemitism in Italy 2025*. <https://www.osservatorioantisemitismo.it/wp-content/uploads/2026/03/Annual-report-on-antisemitism-in-Italy-2025.pdf>
- Chapelan, A., & Becker, M. J. (2024, June). *Israel, Gaza and the Holocaust: The growing online discourse of competitive martyrology*. Fathom. <https://fathomjournal.org/israel-gaza-and-the-holocaust-the-growing-online-discourse-of-competitive-martyrology/>
- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8440–8451). <https://doi.org/10.18653/v1/2020.acl-main.747>

- DAIR.AI (n.d.-a). *Zero-shot prompting*. *Prompt Engineering Guide*. <https://www.promptingguide.ai/techniques/zeroshot>
- DAIR.AI (n.d.-b). *Few-shot prompting*. *Prompt Engineering Guide*. <https://www.promptingguide.ai/techniques/fewshot>
- Eargle, J. C., & Rubin, D. I. (2025). Countering Holocaust trivialization: How educational resources support teachers in the context of political extremism. *Journal of Social Studies Research*, 49(1), 3–23. <https://doi.org/10.1177/23522798241292635>
- European Commission, Directorate-General for Research and Innovation. (2021). *Ethics in social science and humanities*. European Commission.
- Fogu, C. (2006). Italiani brava gente: The legacy of fascist historical culture on Italian politics of memory. In R. N. Lebow, W. Kansteiner, & C. Fogu (Eds.), *The politics of memory in postwar Europe* (pp. 147–176). Duke University Press. <https://doi.org/10.1515/9780822388333-006>
- Foot, J. (2009). *Italy's divided memory*. Palgrave Macmillan.
- Gerstenfeld, M. (2007). The multiple distortions of Holocaust memory. *Jewish Political Studies Review*, 19(3/4), 35–55.
- Giesel, L. (2021). Comparisons between Israel and Nazi Germany in contemporary German discourse. In A. Lange, K. Mayerhofer, D. Porat, & L. Schiffman (Eds.), *Comprehending antisemitism through the ages: A historical perspective*. De Gruyter. <https://doi.org/10.1515/9783110671995-022>
- González-Aguilar, J. M., & Makhortykh, M. (2022). Laughing to forget or to remember? Anne Frank memes and mediatization of Holocaust memory. *Media Culture & Society*, 44(7), 1307–1329. <https://doi.org/10.1177/01634437221088951>
- Google (2025). *Gemini API models*. Google AI for Developers. <https://ai.google.dev/gemini-api/docs/models>
- Gordon, R. S. C. (2012). *The Holocaust in Italian culture, 1944–2010*. Stanford University Press.
- Grabowski, D. (2026). From postmemory to posts: Hirsch's concept in the digital age and the case of the Gaza 2023 war on Instagram. *Heritage Memory and Conflict*, 1(1), 210–226. <https://doi.org/10.5117/HMC2026.1.011.GRAB>
- Grabowski, J., & Klein, S. (2023). Wikipedia's intentional distortion of the history of the Holocaust. *The Journal of Holocaust Research*, 37(2), 133–190. <https://doi.org/10.1080/25785648.2023.2168939>
- Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). spaCy: Industrial-strength natural language processing in Python. *Zenodo*. <https://doi.org/10.5281/zenodo.1212303>
- IBM (n.d.-a). *The basics of prompt engineering*. IBM watsonx Prompt Lab. <https://ibm.github.io/watsonx-prompt-lab/lab-1/>
- IBM. (n.d.-b). What is ground truth in machine learning?, & Think, I. B. M. <https://www.ibm.com/think/topics/ground-truth>
- IHRA (2021). *Recognizing and countering Holocaust distortion: Recommendations for policy and decision makers*. <https://holocaustremembrance.com/resources/countering-holocaust-distortion-policy-recommendations>
- Jumle, V., & Makhortykh, M. (2026). Protean memorialisation: International Holocaust Remembrance Day as a digital site of memory on Twitter. *Politics and Governance*, 14, 11830. <https://doi.org/10.17645/pag.11830>
- Kahn, R. A. (2011). Holocaust denial and hate speech. In L. Hennebel, & T. Hochmann (Eds.), *Genocide denials and the law* (pp. 77–108). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199738922.003.0004>
- Klafl, L. (2023). Holocaust inversion and contemporary antisemitism. In A. Johnson (Ed.), *Mapping the new left antisemitism: The Fathom essays* (1st ed., pp. 141–144). Routledge. <https://doi.org/10.4324/9781003322320-20>
- Levis Sullam, S., Minello, G., Tripodi, R., & Warglien, M. (2022). Representation of Jews and anti-Jewish bias in 19th century French public discourse: Distant and close reading. *Frontiers in Big Data*, 4, 723043. <https://doi.org/10.3389/fdata.2021.723043>
- Levy, D., & Sznajder, N. (2002). Memory unbound: The Holocaust and the formation of cosmopolitan memory. *European Journal of Social Theory*, 5(1), 87–106. <https://doi.org/10.1177/1368431002005001002>
- Lipstadt, D. E. (2019). The three-legged antisemitic stool of Holocaust denial: Illogic, wilful distortions, and camouflaged discourse. In W. W. Beorn, T. Pegelow Kaplan, & G. L. Weinberg (Eds.), *Beyond ordinary men: Christopher R. Browning and Holocaust historiography* (pp. 190–201). Brill Schöningh. https://doi.org/10.30965/9783657792665_016
- Manca, S. (in press). Countering Holocaust distortion in Italy: Education in the age of social media and AI. *Digital Hate Review*.

- Manca, S., & Guetta, S. (2023). Storia e memoria per contrastare la distorsione della Shoah. *Essere a Scuola*, 1/2023, 11–15.
- Manca, S., Marino, F., & Taibi, D. (2025). “Could we possibly see your tattoo? If not that’s totally fine!” Holocaust survivors’ playful activism on TikTok. *Memory Studies*, 18(1), 237–271. <https://doi.org/10.1177/17506980241247266>
- Milvus (2025). *How does vector normalization affect embeddings?*<https://milvus.io/ai-quick-reference/how-does-vector-normalization-affect-embeddings>
- Neiger, M. (2026). New media memory: The evolution of Holocaust remembrance across media platforms. In J. Baumele-Schwartz, & M. Herskovitz (Eds.), *Creating Holocaust memory: Reworking trauma for future generations*. Routledge. <https://doi.org/10.4324/9781003601920-11>
- Neiger, M., Meyers, O., & Ben-David, A. (2023). Tweeting the Holocaust: Social media discourse between reverence, exploitation, and simulacra. *Journal of Communication*, 73(3), 222–234. <https://doi.org/10.1093/joc/jqad010>
- Novis-Deutsch, N. S., Lederman, S., Adams, T., & Kochavi, A. J. (2023). *Sites of tension: Shifts in Holocaust memory in relation to antisemitism and political contestation in Europe*. The Weiss-Livnat International Center for Holocaust Research and Education. <https://archive.jpr.org.uk/object-3605>
- Parnell, M., & Stuckey, M. E. (2023). Holocaust distortion during the global pandemic: An exercise in anti-democratic demagoguery. *Javnost - The Public*, 30(1), 128–144. <https://doi.org/10.1080/13183222.2023.2168959>
- Řehůřek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 45–50).
- Rich, J., & Dack, M. (2022). Forum: The Holocaust in virtual reality: Ethics and possibilities. *The Journal of Holocaust Research*, 36(2–3), 201–211. <https://doi.org/10.1080/25785648.2022.2062920>
- Rothberg, M. (2009). *Multidirectional memory: Remembering the Holocaust in the age of decolonization*. Stanford University Press.
- Sarfatti, M. (2006). *The Jews in Mussolini’s Italy: From equality to persecution*. University of Wisconsin.
- Sarfatti, M. (2017). Notes and reflections on the Italian law instituting Remembrance Day: History, remembrance and the present. *Quest Issues in Contemporary Jewish History*, 12, 112–134.
- Schwarz, G. (2021). Il 27 gennaio e le aporie della memoria. *Italia contemporanea*, 296, 100–123. <https://doi.org/10.3280/IC2021-296005>
- Sierp, A. (2012). Italy’s struggle with history and the Europeanisation of national memory. In U. Engel, M. Middell, & S. Troebst (Eds.), *Erinnerungskulturen in transnationaler Perspektive* (pp. 212–234). Leipziger Universitätsverlag.
- Sierp, A. (2014). *History, memory, and trans-European identity: Unifying divisions*. Routledge. <https://doi.org/10.4324/9781315766973>
- Stevens, G., O’Donnell, V. L., & Williams, L. (2015). Public domain or private data? Developing an ethical approach to social media research in an interdisciplinary project. *Educational Research and Evaluation*, 21(2), 154–167. <https://doi.org/10.1080/13803611.2015.1024010>
- Subotić, J. (2023). Holocaust memory and political legitimacy in contemporary Europe. *Holocaust Studies*, 29(4), 502–519. <https://doi.org/10.1080/17504902.2022.2116539>
- Topor, L. (2025). Terrorism denial: 7/10 and the reasons to deny a pogrom. *Israel Affairs*, 31(5), 716–734. <https://doi.org/10.1080/13537121.2025.2516864>
- Tripodi, R., Warglien, M., Sullam, L. S., & Paci, D. (2019). Tracing antisemitic language through diachronic embedding projections: France 1789–1914. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change* (pp. 115–125). <https://doi.org/10.18653/v1/W19-4715>
- Trzoss, A., Werner, W., Kleist, C., Kwiatkowska-Moskalewicz, K., & Moskalewicz, M. (2023). The politicisation of historical memory on Twitter: Positive antisemitism in the Holocaust debate in Poland. *Rethinking History*, 27(4), 780–810. <https://doi.org/10.1080/13642529.2023.2248451>
- Ulloa, R., Zucker, E. M., Bultmann, D., Simon, D. J., & Makhortykh, M. (2026). From prosthetic memory to prosthetic denial: Auditing whether large language models are prone to mass atrocity denialism. *AI & Society*, 41, 3889–3903. <https://doi.org/10.1007/s00146-025-02719-7>
- USC Shoah Foundation (n.d). Deconstruct: Deconstructing and countering holocaust distortion via campaign, education and training. *IIWitness*. Retrieved from <https://iwwitness.usc.edu/sites/deconstruct?at=ln>

- Vziatysheva, V., & Makhortykh, M. (2024). Take (no) chances: How prompt formulation and stochasticity affect the accuracy of LLMs for fact-checking. In *EKAU 2024: EKAU 2024 Workshops, Tutorials, Posters and Demos, 24th International Conference on Knowledge Engineering and Knowledge Management*.
- Walden, V. G. (2021). *Digital Holocaust memory, education and research*. Palgrave Macmillan.
- Walden, V. G. (2022). Understanding Holocaust memory and education in the digital age: Before and after Covid-19. *Holocaust Studies*, 28(3), 257–278. <https://doi.org/10.1080/17504902.2021.1979175>
- Weitzman, M. (2023). Holocaust denial and distortion. In M. Weitzman, R. J. Williams, & J. Wald (Eds.), *The Routledge history of antisemitism* (1st ed., pp. 373–382). Routledge. <https://doi.org/10.4324/9780429428616-41>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., & Rush, A. M. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 38–45). <https://doi.org/10.18653/v1/2020.emnlp-demos.6>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Priscilla Dionesalvi¹ · Stefania Manca³  · Giorgia Minello² 

✉ Stefania Manca
stefania.manca@cnr.it

Priscilla Dionesalvi
priscilladionesalvi@gmail.com

Giorgia Minello
giorgia.minello@essex.ac.uk

¹ Department of Humanities, Ca' Foscari University of Venice, Dorsoduro, Venice 3246, 30123, Italy

² School of Computer Science and Electronic Engineering (CSEE), University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK

³ Institute of Educational Technology, Italian National Research Council, Via De Marini 6, Genova 16151, Italy