Beyond Participation Indicators: Using Data-Driven Storytelling to Reveal Contested Audience Reception

Priscilla Dionesalvi[1]

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

Abstract

This study examines the responses of the audience to the Italian Facebook page "Proteggi i Fatti," developed within the European project *Deconstruct: Deconstructing and Countering Holocaust Distortion through Campaign, Education and Training*. The corpus consists of 10,160 public comments collected from the comment sections of nine posts published between December 2024 and May 2025. Its objective is to study how educational messages are received and interpreted by users using a data-driven storytelling approach. The methodology adopts a multi-layered analytical approach that combines indicators of participation, sentiment and toxicity analysis, lexical frequencies, semantic associations, and AI-assisted narrative. The results show that engagement and aggregated indicators can describe the intensity and tone of the public's response, but do not fully explain how the original message is interpreted, re-framed and transformed. The integration of the findings reveals recurring forms of contestation and counter-narration within visible participation.

Keywords

Holocaust Distortion, Data-driven Storytelling, Large Language Models,

1. Introduction

Due to social media, content related to the Holocaust can reach larger audiences through forms of remembrance activism and educational campaigns (Manca et al., 2025; Neiger, 2026). However, such content is also vulnerable to trivialization, distortion, instrumentalization, and appropriation when placed in digital environments (González-Aguilar & Makhortykh, 2022; Parnell & Stuckey, 2023). Platforms are not neutral channels of transmission. Even when a message is presented within an educational or commemorative context, users can reinterpret it for different purposes (Ben-David et al., 2024).

This dynamic is relevant to the distortion of the Holocaust, which manifests itself through minimization, relativization, Nazi analogies, and the reversal of the roles of victims and perpetrators (Becker et al., 2024). For this reason, it is challenging to study how users reinterpret and transform the meanings of such content.

This study examines public responses to the Italian Facebook page "Proteggi i Fatti," part of the European project *Deconstruct: Deconstructing and Countering Holocaust Distortion through Campaign, Education, and Training*. The corpus consists of 10,160 comments collected from the comment sections of nine posts published between December 2024 and May 2025. The purpose of this study is to discover the narratives generated by the users themselves while focusing on how they react to and reinterpret the original messages of the posts. The analysis integrates several levels: participation metrics, lexical frequencies, sentiment and toxicity analysis, semantic associations, and AI-assisted narrative analysis.

The approach adopted is based on data-driven storytelling, which allows isolated findings to be linked into an interpretive narrative (Zhao & Elmqvist, 2023). In this process, artificial intelligence can help identify patterns and narratives, while qualitative interpretation remains the responsibility of the researcher (Li et al., 2024).

The research question guiding this study is: what can a data story reveal about audience reception that participation metrics, quantitative, and AI-assisted analyses fail to capture when considered in

Data-Driven Storytelling: Bridging Knowledge Graphs, GenAI, and Narrative (DDS2026), in ISWC 2026 Workshops Joint Proceedings, October 25–26, 2026, Bari, Italy

priscilladionesalvi@gmail.com (P. Dionesalvi)



© 2026 Copyright © 2026 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-11-05

CEUR-WS.org/Vol-4313/dds-paper2.pdf

isolation?

To address this question, the study implements these different levels of analysis to construct a data story about the contested reception of an educational campaign on the Holocaust, aimed at an Italian audience and developed in the political context following October 7, 2023. The computational pipeline begins with data cleaning and extraction of lexical frequency, assesses tone and toxicity, explores semantic associations, and concludes with an AI-assisted qualitative classification employing Gemini 2.5 Pro.

The contribution lies in showing how meaningful participation can also involve processes of challenging and transforming the original message.

2. Background: Data-Driven Storytelling and Contested Holocaust Memory

Zhao and Elmqvist (2023) define data-driven storytelling as the use of narrative techniques to explain complex phenomena through a data-based interpretive framework (Zhao & Elmqvist, 2023). In the fields of digital humanities and cultural heritage studies, historical and cultural data can be characterized by uncertainty and multiple interpretations. In such cases, visualization-based storytelling allows the integration of visual representation and narration to connect different levels of analysis without reducing the complexity of the phenomena under study (Kusnick et al., 2026).

Artificial intelligence is increasingly contributing to the various stages of storytelling, including identifying relevant elements, organizing results, and generating visualizations. However, human input is still necessary to critically evaluate the AI's results (Li et al., 2024). These considerations are relevant to the present study, which uses AI-assisted analysis techniques to support narrative analysis, while the interpretation of the results is guided by human intervention.

This perspective becomes crucial in the study of Holocaust remembrance online. On digital platforms, even when commemorative content is shared by institutional entities, users can re-appropriate and transform it through their interactions. The comments section is the space where commemorative content can be interpreted in different ways and in relation to social and political events. This dynamic is defined as "directionality" and refers to the gap between the meaning attributed by those who post the content and the meaning constructed by those who respond (Ben-David et al., 2024).

When it comes to the memory of the Holocaust, previous studies have shown that the distortion of the Holocaust is often an indirect phenomenon that is difficult to identify. Unlike explicit denial, distortion can occur through minimization, relativization, accusations of exploiting memory, analogies with Nazism, decontextualized comparisons with other historical events, and the reversal of the roles of victims and perpetrators (Becker et al., 2024; Giesel, 2021; Weitzman, 2023). These dynamics intensify significantly on social networks, where such forms circulate more easily and where the memory of the Holocaust is continuously mobilized to interpret contemporary issues (Neiger et al., 2023; Walden, 2021). Furthermore, since these forms do not necessarily require the use of overtly offensive or violent language, understanding them requires qualitative interpretation. As highlighted by Kusnick et al. (2026), the researcher's role is crucial to connecting data and findings and building a narrative that guides the reader through a complex phenomenon such as the distortion of the Holocaust.

The political context in which users produce such discourse, that is the period following the attack on October 7, 2023, plays a central role. Following the Hamas attack and the escalation of the war in Gaza, the memory of the Holocaust has often been invoked in online discussions on the Israeli-Palestinian conflict. Boyd and Merrill (2026) show how commemoration, solidarity, protest, and polarization often overlap in digital spaces. To understand when a comparison between the Holocaust and other events can be considered a form of distortion, it is important to assess whether such a comparison flattens historical differences through inversions or decontextualized equivalences (Boyd & Merrill, 2026).

Against this backdrop, this study integrates different levels of analysis to construct a data story of the campaign's reception. Participation indicators show the volume of response, lexical and semantic frequencies reveal shifts in the discourse themes, while sentiment and toxicity metrics indicate the

characteristics of the tone. AI-assisted analysis helps to identify the most complex narratives. The combination of these levels with human interpretation allows the construction of a data story consistent with forms of memory mobilization already documented in the literature.

3. Data and Methods

3.1. Data Collection and Corpus

This study focuses on the Italian Facebook page "Proteggi i Fatti." The page is part of the European project *Deconstruct: Deconstructing and Countering Holocaust Distortion through Campaign, Education, and Training*, whose goal is to provide educational content, online learning activities, and critical thinking tools to combat Holocaust distortion and denial (USC Shoah Foundation, n.d.). The campaign ran from December 2024 to May 2025. During this period, nine posts with identical structure were published: a brief description introducing the content, one or more images with a neutral background, and a short text. The topics addressed include Holocaust denial and distortion, the role of educators, and the importance of verifying the accuracy of facts before posting online. Starting in February 2025, an advertising campaign was organized to increase the visibility of the content. As an example, fig. 1 represents the content of the ninth post.



Figure 1: Post Example

A total of 10,160 comments were collected from the comment sections of the nine posts and included in the corpus analyzed. The comments were collected through the Facebook Meta Graph API, from the publication date of each post until September 2025.

Data minimization and anonymization procedures were implemented to exclude users' personal information. Specifically, nicknames, personal names, tags and URLs were removed. The analysis did not focus on individual users, but rather on the narratives emerging from the comments, which have been anonymized and slightly paraphrased or translated to prevent them from being traceable. The study followed established guidelines for social media research (European Commission, Directorate-General for Research and Innovation, 2021; Stevens et al., 2015).

3.2. Multi-layer Analytical Approach

The analysis was organized into complementary analytical levels, starting with indicators of user engagement with the campaign and then progressively moving toward more interpretive analyses. This

methodological framework follows the data-driven storytelling approach.

Facebook Insights provided data on engagement indicators, such as: the number of followers, page visits, likes, unique users reached, paid and organic reach, page visits, and content interactions, namely reactions, comments, shares, and clicks.

Lexical analysis was conducted using Natural Language Processing (NLP) techniques. Using the Italian version of spaCy, it was possible to clean and tokenize the comments (Honnibal et al., 2020). In the cleaning phase, the text was converted to lowercase; URLs, special characters, and stopwords were removed. Next, to identify the main themes in the user-generated discussions, the ten most frequent words and bi-grams were extracted.

Sentiment analysis was conducted to identify the tone of the comments. Sentiment was classified as positive or negative using the Italian MilaNLP/feel-it-italian-sentiment model (Bianchi et al., 2021), through the Hugging Face Transformers library (Wolf et al., 2020). Toxicity was calculated using the multilingual Detoxify model, based on XLM-RoBERTa (Conneau et al., 2020). Scores ranging from 0 to 1 were grouped into five categories: very low, low, moderate, high, and very high.

Subsequently, another level of analysis focused on the semantic associations surrounding the term "Shoah," using the Gensim Word2Vec model (Řehůřek & Sojka, 2010). The 20 terms most closely related to the target word were identified using cosine similarity and then projected using Principal Component Analysis (PCA) to visually observe the semantic relationships.

The AI-assisted analyses were conducted using Gemini 2.5 Pro through the Gemini API, in the final layer. The model was run at a temperature of 0.0 to reduce variability. In the first step, a zero-shot prompt approach was used in comments that explicitly mentioned the term "Shoah." The model was asked to identify the dominant narratives associated with the Shoah without being provided with predefined categories. The second step followed a few-shot approach to identify possible forms of distortion and antisemitism using eight predefined categories derived from the "Decoding Antisemitism" lexicon (Becker et al., 2024): demands for a clean break, Holocaust distortion, instrumentalization, Nazi–Jewish collaboration, "Jews have not learned from the past," admonishers, affirmation of Nazis or Hitler, and Nazi analogy. These categories were chosen because they align with the IHRA definition of Holocaust distortion (IHRA, 2021). Then a refinement prompt was used to review the initial output. All the prompts used are provided in appendix A.

The author manually evaluated the final classification using two complementary samples of 100 comments each. The first sample contained comments classified by Gemini as antisemitic or as containing Holocaust distortion, while the second contained comments randomly selected from the entire corpus. This phase served to evaluate false positives and false negatives and to determine which categories were most difficult for Gemini to distinguish.

Among the 100 comments classified by Gemini as problematic, 84 were confirmed as such through manual review, while 16 were identified as false positives. In the random sample, 91 comments were confirmed as true negatives and 9 as false negatives. Considering both samples, precision is 84%, recall is 90.3%, and accuracy is 87.5%. However, these values represent only a small portion of the corpus and it is important to recognize the limitation of having a single annotator to evaluate the outputs.

4. Findings

The analyses described in the previous section yielded different levels of evidence on public reception. The first level concerns campaign participation indicators, which provide information on visibility and the volume of interactions with the content. The quantitative phase focused on identifying lexical frequencies, sentiment, toxicity, and semantic proximity. The AI-assisted qualitative phase was necessary to identify implicit forms of distorted narratives that quantitative analyses alone were unable to detect. The results are presented in accordance with the methodology adopted to construct a multi-layered data story. Tab.1 summarizes the different levels of analysis, explaining how each contributes to the development of the data story.

Table 1
Analytical layers, their contribution to the data story, and main findings.

Analytical layer	Contribution to the data story	Summary of findings
Campaign visibility and participation	Provides information on visibility and the volume of interactions with the content	The campaign reached 1,088 followers, 1,089 likes, while the published content generated 355,409 views and reached 99,893 unique users.
Lexical patterns	Identifies the dominant topics and the thematic direction of the discussion.	The most frequent terms include: *Israele* (1,385), *palestinesi* (1,222), *Gaza* (1,099), *genocidio* (938), *Hamas* (911), and *Shoah* (770). Frequent bi-grams include: *genocidio Gaza* and *pulizia etnica*.
Sentiment and toxicity	Describes the tone and the presence of overt verbal aggression.	89.8% are classified as negative. However, "very low" toxicity was the largest category across all posts, ranging from 57.1% to 69.6%.
Semantic proximity	Explores the discursive contexts in which the term *Shoah* is used.	*Shoah* is associated both with memory-related terms, such as *memoria*, *sterminati*, and *superstiti*, and with terms connected to the contemporary conflict, including *Gaza*, *genocidio*, *sionisti*, *alibi*, and *giustificazione*.
AI-assisted narrative analysis	Identifies and organizes recurring interpretive frames and potential forms of Holocaust distortion.	Political instrumentalization is the most frequent inductive category, followed by comparison and minimization. The most frequent antisemitic categories are Holocaust distortion (1,194 occurrences), "Jews have not learned from the past" (695), Nazi analogy (606), and instrumentalization (533).

4.1. Campaign Participation Indicators

The data collected through Facebook Insights during the advertising campaign period show that the page reached 1,088 followers and 1,089 likes, while the published content generated 355,409 views and reached 99,893 unique users. These metrics provide a numerical measure of the campaign's visibility and reach but do not indicate how the content was received by users.

4.2. Quantitative Findings

An analysis of lexical frequencies shows that the most frequent terms are primarily related to the Israeli-Palestinian conflict. The ten most frequent words 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), palestina (Palestine, 531), israeliani (Israelis, 527), and bambini (children, 417).

The bi-gram analysis specifies the usage of these words. The ten most frequent bi-grams 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 cisgiordania (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 performed on 9,619 comments, while the rest was excluded due to a lack of analyzable content. The model classified 8,638 comments (89.8%) as negative and 981 comments (10.2%) as positive. There is no neutral category, as this is a binary classification.

The analysis of toxicity also provided insight into the general tone of the discussion. Most of the comments in all nine posts fall into the category of "very low" toxicity, with percentages ranging from 57.1% to 69.6%. The "very high" category is the second largest, ranging from 9.3% to 18.5%. The remaining categories show lower values: "low" toxicity ranges from 6.9% to 10.1%, "moderate" toxicity from 4.5% to 8.1%, and "high" toxicity from 4.0% to 7.0%. A limitation of the Detoxify model is its sensitivity to explicit insults, profanity, and other forms of vulgar language; therefore, the scores are indicators of overt aggression and not of antisemitism or Holocaust denial.

The semantic proximity analysis conducted on the nine posts shows that the target term "Shoah"

tends to fall between two lexical fields: one associated with terms related to historical memory, such as “memory,” “exterminated,” “survivors,” and “tragedy”; the other associated with terms related to the Israeli-Palestinian conflict, such as “Gaza,” “Zionists,” “settlers,” “genocide,” “alibi,” “shield,” and “justification.” This analysis reveals the discursive contexts in which the Shoah is situated within user discussions.

4.3. AI-Assisted Qualitative Findings

The first phase of the AI-assisted analysis identified five recurring narratives associated with the Holocaust: political instrumentalization, comparison, minimization, denial, and historical reference. The most frequent category is political instrumentalization, ranging from 25% to 50% in the nine posts. Comparisons ranged from 12% to 43%, while minimization ranged from 10% to 25%. The denial category ranged from 0% to 13%, while historical references ranged from 0% to 17%. The percentages refer to the distribution of categories among comments that directly mention the Holocaust and across the nine posts. These percentages do not describe the entire corpus.

In the second phase, by applying the categories from the “Decoding Antisemitism” lexicon, it was possible to classify the comments according to possible forms of Holocaust distortion and antisemitism. Some comments were assigned to multiple categories because they contained more than one narrative. Holocaust distortion is the most frequent category, with 1,194 occurrences, followed by “Jews have not learned from the past,” Nazi analogy, and instrumentalization. The categories demand for a clean break, Nazi-Jewish collaboration, admonishers, and affirmation of Nazis or Hitler were less frequent.

The Holocaust distortion category includes comments that compare the Shoah with the war in Gaza, apply terminology from the past to the present context, or downplay certain aspects of the persecution of Jews. The “Jews have not learned from the past” category includes comments that portray the Holocaust as a moral lesson that Jews or Israelis should have learned from. This narrative often overlaps with Nazi analogy, in which Israel and its political representatives are compared to Nazi Germany.

Comments classified as instrumentalization argue that the memory of the Holocaust is being used as a tool to legitimize Israeli policies. A smaller number of comments explicitly express approval of Hitler or Nazism. The model encountered greater difficulty in classifying the Holocaust Distortion category, as some comments were situated between political criticism, inappropriate historical comparison, and outright distortion. Tab.2 provides examples of the antisemitic narratives identified in the dataset to show how users reinterpreted the educational messages in the posts.

Table 2
Examples of recurring antisemitic narratives.

Category	Representative Example (Paraphrased)
Holocaust distortion	“The Gaza Holocaust is, in fact, being manipulated and distorted a bit too much by the Zionists”
Jews have not learned from the past	”Never doubted the Holocaust, it’s amazing that Israel is the one applying it to others...”
Nazi analogy	“I agree, the Netanyahu government isn’t Jewish, but rather Khazar-Ashkenazi Zionist, worse than in 1943–45.”
Instrumentalization	”Let’s start with isr@el, who uses and manipulates it as he pleases, tarnishing its reputation...”
Demand for a clean break	“People always talk about the Holocaust far in the past and not about what’s happening right now in Gaza... to prevent the genocide from continuing!!!!”
Nazi-Jewish collaboration	“History is always written by the victors. Who knows, maybe Hitler was just a pawn... to establish Israel right after the war and put Europe and the world at the feet of American Jews.”
Admonishers	“...too much Jewishness will give rise to antisemitism, and that’s exactly what happened... Now go ahead and keep busting our balls with this nonsense, and you’ll see that sooner or later someone will fire up the ovens again... but not to bake bread!!!”
Affirmation of Nazis/Hitler	“The mistake was closing the camps”

5. Discussion

The various integrated analytical tools made it possible to construct a data story that begins with engagement metrics and moves on to more interpretive dimensions. The metrics obtained from Facebook Insights show the breadth of the campaign's reach, but do not allow the interpretation of how users received the educational messages. Visibility does not necessarily equate to acceptance of the message, as greater exposure can also lead to forms of opposition.

The frequency results show that discussions frequently shift to the Israeli-Palestinian conflict, given the recurrence of terms related to Israel, Gaza, Palestine, genocide, and Hamas. This thematic shift is consistent with the concept of "directionality" described by Ben-David et al. (2024), according to whom there is a separation between the content proposed by the broadcaster and the ways in which users reinterpret it, giving it different purposes. The memory of the Holocaust, therefore, does not retain a fixed meaning but is continuously reworked by users on digital platforms and mobilized to interpret contemporary issues (Neiger et al., 2023; Walden, 2021). Platforms can enable greater dissemination of content while also exposing it to polarization and distortion (González-Aguilar & Makhortykh, 2022; Manca et al., 2025; Walden, 2022).

Sentiment and toxicity analyses add another layer to the story. Sentiment is overwhelmingly negative, but toxicity levels are generally low. This discrepancy shows that dissent does not necessarily take the form of explicit insults. In fact, forms of relativization, comparison, and moral reversal often remain invisible to tools designed to detect explicitly aggressive content.

This thematic shift is further explored through semantic proximity analysis. The term "Shoah" appears close to words related to historical memory but also to terms related to the Israeli-Palestinian conflict. This proximity indicates that the Shoah is embedded in discourses in which historical memory, contemporary conflict, and accusations of exploitation intertwine.

The first step of the Gemini analysis identified the recurring narratives surrounding the Holocaust: political instrumentalization, comparison, minimization, denial, and historical reference. These results show that references to the Holocaust serve multiple functions within discussions, such as recalling historical facts but also making judgments about the present or challenging the use of memory itself.

The final step of the Gemini analysis, based on the categories from "Decoding Antisemitism", revealed other forms through which distortion manifests itself. Several comments employed recurring antisemitic and distorting tropes, such as Nazi analogies and forms of Holocaust distortion and inversion (Becker et al., 2024). These results are consistent with patterns described in the literature. The Holocaust is not denied but may be relativized, decontextualized, or used as a tool for political accusation (Becker et al., 2024; Giesel, 2021; Weitzman, 2023).

However, the results also show how complex it is to interpret these phenomena. Not every reference to the Israeli-Palestinian conflict constitutes a form of distortion (Boyd & Merrill, 2026). In some cases, comments are situated on the fine line between hostility and legitimate criticism, appropriate historical comparison, and distortion of the Holocaust. It is precisely this ambiguity that has created difficulties for Gemini, and for this reason, human review is still necessary to evaluate context and refine outputs. AI can support automatic classification but cannot replace human work in attributing meaning to the results (Li et al., 2024).

Finally, the broader context of October 7, 2023, is a crucial element for understanding the results. Holocaust remembrance and the conflict in Gaza are closely intertwined within the communicative space where the campaign took place. Commemoration and protest often overlap in digital spaces, and the data story constructed in this study makes this intersection visible (Boyd and Merrill, 2026). Each level of analysis offers a partial interpretation of audience reception. Participation indicators describe users' interaction with the content and the page itself; frequency analysis shows shifts in topics; sentiment and toxicity analyses indicate the tone of discussions; and semantic associations identify the contexts in which the Holocaust is mentioned. The contribution of AI has made it possible to add a qualitative layer of interpretation by organizing the comments into recurring narratives. Integrating these elements made it possible to construct the narration of the contested reception of the educational campaign. The data-driven storytelling approach was fundamental to integrate separate outputs and

construct a narrative path to explain a complex phenomenon (Zhao & Elmqvist, 2023).

6. Conclusion

The study demonstrates how isolated data points can be integrated to construct a data story. In the specific case of comments collected from the "Proteggi i Fatti" page, the data-driven storytelling approach allowed one to develop a narrative that moves from visibility metrics to understanding how users interpreted the educational message.

The study focuses on a single campaign situated within a specific Italian context following October 7, 2023. It considers only public comments and does not account for silent users or private shares. AI-assisted analysis supported the classification of distorted narratives, but still has limitations that require human review.

Future research could expand this storytelling framework by integrating more structured semantic representations alongside LLMs to explore how narratives evolve across different contexts. Furthermore, the results obtained regarding the extracted distortive narratives could be used as foundational categories to build ontologies and Knowledge Graphs to track Holocaust memory online.

In general, the results suggest that digital Holocaust campaigns should be evaluated not only in terms of visibility and engagement, but also in terms of the meanings produced in the spaces where users interact and respond.

7. Acknowledgments

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 analyzed in this study was developed within the project "DECONSTRUCT - Deconstructing Distortion and disinformation through campaign and Digital Education" (CERV-2023-EQUAL, anti-racism and non-discrimination).

8. Declaration on Generative AI

During the preparation of this work, the author used ChatGPT for language editing and translation support.

9. References

Becker, M. J., Troschke, H., Bolton, M., & Chapelan, A. (Eds.). (2024). *Decoding antisemitism: A guide to identifying antisemitism online*. Palgrave Macmillan. doi:10.1007/978-3-031-49238-9.

Ben-David, A., Meyers, O., & Neiger, M. (2024). How social memory works on social media: A methodological framework. *Memory, Mind & Media*, 3, e20, 1–23. doi:10.1017/mem.2024.18.

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*, 76–83. doi:10.18653/v1/2021.wassa-1.8.

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. doi:10.1017/mem.2026.10032.

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*, 8440–8451. doi:10.18653/v1/2020.acl-main.747.

European Commission, Directorate-General for Research and Innovation. (2021). *Ethics in social science and humanities*. European Commission.

Giesel, L. (2021). Comparisons between Israel and Nazi Germany in contemporary German discourse. In A. Lange, K. Mayerhofer, D. Porat, & L. H. Schiffman (Eds.), *Comprehending antisemitism through the ages: A historical perspective* (pp. 443–464). De Gruyter. doi: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. doi:10.1177/01634437221088951.

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). *spaCy: Industrial-strength natural language processing in Python*. Zenodo. doi:10.5281/zenodo.1212303.

Kusnick, J., Andersen, N. S., Liem, J., Mayr, E., Beck, S., Koch, S., Doppler, C., Seirafi, K., Jänicke, S., & Windhager, F. (2026). A survey on visualization-based storytelling in digital humanities and cultural heritage. *Visual Informatics*, 10, 100312. doi:10.1016/j.visinf.2026.100312.

Li, H., Wang, Y., & Qu, H. (2024). Where are we so far? Understanding data storytelling tools from the perspective of human–AI collaboration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, 1–19. Association for Computing Machinery. doi:10.1145/3613904.3642726.

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. doi:10.1177/17506980241247266.

Neiger, M. (2026). New media memory: The evolution of Holocaust remembrance across media platforms. In J. Baumel-Schwartz & M. Herskovitz (Eds.), *Creating Holocaust memory: Reworking trauma for future generations*. Routledge. doi: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. doi:10.1093/joc/jqad010.

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. doi: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*, 45–50.

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. doi:10.1080/13803611.2015.1024010.

USC Shoah Foundation. (n.d.). *Deconstruct: Deconstructing and countering Holocaust distortion via campaign, education and training*. IWitness. Retrieved from \url{https://iwitness.usc.edu/sites/deconstruct?at=ln}

Walden, V. G. (2021). *Digital Holocaust memory, education and research*. Palgrave Macmillan.

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. doi: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*, 38–45. doi:10.18653/v1/2020.emnlp-demos.6.

Zhao, Z., & Elmqvist, N. (2023). The stories we tell about data: Surveying data-driven storytelling using visualization. *IEEE Computer Graphics and Applications*, 43(4), 97–110. doi:10.1109/MCG.2023.3269850.

Appendix A: Prompts Used for the AI-Assisted Analysis

This appendix reports the prompts employed during the two phases of the AI-assisted analysis.

Prompt 1

```
I will provide you with a list of comments that mention the word
"{TARGET_WORD}".

Your task is to analyze the "vision" of the Shoah that emerges from
these comments.

Do not limit yourself to counting words, but interpret the contexts
of use.

Return an analysis in JSON format with the following structure:

1. "general_vision":
   A summary paragraph describing how the Shoah is perceived and
   discussed in this corpus.

2. "prevalent_sentiment":
   The general tone of the comments.

3. "main_use_themes":
   A list of objects that describe the primary ways the topic is used.

   Each object must have:
   - "theme": The context of use.
   - "frequency": How many times this theme emerged.

4. "example_comments_with_sentiment":
   A list of 3-5 representative comments, each with its own sentiment.

   Each object must have:
   - "comment": The comment text (anonymized if necessary).
   - "sentiment": The sentiment of that specific comment.
```

Prompt 2: Classification According to the Decoding Antisemitism Taxonomy

```
You are an expert analyst of Holocaust memory and antisemitism based on
the "Decoding Antisemitism" handbook.

Analyze the following comments from a Facebook page dedicated to the
Shoah.

Identify ONLY comments that match the following 8 SPECIFIC CATEGORIES.

Use the definitions below as strict guidelines.

1. DEMANDS FOR A CLEAN BREAK
* Demands to end dealing with the Holocaust.
* Demands to end dealing with (Nazi) antisemitism.
* Demands to stop payments to victims of the Holocaust.
```

```
2. HOLOCAUST DISTORTION
* Trivialising or denying the complicity of perpetrator groups
  (e.g. Wehrmacht, collaborators) or local populations.
* Reducing the blame to a small selection of people
  (e.g. "only Hitler").
* Minimising the number of victims.
* Holding Jews responsible for the genocide.
* Equating the Holocaust with other crimes against humanity.
* Equating the Holocaust with scenarios of violence/conflict
  for political aims.

3. INSTRUMENTALISATION
* Accusation that Jews/Israel exploit the Holocaust for financial
  or political terms.
* Suggestion that antisemitism prevention laws/education serve a
  hidden agenda.
* Reduction of the Holocaust to a mere rhetorical tool devoid of
  historical significance.

4. NAZI-JEWISH COLLABORATION
* Claiming the Holocaust was co-organised/instigated by Jews/Zionists.
* Alleging Jewish councils/Ghetto Police/Sonderkommando were
  willing collaborators.
* Claiming prominent Jews or Nazi perpetrators were actually Jewish.
* Claiming the Haavara agreement was a plan for Arab extermination.
* Claiming the Holocaust was a trade-off for the establishment of Israel.
* Idea that Zionists didn't care about European Jews.

5. ADMONISHERS
* Characterises Jews/Israelis as "troublemakers of remembrance".
* Characterises Jews/Israelis as irreconcilable.

6. JEWS HAVE NOT LEARNED FROM THE PAST
* Direct/implied mention of Jewish suffering/Holocaust.
* Suggestion that this should be a lesson for Jews.
* Suggestion that the lesson has NOT been learned.
* Mentions of short memory, forgetfulness, lack of respect
  for ancestors.

7. AFFIRMATION OF NAZIS/HITLER
* Expressing positive opinion about Hitler/Nazism.
* Attempts to endorse/celebrate the Holocaust.
* Upholding the legacy of the Nazi regime/allies as inspiration.
* Jokes trivialising Jewish suffering.
* Expressing desire for Nazi ideologies to be reinstated.

8. NAZI ANALOGY
* All statements comparing Jews, Israel, Israelis or Zionism
  with Nazi Germany entities.
* Comparing Israeli political processes/actions with Nazi actions.
* Comparisons between Jews/Israel and present-day movements
  supporting Nazism.

INPUT DATA:
{text_batch}
```

```
OUTPUT FORMAT (JSON ONLY):

{
  "identified_distortions": [
    {
      "category_name": "One of the 8 categories above",
      "explanation":
      "Quote the specific bullet point from the definition
      that matches",
      "extracted_comment": "Original text"
    }
  ]
}
```

Prompt 3: Refinement and Validation

```
You are an expert reviewer verifying comments previously flagged as
distorted or antisemitic.

Your task is to analyze the batch of comments below
(format: [INDEX] CATEGORY | COMMENT) and determine, for each, if it
should be KEPT (True) or DISCARDED (False) based on the provided
rules, derived from the "Decoding Antisemitism" handbook.

BATCH OF COMMENTS TO VERIFY:
---
{comments_list_str}
---

FINAL EVALUATION RULES

RULES FOR ACCEPTANCE (TRUE POSITIVES):

1. KEEP if it explicitly compares Israel/Gaza/IDF to
   Nazis/Hitler/SS/Auschwitz or uses terms like "lager"
   to describe Gaza.

2. KEEP if it denies/minimizes the Holocaust facts or numbers
   (e.g., "only 300,000 dead").

3. KEEP if it blames Jews for the Holocaust, or claims they
   instrumentalize it for profit or political protection.

4. KEEP "Clean Break" demands
   (e.g., "Stop talking about the Shoah").

5. KEEP if it uses common and implicit tropes:
   "From victim to executioner",
   "Palestinian Shoah",
   or the analogy of the Jew acting like a Nazi.

RULES FOR REJECTION (FALSE POSITIVES):

1. REJECT if it only criticizes the Netanyahu government,
   the IDF, or war tactics without comparisons to Nazis
```

```
   or the Holocaust.

2. REJECT if it only claims that "all deaths are equal"
   or generally compares suffering.

3. REJECT if it attacks Hamas or specific Jihadist groups.

4. REJECT if it is pure political cynicism without explicit
   elements of Holocaust distortion.

RESPONSE FORMAT

Return ONLY a JSON array.

[
  {
    "original_index": 0,
    "is_valid": true,
    "reason": "Brief reason (English)"
  },
  {
    "original_index": 1,
    "is_valid": false,
    "reason": "Brief reason (English)"
  }
]
```