Discourse Report | 6

Pilot Project

Decoding Antisemitism: An AI-driven Study on Hate Speech and Imagery Online

Centre for Research on Antisemitism,
Technical University Berlin

Decoding Antisemitism

Decoding Antisemitism

Principal Investigator:
Dr Matthias J. Becker
Centre for Research on Antisemitism, Technical University (TU) Berlin

Co-Investigator:
Prof. Helena Mihaljević
Data Science, University of Applied Sciences (HTW) Berlin

Research Team TU, HTW & WZB Berlin:
Dr Laura Ascone
Dr Matthew Bolton
Alexis Chapelan
Pia Haupeltshofer
Alexa Krugel
Stefan Munnes
Karolina Placzynta
Milena Pustet
Marcus Scheiber
Victor Tschiskale

Project Manager:
Prof. Uffa Jensen
Centre for Research on Antisemitism, TU Berlin

Project Management:
Dr Laura Ascone

Secretary:
Andrea Rellin

Project Affiliates:
Veronika Bundzíková (Masaryk University, Brno, Czechia)
Iael Kurjan (Federal University of Santa Catarina, Brazil)
Mohamed Salhi (Goethe University Frankfurt, Germany)

Citation recommendation:

Table of Contents

Executive summary

1. This report examines antisemitic discourse online during an extraordinarily tumultuous time. The Hamas attacks of 7 October and the subsequent Israeli military response in Gaza generated a huge level of online discussion as well as incitement, unprecedented in the project's timeframe, even in politically moderate contexts. The report shares our findings from an analysis of over 19,500 comments posted online in the UK, France and Germany since the attacks, with an additional 6,000 comments presented in a six-country comparative study.

2. Antisemitic content posted in the first week following the attacks (7–13 October) represents a turning point in antisemitic online communication, characterised by open celebration and AFFIRMATION OF VIOLENCE inflicted on Israeli civilians by Hamas.[1] In the UK corpus, on average 27% of antisemitic comments affirmed the attacks; in the French data, it was a staggering 55%. By contrast, the German corpora prominently featured the antisemitic concept of ISRAEL'S SOLE GUILT FOR THE CONFLICT (29.6% of all antisemitic comments), pinning responsibility for the 7 October attacks on Israel itself.

3. To get a better sense of the novelty of these immediate reactions, we compared them to our previous analysis. While the overall antisemitism levels remained similar, a much wider range of antisemitic concepts was employed in the 2021 corpus. One reason for this may be a rise in social acceptance of antisemitic expression online between the two events, reducing the need to use a variety of stereotypes or analogies to demonise – or justify action against – Israel. In the 2023 corpus, expressions of antisemitic resentment appear to be more overtly stated.

4. In the subsequent phase of the 2023 conflict (13–31 October), by which time the Israeli military response in Gaza had become the centre of media attention, the pattern of antisemitic discourse gradually reverted to that seen in previous escalation phases in the Middle East. Across all three language communities, claims that Israel was EVIL, a TERRORIST STATE or GENOCIDAL, various forms of the NAZI ANALOGY, as well as notions of Jewish POWER and INFLUENCE, replaced the prior AFFIRMATION OF VIOLENCE as the top concept.

5. Throughout the aftermath of the Hamas attacks, *Instagram* played a crucial role in the online discourse, particularly amongst younger web users. Analysis of responses to posts by both pro-Israeli and pro-Palestinian *Instagram* influencers revealed that the most frequent antisemitic ideas were claims that Israel is inherently EVIL, is committing GENOCIDE and has NO RIGHT TO EXIST. These were often entangled in CONSPIRATORIAL narratives pointing to supposed Jewish political or economic POWER and INFLUENCE.

6. Our analysis of 2,000 social media comments related to Elon Musk's controversial statements throughout 2023 highlighted the phenomena of AFFIRMATION and DENIAL OF ANTISEMITISM, here triggered by the use of classical tropes of EVIL and CONSPIRACY by influential public figures.

7. For the first time, the project expanded its focus to other language communities, examining web user responses to 7 October in Italy, Spain, Poland, Slovakia, Romania and Morocco. This analysis provided a glimpse into the distinct discourses of each language community: from a notably high level of antisemitic slurs in the Romanian corpus to the prominence of conspiracy theories in the Polish data.

8. In the autumn of 2023, we ran a series of online workshops which reflected on the methods and aims of our project, established a dialogue with other expert approaches, and opened a debate on the practical challenges of confronting hate speech. This endeavour illuminated both the advancements and shortcomings in surrounding research areas.

9. As demonstrated by our work on the AI-based Step 2, automated detection of antisemitic content online continues to pose significant challenges and reaffirm the importance of contextuality. Even state-of-the-art models like ChatGPT, which have processed vast portions of the internet, encounter substantial difficulties. A reconsideration of the approach and examination of entire discussion threads might improve the current results.

10. As part of the quantitatively oriented Step 3 of the project, we developed an interactive data visualisation, accessible on our homepage. Using these tools, various discourse events can be compared in detail, allowing users to inspect the frequencies and co-occurrences of labels assigned in qualitative annotation during Step 1, as well as text specific antisemitic keywords.

1. Decoding Antisemitism: the journey so far

Matthias J. Becker

The imperative for investigating online antisemitism has seldom been as compelling as at the time of preparing this report. The events of 7 October plunged the Middle East into chaos once again and have brought the issue of global antisemitism, both on- and offline, to the forefront of public attention. The emotional urgency with which anti-Israel demonstrators poured onto the streets following the Hamas atrocities, sometimes even before any substantial Israeli military response, the rallies at Western universities often explicitly expressing solidarity with Hamas, calls for the elimination of Israel and physical attacks on Jewish people, symbols and synagogues, was matched by an eerily cold detachment from the physical violence, often laced with misogyny, inflicted on Israeli civilians on 7 October. For some, the Hamas atrocities were a cause for outright celebration.[2] Others sought to distort or deny the reality of the attacks, refusing to give Israelis the status of victims, and treating demands for the release of hostages as little more than confected Israeli propaganda.[3]

The central role played by social media in facilitating the wave of Israel-related antisemitism that has engulfed global societies over the last months emphasises once more the singular importance of examining the particular forms of antisemitic expression and conceptual content found online. This is true not only for distorted debates about Israel but also for another deeply concerning phenomenon, namely, indications that young people are increasingly susceptible to distortions and even doubt the veracity of the Holocaust altogether.[4]

While the politically moderate legacy media is by no means devoid of anti-Israel bias, particularly within left-leaning outlets, it is still thankfully rare to find antisemitic concepts, such as stereotypes and overt demonising analogies, within mainstream reporting. By contrast, bottom-up narratives emerging from web communities, charged by anonymity and algorithmic processes of radicalisation, have created the conditions for the free articulation of a discourse of antisemitism laden with (ancient and updated) stereotypes and history-distorting analogies.[5] In addition to the decline of the professional safeguarding once offered by mainstream journalists and media outlets, web users rarely face serious consequences when rehabilitating antisemitic tropes in public discourse, or when pursuing the ruthlessly simplistic evaluation of the highly complex Arab-Israeli conflict in a black-and-white dichotomy.

---

2 – Becker, Matthias J. et al., 2023.

3 – Rich, Dave, 2023. The poster war. Everyday Hate, 29 October 2023 https://everydayhate.substack.com/p/the-poster-war (last accessed on 15 December 2023).

4 – The Economist, 2023. One in five young Americans thinks the Holocaust is a myth. *The Economist*, 7 December 2023, https://www.economist.com/united-states/2023/12/07/one-in-five-young-americans-thinks-the-holocaust-is-a-myth (last accessed on 18 December 2023); Ramgopal, Kit, 2023. Survey finds 'shocking' lack of Holocaust knowledge among millennials and Gen Z. *NBC News*, 16 September 2023, https://www.nbcnews.com/news/world/survey-finds-shocking-lack-holocaust-knowledge-among-millennials-gen-z-n1240031 (last accessed on 15 December 2023); Weisman, Jonathan/Igielnik, Ruth/McFadden, Alyce, 2023. Poll Finds Wide Disapproval of Biden on Gaza, and Little Room to Shift Gears. *New York Times*, 19 December 2023, https://www.nytimes.com/2023/12/19/us/politics/biden-israel-gaza-poll.html (last accessed on 13 January 2024).

5 – Becker, Matthias J., 2021. Antisemitism in Reader Comments. London: Palgrave Macmillan; Hübscher, Monika/von Mering, Sabine (eds.), 2022. Antisemitism on Social Media. London: Routledge.

All these are serious trends whose significance is likely to grow, as the web becomes ever more central to public debate and the formation of political worldviews. The consequences of this epochal shift in communicative conditions associated with the interactive web[6] are still to be fully reckoned with. Still, as the events following 7 October have shown, their impact will affect forms of thought and behaviour in the offline world as much as it does online.

The Decoding Antisemitism project, which commenced its activities in Berlin, London, and Paris in the summer of 2020, has placed this need for a comprehensive assessment of antisemitic discourses on the internet at its core. On the one hand, the observable trends of the normalisation of antisemitic hate communication in recent years justify this research interest. On the other, research initiatives like ours bring numerous opportunities to explore the emergence, nature and trends of social discrimination and hate. The ongoing shift of socio-political debate to social media platforms opens up unprecedented possibilities for real-time analysis of public discourse. Reconstructing cross-societal attitudes towards Jews is no longer limited to surveys that can only reveal a small part of such thinking, and which are often affected by social acceptability bias.

Apart from this, many previous projects on antisemitism or on online hate speech in general, have often relied on a single disciplinary approach and therefore artificially narrowed down the subject. However, well-balanced measurement instruments, drawn from a range of scientific disciplines, are a central requirement for reliable results when investigating a phenomenon as complex and historically divergent as antisemitism. This is particularly so when seeking to build programmes of prevention and intervention on the back of scientific exploration of the subject.

The conception of the Decoding Antisemitism project has sought to overcome the problem of a singular disciplinary focus by bringing together researchers from a variety of **scientific disciplines**,[7] guided by the contention that any understanding of this ideology of hatred on the web must follow the principles of scientific best practice. One of these principles is that past and present antisemitic communication must be seen in context, in order to draw out continuities and innovations and to interpret current communication patterns. Such an approach also means that social media researchers can potentially predict trends of this highly complex worldview in the near future, since the progressively conventionalised patterns in social media will continue to significantly influence thinking and behaviour even in offline contexts. This potential outcome is all too often ignored in public discussions on the uses of social media analysis, despite being central to the attempts to tackle antisemitism.

In the project team, history and diachronic antisemitism research merge with contemporary linguistics, semiotics, image and discourse research. This qualitatively oriented team is, in turn, juxtaposed with a quantitatively oriented second team, which uses the detailed analyses of the former as a foundation for their data science work packages – we delve into this further below.[8]

6 – Becker, Matthias J./Bolton, Matthew, 2022. The Decoding Antisemitism Project – Reflections, Methods, and Goals. In: *Journal of Contemporary Antisemitism* (JCA), 5 (1), 121–126, https://doi.org/10.26613/jca/5.1.105.

7 – Decoding Antisemitism team: https://decoding-antisemitism.eu/team.

8 – Becker, Matthias J., 2019. Understanding Online Antisemitism: Towards a New Qualitative Approach. Fathom Online, http://fathomjournal.org/understanding-online-antisemitism-towards-a-new-qualitative-approach/?highlight=Matthias%20Becker.

Step 1

The Decoding Antisemitism project aimed at developing a **research design** that meets the high standards of web-related antisemitism studies.[9] The nature of antisemitism, characterised by high adaptability, historical diversity and – especially since 1945 – a tendency towards implicit or coded expression, can only be fully captured through the arduous and time-consuming path of granular expert analysis. This task has required the collective development of guidelines to assist in consistently identifying and categorising antisemitic expression in all its myriad forms by the whole team. The result of this work is a **guidebook**, based on an expanded and operationalised version of the IHRA definition, and thus better equipped to grapple with the challenges of authentic web discourse than the broad brushstrokes of the original IHRA document. The guidebook consists of about 175 categories – comprising definitions and examples for explicit and implicit antisemitic hate speech – and covers all concepts and communicative patterns observed in different language communities and web communities during the project period. Furthermore, the guidebook offers delineations of grey areas, such as the differences between antisemitism and legitimate criticism of Israel, or antisemitic and non-antisemitic anti-capitalist and anti-elitist rhetoric (in the context of discussions about, e.g. George Soros or the COVID-19 pandemic).

To make these theoretically and empirically grounded tools available to a broader audience – academics and non-academics alike – in 2024 we will be publishing a **Lexicon** presenting the central antisemitic concepts and strategies in the politically moderate online discourse of the three language communities we examined. This lexicon provides historical context and, through numerous examples, clearly explains their explicit and implicit manifestations.[10]

Over the course of the pilot project, Decoding Antisemitism has analysed more than 125,000 user comments responding to **27 discursive events** (i.e. incidents with the potential to provoke antisemitic communication). Importantly, the discourses analysed were not those of societal or political fringe groups (i.e. the radical right), but contained within the readership of **politically moderate** mainstream media in the UK, France, and Germany. This starting point (which must be carefully considered in qualitative analyses due to the mass of online data) illustrates how antisemitism is, to an extent, country- and topic-dependent. Nor is this ideology of hatred limited to 'classic' ideas of Jewish power and greed alone, but has incorporated a whole spectrum of concepts in history and the present, communicated in countless (verbal and visual) patterns.

We compiled regular updates of our case studies in biannual **Discourse Reports**, explaining reactions to scandals in the three countries (e.g. Dieudonné in France, David Miller in the UK, Hans-Georg Maaßen and *the documenta 15* art festival in Germany) or in the USA (Kanye West and, in this report, Elon Musk). Special attention was also given to contrast analyses of international events, such as the war in Ukraine and the various escalation phases observed in the Middle East over the last four years.

These **detailed analyses** are essential. The images of Jews and Israel circulating in social media in recent years seem to have intensified over time, thereby laying the groundwork for what was, even for experienced observers, the surprising step change represented by the wave of hatred and lack of empathy on 7 October. The conceptual arsenal and communicative patterns of antisemitic hate speech are likely not limited to the forms taken in this recent manifestation but will continue to evolve in the future, adapting to the events discussed and evaluated online.

9 – cf. Becker, Matthias J./Bolton, Matthew, 2022; see also the Decoding Antisemitism research design: https://decoding-antisemitism.eu/about.

10 – Becker, Matthias J./Troschke, Hagen/ Bolton, Matthew/ Chapelan, Alexis (eds.), 2024. Decoding Antisemitism: A Guide to Identifying Antisemitism Online. London: Palgrave Macmillan/ Springer Nature.

One important insight from the pilot phase is a concrete reaffirmation of the diversity and flexibility of antisemitic expression. Antisemitism online is likely even more diverse than antisemitism offline, as the latter does not need to dynamically adapt to a perpetually fluid discursive field and changes to moderation methods by social media platforms. The online milieus are numerous, as are the constructions of in- and out-groups flourishing in these spaces. Additionally, there are different language communities and cultural or social groups, each with their own collective codes and shared memories. All these factors make antisemitism almost entirely disappear in some contexts (explaining the absence of antisemitic conspiracy theories in the context of the Ukraine war in German mainstream media milieus), appear in coded form (as in reactions to concentration camp trials in Germany or COVID-19 conspiracy theories in France), or openly escalate in other places (e. g. the 7 October terror attacks in UK online contexts). Different conditions trigger specific patterns of this highly complex ideology. The idea of a single face of antisemitism, as in a political movement or party, would distort and simplify the subject. The difficulty associated with such content-structural diversity is also evident in the challenges posed by automatic detection of antisemitic communication (see below). In principle, however, this means that expert analyses of web content will continue to be a fundamental prerequisite for research, prevention and intervention in the future.

Step 2

The second goal of the research project – the **training of large language models** (LLMs) based on our annotated datasets – initially proved more challenging than anticipated. After an explorative cooperation between TU Berlin and King's College London, a comprehensive and fruitful collaboration began with Professor Helena Mihaljević and her team at HTW Berlin. In addition to ongoing exchanges with the University of Washington in Seattle, HateLab at Cardiff University, and the European Online Hate Lab, an intensive feedback loop between the team responsible for qualitative analyses and HTW Berlin has been ongoing since the summer of 2022. The first results were compiled in Discourse Report 5 (April 2023) and critically compared with the findings from the frequently used Perspective API by Google. Chapter 5 of this report presents the results after additional training sessions of multiple large language models by Milena Pustet and introduces a web application that enables users to generate predictions based on the best-performing model and subsequently evaluate them.

Step 3

Following the feedback loop between the humanities and social science working groups and the data science domain, the pilot project aimed for **quantitative analyses** of vast datasets (Step 3). The pilot project has primarily focused on the collaboration of Steps 1 and 2 and the training of the LLMs. However, one important quantitative output here are vector analyses showing the frequencies of words and phrases, as well as the visualisation of our results in both steps.

This collaborative approach means that the Decoding Antisemitism project is, in the truest sense, an **interdisciplinary initiative**. The pilot project has broken new ground by making such intense collaborative work a reality, with all the successes and challenges that it entails. Its methods and results have thus laid the groundwork for a new approach to the study of antisemitism and other hate ideologies online that can be used as a basis for a host of future studies.

Beyond the academic community, we have engaged in substantive dialogue with and provided training for **civil society projects** in the education sector. We collaborate with the Tel Aviv Institute, dedicated to strengthening those **influencers** seeking to challenge antisemitism on social media. Moreover, we conduct workshops in various settings, including schools, universities, vocational schools, and within conferences and meetings. Our consistent goal is to introduce participants to the linguistic and multimodal deconstruction of implicit hate communication and shed light on the grey areas between criticism and hate speech.

In addition to our cooperation partners, we have collaborated with the European Commission, UNESCO, JPR, ISGAP, the London Centre for the Study of Contemporary Antisemitism, the UK Antisemitism Policy Trust, the AJC, the Wiener Holocaust Library, the Holocaust Educational Trust, and initiatives such as the reporting centre Respect, Hate Aid, Drudel 11 (for the deradicalisation of right-wing extremists), The European Observatory of Online Hate, Respond! (led by Touro College Berlin) and Terraforming. We remain hopeful for increased and improved avenues of knowledge dissemination from empirical science to civil society, education, politics, and media, including social media platforms and influencers, as the findings from social media studies can offer entirely new perspectives for understanding and addressing the issue.

Future goals

Our goal is to utilise the scientific insights and best practices for coordinating various work packages from the pilot phase as a foundation for a new project phase that encompasses a larger geographical scope and multiple language communities, including the USA and Canada. Furthermore, the extension in scope will also include extremist discourse, through a contrastive approach, highlighting the processes of ideological transfer and percolation (often referred to as 'mainstreaming') between radicalised and mainstream milieus.

Alongside an established team for qualitative analyses, intensified collaboration with the data science domain will allow us to not only cover textual antisemitism but also to identify multimodal (text-image) patterns of antisemitic communication. This involves training models not solely based on texts. Our long-term goal is to fine-tune models for each language community to the extent that we can track trends on various social media platforms in real time during a discourse event. This is not only a crucial pillar in security responses to specific threats aimed at Jewish communities but also the basis for scientifically reconstructing the manifestations of antisemitic communication. In turn, this will enable the design and rapid implementation of further preventative and intervention measures, a task that becomes ever more important as the tidal wave of antisemitic expression online continues to rise.

The Sixth Discourse Report

This final Discourse Report of the pilot phase provides a detailed update on the results of Steps 1–3. Readers can gain a clear understanding of the current status of the project and what we have learned in its course. Section 2 presents a summary of our analyses of the escalation phase from 7 October. This extends the findings presented in our preliminary study, published on 26 October, in which we focused on *YouTube* and *Facebook* profiles of UK, German and French media (see footnote 1). Additionally, we have contrasted online responses in October 2023 with reactions to the escalation phase in May 2021 (Sections 2.1 and 2.2), examined *Facebook* responses following the start of the Israeli bombing and invasion of Gaza soon after Hamas attacks, as well as a study specifically on *Instagram* discourses (2.3 and 2.4). We also expanded our analysis to six other countries, each with its distinct historical, cultural and political context that influences patterns of antisemitic discourse: Italy, Spain, Poland, Slovakia, Romania and Morocco.

Section 3 moves away from Israel-related antisemitism and focuses on the online debates triggered by Elon Musk's posts about George Soros, the ADL, and further explicit antisemitic online statements.

In Section 4, we summarise our workshop series which took place across autumn 2023, and set out some of the take-aways from our engagement with the diversity of methodological approaches used by other research projects dealing with a broad spectrum of hate ideologies in the digital age.

Section 5 presents the current findings from the AI-based efforts of our collaborative partner, HTW Berlin, where the authors present various approaches for automatically detecting antisemitic content online. This includes initial experiments with state-of-the-art conversational models such as ChatGPT. The section concludes with a discussion of the challenges associated with the task and explores potential future strategies to address these hurdles.

Section 6 provides an overview of our quantitative analyses, in which Stefan Munnes (WZB Berlin Social Science Center) highlights the potential and difficulties of conventional quantitative text analysis methods in researching language-fluctuating phenomena such as antisemitism. The detailed and comprehensively labelled data from Step 1 allow for a more precise representation of nuances and changes over time and between discourse events. The section presents specially developed interactive visualisations of these differentiations.

2. The Hamas terror attacks on 7 October

Laura Ascone, Matthias J. Becker, Matthew Bolton, Alexis Chapelan, Karolina Placzynta, Marcus Scheiber

On 7 October 2023, Hamas terrorists broke into southern Israel from Gaza and unleashed an unprecedented violent assault on Israeli civilians. More than 1,200 civilians were killed, with hundreds more attacked, injured, and raped. 240 Israelis were taken hostage. The bodies of some of the victims were paraded through the streets of Gaza. In response, Israeli military forces mounted a fierce bombing campaign on Gaza, followed by a full-scale ground invasion, with many thousands of Palestinian casualties.

The events of 7 October – the deadliest attack on Jews since the Holocaust – and the subsequent Israeli invasion represent a step change in the Middle East conflict, generating a huge media response and frenzied reactions on social media, from all sides. Due to the significance of the events, we have examined various platforms, measurement periods, and language communities to assess the reactions provoked by the occurrences in the Middle East and to discern both commonalities and distinctions.

2.1 Immediate online responses to the events

In the first instance, we examined the *Facebook* profiles of mainstream media outlets in the United Kingdom, France, and Germany during the initial days of the escalation. The corpus for this study was built from threads responding to reports on the Hamas atrocities posted from across the politically moderate spectrum. Each of the threads was published between 7 and 10 October. These reports were all focused on the Hamas attacks themselves, or the immediate aftermath, rather than the retaliations of the Israeli military in Gaza. For each language community, we analysed a total of 1,500 comments.[11]

2.1.1 UK

The Hamas attacks on Israeli civilians on 7 October 2023 unsurprisingly generated a huge amount of coverage and reader response in the UK media. Across the corpus[12] as a whole, the average level of antisemitic comments was 21.6%, and therefore similar to other escalation phases of the Israel-Hamas conflict. However, levels of antisemitic expression ranged widely across the threads, from less than 10% in some threads to an unprecedented 44% in others.

The most striking feature of the antisemitic comments within this dataset was therefore at the level of concept choice rather than the frequency of antisemitism. **By far the most frequently expressed form of antisemitism was the affirmation and even outright celebration of the Hamas attacks (27%).**[13]

Their frequency jumped notably in responses to news articles reporting directly, often in graphic detail, on the Hamas atrocities, as opposed to more general reports of the conflict. This surge was particularly

---

11 – The structure and size of this dataset differs somewhat from that presented in our preliminary study of 26 October 2023 – cf. Becker, Matthias J. et al., 2023 – as direct comparability of samples from each language community became a central prerequisite here. However, the principle of random sample selection was still preserved.

12 – The dataset consists of ten *Facebook* profiles of *The BBC, The Telegraph, The Independent* and *The Guardian*. We analysed the first 150 comments from each thread.

13 – Given the extreme nature of the attacks, responses to direct reports of the violence which simply stated 'Free Palestine' or posted a Palestine flag were categorised here as AFFIRMATION OF THE VIOLENCE. This is in contrast to similar comments posted in response to reports of Israeli military reprisals, where they would not be categorised as affirmation of the Hamas attacks, and therefore not antisemitic.

visible below the line of news stories portraying Israelis and/or Jews as victims, and/or potentially vulnerable.[14] As noted in the following comparative study below (2.2), previous escalation phases of the conflict had not featured high levels of explicit self-positionings by web users openly endorsing, applauding, and celebrating the terror attacks. The centrality of such celebration within the online response to the 7 October attacks therefore seems to indicate a turning point in antisemitic web discourse regarding Israel.

Users often expressed such affirmation directly:

> "way to go HAMAS.. we support you.." (TELEG-FB[20231009]); "Very good Palestine 🇵🇸 👏 good job God bless you (TELEG-FB[20231007])",

> "long live HammasLong live Palestine 🇵🇸 We stand with them and appreciate their every step they take israel is a terrorist state and deserve such kind of destruction" (BBC-FB[20231009]).

Alternatively, they supported the attacks by depicting them as the actions of 'freedom fighters' – "Fighting against an occupying force is not terrorism, it's fighting for freedom" (INDEP-FB[20231008]); "Freedom Fighters aren't terrorists" (TELEG-FB[20231010]). Other commenters gleefully predicted that the attacks were only the first step in a long campaign that would lead to Israel's destruction: "do you think they entered it with no plan B,C,D...enjoy the last days of is-not-real" (BBC-FB[20231008]). This comment combines **AFFIRMATION** with the **DENIAL OF ISRAEL'S RIGHT TO EXIST** through the use of wordplay.

Such denials made up the second-largest category of antisemitic comments in the corpus (18%). Commenters often combined a delegitimisation of Israel with a threat of Israel's future eradication: "there is no such thing as israel and no matter how many years will take they will be expeled from palestine" (BBC-FB[20231009]); "All Islamic states must join hands with Palestine. Israel must permanently deleted" (BBC-FB[20231008]). Accusations that Israel was entirely responsible and has **SOLE GUILT FOR THE CONFLICT** – including the 7 October attacks themselves – were also common (16.9%):

"But doing what the Israeli's have been doing for 75 years in Palestine, and expecting no comeback as a result of their actions is outrageous.Kick a dog for long enough, it will bite you.I cant say that the dog is to blame because the dog is the one mistreated, and most people would say good for the dog" (TELEG-FB[20231009]).

Other commenters took the reports of the Hamas attacks as an opportunity to express their condemnation of Israel in ever more extreme ways. **Comparisons between Nazi Germany and Israel** were prominent – "self-proclaimed Fascist government which believeds in shooting and terrifying chuildren and has as it's xcentral policy a 'Final solution' of extermination of the natives.70 years ago, we were FIGHTING People like that.What changed?" (GUARD-FB[20231012]).

---

14 – In addition, the strikingly comparable observations made by Morris should be mentioned here: "The curious thing is that antisemitic incidents increased exponentially after the massive Hamas slaughter of 7 October, even before the Israelis launched their counterattack on Gaza. In other words, it is signs of Jewish – meaning Israeli – weakness that initially triggered the recent antisemitism in the United States, rather than perceived Israeli wrongdoing against Arabs, though that surely didn't help. In this, there are echoes of 1930s Poland, where antisemitism increased in response to German attacks on Jews and signs of Jewish weakness." (Morris, Benny, 2023. Muslim Antisemitism and the Western Left. Quillette, 7 November 2023, https://archive.ph/ubbeN#selection-455.0-455.40).

Others extended the 'freedom fighter' image to make comparisons between Hamas and resistance fighters to the Nazis, thereby indirectly activating the NAZI ANALOGY between Israel and the Third Reich: "Palestine stands condemned, for what? Doing what the British would have done if Germany had made it across the channel, what the French resistance did do" (GUARD-FB[20231011]). Some web users activated **classic antisemitic associations of Jews with THE DEVIL**: "Its Israel who has been committing genocide of Muslims since decades and all the devil powers are with her... 😡" (TELEG-FB[20231012]), or made DEHUMANISING claims that "Zionism is a cancer" (TELEG-FB[20231010]).

2.1.2 France

As in the UK, Hamas's attacks were widely covered by the French media. Within the French dataset,[15] **16% of comments were considered antisemitic**. Similarly to the British corpus, direct or indirect **AFFIRMATION OF VIOLENCE carried out by Hamas represented by far the most prominent form of antisemitism (55.2% of all antisemitic comments)**. This again suggests that while overall levels of antisemitism remained stable, its severity and intensity increased. Explicit support for Hamas's actions – such as "Long live Hamas" ["Vive le Hamas"] – was often diluted by more generic support for the Palestinian cause: "Total support to Palestine" ["Soutien total à la Palestine"] (FRA-FB[20231009]). That notion was also articulated through the allegation that the attack was morally justified as self-defence or retribution for Israeli violence: "It is only payback for the Zionist parasites" ["Ils n'ont fait que rendre la monnaie de leur pièce aux parasites sionistes"] (MONDE-FB[20231008]).

The NAZI ANALOGY was also a popular means of conveying support for Hamas (10.4%). Hamas was regularly compared to the French Resistance through comments such as "How did we end up turning resistance into terrorism? During the occupation in the 1940s, the Resistance took up arms against the Germans and they were glorified?" ["Comment on arrive à transformer la résistance en terrorisme? Pendant l'occupation dans les années 40 la résistance ont bien pris les armes contre les allmands et on les a glorifié?"] (LEFIG-FB[20231008a]). Another comment activates the NAZI ANALOGY but also maps onto negationist narratives:

> "Hamas is paying it back for their people killed since 48 by ✡ committing genocide (Holocaust), but a real one, and the perpetrators are the survivors of 45 🤔"

["Hamas rend la monnaies de son peuple qui se tue par les ✡ depuis 48 faisant un / génocide (holocauste) mais un vrais et les bourreaux sont les rescapés du 45 🤔"] (LEFIG-FB[20231008b]).

By contrasting the "real" genocide of Palestinians with the supposedly 'fabricated' one of the Jews, this comment interestingly represents a projection of Nazi atrocities onto the Jewish state, while denying the existence of the Nazi genocide. In addition to these direct forms of reference, the analogy can also be articulated through puns, such as in "Isra Heil" (MONDE-FB[20231008]).

15 – The dataset stems from the *Facebook* profiles of *Le Monde, Le Figaro, Le Parisien, Libération, BFMTV, France24* and *20Minutes*. We analysed the first 150 comments from each thread.

Another frequently activated concept is the DENIAL OF ISRAEL'S RIGHT TO EXIST **(9.5%)**: "Israel has no right to peace and security. For the simple reason it has no right to exist" ["Israel n'a droit ni a la paix ni a la securite. Pour la simple raison q'il n'a pas droit a l'existence"] (LEFIG-FB[20231008a]). **This concept is often implied through the allegation that Israel is a TERRORIST STATE (6.5%):**

> "How many Palestinians killed by the terrorist state of Israel !?!!🤔"

["Combien de palestiniens mort par l'état terroriste d'Israël !?!!🤔"] (LEFIG-FB[20231008b]).

In other terms, the reactions to the attacks perpetrated by Hamas in October 2023 seem to focus more on the appreciation for Hamas (or Palestine, when contextually understood as a metonymy for the terrorist group), than on the demonisation of Israel. This shows that even positive statements – positive from a lexical perspective – can actually represent a form of hate speech and, in this case, of antisemitic discourse. The context of the statements is therefore crucial to determine their nature.

2.1.3 Germany

The Hamas attacks on Israeli civilians generated a highly amplified response also on German social media. Within the dataset,[16] a total of **12.6% of all comments were identified as antisemitic**, although there was a wide range of 8% to 29% of antisemitic comments in the individual threads.

Even though the average proportion of antisemitic comments is comparable to other discourse events in German-speaking countries (see previous discourse reports and section 2.2), some thematic and conceptual peculiarities are striking. First of all, the antisemitic statements articulated within the first phase of the conflict justify the cause of Hamas' attack by reference to Israel's behaviour, **attributing to Israel SOLE GUILT for the conflict (29.6% of all antisemitic comments)**: "What has isreal done with all the Palestinians?! Think first!!! What happened before. Where all this anger comes from..." ["Was hat isreal mit den ganzen Palästinenser gemacht?!Erstmal nachdenken!!! Was vorher geschehen ist. Woher diese ganze Wut kommt...."] (WELT-FB[20231008]). This is legitimised by the accusation that the attack is due to **Israel being a TERRORIST STATE (12.7%)** – "The Palestinians are fighting against the terrorist state of Israel" ["Die Palästinenser kämpfen gegen den Terrorstaat Israel"] (BILD-FB[20231009a]), **or committing GENOCIDE[17] (5.8%)** against the Palestinians – "If someone wants to liberate their country and save their people from genocide, they are not a terrorist but a freedom fighter!" ["Wenn jemand sein Land befreien will und sein Volk

---

16 – The dataset contains ten threads, with 150 comments analysed from each and stems from the *Facebook* profiles of *Der Spiegel*, *Die Welt*, *Bild* and *n-tv*.

17 – In identifying distorted statements, such as the claim that Israel would commit a genocide or structurally resemble South Africa's apartheid, we follow the principle that assertions predicting similar scenarios in and around Israel in the future are not to be classified as antisemitic (see Becker, Matthias J./Troschke, Hagen/Bolton, Matthew/Chapelan, Alexis (eds.), 2024. Decoding Antisemitism: A Guide to Identifying Antisemitism Online. London: Palgrave Macmillan/Springer Nature. However, while there is no international political and legal consensus yet that would fully legitimise the genocide accusation, we acknowledge the ongoing, urgent debate around this topic.

vor einem Genozid reten will ist kein Terrorist sondern ein Freiheitskämpfer!"] (BILD-FB[20231009b]). The presentation and legitimation of Hamas's actions as a fight for freedom pervaded the entire discourse: "In Ukraine, these people are called freedom fighters" ["In der Ukraine nennt man die Leute Freiheitskämpfer"] (BILD-FB[20231009c]). As a result, **Israel's RIGHT TO SELF-DETERMINATION is often completely denied (24% of all antisemitic comments)**: "The Jews are the problem. We all know that. They don't own the country" ["Die Juden sind das Problem. Wissen wir doch alle. Ihnen gehoert das Land nicht"] (WELT-FB[20231008]); "question ,history who owns israel ,freedom for palestine" ["frage ,geschichte wem gehört israel ,freiheit für palästina"] (BILD-FB[20231009b]).

Moreover, these attempts to legitimise Hamas's violence through Israel's actions culminate in open **AFFIRMATION** of the terrorist attacks – "Awesome what the Palestinians have pulled off, cheers to Hamas for a great action" ["Geil was die Palästinenser da abgezogen haben ein hoch auf Hamas geile Aktion"] (BILD-FB[20231009b]) – or articulation of the desire for future and even more extensive violence against Israel and/or Jews:

"I hope Israel becomes history like Ukraine" ["ich hoffe Israel wird Geschichte so wie die Ukraine] (BILD-FB[20231009c]).

2.2 Comparison of October 2023 and May 2021

In order to evaluate the novel aspects of the 7 October reactions, we sought to compare it with online responses to a previous outbreak of violence in the Middle East. The escalation phase of the Israel-Hamas conflict in May 2021 was the most serious outbreak of violence in the years preceding the 7 October attacks. At that time, Hamas fired hundreds of rockets from Gaza into Israel after clashes between Palestinians and Israeli police at the Al-Aqsa Mosque, followed by Israeli military retaliations in Gaza.

While neither the Hamas attacks nor the Israeli response came close to the severity of recent events, May 2021 nevertheless represents the closest chronological point of comparison for an evaluation of the similarities and differences in online responses.

The corpus for the May 2021 case study was again made up of 4,500 comments (1,500 from each language community) taken from the *Facebook* pages of leading media outlets.[18]

2.2.1 UK

The total level of antisemitism within the May 2021 threads[19] **was, at 26.9%, similar to that found in the October 2023 threads.** However, the range of the antisemitism levels was much narrower, with all threads bunching around the 20% mark and no high outliers. The analysed threads were divided into two main categories, one focusing on Hamas's rocket attacks and the other on the IDF's retaliatory actions. The pattern of antisemitic stereotypes in the comments sections shifted depending on the emphasis in the articles. When the media highlighted Hamas's activities, there was a notable increase in comments accusing the media of favouritism towards Israel. This response was clearly distinct from reactions to media reporting of the 7 October attacks – in the latter, there was little sign of concern about media bias, but rather, as noted above, direct AFFIRMATION of the attacks (27% of antisemitic comments).

Conversely, **few if any responses to the Hamas rockets of May 2021 sought to affirm or celebrate the attacks**. Speculatively, this might be due to the relative failure of the rocket attacks in comparison with the violence of 7 October: had the May 2021 rocket attacks resulted in the same physical, emotional and psychological impact of the recent Hamas atrocities, perhaps the response would have been more celebratory than conspiratorial. Another possibility for this difference may be an increase in social acceptability of antisemitic expression online between the two events. If overt antisemitism is finding greater social acceptance, there is less need for the use of stereotypes or analogies as means of justification. Alternatively, while the May 2021 escalation was not preceded by a build-up of tension, but seemed to come out of the blue, the October 2023 attacks followed a year in which the inclusion of far-right extremists in the new Netanyahu government, internal Israeli political conflicts over judicial reform, and sporadic but regular violence on the West Bank,

received low-level but persistent media coverage in all three countries. This may have had a radicalising effect on global perceptions of Israel or Jews as whole, which were then activated in the responses to the violence of 7 October.

On the other hand, when the May 2021 articles centred on Israel's response to the rocket attacks, **Israelis and Jews themselves became the target of various antisemitic stereotypes, without reference to media favouritism**. The most frequently expressed stereotypes in the corpus were notions of **Jewish or Israeli EVIL (39% of all antisemitic comments)** and attributing SOLE BLAME FOR THE CONFLICT to **Israel (27%)**. The latter concept was strongly represented in the October 2023 corpus as well (16.9%), as commenters who may be uneasy with direct affirmations of the Hamas attacks sought to justify them.

The concept of EVIL was relatively infrequent in the October 2023 corpus (6.4%). This may be due to the focus of the corpus being on Hamas attacking Israel, rather than vice versa. It may also be due to users preferring to condemn Israel by means of the closely related TERRORIST STATE accusation (8.3%), which often features as a response to media articles which depict Hamas as a terrorist organisation. Another concept to feature prominently in both corpora was the DENIAL OF ISRAEL'S RIGHT TO EXIST, although this was more common in the October 2023 comments (18%) than in the May 2021 corpus (7.7%). Again, the difference for this disparity may lie in the relative success of the respective Hamas attacks. As seen above, many web users saw the 7 October attacks as the first step towards the final destruction of Israel, and their euphoria may have motivated justification for, predictions or threats of Israel's ultimate demise.

---

19 – The dataset contains ten threads, with 150 comments analysed from each, and stems from the *Facebook* profiles of *BBC News, Daily Express, Daily Mail, Financial Times, The Guardian, The Independent, Daily Mirror, The Spectator, The Telegraph* and *The Times.*

2.2.2 France

Compared to October 2023, the online commentary on May 2021[20] exhibited slightly **lower antisemitism levels: 12.9%, as opposed to 16% in 2023**. More importantly, differences emerged in the content of these antisemitic web comments. In 2021, most of the antisemitic comments targeted Israel and Israelis rather than Jews. Here, the most frequent antisemitic stereotype was the one presenting **Israel as EVIL**:

> ```
>     “The barbarism of this
>  State cannot be hidden by
>      2—3 lies of traditional
>  media” [“Ce n'est pas 2 - 3
>  mensonges des médias tradi-
>  tionnels qui pourront cacher
>  la barbarie de cette état”]
>          (MONDE-FB[20210511]);
>      “It's the only people in
>    the world that boasts about
>          killing human beings”
> ```

[“C'est le seul peuple au monde qui se vante de tuer des êtres humains”] (MONDE-FB[20210511]).

This stereotype appeared **in 49.7% of antisemitic comments** in the 2021 corpus, but only 13.9% in 2023. The same tendency was identified for the allegation that Jews are **CHILD MURDERERS and that ISRAEL HAS NO RIGHT TO EXIST**: "They're resisting and defending themselves against the real terrorists who bomb children" ["Ils font de la résistance et se défendent contre les vrais terroriste qui bombarde des enfants"] (LEPAR-FB[20210511]); "Israel doesn't exist and will never exist" ["Israël n'existe et n'existera jamais"] (LEPAR-FB[20210511]).

The 2021 corpus was also characterised by comments accusing Israelis of holding **POWER (5.2%)** over the media and being **LIARS (4.7%)** and: "It's clear that le monde is paid by these demons" ["on voit bien que le monde est payé par ces demons"] (MONDE-FB[20210510]); "You're the champions in lying and falsifying history" ["Vous êtes champions dans les mensonges et la falsification de l histoire"] (LIB-FB[20210512]).

Meanwhile, the 2023 corpus presents many more references to **ISRAEL'S SOLE GUILT IN THE CONFLICT** (9.9% in 2023 and **only 2.6% in 2021**): "Israel never wanted peace, they're just reaping the benefits of their politics" ["Israël n'a jamais voulu la paix, ils récoltent juste les fruits de leur politique"] (LEPAR-FB[20231009]). As this comment shows, the concept is often used to legitimise and support Hamas's attacks as predictable and deserved due to Israel's own actions. The **AFFIRMATION OF VIOLENCE** perpetrated by Hamas, which as stated before constitutes the most prominent form of antisemitism in the 2023 corpus (53%), **was absent in 2021**. There, commenters did not express their support for the hundreds of rockets launched from Gaza but rather CALLED FOR A FUTURE VIOLENT ACTION against Israel:

> ```
> “your day will come and you
> will be mutilated, massacred,
> killed like the thousands of
> innocents Palestinians killed
> by Israel. This day will come”
> ```

["ton jour arrivera et tu seras mutilé masacré tué comme les milliers de palestiniens innocents tué par Israël Ce jour arrivera"] (LEPOI-FB[20210512]).

2.2.3 Germany

While a significant or slight increase in antisemitic comments can be observed within the British and French media, **the level of antisemitic comments in the German discourse across the two escalation phases remained quite stable.**[21] Within the 2023 measurement period, 12.6% of all coded comments were identified as antisemitic – in contrast to 13.4% in 2021. While its quantity was relatively constant, **the quality of antisemitic communication has changed strikingly**. In May 2021, Israel was demonised primarily via two antisemitic concepts.

Firstly, it was characterised as the **EVIL of the world** by implying that Israel's actions are essentially wicked or that it would purposefully and proactively cause extensive harm to others. While **39.7% of all antisemitic comments in 2021** contained the concept, in 2023 it was only 4.2%. As in May 2021, comments in the October 2023 corpus continue to portray Israel negatively, or **BLAME it unilaterally for the conflict**: "Yeah right, it's Hamas's fault. It's Israel's fault, they've been terrorising Hamas for years. But nobody cares. Now all of a sudden it's poor Israel. Hamas is right. Tit for tat" ["Aha die Hamas ist schuld.Israel ist schuld die machen schon Jahre lang Terror gegen die Hamas. Nur das interessiert keinen. Jetzt auf einmal heißt es Armes Israel. Die Hamas hat Recht.Wie du mir so ich dir"] (WELT-FB[20231007]). Instead of essentialising Israel across the board with the concept of evil, commenters are resorting to more nuanced justifications – which explains the decline in the use of the term compared to 2021, especially since the articulation of other stereotypes more than doubled in 2023: while in 2021 **11.3% of all antisemitic comments expressed ISRAEL'S SOLE GUILT IN THE CONFLICT**, 29.6% realised this stereotype in 2023: "Israel has been starting this for 70 years" ["Israel fängt seit 70 Jahre an"] (BILD-FB[20231009d]).

Secondly, in 2021, Israel was increasingly accused of exerting **INFLUENCE on both the media and politics (11.3%)** – a concept area that is less pronounced, but still consistently used in 2023: 7.9% of antisemitic comments in 2023 made accusations such as

"No wonder Bild reports all this about Israel, since Bild GmbH belongs to Jews"

["Kein Wunder das Bild das alles über Israel berichtet weil doch das Bild GmbH Juden gehört"] (BILD-FB[20231009e]).

In contrast, users in 2023 express their AFFIRMATION OF VIOLENCE more clearly. While in 2021 only 3.4% of all antisemitic comments convey such attitudes towards Israel and/or Jews, or legitimise the violent actions of Hamas: "Action ~ Reaction" ["Aktion ~ Reaktion"] (WELT-FB[20231008c]). The idea of ISRAEL'S SOLE GUILT IN THE CONFLICT therefore pervades both discourse events.

21 – In 2021, we examined 150 comments each from ten threads taken from the *Facebook* profiles of *Bild, FAZ, Focus, n-tv, RP Online, Der Spiegel, Süddeutsche Zeitung, taz, Die Welt* and *Die Zeit.*

2.3 Reactions to the Israeli offensive after 13 October

In order to continue tracking responses to the unfolding situation in the Middle East, a further corpus of 1,000 comments per language community was built from responses to reports (published 13 to 31 October) of the Israeli military bombing campaign and the then-prospective invasion of Gaza that followed the 7 October attacks.

2.3.1 UK

The UK dataset ranged from accounts of Israeli bombing raids to UK Prime Minister Rishi Sunak's visit to Israel.[22] Among the responses from the readership, **19.9% were classed as antisemitic**. Compared with the comments posted in the immediate aftermath of the Hamas attacks, there was a **marked drop in the level of AFFIRMATION and/or celebration of ANTISEMITIC VIOLENCE**. As the coverage returned to the standard pattern following previous escalation phases (with the central focus now being Israeli military actions rather than Hamas violence), so too did the pattern of antisemitic expression return to the mean. The most frequently expressed antisemitic concept was that of **Israeli/Jewish EVIL (27% of antisemitic comments)**. Users declared that "humanity will win the fight against the Zionist lobby,we will," positioning Israel as outside of and opposed to humanity, and added that "you might own the politicians but you don't own the people" – thereby making a claim of "Zionist" POWER/INFLUENCE over non-

Israeli political systems (BBC-FB[20231018]). **The second most frequent concept (17.5% of all antisemitic comments) was the accusation that Israel was committing GENOCIDE against the Gazan population.** Some commenters claimed that the Israeli genocidal attack on Gaza was unprecedented in its savagery, using language that evoked images of the Holocaust and thus expressed an indirect NAZI ANALOGY: "2.5 million people are subjected to a war of extermination and a siege that has never happened in history" (BBC-FB[20231018]); "Shame on the world! They kept watching when Hitler did it and now when Netanyahu is doing it! A genocide of Palestinians!" (INDEP-FB[20231013]).

Users made frequent references to **Israel as a TERRORIST STATE (11%)**, a common response to the depiction of Hamas as a terrorist organisation or to comparisons between the Hamas atrocities of 7 October and those of Da'esh: "Zionists is evil Zionists is the first terrorist Organization in the world" (DAILY-FB[20231020]); "israel the real isis" (BBC-FB[20231017]). The **DENIAL OF ISRAEL'S RIGHT TO EXIST was often the ultimate conclusion of antisemitic comments (10%)** – "actually what is not easy to believe Israel when it's whole existence is a lie" (BBC-FB[20231017]) – while users regularly resorted to the **NAZI ANALOGY (6.5%)** to underline their condemnation of Israeli actions: "Israel is the new nazism" (INDEP-FB[20231013]); "there no better then the Nazis" (DAILY-FB[20231020]); "You're worse than the very same people you cry about every year in your self-pitying memorial services – without having learnt f-all. Not even the Nazis mastered propaganda and harnessed the power of the media in the way that Zionists have in the modern age" (INDEP-FB[20231017]); "And you think hitler is the worst .. now i understand why he did the Holocaust" ([INDEP-FB[20231013]).

2.3.2 France

In the French corpus,[23] **8.9% of the comments presented an antisemitic statement**, meaning that the proportion of antisemitic comments decreased compared to the first days after the attacks. As with the UK results, analysis revealed that the severity and intensity of the antisemitic comments diminished somewhat: AFFIRMATIONS OF VIOLENCE perpetrated by Hamas were very prominent in reactions to the media articles published immediately after Hamas's attacks (53%), while in this second phase of the conflict they constituted only **10% of the antisemitic comments**.

In this corpus, the most prevalent concept is the **stereotype of EVIL (20.2%)**. The comments demonise actions or the nature of Israel itself, for example by claiming it is "a very murderous country" ["Israël est un pays très meurtrier"] (MONDE-FB[20231101]) or "an invasive metastasised cancer" ["un cancer invasif ayant métastasé"] (LEFIG-FB[20231101]). The last example, in which Israel is demonised through the disease metaphor, portraying it as an entity with a disintegrating effect, is reminiscent of the 19th- and 20th-century forms of antisemitism.

Although in much smaller proportion than in the first phase, one of the most frequent antisemitic concepts is assigning Israel SOLE BLAME FOR THE CONFLICT **(12.3%)**. Comments tend to explain this by listing the actions undertaken by Israel that have allegedly led to such a dramatic situation as "the expansion of the illegal Israeli colonisation" ["l'expansion des colonies israéliennes illegals"] (MONDE-FB[20231101]). The example illustrates the linkage of the concept of ISRAEL'S GUILT IN THE CONFLICT with COLONIALISM ANALOGIES, a phenomenon which usually co-occurs with the former. Implicitly, this consistently leads to a DENIAL OF ISRAEL'S RIGHT TO EXIST, underpinned by the false claim that Israel was established on Palestine: "the Zionist nation came with an army to settle by force on Palestinians' land" ["la nation sioniste est venue avec une armée s'installer de force sur la terre des palestiniens"] (LEFIG-FB[20231101]).

23 – The French corpus consists of the first 100 comments posted in reaction to ten articles published on the *Facebook* pages of the following media outlets: *Le Monde, Le Figaro, Le Parisien, France24* and *CNEWS*.

2.3.3 Germany

While the first phase of German reporting focused on Hamas's atrocities, the second centred on domestic reactions to the events. The corpus contains comments sections responding to a range of the latter's media reports: from pro-Palestinian demonstrations in Germany to assessments of these by the Office for the Protection of the Constitution, to the Federal Chancellor's visit to Israel.[24] The second phase of the study period is characterised by a significantly lower proportion of antisemitic comments: **only 8.5% of all analysed comments were identified as antisemitic**. This may be due to the change in reporting, as web users often used reports on the actions of the German government as an opportunity to criticise or express displeasure about domestic politics – in other words, to shift the topic of discourse.

It is particularly noticeable that **ISRAEL'S SOLE GUILT IN THE CONFLICT** is attributed significantly less frequently than in the initial phase – only 8.2% of all antisemitic comments realise this stereotype. The high levels of **AFFIRMATION OF VIOLENCE**, which accounted for 20.1% of antisemitic comments in the first phase, has also fallen significantly, to 12.5%. Instead, users return to **familiar patterns and characterise Israel – similar to the findings from May 2021 (see 2.2) – as a state characterised by essential wickedness, most frequently described as a TERRORIST STATE (17.6%)**: "For over 70 years, the terrorist State of Israel has trampled on human rights" ["Seit über 70 Jahren tritt der Terrorstaat Israel das Menschenrecht mit Füßen"] (WELT-FB[20231031a]), or as the "rogue State of Israel" ["Schurkenstaat Israel"] (SPIEGEL-FB[20231031]).

In addition, **the concept of INFLUENCE is often invoked (14.1% of all antisemitic comments)** when depictions of Hamas atrocities are negated – "Either they were fake pictures or pictures of murdered children from Palestine. Israel has some really nice fairy tales in store" ["Entweder waren es fakebilder oder Bilder von ermordeten Kinder aus Palästina. Israel hat richtige schöne Märchen auf Lager"] (WELT-FB[20231031b]) – or presented as disinformation instrumentalised by Israel in order to completely annex Gaza:

> "This whole report has no truth all without bodies only a few suspect pictures all nicely planned by israel so they can occupy the last open prison"

["Dieser ganze Bericht ist schon keine Wahrheit alles ohne Leichen nur ein paar Verdächtigen bilder alles schön von israel geplant damit sie das letzte offen Gefängnis besetzen können"] (FB-WELT[20231031c].

This idea – of disinformation arising from Israeli/Jewish **POWER OVER PUBLIC OPINION** (16.4%) – was expressed much less frequently in the first phase. This can probably be explained by the fact that a justification for demonising Israel's actions is still needed in order to rob Israel of any legitimacy for military action. However, this justification must not (any longer) glorify Hamas's violence, as such glorification is socially unacceptable.

24 – The dataset contains ten threads, with 100 comments analysed from each, and stems from the *Facebook* profiles of *Der Spiegel* and *Die Welt.*

2.4 Antisemitic reactions on *Instagram*

As events in Israel and Gaza unfolded, it became clear that *Instagram* was increasingly becoming a central terrain on which the online battle over narrative and interpretation was taking place, particularly amongst younger web users. As such, we constructed corpora made up of comment threads from popular *Instagram* accounts posting on the conflict in each language community. Both pro-Palestinian and pro-Israeli accounts were included, with a total of 7,500 comments analysed.

2.4.1 UK

The thirty comment threads we examined had been posted between 11 October and 4 November 2023; their focus was therefore not on the initial Hamas attacks, but rather on the Israeli military response, often using emotive language to condemn civilian deaths.[25] Some accounts sought to provide 'context' for the Israeli retaliation; others simply expressed support for one of the sides in the conflict or highlighted the suffering of all its victims. Unlike in our analyses shortly after the terror attacks on 7 October,[26] direct affirmation of the violence was noticeably absent; however, self-positioning expressions like threats, curses and death wishes increased. Out of 3,000 analysed comments, **21% were classed as antisemitic**. The percentage remained similar across both pro-Palestinian and pro-Israeli threads, though within each category there were serious outliers, with antisemitism levels around 40% in some cases. **The most frequent antisemitic concepts were accusations of** GENOCIDE **(20% of antisemitic comments),** DENIAL OF ISRAEL'S RIGHT TO EXIST **(16.8%),** EVIL **(16.4%), antisemitic** CONSPIRACY THEORIES **(9.5%) and** DENIAL OF ANTISEMITISM **(i. e. denying the nature, extent or existence of the Hamas attacks, 7.7%).**

However, the concepts often intertwined and their frequency varied according to the framing of the original post.

One post from the pro-Palestinian activist account *Let's Talk Palestine*, for example, was centred around an infographic claiming to provide an 'Israeli apartheid for beginners' guide. Unsurprisingly, the APARTHEID ANALOGY **was the most frequent antisemitic concept in the thread**, but many web users took the opportunity to fire off a series of concepts:

"Israel is a t3 rror !st genocidal state from it's inception […] Shame on you for defending the slaughter of thousands of children, shame on you for defending a 75yr ongoing genocide and ethnic cleansing campaign. Shame on you for cheering on a racist supremacist apartheid regime! Shame on you for your racist supremacist ideology! The blood you have on your hands can never be washed off!"

(LET-IN[20231022]).

25 – The corpus comprises 30 *Instagram* threads and a total of 3,000 comments. The accounts were primarily those of leading independent influencers (with the number of followers ranging from around 100,000 to 3 million). Eleven accounts were more-or-less explicitly pro-Palestinian (*Ahmed Eldin, Eye On Palestine, IfNotNow, IMEU, key48return, Khaled Beydoun, Let's Talk Palestine, Mohammed El-Kurd, Owen Jones, So Informed, Visualising Palestine*), four were more-or-less pro-Israeli (*Hen Mazzig, Stand With Us, State of Israel, Jews of NY*), and one was ostensibly neutral (*Médecins Sans Frontières*).

26 – Becker, Matthias J. et al., 2023.

Here the apartheid charge is repeated and **combined with the tropes of TERRORIST STATE, GENOCIDE, RACIST STATE, and REFERENCES TO FASCISM**; it seems to act as an enabling concept for the ever-more extreme depictions of Israel.[27]

Similarly, vilifying concepts were often clustered in antisemitic comments prompted by a post from a pro-Israeli account *Stand With Us*, in memory of Israelis who died following the Hamas attacks. The comments accused Israel and its citizens of being inhuman, cowardly, **EVIL and IMMORAL, and sometimes guilty of CHILD MURDER**: "killing babies, kids and woman will destroy Gaza soooooooooo humanity you are......bombs hospitals that only Israel can dooooooo cowardddddddd" (SWU-IN[20231014]). One striking example referenced **ancient anti-Judaic and racist DEHUMANISING stereotypes**, claiming that Jewish CHILD MURDER, DECEIT and CONSPIRACY had led Western countries to support Israel's creation in order to remove Jews from Europe: "in the middle ages the PIGS were living in Europe [...] they were very filthy with you, they live among you and kidnap your children to make bread out of their blood for the Jewish passover [...] you couldn't bear their actions nor that they live among you because they're like animals, their smell is always stinky and unbearable [...] You hated them a lot so you burned 6 million of them. You wrote in your books that the place they're around the percentage of wars is higher and there's no stability when they're around so the solution was is support their presence in our lands [...] that's why you support them because the most important thing to you is never go back to your lands" (LET-IN[20231011]).

Across virtually all the analysed threads, commenters debated the classification of the current events. We were careful to distinguish between the nuanced approaches, concerns about the conflict's future, or empathy for Palestinian victims on the one hand, and blunt and totalising accusations on the other, e.g. "This isn't a conflict and it's NOT A WAR. This is GENOCIDE" (KEY-IN[20231026]). Some of the latter employed the **NAZI ANALOGY** ("you and your kind will be remembered as evil and murderer and traitors and butcher .. just like the world remember hitler!! What he did to all the Jews they are doing the same ! Shame on you and anyone else supporting ethnic cleaning and genocide" (IME-IN[20231013])), sometimes denying Nazi crimes in the process: "Israel is doing to Palestinians what they claim Hitler did to them" (LET-IN[20231011]). Some resorted to HOLOCAUST DISTORTION: "That's nothing compared to the 7 million Palestinians Israelis have murdered over 75 years. Worse than the Holocaust in terms of the death toll" (HEN-IN[20231018]).

Perhaps unsurprisingly, the State of Israel's official *Instagram* account attracted high levels of antisemitic comments, with one thread – calling for the hostages taken by Hamas to be brought home – reaching 33%. One of the most frequent concepts in our dataset, **the DENIAL OF ISRAEL'S RIGHT TO EXIST**, in this thread was again shaped by the nature and word choice (particularly the use of 'home') of the original post. These were often accompanied by comments asserting that Jews were 'homeless' and therefore inherently FOREIGN/ALIEN – not just in the Middle East but globally:

"You don't have home. Hahaha-
haha where's so called home?
The land that you stole? 🤮🤮
🤮;"

"Pack your stuff, u gonna be
the gypsy for the third time
in life . Isn't that coinci-
dence? Nobody loves you in
this world"

(SOI-IN[20231016]).

27 – Bolton, Matthew/ Becker, Matthias J./ Ascone, Laura/ Placzynta, Karolina, 2023. Enabling concepts in Hate Speech: The Function of the Apartheid Analogy in Antisemitic Online Discourse about Israel. In: Ermida, Isabel (ed.). Hate Speech in Social Media. Cham: Palgrave Macmillan, https://doi.org/ 10.1007/978-3-031-38248-2_9.

Elsewhere, a comment asked "Why do Israeli people are in different colors and ethnicities? This indicates that you are like a flock that Britain gathered together and gave you land, but you believed that you are a country with a history and a cultural heritage. You are just a military base for the Americans and nothing more" (LET-IN[20231011]). The DENIAL was also expressed through wordplay: "You do know this has been going on for 70+ years right? And yes, this doesn't have to happen if 'isntreal' gets the tf out of Palestinian land" (HEN-IN[20231015]), or reinforced with emojis: "that is not even a country🤣🤣👋" (SWU-IN[20231019]). Other comments denied the nature and extent of the Hamas attacks or assigned **SOLE BLAME for the conflict** to Israel: "this isn't war, it's genocide. israel started the nakba in 1948" (HEN-IN[20231015]). The theme of allegedly fraudulent behaviour on the part of Israel ran through many comments, suggesting that **Israel had LIED about the hostages captured** by Hamas ("They're all safe…it's just a lie created by Isra-hell" (SWU-IN[20231019])), or that the state is cynically **INSTRUMENTALISING ANTISEMITISM** ("rubbish is powerful Israel being the 'victim' yet killing innocent children for 70+ years" (SWU-IN[20231019])).

The dishonesty argument was also used to support antisemitic **CONSPIRACY THEORIES**, which claimed "Hamas terrorists are funded by Israeli Government, such that they create chaos and suggest that's its Palestinians, barbaric animals" (EYE-IN[20231013]), presenting Israel as a destabilising force and blaming it for Hamas's actions, or for the entire conflict. Others insinuated **Israel's INFLUENCE ON FOREIGN POLITICS**: "American government has been hijacked by Israeli dual citizens and other stooges like both trump and Biden who always put Israel's needs above America's" (KEY-IN[20231026]). Israel and Jews are supposedly "everywhere in high places," they "have powerful allies" (KEY-IN[20231026]) and wield **POWER and INFLUENCE OVER THE MEDIA AND POLITICS** as well. The stereotype seems to be feeding on the growing mistrust towards established sources of information: "apparently that is how journalism works, considering journalists are spreading lies like Hamas beheading babies only for us to find out that there's been

zero proof of that. […] Journalists aren't unbiased" (IME-IN[20231013]). A post from pro-Palestinian influencer Khaled Beydoun targets not Israel as such, but celebrity Muslim artists who had apparently failed to publicly condemn Israel. The vast majority of antisemitic comments again expressed **the idea of JEWISH POWER**: "I'm confused, what exactly did I say that was so racist? Their bosses are literally Jewish as they work in Hollywood and it's 99% Jewish" (KHA-IN[20231013]). Another of his posts triggered similar reactions:

"The head of Insta is Israeli;"

"unfortunately they control everything, literally, the satellites high up in the orbit and everything else…" (KHA-IN[20231020]).

While comments containing aggressive speech acts did not represent the most common linguistic categories in the dataset, their share increased compared to our previous findings. Some direct DEATH WISHES and CURSES appealed to higher powers: "Let hammas finish isreal…!!" (EYE-IN[20231020]), "[A:] Why can't we use nuclear bomb to erase ISR*EL from this planet? [B:] "God will do just wait 👋 no more israel zionist on this beautiful earth" (EYE-IN[20231020]). Others were more opaque: "you should've been in the festival the other day getting your legs in the air real good, but it's okey next one you wouldn't be posting here 😜" (IME-IN[20231013]). Jews, Israelis or Israel were also targets of **THREATS** such as "Your baby killing devil worshipping government is getting what it deserves soon" (IME-IN[20231020]).

2.4.2 France

The French dataset[28] articulated responses to the two distinct scenarios: Israel as a victim of a large-scale and devastating attack, and Israel on the offensive again, in the context of mounting Palestinians casualties and a looming humanitarian crisis in Gaza. These two scenarios map onto traditional core narratives of antisemitism – Jewish guilt for their own suffering, and Jews as ontological aggressors threatening the in-group's existence. **16% of all comments were of antisemitic nature**, with no significant variance between pro-Palestinian and pro-Israeli threads.

Very similar to the British corpus, **the accusation of GENOCIDE (19.5% of antisemitic comments) and the DENIAL OF ISRAEL'S RIGHT TO EXIST (18.3%) are the most frequently articulated by web users, alongside the concept of Jewish EVIL (11.6%). 10.8% of antisemitic comments also included direct or indirect AFFIRMATION OF AND/OR CALLS FOR VIOLENCE** against Israeli civilians. This showcases the shared ideological grammar of Israel-related antisemitism on British and French social media, which coalesces around the demonisation and delegitimation of Israel, ultimately leading to an appeal for total destruction of the Jewish state.

The concept of GENOCIDE, which has deep roots in anti-Israeli rhetoric, is often conveyed explicitly: "We should not be afraid or ashamed to call it a genocide" ["faut pas avoir peur ou honte de dire que c'est un genocide"] (YAEL-IN[20231022]); "It's the year 2023, we are witnessing the biggest genocide and people like you are justifying it???" ["On est en 2023, en train d'assister au plus grand genocide,

et des gens comme vous, êtes en train de justifier ce genocide???"] (YAEL-IN[20231022]); "I scream, end the genocide and the child massacre!" ["je crie stop au genocide et au massacre d'enfants !"] (LEMON-IN[20231103]). Rhetorical questions are also used to emphasise the alleged genocidal nature of the Israeli military offensive: "Yes, genocide. How else can we call the deliberate extermination of an entire population?" ["Oui , génocide. Ou comment peut-on appeler cela l'extermination délibérée d'un peuple entier?"] (LEMON-IN[20231103]).

A similar strategy is enacted through the **NAZI ANALOGY**, which recasts Israel as the perpetrator of a new Holocaust. The comparison allows for implicit genocide accusations such as "Netanyahu is repeating the same actions as Hitler" ["Netanyahou reproduit les mêmes actes qu'Hitler"] (YAEL-IN[20231022]) or "Zionists have the same problem as our moustachioed friend, they won't be able to kill all their targets. Too bad" ["les sionistes ont le même proble que notre ami à moustaches, ils vont pas pouvoir tuer toutes leurs cibles dommage"] (ERIC-IN[20231007]). The latter comment builds on two allusions. Firstly, the moustache is easily understood as a metonymic reference to Adolf Hitler; secondly, the comment implies Israel shares a similar genocidal intent to the Nazi regime, and will run into the same 'logistical' issues. Comments also resort to wordplay and portmanteaus to construct a parallel between Nazi and Israeli actions:

"There would not be enough space for the photos and names of all Palestinians killed by Israelnazi settlers"

["si on devait mettre les photos et nom de morts Palestiniens par les colons Israëlnazii il y aurait pas assez de place"] (LEMON-IN[20231103]).

Others imply that there is a direct ideological continuity between Nazism and Zionism: "It must be said loud and clear that Nazism didn't disappear, it has evolved into Zionism" ["Il faut dire clairement que le nazisme n'a pas disparu, il a muté en sionisme"] (CLEM-IN[20231114]).

COLONIALISM ANALOGIES are another prominent device conjured up to delegitimise Israel. Parallels are drawn between French-occupied Vietnam in the 1950s ("Remember the Vietnam war, when the Việt Minh movement wanted to get rid of the French, were they terrorists also? They just wanted their land back" ["Rappelez vous de la guerre de vietnam quand le mouvement viet minh voulait se débarrasser des colonies françaises ce sont des terroristes aussi ? Ils voulais juste récupérer leurs terre !"] (HUGO-IN[20231025])) or Algeria in the 1960s ("True hypocrites, all those who defend their lands are terrorists, like the FLN, you said they were terrorists too" ["Des vrais hypocrites, tous ceux qui défendent eurs terres on c'est des terroristes, comme la FLN aussi que vous disiez qu'ils étaient terroriste"] (LESRE-IN[20231007])) and the situation in present-day Middle East. This frames the Hamas-Israel conflict as an anti-colonial struggle between an oppressed indigenous population and Western invaders. **The portrayal of Israelis as FOREIGN and rootless**, entrenched in traditional antisemitic imagery of the Wandering Jew, also fits into such framing:

> "Yessss, it's too much because even the Torah says that Jews have no land"

["Ouiiii c'est trop car meme dans la thora (sic) elle-même il est indiqué que les juives n'ont pas de terre"] (ELIE-IN[20231021]).[29]

The comment ends with **a curse conjuring the destruction of "all Israelis"** by a natural catastrophe: "I hope a natural phenomenon wipes out all Israelis and the Palestinians will finally get their land back" ["J'espere qu'un phénomène naturel balayera tout les israélien et les palestiniens pourront enfin retrouver leur terre"] (ELIE-IN[20231021]).

Another comment disparages Jews as coming to the Middle East as "homeless bums," chased out of Germany and Europe ["à la base ils sont venus comme des SDF parce que l'Allemagne les a mit a la porte"] (MELEN-IN[20231027]). The logical conclusion of such language is a **DENIAL OF ISRAEL'S RIGHT TO EXIST**: "Israel is a made-up country. Its people is a made-up people" ["Israël est un pays factice . Son peuple est factice"] (ERIC-IN[20231007]). In this context, commentators dispute the terrorist nature of the 7 October attacks, presenting Hamas's actions as "heroic resistance" ["résistance héroique"] (ELIE-IN[20231021]) against the colonisers. Any parallels with other terrorist attacks are rejected:

> "10/7 is not a 9/11, as they try to make us believe. It happened in a context of colonialism and apartheid"

["Le 7/10 n'est pas un 9/11 comme on essaye de nous le faire gober. On est dans un contexte de colonisation et d'apartheid"] (MDIPLO-IN[20231102]).

Such **AFFIRMATION OF VIOLENCE** against civilians can be coupled with the topos of **COLLECTIVE RESPONSIBILITY of Israelis:** "I don't consider any Israeli to be a civilian: they all serve in the military, oppress Palestinians, the kids are being indoctrinated almost since birth, have access to guns, so no" ["Je ne considère aucun Israélien comme

civil : ils font tous l'armée, tapent tous sur les Palestiniens, les enfants sont endoctrinés quasi dès la naissance, ont accès aux armes, alors non"] (LESRE-IN[20231007]).

There is a clear synergy between anti-Israeli and classical antisemitic tropes such as **JEWISH POWER AND CONSPIRACY**. On French social media, the main focus of such accusations was the French president Emmanuel Macron. Despite France's relatively moderate stance on the conflict and its staunch support for a two-state solution, Macron is portrayed as a servile lackey to shadowy networks of Jewish influence. His former employment as an investment banker at Rothschild & Co is often highlighted, in a manner that activates a century-old antisemitic imagery: "You really believe the former Rothschild banker is going to condemn Israel??" ["Mais sincèrement vous pensez que l'ex-banquier des Rotschild va condamner Israël??"] (AMNE-IN[20231024]); "They are all guilty, all complicit! Of course, he pledges support, a son never turns against his mother, especially if she breastfeeds him!" ["Tous coupables tous complices ! Évidemment qu'il apporte sont soutien un fils se retourne jamais contre sa mère surtout si celle ci lui donne le sein et le bibérons !"] (LESRE-IN[20231007]). The image of Macron as a product of Rothschild, "breastfed" by the powerful family, has particular connotations, as a common trope in antisemitic visual grammar represents the Rothschilds as as a large sow feeding her 'piglets,' namely loyal governments or agencies.[30] Macron is further described as a "pathetic lackey of Zionist terrorism" ["Minable soumis au terrorisme sioniste"] (LESRE-IN[20231007]), while other comments allege he is "too afraid to condemn the real terrorist attacks" against Palestinians ["trop peur de dénoncer les vrais attentats terroristes"] (LESRE-IN[20231007]).

The Zionist Occupied Government (ZOG) conspiracy theory is activated when one web user outlines what they think the 'real' government of France would look like, citing mainly Jewish names or figures known as very pro-Israeli:

"There is a government in France whose existence we are not aware of: I can imagine it very well: Drahi/Bolloré as minister of propaganda, Zemmour – civil war, Bernand Henri-Lévy – foreign affairs… Hanouna as minister of brainwashing the sheeple, Attali as minister of the interior, Mr Rothschild as minister of finances, and Macron undoubtedly as the accountant…"

["Son rôle est clair … ya un gouvernement en France dont on ne connaît même pas l'existence : Je l'imagine très bien : Drahi - Bolloré ministre de la propagande, Zemmour ministre de la guerre civile, BHL ministre des affaires externes…. Hanouna ministre de l'anesthésie cérébrale sur le bétail, Attali ministre de l'intérieur, compagnir Rotchild ministre des finances, et Macron l'expert comptable sans le moindre doute …."] (LESRE-IN[20231007]).

30 – American Jewish Committee – Translate Hate. "Rothschild," https://www.ajc.org/translatehate/Rothschild (last accessed on 16 January 2024).

2.4.3 Germany

While in the past the Hamas attacks on Israeli civilians triggered few reactions in the German-speaking community, the events of 7 October generated a huge response on *Instagram*. Within the dataset of 3,000 comments, **16.8% were identified as antisemitic** – significantly more than on *Facebook* for the entire analysed period. In order to cover the developments of the conflict in its entirety and to ensure comparability with the *Facebook* analyses, the corpus included comments sections from the period between 9 October and 4 November. While their political stance cannot always be clearly determined, we aimed to maintain a balance between the different voices, with 15 pro-Israeli and 15 pro-Palestinian threads examined for our analysis.[31] While some threads attempt to explain and contextualise the conflict and the subsequent debates – such as the position of the German government, or Fridays for Future activist Greta Thunberg – other threads merely expressed support for one of the parties in the conflict or highlighted the suffering of all victims, so that the proportion of antisemitic comments in the individual threads ranged from 1 to 42%. The largest proportion of antisemitic comments could be identified in the pro-Palestinian threads.

Although the percentage distribution differs, the most frequently expressed concepts are similar to the findings in the British and French datasets, as well as the results of the *Facebook* analyses in the previous sections. **The most prevalent antisemitic concepts were accusations of EVIL (26.7% of antisemitic comments), GENOCIDE (16.8%), AFFIRMATION OF VIOLENCE (11.9%), FREE PASS (9.7%) and INFLUENCE (9.7%). However, these tropes often co-occurred with each other, as well as with concepts such as antisemitic CONSPIRACIES (7.5%) and TERRORIST STATE (6.1%),** and their frequency depended on the framing of the original post.

This is particularly evident in the following example, which combines a wide variety of antisemitic ideas: "Supporting Israel's right to exist means that you support imperialism, colonialism, apartheid (Amnesty International, Human Rights Watch, B'Tselem), genocide, expulsion, etc. Shows very well where you stand. Germany must stop blaming Palestinians for their past" ["Existensrecht Israel zu unterstützen bedeutet, dass ihr Imperialismus, Kolonialismus, Apartheid (Amnesty International, Human Rights Watch, B'Tselem), Völkermord, Vertreibung usw. unterstützt. Zeigt sehr gut wo ihr steht. Deutschland muss aufhören ihre Schuld aus der Vergangenheit den Palästinensern in die Schuhe zu schieben"] (FFFG-IN[20231020]). The comment links Israel's right to exist to negative attributions, consequently calling that right into question: **by presupposing the APARTHEID ANALOGY, IMPERIALISM ANALOGIES and GENOCIDE, ITS RIGHT TO EXIST CAN BE NEGATED insofar as the attributions meet with social rejection**. At the same time, the comment alludes to the Holocaust when insinuating that Germany's support for Israel is solely due to historical guilt, which supposedly blinds it to Palestinian suffering in the present.

The narrative of Israel as an **absolute EVIL** pervades the entire discourse and is mostly used to ascribe Israel with wilful and deliberately cruel intentions towards the Palestinians. The Jewish state is either accused of inhumane behaviour:

> `"It is obvious that 'Israel' does not consider human life worth protecting"`
> ["Es ist doch offensichtlich das Israel' menschliches Leben nicht für schützenswert hält"] (VICE-IN[20231018]),

or charged with active interest in causing as much suffering as possible: "Of course the Israelis care about the Palestinians. They think every day how best to make the people in Gaza suffer" ["Natürlich sind die Palästinenser den Israeliten nicht egal. Die machen sich tagtäglich Gedanken darüber, wie man die Menschen im Gaza am besten leiden lassen kann"] (TAREK-IN[20231021]), or simply essentialised as an undemocratic **TERRORIST STATE** under which the Palestinians have to suffer on a daily basis: "It's terror! It is bad! And without mentioning a but, the long list of names of the innocent, slaughtered 3,000 children in Gaza is also terror! Terror of a government that is being sold here as democratic" ["Es ist Terror ! Es ist schlimm! Und ohne ein Aber zu nennen, die lange Namensliste der unschuldigen, abgeschlachteten 3000 Kindern in Gaza ist ebenso Terror! Terror einer Regierung die hier als demokratisch verkauft wird"] (DUZEN-IN[20231020]). These alleged efforts to cause hurt usually culminate in insinuations that represent a modern classic of antisemitic ascriptions, namely **the claim that Israel is carrying out a GENOCIDE**, which, similarly to the findings in the French corpus, is often communicated openly and explicitly "Israel is committing genocide against the Palestinians!" ["Israel begeht einen Genozid an den Palestinänsern!"] (VOLKS-IN[20231020]).

This is supposedly accepted without reservation by the international community. Some comments suggest that "Israel has foolproof freedom with the West. The political elite will not condemn Israel, no matter what it does" ["Israel hat beim Westen narren Freiheit. Die politische Elite wird Israel egal was sie macht nicht verurteilen"] (JUERG-IN[20231017]), reiterating the idea that **Israel and/or Jews have a manipulative INFLUENCE** (or even absolute control) over politics and the public: "We can rename ourselves the Federal Republic of Israel. Or hand over the login details of the government-run social media channels to the IDF. Then you save yourself the work of emails and phone calls"

["Wir können uns umbenennen in Bundesrepublik Israel. Oder übergibt gleich die Login-Daten der regierungsgeführten Social Media Kanäle an die IDF. Dann spart man sich die Arbeit mit Emails und telefonieren"] (TAREK-IN [20231021]). **The prominence of the alleged FREE PASS** in the German corpus, compared to the *Facebook* analyses as well as the British and French analyses, could be explained by the fact that many of the pro-Palestinian threads date from the end of October 2023 – at a time when Israel's military interventions were already advanced. It seems that such comments take this as an opportunity to unrestrictedly attribute exaggeratedly reprehensible behaviour to Israel, and justify these attributions with the concept of a FREE PASS.

All these ascriptions, which are used to condemn the actions of Israel, ultimately serve to legitimise the violence of Hamas, so that the **AFFIRMATION OF VIOLENCE**, which appears prominently in the entire discourse from this period, is fed by the network of concepts described above:

"The years of inaction and ceasefire have only led to innocent Palestinians being murdered and bombed almost daily and further unjustly occupied, while the West has predictably played its three monkey role… Unfortunately there is only one language that Israel and the Zionists understand"

["Die jahrelange tatlosigkeit und Waffenstillstand hat nur dazu geführt das fast täglich unschuldige Palästinenser ermordet bombardiert und weiter zu Unrecht mehr besetzt wurden während der Westen wie erwartet ihre 3 Affen Rolle gespielt haben.. Es gibt leider nur eine Sprache die Israel bzw die Zionisten verstanden werden"] (JUERG-IN[20231017]).

2.5 Qualitative analyses of responses in other countries

In order to broaden, diversify and enrich the scope of our regular research, we carried out analyses of six further language communities and their online reactions in the immediate aftermath of the Hamas attacks.[32] The six datasets were collected from the *Facebook* pages of mainstream media outlets in Italy, Spain, Poland, Slovakia, Romania and Morocco, directing our focus towards articles posted between 7 and 10 October 2023 to examine the spontaneous reactions of web users to these initial events. Each dataset comprises 1,000 user comments (100 first comments from each of ten threads); we followed the usual project methodology to annotate and classify the data. The additional six thousand comments and their analyses presented below are not just a brief glimpse into the unique discourses of each country, but also provide material for observing the commonalities of antisemitic discourse across languages and cultures. With this extended contrastive dimension to the 7 October terror attacks, our corpus analysis represents the largest investigation so far into Israel-related antisemitism online during the conflict's current escalation phase.

2.5.1 Italy

Laura Ascone

With only 9.7%, the Italian data[33] contained significantly fewer antisemitic responses than the other language communities analysed in this report. However, while the percentage of antisemitic comments is lower in the Italian dataset, the distribution of the concepts expressed follows the overall tendency identified in the other countries. The analysed comments present either an antisemitic concept, typically aimed at Israel, or the expression of support for Hamas, which – as stated before – represents a form of antisemitism. In the ten threads

under investigation, the percentage of antisemitic reactions ranges from 5% (*Il Messaggero*) to 17% (*Corriere della Sera*).

Frequently, commenters refer to ISRAEL'S SOLE GUILT IN THE CONFLICT (19.6%), presenting it as the only party responsible for the terrorist attacks. When comments accuse Israel of sole responsibility for the conflict, they tend to refer to the Israeli actions which supposedly led to or justified such a violent attack: "Stop giving money to Israelis, who provoke Muslims on a daily basis" ["Basta dare soldi ai fascisti israeliani che quotidianamente provocano i musulmani"] (LIBER-FB[20231009]); "Israelis are paying for the years of abuses on Palestinians... sorry for the dead... but Israel is the one responsible of all this bloodshed" ["Gli israeliani pagano anni di soprusi contro i Palestinesi ....dispiace per i morti... ma Israele ha tutte le colpe di questo sangue versato..."] (LASTA-FB[20231008a]). As mentioned above, this antisemitic concept sometimes leads to the justification of Hamas's attacks: "The stinkers have been killing the beasts[34] for seventy years... In the end the beasts have reacted. Easy..." ["Da settant'anni che le puzzolente stanno uccidendo le belve ..Alla fine le belve hanno reagito .facile"] (LIBER-FB[20231009]); "what goes around comes around, it's called karma" ["quello che di fa si riceve, si chiama Karma"] (LASTA-FB[20231008b]).

---

32 – Schwarz-Friesel and Reinharz analysed antisemitic statements in letters and emails to Jewish and Israeli institutions across seven European countries: Austria, Switzerland, The Netherlands, Spain, Belgium, England, Ireland, and Sweden. Schwarz-Friesel, Monika/Reinharz, Jehuda, 2017. Inside the Antisemitic Mind: the language of Jew-Hatred in contemporary Germany. Chicago: University of Chicago Press, 194–203.

33 – The Italian corpus consists of the first 100 comments posted in reaction to 10 articles shared on the *Facebook* pages of five Italian mainstream media (*La Repubblica, Corriere della Sera, Il Messaggero, La Stampa,* and *Libero*). Two threads per outlet were analysed.

34 – The commenter uses the same term used in the title of the article. However, they use it in an ironic way, implying that the real beasts are the Israelis.

As we have seen in the previous sections, **instead of directly attacking Israel some users ᴀꜰꜰɪʀᴍ ᴛʜᴇ ᴠɪᴏʟᴇɴᴄᴇ of Hamas's attacks (18.3%)**. Most of the time, this is done through the expression of support for Palestine, which constitutes here a metonym for Hamas and its attacks and, as a consequence, a form of antisemitism: the slogans "Go Palestine" ["Forza Palestina"] and "Free Palestine" ["Palestina libera"] appear in several threads. In some cases, this form of antisemitism is verbalised in an explicit way: "Go Hamas" ["W Hamas"] (LIBER-FB[20231008]); "if Hamas coordinates the fight against Israel, may it come!" ["se Hamas coordina il contrasto a Israele, ben venga!"].

Commenters justify the Hamas attack by attributing malevolent, ᴇᴠɪʟ ᴄʜᴀʀᴀᴄᴛᴇʀ to Israel (16.5%). They either demonise Israeli actions themselves ("Israel has been bombing Palestine for 70 years" ["Sono 70 anni che Israele colpisce la Palestina a suon di bombe sulla città"] (REPUB-FB[20231008])) or compare them to the Hamas massacre on 7 October, thereby both downplaying and legitimising the latter: "Israel has been doing this for 80 years…" ["Israele lo fa da 80 anni…"] (REPUB-FB[20231009]).

Another way of demonising Israel is through the ɴᴀᴢɪ ᴀɴᴀʟᴏɢʏ (11.3%) which, again, is verbalised in two different ways: focusing on Israel's alleged nature ("Israel is Nazism," ["Israele è il nazismo"] (MESSA-FB[20231007])) or comparing Israel's actions to those of the Nazis ("That's incredible how Jews who have suffered Hitler are behaving like their persecutor" ["Incredibile come gli ebrei che hanno subito Hitler si comportano come il loro persecutore"] (CORRI-FB[20231009])). In some cases, not only is Israel compared to the Nazis, but presented as being even worse: "Even the Nazis didn't" ["Roba che manco i nazisti"] (REPUB-FB[20231008]).

Furthermore, **Israel is also described as a ᴛᴇʀʀᴏʀɪsᴛ sᴛᴀᴛᴇ (10.3%), which in some cases leads to the ᴅᴇɴɪᴀʟ ᴏꜰ ɪᴛs ʀɪɢʜᴛ ᴛᴏ ᴇxɪsᴛ (9.2%)**. The analysis shows that, throughout the dataset, the focus is on the demonisation of Israel more than on the praise of Hamas.

2.5.2 Spain

Iael Kurjan

Within the Spanish *Facebook* dataset, 17% of analysed comments were classed as antisemitic.[35] Much of the commentary both explicitly and implicitly justified the Hamas attacks as a legitimate response to previous Israeli actions – claiming the violence represented Palestinian self-defence, due to Israeli oppression, in order to frame the notions of victim/oppressed vs guilty/oppressor.

By far the most frequently expressed antisemitic concept was the accusation of Israel being a ᴛᴇʀʀᴏʀɪsᴛ sᴛᴀᴛᴇ (16.5% of all antisemitic comments). According to such comments, Israel massacres or exterminates Palestinians: "The terrorist state of Israel exercises the occupation and extermination of Palestine" ["el estado terrorista de Israel ejerce la ocupación y exterminio de Palestina"] (ABC-FB[20231010]). **Some web users combined the ᴛᴇʀʀᴏʀɪsᴛ sᴛᴀᴛᴇ accusation with indirect forms of the ɴᴀᴢɪ ᴀɴᴀʟᴏɢʏ,** e.g. by equating Gaza with Auschwitz:

"the terrorist state of Israel is carrying out a total extermination in Gaza, Gaza has been converted into Auschwitz"

["el estado terrorista de Israel, ejerce un total exterminio sobre Gaza, Gaza la han convertido en un Auschwitz"] (ABC-FB[20231010]).

35 – 1,000 user comments posted on *Facebook* profiles of five Spanish mainstream media outlets were annotated: *El País, CNN en Español, ABC España, El Mundo* and *La Vanguardia*.

Instances of the NAZI ANALOGY comprised 9.4% of all antisemitic comments: "Jewish Zionists repeat with the Palestinians what they suffered at the hands of the Nazis" ["Los sionistas judíos repiten con los palestinos lo que ellos sufrieron a manos de los nazis"] (ABC-FB[20231007]), or "You can tell that Israelites have learned a lot about Nazism during the Second World War. Then they play victims around the world" ["Se ve que los israelitas aprendieron mucho del nazismo de la Segunda Guerra Mundial. Luego van de víctimas por el mundo"] (ELPAI-FB[20231007]). This comment uses irony to imply that JEWS HAVE NOT LEARNED FROM THE PAST, while others state that Israelis' "DNA was contaminated with Nazism" ["su ADN se contaminó con el nazismo"] (LAVAN-FB[20231010]).

The third most frequent concept was the accusation that Israel is committing GENOCIDE on the Palestinians (8.8%): "cowards are those who enter a land that is not theirs and commit genocide, they corner them and expel them, Jews will always be hated by humanity, very brave now but in Berlin in 1944 they pooped their pants" ["cobardes es el que se mete en una tierra que no es suya y le hacen genocidio, lo acorralan y los expulsa, los judíos siempre serán odiados por la humanidad, muy valientes ahora pero en Berlín en 1944 se hacían caca en los pantalones"] (ABC-FB[20231010]). As in this case, many GENOCIDE comments incorporated the NAZI ANALOGY, along with references to child abduction or even deliberate CHILD MURDER (6.5%):

"It is what the 'chosen ones' sought as God's people; massacring our Palestinian brothers and what is worse, imprisoning children. No one justifies violence, but these Israelis are worse than Nazis"

["Es lo que buscaron los 'elegidos' como pueblo de dios; masacrando a nuestros hermanos palestinos y lo que es peor encarcelando a los niños. Nadie justifica la violencia, pero estos israelíes son peores que los nazis"] (ELPAI-FB[20231007]).

Although this *Facebook* comment does not per se amount to CHILD MURDER, the commenter claims the deliberate imprisonment of children and thus perpetuates a centuries-old prejudice.

It is also worth noting that a **considerable number of comments expressed indirect AFFIRMATION (15.3%) of the Hamas attacks via generalised support for Palestine**, with slogans such as: "Resist, free Palestine" ["Resiste Palestina libre"] (CNN-FB[20231007a]). **Others used the Hamas attacks as an opportunity to deny ISRAEL'S RIGHT TO SELF-DETERMINATION (5.3%).** Antisemitic comments often conveyed the idea that Zionists are invaders and that they must leave Palestinian lands – not just the West Bank but the entire region. Israel was widely referred to, in essentialised terms, as an "occupier." Throughout the dataset, there was a clear association between the generalised use of 'occupier' and terms such as 'invasion': "Israel, the invader and colonialist in the 21st century" ["Israel invasor y colonialista en pleno siglo XXI"] (ELPAI-FB[20231009]), or: "Israel is the country that invaded Palestine" ["Israel es el país que ha invadido Palestina"] (ELMUN-FB[20231009]).

2.5.3 Poland

Karolina Placzynta

In the *Facebook* comments sections of Polish news outlets, antisemitism levels averaged 17.7%, reaching a high of 38% under a news post describing the events at the Supernova festival.[36] In both this thread and the entire dataset, evenly distributed across all outlets, **the top concept was CONSPIRACY (19.8% of all antisemitic comments), as web users refused to accept Hamas attacks had happened without Israel's tacit agreement or even instigation**: "This action is a provocation by Mossad, there was a threat of civil war in Israel so they had to react" ["Ta akcja to prowokacja Mosadu, w Izraelu groziła wojna domowa więc trzeba było zareagować"] (FAKT-FB[20231008a]). Such comments also linked the supposed "provocation" with a wish to "destroy Palestine... who created hamas if not israel..." ["zniszczyc palestyne... kto stworzył hamas jak nie izrael...] (INTER-FB[20231008]). Other comments in this category often referenced the 'Heavenly Jerusalem' project – a Jewish resettlement conspiracy theory:[37]

"Just as planned. Ukrainians will flood Poland and Jews will settle in Ukraine. The global plan is slowly coming to fruition 🤬🤬🤬"

["Czyli tak jak miało być . Ukraińcy zalewają Polskę, a Żydzi zamieszkają na Ukrainie.Globalny plan pomału się realizuje 🤬🤬🤬"] (RMF-FB[20231008]);

36 – The corpus contains comments from the *Facebook* accounts of the daily newspaper *Fakt*, the weekly magazine *Wprost*, online news platform *Interia*, investigative journalism platform *Oko.press*, TV station *Polsat* and two radio stations: *RMF FM* and *Radio Zet*.

37 – Similar claims have appeared in e. g. Romanian discourse (cf. 2.5.5) and German far-right spaces online: https://www.isd-global.org/digital_dispatches/an-antisemitic-conspiracy-theory-is-being-shared-on-telegram-to-justify-russias-invasion-of-ukraine/ (last accessed on 15 January 2024).

"Another false flag performed by the sidelocks. I reckon they'll make us into a second Palestine now" ["Kolejna fals flag w wykonaniu pejsatych. Mysle że z nas teraz zrobią 2 Palestyne"] (FAKT-FB[20231008b]). The condescending reference to Jews/Israelis used in the latter example reflects the **relatively high share of slurs present** in the corpus (4.5% of antisemitic comments), compared to the UK, France or Germany.

The second most common antisemitic trope in the corpus (18.6%) also accused Israel of nefarious intent, portraying it as EVIL or IMMORAL – either throughout its entire history: "Israel has been murdering Palestinians for years, the world is silent!!" ["Izrael od lat morduje Palestyńczyków, świat milczy!!"] (POLSAT-FB[20231007]), or on a global scale: "Israel is the cause of all evil in the world" ["Izrael to przyczyna całego zła na świecie"] (ZET-FB[20231008]). This sometimes spilled over into claims that **Israel is a TERRORIST STATE**: "There are no bigger terrorists than those under the Star of David. A viper tribe" ["wiekszych terrorystów niż spod gwiazdy Dawida to nie ma. Żmijowe plemię"] (RMF-FB[20231008]), here reinforced with a **DEHUMANISING insult and its biblical connotations of TREACHERY or DISLOYALTY** (though such religious references were rare, with comments often attacking religion as a source of conflict or ignorance, echoing the current anti-Catholic sentiments in the society). Some constructed **Israel as GENOCIDAL** "on a par with the Hitlerites. The same methods of extermination" ["na rowni z hitlerowcami. Te same metody eksterminacji"] (POLSAT-FB[20231007]). **The NAZI ANALOGY (9% of antisemitic comments)** was expressed with a mere allusion by this reader of the left-leaning *OKO.press* "Adolf's diligent students" ["Pilni uczniowie Adolfa"] (OKO-FB[20231008]), and a direct accusation in a user comment from the centre-right *Wprost*: "Literally a slow Holocaust. But they're Jews, and Jews are allowed to" ["Dosłownie polowolny holokaust. Ale to żydzi, a żydom

wolno"] (WPROST-FB[20231008]). Here, the sarcastic punchline hints at the **stereotype of a FREE PASS supposedly enjoyed by Jews and/or Israel (6%)**; such comments are a clue to the lingering perceptions of Jews as untrustworthy and privileged, despite the contemporary attempts to dismiss antisemitism as either a thing of Poland's past, or only ever a problem of other nations.

Alongside the conspiratorial explanation behind Hamas attacks, some comments framed them as simply well-deserved, BLAMING Israel either for the attacks or FOR THE WHOLE CONFLICT (13% of all antisemitic comments). This category particularly strongly correlated with a lack of empathy (although this was found in non-antisemitic comments too), here succinctly conveyed via the idiom "[w]ho sows the wind reaps the storm" ["Kto sieje wiatr,ten zbiera burze"] (FAKT-FB[20231009]). **Some web users cynically AFFIRMED THE ANTISEMITIC VIOLENCE (7% of antisemitic comments):** "Israel has lived to regret these decades of murdering Palestinians. Lord, you're not swift, but you are righteous!" ["Doczekał się Izrael za te dziesiątki lat mordowania Palestyńczyków. Panie Boże jesteś nie rychliwy ale za to sprawiedliwy!"] (POLSAT-FB[20231007]). Some comments in this category suggested the attacks would not have happened had Israel not "created a fictitious state on Palestine's land, this illegal creation was legalised thanks to the USSR and the USA" ["stworzył fikcyjne państwo na terenach nalezacych do palestyny I dzięki ZSRR oraz usa ten nielegalny twór został zalegalizowany"] (INTER-FB[20231008]). **Another comment similarly DENIES ISRAEL'S RIGHT TO EXIST, painting a contemptuous picture of its citizens as eternally homeless and ALIEN** "[t]here are no Israelis, there are Jews. They have nothing of their own, because wherever they appeared they have been chased away" ["nie ma Izraelczyków, są Żydzi. Nie mają nic swojego, bo gdziekolwiek się zjawili to ich przeganiano"] (ZET-FB[20231008]).

2.5.4 Slovakia

Veronika Bundzíková

The average level of antisemitism in the *Facebook* comments of Slovak media outlets stands at 14.9%, even though it reaches up to 23.7% in some threads.[38] Compared with some West European countries, pro-Palestinian comments that would in fact disguise the support of Hamas's attack only rarely occur in the comments sections of the analysed Slovak media outlets. Also, the explicit support of the atrocities perpetrated by Hamas is present in a very limited scope; only one user reacts by posting: "Great, fingers crossed for Palestine, vivat Palestine" ["Super držím palce Palestíne, vivat Palestína"] (NOVYC-FB[20231008]). Instead, the pro-Palestinian sympathies are disclosed as a by-product of Israel's vilification. They usually suffer from oversimplification of the Middle Eastern political context, framing the Israeli-Palestine conflict in a black-and-white manner. **Israel is portrayed as the IMMORAL aggressor (6%) or a PRIVILEGED state that evades any critique for its misconduct thanks to holding a FREE PASS for its actions (10%).**

Occasionally, **parallels between the Russian-Ukrainian war and the situation in Israel** are drawn in the Slovak antisemitic comments. This might be enhanced by the fact that Slovakia is Ukraine's neighbour and has been affected by the war looming at its borders. One user states: "When the Russians took over Ukrainian soil, the whole world condemned them but nobody condemned Israel, something stinks here" ["Keď Rusi zabrali Ukrajine územie skoro cely svet ich odsúdil Izrael nie niečo tu smrdí"] (NOVYC-FB[20231008]). Also, a handful of comments (5%) equate the situation in Palestine right after the Israeli counterattack with the suffering of Jews during the Holocaust, which represents an **indirect NAZI ANALOGY**: "One French journalist noted that it reminds him of the

38 – The Slovak dataset contains comments replying to posts published by the following media outlets: *Denník N, SME, Aktuality, Pravda, Plus Jeden Deň, Topky* and *Nový Čas.*

liquidation of the Jewish ghetto in Warsaw by the fascists!" ["Jeden francúzsky novinár sa vyjadril, že mu to pripomína likvidáciu židovského varšavského gheta fašistami!"] (AKTUA-FB[20231009]).

At the conceptual level, strongly represented are the attacks based on denying Israel's statehood and the Jewish RIGHT FOR SELF-DETERMINATION. This idea is expressed in 40% of the comments classified as antisemitic. The users either deny Israel's historical right to the land ("Unfortunately, Israel stole this land from Muslims. They used to live there in harmony with Christians. Then Jews came, and that was the end" ["bohužiaľ Izrael moslimom ukradol tuto zem. Oni si tam kedysi s kresťanmi v súlade žili. Potom prišli židia a bol to ich koniec"] (PLUSJ-FB[20231009])), or frame the Israelis as occupiers ("Jews have been occupying these lands and killing Palestinians for decades, and the world is quiet") ["Židia desaťročia okupujú územia a vraždia Palestínčanov a svet je ticho"] (NOVYC-FB[20211008]). **Some CONSPIRACY-led comments (4%) highlight the role of the USA in the establishment of Israel,** which would not have existed otherwise. **Such comments, though rare, are closely related to the accusation that Israel is a FOREIGN or DISINTEGRATING element in the region (10%):** "because they are native, in contrast to the few millions of these Talmudic pseudo-Hebrews gathered from all around the world, who occupy it there with their artificially established state Israhell" ["lebo odtiaľ pochádzajú, na rozdiel od tých niekoľko miliónov talmudických pseudohebrejcov pozbieraných z celého sveta, ktorí to tam okupujú ako bol umelo vytvorený štát Izrahell"] (PRAVD-FB[20231008]).

For users posting such comments, Israel has no justification for its existence and hence is not worth any pity after the attack on 7 October. Such argumentation is augmented by **the belief that the Jews (or Israel) are the ultimate EVIL (18.8%).** These statements overblow the duration or the scope of their misdeeds, with the following example demonstrating how anti-Israeli and anti-Jewish sentiments may overlap: "They occupy this land and oppress the Palestinians since 1947! This should stop, and Jews should stop spreading evil in the world as they have been doing since the Old Testament until nowadays!" ["Od roku 1947 okupuju toto uzemie a utlačaju palestinčanov! Toto by mali skincit a židia by msli prestat širit zlo po svete od stareho zakona po dnes!"] (PLUSJ-FB[20231009]). This goes hand in hand with another antisemitic concept that appears in the comments: **BLAMING JEWS FOR ANTISEMITISM (16%) or SOLELY ISRAEL FOR THE CONFLICT (12.8%), usually accompanied by lack of empathy and schadenfreude.** In such comments, users pretend to put the Hamas attack into the bigger picture but end up swapping the perpetrators for the victims: "It's horrible. But honestly, Israel has been perpetrating genocide of the Palestinian nation for decades. So, unfortunately, we could have expected this to happen" ["Je to strašné. Ale popravde, Izrael pácha genocídu palestínskeho národa celé desaťročia. Takže sa to žiaľ dalo očakávať"] (PLUSJ-FB[20231009]). Lastly, some comments suggesting a secret conspiracy behind the Hamas attack harbour antisemitism: **they accuse Israel or Jews of being behind the atrocities to DECEITFULLY profit from them (9.4%):**

"Maybe it's no failure of the intelligence services but a perfect accomplishment. Now, they can use the attack to eradicate 'with impunity' the Gaza Strip from the map forever"

["Mozno to vobec neni ziadne zlyhanie tajnych sluzieb, ale naopak dokonaly uspech. Teraz pod zamienkou napadnutia mozu 'beztrestne' vygumovat pasmo gazy z mapy navzdy"] (PROJE-FB[20231007]).

2.5.5 Romania

Alexis Chapelan

In the Romanian dataset,[39] **18.2% of all comments were classified as antisemitic.** This percentage showcases that the Middle East conflict is a significant trigger for antisemitism on Romanian social media. However, the nature of the discourse differs from other European contexts: Israel-related concepts are less prominent; instead, classical antisemitic tropes as well as slurs highlight a fundamental continuity between older patterns of demonisation and othering of Jews and perceptions of Israel as a 'collective Jew.'

This is reflected, for example, in the widespread use (in almost 30% of all antisemitic comments) of the term 'Jews' to describe Israelis. **This conflation of political and ethnic/religious identity not only enacts a form of COLLECTIVE RESPONSIBILITY of Jews for the Israeli government's policy,** but transfers the negative imagery linked to Jews onto Israel. **Slurs historically referring to Jews** (such as 'jidan,' roughly translated as 'kikes') are also applied to Israelis more frequently than in other countries (around 5% of all antisemitic comments). Concepts of DEHUMANISATION and REPULSIVENESS often work in synergy: "History taught us they are nothing but a bunch of parasites, even viruses have their pride, Jews don't" ["Iistoria ne-a invatat ca nu sunt doar o sleahta de paraziti!In comparatie,virusii au mandria lor,evreii nu!"] (STIRI-FB[20231007]).

One user describes Israelis in ways reminiscent of 19th century eugenics, focusing on **GREED, LACK OF CLEANLINESS** and **DISORDERLY SEXUALITY**: "They make 6–7 kids, they are loud, dirty, they litter everywhere they go, they want everything cheaper or even for free!! You should actually listen to those who had to deal with this people called Jews!!" ["Fac cate 6 -7 copii, sant galagiosi, sant murdari

si lasa mizerie pe unde trec, si vor totul cat mai ieftin sau chiar pe degeaba!! Voi cei care vorbiti din tribunal ar fi bine sa ascultati oamenii care au avut tangenta cu acest neam numit evrei!!"] (STIRI-FB[20231007]).

Jews are also accused of having had a DISINTEGRATING impact on society throughout history: "Wipe Jews out the face of the earth, they destroyed Romania" ["Săi radă de pe fața pământului pe evrei au distrus România"] (RO.EU-FB[20231007]); "Fuck off, scum! You destroyed entire countries and peoples! You will pay for it, scumbags!" ["Să vă ia dracul de nenorociți ! Ați distrus țări și popoare ! O să plătiți voi nenorociților !] (RO.EU-FB[20231007]); "They are not with God but with the Devil, that's why they only want wars to destroy humanity" ["Adtia nu mai sunt cu Dumnezeu și cu dracu deaia vor numa războaie și sa distrugă omenirea..."] (FB-DIGI[20231009]).

Current actions of the Israeli government are being resituated into a broader history of alleged **Jewish EVIL and cruelty, often through religious references**, thus creating a clear narrative that such violence is ontologically tied to Jewish nature: "You will see how generous Jews are, cheap propaganda, ever since Moses brought them from Egypt they massacred all nations living in the Promised Land, at least that's what the Bible says" ["o sa vedeți bunătatea evreilor, propaganda de doi lei, de când i-a adus Moise din Egipt au macelarit toate popoarele care trăiau in tara făgăduinței cel puțin asa scrie în Biblie"] (STIRI-FB[20231009]); "You are a clueless fool, Israel is and will forever be a bunch of murderous immigrants. [...] Don't forget who murdered Jesus! It wasn't the Palestinians!" ["Nu stiti nimic vorbiti multi sa va aflati in treaba, israel va fi si va ramane niste imigranti criminali [...] Si nu uita cine la omorat pe isus! Nu ai fost palestine!"] (STIRI-FB[20231007]).

Similar to the Polish data, the analysed threads also showcase **a wide range of antisemitic conspiracy theories (29% of all antisemitic comments)**, from the traditional allegation of global domination to the Khazar claim: "This is what Israel always did, find ways to rule the world – it's in their religion – and stoke up wars" ["Asta a stiut Israelul sa faca toata viata,sa caute mijloace sa conduca lumea – e in religia lor – si sa provoace razboaie !"] (RO.EU-FB[20231007]); "That's why they provoked the war in Ukraine, so it is abandoned by Ukrainians and these Khazars come and build the new Jerusalem" ["De aceia au provocat așa zisul război în Ukraina ca să fie părăsită de urceainieni și să vină ei să construiască noul Ierusalim acești kazari."] (STIRI-FB [20231009]). **A prominent conspiracy theory also asserts that the 7 October massacre was an Israeli false flag operation** to manipulate public opinion into accepting the war: "A cheap set-up, to justify to the world the expansion of satanism" ["O înscenare ordinară care să justifice în fața omenirii expansiunea satanismului"] (RO.EU-FB[20231007])".

One user draws a parallel with Pearl Harbor, thus implying Jews were also behind America's involvement in World War II: "To justify their attack on Iran and prepare and manipulate public opinion the Jews needed that. Something like Pearl Harbor in the Second World War" ["Pentru a motiva atacul impotriva Iranului si a pregati si manipula opinia publica in favoarea razboiului evrei aveau nevoie de asta. Ceva ca in al II -lea razboi mondial la Pearl Harbor!"] (STIRI-FB[20231007]). Some comments also express the fear that if Israel is defeated Jewish people will resettle and occupy Romania or Ukraine: "I see them massively coming to Romania, with the 600,000 which received Romanian citizenship from the traitor [President of Romanian Senate] Ciucă. I wonder if all of them are going to come here eventually, it's hard with the Arabs but Romanians are awfully docile" ["Vad ca vin masiv in Romania dintre cei 600 000 de evrei ce au primit cetatenie romana de la generalul tradator Ciuca. Ma intreb daca pina la urma vin toti, ca e greu cu arabii, romanii insa sint prosti de docili"] (STIRI-FB[20231007]).

2.5.6 Morocco

Mohamed Salhi

While Morocco does not share the same Jewish persecution history as Europe, and despite the rich Jewish culture in Morocco and the institutional efforts to engage with this history,[40] it does suffer from a significant antisemitism problem. Among the explanations for contemporary antisemitism in Morocco, argues Zhor Rehihil, the director of the Museum of Moroccan Judaism, are the lack of understanding and ignorance of Jewish history, heritage, and culture in Morocco.[41] Other factors include the centrality of the Palestinian cause to the popular and political memory,[42] the historically imported[43] and post-Zionist antisemitism (i. e. memory of the Nakba and Intifada),[44] and the frowned-upon Abrahamic accords.[45]

40 – "U.S. Special Envoy for Combatting Anti-Semitism Commends Moroccan Model in Terms of Interfaith Dialogue," *Agence Marocaine de Press*, 15 November 2023 https://www.mapnews.ma/en/actualites/social/us-special-envoy-combatting-anti-semitism-commends-moroccan-model-terms-interfaith (last accessed on 8 January 2024).

41 – Bourchachene Wail, "Anti-Semitism in Morocco: A Complicated Issue," *Morocco World News*, 11 April 2016 https://www.moroccoworldnews.com/2016/04/184021/anti-semitism-in-morocco-a-complicated-issue (last accessed on 8 January 2024).

42 – Alexandre Aublanc, "Au Maroc, l'ancien premier ministre islamiste Abdelilah Benkirane dérape dans l'antisémitisme," *Le Monde*, 24 November 2023 https://www.lemonde.fr/afrique/article/2023/11/24/au-maroc-l-ancien-premier-ministre-islamiste-abdelilah-benkirane-derape-dans-l-antisemitisme_6202163_3212.html (last accessed on 8 January 2024).

43 – The Documentation Center for North African Jews, The Jews of Morocco during World War II, https://northafricanjews-ww2.org.il/en/jews-morocco-during-world-war-ii (last accessed on 8 January 2024).

44 – Dominique Vidal, "Quand « Le Journal » dénonce l'antisémitisme," *Le Monde Diplomatique*, June 2022 https://www.monde-diplomatique.fr/2002/06/VIDAL/8946 (last accessed on 8 January 2024).

45 – Alexandre Aublanc, "Guerre Israël-Hamas : au Maroc, les tiraillements de la communauté juive," Le Monde, 5 December 2023 https://www.lemonde.fr/afrique/article/2023/12/05/guerre-israel-hamas-au-maroc-les-tiraillements-de-la-communaute-juive_6204047_3212.html (last accessed on 8 January 2024).

Of the 1,000 comments analysed in the Moroccan dataset,[46] 31.3% were classified as antisemitic. As was common in the other language communities, here again **the AFFIRMATION of Hamas's violence was the most frequently expressed antisemitic statement (47.6% of all antisemitic comments). CALLS FOR FURTHER VIOLENCE (47.6%) and DEATH WISHES (11.5%)** were also prominent, alongside antisemitic forms of **SCHADENFREUDE (12.1%)**. Examples include statements such as "I swear this is not much we want more" ["والله العظيم هذا عدد قليل نحن نريد اكثر"] ([HESPR-FB[20231009]) and the highly recurrent soundbite "May God bless and increase [the number of deaths]" ["اللهم زد وبارك"] (HESPR-FB20231007]). More extreme cases involved in-group calls for violent action against Israel and Israelis; "A call to our free and honourable brothers in the land of Morocco. Whoever among you sees a Zionist of Israeli nationality, stab him [...]" ["نداء لاخوننا الاحرار الشرفاء بارض المغرب من رأي منكم صهيونيا من حملة جنشسية الكيان فاليطعنه و ليحكم الطعنة"] (HESPR-FB[20231008a]).

Other users **expressed their joy at the events of 7 October**: "Good news 😍😍" (MWN-IN [20231007]); "ORGASM ❤️❤️❤️", and "Lets gooo" (MWN-IN [20231009]). Some made clear their **lack of empathy towards Jewish victims**: "Fock Israel, we don't care about them like they don't care about the Palestinians" (MWN-IN [20231009]). The **DEHUMANISATION of the victims** was a common feature, namely in referring to victims as animals or corpses in comments such as "600 Zionist rats were annihilated God bless it" ["تم إبادة 600 جرد صهيوني اللهم زد وبارك في ذلك"] (HESPR-FB[20231008a]) and "The hunting season started with the killing of 600 pigs" ["قد بدأ موسم القنص بقتل 600 خنزير"] (HESPR-FB[20231008a]).

Commenters sought to delegitimate Israel and to DENY ITS RIGHT TO SELF-DETERMINATION (3.2%), largely using the nominations 'entity' 14 instead of 'state' to describe Israel, or using the name Israel in quotation marks, implying its non-existence. Examples include:

"keep our flag from your false state and flag!" (MWN-IN[20231008b])

"Israel ❌ The Zionist entity ✅" ["إسرائيل ✅ ❌ الكيان الصهيوني"] (HESPR-FB[20231008b]).

Users sought to foreground the responsibility of Israel, and in some cases Jews as a whole, for the 7 October attacks: ["[...] do not mix up between jews and Zionists but Jews in Israel are not innocent") (IN-MWN 20231009]) and "There are no innocent citizens in the Israeli occupying entity, they are all armed settlers" ["لا يوجد مواطن بريء في كيان الإحتلال الإسرائيلي كلهم مستوطنين مسلحين"] (HESPR-FB[20231008b]). Additionally, there were some celebratory references to the Holocaust: "7 million + 800 😍😍😍😍😍" (IN-MWN[20231009]) and "8,000,000 gazollian jews died" (MWN-IN[20231009]). Interestingly, however, little to no references have been made to concepts which routinely appear in other language communities, such as **CONSPIRACY THEORIES**, **APARTHEID** and **NAZI ANALOGIES** and **BLOOD LIBEL**.

An interesting qualitative observation of the corpus reflects the use of the term 'Zionist' instead of 'Jew,' which may reflect the progressive awareness of the Jewish identity and the centrality of the Arab-Israeli conflict in Moroccan memory.[47] Commenters, nevertheless, tend to fail to make such distinction clear and end up in antisemitic tropes.

3. Elon Musk's online statements about Jewish individuals and organisations

Matthias J. Becker and Matthew Bolton

The following section presents four case studies involving billionaire tech mogul Elon Musk, spanning from May to November 2023. The pattern of these four micro-discourse events was similar: a) Elon Musk shared – or confirmed – a problematic statement about Jewish individuals and groups online, b) there was a backlash, c) Musk defended himself. All of this occurs at a time when the social media platform X, owned by Musk, is publicly accused of having morphed into a hotbed for antisemitic and racist sentiments and worldviews. In contrast to the case studies surrounding the current escalation phase in the Middle East, which focus (in line with the primary research focus of the Decoding Antisemitism pilot phase) on web discourse on social media profiles of European media, this investigation revolves primarily around user reactions within US media channels on YouTube, due to the American specificity of this discourse event.

3.1 George Soros

In May 2023, Elon Musk launched an incendiary attack on George Soros, drawing directly on many of the antisemitic conspiracy theories that have circulated around the latter for decades.[48] Musk accused the financier of "hat[ing] humanity" just days after Soros announced selling a modest stake in Tesla. Given that Soros's Jewish identity is common knowledge, Musk was then criticised for his dangerous rhetoric that could potentially encourage further attacks on Soros, even though Musk had not explicitly mentioned Soros's ethnicity in his statements. Later on, Musk likened Soros to the fictional villain Magneto, who, like Soros, was a survivor of the Holocaust, according to the character's comic book backstory. When another user defended Soros as having good intentions that are criticised by those who disagree with his politics, Musk responded "You assume they are good intentions. They are not. He wants to erode the very fabric of civilization. Soros hates humanity." Jonathan Greenblatt, CEO of the Anti-Defamation League (ADL), a civil rights organisation that tracks and combats incidents of antisemitism, criticised Musk's comment and said it "will embolden extremists."[49]

---

48 – For detailed analysis of Soros conspiracy theories online, see Becker, Matthias J./Troschke, Hagen/Allington, Daniel, 2021. Decoding Antisemitism: An AI-driven Study on Hate Speech and Imagery Online. Discourse Report 1. Berlin: Technical University Berlin. Centre for Research on Antisemitism.

49 – Isodore, Chris, 2023. Elon Musk claims George Soros "hates humanity." The ADL says Musk's attacks "will embolden extremists'," *CNN*, 17 May 2023, https://edition.cnn.com/2023/05/16/business/elon-musk-george-soros/index.html (last accessed on 20 December 2023).

This corpus was built from five *YouTube* threads responding to reports of Musk's comments and the reactions from the ADL and others.[50] In this dataset, **14.2% of comments were classed as antisemitic. The most frequently articulated antisemitic** expression was the **depiction of Soros as an EVIL force within the global society (23.6% of all antisemitic comments)**. Users followed Musk's lead by making comparisons between Soros and a variety of comic book and movie villains, from "Emperor Palpatine" to "Mr Sinister," "Kang" (HILL-YT[20230516]), and "The Gargamel" (FOX-YT [20230612]). These parallels are building on both perceived moral characteristics and physical features, evoking **the concept of REPULSIVENESS**, historically anchored in antisemitic imagery. Others directly claimed that Soros was "in league with Satan" (FLAG-YT[20230614]), or was "Satan spawned. The nightmare continues. Sad for all humanity" [FOX-YT[20230612]].

Direct AFFIRMATION of Musk's antisemitic comments was also prominent (22%). Web users described Musk's statements as being "spot on" (CNN-YT[20230517]) and stated that he was "not wrong," backing up their agreement with references to other "great minds" who had "eventually come to this conclusion" such as "ye michael jackson bobby fischer(a j3w) and now elon" (CNN-YT[20230517]). Other users quickly resorted to the common antisemitic claim that the young Soros had **COLLABORATED WITH THE NAZIS (12.5%)** in

order to persecute Jews. This charge was particularly prominent in response to a *YouTube* clip from the *Flagrant Clips* channel, which itself included vague allegations of Soros's power and intent to undermine American society, with comments moving beyond mere claims of collaboration to direct comparisons with Hitler and the SS:

"Anyone who thinks SS
Soros is a good guy needs
a deep down self appraisal;"

"Inexplicably, you can
criticize Adolph, but you
can't criticize George,
even though they played for
the same ghoulish team.
Weird, I know"

(FLAG-YT[20230614]).

The theme of **DISINTEGRATION** was also articulated (5%), through accusations that Soros was working **to destroy social cohesion in the US and profit from the result**: "Sow chaos, cause division, start fights between groups and own everything that they need to hate each other. Then make f**k loads of money" (FLAG-YT[20230614]).

3.2 ADL

The second event pertains to the accusation from the Anti-Defamation League (ADL) that the lack of content moderation on *X* was fuelling antisemitism. The content moderation council that Musk had promised in late 2022 has never been implemented and content moderators dismissed around the same time have not been rehired; *X* has disposed of the function allowing users to flag false information and reinstated the accounts of several figures known for posting extreme claims. As a result of these events and the subsequent rise in hate speech levels on the platform, the ADL was alleged to have reached out to advertisers. Musk threatened legal action in August 2023 and made statements suggesting that the organisation was ruining his business, claiming they were the "biggest generators of anti-Semitism" on *X*.[51] A hashtag campaign #BanTheADL sprung up, originating with the radical right, with Musk liking a post and asking his followers if he should run a poll to decide whether the ADL should be banned from *X*.[52]

The 500 analysed comments respond to two phases: firstly, Musk's threat to sue the ADL, and secondly, the reactions to this threat from ADL's CEO Jonathan Greenblatt.[53] From a quantitative perspective, it is noticeable that within the **user comments discussing Musk's threat, antisemitism is relatively low (2–12%), whereas – in contrast – the threads related to Greenblatt's reaction are rife with antisemitism (39–55%).** This implies that the media focus on a known Jewish individual was more of a catalyst for antisemitic reactions. What stands out – particularly in the comments sections with a high frequency of antisemitic content – are the contributions that AFFIRM ANTISEMITISM: "Speaking truth to liars is hateful and offensive," with statements reminiscent of debates about Kanye West:[54] "They'll say things like hate speech and call others racists but they'll never call them liars. Just because what's being said makes the ADL uncomfortable doesn't mean it's untrue" (MSNBC-YT[20230906]).

51 – Kornbluh, Jacob, 2023. Elon Musk accuses ADL of being "biggest generators of anti-Semitism" on Twitter. *Forward*, 4 September 2023, https://forward.com/fast-forward/559523/elon-musk-adl-jonathan-greenblatt-antisemitism-twitter (last accessed on 11 January 2024).

52 – Kampeas, Ron, 2023. Elon Musk amplifies call by antisemites to ban the ADL from *X. Times of Israel*, 4 September 2023, https://www.timesofisrael.com/elon-musk-amplifies-call-by-antisemites-to-ban-the-adl-from-x (last accessed on 11 January 2024).

53 – The threads were taken from the *YouTube* channels of *CNBC, The Hill, The Majority Report* and *MSNBC*.

54 – Chapelan, Alexis, et al., 2023. Decoding Antisemitism: An AI-driven Study on Hate Speech and Imagery Online. Discourse Report 5. Berlin: Technical University Berlin. Centre for Research on Antisemitism.

Users go so far as to not only AFFIRM ANTI-SEMITISM (22.5% of all antisemitic comments) but also NAZI CRIMES (5%), and express DEATH WISHES (4%): "sometimes I wish a certain Austrian painter would have won the war" (CNBC-YT[20230906]); "Tiny mustache man gassed the wrong spews" (MSNBC-YT [20230906]).

Another concept that users bring into the discourse is the INSTRUMENTALISATION OF ANTISEMITISM (19%): "Professional victim..." or "Juice will cry while striking you" (CNBC-YT[20230906]). Others rehash the notion of a TABOO OF CRITICISM (11%):

> "Mind Begs the Question:
> If Anti-Semitic is a term used
> Even against Semitic People
> If they speak against a Govt,
> Policies Anti-Semitic is term
> used for Suppression, no?"
> (HILL-YT[20230905]),

in one case alluding to the post title "We will be ferocious in fighting hate" with "Translation: We will be ferocious in fighting criticism" (MSNBC-YT[20230906]).

Also very popular is the idea of BLAMING JEWS FOR ANTISEMITISM (10%), sometimes conveyed through sarcasm and irony:

> "As the old Russian proverb goes, 'The Jew will always tell you what happened to him, but he will never tell you why'"
> (CNBC-YT[20230906]);

"People really need to stop blaming God's Chosen for the things they keep on doing." Users repeatedly mention the number 109, a reference to the "109 countries they've been kicked out of but it was never their fault. 🤮🤮🤮" - an idea originating with the white supremicist far right - **sometimes adding a threat:** "Make America 110!!!" (MSNBC-YT[20230906]).

3.3 "The Artist Formerly Known as Eric"

Elon Musk sparked another controversy by expressing support for an antisemitic post from a user calling themselves "The Artist Formerly Known as Eric," who in November 2023 tweeted

> "Jewish communities have been pushing the exact kind of dialectical hatred against whites that they claim to want people to stop using against them. I'm deeply disinterested in giving the tiniest shit now about western Jewish populations coming to the disturbing realization that those hordes of minorities that support flooding their country don't exactly like them too much."

Musk responded, saying,

> "You have said the actual truth."[55]

In subsequent statements, Musk did not retract his position but only softened it slightly, stating that he meant "some groups" like the ADL, which he claimed spread "de facto anti-white racism or anti-Asian racism or racism of any kind" – a statement playing on associations with the Great Replacement Theory, according to which Jewish elites seek to destroy the white, Christian society through mass immigration.[56]

Our dataset sampling these online debates contains **relatively high levels of antisemitism (33–38%)**, aside from two threads with 9–10% antisemitic comments.[57] The most frequent concept articulated is the **affirmation of Musk's comment, tantamount to the reproduction of an ANTISEMITIC CONSPIRACY THEORY (39% of all antisemitic comments)**: "Elon for President" (TODAY-YT[20231116]); "Elon Musk stood between us and the deep state control freaks" (TODAY-YT[20231116]). In an equally telling interaction, the comment "If you wish to know who wants to rule over you, find out who you're not allowed to criticise" triggered "✡️✡️✡️" (GMA-YT[20231116]) as a reply. Commenters also tended to **AFFIRM ANTISEMITISM in general (11%)** and elevate it as fact-based criticism: "He's facing backlash for spreading antisemitic truths" (ABC-YT[20231117]), or make approving references to Kanye West: "Ye was right!" (ABC-YT[20231117]). Formulations such as "They call you antisemitic but they never call you liar" (GMA-YT[20231116]) appeared consistently in West-related discourse, justifying the latter's antisemitic conspiracy theories.[58]

The concept of AFFIRMATION can also refer to the ANTISEMTISM OF NAZI GERMANY:

> "Hitler's overreaction doesn't mean there wasn't an issue to begin with, just like Netanyahu's overreaction doesn't mean there wasn't an issue" (MSNBC-YT[20231117])

– not only equates Nazi Germany and Israel via an **indirect NAZI ANALOGY** but also suggests that Hitler was not entirely wrong.

55 – Hull, Dana, 2023. Elon Musk Calls Antisemitic Post on *X* the "Actual Truth," Bloomberg, 16 November 2023, https://www.bloomberg.com/news/articles/2023-11-16/elon-musk-calls-anti-semitic-post-on-x-the-actual-truth (last accessed on 11 January 2024).

56 – See Chapelan, Alexis, 2024. 'Conspiracy' in: Becker, Matthias J. et al., 2024.

57 – Threads were taken from the *YouTube* channels of *MSNBC, ABC News, The Majority Report, Good Morning America* and *Today.*

58 – Chapelan, Alexis et al., 2023.

3.4 Media Matters for America, IBM, The White House

In November 2023, IBM announced that its advertisements had been found alongside Nazi posts and antisemitic comments on *X*.[59] The organisation Media Matters for America highlighted that ads from software giants such as IBM, Apple, and Oracle appeared on *X* next to posts with positive statements about Adolf Hitler and Nazi ideology. IBM decided to pause advertising on *X* – a decision that was followed by Apple and the media companies Disney, Paramount, Warner Bros and Lionsgate. The EU Commission also announced that it would temporarily suspend advertising on *X*, with sharp criticism coming also from the White House. Musk referred to Media Matters as "pure evil" and threatened a "thermonuclear lawsuit."[60]

This corpus consisted of 500 comments taken from the YouTube channels of both mainstream and alternative media.[61] Of these, **16.8% were classed as antisemitic**. The two most prominent concepts were the DENIAL OF THE ANTISEMITIC CHARACTER of **Musk's comments (35.7%) and their AFFIRMATION (35.7%)**. At times, these two reactions occur in combination: "Well, he does not care, and he is NOT anti-semitic just because he wants justice and not singing the Song of the Zionists" (SKY-YT[20231118]). Elsewhere, a comment combines a DENIAL with a blunt allegation that ANTISEMITISM IS BEING INSTRUMENTALISED **(7%)**: "An onslaught of exaggerated, sometimes wholly baseless accusations of antisemitism has proven to be brutally effective" (MSNBC-YT[20231118]). However, AFFIRMATION was often expressed indirectly, through wider approval of Musk or his companies: "I never thought of buying Tesla but I think I want one now' (CBSE-YT[20231118]).

Conclusion

These exploratory studies once again underscore the pivotal role of opinion leaders like Elon Musk in the resurgence of commonly expressed antisemitism. Most of the comments exhibit an affirmation of the antisemitic worldview expressed in the initial post, manifested through stereotypes or conspiracy theories. Alternatively, commenters DENY THE ANTISEMITIC CHARACTER of the worldview, thereby seeking to legitimate and affirm it in an inverted sense. Both can be seen as a bridge to prejudices against Jews, independently expressed by the commenters, as observed in our analyses in the form of stereotypes, such as EVIL, DISINTEGRATION, POWER, or the INSTRUMENTALISATION OF ANTISEMITISM. In some of the examined YouTube comments sections (notably channels from mainstream news outlets), the language used by commenters could hardly be distinguished from rhetoric within the White Supremacy spectrum, studied by other research projects, illustrating the conditions for 'mainstreaming' openly hateful and exclusionary ideas.

59 – Wendling, Mike, 2023. IBM suspends ads on *X* after they appeared next to Nazi posts, *BBC*, 17 November 2023, https://www.bbc.co.uk/news/world-us-canada-67446797 (last accessed on 21 December 2023).

60 – Kolodny, Lora, 2023. Musk threatens "thermo-nuclear lawsuit" against media watchdog, calls advertisers "oppressors." *CNBC*, 18 November 2023, https://www.cnbc.com/2023/11/18/musk-threatens-thermonuclear-lawsuit-against-media-watchdog-calls-advertisers-oppressors.html (last accessed on 21 December 2023).

61 – Threads were taken from the *YouTube* channels of *BBC, MSNBC, Sky News, Fox News, CSN Evening News*.

4. Our workshop series on web-related hate studies: conclusions and takeaways

Karolina Placzynta and Marcus Scheiber

In the autumn of 2023, the Decoding Antisemitism project team ran seven online workshops on contemporary online hate speech. The series was conceived as a means to reflect the methods and aims of our interdisciplinary project, establish a dialogue between our research experiences and other expert approaches, and open the debate on the challenges of hate speech in various areas of academic and non-academic professional practice.

Workshop 1:
Qualitative analyses of online hate speech and the challenge of grey areas of meaning

In the opening event of the series, moderated by Alexis Chapelan (Decoding Antisemitism), we focused on the question of distinguishing between hate speech in its heterogenous forms on the one hand, and legitimate critique of social structure, inequality, or human rights violations on the other. In his talk, Professor **Joachim Scharloth** (Waseda University) presented the key concepts of denigration and invectivity, discussing the fear of social devaluation, exclusion and losing face which drive reproduction of the social order, but also result in disparaging speech towards the other. Professor **Victoria Guillén Nieto** (University of Alicante) examined legally actionable intentionality through Speech Act Theory and Relevance Theory, using the case of United States v. Wilcox to illustrate the responsibility of both the speaker and the hearer in (re)constructing the intended meaning. Dr **Matthias J. Becker** (project lead of Decoding Antisemitism) presented the project's taxonomies, which help distinguish between legitimate critique and essentialising or generalising hate speech, using as an example the relationship between antisemitism and anticapitalism in the discourse

around George Soros, as well as the distinction between criticism of Israel and Israel-related antisemitism. In deciphering implicitness, pragmatic linguistic approaches involve language, context, and world knowledge. Agreement among the experts extended to observations that hate speech conveyed through indirectness and implicitness poses a significantly greater threat than its direct and explicit counterpart, as it can penetrate different milieus.

Workshop 2:
Discourse analytical approaches to online hate speech

The second workshop, moderated by Marcus Scheiber (Decoding Antisemitism) addressed online hate speech from the post-disciplinary perspective of discourse analytical approaches, considering its forms as well as the social changes impacted by and impacting on its spread in the fields of art, culture or politics. Professor **Benno Herzog** (University of Valencia) used the Spanish context to discuss the continuity and change in antisemitic metaphors and semantics in the information architecture of online spaces: potentially more democratic but with lower thresholds for truth. Discourse studies examine online antisemitism on symbolical and material levels and offer strategies to counter it. Professor **Johannes Angermuller** (The Open University) used the UK case of antisemitism accusations against Jeremy Corbyn to demonstrate antisemitism as a product of linguistic and non-linguistic (social) practices and systems. He highlighted how a prestigious and highly visible position not only enables and amplifies the impact of a political intervention but is also crucial in constituting an ideological effect such as antisemitism. In his view, antisemitism is not inherent in texts, but results from texts used in a system of social practices and relationships; therefore, antisemitism cannot be understood

independently from the social dynamics in which the ideological meanings, contents and intentions of the actors are constructed. Dr **Monika Kopytowska** (University of Łódź) explored the effects of technology and distance dynamics on digital meaning-making process, and the implications for the production, distribution and consumption of hate speech. She concluded that the trans-spatial, trans-temporal and techno-discursive character of online discourse shapes the social impact of hate speech, giving it a larger pool of arguments and a strong emotional potential. All speakers emphasised that discourse analysis as a meta-reflexive approach lends a nuanced quality to conventional qualitative analyses, as the researcher's own standpoint is also included in the analysis.

Workshop 3:
Quantitative and statistical analysis of online hate speech

In the third workshop, moderated by Dr Ildikó Barna (Eötvös Loránd University), we explored how quantitative and statistical analysis can shed light on, and expand the focus of, qualitative analyses based on annotated corpora of online data. **Daniel Miehling** (Technical University Berlin) opened the discussion by presenting quantitative and statistical analyses of conversations on the mainstream platform *Twitter* (now *X*) as well as in fringe communities on Telegram. Using a range of tools, including regression models, he showed that antisemitism has increased over time on both platforms, although they differ significantly in terms of the number of users, size and volume of content generated. Antisemitic conversations about Jews and Israel on *X* show a wider range of different users disseminating such content, each producing only a few messages, while

Telegram users belonging to fringe communities are not only more likely to produce multiple messages, but also disseminate these messages more frequently. Dr **Julian Hargreaves** (Woolf Institute) presented measurements of antisemitism across three studies into *Google*, *Instagram* and *Twitter* in 2020–2021, commissioned by the Community Security Trust and Antisemitism Policy Trust. His team found high presence of visual and textual antisemitism on all three platforms, including a proliferation of Israel-related antisemitism (particularly conspiracy theories) on *Instagram*. They concluded that services such as *Google's SafeSearch* were not sufficiently effective, producing both false negatives and false positives. **Stefan Munnes** (WZB Berlin Social Science Center) explored antisemitic dynamics in online discussions using Decoding Antisemitism's annotated datasets, comparing and contrasting data from different platforms, periods and countries. His analyses revealed a pattern in which initial comments tend to be more implicit, increasingly leading to explicit antisemitism and triggering further and more threatening replies. The complimentary research approaches of the three speakers showed the promise of quantitative and statistical analysis when combined with a qualitative element in the process of human annotation before the analyses, both expert and non-expert, in order to capture the implicit and contextual expressions of antisemitism.

Workshop 4:
Normalisation of hate speech in online spaces

The fourth meeting, moderated by Dr Matthew Bolton (Decoding Antisemitism), considered the acceptability and mainstreaming of hate speech, its evolution in the recent decades, the increasing polarisation, and the role of interactive media in these processes.[62] **Susan Benesch** (Harvard University & Dangerous Speech Project) showcased the phenomenon of virtue talk: language that convinces the in-group that they are crucial, honourable members of society duty-bound to protect women, children, cultural identity etc. (using as an example the social media discourse of participants in the attack on the US Capitol on 6th January 2021). As a positive, inwardly directed discourse, it is not immediately identified as harmful, and therefore is permissible in public discourse. **Murilo Cambruzzi** (Observatory on Antisemitism at Fondazione CDEC) focused on humour as a powerful vehicle for not just in-group solidarity but also for ridiculing and belittling the out-group. Jokes, irony or schadenfreude discourse shift the line of what is acceptable, obscure the speaker's intention, and offer a way to backtrack if confronted. Both speakers emphasised that it is misleading to equate hate speech with angry, emotional hateful language, as it aids gradual normalisation of the former, and the attitudes that underlie it, echoing the threat of implicitness debated in the first workshop, and calling attention to the urgent need for improved tools measuring of hate ideologies in their whole diversity in order to answer larger questions of normalisation.

Workshop 5:
Multimodal analysis of online hate speech

At the centre of the fifth event, moderated by Dr Laura Ascone (Decoding Antisemitism), were talks focusing on the growing instrumentalisation of memes, GIFs or videos in enabling hate ideologies to infiltrate digital culture. Dr **Janina Wildfeuer** and Dr **Dimitris Serafis** (University of Groningen) discussed both 'soft' and 'hard' examples of hate speech in online environments, explored with two-step multimodal analysis and argumentation analysis to reveal the interplay of several expressive forms in the meaning making, and the challenge of delimiting contextuality. **Marcus Scheiber** (Decoding Antisemitism) introduced the communication format meme and showed how it is utilised for antisemitic concepts. In particular, he highlighted the template character and demonstrated how memes can be analysed from a discourse-semiotic perspective. Dr **Inari Sakki** (University of Helsinki) explored multimodal artefacts in the form of a *TikTok* video created in the context of political hate communication. She placed emphasis on the complex interplay of modalities (visual, verbal, aural) and the need for methodological work in the field of music, as well as the contribution that music can make to the construction of hate. All speakers referenced the importance of including all modalities in qualitative analyses as the only way to understand the construction of hate ideologies online, insofar as online communication is always multimodal. They stressed the need for further systematic and fine-grained analyses, large corpora approaches and computational elements in the field.

62 – One of the scheduled speakers was unable to join, hence only two presentations are summarised here.

Workshop 6:
Automated detection of online hate speech

In the last two events of the series, we wanted to address the application of research findings to professional practice outside of academia. In the sixth workshop of the series, moderated by Jan Fillies (Institute for Applied Informatics, University of Leipzig) experts considered **means of automatically detecting hate speech** and the current possibilities and limitations of the iterative exchange between human coders and AI models, with the view to significantly advance social media studies, but also to improve content moderation. Professor **Helena Mihaljević** (HTW Berlin & Decoding Antisemitism) talked about her team's work on detection of online conspiracy theories and antisemitism, based partly on data shared by Decoding Antisemitism, using a variety of models, including BERT, Llama and GPT-4. She hypothesised that challenges of detecting online antisemitism may be due to the more elusive nature of antisemitic expressions and practical aspects such as the length of online comments, world knowledge gap, or contextuality. Dr **Savvas Zannettou** (University of Delft & Max Planck Institute for Informatics) described his work on detecting online antisemitism, in text and images, on platforms such as 4Chan. He concluded that his model can play a role not just in identifying antisemitic content but also understanding its evolution over time, but that human involvement should still be part of the moderation process. The work **Julia Mendelsohn** (University of Michigan) presented in her talk focuses on coded language often present in toxic speech within the public sphere, and on introducing this problem to the field of hate speech detection and computational linguistics. Together with her team, she developed a typology of dog whistles, prepared a living glossary with rich contextual information, analysed political speeches, and evaluated recommendations by large language models, working with GPT3 prompts and Google Perspective API. The speakers pointed out the relevance of the shift away from traditional machine learning models towards new technologies, importance of detecting existing toxic language as well as predicting subtle and innovative expressions of hate ideologies, and the urgent need to extend the work from English to other languages.

Workshop 7:
Educational strategies for online hate speech awareness, prevention and reaction

In the seventh workshop, moderated by Karolina Placzynta (Decoding Antisemitism), we examined the lessons from **educational interventions and practices**. Professor **Özen Odağ** (Touro University Berlin) talked about empowerment of emerging adults and creation of social change through training them to recognise and react to online antisemitism within a participatory framework, informed by young people's own perspectives. Drawing on a series of studies as part of her RESPOND! project, she identified the challenge of implicit, low-intensity antisemitism that is difficult for young people to recognize and that triggers disengagement. She emphasized young people's identities and loyalties as crucial elements of media competence trainings to combat antisemitism. **Miško Stanišić** (Terraforming) discussed the educational resources his organisation provides to counter state-sponsored Holocaust distortion in Serbia (i. e. the hijacking of this issue for propaganda purposes). He also elaborated on the project mapping antisemitism identification capacities in Serbia, Croatia and Romania within professional and social communities that shape the public discourse. **Malte Holler** (Bildung in Widerspruch) gave a preview of his organisation's new website "An Allem Schuld" (launched soon after the workshop), a virtual place of information and learning where teenagers and young adults can independently deal with current forms of antisemitism. The resource provides basic knowledge about antisemitism and its manifestations, insights into Jewish experiences and Jewish diversity, but also ways to manage emotions and psychological functions of antisemitism, promoting tolerance for ambiguity and, again, stimulating reflection. All three speakers touched on the topics of young people grappling with the difficulty and responsibility of confronting antisemitism, stressing the need for research that tracks the types of antisemitic speech they are exposed to online.

5. Automated detection of antisemitic texts: is context all we need?

Milena Pustet and Helena Mihaljević

In our previous report, we detailed our experiments involving fine-tuning a BERT-based large language model (LLM) to classify texts as either antisemitic or non-antisemitic. Our focus was on the English-language texts and the utilisation of a pre-trained English-language model. However, the **imbalanced nature of the data**, with nearly 10 non-antisemitic statements for every antisemitic one, presented a challenge common when attempting to identify offensive language or hate speech in mainstream online media or social media platforms.

Our initial model achieved an F1 score of 0.69 for antisemitic texts (positive class/class 1) and 0.96 for non-antisemitic texts (negative class/class 0).[63] **Our objective was to employ various strategies to further enhance the performance of the existing model.** Moreover, we utilised the data that had been annotated in the meantime. **Additionally, we aimed to assess the robustness of the model by training and evaluating it on different splits of the available data.** Alongside this, **we explored the capabilities of conversational models such as ChatGPT** that have shown solid performance across various text classification tasks.

5.1 Attempts to optimise existing BERT-based model

Working with only approximately 2,400 English-language texts labelled as antisemitic posed a challenge due to the limited data available for training. This is especially the case given the brevity, frequent usage of implicit, coded speech, and fragmentation of the texts in our dataset. At the same time, antisemitism manifests in various facets within these texts, as reflected by the extensive annotation scheme of the project. Thus, for effective training, a substantially larger volume of data would be beneficial.

One strategy to enhance the model involved **augmenting the data in the positive class** as this is significantly underrepresented. The first approach involved the replacement of some words by others with a supposedly similar meaning in the sense of a pre-trained language model, and the insertion of words that supposedly do not significantly alter the overall meaning of the sentence. The second approach involved translation as a strategy. Texts labelled as antisemitic were translated from English into German and then back to English. Additionally, positive data from other languages were translated into English. A random sample of translations was manually checked.

However, none of these strategies resulted in a significant improvement in the F1 score for class 1. A notable challenge stems from the fact that standard pre-trained models may not effectively capture the sensitive context in corresponding messages. For instance, the words 'Israel' and 'Palestine' might be

considered close from the perspective of a general language model, but they are not suitable replacements for each other in the context of the Middle East conflict. Models that have been further trained or fine-tuned on a corpus reflecting such nuances are likely more suitable, as well as other more sophisticated text augmentation strategies that we attempt to explore in the future.

Numerous experiments were conducted to **reduce the so-called 'easy negatives,'** referring to examples from the negative (i. e. non-antisemitic) class that are easily correctly classified. While a model is being trained, all data points contribute in the same way to the computation of the error which guides the training process. Thus, reducing such examples has the potential to enhance the detection of the positive class. While this strategy led to a higher recall for the positive class, and thus an increased identification of antisemitic texts, it came at the cost of precision, so no substantial overall improvement was achieved.

Furthermore, **language-specific models were trained for German and French, along with a multilingual model**. However, the results fell short of those obtained by the English-language monolingual model. The worse performance of the German and French models can be mainly attributed to the substantially smaller number of positive examples in these languages compared to the English corpus. A qualitative analysis of the model's output with regard to discourse countries would be necessary to make more precise statements in this regard.

In the past six months, we have **incorporated new English-language data** from both existing and emerging discourses into our training of new models. These new discourses included, for example, Elon Musk's tweet about George Soros and his negative remarks about the ADL, as well as the reactions to the 7 October massacre in Israel and the subsequent escalation in the Arab-Israeli conflict. In total, 676 additional positive data points were included.

With the infusion of these new data points, we proceeded to train various models. As outlined earlier, our objective was not only to enhance model performance but also to **ensure stable performance** when implementing diverse data splits for training, validation, and testing purposes. More precisely, we created five splits of the entire dataset, always keeping 10% of the data for testing and assuring that the percentages of both classes are comparable across the different splits.

As shown in Table 1, the overall results are less accurate than in the previous model. **The recall of the positive class, which measures the fraction of antisemitic instances that were effectively retrieved by the model, was especially low.** Moreover, the results for the positive class vary to a relatively high degree between the different splits, **indicating a high dependence of the model on data used for training and evaluation**. There are various plausible explanations for this. The new data contains additional discourse events in which antisemitism might have been expressed differently (see datasets in the report's chapters 2 and 3). Moreover, a qualitative inspection of records in the positive class suggests that annotators tend to struggle with recognising the level of outside context they draw upon to annotate an antisemitic text as such. The particular nature of the annotation process – where messages are considered within the specific discourse event and surrounding messages – makes it very difficult to simulate the situation where this type of contextual knowledge – knowledge that is not provided to the trained model – is unavailable.

	Precision	Recall	F1-score	Number of records	Accuracy
class 0 (non-AS)	0.96 ± 0.00 (0.97 ± 0.00)	0.98 ± 0.01	0.97 ± 0.00 (0.97 ± 0.01)	3.533	0.94 ± 0.01 (0.95 ± 0.01)
class 1 (AS)	0.71 ± 0.05 (0.73 ± 0.04)	0.52 ± 0.04 (0.63 ± 0.03)	0.6 ± 0.02 (0.68 ± 0.02)	309	

Table 1: Mean ± standard deviation of the best-performing models per split on test data (and validation data in brackets if different).

5.2 Experiments with prompt-based models

OpenAI's recent strides in the field of generative pre-trained transformers (GPT), particularly with the introduction of GPT-3 and the multi-modal GPT-4, have attracted considerable attention owing to their noteworthy proficiency in generating responses that closely resemble human-like expressions. These models, accessible through platforms such as ChatGPT, also present novel opportunities for text and image classification. Meanwhile, openly accessible models such as Llama 2, developed by *Meta*, have been provided, achieving similar performance as GPT-3 models in various tasks.

These models can be applied in both **zero-shot** and **few-shot** settings. In the former, the model is presented solely with a task description, whereas the latter permits the model to glean knowledge from a limited set of so-called in-context examples, whose role is thus comparable to that of training data for supervised models. The formulation of textual instructions, referred to as prompting, significantly influences the model's behaviour. It serves as a critical task, as it shapes the model's focus and determines its responses, underscoring the importance of precision and nuance in the design of prompts.

Initial investigations highlight the potential of these models in text classification applications, e.g. in detecting hateful or insulting language.[64] However, recent research also sheds light on the varied effectiveness of GPT-3 in identifying "coded expressions that simultaneously convey one meaning to a broad audience and a second one, often hateful or provocative, to a narrow in-group"[65] – so-called dog whistles – including antisemitic ones. Further investigation into the capabilities and limitations of generative language models in tackling intricate classification tasks are needed, particularly in the context of antisemitic texts. The experiments we present in the following section represent an initial stride in this direction.

64 – Chiu, Ke-Li/Collins, Annie/Alexander, Rohan, 2021. Detecting Hate Speech with GPT-3. Preprint. arXiv:2103.12407; Wang, Yau-Shian/Chang, Yingshan, 2022. Toxicity Detection with Generative Prompt-based Inference. Preprint. arXiv:2205.12390; Huang, Fan/Kwak, Haewoon/An, Jisun, 2023. Is ChatGPT better than Human Annotators? Potential and Limitations of ChatGPT in Explaining Implicit Hate Speech. In Companion Proceedings of the ACM Web Conference 2023. ACM, Austin TX USA, 294–297, doi:10.1145/3543873.3587368.

65 – Mendelsohn, Julia/Le Bras, Ronan/Choi, Yejin/Sap, Maarten, 2023. From Dogwhistles to Bullhorns: Unveiling Coded Rhetoric with Language Models. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). doi:10.18653/v1/2023.acl-long.845.

5.3 Experimental results

In our experiments, we **kept to the zero-shot setting**, primarily for two reasons. Prior research indicates that the performance of models in few-shot scenarios may exhibit a decline compared to zero-shot settings.[66] The incurred cost is contingent on the prompt length, which tends to be significantly longer when incorporating in-context examples, particularly when coupled with a custom definition. This is especially pronounced for GPT-4, which is already twice as expensive as GPT-3.5.

We conducted zero-shot experiments employing GPT-3.5, GPT-4 and Llama 2,[67] wherein we compared performance under two conditions: one where the prompt solely outlines the task, and another where a **custom definition** of antisemitic speech is provided. Specifically for GPT-3.5, we further investigated the **impact of constraints on model output**. In the 'Yes/No' configuration, the model was tasked with providing a binary decision about the presence of antisemitic content in the message. In the 'Yes/No with Reason' setting, an additional short reasoning was requested. The third configuration, 'Probability Scoring,' mandated the model to provide a score between 0 and 1 that indicates the likelihood of a message being antisemitic.

The following shows what the prompt looked like for the 'Yes/No' experiment with the provision of an expert definition:

```
Consider the following statement:
'{STATEMENT}'

You have to decide whether the
statement communicates antisemitism
considering the following definition:
'{DEFINITION}'

Give your answer using one of the
two options:
    A) Yes
    B) No

Do not provide any other output or
explanation in your output.
```

Evaluations were run on validation data of the model presented in the previous report, comprising 249 positive and 2,061 negative samples. As shown in Table 2, the conversational models perform even worse than our best BERT-based approach (see bottom row in Table 2). With an F1 score on class 1 of 0.55, **GPT-4 is clearly the best-performing model, albeit achieving a significantly lower score than previous attempts using BERT-like models**. The F1 score for class 0 also remains below the previous one. Both GPT-4 and Llama 2 perform better when equipped with an expert definition of antisemitic speech. However, GPT-3.5 shows different behaviour depending on the output constraint, with performance depending on the provided definition only in the binary 'Yes/No' setting. This indicates that GPT-3.5 highly depends on the chosen configuration.

66 – Chae, Youngjin (YJ)/Davidson, Thomas, 2023. Large Language Models for Text Classification: From Zero-Shot Learning to Fine-Tuning. Preprint, https://osf.io/preprints/socarxiv/sthwk; Pustet, Milena/ Steffen, Elisabeth/Mihaljević, Helena, 2023. Conspiracy Narrative Detection Beyond Keyword Bias in German-language *Telegram*. Preprint.

67 – The exact model versions are gpt-3.5-turbo-0613, gpt-4-061 and llama 2 70b.

5. Automated detection of antisemitic texts: is context all we need?

Model	Output constraint	Definition	F1-class 0	F1-class 1	Accuracy
GPT-3.5	Yes/No	-	0.93	0.40	0.88
		X	0.95	0.19	0.90
	Yes/No + Reasoning	-	0.85	0.42	0.76
		X	0.86	0.43	0.77
	Probability Scoring (threshold = 0.5)	-	0.94	0.42	0.89
		X	0.94	0.42	0.88
GPT-4	Yes/No	-	0.95	0.46	0.91
		X	0.95	0.55	0.92
Llama 2	Yes/No	-	0.78	0.23	0.65
		X	0.77	0.30	0.65
fine-tuned bert-base-uncased			0.96	0.69	0.94

Table 2: Performance of the generative models on validation data from UK discourses, comprising 249 positively and 2,061 negatively labelled samples.

5.4 Conclusion and future work

As demonstrated, both BERT-like models and conversational models like GPT-4 exhibit limited proficiency in accurately identifying antisemitic texts. Our comprehensive experiments lead us to various possible explanations and possible paths forward: first, the data basis we utilised might not be conducive to effectively classifying individual, typically concise messages for the presence of antisemitism. The intricate relationships between the comments within a thread, the corresponding article or post, and the prevailing (political) context seem to pose significant challenges. Human experts excel in considering these nuanced implications and dependencies, enabling them to provide well-founded assessments of the respective texts. Conversely, models, especially those lacking a comprehensive understanding of current political knowledge and the context of messages from other users and commenters, struggle with this task. Hence, it seems advisable to reassess the meaningfulness of distinguishing annotations between contextualised and non-contextualised antisemitism. **For the training of prediction models, it might make more sense to assume the inherent relevance of context, especially in terms of direct references to preceding messages or the original text. Consequently, rather than classifying individual messages in isolation, a more fruitful approach appears to involve the examination of entire discussion threads.** Such a framework could leverage the capabilities of LLMs more effectively than previous attempts. We intend to delve into this challenge in future research, with a particular focus on harnessing the potential of prompt-based models for improved results.

Second, antisemitism is a particularly challenging field compared to other hate ideologies. It is often expressed using coded language with a specific meaning to one audience while appearing innocuous to others. Antisemitic statements can be conveyed through historical references, making them challenging to identify without contextual understanding. Moreover, antisemitic statements can be embedded in various political spaces or online subcultures,

with different rhetorical nuances and argumentation strands. Also, antisemitism is less addressed in datasets and modelling endeavours related to hate speech detection, which contributes to an explanation why models such as GPT-4 struggle substantially more to recognise antisemitic content.

The substantial challenges posed by antisemitism to prompt-based models such as GPT-4 mean further research is imperative in this domain. For instance, one avenue of investigation could involve assessing the impact of different definitions, the provision of examples, or information regarding the discourse event on model performance. A qualitative analysis of the reasoning generated by the model could prove insightful in identifying concepts that the model struggles to recognise adequately. This information could, in turn, guide modifications to the definitions, enriching them with relevant in-context examples if deemed necessary.

Ensuring the practical utility and ongoing evolution of models such as the ones we have developed is paramount. To facilitate real-world application, **we have established an inference service featuring the best BERT-based model within a web app**. This service enables users to input text, receive predictions, and view corresponding scores. A crucial feedback loop has been implemented, allowing users to express agreement or disagreement, thereby enhancing our understanding of the model's performance and aiding in the collection of additional training data.[68] The trained models can be provided upon request.

---

68 – All data is securely stored in a MariaDB database, with HTTPS encryption ensuring robust security measures. Our technical stack includes AWS, Angular, Node.js, and MariaDB, and the system is hosted on a virtual Linux server. Currently operational within our team, the code is available for sharing, facilitating the implementation of similar setups in other projects.

6. Quantifying antisemitism: approaches, challenges and promises

Stefan Munnes

Because of the dramatic increase in the volume of discourse on social media, hate speech research is undergoing a process of "datafication"[69] which relies on transparent, reproducible and statistically verified methods of quantitative text analysis. However, modern antisemitism, characterised by a strong diversity of forms and mutability in the context of public condemnation, is a challenging object for quantitative investigation.[70] The existing literature focuses mainly on frequency analyses based on keywords and slurs, but also include image and vector analysis.[71]

Grasping the importance of combining granular qualitative analysis and quantitative approaches, the Decoding Antisemitism project has developed an integrated three-step framework to address this challenge. Analysis begins with detailed qualitative examination, focusing on the content-conceptual and linguistic-structural composition of antisemitic communication online. This labelled data is then used as a basis for the subsequent development of large language models (LLMs) and then broad-reaching statistical procedures.[72]

This final step of the project uses quantitative methods to quantify and visualise the annotated data: the codes, textual features and relationships between them. In Discourse Report 2, Daniel Allington presented a lexical network analysis of three discourse events, showing connections between main keywords used in the corpus.[73] To draw out the specific keywords for the discourses under study, an additional reference corpus had to be created for comparative purposes. As unlabelled text was used, these keywords could not be distinguished according to antisemitic speech.

To utilise the full potential of the annotated dataset, two further methodological approaches were applied to analyse the specifics of the antisemitic discourses. The results of these two approaches can be practically experienced by the reader through various interactive visualisations that are accessible on the project website.[74]

First, the **frequencies of the different labels** (such as stereotypes and analogies) used to annotate antisemitic comments are visualised for the different discourse events and therefore made comparable. This gives a first impression of the importance of certain antisemitic concepts and their linguistic form.

69 – Laaksonen, Salla-Mariaa et al., 2020. The Datafication of Hate: Expectations and Challenges in Automated Hate Speech Monitoring. In: *Frontiers in Big Data*, 3 (1), 10.3389/fdata.2020.00003.

70 – The problem of communication latency has already been discussed in the classic survey research and requires adapted survey instruments (cf. Beyer, Heiko/Krumpal, Ivar, 2013. The Communication Latency of Antisemitic Attitudes: An Experimental Study. In: Small, Charles A. (ed.). Global Antisemitism: A Crisis of Modernity. Brill/Nijhoff, 83–96, https://doi.org/10.1163/9789004265561_010; Beyer, Heiko/Liebe, Ulf, 2013. Anti-Semitism Today Measuring Current Manifestations of Hostility towards Jews Using Factorial Surveys. In: *Zeitschrift für Soziologie*, 42 (3), 186–200.

71 – See Anti-Defamation League, 2018. Quantifying Hate: A Year of Anti-Semitism on Twitter, https://www.adl.org/resources/report/quantifying-hate-year-anti-semitism-twitter; Anti-Defamation League, 2019. Gab and 8chan: Home to Terrorist Plots Hiding in Plain Sight, https://www.adl.org/resources/report/gab-and-8chan-home-terrorist-plots-hiding-plain-sight#_ftn1; Finkelstein, Joel et al., 2020. Antisemitic Disinformation: A Study of the Online Dissemination of Anti-Jewish Conspiracy Theories, https://networkcontagion.us/reports/antisemitic-disinformation-a-study-of-the-online-dissemination-of-anti-jewish-conspiracy-theories (last accessed on 17 January 2024).

72 – See the research design of the Decoding Antisemitism project: https://decoding-antisemitism.eu/about.

73 – See Becker, Matthias J. et al., 2022. Decoding Antisemitism: An AI-driven Study on Hate Speech and Imagery Online. Discourse Report 2. Berlin: Technical University Berlin. Centre for Research on Antisemitism.

74 – The interactive visualisations are located in the "Publications" section, via the description of Discourse Report 6.

These distributions can also be compared between the three studied countries. In addition, **co-occurrence** diagrams show the connections between the different labels representing antisemitic concepts and the linguistic features with which those concepts are expressed by web users (e. g. the concept of EVIL conveyed through a pun, or the NAZI ANALOGY depicted through an allusion). The focus on the occurrence of the labels in the same comment allows for both the interaction of the different concepts and their specific linguistic expression to be captured. Here, too, the discourse-specific antisemitic framings can be shown both across countries and in their development over time.

Second, further visualisations of **frequencies and relationships between antisemitic keywords** were created. While a keyword approach may risk missing some elements of implicit antisemitism which does not use such terms, the large amount of annotated text data means that it is possible to extract keywords that are generally characteristic of texts classified as antisemitic, compared to neutral or non-antisemitic texts. Using a keyness measure, words were extracted that were more likely to occur in comments labelled as antisemitic. This allows for the large-scale examination of the linguistic characteristics of antisemitic utterances in different discourse events and means results can be presented in an interactive manner so that users can compare the use and frequency of specific terms across discourses and time. However, inter-country comparison is limited due to the different languages.

Both approaches demonstrate the possibilities for analysing large amounts of well-annotated data and help to gain a better understanding of the data and quantify patterns that have already been described in detail in the qualitative reports.[75] Nevertheless, despite the large size of the dataset, the various forms of antisemitism and the relatively low number of occurrences of most antisemitic concept constitute a challenge to the quantitative analysis.

To build on these results, a future project will seek to further integrate quantitative hypotheses and research questions into the initial data collection process. For regression-based approaches that aim to model relationships between an outcome variable and different features – i. e. the prediction of antisemitic comments based on text features, given keywords, comment metadata or the news outlet – an even broader sample would be necessary. Oversampling of discourse events triggering antisemitism makes sense from a labelling perspective but makes it difficult to estimate valid effect sizes due to the overrepresentation of antisemitic comments. The need for larger datasets again underlines the importance of the training of large language models able to accurately classify comments automatically.

75 – For another approach that abstracts from the content level and focuses on the metadata of comments to analyse the dynamics of (antisemitic) discussions and threats, see Vincent, Chloé, forthcoming. Discussion trees on social media: a new approach to detecting antisemitism online. In: Becker, Matthias J./ Ascone, Laura/Placzynta, Karolina/ Vincent, Chloé (eds.). Antisemitism in Online Communication, Transdisciplinary Answers from Research, https://decoding-antisemitism.eu/publications/announcement-anthology; see also Stefan Munnes in our Workshop Series presented in Chapter 4 of this report.