حملة - المركز العربي
لتطوير الإعلام الاجتماعي
7amleh - The Arab Center for
the Advancement of Social Media

Racism and Incitement Index 2023

Racism & Incitement Index 2023
Author: Rama Salahat

Revised and edited by Jalal Abu Khater & Ahmad Qadi

Translated into English Language by: Ritaj Managerial Solutions

Designed by: khalil kahwaji

This version is licensed under the following International License: Attribution-NonCommercial-NoDerivs 4.0 International
To view a copy of the license, please visit the following link:
https://creativecommons.org/licenses/by-nc-nd/4.0/

Contact Us:
Email: info@7amleh.org
Website: www.7amleh.org
Telephone: +972 (0)774020670

Follow us on social media: **7amleh**

Content Table:

Introduction/background	4
Contextualizing Violent Content	6
Methodology	11
Classifier	11
Data	11
Findings	12
Conclusion	16
Recommendations	17

Introduction/background

During the past year, the proliferation of hate speech and incitement to violence towards the Palestinian people on online platforms reached alarming levels. During the first half of 2023, 7amleh reported on extensive Israeli campaigns on social media platforms, prompted and supported by Israeli ministers and officials, where inciting violence against Palestinians was widespread. In one case, the patterns of incitement online translated into real-world harm with fatal implications as they contributed to the organisation of an attack by hundreds of settlers on the village of Huwara and its residents, on the evening of February 26. During the latter part of 2023, and particularly after October 7th, 7amleh reported an unprecedented increase in hate speech and incitement in the Hebrew language, directed towards Palestinians, perpetuating violence, justifying collective punishment, and exacerbating the dehumanisation of the Palestinian people.

On 27 October 2023, the United Nations Committee on the Elimination of Racial Discrimination expressed serious concern about "the sharp increase in racist hate speech and dehumanisation directed at Palestinians since 7 October, particularly on the Internet and in social media." It is evident from the mounting evidence and reports through the Palestinian Observatory for Digital Rights Violations (7or) that hate speech, incitement to violence, and dehumanisation of Palestinians are already at an alarming level. For instance, a post from the Israeli Deputy Mayor of Jerusalem on December 8 described stripped and blindfolded Palestinian detainees in Gaza as "hundreds of ants" that he wanted to bury alive; "they are neither human beings nor human animals, they are sub-humans and that's how it should be." The post was removed after it was reported but many others remain. Such rhetoric, allowed to persist unchecked, not only perpetuates the dehumanisation of Palestinians but also fuels an environment where violence is normalised and celebrated.

Data provided by the Palestinian Observatory for Digital Rights Violations indicate a total of 2749 violations of harmful content were documented on platforms like Facebook, X (formerly Twitter), and Telegram between 1/1/2023 and 31/12/2023. The three aforementioned platforms are considered frequent sites of such violations.

7amleh's documentation efforts increased with the launch of the «Violence Indicator», an AI-based language model, a real-time indicator that monitors the spread of hate speech and violence in Hebrew against Palestinians and their advocates on social media platforms. Since it began recording on October 6, and by the end of the year, the Violence Indicator recorded nearly 3 million instances of hate speech and violent content against Palestinians on platforms like X (formerly Twitter) predominantly, in addition to Facebook and Telegram.

On 26 January 2024, the International Court of Justice (ICJ) ordered provisional measures in the case of South Africa v. Israel, determining the plausibility that Israel is carrying out genocide against the Palestinian people in Gaza. The Court adopted legally binding orders that include requiring Israel to prevent genocide against Palestinians in Gaza as well as to prevent and punish direct and public incitement to commit genocide, as foreseen in Article III(e) of the Convention on the Prevention and Punishment of the Crime of Genocide.

This ICJ order underscores the gravity of the situation, especially considering the documented use of online platforms to incite genocide against Palestinians in Gaza, including by the highest levels of the Israeli leadership. The ICJ took note of statements made by Isaac Herzog, President of Israel, and Israeli Minister Yoav Gallant, and specifically referred to a post shared by the Israeli Minister Israel Katz on X, which read: "The line has been crossed. We will fight the terrorist organization Hamas and destroy it. All the civilian population in Gaza is ordered to leave immediately. We will win. They will not receive a drop of water or a single battery until they leave the world."

We note that all of these Israeli officials, as well as many others who have made similar genocidal statements, continue to disseminate harmful content across online platforms. The urgency of the matter cannot be understated. The widespread dissemination of hate speech, incitement and dehumanization of Palestinians on online platforms contribute to a narrative that condones and supports actions amounting to war crimes.

Online platforms must fulfill their legal and moral responsibility to uphold human rights, promote accountability, and prevent the dissemination of harmful content on their platforms including incitement to genocide.

Contextualizing Violent Content

As 7amleh endeavors to create a free, fair, and safe digital environment for all users, it is notable that social media platforms haven't invested in establishing Hebrew classifiers specifically designed to identify and address violent and hateful content. It also seems to be a lack of commitment on their part to initiate the development of such mechanisms. Given the platforms' complicity, Israeli authorities and organizations have failed to actively pursue legal action against individuals disseminating violent content, contributing to a sense of impunity that led to exacerbating the issue.

7amleh employs a comprehensive approach to document and combat hateful and violent content proliferating across social media platforms. This involves a dual-level strategy, utilizing both manual and AI model interventions.

At the manual level, the Palestinian Observatory of Digital Rights Violations (7or) plays a pivotal role. This observatory meticulously collects content and account information, captures screenshots, and subjects the data to rigorous review and verification processes. Subsequently, validated cases are escalated to social media companies, urging them to promptly remove the offensive content.

The significance of this manual documentation extends beyond mere content removal. 7or's efforts result in the creation of a robust dataset, meticulously classified and categorized. This dataset becomes a valuable resource for in-depth analysis and advocacy initiatives. In the year 2023 alone, 7or successfully documented a staggering 4400 digital violations. Within this grim total, 2749 instances of violent and hateful content in Hebrew were identified.

A discernible pattern emerges from the analysis of this data, revealing a surge in incitement against Palestinians by general Israeli users, particularly in the aftermath of the October war eruption. This disturbing trend not only perpetuates the ongoing conflict but also intensifies the pressure for its continuation. Moreover, 7amleh's documentation sheds light on a concerning aspect: public figures engaging in incitement and hate speech. These influential figures wield considerable impact, further compromising the lives and rights of Palestinians.

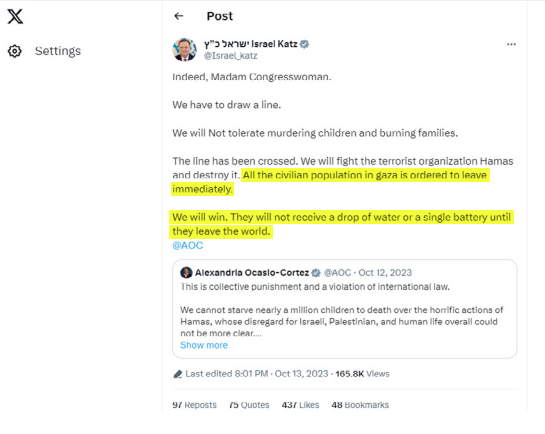
Importantly, 7amleh recognizes the interconnection between digital incitement, hate speech, and on-the-ground attacks. The documentation efforts not only serve as a means to remove offensive content but also as a crucial tool to understand and address the broader socio-political implications of these online activities. By revealing the nexus between online rhetoric and real-world consequences, 7amleh strives to create awareness and advocate for a safer and more inclusive digital space for all.

The following are a few examples of violent content posted by Israeli officials and public figures targeting Palestinian civilians:

X (formerly Twitter)

"Erase Gaza. Nothing else will satisfy us. It is not acceptable that we have a terrorist authority next to Israel. Not to leave a child there, expel all the remaining ones till they are out, so that they will not have a resurrection."

"Invest this energy in one thing;
Erasing all of Gaza from the face of the earth.
That the Gazan monsters will fly to the southern fence and try to enter Egyptian territory.
or they will die.
And they will go to hell after death.
Gaza should be erased.
And we'll see fire and smoke of the heads of the Nazis in Judea and Samaria.
The Jewish rage will shake the whole world.
A vengeful and cruel IDF is needed here.
Anything less is immoral.
Just immoral."



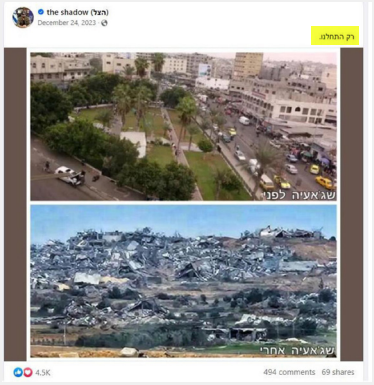
. . .“All the civilian population in Gaza is ordered to leave immediately.

We will win. They will not receive a drop of water or a single battery until they leave the world.” @AOC

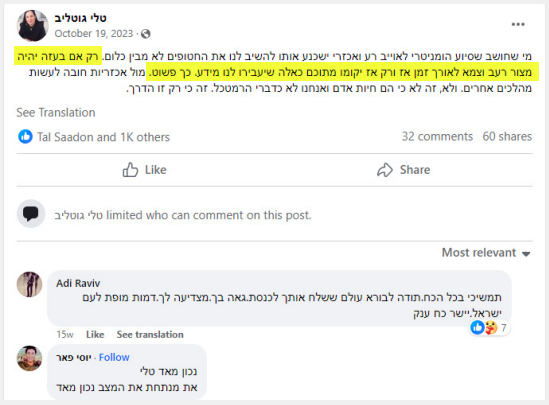
. . . “The Hague judges who care about the situation of the Gazans can call on the countries of the world to open their doors and assist in the reception and rehabilitation of the residents of Gaza.”

. . . “Right now, we have one goal: Nakba!
A Nakba that will overshadow the Nakba of 48.
Nakba in Gaza and Nakba to anyone who dares to join Gazans!”

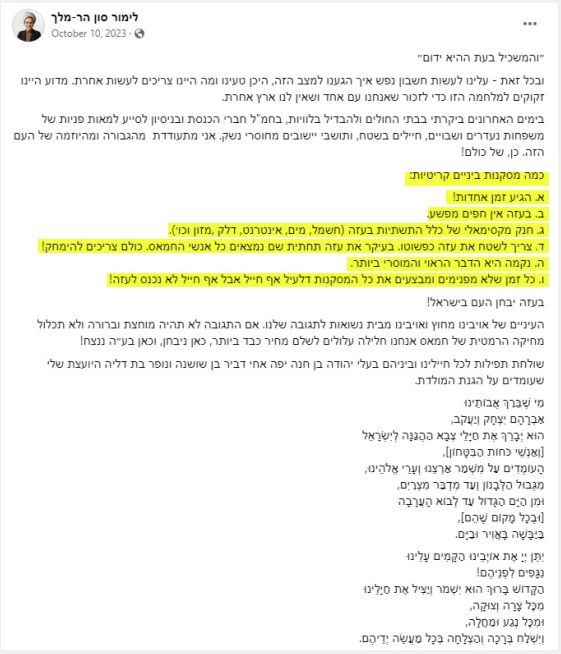
Meta



We just began.



. . . “Only if there is a hunger and thirst siege in Gaza for a long time, then and only then will some of them arise who will pass us information. That simple.”



. . . “Some critical interim conclusions:

A. It’s time for unity!

B. There are no innocents in Gaza.

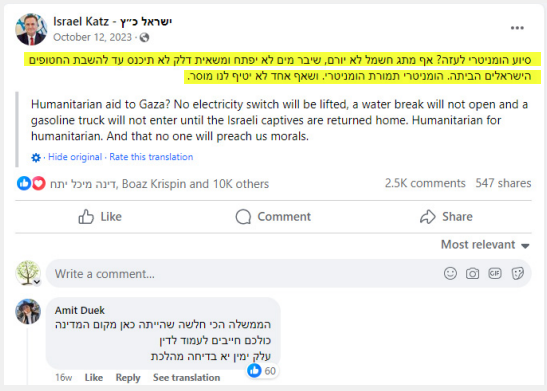
C. Maximum strangulation of all infrastructure in Gaza (electricity, water, internet, fuel, food, etc.).

D. Gaza needs to be flattened. Mainly lower Gaza where all the Hamas people are. They all need to be wiped from the earth!

E. Revenge is the most appropriate and moral thing.

F. As long as all the above conclusions are not internalized and carried out, not a single soldier enters Gaza!

The people of Israel will be tested in Gaza!”

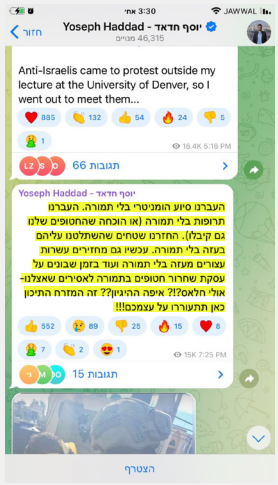


“Humanitarian aid to Gaza? No electrical switch will be turned on, no water hydrant will be opened and no fuel truck will enter until the Israeli abductees are returned home. Humanitarian for humanitarian. And no one will preach us morals.”

Telegram

"The Gazan journalist Muthana Al-Najjar, who broadcast from Gaza envelope since 7.10 laments that 20 members of his family have been buried under the ruins of a building for over two months.

How is he not buried yet???"



"We transferred humanitarian aid for nothing. We transferred medicine for nothing (not even a proof that our abductees also received it). We returned territories we took over in Gaza for nothing. Now they are also getting dozens of detainees from Gaza for nothing and more while they are building a deal to release hostages in exchange for our prisoners – That's enough! Where is the logic?? This is the Middle East here, stand up for yourselves!!!"

"There is a solution, migration now!"

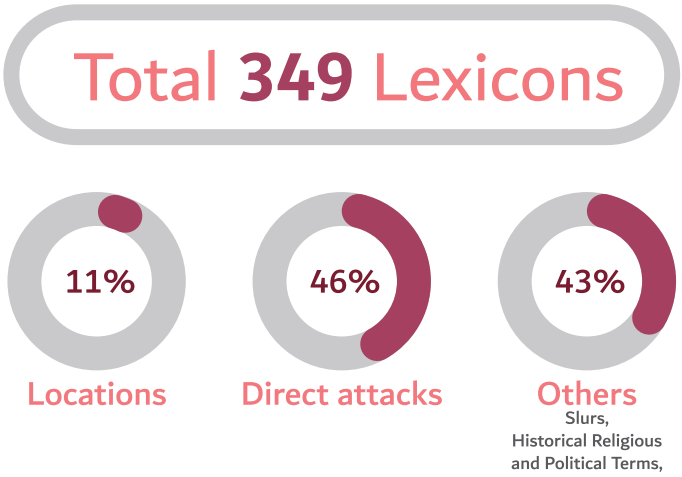
Methodology

Classifier

This report is powered by a series of advanced Large Language Model (LLM) classifiers, that were internally developed in 2023 and trained with human-labeled data by 7amleh's team. The data was annotated by two individuals separately, with a third person consulted to resolve any discrepancies between labels.
The violence binary classifier achieved an F1 score of 92% and a Recall of 87%, indicating that it accurately detects violent text in 92% of cases when taking both relevant and irrelevant texts into account, while also correctly identifying truly violent texts with an accuracy rate of 87%.
This model was also compared against Azure's Content Moderation API to ensure consistent and accurate results, and the results showed an 83% agreement rate between our LLM classifier and the API's classification.

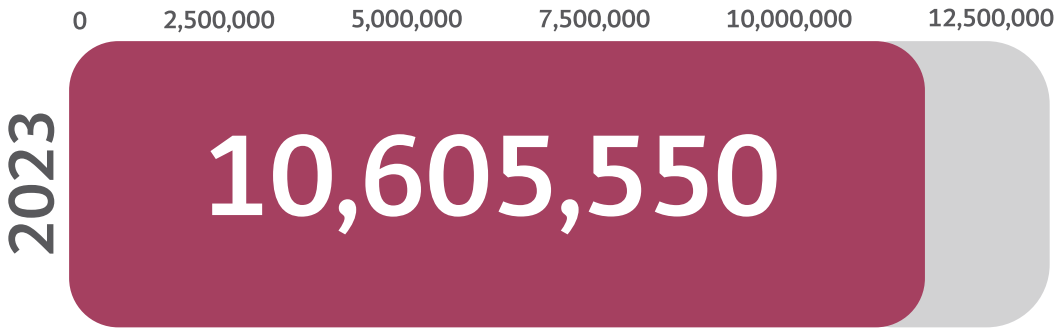
Data

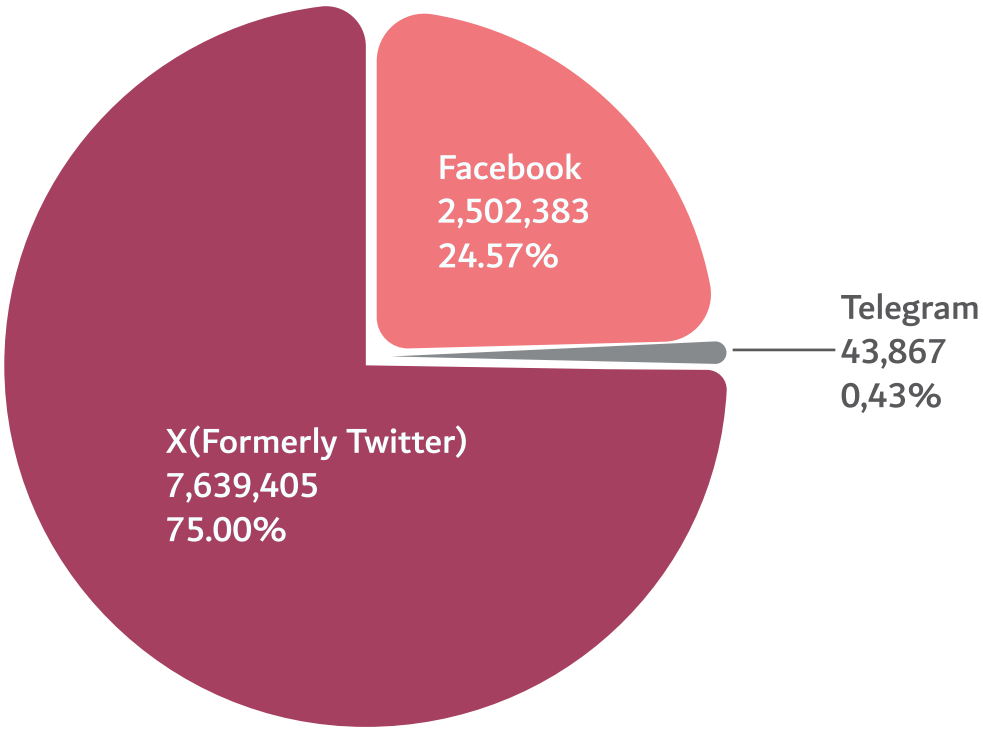
Data was gathered from three different platforms: X (previously known as Twitter), Facebook, and Telegram. For X, 7amleh conducted searches using a curated list of 675 words associated with hate speech and violence. This list was manually compiled. Facebook and Telegram data were collected using specific Pages/Channels identified by us.
Subsequently, the collected data underwent processing before being input into 7amleh's internal classifiers for in-depth analysis. 7amleh created a set of lexicons to extract data based on previously reported incidents through the 7or platform. These lexicons included names of Palestinian areas in Hebrew and settlements where events occurred. Additionally, specific lexicons focused on direct attacks, using words or sub-sentences such as "kill Palestinians," "second Nakbah," and "destroy." The category labeled as "Others" encompassed terms like "Hamas," "water," and slurs.



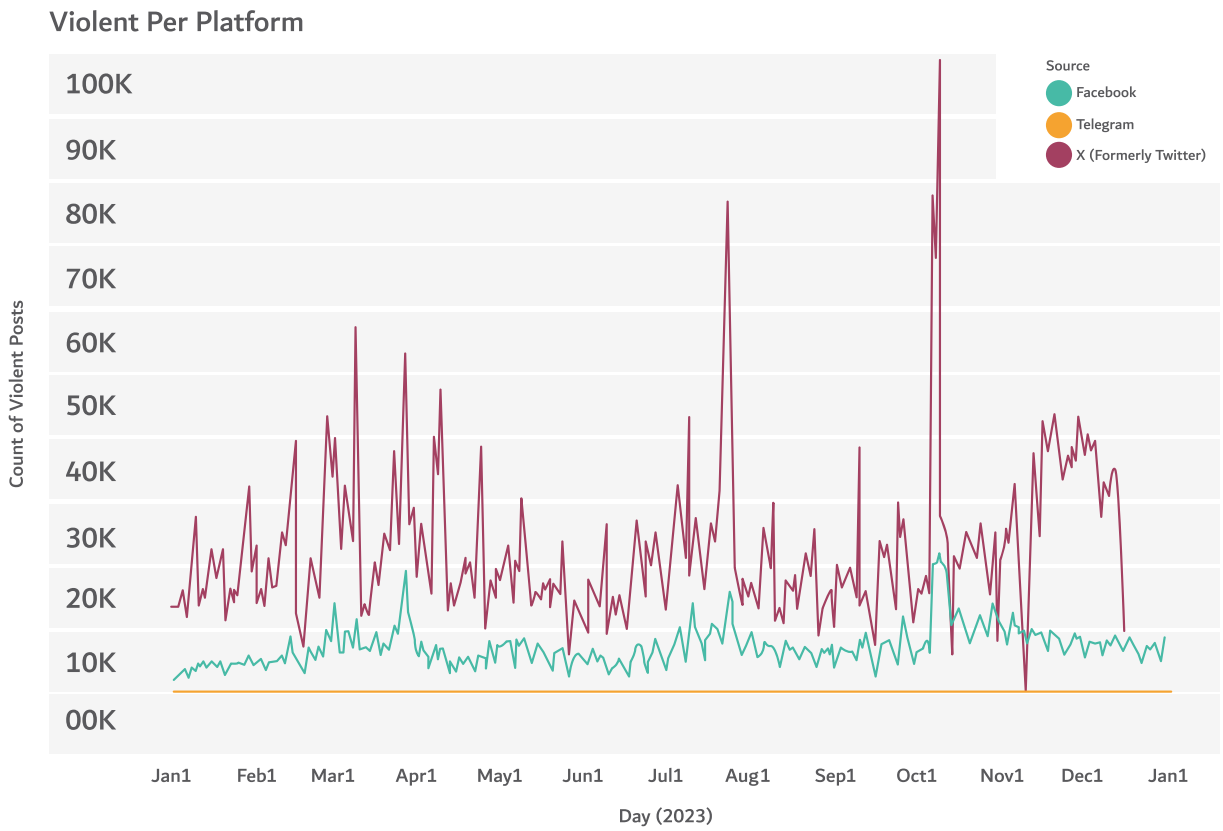
Findings

The AI language model has detected **10,605,550** violent pieces of content in Hebrew across three main social media platforms.

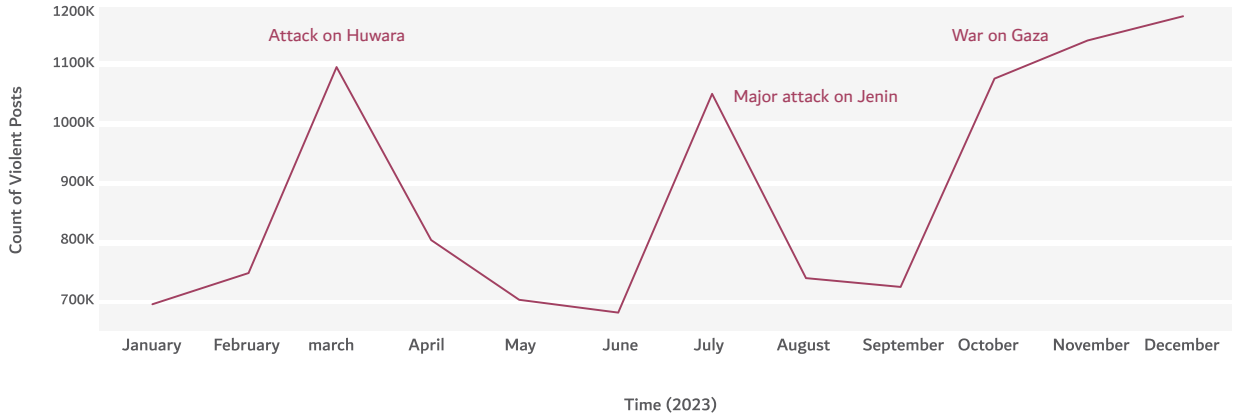




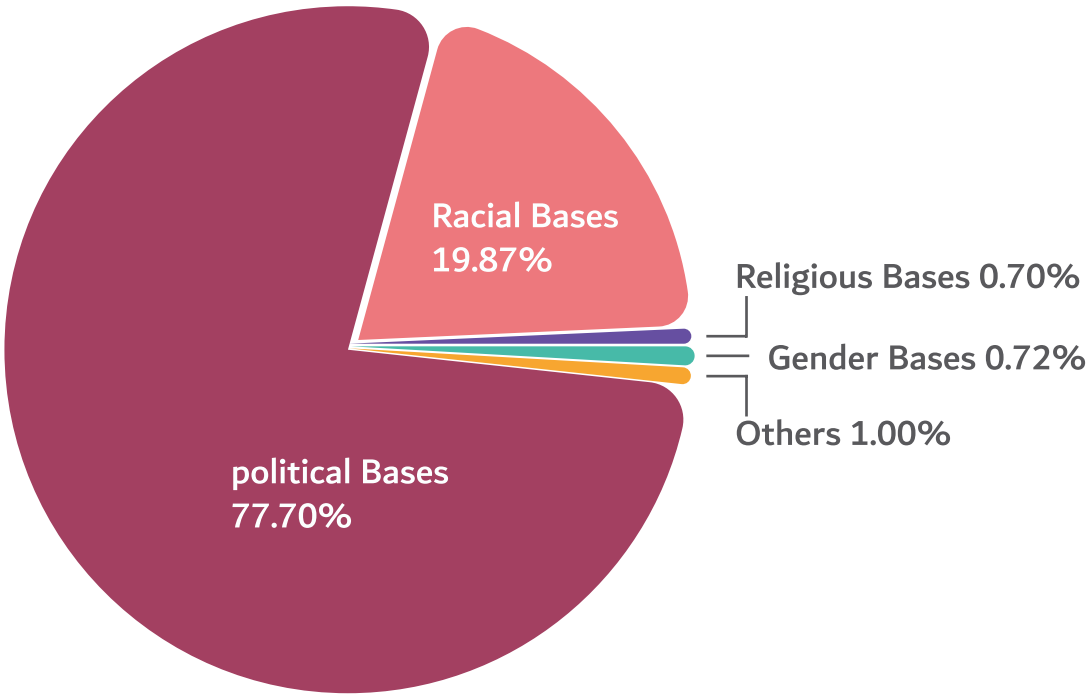
Three-quarters **(76%)** of the collected violent and hateful pieces of content originate from X, highlighting the significant prevalence of such content on this platform. Facebook accounts for **23.6%** of these records, emphasizing a substantial presence as well. This distribution underscores the rampant nature of violent content in Hebrew, particularly on X, while Meta's platforms exhibit comparatively lower prevalence. It's worth noting that Telegram also serves as a platform for the substantial proliferation of hateful content, though a more comprehensive dataset could not be obtained due to technical challenges.



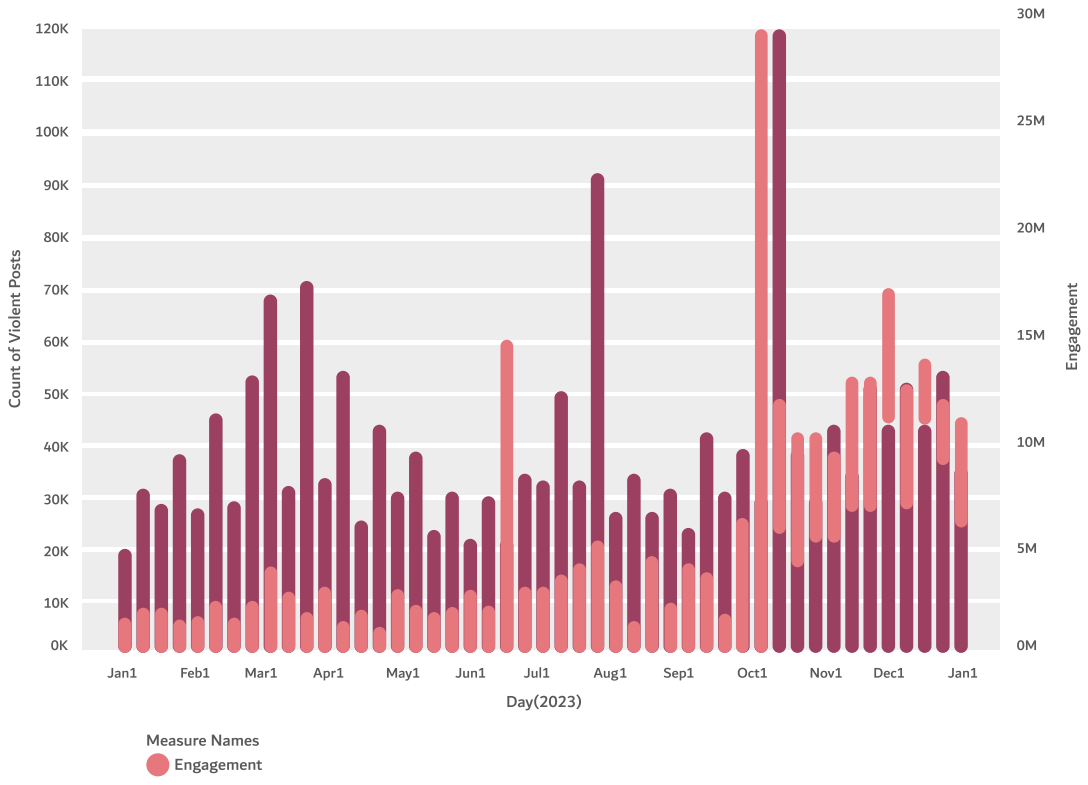
It appears that the proliferation of violent content on platform X is particularly high, with Meta's platforms, specifically Facebook and Instagram, following closely behind in terms of prevalence.



Based on the data collected, there was an average of 19 hateful/violent pieces of content posted per minute. From Oct 7th on, the average increased to around 23 posts/comments per minute.

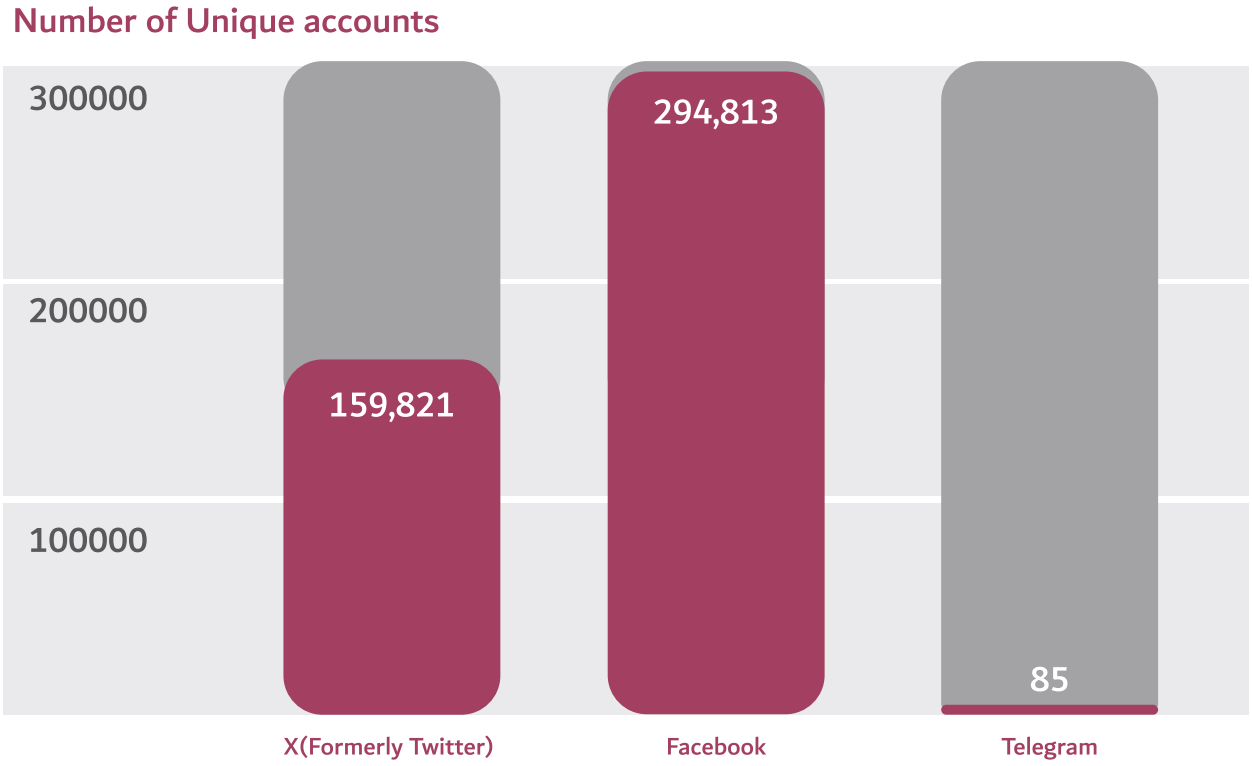


77.7% of these posts were hateful/violent on political bases, while **19.8%** of the content were on racial bases, it's also clear that racial basis violent content has tripled after October 7 2023.



After Oct 7th, there was an average of **284** "engaging actions" **(like, comment, share...)** per post compared to only **70.7** average engaging actions before Oct 7th. It is almost 4 times the amount of engagement.

יהודים
עזה
חמאס
מחבלים
מלחמה
טיפש
ערבים
אידיוט
אנשים
הטרור
טרור
שנאה
החמאס
החטופים
בעזה
צהל
חיילים
להרוג
המלחמה
להעיף
לעזה
מחבל
טמבל
לעזאזל
רוצחים
מוות
מסריח
המחבלים
ישראלים
נשים
הערבים
להוציא
כיבוש
עונש
בוגדים
במלחמה
טילים
מכות
גרוע
חיסול
למות
בדם
באוקטובר
שחרור
מפגר
בוגד
לרצוח
לכבוש
להשמיד
מטומטמת
לשרוף
תחזור
מכל
אבן
רצח
דבר
בטרור
לחמאס
מטומטמים
אשפה
טיפשה
בגץ
נאצי
אינם
פושעים
לעומת
הרג
הנאצים
מדהים
האחרונה
להביא
נורא
אויב
מטורף
נרצח
לבטל
מטונף
להחליף
פייק
ערבית
רקטות
ברחובות
רעל
אבנים
שחמאס
מוחמד
עבריין
לזרוק
טרוריסט
פגיעה
שונאים
הזבל
למחבלים
אנטישמיות
הרצועה
לירות
המחבל
בכלא
מטומטם
לפגוע
אסור
ערבים
חטופים
אידיוט
אנרכיסטים
קרעי
ליהודים
גירוש
זוועה
לערבים
הדם
יורים
חופש
להילחם
לגרש
להורג
קיצוניים
בחוארהדאעש
מוסלמי
רשע
המחנה
הורגים
לטרור
טיפשים
חלאה
קיצוני
ההיסטוריה
מעבר
מאופס
עלובים
מורשע
והסתה
שונא
מתים
להקיא
חמור
לפטר
רוע
אונס
איכס
האוייב
אישום
נשכח
רקובה
גוש
האש
הנפש
בכוח
לחטוף
מגעיל
שורפים
הרעל
מכוער
לשחוט
אוכל
אמסלם
התחת
שרת
יימח
ישלם
מאיים
קרב
השטן
זין
אשפה
השמדה
שרמוטה
לעזה
פיגוע
המלחמה
חרא
הפלסטינים
רצח
גבר
חיל
מכה
לזיין
לפח
האויב
העזתים
לטבח
טרוריסטים
לרצוח
הערבים
להוציא
כיבוש
עונש
מגעיל
שורפים
הרעל
גוש
האש
הנפש
בכוח
לחטוף
מפגר
בוגדים
שחרור
באוקטובר
מכות
גרוע
חיסול
למות
בדם
מיני
טילים
במלחמה
מין
הורג
מטומטמת
לשרוף
תחזור
מכל
אבן
רצח
דבר
בטרור
לחמאס
מטומטמים
אשפה
טיפשה
בגץ
נאצי
אינם
פושעים
לעומת
הרג
הנאצים
מדהים
האחרונה
להביא
נורא
אויב
מטורף
נרצח
לבטל
מטונף
להחליף
פייק
ערבית
רקטות
ברחובות
רעל
אבנים
שחמאס
מוחמד
עבריין
לזרוק
טרוריסט
פגיעה
שונאים
הזבל
למחבלים
אנטישמיות
הרצועה
לירות
המחבל
בכלא
מטומטם
לפגוע
אסור
ערבים
חטופים
אידיוט
אנרכיסטים
קרעי
ליהודים
גירוש
זוועה
לערבים
הדם
יורים
חופש
להילחם
לגרש
להורג
קיצוניים
בחוארהדאעש
מוסלמי
רשע
המחנה
הורגים
לטרור
טיפשים
חלאה
קיצוני
ההיסטוריה
מעבר
מאופס
עלובים
מורשע
והסתה
שונא
מתים
להקיא
חמור
לפטר
רוע
אונס
איכס
האוייב
אישום
נשכח
רקובה
גוש
האש
הנפש
בכוח
לחטוף
מגעיל
שורפים
הרעל
מכוער
לשחוט
אוכל
אמסלם
התחת
שרת
יימח
ישלם
מאיים
קרב
השטן
זין



The data collected from a multitude of unique accounts suggests the active participation of hundreds of thousands of users in creating and spreading profoundly violent content in Hebrew. This indicates that these violent narratives are not confined to marginalized discourses; instead, they reflect the sentiments of a substantial sector within society.

Conclusion

The surge in the dissemination of violent content in Hebrew over the past year is noticeable when compared to previous years, particularly gaining momentum after the eruption of the Israeli war on Gaza on October 7. A materialized correlation emerges between digital violent content and real-world attacks by settlers and the Israeli Military, exemplified in both the recent conflict in Gaza and the preceding assault on Huwara. This underscores a significant connection between online narratives and real-world events, shaping a complex dynamic worth further examination.

Therefore, 7amleh underscores the need for social media companies to take decisive actions to curb the dissemination of violent and hateful content targeting Palestinians across various platforms. Such content contributes to a surge in real-world attacks against Palestinians.

Recommendations

Companies need to:

- Develop effective Hebrew language classifiers to combat the surge in hate speech and incitement against Palestinians online, alongside a commitment to fighting hate speech and incitement in all languages to safeguard users globally.
- Strengthen the platform's trust and safety, and human rights and policy teams, to be better equipped to tackle hate speech and incitement arising during crisis situations.
- Increase native-speaking Hebrew and Arabic human content reviewers on their platforms.
- Commit to ongoing co-design with civil society to enhance policies and processes related to negative and illegal content online.

حملة - المركز العربي
لتطوير الإعلام الاجتماعي
7amleh - The Arab Center for
the Advancement of Social Media

2023