What Distinguishes Disseminators of Antisemitic Tweets and What Themes Do They Use?

Gunther Jikeli[1][0000-0002-6897-2565] and Rhonda Fischer[1][0000-0003-1163-8563]

[1]Indiana University, Bloomington IN 47405, USA
{gjikeli, rkfische}@iu.edu

Abstract. Social media is the largest single disseminator of antisemitism. It provides a breeding ground for radical terrorist whose ideological glue are antisemitic conspiracy theories. While the moist radical groups are on the fringes, many fragments of antisemitic ideologies and myths are disseminated on mainstream social media, such as Twitter. Tracking and monitoring antisemitic messages on large platforms require computer assisted methods. We built a labeled dataset of 4,137 live tweets from representative samples of tweets of four separate keywords, "Jews, Israel, kikes, and ZioNazi*." Our data shows an increase in the percentage of antisemitic messages in conversations about Jews from 2019 to summer 2020, despite some efforts by Twitter to delete or suspend hateful tweets and accounts. Our statistical analysis of word frequencies, hashtags, average number of friends and followers, and the percentage of repeat users reveal interesting differences between antisemitic and non-antisemitic tweets. The paper provides initial insights into the question of who the average Twitter users are who are sending antisemitic tweets and their primary themes.

Keywords: labeled dataset, antisemitic users, Twitter.

1 Introduction

Conspiracy theories are a core element of modern antisemitism, that is the belief that Jews run the world and are responsible for disasters, wars, economic hardship, or the pandemic. Antisemitism can be understood as a particularly vicious form of disinformation that can quickly turn violent. The most lethal forms of terrorism of the last few years, white nationalism and radical Jihadism, are inherently antisemitic at its core. (Rickenbacher, 2021; Ward, 2017) Many terrorists have been radicalized online in environments where hatred against Jews and other minorities are the norm. (Kursuncu et al., 2020) Social media has become the largest single disseminator of antisemitic propaganda and social media provide breeding ground for antisemitic ideas that can go viral and/or push Jews out of these spaces. (Barlow, 2021) However, tracking antisemitic messages on social media is challenging. Marginal radical groups might be monitored manually. Monitoring the dissemination of antisemitic messages to larger audiences and on mainstream platforms, such as Twitter, with millions of daily messages require methods that are automated assisted. (Bruns, 2020; Davidson et al., 2017; Malmasi & Zampieri, 2017; Zannettou et al., 2020) We built a labeled dataset of tweets that can serve as an initial Gold Standard to train algorithms to identify antisemitic tweets in larger datasets. The dataset is based on representative samples of relevant keywords. This paper provides a basic statistical analysis of the metadata and the text of the antisemitic and non-antisemitic tweets. It provides first indications who average disseminators of antisemitic content on Twitter are and what distinguishes them and the themes they promote from non-antisemitic users.

2 Labeled Corpus

Our corpus draws on Indiana University's Observatory on Social Media (OSoMe) database that includes 10 percent of all live tweets on an ongoing basis, going back 36 months from the time of a query. This allows us to build statistically relevant ("representative") subsamples that can then be annotated manually. We used two keywords that are likely to result in a wide spectrum of conversations about Jews as a religious, ethnic, or political community: "Jews" and "Israel." We then added samples with more targeted keywords likely to generate a high percentage of antisemitic tweets, that is the insults: "kikes" and "ZioNazi*." We drew 11 representative samples of different timeframes within the period January 2019 and August 2020. We built an Annotation Portal (https://annotationportal.com) for easier and consistent labeling and had two out of a team of five expert annotators label every tweet and discuss their discrepancies, if any. Table 1 shows the annotation results. The classification as antisemitic means that at least one paragraph of the Working Definition of Antisemitism of the International Holocaust Remembrance Alliance (IHRA) applies. If annotators felt that the tweet was antisemitic but no paragraph of the IHRA definition applies, then they would classify the tweet as not antisemitic according to IHRA or vice versa.

Table 1. Labeled dataset comprised 11 representative samples for keyword and time period.

	Keyword	Timespan	Number of tweets in Gold Standard corpus	Percentage of antisemitic tweets
1	Jews	Jan.-Dec. 2019	439	6.2 %
2	Jews	Jan.-Dec. 2019	414	7.5 %
3	Jews	Jan.-Apr. 2020	469	11.9 %
4	Jews	Jan.-Apr. 2020	429	11.4 %
5	Jews	May-Aug. 2020	394	14.0 %
6	Jews	May-Aug. 2020	388	16.2 %
7	ZioNazi*	Jan.-Dec. 2019	374	88.8 %
8	ZioNazi*	Jan.-Apr. 2020	158	85.4 %
9	Israel	Jan.-Apr. 2020	344	10.2 %
10	Israel	May-Aug. 2020	431	13.0 %
11	kikes	Jan.-Dec. 2019	297	31.6 %
SUM		**Jan. 2019 to Aug. 2020**	**4,137**	**22.55** %

The representative samples of live tweets show an increase of the percentage of antisemitic tweets in conversations on Jews from 2019 to summer 2020. The percentage of antisemitic tweets rose only slightly from the first four month of 2020 to summer 2020. As expected, the percentage of antisemitic tweets was high for the insult "ZioNazi*," but relatively low for the insult "kikes," due to a high percentage of tweets calling out the use of the latter and due to references to two famous sportsmen with the nickname kiké.

3 Statistical Differences of Antisemitism and Non-Antisemitic Tweets

We tried to avoid internal bias as much as possible and ran the analysis separately for each of the four keywords. The keywords relate to different contexts and, perhaps more importantly, while the keywords "Jews" and "Israel" can be found in millions of tweets within our time frame, the insults "kikes" and "ZioNazi*" are used on the margins only.

3.1 Comparing Top Words

Notwithstanding the above, it is striking that the top 25 words in the overall dataset are different for the antisemitic and non-antisemitic tweets. Cleared of common English words and the query keywords, antisemitic tweets often include the words "apartheid" (0.41), "Palestine" (0.36), "state" (0.32), "Israeli" (0.3), "Jewish" (0.3), and "world" (0.29). Non-antisemitic tweets on the other hand most often include "Trump" (0.33), "hate" (0.32), "against" (0.31), "Nazi" (0.29), and "Muslims" (0.28). Antisemitic tweets often include words that relate to the Israeli-Palestinian conflict, indicating Israel-related forms of antisemitism. They also often include the words "world" and "Jewish," indicating references to one of the core myths of modern antisemitism, alleged Jewish world domination. The frequent appearance of the word "Trump" reflects the fact that Trump is the most mentioned user in our dataset. He is also frequently mentioned in antisemitic tweets, albeit a little further down the list (0.22). Other frequent words in non-antisemitic tweets indicate that many of such tweets are calling out hate and bias. Word clusters vary depending on the method used as the image below shows comparing two simple methods.

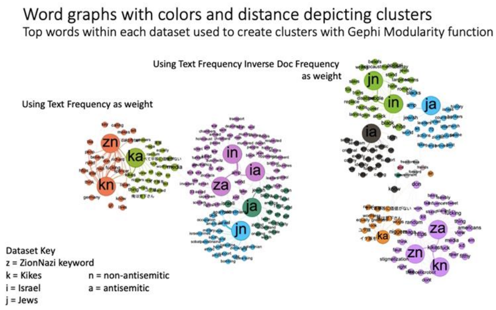

Fig. 1. Word clouds compare using either text frequency or tf-idf for the cluster grouping.

Conversations on Jews. Our dataset includes 2,533 tweets of conversations about Jews, using that keyword. 281 (11 percent) are antisemitic. They are representative for live tweets with that keyword for three consecutive periods from January 2019 to August 2020. Antisemitic and non-antisemitic tweets had some frequent words in common, including “people, Jewish, Israel, and Christians.” It shows that conversations about Jews are often about the Jewish people, Israel, and relations to Christians. However, the words “kill” (0.58), “Palestine” (0.44), and “world” (0.3) were found exclusively in the antisemitic tweets in the top 25 words, while “'hate” (0.38), (“Muslims (0.37), “Holocaust” (0.36), and “Trump” (0.33) were exclusively among the top 25 words of the non-antisemitic words. As in the overall sample, antisemitic tweets include words that indicate references to antisemitic tropes related to Israel and world domination. The frequent use of the word “kill” is mostly part of accusations against Jews being murderers. Non-antisemitic tweets on the other hand frequently include words that indicate that they are calling out antisemitism or that they are related to the Holocaust. References to Trump can be positive or negative reactions to activities and tweets by the president.

Conversations on Israel. Our dataset includes 775 tweets of conversations about Israel, using that keyword. 91 (12 percent) are antisemitic. This is a relatively small subset of tweets and some important themes might be missing, especially from antisemitic conversations about Israel. Unsurprisingly, common words in all tweets about Israel were “Palestine and Palestinian,” and also “against.” Common words in antisemitic tweets were "state" (1.27), "Apartheid" (0.87), "Zionist" (0.63), "world" (0.56), "people" (0.48), "swilkinsonbc" (0.48), "destroy" (0.4), “years" (0.4), "terrorist" (0.4), "terrorism" (0.4), and “American" (0.32), indicating that many of these tweets accuse the Jewish State of Apartheid, destruction, and terrorism. Others include phrase such as “the world is silent/sleeping” in face of Israel’s alleged crimes or references to Israel “influencing the world.” "Swilkinsonbc" is a very active user who often posts negative tweets about Israel. Non-antisemitic tweets often refer to news reports about Israel. Frequent words are “UAE” (0.49), "peace" (0.47), “deal” (0.42), “Iran” (0.4), “Trump” (0.39), “Gaza” (0.36), “Coronavirus” (0.32), “united” (0.29), “Arab” (0.29), and “breaking” (0.28).

Messages That Include the Insult “Kikes”. Our dataset includes 297 tweets with the insult “kikes,” of which 94 (32 percent) are antisemitic and use the term approvingly. Most of the non-antisemitic tweets are calling out those who use that term approvingly or refer to one of two famous sportsmen with the nickname “kiké.” Both categories often include some variation of the word f**k, reflecting the profanity of this keyword. Antisemitic tweets often include the words "medias" (1.2), "MAGA" (1.1), "randum" (1.0), "niggers" (1.0), "syria", (0.5), "influence" (0.5) and signs of approval or disapproval (rawr, lol, wops, fad, look, greaser). Popular words in non-antisemitic tweets are "Spencer" (2.23), "ruled" (2.23), “AynRandPaulRyan" (2.19), "believe" (2.19), "listen" (2.19), "Jews" (0.99), "antisemitism" (0.49), "hate" (0.41), "calling" (0.37), "used" (0.37), "mean" (0.37), and "think" (0.37). The user “AynRandPaulRyan” sent a popular tweet denouncing Richard Spencer for using this insult.

Messages That Include the Insult “ZioNazi*”. Our dataset includes 532 tweets with the word “ZioNazi” or “ZioNazis.” Almost all of them, 88 percent, are antisemitic. The percentage of users denouncing this insult which is related to Israel is lower than for those who denounce using the insult “kikes.” Unsurprisingly, antisemitic tweets with this keyword often include words that are used to denounce Israel or that are related to the Israeli-Palestinian conflict, such as “Israel" (1.21), "Apartheid" (0.64), “Israeli" (0.44), "land" (0.42), "Gaza" (0.35), "zionazism" (0.33), "Palestinian" (0.33), "BDS” (0.31), "Trump" (0.29), "zionazist" (0.29), and “Palestine" (0.27). The user "swilkinsonbc (0.44) was often mentioned.

3.2 Hashtags

Hashtags are another indication of frequent themes of tweets. The top 25 hashtags of antisemitic tweets of the overall dataset are ‘Israel, IsraeliCrimes, BDS, Putin, Syria, Palestine, FoxNews, Arab, randum, Russia, ZioNazi, Trump, FreePalestine, SkyNews, Assad, unpopularopinion, Group4Palestine, BlackLivesMatter, sellout, terrorUSt, USrael, NYTmes, China, Iran, and Thieves.” The top 25 hashtags of non-antisemitic tweets of the overall dataset are “Auschwitz, goodmorning, BuenosDias, bonjour, coronavirus, Drancy, antisemitic, UAE, Iran, Israel, WeRemember, HolocaustRemembranceDay, StopBiafraKillings, Bible, Democrats, Covid19, USA, Jesus, UAEIsrael, Passover, NeverAgain, 'NeverForget, JustJews, BoardofDeputies, and ManyJewishVoices.” Many of the hashtags of the antisemitic tweets are related to international politics or the Israeli-Palestinian conflict, whereas popular hashtags in the non-antisemitic tweets are about remembrance of the Holocaust, calling out antisemitism, Christian themes, Jewish organizations, or just everyday themes, such as #goodmorning. Interestingly, there were only two overlapping hashtags among the top 25, #Israel and #Iran. These different themes of hashtags in antisemitic and non-antisemitic tweets are reflected in tweets of all four keywords.

3.3 Followers and Friends

Users of antisemitic tweets have less followers and less friends on average than users who send non-antisemitic tweets, see tables below.

Table 2. Number of friends on average.

Number of Friends on Average				
Keyword	Jews	Israel	kikes	ZioNazi*
Antisemitic Tweets	2,036	1,755	174	1,681
Non-Antisemitic Tweets	2,579	2,789	1,965	1,814

Table 3. Number of followers on average.

Number of Followers on Average				
Keyword	Jews	Israel	kikes	ZioNazi*
Antisemitic Tweets	3,200	3,160	133	1,456
Non-Antisemitic Tweets	28,315	27,401	2,836	2,622

3.4 Repeat Users

Antisemitic tweets have a lower percentage of unique users; in other words, there are more repeat users in the antisemitic data set.

Table 4. Number of tweets vs. unique users.

Number of tweets vs unique users			
Keyword	Total tweets	Unique users	% tweets with unique users
Antisemitic Tweets	933	672	72%
Non-Antisemitic Tweets	3204	3094	97%

4 Conclusions

Comparing tweets with common key words where tweets have been classified as antisemitic or non-antisemitic, enables quantifying differences between tweet sets. Differences in the tweet sets include the frequency of specific words and hashtags due to differences in their themes. And users who tweet antisemitic messages average fewer followers and friends and have a higher tendency to be repeat offenders. These findings encourage expanded use of exploratory data analysis and machine learning as part of distinguishing antisemitic messaging themes within social media.

Acknowledgements

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. We are grateful that we were able to use Indiana University‘s Observatory on Social Media (OSoMe) tool and data. (Davis et al., 2016)

References

1. Barlow, E. (2021, May 25). *The Social Media Pogrom*. Tablet Magazine. /sections/news/articles/the-social-media-pogrom
2. Bruns, A. (2020). Big Social Data Approaches in Internet Studies: The Case of Twitter. In J. Hunsinger, M. M. Allen, & L. Klastrup (Eds.), *Second International Handbook of Internet Research* (pp. 65–81). Springer Netherlands. https://doi.org/10.1007/978-94-024-1555-1_3
3. Davidson, T., Warmskey, D., Macy, M., & Weber, I. (2017). Automated Hate Speech Detection and the Problem of Offensive Language. *Proceedings of the Eleventh International Conference on Web and Social Media 2017, Montreal, Canada*. https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15665
4. Davis, C. A., Ciampaglia, G. L., Aiello, L. M., Chung, K., Conover, M. D., Ferrara, E., Flammini, A., Fox, G. C., Gao, X., Gonçalves, B., Grabowicz, P. A., Hong, K., Hui, P.-M., McCaulay, S., McKelvey, K., Meiss, M. R., Patil, S., Peli Kankanamalage, C., Pentchev,

V., … Menczer, F. (2016). OSoMe: The IUNI observatory on social media. *PeerJ Computer Science*, *2*, e87. https://doi.org/10.7717/peerj-cs.87

5. Kursuncu, U., Gaur, M., Castillo, C., Alambo, A., Thirunarayan, K., Shalin, V., Achilov, D., Arpinar, I. B., & Sheth, A. (2020). Modeling Islamist Extremist Communications on Social Media using Contextual Dimensions: Religion, Ideology, and Hate. *ArXiv:1908.06520 [Cs]*. http://arxiv.org/abs/1908.06520
6. Malmasi, S., & Zampieri, M. (2017). Detecting Hate Speech in Social Media. *ArXiv:1712.06427 [Cs]*. http://arxiv.org/abs/1712.06427
7. Rickenbacher, D. (2021). The Centrality of Antisemitism in the Islamic State's Ideology and Its Connection to Anti-Shiism. In G. Jikeli (Ed.), *The Return of Religious Antisemitism?* (pp. 93–102). MDPI. https://doi.org/10.3390/books978-3-03943-498-5
8. Ward, E. (2017, June 29). *Skin in the Game. How Antisemitism Animates White Nationalism*. Political Research Associates. http://www.politicalresearch.org/2017/06/29/skin-in-the-game-how-antisemitism-animates-white-nationalism
9. Zannettou, S., Finkelstein, J., Bradlyn, B., & Blackburn, J. (2020). A Quantitative Approach to Understanding Online Antisemitism. *Proceedings of the International AAAI Conference on Web and Social Media*, *14*, 786–797.