More data, less hate.

Policy Advisory Opinion 2024-05-07
Regarding New Cases on Posts that Include "From the River to the Sea"
Comments Submitted by Tal-Or Cohen Montemayor on behalf of CyberWell Ltd. (CC)

Executive Summary

As the Oversight Board considers recommendations regarding Meta's policies and practices when addressing content citing the phrase, "From the River to the Sea, Palestine will be Free", foremost in the deliberations must be consideration for the real-world harm that has been caused and continues to be perpetuated by the use of this term as a call for violence and vandalism against Jewish communities, their institutions, and monuments dedicated to preserving the memory of the Holocaust around the world. CyberWell recommends that, similar to Meta's approach to the seemingly 'innocuous' term "Where we go one, we go all", the use of "From the River to the Sea" to target and attack Jewish individuals and communities has met a threshold of real-world harm and risk and, as such, should be flagged and actioned as a sensitive term that causes violence until the current wave of the antisemitism against Jews, offline and online, subsides.

Introduction

As a nonprofit organization committed to eradicating online Jew-hatred through driving the enforcement and improvement of community guidelines and safety policies of digital platforms, it is important for CyberWell to provide guidance on this subject, as the phrase "From the River to the Sea, Palestine will be free" is currently frequently being used as a mobilization cry in various flashpoints of antisemitism and violence against Jews online and in their communities, on college campuses, and against Jewish institutions worldwide.

We further seek to offer solutions rooted in content moderation best practices that balance freedom of expression with Meta's obligation to protect their users with adequate responses and protocol to prevent the spread of additional violence against Jews. As such, we refer to the actions taken by Meta against a similarly 'innocuous' term at scale on Facebook and Instagram, "Where we go one, we go all", due to the risk of real-world harm, as a framework for sourcing recommended solutions on how to treat "From the River to the Sea".

Antisemitism Expressed by "From the River to the Sea"

As Jews constitute a protected category on Meta platforms, and as CyberWell's ongoing methodology focuses on identifying antisemitic content and trends on digital platforms, we first conduct an analysis on "From the River to the Sea" and whether it can hold antisemitic

implications. The International Holocaust Remembrance Alliance's (IHRA)[1] working definition of antisemitism is a consensus definition of anti-Jewish bias that is adopted formally by over 40 countries and over 1,000 universities, municipalities, and sports leagues, and CyberWell uses this definition when analyzing our data.

The implementation of the call to dissolve the State of Israel "From the River to the Sea", effectively barring Jews from their right to self-determination in their ancestral homeland of modern-day Israel, would be consistent with the seventh example featured in the IHRA working definition.

However, since Meta has not adopted the IHRA working definition in full and the words of this phrase do not explicitly target a protected category of users on Meta's platforms (i.e. Jews and Israelis), we recognize that it would be difficult to identify this phrase on its own as violating existing digital policy without additional context.

Two examples from the IHRA definition are more appropriate in this case, providing the necessary additional context on when use of this phrase would be a violation of Meta's existing policies. These include the first example, "calling for, aiding, or justifying the killing or harming of Jews", and the eleventh example, "Holding Jews responsible for the actions of the State of Israel".

Weaponizing this phrase, or any other phrase, and using it to harass, target, or commit violence against Jews, would be against Meta's existing policies, and moreover, would be consistent with real world harms that we see occurring today such as violations of law, proliferation of hate crimes, and danger to public safety.

Real World Harm

During the current unprecedented surge in antisemitism worldwide following the violent October 7 attacks against Israeli civilians, this phrase is being used in just such a way. Synagogues from Philadelphia[2] to Barcelona[3] were vandalized with this phrase. "From the River to the Sea" was spray-painted on the site where the Palestinian terror group, Black September, massacred and took hostage 11 Israelis and a German police officer during the 1972 Munich Olympic games.[4] This phrase was used as a rallying cry in multiple illegal college solidarity encampments when physically restricting, harassing, threatening, spitting on, and physically harming Jewish students, press, and members of the public.[5] It was also chanted as

---

[1] https://holocaustremembrance.com/resources/working-definition-antisemitism
[2] https://www.jpost.com/diaspora/article-794864
[3] https://www.jta.org/2023/04/18/global/barcelona-synagogue-defaced-with-anti-israel-graffiti
[4] https://www.israelhayom.com/2024/03/15/from-the-river-to-the-see-spray-painted-on-scene-of-1972-munich-massacre/
[5] https://www.foxnews.com/media/arab-israeli-journalist-speaks-attacked-anti-israel-agitator-columbia-university

eggs and fireworks[6] were hurled during the opening of the new National Holocaust Museum in Amsterdam.[7] The phrase was spay-painted on multiple university and college buildings, properties, and monuments that were sites of the encampments in the US and Europe, with the University of Amsterdam estimating damages of 1.5 million Euros[8] that extended to municipality and public neighborhoods.

As record-numbers of violent antisemitic incidents against Jews continue to soar, including 'revenge' murders, kidnappings, and rapes of perceived Jewish people in the name of suffering in Gaza (see cases in UK[9] and France[10]), it is safe to say that the burden of proof of 'real world harm' that would require additional content moderation restrictions around incendiary rallying cries meant for mobilization has been met.

Online Harm

Online antisemitism is cited in multiple surveys as the top form of Jew-hatred that effects Jewish people on a day-to-day basis, causing them fear for their physical safety, depression, lack of sleep, and even causing them to hide their Jewish identity online and offline. The number one social media platform cited by survey participants as where they most experience digital Jew-hatred was Facebook in the American Jewish Council's State of Online Antisemitism survey in 2023.[11]

While the Oversight Board attempts to offer important guidance on the use of this phrase by bringing content moderation cases for review, these cases do not address the way this phrase is weaponized on Meta's platforms in the comment section. Aside from the experience of Jewish and Israeli users seeing this phrase calling for the destruction of their homeland normalized on their favorite social media apps, the weaponization of this phrase, alongside the use of the Palestinian flag, the inverted red triangle (violent Hamas propaganda symbol signaling targets for murder), and many pro-Palestinian hashtags that do not inherently violate community standards, are being used to promote online bullying and harassment by flooding the comments section of non-political Jewish and Israeli accounts. This harassment continues with zero systematic intervention on the part of Meta.

---

[6] https://uk.news.yahoo.com/protests-outside-holocaust-museum-israeli-133500801.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAAeCGQHwayMgQNXE8WkO2ECApf6E4ZI32m23z90NkalTdkt67TCi51-szo2xiVwbvLA_7oXdJHlY09WHi9G61wNkGvW4faEuiKi6urzmhEtOQZQauJCGTewPdvzmgEn3ng-VDRC7qG1vXBgYmpTqI6RSDQNqttEuTDVw-lkj-Hqb

[7] https://www.dutchnews.nl/2024/03/hundreds-gather-to-protest-at-israeli-presidents-presence-in-nl/

[8] https://www.dutchnews.nl/2024/05/amsterdam-university-shuts-for-two-days-after-more-protests/

[9] https://apnews.com/article/asylum-seeker-britain-stabbed-retiree-revenge-gaza-f9594df1f1cebb32cb80237c728f3afa

[10] https://www.jpost.com/breaking-news/article-798362

[11] https://www.ajc.org/AntisemitismReport2023/AmericanJews

Learning from Precedent

In 2019-2020 after the QAnon conspiracy theory became popular online, 'jumping' from darker messaging platforms like 4chan to larger mainstream social media platforms via YouTube, such as Facebook, Instagram, and X, companies, including Meta, began taking strong systematic content moderation actions against QAnon-related content and groups, including seemingly 'innocuous' and non-violating content like the slogan, "Where we go one, we go all," and its affiliated hashtag #WWG1WGA, which was often touted by QAnon believers.

Making initial moves against QAnon and militarized Antifa content in August 2020, which was later expanded and formalized to include a new category called militarized social movements, Meta's removal actions resulted in a total removal of 127,000 Facebook accounts, 8,400 Facebook pages, 32,800 Facebook groups, 47,700 Instagram accounts, and over 1,000 Facebook events between August 2020-October 2022. Meta also targeted over 300 affiliated hashtags across both Facebook and Instagram for enhanced moderation.

In an effort to suggest a framework as to how "From the river to the sea" and other seemingly non-violative content against Israelis should be treated given the unprecedented amount of damage and real world harm that has been caused to Jewish communities, Israeli communities, public institutions, places of worship, and Holocaust memorial sites, CyberWell referred to Meta's initial statement, "How we address movements and organizations tied to violence," published on August 19, 2020 (**hereinafter: Militarized Groups Approach**), and asks the Oversight Board to suggest the same in this case.

The Militarized Groups Approach is meant to be applied to individuals who "Have celebrated violent acts, [and] individual followers with patterns of violent behavior." This systematic approach was not only applied to the individuals affiliated with these groups, but also to identifiable phrases and hashtags, even when they did not inherently violate community standards.

There is ample evidence to suggest that the current anti-Israel movement is being actively supported by radical militant groups, who are using pro-Palestinian symbols and slogans to perpetuate violence. During the clearing of Columbia University's solidarity encampment, over half of the arrests made by the NYPD were individuals who were not students or faculty, with law enforcement reporting that many of them were part of known militant activist groups. In Portland, only six of the thirty arrests made at the Portland State University encampment were students. An Antifa affiliated group known as Rachel Corrie's Ghost Brigade participated in the PSU encampment, encouraged students to throw fireworks at 'Zionist settlers', and burned 15 police cars.

The Militarized Groups Approach further imposes restrictions to limit the spread of content from Facebook Pages, Groups, and Instagram accounts where there are discussions of potential violence including the use of **veiled language and symbols particular to the movement to do so.**

Therefore, CyberWell suggests the following content moderation measures be applied from the Militarized Groups Approach to the phrase "From the River to the Sea" due to real-world harm being committed under this slogan.

- **Reduce ranking in news feed**
- **Reduce in search**
- **Reviewing related hashtags that are sparking or mobilizing real world harm**
- **Prohibit use of ads, fundraisers, commerce surfaces and monetization tools when using this phrase**

<u>Appendix</u>

IHRA Type 1

1. https://www.facebook.com/rafiq.ibrahim.9849/posts/pfbid021Lgxe4WrG2qW2yJoKj9iKPZukzspVh85442CqrPXEeyCuESftP3UjrtiwsGryF2Ul
2. https://www.facebook.com/100081437006283/videos/2078717669155119/
3. https://www.facebook.com/hani.ayyash.16/posts/pfbid07yKzuqx9Zu8gw3fYxtEcQdx3qvc61ZyYEKCxWMAJ1FWZJNdp82zzvNsVZFqCmSqWl

IHRA Type 11

1. https://www.facebook.com/rahima.ema/posts/pfbid0p7oBgveyLP7cYyA6RyKBtHaBetSvtxTznCShX67qvueSCbcH1DTLifEsgN1Pzmzil
2. https://www.facebook.com/permalink.php?story_fbid=pfbid035axTxcuNTYNYEd5yAhERLSJJ5bDB223N2jmtW55st2XpFUStLj8i7Grj5wtxoSwSl&id=100090551333293
3. https://www.facebook.com/1169346037/posts/10229418315500206
4. https://www.facebook.com/amit.ruparelia.5/posts/pfbid0msQC4B9AhgqpLaFCq3v11wbjwX9futkm6W68tMrat3rzVjG3Z21wXgSiZ7gzQup9l

Antisemitic comment on a post with the slogan "From the River to the Sea"

1. https://www.facebook.com/permalink.php?story_fbid=pfbid0YjDhysjEtjj8duna8EMDD8hayX6vgpYgMczFH2Pf6Jj3BdfeW8YRsYFJ311iCsJLl&id=100072649719754&comment_id=943546897776831