

אנטישמיות ברשתות החברתיות בערבית ובאנגלית אחרי ה-7 באוקטובר 2023: ניתוח השיח ודרכי מניעה

ורד אנדרייב, לארה פורטנוי | 16 בפברואר, 2025

מחברים תורמים: צוות המחקר של CyberWell, טל-אור כהן מונטמאיור

ב-7 באוקטובר 2023, בזמן שמתקפת חמאס התרחשה בשטח ובמקביל ברשתות החברתיות, טכנולוגיית הניטור של CyberWell זיהתה עד מהרה נסיקה בתוכן אנטישמי. ב-25 הימים שאחרי ה-7 באוקטובר חלה עלייה של 86% בהשוואה ל-25 הימים שקדמו לו. הטכנולוגיה של CyberWell זיהתה גם שינוי ניכר ומדאיג בשיח בערבית, שעבר מקידום סטראוטיפים ישנים ומוכרים על יהודים, למשל היותם חורשי מזימות השולטים בעולם, לעידוד והצדקה של אלימות ישירה נגד יהודים בכלל. במאמר זה אנחנו מציגים את הממצאים וההמלצות שלנו הן לפלטפורמות של מדיה חברתית והן לממשלות – כיצד להתמודד עם הסכנה הברורה, הנשקפת מדיבור אלים ומרטוריקת שטנה בלתי מרוסנים כלפי יהודים, לביטחון הלאומי ולבטיחותן של קבוצות מיעוטים בכלל.

המתקפה האכזרית של ארגון הטרור חמאס ב-7 באוקטובר 2023 התאפיינה בהצפה בו-זמנית של תוכן גרפי ואלים, שהועלה בזמן אמת ובשידור חי ברחבי הרשתות החברתיות. למרבה הצער הפלטפורמות הגדולות נתפסו בלתי ערוכות לטיפול בנפח העצום של תוכן המפר את המדיניות שלהן, והותירו אותו נגיש ברשת במשך שעות ואפילו ימים. הדימויים המחרידים וחגיגת הטבח שחברי ארגון חמאס עצמם העלו זכו אומנם בתחילה לגינוי נרחב, אולם הנרטיב השתנה במהירות לאנטישמי ואלים עד מאוד.

עד מהרה זיהתה טכנולוגיית הניטור של CyberWell זינוק חד באנטישמיות בפלטפורמות הרשתות החברתיות, הן באנגלית והן בערבית.¹ האנליסטים שלנו יישמו שיטות מקצועיות לסיווג פוסטים שונים נגד יהודים, שכללו האדרה והסתה לאלימות, דה-הומניזציה,

¹ CyberWell הוא ארגון ללא מטרות רווח שייעודו מיגור אנטישמיות ברשת באמצעות קידום האכיפה ושיפור כללי קהילה ומדיניות בעניין דברי שטנה במרחב הדיגיטלי. על סמך נתונים אנחנו מבקשים לזהות היכן המדיניות אינה נאכפת והיכן היא אינה מצליחה להגן על משתמשים יהודים מפני הטרדה ושנאה. האנליסטים המקצועיים שלנו הוכשרו בתחומי האנטישמיות, הבלשנות והמדיניות הדיגיטלית. כחלק מהאסטרטגיה שלנו לדמוקרטיזציה של נתונים השקנו במאי 2022 את הפלטפורמה המאפשרת גישה למאגר נתונים פתוח הראשון בעולם לתוכן אנטישמי ברשת.

סטראוטיפים של שנאה, הפצת מידע פוגעני שגוי ואפילו הכחשה שהאירועים הקשים בכלל התרחשו – כחלק מהנרטיב שהיהודים הם שקרנים.

הרשתות החברתיות והמרחב הדיגיטלי בכלל אינם כפופים לפיקוח בדומה לאמצעי התקשורת המסורתיים. הדבר מאפשר חופש תקשורתי, חופש ביטוי והתפשטות מהירה של נרטיבים באופן חסר תקדים. למרבה הצער מרחב זה גם מנוצל לרעה לשם הפצת דברי שטנה, מידע כוזב וקריאות לאלימות נגד יחידים וקבוצות מיעוטים, לרבות יהודים. למרות שרוב פלטפורמות המדיה החברתית הגדולות מבטיחות חוויית משתמש נטולת הטרדות לקבוצות בעלות "מאפיינים מוגנים", ברוב המקרים הן אינן מצליחות לעמוד בהבטחה.²

במאמר זה אנחנו מציגים מערכי נתונים שנאספו ונותחו לפני ואחרי ה-7 באוקטובר 2023, מציגים תובנות ראשוניות על השינוי שחל בשיח האנטישמי בשפה הערבית ודנים בפערי האכיפה הגדולים בפלטפורמות המדיה החברתית באשר לטיפול יעיל באנטישמיות. כמו כן ריכזנו את ההמלצות העיקריות שהעברנו לאותן פלטפורמות בזמן אמת, בחלוקה לארבע קטגוריות של הפרת מדיניות על פי הגדרת הפלטפורמות עצמן.³ ההמלצות שלנו נועדו לסייע לצוותי בדיקת התוכן בזיהוי ובהסרה של תוכן אנטישמי אסור בקטגוריות של אלימות, דברי שטנה, טרור ומידע כוזב. לסיכום מובאות המלצות לממשלות – כיצד לפקח בצורה טובה יותר על פלטפורמות המדיה החברתית כדי להפחית הפצת שנאה בקנה מידה נרחב.

מערכי נתונים ומתודולוגיה

המתודולוגיה של CyberWell לזיהוי תוכן אנטישמי כוללת שימוש בטכנולוגיה המנטרת פלטפורמות דיגיטליות לשם זיהוי מילות מפתח ספציפיות, וזאת תוך שימוש בלקסיקון מיוחד וקנייני התואם את הגדרת העבודה לאנטישמיות של [הברית הבינלאומית לשימור זכר השואה](#) (IHRA). ההגדרה כוללת 11 קריטריונים (קטגוריות) של אנטישמיות. השיטה: קטע תוכן שטכנולוגיית הניטור של CyberWell מסמנת כי סביר להניח שהוא אנטישמי עובר תהליך אימות, הכולל סבב אחד או שניים של בדיקה אנושית. הבדיקה כוללת תיוג התוכן בהתאם ל-11 הקריטריונים של IHRA, ניתוח הפרות מדיניות של הפלטפורמה וזיהוי ביטויים או דימויים אנטישמיים ספציפיים שהפוסט מקדם, לצד קישור לאירועים בעולם האמיתי. מידע נוסף על המתודולוגיה של CyberWell זמין [בהנחיות המדיניות](#).

התובנות המובאות במאמר זה מבוססות על פוסטים שסומנו כאנטישמיים בסבירות גבוהה, וכן על תוכן בערבית ובאנגלית שאומת כאנטישמי לפי המתודולוגיה שהוצגה לעיל. מערכי

² כל חמש פלטפורמות המדיה החברתיות הגדולות ש-CyberWell מנטרת מגינות על קבוצות בהתאם ל"מאפיינים מוגנים" כמו גזע, מוצא אתני, מגדר, השתייכות דתית וכן הלאה. למידע נוסף ראו: <https://tinyurl.com/bddv9etw>; <https://tinyurl.com/2ny7dr7j>; <https://tinyurl.com/2vyzx2t9>; <https://tinyurl.com/47ymrzrr>

³ למידע נוסף על החלוקה לקטגוריות של הפרות מדיניות במתודולוגיה של CyberWell ראו: <https://tinyurl.com/yc86s72j>

הנתונים התמקדו בעיקר בתוכן מחמש פלטפורמות המדיה החברתיות הגדולות המנוטרות ברציפות על ידי CyberWell: פייסבוק, אינסטגרם, X (לשעבר טוויטר), טיקטוק ויוטיוב.⁴

לצורך הדיון הרחב חשוב לציין שזיהוי המיקום הגיאוגרפי של משתמשי רשת חברתית כאשר הם מעלים פוסטים הוא מוגבל מאוד, ולרוב מבוסס על נתונים שהמשתמשים מספקים בתיאור הביו (Bio) בפרופיל שלהם. בין ספטמבר 2023 למארס 2024 זוהה בהצלחה המיקום הגיאוגרפי של 9% מהפוסטים האנטישמיים בערבית שנכללו במערך הנתונים שלנו. בחלק מהמדינות שזוהו, כמו מצרים, ירדן ומרוקו, האוכלוסיות אכן דוברות ערבית. עם זאת זיהינו גם פוסטים בערבית שמקורם בגרמניה, באיטליה ובארצות הברית. תובנה זו מלמדת על זרימת תוכן ורעיונות בין שפות ומיקומים שונים ומדגישה את הצורך בניתוח מעמיק יותר.

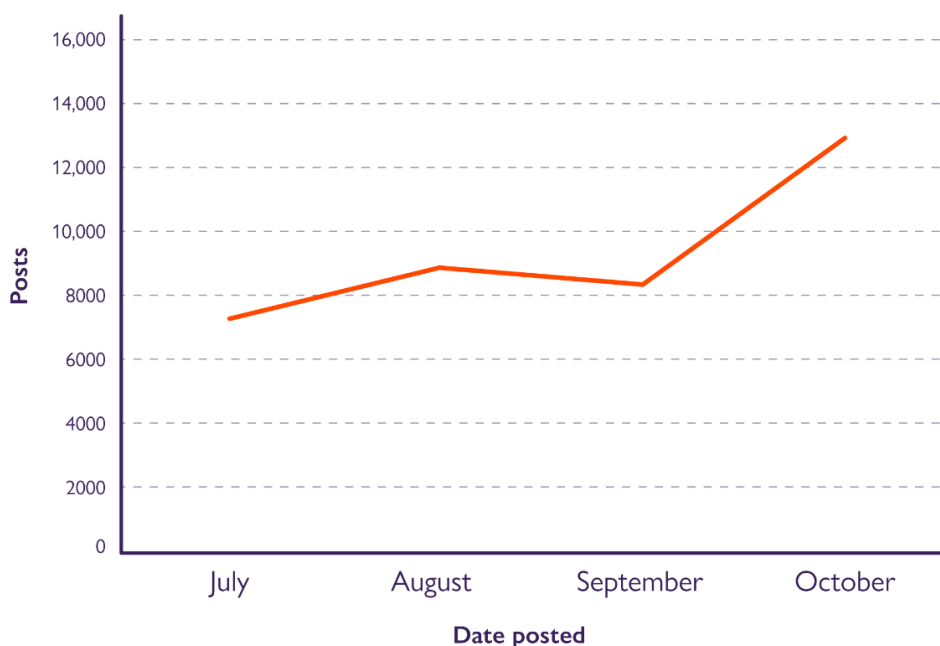
תובנות מהנתונים

גידול כמותי בתוכן ברשתות החברתיות המסומן כאנטישמי בסבירות גבוהה

לפני שנתעמק בניתוח הנרטיב ובשינוי שחל באופי התוכן האנטישמי המאומת מאז ה-7 באוקטובר, חשוב להתעכב על הגידול הכמותי בתוכן שזוהה כאנטישמי בסבירות גבוהה. ב-25 הימים שקדמו למתקפת ה-7 באוקטובר (מ-11 בספטמבר עד 6 באוקטובר), סומנו 6,959 קטעי תוכן כאנטישמיים בסבירות גבוהה. בין 7 באוקטובר ל-31 באוקטובר עלה מספרם ל-12,949. זוהי עלייה חדה של 86% בתוכן המסומן בכל חמש פלטפורמות המדיה החברתית שאותן מנטרת CyberWell באנגלית ובערבית (תרשים 1). משתמשים יהודים ברשתות החברתיות סופגים בקביעות כמות לא-פרופורציונלית של דברי שטנה והטרדה, בייחוד במהלך אירועים מרכזיים שבהם מעורבת ישראל. תופעה זו בלטה בעיקר אחרי הטבח האלים ב-7 באוקטובר, שהיה מכוון במיוחד ליהודים אך גם גרם לשיח האנטישמי להרקיע שחקים.

⁴ CyberWell היא כיום במעמד שותפה מהימנה (trusted partner) הן של מטא (החברה שבבעלותה פייסבוק, אינסטגרם ו-Threads) והן של טיקטוק, ולכן היא יכולה לנקוט פעולות הסלמה במקרה של תוכן מפר מדיניות ולייצג ישירות לצוותי הגבלת התוכן. מאז ה-7 באוקטובר הגישה CyberWell שמונה דוחות למטא, X, טיקטוק, וכן מידע נוסף שנשלח לגוגל.

תרשים 1. פוסטים שסומנו על ידי טכנולוגיית ה-AI של CyberWell כאנטישמיים בסבירות גבוהה



2023 * These numbers represent the posts across all social media platforms that CyberWell's monitoring tech flagged as highly likely to be antisemitic.



* התרשים מבוסס על מספר הפוסטים בכל פלטפורמות המדיה החברתית שטכנולוגיית הניטור של CyberWell מסמנת כאנטישמיים בסבירות גבוהה | **מקור:** CyberWell

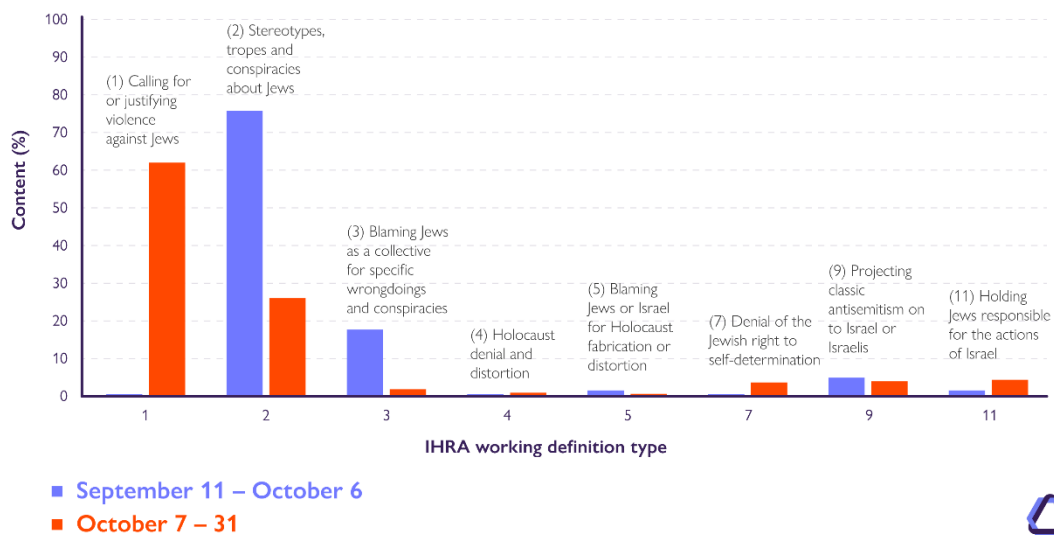
בעת בחינת התוכן של פוסטים בערבית, שהועלו בפרק הזמן שנחקר ואומתו כאנטישמיים, הבחנו בשינויים בשיח שהתרחשו בשני ממדים: בקריטריונים המרכזיים של IHRA ובנרטיבים האנטישמיים המובילים.⁵

תרשים 2 מציג את השינוי הברור בנרטיב, מאנטישמיות על פי קריטריון 2 בהגדרת IHRA (דה-הומניזציה והאשמות סטראוטיפיות) לפני ה-7 באוקטובר לאנטישמיות על פי קריטריון 1 בהגדרת IHRA (קריאה לאלימות נגד יהודים או הצדקה של אלימות כזו) לאחר מתקפת חמאס. התרשים מצביע גם על עלייה באנטישמיות על פי קריטריון 11 בהגדרת IHRA – הטלת אחריות קולקטיבית על יהודים לפעולות של מדינת ישראל או צה"ל. להאשמות כאלה יש השפעה

⁵ מאמצי הניטור של CyberWell מתעדכנים באופן קבוע ומותאמים לאירועים בעולם האמיתי. ההשוואה אומנם אינה מדויקת, אך היא מספקת אינדיקציה לשינוי בשיח הכללי. למידע נוסף על הגדרת העבודה של IHRA ראו: International Holocaust Remembrance Alliance, "Working Definition of Antisemitism". <https://tinyurl.com/uabx9rmc>

ישירה על יחידים וקהילות של יהודים החיים מחוץ לישראל, כולל מי שאינם בהכרח מזדהים עם ישראל או עם מדיניותה ופעולותיה.

תרשים 2. פירוט קריטריוני IHRA בערבית לפני ואחרי ה-7 באוקטובר 2023



מקור: CyberWell

להלן שלוש תמונות שנלקחו מסרטון בטיקטוק שפורסם באוקטובר 2023. הסרטון קורא לשואה נגד יהודים בעקבות המלחמה בעזה ומסווג כאנטישמי על פי קריטריון 11 בהגדרת IHRA (תפיסת היהודים כאחראים קולקטיבית לפעולותיה של מדינת ישראל) ועל פי קריטריון 1 (קריאה להרג או לפגיעה ביהודים, או סיוע או הצדקה לכך).

תמונה 1. מזרח ירושלים ודגל פלסטין. הכיתוב: היכן אתה היטלר?



מקור: מתוך סרטון בטיקטוק

תמונה 2. מזרח ירושלים ודגל פלסטין. הכיתוב: עזה זועקת לעזרה



מקור: מתוך סרטון בטיקטוק

תמונה 3. חיל ליד משרפה. הכיתוב: משפחה יהודית 1941



מקור: מתוך סרטון בטיקטוק

שינויים בנרטיבים של אנטישמיות בשפה הערבית

מהשוואת הנרטיבים האנטישמיים הבולטים בתוכן הערבי לפני ואחרי ה-7 באוקטובר, הנכללים במערך נתוני המחקר, עולה שחל שינוי כולל בשיח (ראו טבלה 1). לפני ה-7 באוקטובר התמקדו התמות האנטישמיות הבולטות בתאוריות קונספירציה על שאיפות יהודיות לשליטה עולמית, כולל טענות בדבר שליטה ציונית,⁶ או חורשי מזימות סודיים כפי שמופיע ב'פרוטוקולים

⁶ אחת מתאוריות הקונספירציה האנטישמיות הנפוצות ביותר טוענת שיהודים חותרים לשליטה בעולם דרך השתלטות על הכלכלה, התקשורת והממשלות. בהקשר זה חשוב מאוד לציין שגם ברשת וגם מחוץ לרשת, המונח ציוני משמש לעיתים קרובות כחלופה ליהודי או לעם היהודי, כדי להתחמק מגילוי של מגבילי תוכן בפלטפורמות מדיה חברתית ולהימנע מביקורת בינלאומית. על ידי הימנעות משימוש ישיר במילה יהודי, פוסטים אנטישמיים עלולים לחמוק מתחת לרדאר ולהישאר ברשת למרות הפרה של כללי הקהילה והמדיניות לגבי דברי שטנה.

של זקני ציון⁷, לצד ההאשמה שיהודים הסתננו לארגון הבונים החופשיים והשתלטו עליו.⁸ אחרי ה-7 באוקטובר נעשו הנרטיבים שטופי שנאה ומפלגים במידה רבה יותר, תוך מסגור בנימה של "אנחנו והם". השינוי כלל תוכן העושה דה-הומניזציה ליהודים דרך השוואתם לבעלי חיים כמו עכברושים, קופים וחזירים,⁹ הצגת כלל היהודים כרשעים,¹⁰ ובעיקר תיאור היהודים באופן קולקטיבי כ"האויב", לפעמים על רקע אנטישמיות דתית. האחוזים בטבלה מייצגים את שיעור הפוסטים במערך הנתונים שסווגו כמקדמים את האמירות הספציפיות, מתוך סך הפוסטים בערבית שאומתו כאנטישמיים וסווגו על ידי CyberWell במסגרת הזמן שנחקרה. כל פוסט עשוי להתייחס ליותר מנרטיב אנטישמי אחד.

⁷ הפרוטוקולים של זקני ציון' (הפרוטוקולים) הוא מסמך אנטישמי מזויף המפרט לכאורה מזימה יהודית להשתלטות על העולם. הוא נכתב כביכול על ידי חורשי מזימות יהודים שתכננו להשיג שליטה בעולם. ככל הנראה הפרוטוקולים נכתבו בתחילת המאה ה-20 באימפריה הרוסית, לפני סדרה אלימה במיוחד של פוגרומים אנטי-יהודיים שהתרחשו בין השנים 1903 ל-1906. למרות שהוכח כי מדובר במסמך מזויף שמטרתו להסית לשנאת יהודים, הטקסט ממשיך לשמש בסיס לקונספירציות אנטישמיות רבות, כולל תעמולה נאצית. למידע נוסף על האופן שבו נעשה שימוש בפרוטוקולים כדי להאשים יהודים באירועים בעולם האמיתי, ראו:

CyberWell, "World Cup 2022: Sports Uniting Peoples or a Perfect Storm of Antisemitism?" blog, January 10, 2023. <https://tinyurl.com/2s44bwz9>

⁸ ארגון הבונים החופשיים במתכונתו המודרנית נוסד רשמית ב-24 ביוני 1717, לאחר שכמה לשכות קטנות התאחדו כדי ליצור את "הלשכה הגדולה" הראשונה בלונדון. כיום ניתן למצוא לשכות של הבונים החופשיים בכל העולם. הארגון הידוע במבנה ההיררכי שלו ובחשאיותו הוא נושא אהוב על העוסקים בתאוריות קונספירציה, והחברים בו מוצגים לא פעם כרשעים זוממים לשלוט בעולם. נגזרת של תאוריות אלו היא תאוריית הקונספירציה המשלבת את הבונים עם היהודים ומאשימה את העם היהודי בהסתננות לבונים החופשיים כדי לקדם מטרה לכאורה של שליטה בעולם, ובכך מנציחה סטראוטיפ אנטישמי קלאסי. מעניין לציין שלשכות רבות של הארגון סירבו לקבל לשורותיהן חברים יהודים במשך יותר מ-200 שנה. ראו:

CyberWell, "The Judeo-Masonic Conspiracy Theory on Meta," Report, June 2023. <https://tinyurl.com/3envjskv>

⁹ ההתייחסות ליהודים או ל"עם הספר" כאל קופים וחזירים מקורה בכמה פסוקים בקוראן (לדוגמה, 5:59-60) ובחדית', והיא נפוצה בשיח האסלאמיסטי כחלק מפרשנות שנויה במחלוקת. ראו:

Ofir Winter, *Peace in the Name of Allah: Islamic Discourses on Treaties with Israel* (De Gruyter, 2022), 120; Neil J. Kressel, 'The Sons of Pigs and Apes': Muslim Antisemitism and the Conspiracy of Silence, (Potomac Books, 2012), 23-27.

¹⁰ "יהודים הם רוע" הוא אחד הנרטיבים האנטישמיים הנפוצים. יש לו צורות שונות, אך בדרך כלל הוא מתמקד במאפיינים שלכאורה הם חלק בלתי נפרד מהאופי היהודי, ומקדם את התפיסה שיהודים הם רעים מטבעם.

טבלה 1. השינוי בשיח בערבית לפני ואחרי ה-7 באוקטובר 2023

7 באוקטובר - 31 באוקטובר		11 בספטמבר - 6 באוקטובר	
נושא	שיעור	נושא	שיעור
הפרוטוקולים של זקני ציון	35%	יהודים הם האויב	94%
הבונים החופשיים	35%	היהודים כחיות	4%
שליטה/השתלטות של הציונים	35%	יהודים הם רוע	3%

תמונה 4 נלקחה מהטיקטוק אחרי ה-7 באוקטובר. הוא מתאר תמונת AI של נחש שעורו בצבעי דגל ישראל, המלפף את כדור הארץ. הכיתוב הוא (בתרגום חופשי): זכרו את חייבר, יהודים, צבא מוחמד עוד ישוב.¹¹ הפוסט מקדם את הנרטיב שיהודים וישראל הם אויב מרושע, והטקסט קורא לאלימות נגד יהודים.

¹¹ הקריאה "זכרו את חייבר, יהודים, צבא מוחמד עוד ישוב" מתייחסת לקרבות שבהם לחם הנביא מוחמד נגד קהילות יהודיות ספציפיות במהלך המאה ה-7 לספירה, אשר הביאו לטבח המוני של יהודים. עוד על כך ראו להלן בתת-הפרק 'פער אכיפה ספציפי 1, דוגמה 2'.

תמונה 4. פוסט המציג את היהודים ואת ישראל כאויב מרושע



מקור: פוסט בטיקטוק (שהוסר בהמשך)

תמונה 5 מציגה פוסט בערבית שעלה ב-X (לשעבר טוויטר) ובו כמה קונספירציות אנטישמיות, כולל האשמת היהודים בעוולות שונות, השוואות לבעלי חיים באופן שעושה דה-הומניזציה וטענה שיהודים הם צאצאי השטן.

תמונה 5. פוסט המשתף בתאוריות קונספירציה אנטישמיות



מקור: חשבון @UmLucaM ב-X (טוויטר), 12 באוקטובר 2023

תמונה 6 מציגה פוסט בשפה הערבית שעלה ב-X ומצדיק את השואה, הוא פורסם 10 ימים לאחר טבח ה-7 באוקטובר וצבר יותר משני מיליון צפיות.

תמונה 6. פוסט בערבית המצדיק את השואה



מקור: חשבון @mashhad0 (טוויטר), 18 באוקטובר 2023

מאמצים גלובליים להסדרת פיקוח על פלטפורמות מקוונות

הטיפול בדברי שטנה בפלטפורמות המדיה החברתית נעשה ברמות שונות ועל ידי רשת מורכבת של שחקנים, ובכלל זה קובעי מדיניות, ארגונים ללא מטרות רווח ומותגים גלובליים.

פיקוח ממשלתי

ברמת המדיניות, ההחלה של חוק השירותים הדיגיטליים (DSA) של האיחוד האירופי וחוק הבטיחות ברשת שחוקקה בריטניה מאתגרים כעת את הסלחנות כלפי ניטור והסרה של תוכן שנאה, שהפלטפורמות נהנו ממנה בעבר. אחרי ה-7 באוקטובר שיגר האיחוד האירופי מכתבים לאילון מאסק ולמארק צוקרברג על סמך חוק ה-DSA. המכתבים התריעו מפני אי-ציות לתקנות ה-DSA בדבר הסרת דברי שטנה בלתי חוקיים. הפרה כזו יכולה לגרור עונשים חמורים בדמות קנסות של עד 6% מההכנסה הגלובלית של החברות.

פיקוח עצמי של הרשתות החברתיות

לזכותן של פלטפורמות המדיה החברתית ייאמר כי במידה רבה יש להן מדיניות ברורה לגבי דברי שטנה והשלכות שמשמשות עלולים לספוג בנין העלאת תוכן שנאה אסור.¹² בעקבות בחינת מצב האנטישמיות בפלטפורמות המדיה החברתית, CyberWell מזהה ומספקת המלצות לגבי פערים שזוהו הן באכיפת המדיניות והן בטיפול בצורות מורכבות, רבות דקויות או מודרניות של אנטישמיות. הצוות שלנו זיהה פערים בין המדיניות הכתובה לאכיפתה בפועל בכל חמש פלטפורמות המדיה החברתית המנוטרות. הסבר חלקי אפשרי לפער הזה הוא חוסר השקעה בצוותי הגבלת תוכן (content moderation). לדוגמה, [בדוח השקיפות האחרון](#) [מכוח חוק DSA](#) טענה רשת X (לשעבר טוויטר) כי צוות הגבלת התוכן שלה כולל רק 12 איש דוברי ערבית, לעומת 2,294 אנשי צוות דוברי אנגלית.

כחלק מהמתודולוגיה השוטפת שלנו, CyberWell מעריכה את הפעולות שנוקטות הפלטפורמות בהתאם למדיניות ולאסטרטגיות של כל אחת מהן. באופן כללי הפלטפורמות אוכפות מדיניות כלפי דברי שטנה בשתי דרכים עיקריות: הראשונה היא הסרת התוכן; השנייה היא צמצום הפצתו כדי להגביל את הנראות שלו. המתודולוגיה השנייה היא השיטה העיקרית שמיושמת ברשת X, כחלק מגישה שהם מכנים ["Freedom of Speech, Not Reach"](#) ("חופש הביטוי, לא חופש ההפצה"), המדגישה את צמצום ההפצה של דברי שטנה במקום הסרתם. דרך שלישית שחלק מהפלטפורמות מאמצות היא חסימת היכולת לאתר האשטאג מסוים או מילות מפתח ספציפיות בחיפוש. אומנם אמצעי זה אינו מונע ממשתמשים לפרסם תוכן עם ביטויים מסוימים, וגם לא את האפשרות שמשתמשים ייתקלו בתוכן כזה שלא מרצונם, אך הוא מונע ממשתמשים לחפש באופן יזום תוכן שתויג עם אותם ביטויים. לאחרונה גם החלו כמה פלטפורמות להציג הערות למשתמשים המחפשים תוכן שנוי במחלוקת. לדוגמה, משתמשים ביוטיוב שיעשו חיפוש למילים "הפרוטוקולים של זקני ציון" יראו בחלק העליון של

¹² בעוד שאנטישמיות נחשבת לסוג של דברי שטנה, תוכן אנטישמי במדיה החברתית יכול להפר גם מדיניות הקשורה להסתה ל אלימות, להטרדה, למידע כוזב ועוד.

שורת החיפוש קישור לערך בוויקיפדיה עם הכיתוב "הפרוטוקולים של זקני ציון הוא מסמך אנטישמי מזויף שעומד במרכזן של תאוריות קשר רבות, הטוענות שהעם היהודי שולט בעולם".

פערי אכיפה

פיקוח ממשלתי

המאמצים של האיחוד האירופי ובריטניה לרסן דברי שטנה ברשת אומנם ראויים לשבח, אולם יש צורך בחקיקה דומה גם בארצות הברית. סעיף 230 בחוק האמריקאי העניק חסינות לפלטפורמות מדיה חברתית במשך זמן רב מדי, ופטר אותן מאחריות לתוכן של צד שלישי המתארח בפלטפורמות שלהן. CyberWell מעודדת קובעי מדיניות בארצות הברית ללכת בעקבות האיחוד האירופי ובריטניה ולהציב את התיקון לסעיף 230 בראש סדר העדיפויות. ניתן וצריך לעשות יותר כדי לגשר על פערי האכיפה ולהגן על המשתמשים מפני ההשפעות המזיקות של דברי שטנה ברשת. נוסף על תיקון סעיף 230 יש צורך בתקנות מקיפות, שיחייבו את פלטפורמות המדיה החברתית להטמיע אסטרטגיות הגבלת תוכן יעילות יותר ולהבטיח שקיפות בפעולות האכיפה שלהן, כפי שיפורט בפרק המסקנות של המאמר.

פיקוח על פלטפורמות

למרות שאף אחת מהפלטפורמות הגדולות לא אימצה רשמית את הגדרת העבודה של IHRA לאנטישמיות כמדיניות, ולא שילבה אותה בהנחיות האכיפה שלה, בדרך כלל ייאסרו אמירות אנטישמיות שמתאימות לקריטריונים השונים של IHRA בשל כללי הקהילה של הפלטפורמות. כחלק מהמתודולוגיה שלנו בדקנו כל פריט תוכן לא רק לצורך איתור אנטישמיות אלא גם לבחינה אם יש בו הפרה כלשהי של מדיניות הפלטפורמה שבה הוא עלה, וכן נבדקה אכיפת כללי הקהילה של הפלטפורמה. מאז תחילת הפעילות זיהתה CyberWell בכל הפלטפורמות שני פערי אכיפה מבוססי שפה, והיא ממשיכה לעקוב אחריהם.

פער אכיפה כללי 1: שיעור הסרה של תוכן בערבית לעומת תוכן באנגלית

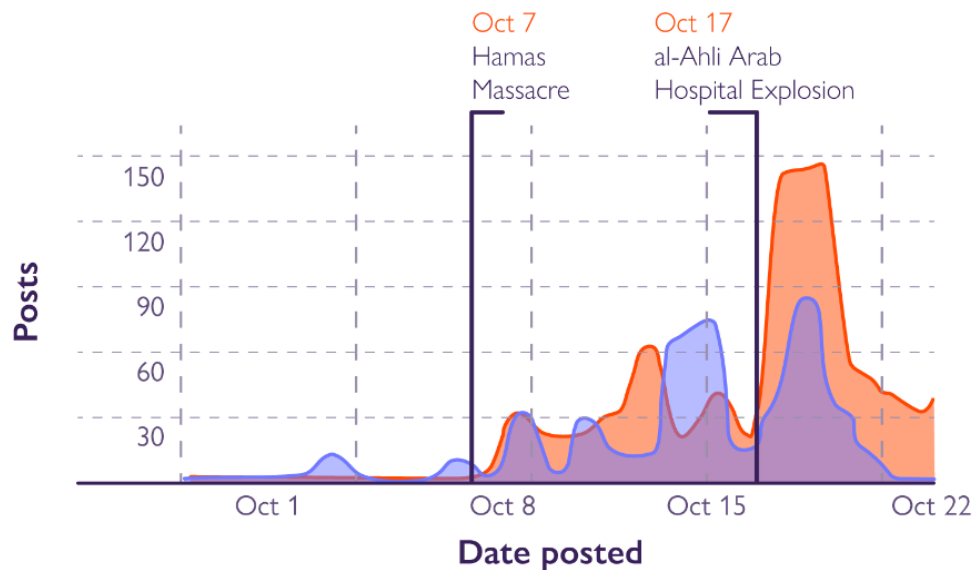
אחת מהתובנות הראשונות של CyberWell שהועברו לפלטפורמות המדיה החברתית הייתה ההבדל הבולט בשיעור ההסרה של תוכן מפר מדיניות בשפה האנגלית, בהשוואה לשפה הערבית. לפי בדיקה שבוצעה בספטמבר 2023, שיעור ההסרה של תוכן שדווח כמפר מדיניות בכלל הפלטפורמות עמד על 24% באנגלית, לעומת 13% בלבד בערבית. ביקשנו מהפלטפורמות לטפל בפער הזה, ונכון לבדיקה עוקבת במאוס 2024 המצב השתפר. ביוני 2024 עלו שיעורי ההסרה הכוללים בפלטפורמות ל-29% עבור תוכן באנגלית המפר מדיניות, ול-21% בשפה הערבית. המאמצים שנקטו הפלטפורמות כדי לשפר את הגבלת התוכן בערבית ראויים לשבח. יש לעודד אותן להמשיך, כדי לצמצם את פער השפה ולהעלות את שיעורי ההסרה של תוכן אסור בכל השפות, הן בעיתות מלחמה והן בעיתות שלום.

פער אכיפה כללי 2: פעולת הגבלה לגבי ביטויים דומים

השוואה בין ביצוע או אי-ביצוע פעולת הגבלה של פלטפורמות לגבי תוכן עם ביטויים דומים בערבית, לעומת תוכן כזה באנגלית, חושפת פער בולט. דוגמה לפער הגבלה כזה היא למשל בביטוי "היטלר צדק" באנגלית ובערבית. בכלל המרחב הדיגיטלי זהו שתי עליות חדות בשימוש בביטוי זה, הן באנגלית והן בערבית – אחת בעקבות מתקפת חמאס ב-7 באוקטובר 2023 והשנייה לאחר הפיצוץ בבית החולים אל-אהלי בעזה ב-17 באוקטובר, שגבה קורבנות רבים.¹³ באנגלית הביטוי צבר פופולריות בהאשטאג #hitlerwasright, עם טווח הגעה פוטנציאלי של יותר מ-25 מיליון משתמשי רשתות חברתיות ב-18 הימים שאחרי ה-7 באוקטובר.¹⁴ בערבית חלה עלייה חדה בתפוצת הביטוי כשילוב של מילים ולא כהאשטאג, שהגיע לקהל פוטנציאלי של כחצי מיליון משתמשים באותו פרק זמן.

¹³ בהתבסס על טענות חמאס ומשרד הבריאות בעזה, ערוצי חדשות דיווחו תחילה כי הפיצוץ נגרם מתקיפה אווירית ישראלית, אולם ישראל הכחישה. ברשת הייתה התפרצות זעם שהופנתה נגד ישראל, כאשר חלק אף קראו לאלימות והצדיקו אלימות לא רק נגד ישראלים אלא נגד יהודים בכל מקום. מספר ימים לאחר מכן הגיעו מומחים בלתי תלויים למסקנה כי הנזק נגרם משיגור כושל של טיל על ידי הג'האד האסלאמי. ראו: CyberWell, "Fighting Hitler- Inspired Jew-Hatred on the Digital Frontlines," October 20, 2023. <https://tinyurl.com/42nrcjpt>
Michael Biesecker, "New AP Analysis of Last Month's Deadly Gaza Hospital Explosion Rules Out Widely Cited Video," November 22, 2023. <https://tinyurl.com/5aur355p>
¹⁴ טווח הגעה פוטנציאלי של מילות חיפוש או של ביטויים מחושב באמצעות הכלים שלנו להאזנה לרשתות חברתיות, שעוקבים אחר פורומים וחלק מהרשתות החברתיות, בהתבסס על מדגם של פוסטים עם מעורבות (engagement) גבוהה. מספרים אלו מצביעים על שינויי דפוס אך אינם מוחלטים.

תרשים 3. עליות חדות ברשת של ההאשטאג #HitlerWasRight באנגלית ובערבית



המונח בערבית שנוטר הוא: هتلر كان على حق (היטלר צדק). מספר הפוסטים מבוסס על מדגם של פוסטים עם שיעורי מעורבות גבוהים, אשר מופו באמצעות כלי האזנה למדיה חברתית.

מקרא: כחול/סגול מציין אנגלית; כתום מציין ערבית.

מקור: CyberWell

ההאשטאג #hitlerwasright כבר זינק בעבר בעקבות ההצהרות האנטישמיות של הראפר Ye (קניה ווסט) באוקטובר 2022. בעקבות [הדוח של CyberWell](#) על אנטישמיות כתגובה להצהרות האנטישמיות של Ye או הגנה עליהן, זיהו הכלים שלנו ירידה של 91% בשימוש בהאשטאג זה ב-X. המשמעות היא שהפצתו צומצמה, אם כי הוא לא הוסר.

אחרי ה-7 באוקטובר התריעה CyberWell בפני פלטפורמות המדיה החברתית על התעוררות מחודשת של הביטוי, ובעקבות זאת ההאשטאג באנגלית נחסם לחיפוש ב-X; עם זאת לא ננקטו צעדים כדי לחסום את אותו ביטוי בערבית.¹⁵ ניתן להבין מכך כי מגבילי תוכן זיהו את

¹⁵ הגישות השונות של פלטפורמות המדיה החברתית להאשטאגים לעומת ביטויים שמקדמים את אותו תוכן ניכרו במקרה של האשטאג השנאה #thenoticing. בעקבות ההתרעות החוזרות ונשנות של CyberWell לגבי התוכן וההיקף המטרידים של הביטוי הזה, X ומטא חסמו שתיהן את ההאשטאג לחיפוש, אך לא חסמו את הביטוי אם הוא הופיע בתוך שילוב של מילים. הביטוי ממשיך להפיץ קונספירציות אנטישמיות וגזעניות. ההאשטאג #thenoticing קורא לציבור "להתעורר" ל"אמת". "אמת" זו כוללת תאוריות קונספירציה שונות על יהודים החודרים לעמדות מנהיגות עולמיות מרכזיות כדי להשיג שליטה בעולם – קונספירציה אנטישמית קלאסית עם שינוי מודרני. ראו:

CyberWell, "Jewish People Take the Blame for Violent Protests in France," July 17, 2023.
<https://tinyurl.com/yj3ibbz7>

הביטוי כמפר את מדיניות הפלטפורמה אך נקטו פעולה בשפה אחת בלבד, למרות המלצתנו להתייחס אליו בשתי השפות.

אף כי רשת X מעדיפה צמצום חשיפה¹⁶ על פני הסרה, הביטוי האלים ומפר המדיניות (בערבית) הגיע לקהל פוטנציאלי של חצי מיליון משתמשים דוברי ערבית. אומנם טווח ההגעה של ההאשטאג באנגלית היה גדול יותר, אבל ראוי היה שהגרסה הערבית תזכה לכל הפחות טיפול דומה. טווח הגעה פוטנציאלי כה גדול למדגם כל כך קטן של פוסטים, המסיתים במפורש לאלימות נגד יהודים ומקדמים דברי שטנה המתייחסים לשואה, מעמיד בספק את יעילותה של גישת "חופש הביטוי, לא חופש ההפצה".¹⁷

פער אכיפה ספציפי 1: אלימות

דוגמה 1: "PUBG the Jews". בשנת 2020 נשא מחמוד אל-חסנאת, דרשן וסופר פלסטיני,¹⁸ דרשה ידועה לשמצה (ראו תמונה 7) שבה הצהיר: "אומרים שהדור הזה הוא דור ה-PUBG, נכון? אבל הסירו להם את הגבולות, ובאלוהים, הם יעשו PUBG ליהודים".¹⁹ PUBG מתייחס כאן למשחק הווידאו PUBG: Battlegrounds. זהו משחק ירי אלים בגוף ראשון,²⁰ שבו שחקנים צונחים על אי, אוספים כלי נשק ויורים באנשים. בעקבות הטבח ב-7 באוקטובר הפכה הדרשה לווריאנט, ומשתמשים שתלו את קטע השמע בתוך סרטונים של מצנחי הרחיפה של חמאס כשהם חוזרים לשטח ישראל, כפי שקרה ב-7 באוקטובר, וכן בתוך סרטוני משחק PUBG המדמים ירי ובסרטונים המציגים כלי נשק.

¹⁶ צמצום חשיפה (De-amplification) הוא סוג של אכיפה המשמשת פלטפורמות מדיה חברתית להפחתת החשיפה של פוסטים מבלי להסיר אותם ישירות. כאשר מצמצמים חשיפה של פוסט הוא נשאר בפלטפורמה, אך טווח ההגעה שלו מוגבל באופן ניכר. המשמעות היא שפחות משתמשים ייתקלו בתוכן בפייד שלהם, בתוצאות החיפוש או בהמלצות האלגוריתם שיקבלו. באופן עקרוני הפלטפורמה ממזערת את פוטנציאל התפשטות הפוסט אך גם מאפשרת לו להתקיים.
¹⁷ מדיניות "חופש הביטוי, לא חופש ההפצה", שהוצגה על ידי טוויטר (כיום X), מתמקדת במתן אפשרות למשתמשים להתבטא בחופשיות תוך הגבלת עוצמת החשיפה של תוכן מזיק או מפר מדיניות. למידע נוסף ראו:

<https://tinyurl.com/mr3axrf4>

¹⁸ מחמוד סלמאן ג'בריל אל-חסנאת נולד במחנה הפליטים ג'באליה ב-1989, גדל בעזה וכיום מתגורר בטורקיה. הוא משפיען בולט בעולם הערבי, זכה בתחילה לתהילה ביוטיוב ובהמשך התרחב לרוב פלטפורמות המדיה החברתית. בתוכן שלו מופיעות לעיתים קרובות דרשות על הסכסוך הישראלי-פלסטיני, שחלקן כוללות אמירות אנטישמיות בוטות ומביעות תמיכה בטורר או בארגוני טרור. CyberWell מנטרת את דברי אל-חסנאת מאז תחילת מלחמת ישראל-חמאס, ומספר העוקבים שלו בחשבונות המדיה החברתית השונים ממשיך לגדול במהירות.

¹⁹ הפוסט המקורי עם הדרשה הוסר מהחשבונות הרשמיים של אל-חסנאת, אך הסרטון של הדרשה פורסם בהרחבה על ידי משתמשים אחרים. ראו למשל תמונה 7 – צילום מסך מפוסט 1 ב-X (לשעבר טוויטר), 13 באוקטובר 2023.

²⁰ משחק בגוף ראשון הוא סוג של משחק וידאו שבו השחקן חווה את עולם המשחק דרך עיני הגיבור או הדמות המשחקת.

תמונה 7. צילום מסך מתוך סרטון הדרשה של מחמוד אל-חסנאת



מקור: מתוך חשבון @lamia_g5k, X (לשעבר טוויטר), 13 באוקטובר 2023

על פי הניתוח של CyberWell, לשילוב של המונחים בערבית "בجي" [PUBG] עם "יהוד" [יהודים] או עם תמונות של גלשונים/טרוריסטים חמושים היה סיכוי גבוה לכלול תוכן אלים הקורא לפגיעה ביהודים. בין 7 ל-29 באוקטובר היה לביטוי הזה טווח הגעה פוטנציאלי של 53,800 משתמשים, על פי כלי האזנה לרשתות חברתיות שמשמשים את CyberWell. פלטפורמות המדיה החברתית כשלו תחילה בזיהוי הקריאה הברורה הזו לאלימות, ואפשרו לפוסטים הללו להישאר זמינים. בעקבות ההתרעה מ-CyberWell, שכללה המלצות מעשיות, הוסרו מטיקטוק מספר לא מבוטל של סרטונים עם מגמה זו.

דוגמה 2: "זכרו את ח'ייבר". הקריאה "זכרו את ח'ייבר, יהודים, צבא מוחמד עוד ישוב" מתייחסת לקרבות שבהם לחם הנביא מוחמד נגד קהילות יהודיות מסוימות במאה ה-7 לספירה. מכיוון שקרבות אלו הסתיימו בטבח המוני של יהודים, משמעות השימוש בסיסמה זו כיום, בייחוד לצד תמונות של אנשים חמושים, היא הסתה לאלימות נגד היהודים כקולקטיב.²¹

²¹ לדיווח מפורט בנושא זה ראו:

בין 7 באוקטובר ל-30 בנובמבר זכה הסלוגן הזה בערבית [לטווח הגעה פוטנציאלי](#) של יותר מ-596 מיליון משתמשים, בהשוואה ל-1.8 מיליון בחודשיים שקדמו ל-7 באוקטובר.

דוגמה 3: הסיסמה של החות'ים. סיסמה דתית נוספת שמסיתה באופן תדיר לאלימות נגד יהודים היא "מוות לאמריקה, מוות לישראל, קללה על היהודים, ניצחון לאסלאם" של החות'ים, שהוכרו כארגון טרור על ידי [משרד החוץ האמריקאי](#) בתחילת 2024. עלייה חדה בשימוש בסיסמת השנאה והאלימות הזו בערבית אותרה אחרי ה-7 באוקטובר, לפני מתקפת הטילים של החות'ים (ראו תרשים 4),²² ומ-24 בספטמבר עד 24 באוקטובר עמד טווח ההגעה הפוטנציאלי שלה על שישה מיליון משתמשי רשתות חברתיות. ביטוי זה ראוי לניטור צמוד משום שלעיתים קרובות הוא משמש כקריאה לאלימות. CyberWell העבירה המלצות בנושא זה לפלטפורמות מדיה חברתית.

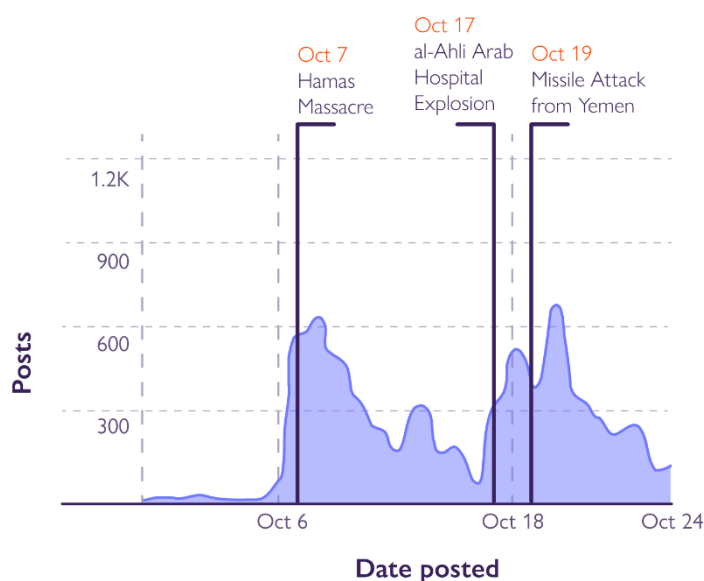
CyberWell, "Jenin, Jew-Hatred, and Incitement to Violence on Social Media," July 7, 2023.

<https://tinyurl.com/36ydr6ub>

²² ארגון הטרור החות'י בתימן שיגר מספר התקפות טילים וכטב"מים לעבר ישראל והים האדום כאקט של סולידריות ותמיכה בחמאס בעזה. תרשים 4 מתייחס למתקפת הטילים ששוגרה ב-19 באוקטובר 2023. ראו:

Maha El Dahan, "Who Are Yemen's Houthis and Why Did They Attack Israel," Reuters, November 1, 2023. <https://tinyurl.com/e52cc2td>

תרשים 4. הפופולריות ברשת של הסיסמה הח'ותית "מוות לאמריקה, מוות לישראל, קללה על היהודים, ניצחון לאסלאם"



מקור: CyberWell

פער אכיפה ספציפי 2: דברי שטנה

הניתוח זיהה שני ביטויים אנטישמיים מרכזיים המפירים את מדיניות פלטפורמות המדיה החברתית לגבי דברי שטנה:

1. הצגת יהודים באופן קולקטיבי כאויב.

2. השוואת יהודים לבעלי חיים (למשל, חולדות, חזירים, קופים, כלבים) – צורה של דה-הומניזציה האסורה בכל פלטפורמות המדיה החברתיות הגדולות.

CyberWell ממליצה כי הפלטפורמות ינטרו מקרוב את השילוב של הביטוי בערבית "يهود" (יהודים) עם "عدو" (אויב), שכן יש סבירות גבוהה לכך שהוא בבחינת הפרה של המדיניות. אנחנו ממליצים גם כי פלטפורמות ינטרו את השימוש במונחים בערבית "قرد, خنازير, كلاب" (קופים, חזירים, כלבים) יחד עם "يهود" (יהודים), בערבית ובאנגלית. לתוכן זה סיכוי גבוה להפר את מדיניות דברי השטנה ולחייב הסרה גורפת.

פער אכיפה ספציפי 3: תמיכה בטרור

נוסף על מתן הנחיות לגבי תוכן אנטישמי, אחרי ה-7 באוקטובר החלה CyberWell לנטר ולהציע המלצות גם בנושאים קשורים, כגון דברי תמיכה ושבח לטרור. להלן דוגמאות לפעולות שהומלצו לצוותי הגבלת תוכן ברשתות החברתיות:

1. מעקב מקרוב אחר שילובים של שמות של ארגונים מסוכנים,²³ כמו למשל "חמאס", עם צורות שונות של שבחים. לדוגמה, יש לסמן מילים כמו "بطل" (גיבור) או "شهيد" (שהיד, מת מות קדושים), שאינן מחייבות ניטור בפני עצמן, כאשר הן מופיעות בהקשר של ארגון ספציפי, מתקפת טרור או אדם ספציפי שקשור לטרור.²⁴

2. מעקב אחר האשטאגים והפניות ישירות לארגוני טרור או לפיגועי טרור. כמו ניטור ההאשטג #طوفان_الاقصى (מבול אל-אקצא) יחד עם "يهود" (יהודים) או עם הפניות ישירות כגון #القسام (אל-קסאם).²⁵ ניטור ומיפוי שילובים כאלה בזמן אמת, כאשר הם נעשים פופולריים עם נראות רבה, יסייע לפלטפורמות מדיה חברתית לתת מענה לפערים בהגבלת התוכן בצורה מהירה ואפקטיבית ככל האפשר.

פער אכיפה ספציפי 4: הכחשת אירוע אלים כסוג של מידע אנטישמי כוזב

מידע כוזב (מיס-אינפורמציה) ומידע מסולף (דיס-אינפורמציה) מילאו תפקיד מרכזי בזריעת כאוס, יצרו אווירה של חוסר אמון ושנאה ועיצבו תפיסות גלובליות תוך כדי ההתרחשויות ב-7 באוקטובר. CyberWell מתמקדת כיום בניטור ובניתוח סוג ספציפי של מידע כוזב הקשור ל-7 באוקטובר – הכחשה שהטבח ומעשי האלימות שביצע חמאס, אשר תועדו בהרחבה על ידי המחבלים עצמם, בכלל התרחשו.²⁶ ההכחשה לא רק מספקת הצדקה פוליטית להתקפות נוספות של חמאס אלא גם משמשת ביטוי ברור לאנטישמיות. יש לה הקבלה ברורה להתפתחות שיח הכחשת השואה, כולל ניסיונות להצדיק את הטבח.

נוסף על נרטיב ההכחשה שמחבלי חמאס תקפו באלימות אזרחים ישראלים, המופיע באנגלית ובערבית, זוהו שלושה תת-נרטיבים מרכזיים של הכחשה ועיוות: (1) הכחשה שהתרחשו מעשי תקיפה מינית ו/או אונס; (2) טענה שהמעשים האלימים אומנם אירעו, אך ישראל היא שעומדת

²³ כל פלטפורמה מגדירה לעצמה את רשימת הארגונים והאנשים הנחשבים למסוכנים על סמך הפניות חיצוניות, בדרך כלל ממשרד החוץ האמריקאי או ממקורות אחרים. לדוגמה ראו: Meta's policy on dangerous individuals and organizations <https://tinyurl.com/unksr8wm>

²⁴ סביר מאוד להניח כי כאשר המונח "شهيد" משמש להתייחסות לטרוריסט שמת, הוא משמש כדי להלל ולהאדיר את מעשיו האלימים. CyberWell מנטרת את השימוש במונח בפרט כאשר הוא מפאר את מותם של אנשים שביצעו מעשים אלימים. מסורות מוסלמיות רבות מתארות את הגמול הממתין ל"שהיד", כולל עלייה ישירה לגן העדן, מחילה על כל החטאים וכתר תהילה שיונח על ראשו. מוקדם יותר השנה הגישה CyberWell למועצת הפיקוח העצמאית של Meta <https://tinyurl.com/3xwefy4v> חוות דעת מייעצת למדיניות בנושא זה.

²⁵ מתייחס לגדודי אל קסאם, הזרוע הצבאית של חמאס.

²⁶ CyberWell, "Denial of the October 7 Massacre on Social Media Platforms," Antisemitism Trend Alert. <https://tinyurl.com/3xwefy4v>

מאחוריהם; 3) טענה שישראל והיהודים משקרים באשר למעשי האלימות ומרוויחים מסיפור הטבח.

חשוב לשים לב לכך שלכל פלטפורמות המדיה החברתית הגדולות יש מדיניות מפורשת נגד הכחשת שואה, ולרובן יש גם מדיניות נגד הכחשת העובדה שנעשו מעשים אלימים המתועדים היטב, או באשר להיקף המעשים בכלל. זאת ועוד, לפלטפורמות הללו יש כללי מדיניות שונים שנועדו לטפל במידע מסולף ובמידע כוזב, אם כי כללים אלה משתנים בין הפלטפורמות.

מערך הנתונים שלנו בנושא הכחשת ה-7 באוקטובר כולל קצת יותר מ-300 פוסטים שנאספו מחמש פלטפורמות וצברו יחד יותר מ-25 מיליון צפיות. לאחר הדיווח לפלטפורמות הוסר רק 6% מהתוכן, נכון למאוס 2024. שיעור ההסרה הנמוך חושף פערים גדולים באכיפת המדיניות של הפלטפורמות, או כישלון של הפלטפורמות בהכללת טבח ה-7 באוקטובר כ"מעשה אלימות מוכר". פלטפורמות המדיה החברתית חייבות להכיר בתופעה של הכחשת המעשים שנעשו ב-7 באוקטובר כתוכן המפר מדיניות, ולוודא את הסרתו הגורפת.²⁷

²⁷ מאז כתיבת מאמר זה טיקטוק הכירה בפומבי בהכחשת תקיפה מינית ב-7 באוקטובר כתוכן אסור. שיעורי ההסרה עלו בהתאם – הוכחה לכך שמדיניות ברורה מובילה לפעולה קונקרטית.

סיכום ומסקנות

ההתרחשויות ב-7 באוקטובר הביאו למספר אבחנות מרכזיות. ראשית, מספר הפוסטים באנגלית ובערבית שסומנו כאנטישמיים בסבירות גבוהה הרקיע שחקים, והדבר ממחיש עלייה כללית ברטוריקה האנטישמית (ראו תרשים 1). שנית, בפוסטים בערבית שאומתו כאנטישמים חל שינוי בולט – מנרטיבים המנציחים סטראוטיפים כלליים על יהודים לעבר שיח אלים יותר נגד יהודים (ראו תרשים 2 ותמונות 1-3). שלישית, חלה עלייה גדולה בפוסטים בערבית המאשימים את היהודים כקולקטיב בפעולות של מדינת ישראל וצה"ל (ראו תרשים 2 ותמונות 1-3). אף שהשינוי היה צפוי, הוא מטיל אחריות קולקטיבית על היהודים כקבוצה ומעמיד בסיכון את הקהילות היהודיות ברחבי העולם.

למרבה הצער, פלטפורמות המדיה החברתית לא היו ערוכות לעלייה החדה הזו. בימים ובשבועות שאחרי ה-7 באוקטובר הוצפו צוותי בדיקת התוכן בשטף של תוכן אלים, גרפי, תומך טרור, אנטישמי וכוזב. הפלטפורמות הללו לא השקיעו את המשאבים המתאימים ולא פיתחו תשתית מספקת לניהול משברים, והדבר בלט לעיני כול, למרות שמדובר בסוגיה שהפכה ברורה יותר ויותר במהלך אירועים קודמים, כמו הבחירות בארצות הברית או התפשטות מידע מסולף בעניין הקורונה. המקרה הנוכחי הוא בבחינת "חטיפה" של פלטפורמות מדיה חברתית על ידי ארגון טרור ותומכיו. לפיכך, בהיעדר שימוש עקבי בהון אנושי ובטכנולוגיה לניטור, לסימון ולהסרה של תוכן מזיק, פלטפורמות גדולות של מדיה חברתית כשלו במתן מענה יעיל להתפשטות ההמונית של רעל אנטישמי, שכלל האשטאגים, מונחי חיפוש, תמונות, סרטונים, נרטיבים וכן ארגונים קיימים שהורשו להפיץ את הרעל, ובכך לטשטש את האמת ולשטוף במהירות את הרשת בשנאת יהודים.

לנורמליזציה של האנטישמיות בפלטפורמות המדיה החברתית של הזרם המרכזי יש השלכות חמורות בעולם האמיתי. למשל, תאוריות קונספירציה כמו QAnon הצליחו לגייס אליהן המונים, בתהליך שהגיע לשיאו בהסתערות על בניין הקפיטול האמריקאי ב-6 בינואר 2021. בשבועות ובחודשים שאחרי ה-7 באוקטובר הוכיחה החגיגה הגלויה של אלימות נגד יהודים וישראלים כיצד תשתית לקויה והשקעה בלתי מספקת בהגבלת תוכן הן סכנה לביטחון הלאומי. חיזוק לכך היו אירועים כמו השתלטות אלימות על מבני ממשלה, ביניהם [נבעת הקפיטול](#), על מרכזי תחבורה גדולים, וכן מאהלי המחאה שהוקמו באופן לא חוקי [בקמפוסים של מכללות בארצות הברית](#). משתמשים שהוסתו נגד ישראל ברשת המשיכו הלאה להדהד קריאות לאלימות נגד יהודים תוך היותם במרחב הזהוד תקשורתי (echo chamber) שאותו ליבו האלגוריתמים של הרשתות החברתיות, המעודדים תוכן משלהב וכך יוצרים לולאת משוב המנציחה את עצמה.

פלטפורמות המדיה לא הצליחו לפעול, והדבר מעלה סימני שאלה לגבי האפקטיביות של האסטרטגיה שלהן בהתמודדות עם דברי שטנה וביטויי אנטישמיות, בייחוד אחרי שחמאס הוכיח כי הוא יודע לנצל את הפלטפורמות הללו ככלי ללוחמת טרור פסיכולוגית. חוסר היעילות של השיטות להגבלת תוכן ניכר בעיקר בפערים הגדולים באכיפה עקבית של מדיניות, בייחוד בשפה הערבית ובחודשים שאחרי ה-7 באוקטובר. אינדיקציה כמותית ברורה לכישלון האכיפה היא הפער המתמשך בשיעורי ההסרה של תוכן המפר מדיניות בערבית, לעומת תוכן באנגלית. כדי לטפל בבעיה זו יש להקצות משאבים נוספים באופן נרחב, שיחזקו את צוותי הגבלת התוכן דוברי הערבית.

המלצות למדיניות

ראשית, יש צורך דחוף בפיקוח ממשלתי מוגבר על פלטפורמות מדיה חברתית. בעקבות אירועי ה-7 באוקטובר, על קובעי המדיניות לשקול אסדרה של פלטפורמות מדיה חברתית וחברות הביג טק, בדומה למגזרים ציבוריים אחרים הנמצאים בפיקוח הדוק, כגון בנקאות, תעשיית התרופות או ניהול פסולת. אחת ההמלצות היא לחייב פלטפורמות מדיה חברתית לעמוד בסף מוכח של ניתוח אוטומטי של תמונות וקטעי שמע שמאדירים אלימות. מהלך כזה יחייב חברות גדולות של מדיה חברתית וחברות ביג טק להשקיע כמות מינימלית של משאבים במערכות אוטומטיות שנועדו למנוע הפצת תוכני שטנה ואלימות.

צעד נוסף שארגונים וגורמים ממשלתיים בכל הרמות יכולים לעשות הוא לאמץ את הגדרת העבודה של IHRA לאנטישמיות. ההגדרה היא תוצר של יותר מעשור של מאמצי שיתוף פעולה מקצועיים של פרקליטים לענייני מדיניות והחוק הבינלאומי, היסטוריונים, אקדמאים ומומחים לשיח שנאה. ההגדרה אומצה על ידי יותר מ-40 מדינות ויותר מ-1,000 ארגונים ומוסדות, ביניהם בתי מחוקקים, עיריות, מוסדות חינוך וליגות ספורט. ממשלות וגופים שיאמצו את ההגדרה המפורטת והמדויקת יוכלו להציע כלי מנחה למאבק בשנאת יהודים, שירחיב את ההשפעה מעבר למדיניות – לתחומים כמו מחקר ואקדמיה ולתחומים מפציעים של טכנולוגיה חדשה.

שנית, בהמשך לצורך בשיפור האסטרטגיות להגבלת תוכן, ובפרט עקב מתקפת ה-7 באוקטובר, על פלטפורמות מדיה חברתית לנקוט מהלכים ספציפיים שישפרו את בטיחות המשתמשים. בעקבות זאת העבירה CyberWell לפלטפורמות המדיה החברתית את ההמלצות הבאות והתמקדה בייחוד בתוכן בערבית:

- סגירת פער האכיפה בין פוסטים באנגלית לפוסטים בשפות שאינן אנגלית, ובפרט בערבית, על ידי גיוס בודקי תוכן מיומנים נוספים.
- הטמעת גישה איתנה יותר לסימון אוטומטי של דברי שטנה בתוכן של תמונות, סרטונים ושמע.

- בחינה מחודשת של המדיניות הנוגעת לאנטישמיות, תוך התייחסות למקורות דתיים בשפה הערבית.

בסופו של דבר, התפקיד שממלאות פלטפורמות המדיה החברתית בעיצוב נרטיבים מזיקים ושקריים, וכן בהפצת תוכן גרפי ואלים שנועד לעורר פחד בקרב יהודים וישראלים, צריך להדאיג את כולם. תקריות אנטישמיות בארצות הברית כבר רשמו שיאים היסטוריים בשנת 2023 לפני ה-7 באוקטובר,²⁸ ומאז התרחבו ברחבי העולם. הפלטפורמות כשלו באחריותן המוסרית להגן על המשתמשים בגלל תשתית לקויה ועיכוב בלתי נסלח בזמני תגובה. הגיעה העת שהן יעשו מעשה ושממשלות וגופים רגולטוריים יטילו עליהן אחריות משפטית, שתחייב אותן לסגור את הפער ולמנוע חטיפה עתידית של הרשתות החברתיות שלנו על ידי ארגוני טרור.

* מאמר זה נכתב כחלק מפרויקט משותף של המכון למחקרי ביטחון לאומי (INSS) ומשרד החוץ בנושא 'תפיסות ביחס ליהודים ולישראל במרחב הערבי-מוסלמי והשפעתן על המערב'. לפרסומים נוספים בנושא ראו: <https://tinyurl.com/4nh569w3>

²⁸ ADL, "Audit of Antisemitic Incidents 2023," April 16, 2024. <https://tinyurl.com/ymxvp855>