

ט"ז באלול התשפ"ב
12 בספטמבר, 2022

לכבוד
עדי סופר-תאני
מנכ"לית
פייסבוק ישראל

שלום רב,

הנדון: מנגנונים להתמודדות עם כשלי אכיפה של כללי הקהילה בפלטפורמות פייסבוק ואינסטגרם בישראל

אין חולק כי בעשורים האחרונים, פלטפורמות תוכן דיגיטליות, ובראשן הרשתות החברתיות, הפכו לזירה מרכזית בה מתעצב המרקם החברתי והכלכלי בעולם כולו, ובישראל בפרט. בצד תרומתן של רשתות חברתיות לבניית קהילות והעשרת השיח הציבורי, גם אין חולק על האחריות של הפלטפורמות בשמירה על בטיחות המשתמשים והגנה על הציבור מפני שימוש לרעה בשירותיהן.

כללי הקהילה, אותן קובעות הרשתות החברתיות, הם אחת הדרכים המרכזיות שלהן להגן על המשתמשים בהן, בין היתר באמצעות קביעה ואכיפה של כללים האוסרים על תכנים וחשבונות מסוגים שונים. למשל, פייסבוק קובעת איסורים על תוכן המעודד התאבדות או פגיעה עצמית, על יצירת חשבונות מזויפים או בלתי-מזוהים ומגבילה את התפוצה של תכנים מסוג חדשות כזב אשר עלולים לפגוע בשלומם של המשתמשים. אכיפה של כללי הקהילה האמורים מקבלת משנה חשיבות בתקופות רגישות כגון מערכת בחירות לאומית, שבהן קיימת סבירות גבוהה להפצה ופרסום של תכנים מסוכנים או בעייתיים.

אנו פונים אלייך, כנציגה הבכירה בישראל של קבוצת מטא, על רקע ממצאים ופרסומים מדאיגים המלמדים על כשלים מהותיים באכיפה עצמית של כללי הקהילה של קבוצת מטא במדינות ובשפות לא-מערביות, ולענייננו – בישראל, בנוגע לתוכן בשפות העברית, הערבית והרוסית.

א. הצורך באכיפת כללי הקהילה בשפה שאינה אנגלית

רק לאחרונה קבע ה-Oversight Board - גוף הביקורת העצמאי של חברת Meta – כי ההשקעה בניהול ובפיקוח על תוכן בשפות שאינן אנגלית אינה מספקת, והצביע על המספר הקטן של פניות שהגיעו אליו ממשתמשים מהודו וממדינות אחרות שאינן דוברות אנגלית.¹ לדאגות אלו מצטרפים גם נתונים שהתפרסמו בסוף שנת 2021, לפיהם 87% מהתקציב העולמי של פייסבוק לצורך התמודדות עם חדשות כוזבות (fake news) מושקע בתוכן שמפורסם בארה"ב.² לכך מצטרפים מסמכים שהציגה עובדת פייסבוק לשעבר, פרנסס האוגן, לממשל האמריקאי החושפים כי מערכות הבינה המלאכותית של פייסבוק אינן מסוגלות לזהות תוכן פוגעני במספר רב של שפות

¹ Oversight Board First Annual Report, page 19, June 2022 ([link](#)).

² Nick Popley, "The 5 Most Important Revelations From the 'Facebook Papers'", 26.10.2021 ([link](#)).

אשר משמשות משתמשים בפלטפורמה ברחבי העולם. כך, למשל לפייסבוק לא היו בשנת 2020 אלגוריתמי סינון המשמשים לצורך מציאת מידע שגוי בשפות של מיאנמר או אתיופיה.³ גם המחקר האקדמי זיהה את הסכנות המוגברות לבטיחות המשתמשים במדינות בהן השפה העיקרית אינה אנגלית. אוולין דואק, חוקרת בכירה במכון "Knight First Amendment" מסבירה שהסכנה המוגברת לבטיחות משתמשים במדינות בהן השפה העיקרית אינה אנגלית היא הזמן שנדרש לתת לבינה המלאכותית ללמוד את השפה ואת הזמן הדרוש לצורך תרגום תנאי הקהילה והתאמת מנגנוני אכיפת התוכן, לפני שפייסבוק נכנסת לשוק במדינה זרה.⁴

בנוסף, ממצאים של ארגון Global Witness מהחודש האחרון מלמדים כי מטא עדיין נכשלת באופן משמעותי בהתמודדות עם תוכן המוגדר כפוגע בטוהר הבחירות בשפות שאינן אנגלית.⁵ מחקר ניסויי שערך הארגון הראה כי חברת מטא אישרה 10 מתוך 10 פרסומות שפורסמו בפורטוגזית על רקע הבחירות המתקרבות שם, על אף שהכילו תוכן מטעה באופן ברור הנוגע לבחירות אשר עלול היה להשפיע עליהן בצורה משמעותית. בנוסף, כחלק מהניסוי, המודעות פורסמו ממקום מחוץ לברזיל, ואף לא סומנו כמודעות או תוכן ממומן כפי שדורשת מטא.

המגבלות של הרשת החברתית בזיהוי והסרה יעילים של תוכן וחשבונות המפרים את כללי הקהילה שלה נוכחות במלוא עצמתן גם בזירה הישראלית ובשפות העברית, הערבית והרוסית ומחייבות מענה מיידי על ידי קבוצת מטא.

ב. הצורך בחיזוק האכיפה של Meta את כללי הקהילה שלה בישראל

בדיקה נרחבת שנערכה בישראל בחודשים האחרונים על ידי ארגון FakeReporter, המבצע מחקר למאבק בפעילות זדונית ברשת, ממחישה ומדגימה היטב כיצד הכשלים הסיסטמיים שזוהו לגבי אכיפת כללי הקהילה ביחס לשפות ומדינות לא מערביות באים לידי ביטוי ביתר שאת גם בשפה העברית.⁶ הממצאים המטרידים שעולים מהבדיקה של FakeReporter מצביעים על כשלים רבים במנגנוני האכיפה והדיווח הדורשים תיקון דחוף:

קבוצה אחת של דוגמאות כוללת כשלי אכיפה של כללי הקהילה בנושא אמינות ואותנטיות שמות המשתמשים ברשת החברתית (**Account Integrity and Authentic Use**) בשפה העברית. בכלל זה, כללי המדיניות של פייסבוק בנושא אותנטיות החשבון אוסרים על שימוש בסמלים, מספרים, רצף של תווים זהים, שילוב של מספר

³ שם.

Rishy Iyengar, "Facebook has language blind spots around the world that allow hate speech to flourish", CNN, ⁴ 26.10.2021 ([link](#)).

<https://www.globalwitness.org/en/campaigns/digital-threats/facebook-fails-tackle-election-disinformation-ads-ahead-tense-brazilian-election/>⁵

⁶ המחקר בדק את הנושא בשפה העברית בלבד. אנו יוצאים מנקודת הנחה שאין שינוי מהותי בשפות אחרות הנהוגות בישראל, ובפרט ערבית ורוסית.

שפות ואפילו איסור על שילוב שמות או ביטויים של מקומות.⁷ למרות קיומם של קריטריונים מפורשים אלה, תחקיר ארגון FakeReporter בשנה האחרונה תיעד מאות חשבונות בשפה העברית המפירים באופן מובהק את הכללים בנושא אמינות ואותנטיות, אשר לא מוסרים על ידי הרשת החברתית גם לאחר דיווח אקטיבי של משתמשים.⁸

כמובן, לא סביר לצפות שמנגנוני האכיפה של פייסבוק לגבי כללים אלו יזוהו ויטופלו באופן מיידי בכל חשבון המפר את כללי הקהילה. עם זאת, ברור כי הפרות אלו ניתנות לאכיפה אוטומטית פשוטה יחסית, למשל סריקות אוטומטיות תדירות של שמות המשתמשים תוך איתור שילוב שפות, רצף תווים או הפרות נפוצות כגון שימוש בשם "אנונימי" בשפות השונות. כדי להמחיש את היקף ההפרה של הכללים בעניין זה בשפה העברית, מספיק להביט בתוצאות החיפוש של פייסבוק עבור משתמשים ששםם הוא "שם בדוי" ([קישור](#)), כהפרה ברורה ונפוצה של כללי הקהילה, שניתן לזהות בקלות ובאופן אוטומטי.

דוגמה נוספת היא כשלי אכיפה בנוגע לתוכן בעברית שמפר באופן מובהק את כללי הקהילה הברורים נגד תכנים בנושא סמים או שימוש בפלטפורמה לקידום פעילות בלתי-חוקית. כללי הקהילה של Meta אוסרים באופן מפורש על תוכן אשר עוסק בסחר או קידום של סמים מסוכנים ועל פעילות או תוכן שנועדו לתיאום או קידום פשיעה (כגון הפצה של מטבע מזויף או שירותי זנות). יודגש כי איסור תוכן ספציפי זה הוא פשוט יחסית לזיהוי ביחס לאיסורי תוכן מופשטים או סובייקטיביים יותר כגון "דברי שנאה" או "הסתה".

למרות האיסור המפורש על תכנים מהסוג האמור, ארגון FakeReporter תיעד מקרים רבים של חשבונות ופוסטים בשפה העברית העוסקים באופן ברור ומפורש בסחר של סמים מסוכנים המפרטים בעברית ובאופן מפורש את שלל החומרים אותם הם מציעים למכירה.⁹ כפי שתיעד הארגון, לא רק שפוסטים אלו לא מאותרים אוטומטית על ידי הרשת החברתית באמצעות חיפוש פשוט של שמות הסמים בעברית, גם דיווחים יזומים על התכנים נדחים על ידי הרשת החברתית (באופן אוטומטי, ככל הנראה) בקביעה שהתוכן אינו מפר את כללי הקהילה, למרות שלא יכולה להיות מחלוקת, עבור מי שמבין עברית, במה עוסקים התכנים שדווחו.

כשלי האכיפה הללו מעוררים חשש מיידי לבטיחות המשתמשים, בפרט על רקע ההכנות והקמפיינים לקראת סבב הבחירות הקרוב בישראל שיתקיים בחודש נובמבר הקרוב. אנו פונים אליכם בעת הזאת כדי להתריע מבעוד מועד על הכשלים שפורטו לעיל ולדרוש שיקבלו מענה הולם בהקדם האפשרי.

⁷ <https://www.facebook.com/help/112146705538576>

⁸ [בקישור זה](#) ניתן להתרשם ממספר דוגמאות מייצגות לחשבונות בשפה העברית המפירים באופן חד-משמעי את כללי פייסבוק נגד שמות לא אותנטיים.

⁹ [בקישור זה](#) ניתן להתרשם ממספר רב של דוגמאות לתכנים בעברית המפירים את כללי הקהילה בנושא סמים ופעילות לא חוקית אשר פועלים תקופות ארוכות בפייסבוק ובאינסטגרם.

זאת, על רקע הצהרות חברת מטא כי היא מפעילה מרכז פעולה המיועד לתת מענה הולם לשמירה על תקינות תהליך הבחירות שיתקיימו במספר מדינות נבחרות (בהן ארה"ב, הודו, ברזיל, מדינות אירופה ואחרות).¹⁰ רק לאחרונה, בחודש אוגוסט 2022, הודיעה חברת מטא כי תנקוט מספר צעדים על מנת לעזור לשמור על טוהר הבחירות המתקרבות בברזיל, לרבות: השקעה בטכנולוגיה ובצוותים אשר יגנו על הבחירות; יצירת שיתוף פעולה עם הרשויות ובודקי עובדות עצמאיים לצורך אספקת מידע אמין בנושא הבחירות ובכדי להילחם במיסאינפורמציה; והקמת מרכז פעולה לטיפול בבחירות אשר יזהה סכנות בזמן אמת ויקצר את זמני הטיפול.¹¹ החברה מדגישה כי מעבר לדאגה כי הפלטפורמות שלה לא יובילו לפגיעה בבחירות, היא מתחייבת גם לספק שקיפות בנוגע הבחירות ולעודד אנשים להצביע.¹²

אנו כמובן מעריכים צעדים אלו שנוקטת חברת מטא, אך על רקע הערכות החברה לקראת אירועים דומים במדינות מסוימות ובשפות מסוימות, אנו מבקשים לפעול על מנת להגן בצורה זהה ושוויונית על כלל המשתמשים, ללא קשר למקום מוצאם או שפתם.

כפי שקבוצת מטא עצמה חוזרת ומציינת, כללי הקהילה בנוגע לתוכן או משתמשים אסורים נועדו בראש ובראשונה להבטיח את בטיחותם ושלומם של המשתמשים מפני ניצול לרעה של הפלטפורמה על ידי משתמשים אחרים. הדוגמאות הנקודתיות שהובאו לעיל ביחס לשפה העברית, מצטרפות כאמור לכלול העדויות המצטרבות לכשלים הייחודיים של הרשת החברתית באכיפת כללי הקהילה שלה כנגד תוכן ומשתמשים מפריס. אם לא נבטיח שלכללי הקהילה ולתנאי השימוש של הפלטפורמות יש משמעות בפועל גם בשפות העברית, הערבית והרוסית בישראל – אלו יהיו חסרי משמעות.

על רקע זה, אנו קוראים לכם לפתוח בדיאלוג מיידי עם רשויות המדינה וגופי החברה האזרחית הרלוונטיים, על מנת לגבש מענה אפקטיבי ומיידי שיבטיח כי המשתמשים בישראל יזכו למרחב מקוון בטוח. כאמור, **צורך זה הופך דחוף וקריטי אף יותר עם פתיחת הקמפיינים לקראת הבחירות לכנסת והצורך להבטיח אכיפה אפקטיבית של כללי הקהילה הרלוונטיים למניעת חשבונות מזוייפים או תכני הסתה.**

תחילת הדיאלוג והשותפות במטרה לוודא אכיפה אפקטיבית של כללי ההתנהגות והבטיחות בפלטפורמה עבור תוכן ומשתמשים בשפה העברית צריכה להיעשות באמצעות שיתוף במידע חיוני בהבנה ולמידה של כשלי האכיפה הספציפיים ביחס לתוכן ומשתמשים בעברית או בישראל. למשל, כפי שפייסבוק מפרסמת נתונים מדינתיים על היקף ויעילות פעולות אכיפת התוכן שלה במדינות כגון הודו או גרמניה, אנו קוראים לפייסבוק לפרסם נתונים לגבי ניהול התוכן (הסרה, הפחתת חשיפה, השעיית חשבונות, הפסקת מוניטציה וכו') שביצעה בשנה החולפת ביחס למשתמשי הפלטפורמה שמקורם בישראל או בשפה העברית, הערבית והרוסית בישראל;

¹⁰ <https://www.facebook.com/gpa/resources/election-integrity>

¹¹ <https://about.fb.com/news/2022/08/how-meta-is-preparing-for-brazils-2022-elections/>

¹² <https://about.facebook.com/actions/preparing-for-elections-on-facebook/>

יחד עם היקף ותוצאות הליכי השגה עליהם. בכלל זה, אנו קוראים לקבוצת מטא לשתף אותנו במידע על האופן שבו נבדקים דיווחים ביחס לתוכן או חשבון בשפה העברית ובשפה הערבית, ומאפיינים של הגורמים הטכנולוגיים או האנושיים המטפלים בדיווחי המשתמשים ומכריעים האם לנקוט בפעולת אכיפה.

אנו מקווים כי חברת Meta ופייסבוק ישראל כנציגתה בישראל, יגלו את האחריות המתבקשת כלפי הציבור הישראלי ויגבשו צעדים מיידיים להתמודדות עם כשלי אכיפת התוכן ביחס לתוכן וחשבונות בעברית, ערבית ורוסית, במיוחד על רקע הבחירות הקרבות והצורך לאכוף את כללי התוכן של הרשתות בזירה זו.

נשמח לתאם פגישה בהקדם על מנת לדון בדברים.

בברכה,



עו"ד יורם הכהן, מנכ"ל איגוד האינטרנט הישראלי.

ד"ר שרון בר זיו, נשיאת איגוד האינטרנט הישראלי.

העתקים

הוועד המנהל של איגוד האינטרנט הישראלי (ע"ר).

.David Miles, Director of Safety and Policy EMEA, Meta