

י"ג בשבט התשפ"א

26 בינואר 2021

באמצעות דוא"ל

לכבוד
כבוד שופט בית המשפט העליון, עוזי פוגלמן
יו"ר ועדת הבחירות המרכזית לכנסת ה-24

שלום רב,

הנדון: התמודדות עם השימוש בטכנולוגיית זיוף עמוק (Deep Fake)
במטרה להטעות את ציבור הבוחרים או לשבש את הליך הבחירות לכנסת ה-24

1. מבוא ועיקרי ההמלצות

לאור השימוש שנעשה לאחרונה בטכנולוגיית הזיוף העמוק (deep fake), בישראל ובמדינות דמוקרטיות נוספות, אנו מבקשים להציג בפני ועדת הבחירות המרכזית לכנסת ה-24 מידע וכלים להתמודדות עם הסכנות הדמוקרטיות הטמונות בניצולה לרעה.¹ להערכותינו, ולאור ניסיון העבר ממערכות בחירות אחרות, קיימת אפשרות ממשית כי בבחירות הקרובות ייעשה שימוש בטכנולוגיה זו על מנת להטעות באופן בלתי הוגן את קהל הבוחרים או לשבש את מהלכם התקין של הבחירות.

כמפורט בהרחבה בהמשך התייחסות זו, האתגר המשפטי והיישומי של התמודדות עם סכנותיו הדמוקרטיות של מידע כוזב וזיוף עמוק בפרט נובע מקצב ההתפשטות הגבוה, הנרחב והמגוון שלו ("ויראליות") ויכולת מוגבלות לזהותו ולהגבילו באמצעים טכנולוגיים. על כן, מתחייבת תגובה מיידית ואפקטיבית של ועדת הבחירות – הן בתהליכי קבלת ההחלטות והן בתהליכי אכיפתן במרחב המקוון – ככל שאכן מדובר בתוכן שקרי אשר צפוי להטעות את ציבור הבוחרים הסביר או לשבש את הליך התקין של הבחירות.

בשים לב למאפיינים הייחודיים של מידע כוזב וזיוף עמוק – והאופן בו אלו עשויים להתבטא בהפרעה ממשית לבחירות או לאמון הציבור במוסדות שלטוניים – מסמך זה מבקש להציע לוועדת הבחירות המרכזית "ארגון כלים" מוצע להתמודדות עם האתגרים הדמוקרטיים והמשפטיים שמעוררת טכנולוגיית הזיוף העמוק:

(א) **הסברה והעלאת מודעות בקרב הציבור:** הבהרה לציבור כי יצירה או הפצה של תכנים באמצעות זיוף עמוק (ללא סימון או הבהרה ברורים) עלול להוות במצבים מסוימים עבירה פלילית לפי סעיפים 159 ו-197 לחוק העונשין (פרסום ידיעות כוזבות הגורמות פחד ובהלה או המפריעות את שלום הציבור, הפרעה לציבור), במטרה לממש את ההרתעה אותה מבקש להשיג הדין הפלילי במישורים אלו;

¹ פניה זו הינה תוצר משותף של איגוד האינטרנט הישראלי (ע"ר) והקליניקה למשפט, טכנולוגיה וסייבר בפקולטה למשפטים, אוניברסיטת חיפה ומחבריו הם בעלי מומחיות מחקרית ומעשית בהתמודדות עם ההסדרה המשפטית של המרחב המקוון.

(ב) הנחייה מקדימה לגופי התקשורת הממוסדת הכפופים לדין הישראלי ו/או לרגולציה: הבהרה לגופי התקשורת הממוסדת כי חלה עליהם החובה לסמן באופן ברור הצגה של תוכן מסוג זיוף עמוק, לכל הפחות בתקופת 60 הימים שלפני הבחירות;

(ג) יצירה מוקדמת של ממשקים מול הרשתות החברתיות הדומיננטיות (פייסבוק, גוגל/יוטיוב וטוויטר), במטרה להבטיח את האכיפה המיידית של החלטות ועדת הבחירות המרכזית, ככל שזו תמצא כי יש לטפל בתוכן שקרי המופץ במרחב המקוון. זאת, בשים לב לכך שאכיפה מאוחרת מאיינת את היכולת למנוע תפוצת התוכן השקרי (ובהתאמה, מאיינת את האפשרות למנוע או לצמצם את נזקיו הדמוקרטיים של תוכן שקרי) ושממילא פלטפורמות אלו אוסרות על משתמשיהן להפיץ תכנים מסוג זיוף עמוק, כמפורט בהמשך.

(ד) היכרות ועדת הבחירות עם מומחים וכלים טכנולוגיים-פורנזיים לניתוח תכני מדיה החשודים כעובדו או שונו באמצעות טכנולוגיית הזיוף העמוק, בהם תוכל ועדת הבחירות להסתייע כאשר נדרשת הכרעה עובדתית בתנאי אי-ודאות לגבי אמיתות התכנים העשויים בוודאות קרובה להשפיע על הצבעת הציבור או לפגוע במהלכן התקין של הבחירות.

(ה) התאמה הקשרית של מבחן הוודאות הקרובה, כמבחן הסתברותי, לפוטנציאל התפוצה והאפקט השכנועי הגבוהים של מידע כוזב מסוג זיוף עמוק, תוך בחינת אפשרות הוצאת הנחיות מיוחדות לתקופה של 48 השעות לפני הבחירות.

להלן נבחן את מאפייניה של טכנולוגיית הזיוף העמוק כאיום ממשי על טוהר הבחירות והסדר הדמוקרטי, נסביר כיצד היא מאתגרת את המשפט הקיים, ונציע ארגז כלים להתמודדות אפקטיבית של ועדת הבחירות עם יכולתה חסרת-התקדים של טכנולוגיית הזיוף העמוק להטעות את ציבור הבוחרים או לשבש את ההליך התקין של הבחירות.

2. דיסאינפורמציה וטכנולוגיית הזיוף העמוק כאיום ממשי על טוהר הבחירות והסדר הדמוקרטי

טכנולוגיית זיוף עמוק מהווה קפיצת מדרגה בתחום הדיסאינפורמציה וחדשות הכוזב (fake news), ומבוססת על שימוש בבינה מלאכותית על מנת לייצר תוכן אודיו-ויזואלי (סרטון וידאו, קובץ שמע או תמונות), הנחזה להיות כתיעוד אותנטי של אירוע או מציאות, באופן אשר מקשה על הנחשפים לזהות כי למעשה מדובר בזיוף. בבסיס הטכנולוגיה עומד שימוש בכלים של למידה עמוקה (deep learning) ורשתות נוירוניות (neural networks) המוזנים במידע אותנטי (למשל תמונות ו/או סרטונים של האדם אותו רוצים לזייף), שעל בסיסו ייוצר המידע המזויף. תהליך יצירת הזיוף העמוק כרוך בעיבוד מידע המקור ומיצוי מאפיינים ביומטריים של היעד לזיוף ממנו, ולאחר מכן בניית התוכן המזויף על בסיס אותם מאפיינים. בשל כך, יעדים פוטנציאליים לזיוף עמוק ברמה גבוהה הם דמויות ציבוריות, אשר קיים וזמין תיעוד אודיו-ויזואלי שלהם באיכות דיגיטלית גבוהה, לרוב באינטרנט.

בהחלטת יו"ר ועדת הבחירות המרכזית לכנסת ה-24 מיום 18.1.2021 הוגדרה טכנולוגיית הזיוף העמוק באופן הבא: "טכנולוגיה ליצירת תוכן קולי או חזותי או לשינוי תוכן קיים, כך שהצופה הסביר (ואף הצופה

המתחכם) יסבור כי פלוני ביצע פעולה או העביר מסר, אך התוכן אינו אמיתי. התוכן הוא באיכות גבוהה עד כדי כך שמשתמש מן היישוב יתקשה לרוב לגלות שמדובר בזיוף.²

לאחר שהכירה וועדת הבחירות בכך שתוכן מסוג זיוף עמוק מהווה תופעה ייחודית בה נדרש המשפט לטפל, נבקש להזין את ההכרעות המשפטיות העתידיות של וועדת הבחירות בנתונים רלוונטיים אודות פוטנציאל הנזק הגבוה במיוחד של מידע כוזב וזיוף עמוק בפרט.

הספרות המקצועית מראה שמידע כוזב נוטה להתפשט מהר יותר, רחוק יותר ולקהילות מגוונות יותר.³ כמו כן, בשל היותו מושך תשומת לב בחריגותו, יש לסרטונים כוזבים נטייה להפוך לווראליים יותר מסרטונים אמיתיים.⁴ הדבר נכון באופן כללי כשמדברים על חדשות כוזב (Fake News) ובפרט כשמדובר בסרטון שפוברק בטכניקת הזיוף העמוק. בסקר שנערך על ידי "Deeptrace" (Sensity בשמה הקודם, כאמור), עלה כי נכון לאוקטובר 2019 היו יותר מ-14,678 סרטים מקוונים של זיוף עמוק. מדובר בעליה של 84% מחודש דצמבר 2018, כאשר 96% מהסרטים המקוונים שנוצרו בזיוף עמוק היו סרטים בעלי תוכן פורנוגרפי.⁵ המכון למחקרי ביטחון לאומי בישראל (INSS) הצביע אף הוא על הסכנות הטמונות בזיוף העמוק על הביטחון הלאומי בכלל ובפרט בתקופת בחירות,⁶ אשר עשויות לעלות על אלו של חדשות כוזב "רגילות".

הקושי האנושי והטכנולוגי להבחין בין תוכן אודיו-ויזואלי אותנטי (המתעד את המציאות) לבין כזה שעובד או יוצר באמצעות טכנולוגיית זיוף עמוק נעוץ בכך שטכנולוגיית הזיוף העמוק כוללת כאמור טכניקות של למידת מכונה (בינה מלאכותית), המקשה על זיהוי אנושי או ממוחשב של זיוף עמוק. במסגרת טכניקה זו, פועלים שני אלגוריתמים של בינה מלאכותית, המזינים ומתחרים זה את זה: אלגוריתם אחד, מתבסס על מאגר מידע מסוים ויוצר תוצר חדש "מזויף" על בסיסו, אשר נחזה להיות חלק ממאגר המידע. האלגוריתם השני אמור לנסות לזהות האם מדובר בזיוף או לא, על ידי מציאת "פגמים" בזיופים. הרצה של שני האלגוריתמים במקביל בדרך של היזון חוזר מאפשרת לאלגוריתם המזויף להשתפר במהירות רבה, עד שהאלגוריתם לא יוכל לזהות אם מדובר בזיוף,⁷ כך שקיים קושי ממשי לזהות זיוף עמוק, קושי שייגבר עם השנים.

יתרה מכך, כוחו השכנועי והווראלי של תוכן מסוג זיוף עמוק, להבדיל מחדשות כוזב, נובע מכוח שהראשון אינו מהווה דיווח על אירוע או ביטויים אלא נחזה להיות "תיעוד ישיר" של התרחשותם במציאות, בבחינת

² תב"כ 9/24 יש עתיד נ' עמותת כן לשלום, פסקה 4 (18.1.2021).

³ Soroush Cosoughi, Deb Roy and Sinan Aral, *The Spread of True and False News Online*, 359(6380) SCIENCE 1146 (2018) ראו גם Bobby Chesney and Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CAL. L. REV. 1753, 1755-1757 (2019).

⁴ NAHON KARINE AND JEFF HEMSLEY, GOING VIRAL (2013), Cambridge: Polity Press.

⁵ Henry Ajder, Giorgio Patrini, Francesco Cavalli & Laurence Cullen, *The State of Deepfakes: Landscape, Threats, and Impact*, Deeptrace (Sep. 2019) https://regmedia.co.uk/2019/10/08/deepfake_report.pdf קרין נהון, ורדית אבידן ורן גלעד "זיוף עמוק (DeepFake) בהקשר של ניצול מיני: דרושה פרשנות תכליתית של חוק הסרטונים (הוגש לשיפוט, עותק מצוי אצל הפונים, ניתן לשלוח אם הוועדה הנכבדה תמצא לנכון).

⁶ ענבל אורפז "Deepfake": מקרה בוחר להשפעת פייק ניוז על הביטחון הלאומי INSS (11.9.2020). קישור: <https://www.inss.org.il/he/publication/deepfake>

⁷ Russell Spivak, "Deepfakes": *The Newest Way to Commit One of the Oldest Crimes*, 3 GEO. L. TECH. REV. 339, 341-345 (2019).

טוב מראה עיניים ממשמע אוזניים. כלומר, אנשים נוטים להאמין לדברים שהם רואים בצורה מומחשת וממשית. בהתאמה, פוטנציאל הנזק של תכנים שהופקו באמצעות טכנולוגיית זיוף עמוק אף עולה על זה של חדשות הכזב, באופן המחייב הערכות מוקדמות ובזמן אמת של ועדת הבחירות המרכזית על מנת להבטיח את טוהר ההליך הדמוקרטי.

השימוש בטכנולוגיה זו, לא פסח אף על ישראל. בינתיים, ככל הידוע לנו, נעשה בה שימוש פארודי או שיווקי בלבד,⁸ אך הידע והיכולת לעשות בה שימוש לרעה בהקשרי הבחירות בישראל קיימים.

3. הצורך במענה משפטי ייחודי להתמודדות עם האיום הדמוקרטי של זיוף עמוק

נוכח ההכרה בסכנה הייחודית הגלומה בתכנים מסוג זיוף עמוק ובפרט בנוגע לתהליכים דמוקרטיים והליכי בחירות, רוב פלטפורמות התוכן המקוונות אוסרות על הפצת תכנים מסוג זיוף עמוק והודיעו כי יפעלו להסרתם.⁹ אולם, נוכח חדשנותה של טכנולוגיית הזיוף העמוק ומורכבות האתגרים הדמוקרטיים-משפטיים המתעוררים באשר לאסדרתה, טרם התגבש קונצנזוס באשר לאופן בו נדרשת מדינה דמוקרטית להתמודד עם טכנולוגיה זו, היוצרת איומים חסרי תקדים על ערכים ותהליכים דמוקרטיים.

כיום, קיימים כבר ניצנים ראשוניים למערכי אסדרה משפטיים אשר עוסקים באופן ספציפי בטכנולוגיית הזיוף העמוק וקובעים איסורים ומגבלות על יצירתם או הפצתם (ללא גילוי כי מדובר בזיוף), בשים לב לפוטנציאל הנזק הממשי שלהם עבור חברות דמוקרטיות בסמוך לתקופת הבחירות. בשנת 2019 חוקקה מדינת טקסס שבארצות הברית תיקון לחוק הבחירות אשר קובע איסור פלילי על יצירה או הפצה של תכני וידאו שנוצרו באמצעות טכנולוגיית זיוף, ככל שאלו נועדו לפגוע במועמד או להשפיע על תוצאות הבחירות, במשך שלושים ימים שטרם הבחירות.¹⁰

חקיקה מקיפה יותר התקבלה באותו שנה במדינת קליפורניה, בה נקבעו איסור פליליים על יצירת או הפצה של תוכן מסוג זיוף עמוק בהקשרים מיניים ובהקשרי בחירות,¹¹ כאשר "זיוף עמוק" מוגדר כ: "כל מדיה קולית או חזותית, בפורמט אלקטרוני, אשר נוצר או עובד באופן שיטעה אדם סביר לחשוב שמדובר בתיעוד אותנטי של דיבור או מעשה של אדם, ללא הסכמתו".¹² בכל הנוגע להקשרי בחירות, החוק בקליפורניה אוסר על יצירה או הפצה של תוכן מסוג זיוף עמוק אשר מטרתו להשפיע על הצבעה בעד או נגד מועמד בבחירות, ומאפשר למועמדים שנעשה שימוש בדמותם ליצירת תוכן מסוג זיוף עמוק (שאינו מסומן ככזה) לפתוח בהליכים משפטיים כנגד המעורבים ביצירה או הפצה של התוכן בגין הנזקים שנגרמו להם בעקבות כך; ומסמיך את

⁸ ראו למשל סרטון פארודי בו נראה ראש הממשלה, מר בנימין נתניהו, בסרט של מל ברוקס ריבלין ז"ל, תוך שימוש בטכנולוגיה זו <https://www.youtube.com/watch?v=EXdxWIAqVmw>; וסרטון של האגודה למלחמה בסרטן בכיכובם של ציפי שביט וספי פייסבוק: <https://about.fb.com/news/2020/01/enforcing-against-manipulated-media>; טוויטר: <https://blog.youtube/news-and-events/how-youtube-supports-elections?m=1>; יוטיוב: <https://help.twitter.com/en/rules-and-policies/manipulated-media>

⁹ Texas Senate Bill 751 (2019-2020).

¹⁰ חקיקה כנגד שימוש בטכנולוגיית זיוף עמוק ליצירת תכנים פורנוגרפיים ללא הסכמת מושא התוכן קיימת משנת 2019 גם במדינת וירג'יניה.

¹¹ California Assembly Bill 730, Elections: Deceptive Audio or Visual Media (2019-2020) (התרגום שלנו).

בתי המשפט להעניק צווי מניעה לגופי המדיה השונים, לרבות אתרי אינטרנט, המיועדים לעצור המשך את הפצת התוכן המזויף.

בישראל, כידוע, עדיין אין חקיקה ספציפית המתייחסת לתופעה הזיוף העמוק או להפצה של דיסאינפורמציה באמצעות טכנולוגיות מידע במסגרת חוק הבחירות. עם זאת, הדין הקיים כולל מספר מקורות נורמטיביים המאפשרים לוועדת הבחירות להתמודד עם השימוש לרעה בטכנולוגיית זיוף עמוק והפצת דיסאינפורמציה בניסיון להשפיע על הליך הבחירות ובראשם חוק הבחירות (דרכי תעמולה), התשי"ט-1959 ("חוק הבחירות"), המסמיך את יו"ר ועדת הבחירות להוציא צו מניעה המונע ביצוע מעשה עבירה לפי החוק. אמנם, המגבלות הקבועות בחוק מתייחסות בעיקרן לתעמולת בחירות ברדיו ובטלוויזיה, אולם יו"ר ועדת הבחירות המרכזית לכנסת ה-21, כבוד השופט חנן מלצר, קבע זה מכבר כי על פי תורת הפרשנות התכליתית חובות השקיפות והגילוי הקבועות בחוק חלות גם לגבי פרסומים ברשת האינטרנט,¹³ תוך קביעה מפורשת כי "אין מקום להותיר את רשת האינטרנט במרחב אקס-טריטוריאלי בו המשחק הפוליטי מכשיר כל פרסום והתנהגות".¹⁴

כמו כן, אך לאחרונה קבע כבוד השופט פוגלמן כי הפצת מידע כוזב, זיוף עמוק בפרט, אשר עשוי בוודאות קרובה להטעות את הבוחר הסביר או לשבש את ההליך התקין של הבחירות מפריס את סעיף 13 לחוק הבחירות (וייתכן שאף את סעיף 119 לחוק הבחירות לכנסת [נוסח משולב], התשכ"ט-1969), כך שבסמכותו של יושב ראש ועדת להורות על הסרת התוכן הכוזב או המזויף (לרבות מרשת האינטרנט ופלטפורמות תוכן מקוונות).

4. ערוצי פעולה להתמודדות עם הסכנות הייחודיות של תכני זיוף עמוק בתקופת הבחירות

נוכח מאפייניה הייחודיים של טכנולוגיית הזיוף העמוק וזמינותה עבור גורמים זרים או מקומיים שיבקשו להשפיע באופן בלתי-הוגן על תוצאות הבחירות או על התנהלותם, אנו מציעים לוועדת הבחירות המרכזית "ארגז כלים" ומיפוי מישורי הפעולה המומלצים לשם להתמודדות עם האתגרים הדמוקרטיים והמשפטיים שמעוררת טכנולוגיית הזיוף העמוק:

4.א פניה לציבור ופעילות הסברתית

כפעולת מנע מקדימה, אנו ממליצים כי ועדת הבחירות המרכזית תפרסם הודעה לציבור על כך שיצירה או הפצה של חדשות כזב או תכני זיוף עמוק בסמוך לבחירות בניסיון להשפיע בדרך בלתי הוגנת על דעת הקהל עשויה להוות עבירה פלילית לפי סעיפים 159 ו-197 לחוק העונשין (פרסום ידיעות כוזבות הגורמות פחד ובהלה או המפריעות את שלום הציבור, הפרעה לציבור).

באופן כללי יותר, מומלץ כי ועדת הבחירות המרכזית תנקוט בפעילות הסברתית לציבור אודות תופעת חדשות הכזב בכלל וטכנולוגיית הזיוף העמוק בפרט, כפי שוועדת הבחירות לכנסת ה-22 יצרה עמוד מקוון ובו הסברים על הדרכים להימנע מהשפעתם הבלתי רצויה של חדשות כזב, לרבות סרטוני

¹³ תב"כ 8/21 בן מאיר נ' מפלגת הליכוד (27.2.2019).

¹⁴ שם, בפס' 82.

הסבר על התופעה והדרכים להימנע ולהיזהר ממנה.¹⁵ זו היא יוזמה מבורכת ואני סבורים כי יש להציף אותה מחדש בבחירות הקרובות ולהוסיף עליה נדבך אזהרה מטכנולוגית הזיוף העמוק.

ב.4 פעולות מקדימות: פניה לאמצעי התקשורת הממוסדים

אמצעי התקשורת הממוסדים, כגון העיתונות המודפסת, שידורי הטלוויזיה וגופי התוכן העיתונאי-חדשותי המקוונים ממלאים תפקיד קריטי בתיווך והפצה של מידע לציבור, ובפרט בהקשרים פוליטיים הרלוונטיים להצבעות קהל הבוחרים. להבדיל מפלטפורמות התוכן המקוונות (כגון facebook או YouTube), אמצעי התקשורת הממוסדים מתפקדים כשומרי סף המפיקים את התכנים אותם הם מפיצים ונושאים באחריות המלאה להם. יתרה מכך, גופי התוכן העיתונאי-חדשותי מחויבים לבדוק את נכונות ומהימנות הדיווחים אותם הם מוסרים, בזהירות הראויה לפי נסיבות העניין, כאשר דחיפות הפרסום אינה פוטרת מבדיקת נכונות הידיעה.

על כן, אנו ממליצים כי וועדת הבחירות תפנה לאמצעי התקשורת הממוסדים מבעוד מועד, ותנחה אותם כי בכל פעם בה מבקשים לעשות שימוש בסרטונים מזויפים הקשורים לתחום הפוליטי, במסגרת דיווח חדשותי, יש לסמן באופן ברור, לאורך כל הסרטון או קטע השמע, כי מדובר בזיוף, לכל הפחות בתקופת 60 הימים לפני הבחירות. המלצה זו עולה אף בקנה אחד עם החלטת ועדת הבחירות בעניין תב"כ 9/24 ומשלימה לה. כמו כן, אנו מציעים כי ועדת הבחירות תפנה מבעוד מועד למועצת העיתונות על מנת להוביל לגיבוש קווים מנחים באשר לפרסום או שידור של תכנים אודיו-ויזואליים אשר קיים חשש מבוסס לגבי מהימנותם.¹⁶

ג.4 פעולות מקדימות: יצירת ממשקים בין ועדת הבחירות לרשתות החברתיות ופלטפורמות תוכן

כפי שנוכחנו לדעת, הרשתות החברתיות ופלטפורמות תוכן מקוונות, כגון פייסבוק, יוטיוב (בבעלות גוגל) וטוויטר שינוי באופן יסודי את דרכי ההפצה של מידע להמונים, באופן העוקף את המדיה המסורתיים כ- "שומרי הסף" על המידע שמועבר לכלל הציבור. לאחרונה, בעקבות 'אירועי הקפיטול' בארצות הברית, נוכחנו לדעת כי פלטפורמות תוכן אלו (והאלגוריתמים המארגנים את התוכן שהן מפיצות) מאפשרות להפיץ בקלות ובמהירות מידע כוזב בהקשרי בחירות, אשר עשוי לא רק להשפיע באופן בלתי הוגן על הצבעות הבוחרים אלא גם עשוי לגרום נזקים "בעולם האמיתי" והפרעה להליך ההצבעה או לחילופי השלטון.

למעשה, פלטפורמות התוכן עצמן מכירות כיום בסכנות הייחודיות של תכני זיוף עמוק (ובמיוחד בתקופת בחירות) וקובעות איסורים או תנאים מחמירים באשר לאופן שמשתמשיהן רשאים להפיץ תכנים מסוג זה. לאחרונה, גם וועדת הבחירות הנכבדה היטיבה להכיר בסכנות הדמוקרטיות של הפצת זיוף עמוק בסמוך לתקופת הבחירות, המחייבות אותה לפעול על מנת לעצור פרסומי כזב

¹⁵ <https://bechiro22.bechiro22.gov.il/election/Committees/Pages/fake-news.aspx>

¹⁶ זאת, בהתאם לחובה האתית והמשפטית החלה על עיתונאים לוודא את אמינותו של מידע הנמסר להם טרם פרסומו כידיעה, ולבטח כאשר מידע זה צפוי להזיק או לפגוע בגורם ספציפי.

מקוונים, לרבות ברשתות חברתיות, ככל שאלו עלולים – בוודאות קרובה – להטעות את הבוחר הסביר או לשבש את ההליך התקין של הבחירות.

עם זאת, נוכח המורכבות המשפטית והיישומית הכרוכה בחיוב פלטפורמות התוכן להסיר תוכן ספציפי (ככל שתמצא וועדת הבחירות כי זה צפוי בוודאות קרובה לפגוע באופן ממשי בהליך הדמוקרטי), אנו מציעים לוועדת הבחירות לפנות בהקדם לפלטפורמות התוכן המרכזיות, וביניהן פייסבוק, טוויטר וגוגל, על מנת לבחון את מוכנותן של הפלטפורמות לטיפול מיידי באירועים אלה, כפי שנעשה אף בארה"ב ערב בחירות 2020, על ידי חבר הקונגרס אדם שיף, יו"ר וועדת המודיעין של בית הנבחרים של ארה"ב.¹⁷

כלל זה, מוצע כי ועדת הבחירות תבדוק מבעוד מועד את נכונותן ומוכנותן של הרשתות החברתיות ליישם באופן מיידי את החלטותיה בדבר חסימה או סימון של תכנים מסוג זיוף עמוק, אשר נמצא כי הם צפויים בוודאות קרובה להטות באופן לא הוגן את עמדות המצביעים או להפריע למהלכם התקין של הבחירות.

ד.4 הסתייעות בכלים טכנולוגיים-פורנזיים מתקדמים כאשר קיימת אי-ודאי באשר להיותו של

תוכן מזויף או אמיתי

בצד התפתחותה המואצת של טכנולוגיית הזיוף העמוק ככלי ליצירה או עיבוד של תוכן מזויף, מתפתחת גם טכנולוגיה מתקדמת לזיהוי ממוחשב של תכנים אלו. לאחרונה, ענקיות הטכנולוגיה החלו להציע כלי ניתוח ומאגרי נתונים המתיימרים לזהות את הסבירות שתוכן אודיו-ויזואלי מסוים עובד או טופל באופן מלאכותי, ובפרט בכל הנוגע למאפייני זהות כגון פנים או מידע ביומטרי.¹⁸ זאת, בצד יוזמות פרטיות נוספות המאפשרות למשתמשים להעלות ולנתח סרטונים החשודים כנוצרו או עובדו באמצעות טכנולוגיית הזיוף העמוק, כגון DeepWare, המספקת ממשק אינטרנטי להעלאה של קבצי וידאו וניתוח ההסתברות לכך שאלו עברו עיבוד מסוג זיוף עמוק, באמצעות טכנולוגיות של בינה מלאכותית.¹⁹ יודגש, כי מהימנותם או איכותם של כלים אלו לא נבדקה על ידנו וכן כי קיימות חברות פרטיות נוספות המציעות טכנולוגיה לזיהוי תכני זיוף עמוק, כך שאין בסקירה זו משום רשימה ממצה או המלצה על שירות מסחרי מסוים. עם זאת, נבקש לספק לוועדת הבחירות הסבר טכנולוגי תמציתי על הטכניקות החישוביות שבאמצעותן מקובל לבחון האם תוכן אודיו-ויזואלי עובד או נוצר באמצעות זיוף עמוק.

¹⁷ Congressman Adam Schiff, Schiff Presses Facebook, Google and Twitter for Policies on Deepfakes Ahead of 2020 Election, (15.7.2019) <https://web.archive.org/web/20201104141231/https://schiff.house.gov/news/press-releases/schiff-presses-facebook-google-and-twitter-for-policies-on-deepfakes-ahead-of-2020-election>

¹⁸ בספטמבר 2020 הציגה חברת מיקרוסופט שני כלים טכנולוגיים להתמודדות עם דיסאינפורמציה במרחב המקוון, ותכני זיוף עמוק בפרט. הראשון, Microsoft Video Authenticator, מאפשר לזהות תכנים חזותיים שעברו מניפולציה מלאכותית באמצעות הדגשת פרטים חזותיים שעין בלתי מזוינת אינה יכולה להבחין בהם. השני, הוא מערכת חתימה דיגיטלית ליוצרי תוכן, אשר מאפשרת להתריע לצרכני התוכן במקרה של שינוי או שיבוש התוכן המקורי שנחתם על ידי המפרסם. ראו:

<https://blogs.microsoft.com/on-the-issues/2020/09/01/disinformation-deepfakes-newsguard-video-authenticator>

¹⁹ <https://deepware.ai>

כיום, נהוגות שתי טכניקות מרכזיות לזיהוי תכני וידאו שנוצרו או עובדו באמצעות טכנולוגיות הזיוף העמוק: ²⁰ הראשונה, Low-level forensics, מתמקדת בסריקת הפיקסלים המרכיבים את התוכן החזותי במטרה לאתר "פרטי לוואי" (artifacts) שנוצרים כתוצאה מהפעלת טכנולוגיית הזיוף העמוק. השנייה, High-Level forensics, אינה מבוססת על איתור "חתימות מוכרות" של עיבוד תמונה אלא מנתחת את האמינות או העקביות של מאפייני הדמות האנושית המוצגת בסרטון, כגון מצמוצי עיניים, תנועות ראש, הבעות פנים, דופק או מידע ביומטרי נוסף.²¹ טכניקה זו נחשבת ליעילה יותר בהתמודדות עם השיפור המתמיד בטכנולוגיית הזיוף העמוק, שכן היא מבוססת על טכנולוגיות של למידה עמוקה ורשתות נוירוניות (המשמשות כאמור גם ליצירת זיוף עמוק) על מנת לשפר את יכולותיהן לאורך זמן.²²

העובדה כי במגזר הפרטי קיימים כלים טכנולוגיים לבחון האם תוכן אודיו-ויזואלי הוא "אותנטי" או שמא עובד או נוצר באמצעות זיוף עמוק, מאפשרת להניח כי יכולות אלו נמצאות גם בידי כוחות הבטחון בישראל. על כן, אנו ממליצים כי ועדת הבחירות המרכזית תבחן את שאלת יכולתה להסתייע במומחיות ובכלים הפורנזיים שעומדים לרשות מערכת הביטחון ואכיפת החוק בישראל, ככל שקיימת חוסר-ודאות עובדתית לגבי האותנטיות של תוכן אודיו-ויזואלי, אשר צפוי להטעות את ציבור הבוחרים או לפגוע במהלכם התקין של הבחירות.

4. הצורך בטיפול מיידי במידע כוזב וזיוף עמוק: משמעויות משפטיות נוספות

כאמור, נוכח קצב ההתפשטות המהיר, הנרחב והמגוון של חדשות כוזב בכלל וזיוף עמוק בפרט, אותו מאפשרות פלטפורמות התוכן ורשת האינטרנט בכלל, התגובה של ועדת הבחירות אליהם נדרשת להיות מהירה במיוחד, לאור הסכנה הגדולה הטמונה בויראליות הדברים. **נוכח הכרה בחשיבותה של תגובה מיידי לאיומים הדמוקרטיים של הפצת זיוף עמוק בתקופת הבחירות, אנו ממליצים כי ועדת הבחירות תגבש מבעוד מועד נהלים וסדרי פעולה מהירים לטיפול בתלונות על הפצת מידע כוזב אשר צפוי להטעות את ציבור הבוחרים או לפגוע בהתנהלות הבחירות. זאת, הן ביחס להיבטים מינהליים של זמינות ועדת הבחירות לקליטתן של תלונות על תכני מידע כוזב וזיוף עמוק במרחב המקוון; הן ביחס לדוקטרינה או המסגרת המשפטית אותה מפעיל יו"ר ועדת הבחירות כדי להכריע; והן ביחס לממשקי הפעולה של ועדת הבחירות מול הרשתות החברתיות לשם אכיפה מיידי של החלטות הוועדה.**

כמו כן, אנו מציעים לשקול כי ועדת הבחירות תקבע כללים מיוחדים באשר לפרסום מידע כוזב או זיוף עמוק בתקופה של 48 שעות לפני מועד הבחירות. פרסום כזה בסמוך למועד הבחירות עשוי להשיג את מטרתו מחד גיסא, ולא להותיר זמן למערכת המשפטית או הפוליטית להגיב מאידך גיסא. על כן,

²⁰ Shrutli Agarwal et al., Detecting Deep-Fake Videos from Appearance and Behavior (2022) קישור: <https://farid.berkeley.edu/downloads/publications/wifs20.pdf>

²¹ ראו למשל: FakeCatcher, פרויקט משותף של אוניברסיטת בינגהמטון וחברת אינטל, המספק כלי לזיהוי זיוף עמוק באמצעות ניתוח קצב פעימות הלב של הדמות האנושית המופיעה בסרטון. קישור: <https://ieeexplore.ieee.org/document/9141516>

²² Agarwal et al., לעיל ה"ש 20.

מוצע לשקול גיבוש של אמצעים מיוחדים (הן מבחינת זמן הפעולה, ואולי גם מבחינת עוצמת הכלי המשפטי) במקרה שנעשה בזיוף עמוק בתקופה של 48 שעות לפני מועד הבחירות, או שעולה חשש מהותי לכך.

באופן כללי יותר, אנו קוראים לוועדת הבחירות המרכזית לפתח את ההתייחסות הייחודית לתופעת הזיוף העמוק והמענה לה, לאור התובנות שהצגנו לעיל, לפיהן פוטנציאל הנזק של חדשות כזב וזיוף עמוק המופצות במרחב המקוון הוא גבוה במיוחד. תופעות אלו אותן מאפשרת הטכנולוגיה העכשווית מאתגרות עמוקות גם את החשיבה המשפטית המסורתית אודות עיקרון חופש הביטוי וההצדקות להגנה עליו. טרם עידן הרשתות החברתיות, הצדקות "חקר האמת" ו-"שוק הרעיונות" להגנה על חופש הביטוי של הפרט הניחו כי אין הצדקה לקיים מניעה מוקדמת של ביטוי שקרי, שכן זה יעומת עם ביטויים מתחרים בזירה הציבורית ומתוך כך יתבררו השקר והאמת. כיום, כפי שנוכחנו דמוקרטיות רבות לדעת בשנים האחרונות, לא תמיד ניתן לסמוך על כך שהאמת תגבר על השקר בעידן עודף-המידע הנוכחי.

כאמור, אנו מציעים כי במסגרת הפעלת מבחן הוודאות הקרובה להערכת נזקיו הפוטנציאליים של תוכן מסוג זיוף עמוק, כמבחן הסתברותי מבוסס-עובדות, יש להכיר בכך כי הוויראליות הייחודית של מידע כוזב – והכוח המניפולטיבי הייחודי של זיוף עמוק – מגבירים באופן משמעותי את היקף הנזק הצפוי מהם בדמות הטעיית קהל הבוחרים או שיבוש מהלכם התקין של הבחירות והאמון הציבורי בהם.

נשמח לעמוד לרשותכם בכל הבהרה נוספת שתידרש, לרבות העברת חומרים מקצועיים, היכרות עם מומחים רלוונטיים לזיוף עמוק ובינה מלאכותית או לשם יצירת ממשקים מול פלטפורמות התוכן והרשתות החברתיות הפועלות בישראל. ניתן לפנות אלינו באמצע הדוא"ל: policy@isoc.org.il

בברכה ובכבוד רב,

ד"ר אסף וינר
ראש תחום רגולציה ומדיניות
איגוד האינטרנט הישראלי (ע"ר)

ד"ר דלית קן-דרור פלדמן
מנחה משפטית
הקליניקה למשפט וסייבר
אוניברסיטת חיפה

עו"ד יורם הכהן
מנכ"ל
איגוד האינטרנט הישראלי (ע"ר)

העתיקים:

עו"ד אורלי עדס, מנהלת כללית, ועדת הבחירות המרכזית לכנסת ה-24.
עו"ד דין ליבנה, יועץ משפטי, ועדת הבחירות המרכזית לכנסת ה-24.
הוועד המנהל של איגוד האינטרנט הישראלי (ע"ר).