28 באוגוסט 2022
א' באלול, התשפ"ב

פרטיות ואבטחת מידע בשימוש בטכנולוגיות Deepfake (זיוף עמוק)

נוסח מעודכן בעקבות תיקון 13

פתח דבר

1. אחת התופעות הטכנולוגיות הצוברות פופולריות רבה בעת האחרונה היא תופעת השימוש בטכנולוגיית Deepfake (זיוף עמוק),[1] המאפשרת יצירה ועריכה דיגיטלית של תמונות, קטעי קול וקטעי וידאו מזויפים של אנשים או אירועים.

2. כפי שיפורט, לתופעה זו עלולות להיות השלכות קשות, בין היתר, על היבטי פרטיות. במסמך זה מציגה הרשות להגנת הפרטיות (להלן: 'הרשות') את עמדתה בנושא. המסמך מציג את הקשר בין התופעה לפגיעה בפרטיות, ומפרט על אמצעים והמלצות להתמודדות עם התופעה.

3. הרשות מבהירה כי לעמדתה, **הפצה ללא הסכמה של תמונה או סרטון שזויפו בטכנולוגיית Deepfake - המציגים תוכן משפיל או כזה הנוגע לצנעת חייו האישיים של אדם ויכול להיתפס בציבור כאותנטי - מהווה הפרה של חוק הגנת הפרטיות, התשמ"א-1981 (להלן: 'חוק הגנת הפרטיות')**.

4. הרשות מבקשת להביא לצמצום הפצתם ללא הסכמה של תכנים משפילים או כאלו הנוגעים לצנעת חייו של אדם, המזויפים באיכות גבוהה, ואשר יש בהם כדי לפגוע ביכולתו של אדם לשלוט על מידע אישי הנוגע אליו. כך למשל, פרסום סרטון מזויף בדבר הימצאותו של אדם במקום בעל רגישות מיוחדת (כגון במתחם הימורים) החושף לכאורה מידע על צנעת חייו ועל מצבו האישי (לרבות מצבו הנפשי או הכלכלי).

5. כמו כן, הרשות קובעת שחברות המייצרות תכנים בטכנולוגיית Deepfake מחויבות לעמוד בהוראות תקנות הגנת הפרטיות (אבטחת המידע), התשע"ז-2017 (להלן: 'תקנות אבטחת מידע'), וזאת כל עוד מידע זה עשוי להיתפס בציבור כמידע אישי אותנטי ואמיתי. כלומר, **עמדת הרשות היא כי מידע מזויף, העלול להיתפס בציבור כמידע אישי אותנטי, הוא מידע שיש לפעול לאבטחתו על פי הדין, כאילו היה מדובר במידע אישי אמיתי**. לגישת הרשות, תכנים מזויפים באיכות העלולה לגרום לציבור להניח כי מדובר במידע אישי אותנטי, מהווים 'מידע' לפי חוק הגנת הפרטיות, וככאלו חלות עליהם כל הוראות החוק וחובות אבטחת המידע לפי התקנות.

6. המסמך אינו מבקש למנוע את השימוש בטכנולוגיית Deepfake אלא לחדד את החשיבות בשמירה על פרטיות בשימוש זה. הרשות מציגה מספר המלצות מעשיות להתנהלות הציבור

[1] המילה "Deepfake" היא הלחמה של המינוחים "Deep learning" (למידה מעמיקה) ו "fake" (זיוף).

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

בעניין, ובין היתר ממליצה לחברות המייצרות תכני Deepfake להשתמש בסימני מים דיגיטליים-גלויים, המבהירים כי התכנים עברו עריכה וכי הם אינם אותנטיים.

רקע

7. טכנולוגיית Deepfake היא שם כללי לאלגוריתמים המבוססים על בינה מלאכותית (AI), המאפשרים יצירה של תוכן דיגיטלי מזויף מסוג תמונה, סרטון וידאו, וקובץ שמע או עריכה מניפולטיבית של תוכן אמיתי והפיכתו למזויף. לאור איכותו, התוכן המזויף עשוי להיתפס ברבים כ"אמיתי", כלומר כייצוג אותנטי של המציאות.

8. בהחלטת יו"ר ועדת הבחירות המרכזית לכנסת ה-24, הוגדרה הטכנולוגיה האמורה כ"**טכנולוגיה ליצירת תוכן קולי או חזותי או לשינוי תוכן קיים, כך שהצופה הסביר (ואף הצופה המתוחכם) 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**" (ההדגשות אינן במקור).[2]

9. השימוש בטכנולוגיה מאפשר הפצה של תמונות או קטעי וידיאו שעברו מניפולציה או פוברקו. הטכנולוגיה מאפשרת, בין היתר, "להשתיל" או להחליף (Face swap) דמויות פנים של אנשים בתמונות ובקטעי וידאו. הטכנולוגיה מאפשרת גם לתמרן הבעות פנים ולעוות מלל הנאמר על-ידי אנשים בקטעי וידאו, וזאת באמצעות זיוף קול ואופן תנועת פני הדוברים.

10. הזיוף נעשה, בין היתר, על-ידי שילוב של תמונות וקטעי וידאו מקוריים עם תכנים פיקטיביים. על-פי רוב, המידע האמיתי (או האותנטי) מוזן לאלגוריתם מסוג בינה מלאכותית ולמידה עמוקה (deep learning), ועל בסיס מידע זה מייצר האלגוריתם מידע שונה ומזויף. **ככל שהמידע המוזן (כגון תמונות) הוא רב ומגוון יותר, כך למידת המכונה תהיה איכותית יותר, והמניפולציה של המידע תיעשה באופן אמין ומציאותי יותר**. טכנולוגיה מתקדמת עשויה אף לייצר דמויות וסיטואציות על בסיס מידע מלאכותי (מידע סינטטי) לחלוטין.

11. טכנולוגיית ה-Deepfake עשויה לשמש למטרות לגיטימיות כגון אומנות, פרסום וסאטירה. עם זאת, בשנים האחרונות ניתן לזהות שימוש רב בטכנולוגיה זו למטרות פסולות, כגון השפלה וביוש אנשים ברבים, יצירת תוכן מיני ופורנוגרפי מזויף,[3] והטיה של שיח פוליטי (על-ידי זיוף סרטונים של דמויות פוליטיות וכדומה).[4]

12. **דוגמה לשימוש לרעה בטכנולוגיה ניתן לראות בפרסום לפיו הרשויות בארה"ב עצרו אישה לאחר שנחשף כי היא השתמשה בטכנולוגיית Deepfake כדי לערוך סרטונים ותמונות של יריבותיה של בתה בנבחרת המעודדות, כך שייראו כאילו עישנו, צרכו אלכוהול והצטלמו**

---

[2] תב"כ 24-9 **יש עתיד – בראשות יאיר לפיד נ' עמותת "כן לשלום"** (פורסם בנבו, 18.1.2021), בעמ' 2 להחלטה.

[3] במחקר שנערך בתחום ופורסם בספטמבר 2019 נמצא כי 96% מקטעי הווידאו שנוצרו בטכנולוגיה של Deepfake היו בעלי תוכן פורנוגרפי שנוצר בהיעדר הסכמה.

[4] סוגיית השימוש בטכנולוגיית ה-Deepfake להטיית שיח פוליטי היא סוגיה נפרדת ומסמך זה אינו עוסק בה.

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

בעירום, והכל במטרה להביא לסילוקן מהקבוצה.[5] דוגמה אחרת היא ההצלחה של אתרים המאפשרים "להסיר" באופן דיגיטלי ביגוד מתמונות רגילות של אנשים, באופן שמציג אותם בעירום.[6]

13. בנוסף, טכנולוגיית Deepfake הופכת לאיום אבטחת מידע הולך וגובר לארגונים. גורמים עבריינים משקיעים בבינה מלאכותית ולמידת מכונה כדי ליצור תוכן דיגיטלי סינתטי ו"מזויף" (כגון תמונות, וידאו, אודיו וטקסט) לצורך שימוש במתקפות סייבר והונאות. מתקפות והונאות אלו מתרחשות מכיוון שתוכן אשר משכפל באופן ריאלי את המראה, הקול, הגינונים או אוצר המילים של אדם, מאפשר להערים על מטרות אנושיות ואוטונומיות להאמין שמה שנחזה או מוצג להן הוא אותנטי ואמין. גורמים עבריינים עושים למעשה שימוש בטכנולוגיות מסוג Deep fake כדי להתחזות בפני מערכי זיהוי ביומטריים, פקידים אנושיים או עמיתים לאדם בעל הרשאה במסגרת מתקפה של הנדסה חברתית, וזאת על מנת לקבל גישה פסולה למערכות ומאגרי מידע.

14. ככלל, השימוש בטכנולוגיית ה-Deepfake הופך לקל ופשוט יותר ויותר. נכון להיום קיימות אפליקציות, הניתנות להורדה בחינם, והמאפשרות יצירת תכנים מזויפים ללא ידע טכנולוגי נרחב. **זמינות וקלות השימוש בטכנולוגיה זו, בשילוב עם ריבוי ומגוון התמונות וקטעי וידאו הנמצאים ברשת (ובעיקר ברשתות חברתיות), עלולים להביא לעלייה ביצירה ובהפצה של תכנים מזויפים באיכות גבוהה**. אם בשלהי הפיתוח של הטכנולוגיה נדרשו היקפים נרחבים של תכנים כדי שהמחשב יוכל לנתח אותם ולייצר תוכן דיגיטלי אמין, כיום אפילו מספר תמונות יכולות להספיק לשם כך. בנוסף, מבחינה טכנולוגית, אפשרות זו הפכה לקלה מאי פעם, וכבר אינה דורשת ציוד משוכלל כבעבר. כפי שיפורט להלן, למציאות זו השלכה עצומה על הפרטיות.

Deepfake בראי הזכות לפרטיות

פגיעה בפרטיות על-פי סעיף 2 לחוק הגנת הפרטיות

15. סעיף 1 לחוק הגנת הפרטיות קובע כי "לא יפגע אדם בפרטיות של זולתו ללא הסכמתו". סעיף זה מבטא את אחד העקרונות המרכזיים בדיני הגנת הפרטיות והוא **עיקרון ההסכמה**. מתוך עיקרון זה עולה גם כי הפרט הוא האחראי ביחס לאיזה מידע הנוגע אליו ייחשף, למי, ולאילו מטרות.[7] **לכן, ככלל, איסוף ושימוש במידע אישי ללא הסכמה מהווה פגיעה בפרטיות.**[8]

---

[5] "אימא יצרה דיפ-פייק כדי לסלק בנות מנבחרת המעודדות" **Ynet** (16.3.2021).

[6] Matt Burgess, *The Biggest Deepfake Abuse Site Is Growing in Disturbing Ways*, WIRED (15. Dec 2021) לפי הכתבה, האתר זכה ליותר מ-50 מיליון ביקורים בין ינואר לסוף אוקטובר של שנת 2021. עוד צוין כי "מאות אלפי" תמונות הועלו ביום אחד, וכי בחודש אוגוסט האתר הגיע לשיא עם 6.92 מיליון צפיות.

[7] תפיסה זו ידועה בשם "פרטיות כשליטה". להרחבה ראו מיכאל בירנהק **מרחב פרטי: הזכות לפרטיות בין משפט לטכנולוגיה**, 89-108 (2010).

[8] עם זאת, איסוף ושימוש במידע מכוח הסמכה שבדין עשוי, במקרים רבים, שלא לחייב הסכמה. ככלל, גם עצם איסוף מידע ממקורות פומביים (כגון רשתות חברתיות "פתוחות") אינו מחייב שלעצמו את הסכמת נושא המידע.

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

16. סעיף 2 לחוק הגנת הפרטיות מציין רשימה של מקרים שכל אחד מהם מהווה פגיעה בפרטיות. שלושה מהם רלוונטיים ישירות לתופעת ה- Deepfake:

א. סעיף 2(4) לחוק קובע מפורשות כי פרסום ללא הסכמה של תצלומו של אדם ברבים בנסיבות שבהן עלול הפרסום להשפילו או לבזותו מהווה פגיעה בפרטיות.

ב. סעיף 2(6) מייחס פגיעה בפרטיות בנסיבות של שימוש ללא הסכמה בתמונה או בקול של אדם למטרת רווח.[9]

ג. סעיף 2(11) קובע כי פרסום ללא הסכמה של ענין הנוגע לצנעת חייו האישיים של אדם (בהתייחס בין היתר לעברו המיני, למצבו הבריאותי או להתנהגותו ברשות היחיד) מהווה גם הוא פגיעה בפרטיות.

כלומר, הוראות חוק הגנת הפרטיות אוסרות הפצה ללא הסכמה של תמונה או קטע וידאו אשר יש בהם כדי לבזות או להשפיל את האדם המצולם או המוקלט. כך, גם בנוגע להפצה ללא הסכמה של תוכן דיגיטלי המציג אדם תוך התייחסות לצנעת חייו האישיים, או לשימוש בתמונה וקול של אדם למטרות מסחריות.

17. **לגישת הרשות, הוראות חוק אלו חלות גם על פרסום והפצה של תוכן שזויף באופן דיגיטלי, במידה שהוא נעשה באיכות העלולה להטעות את הציבור ולגרום לו להניח שהוא אותנטי.**

כלומר, **פרסום והפצה ללא הסכמה של תמונה או קטע וידאו שזויפו באמצעים טכנולוגיים כך שהם עלולים להיתפס כאותנטיים, ושיש בהם בכדי להשפיל ולבזות אדם המופיע בהם או להציג מידע הנוגע לכאורה לצנעת חייו האישיים – מהווים הפרה של הוראות חוק הגנת הפרטיות. כמו כן, כל זיוף ללא הסכמה של תמונה, קטע וידאו או קטע קול לצרכים מסחריים מהווה גם הוא פגיעה בפרטיות.**

18. יובהר כי אין הכוונה רק לתמונות או קטעי וידאו בעלי אופי מיני.[10] **גם הפצה של תכנים משפילים או מביכים אחרים עלולה להיחשב הפרה של הוראות החוק, כל עוד הפצה זו נעשית באופן העלול להטעות את הציבור בנוגע לאדם המתועד בהם.**[11] כך לדוגמה, פרסום ללא הסכמה של סרטון מזויף של אדם הנוקט באלימות, או סרטון מזויף של אדם המתבטא באופן פוגעני כלפי אחרים, עשויה להיחשב פגיעה בפרטיותו של האדם, וכפרסום המנוגד להוראות חוק הגנת הפרטיות.

---

[9] הגדרת שימוש בסעיף 3 לחוק הגנת הפרטיות היא "כל פעולה שמבוצעת על מידע אישי, לרבות קבלתו, איסופו, אחסונו, העתקתו, עיון בו, גילויו, חשיפתו, העברתו, מסירתו או מתן גישה אליו".

[10] למען הסר ספק יצוין כי בית המשפט העליון קבע לא פעם שהפצת תמונות בעלות אופי מיני ללא הסכמה (כגון תמונות עירום) מהווה פגיעה בפרטיותו של אדם, כמו גם הלבנת פניו וביזויו ברבים, וכן פגיעה בזכותו לכבוד ולשם טוב. ראו לדוגמה רע"פ 1024/21 **פלוני נ' מדינת ישראל** (פורסם בנבו, 17.3.2021).

[11] מבוכה היא מצב רגשי המבטא לרוב אי-נוחות מהאופן שבו התנהלות מסוימת של האדם נתפסת על-ידי אחרים בחברה. להרחבה ראו: Erving Goffman, Embarrassment and Social Organization, 62 AM. J. SOCIOLOGY 264 (1956).

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

19. הפצה שכזו, מעבר להיבט ההשפלה שבה, יש בה כדי לפגוע ביכולתו של האדם לשלוט על זהותו ותדמיתו בקרב הציבור, ומכאן על האופן שבו הוא נתפס על-ידי הציבור.[12] פגיעה שכזו ביכולת השליטה של אדם על מידע הנוגע אליו מהווה פגיעה בפרטיות.[13] מכאן **שבכל הנוגע לפגיעה בפרטיות אין הבחנה בין פרסום ללא הסכמה של תוכן פוגעני או משפיל שהינו אמיתי ואותנטי, לבין פרסום תוכן מזויף שכזה.**[14] המבחן המרכזי בהקשר זה הוא האם הפרסום המזויף (בהינתן שהוא איכותי דיו) יגרום לציבור להניח שאדם מסוים אכן פעל באופן המתועד או המוקלט.

20. בית המשפט העליון, בהתייחסותו לסוגיית הפרסום וההפצה ללא הסכמה של תכנים בעלי אופי מיני, ציין גם הוא את הקשר בין שליטה על מידע לשליטה על דימוי ציבורי:

> "יסוד ההסכמה חשוב במיוחד, שכן הוא מבטא את האוטונומיה של הפרט במובנה המובהק ביותר, דהיינו בהחלטות של אדם אילו פרטים אינטימיים לחשוף לאחרים, למי לחשוף אותם **ובניסיונו לשלוט בדימוי שלו בקרב מכריו ובציבור בכלל** [...] חומרת התופעה מתחזקת בשל קלות ההפצה שמאפשרת הקִדמה הטכנולוגית" (ההדגשה אינה במקור).[15]

21. עד כאן בכל הנוגע לפגיעה בפרטיות על-פי סעיף 2 לחוק הגנת הפרטיות. לעניין פרסום תמונות וקטעי וידאו בעלי אופי מיני, יודגש כי עניין זה מוסדר גם בהוראות סעיף 3(א)(5א) לחוק מניעת הטרדה מינית, השתנ"ח-1998 (להלן: 'חוק למניעת הטרדה מינית').

סעיף זה קובע כי פרסום תצלום, סרט או הקלטה של אדם (לרבות עריכה או שילוב של כל אחד מהם, ובלבד שבנסיבות העניין ניתן לזהות את האדם), המתמקד במיניותו, בנסיבות שבהן הפרסום עלול להשפיל את האדם או לבזותו, ולא ניתנה הסכמתו לפרסום, מהווה הטרדה מינית (וזאת כאמור, בנוסף לפגיעה בפרטיות). על-פי פרשנות החוק למניעת הטרדה מינית, האחריות

---

[12] הסוציולוג ארווינג גופמן טען שלכל אדם בסיטואציה חברתית יש אינטרס לשלוט על היחס שאחרים יגלו אליו וזאת תוך יצירת רושם מסוים עליהם ושליטה במידע. ERVING GOFFMAN, THE PRESENTATION OF SELF IN EVERYDAY LIFE 3-4, 8 (1956).

[13] על-פי המשפטן צ'ארלס פריד (Fried), פרטיות אינה רק היעדר מידע הנוגע אלינו במחשבותיהם של אחרים, אלא השליטה שיש לנו על מידע הנוגע אלינו. Charles Fried, *Privacy*, 77 YALE L.J.475, 482-483 (1968). ליסה אוסטין טוענת כי צילום אדם במרחב הציבורי והפצת התמונה במגזין עשויים לשנות באופן דרמטי את יחסו של הציבור, מולו האדם מציג את עצמו. שינוי זה, לטענתה של אוסטין, מערער את יכולתו של האדם להציג את עצמו על-פי רצונו, ומקים, לכאורה, את הטענה בדבר פגיעה בפרטיות. לגישת אוסטין, פגיעה זו מתרחשת למרות שהתמונה צולמה במרחב ציבורי ואף בנסיבות בהן הפרסום לא היה משפיל או מסגיר מידע רגיש. ראו: Lisa M. Austin, *Control Yourself, or at Least Your Core Self*, 30 BULL. SCI. TECH. & SOC'Y 26, 29 (2010).

[14] על הקשר בין פרסום מידע שאינו אותנטי לבין פגיעה ביכולת השליטה במידע אישי ופגיעה בפרטיותו של אדם ניתן ללמוד גם מפסק הדין בעניין **וינר**, שם קבע בית המשפט, בעניין פרסום והפצה של תוכן אינטימי מוטעה הנוגע לתובע, כי: "לא רק שתמונתו האינטימית של התובע הועברה לאנשים זרים ללא ידיעתו וממילא ללא הסכמתו, אלא שהופץ ברשת, או לכל הפחות באפליקציית הגריינדר, מידע אינטימי מוטעה אודותיו, כאילו הוא מקיים יחסי מין מסוג מסוים עם אדם מסוים, כאשר מידע זה הוא שקרי. בכך למעשה גזל הנתבע מהתובע את שליטת התובע על המידע אודותיו ופגע בפרטיותו". ראו: ת"א (הרצליה) 18-02-1158 **טל וינר נ' אייל קרול** (פורסם בנבו, 10.1.2021), (להלן: 'עניין **וינר**'), בפסקה 107.

[15] ע"פ 5090/18 **מדינת ישראל נ' פלוני** (פורסם בנבו, 18.11.2018), בפס' 10 לפסק דינו של השופט פוגלמן.

בעניין זה מוטלת לא רק על הגורם המצלם והמפיץ הראשי של התכנים, אלא גם על מי שקיבל אותם והעביר אותם הלאה לקבוצת אנשים, לדוגמה במסגרת רשתות חברתיות או אפליקציות להעברת מסרים מיידים כגון וואטסאפ וסיגנל.[16]

הסכמה במסגרת יצירת ופרסום תוכן דיגיטלי מזויף

22. כאמור, הסכמה היא עיקרון מרכזי בדיני הגנת הפרטיות. ככלל, הפצה של תמונות או קטעי וידאו, הנעשית בהסכמה, אינה מהווה פגיעה בפרטיות, וזאת ללא קשר לתוכנו של הפרסום.

23. סעיף 3 לחוק הגנת הפרטיות מגדיר הסכמה כ"הסכמה מדעת, במפורש או מכללא". מכללא – קרי באופן שאינו מפורש אך משתמע מתוך התנהגות וכדומה.

24. אחד האתגרים המרכזיים הקיימים בנושא עוסק בהסכמה הניתנת במצבים בהם יש אי שוויון מובנה בין הצד המבקש לצד המעניק את ההסכמה. הסכמה זו מוגדרת כהסכמה "חשודה".[17]

25. **באופן דומה, עמדת הרשות היא כי תוכן דיגיטלי שעבר עריכה או מניפולציה, ובמיוחד כזה שיש בו כדי להשפיל, להביך או להביא לפרסום של עניין הנוגע לצנעת הפרט של אדם, הוא תוכן "חשוד" - קרי שקיים חשש כי הוא נוצר ללא הסכמת האנשים המתועדים בו**.

אבטחת מידע מזויף במאגרי מידע

26. פרק ב' לחוק הגנת הפרטיות עוסק בהגנה על פרטיות במאגרי מידע. מידע אישי, כהגדרתו בסעיף 3 לחוק הגנת הפרטיות, הוא "נתון הנוגע לאדם מזוהה או לאדם הניתן לזיהוי", קרי "מי שניתן לזהותו במאמץ סביר, במישרין או בעקיפין, ובכלל זה באמצעות פרט מזהה כגון שם, מספר זהות, מזהה ביומטרי, נתוני מיקום, מזהה מקוון, או נתון אחד או יותר הנוגע למצבו הפיזי, הבריאותי, הכלכלי, החברתי או התרבותי".

27. על-פי סעיף 3 לחוק הגנת הפרטיות, הוגדרו סוגי מידע מסוימים כ"**מידע בעל רגישות מיוחדת**", כגון: מידע על צנעת חיי המשפחה של אדם, על צנעת אישיותו ועל נטייתו המינית; מידע המתייחס למצב בריאותו של אדם; מידע שהוא מזהה ביומטרי המשמש או המיועד לשמש לזיהוי אדם או לאימות זהותו באופן ממוחשב; מידע על מוצאו של אדם; מידע על עברו הפלילי של אדם; מידע על אודות דעותיו הפוליטיות או אמונותיו הדתיות של אדם או השקפת עולמו; נתוני מיקום העלולים לחשוף מידע בעל רגישות מיוחדת מסוגים שונים ביחס לאדם מסוים; ומידע על נתוני שכר של אדם ועל פעילותו הפיננסית. כמו כן, בעלי שליטה במאגרי מידע, שבמאגריהם נשמר מידע שכזה על אודות אדם, מחויבים באבטחת המידע בהתאם לקבוע בתקנות אבטחת מידע.

---

[16] הנחיית פרקליט המדינה מס' 29.2 "חקירה והעמדה לדין בעבירות של פרסום תצלומים, הקלטות או סרטים של אדם, בעלי אופי מיני, ללא הסכמתו", (1 בפברואר 2017).

[17] בירנהק, לעיל ה"ש 107, בעמ' 253. להרחבה בעניין זה בהקשר של מערכת יחסי עובד-מעסיק ראו: ע"א 8189/11 **רפאל דיין נ' מפעל הפיס** (פורסם בנבו, 21.2.2013), בפס' 34 לפסק דינה של השופטת ברק-ארז; ע"ע (ארצי) 90/08 **טלי איסקוב ענבר נ' מדינת ישראל - הממונה על חוק עבודת נשים** (פורסם בנבו, 8.2.2011).

28. חובת אבטחת המידע נועדה למנוע, או לכל הפחות לצמצם, את האפשרות לזליגת מידע על אודות אדם וחשיפתו ברבים, ללא הסכמתו לכך, ותוך פגיעה ביכולתו לשלוט במידע. זליגה וחשיפה שכאלו מהווים, מטבע הדברים, פגיעה בפרטיות.

29. לגישת הרשות, ובדומה לקביעת בית המשפט בעניין **וינר**, זליגה וחשיפה של מידע מזויף הנחזה להיות אותנטי ואמיתי עלולים להביא לפגיעה בפרטיות הדומה לפגיעה העלולה להיגרם מחשיפתו של מידע 'אמיתי'.

כך לדוגמה, תמונה של אדם בעת ביקור במרפאה המעניקה טיפולים רפואיים מסוג מסוים עלולה להיחשב כמידע על מצב בריאותו, בין אם מדובר בתמונה אותנטית שלו, ובין אם מדובר בתמונה שזויפה ברמה ובאיכות העלולים לגרום לציבור להאמין כי המדובר בתמונה אמיתית. במובן זה, זליגת התמונה תביא לפגיעה בפרטיותו של האדם המופיע בה וביכולתו לשלוט במידע הנוגע אליו, באופן דומה לזליגת התמונה האמיתית.

30. לפיכך, לגישת הרשות, **תכנים שעברו מניפולציה דיגיטלית באיכות העלולה לגרום לציבור להבין כי מדובר במידע אותנטי ואמיתי על אודות אדם, מהווים 'מידע' לפי הוראות פרק ב' לחוק הגנת הפרטיות**.[18]

<u>הבהרות והמלצות הרשות להגנת הפרטיות</u>

<u>הבהרות והמלצות לציבור</u>

31. יש להימנע מפרסום, הפצה ושיתוף ללא הסכמה של תמונות, קטעי קול וקטעי וידאו בעלי אופי מיני, וכן כאלו שיש בהם בכדי להשפיל אדם או להביא לפרסום של עניין הנוגע לצנעת חייו. כמו כן, יש להימנע משימוש ללא הסכמה בתמונה או בקול של אדם למטרת רווח. פעולות אלו עלולות להוות הפרה של הוראות חוק הגנת הפרטיות והחוק למניעת הטרדה מינית, לחשוף את המפר לתביעות אזרחיות, וכן להביא לנקיטת הליכים פליליים כנגדו.[19]

כלל זה חל בין אם מדובר בתכנים אותנטיים ואמיתיים, ובין אם מדובר בתכנים שעברו מניפולציה באמצעים טכנולוגיים (או שקיים חשש כי עברו מניפולציה שכזו).[20]

32. לפיכך, על כל מי שמבקש לפרסם ולהפיץ תכנים שכאלו (לרבות בדרך של העלאתם לרשתות חברתיות או העברתם באפליקציות כגון WhatsApp) – מוטלת האחריות לוודא כי פרסום זה נעשה בהסכמת נושא המידע.

---

[18] מסקנה זו עולה בין היתר מסעיף 14 לחוק הגנת הפרטיות, הקובע כי נושא מידע רשאי לפנות לבעל השליטה במאגר בבקשה לתיקון או למחיקת מידע עליו שאינו נכון.

[19] ראו לעניין זה סעיפים 4 ו-5 לחוק הגנת הפרטיות.

[20] כאמור, הנחת המוצא היא כי פרסום תוכן דיגיטלי מזויף, שיש בו להשפיל אדם או להביא לפרסום של עניין הנוגע לצנעת חייו, נעשה שלא בהסכמתו. מכאן, שכל העברה או פרסום נוסף של התוכן (לדוגמה, בעת העברת התוכן באפליקציות להעברת מסרים מיידים) – עלולים להוות הפרה של הוראות חוק הגנת הפרטיות.

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

33. בעת פרסום והפצה ברשת של תכנים אותנטיים "רגילים" (כגון בעת העלאת תמונות וקטעי וידאו שלכם או של אחרים לרשתות חברתיות) קחו בחשבון כי תכנים אלו עשויים לשמש בסיס למניפולציה וליצירה של תוכן מזויף. זכרו שככל שכמות, סוג ומגוון התמונות והסרטונים גבוה יותר – כך איכות הזיוף תהיה גבוהה יותר.

34. תמונות וקטעי וידאו שלכם עלולים לזלוג או להיגנב ממכשירכם, ומכאן לשמש כבסיס למניפולציה. על כן הקפידו על כללי אבטחת המידע, כגון אי-השארת מכשירכם ללא השגחה ושימוש בסיסמאות כניסה חזקות לתוכנות ולמכשירים הדיגיטליים שלכם.

35. בנסיבות בהן תמונות או סרטונים - אותנטיים או מזויפים - המציגים את דמותכם באופן שעלול להשפילכם או לחשוף מידע הנוגע לצנעת חייכם הופצו ברשת, או במקרה שהטלפון הנייד שלכם נגנב או נפרץ, ניתן לפנות ולהגיש תלונה במשטרת ישראל. במקרים של הפצת תמונות או סרטונים של ילדים או בני-נוער ניתן לפנות גם למוקד המטה הלאומי להגנה על ילדים ברשת (מוקד 105). כמו כן, מי שתמונותיו או סרטוניו הופצו ברשת באופן הפוגע בפרטיותו יכול לפנות ישירות אל אתרי האינטרנט בבקשה להסרת התכנים, וכן להגיש תביעה לפיצויים נגד מי שהפיץ אותם. כמובן שניתן גם לפנות לרשות להגנת הפרטיות לסיוע בעניינים אלו או בהפניה לגורמים הרלוונטיים המתאימים.

<u>הבהרות והמלצות לבעלי שליטה במאגרי מידע</u>

36. סעיף 17 לחוק הגנת הפרטיות ותקנות אבטחת מידע קובעים את חובות אבטחת המידע, ומתייחסים גם לניהול אמצעי הזיהוי והאימות לגישה למאגר.

37. **על בעלי שליטה במאגרי מידע לעמוד בחובות האבטחה הקבועות בדין וזאת גם ביחס למאגרים בהם נשמר מידע שזויף באמצעים דיגיטליים, וקיים חשש כי מידע זה ייתפס על-ידי הציבור כאותנטי ואמיתי**.

38. כאשר במאגר מידע נשמרים תכנים דיגיטליים מזויפים העשויים להיתפס על-ידי הציבור כמידע אותנטי על אודות אדם, עומדת לאותו אדם הזכות לפנות לבעל השליטה במאגר בבקשה לתיקון או למחיקת המידע, והכל בהתאם להוראות סעיף 14 לחוק הגנת הפרטיות. **כלומר, לגישת הרשות, תמונות או סרטונים מזויפים שעברו מניפולציה דיגיטלית ושעלולים להיתפס על-ידי הציבור כאותנטיים ואמיתיים, הם בגדר מידע ש"אינו נכון", ועל-פי סעיף 14 לחוק עומדת לאדם האפשרות לפנות לבעל השליטה במאגר בבקשה לתיקונם או למחיקתם מהמאגר**.

39. לשם צמצום הסיכון לפגיעה בפרטיות כתוצאה מדלף מידע או שימוש בו למטרות פסולות, **על בעלי השליטה במאגרים לפעול לצמצום מידע שאינו דרוש להם,**[21] **וחובה זו נכונה גם ביחס למידע מזויף השמור במאגר.** בעניין זה יצוין כי תקנה 2(ג) לתקנות אבטחת מידע, קובעת כי

[21] ככלל, שמירת מידע לאורך זמן מגבירה את הסיכון לפגיעה בפרטיות. ככל שמידע רגיש נשמר לתקופה ארוכה יותר – כך גובר הסיכון כי מידע זה ידלוף או ייחשף ויפגע קשות בפרטיות נושאי המידע.

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

"בעל מאגר מידע יבחן, אחת לשנה, אם אין המידע שהוא שומר במאגר רב מן הנדרש למטרות המאגר".[22]

40. כמו כן, בעידן בו טכנולוגיות Deepfake עלולות לאפשר זיהוי שגוי של אדם, ובכך לאפשר גישה לאדם שנחזה באמצעות טכנולוגיה זו לבעל הרשאה המורשה לגשת למאגר ולמערכות המאגר, אף שהוא אינו בעל ההרשאה האותנטי – **הרי שעל בעל שליטה במאגר לנהל את הסיכון בהתאם לאופי המאגר וטיבו, ולהתייחס גם לסיכונים אלה בעת שהוא בוחר את אמצעי הזיהוי המאפשרים את הגישה לבעלי הרשאה**.

<u>הבהרות והמלצות למפעילי ומפתחי אפליקציות ותוכנות מסוג Deepfake</u>

41. בעלי שליטה במאגרים שבהם מאוחסנים **התכנים הדיגיטליים האותנטיים** שמוזנים למערכות Deepfake על-ידי משתמשים, לצד **התכנים המזויפים** שנוצרים באמצעות המערכות, מחויבים באבטחת תכנים אלו בהתאם להוראות הדין, לרבות תקנות אבטחת מידע.[23] כאמור, לגישת הרשות, בכל הנוגע למידע לפי הוראות פרק ב' לחוק הגנת הפרטיות, אין זה משנה אם מדובר במידע מזויף או אותנטי, והמבחן המרכזי הוא האם קיים חשש שמידע זה ייתפס על-ידי הציבור כמידע אמיתי על אודות אדם.

42. על מנת לצמצם את הסיכון לשימוש לרעה בטכנולוגיה ולפגיעה בפרטיות, **מומלץ כי על התוכן הדיגיטלי המזויף שנוצר על-ידי המערכות ושמועבר למשתמשים יוטמע סימן (כגון 'סימן מים' דיגיטלי גלוי),**[24] **המבהיר לצופה בו כי התוכן עבר עריכה דיגיטלית, וכי הוא אינו אותנטי**. רצוי כי מבחינה טכנולוגית סימן זה יהיה מוטמע באופן שיקשה על הסרתו מהתמונה או מקטע הווידאו.

43. **מומלץ כי בעת השימוש במערכות תוצג למשתמשים אזהרה שלא להשתמש בתוכן הדיגיטלי המזויף לרעה**, ובעיקר שלא להפיץ ולפרסם תמונות וסרטונים שעברו מניפולציה ושעלולים להשפיל אדם, להביכו, או להביא לפרסום של עניין שייתפס על-ידי הציבור כנוגע לצנעת חייו האישיים (עברו המיני, מצבו הבריאותי או התנהגותו ברשות היחיד).

<u>הבהרות והמלצות לגורמים המפרסמים תכנים דיגיטליים</u>

44. סעיפים 2(4), 2(6) ו- 2(11) לחוק הגנת הפרטיות חלים כמובן גם על גופים ציבורים ופרטיים שבמסגרת פעילותם מפרסמים תכנים דיגיטליים שונים, ובכלל זה אתרי חדשות.

---

[22] להרחבה ראו טיוטת מסמך המדיניות של הרשות להגנת הפרטיות בעניין צמצום מידע (Data Minimization).

[23] להרחבה ראו אתר הרשות להגנת הפרטיות.

[24] סימן מים דיגיטלי הוא אמצעי זיהוי מסוג טקסט או תמונה, הניתן לשיבוץ ולהטמעה בתמונה או בקטע וידאו, ואשר נועד, בין היתר, למנוע זיופים ולאמת אותנטיות של תכנים.

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב

45. במצב בו יש חשש כי פרסום תוכן דיגיטלי עלול להוות פגיעה בפרטיות, על הגורם המפרסם לקבל את הסכמת האנשים המתועדים בו. **במצב בו יש סבירות גבוהה כי הפרסום אינו אותנטי אלא מזויף – מומלץ שלא להסתפק בהסכמה מכללא, אלא לקבל הסכמה מפורשת לפרסום.**

46. לשם צמצום הסיכון לפגיעה בפרטיות כתוצאה מפרסום תוכן דיגיטלי מזויף, מומלץ כי גופים המפרסמים ינקטו באמצעים שונים לצורך בחינת מידת האותנטיות של התכנים הדיגיטליים המפורסמים על-ידם.

במצב בו יש חשש כי תוכן מסוים הוא מזויף, מומלץ לעשות שימוש בכלים טכנולוגיים שנועדו לבחינה האם תוכן מסוים עבר עריכה דיגיטלית (Deepfake detection tools), לרבות תוך שימוש בכלי בינה מלאכותית.[25]

47. כמו כן בעת פרסום תוכן דיגיטלי אשר עבר עריכה ומניפולציה דיגיטלית ואשר הוחלט לפרסמו מסיבות לגיטימיות מסוימות (לדוגמה במסגרת כתבה שעוסקת בנושא), מומלץ כי עם הפרסום תוצג הבהרה לפיה התוכן אינו אותנטי.

<u>סיכום</u>

48. תופעת השימוש בטכנולוגיית Deepfake היא תופעה שהולכת ומתרחבת. תופעה זו, על אף יתרונותיה השונים, עלולה להוות סיכון של ממש לפרטיות.

49. כפי שפורט, חוק הגנת הפרטיות אוסר על הפצה ללא הסכמה של תמונה או קטע וידאו אשר יכולים לבזות או להשפיל את האדם המצולם או המוקלט. כך, גם בנוגע להפצה ללא הסכמה של תוכן דיגיטלי המציג אדם תוך התייחסות לצנעת חייו האישיים, או לשימוש בתמונה וקול של אדם למטרות מסחריות.

50. **לגישת הרשות הוראות חוק הגנת הפרטיות אלו חלות גם על פרסום והפצה של תוכן שזויף באופן דיגיטלי, אם יש בו כדי להטעות את הציבור ולגרום לו להניח כי מדובר בתוכן אותנטי.** לעיל פורטו שורה של הבהרות והמלצות שמטרתן צמצום אפשרות הפרסום וההפצה של תוכן דיגיטלי מזויף העלול להביא לפגיעה בפרטיות.

51. כמו כן, **הרשות מבהירה כי נתונים הנכללים בהגדרת 'מידע' לפי חוק הגנת הפרטיות, שעברו מניפולציה דיגיטלית באיכות העלולה לגרום לציבור להבין כי מדובר בנתונים אותנטיים על אודות אדם, הם בבחינת 'מידע' שחלות עליו הוראות פרק ב' לחוק.**

---

[25] להרחבה בנוגע לכלים השונים ראו: Shraddha Goled, *Top AI-Based Tools & Techniques For Deepfake Detection*, AIM (Nov.10, 2020). לכתיבה על טכנולוגיות שונות לזיהוי תכנים דיגיטליים שעברו מניפולציה, ראו: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w39/Guarnera_DeepFake_Detection_by_Analyzing_Convolutional_Traces_CVPRW_2020_paper.pdf; https://arxiv.org/pdf/2008.11363.pdf; https://openaccess.thecvf.com/content_ICCVW_2019/papers/HBU/Amerini_Deepfake_Video_Detection_through_Optical_Flow_Based_CNN_ICCVW_2019_paper.pdf.

52. הרשות מציינת גם כי על בעלי שליטה במאגרי מידע ועל מפעילי ומפתחי אפליקציות ותוכנות מסוג Deepfake בפרט, להיות מודעים להוראות חוק הגנת הפרטיות ותקנותיו בהקשרי פרטיות ואבטחת מידע, ולפעול בהתאם להוראות הדין.

53. בנוסף, הרשות ממליצה להטמיע על גבי תוכן דיגיטלי שזויף במערכות של Deepfake סימן דיגיטלי שאינו ניתן להסרה (כגון 'סימן מים' דיגיטלי גלוי), המבהיר לצופה בו כי התוכן עבר עריכה דיגיטלית, וכי הוא אינו אותנטי. הרשות ממליצה גם כי בעת השימוש במערכות תוצג למשתמשים אזהרה שלא להשתמש בתוכן הדיגיטלי המזויף לרעה.

PPA@justice.gov.il | 02-6467064 | 073-3928555
WWW.PPA.JUSTICE.GOV.IL | 6107202 קרית הממשלה, ת.ד. 7360, תל אביב