

הקדמה

המונח דיפ־פייק (deepfake) הוא הלחם המושגים למידה עמוקה (deep learning) וזיוף (fake). הוא מתאר יישום טכנולוגי מבוסס בינה מלאכותית, המאפשר לשנות או לעבד תוכן של תמונות או סרטונים כך שקשה ולפעמים אף בלתי אפשרי להבחין בזיוף. זאת בשונה מ'צ'יפ־פייק' – זיוף שנוצר בעריכה פשוטה ולא בזיוף מבוסס בינה מלאכותית (Paris & Donovan, 2019). יישום זה, כמו יישומים רבים אחרים בתחום הבינה המלאכותית, נעשה בעשור האחרון זול ופשוט יחסית לשימוש, וחלק מהיישומים המאפשרים לייצר זמינים לכל אדם עם גישה לרשת האינטרנט.

יכולות דיפ־פייק מאפשרות יצירת סרטונים או קטעי שמע סינתטיים בעלי מראה או שמע מציאותיים. כך אפשר לייצר הלכה למעשה "זיוף ביומטרי" הנסמך לרוב על תווי הפנים או הקול, ולהציג מצג שווא כאילו אדם ביצע פעולות או אמר דברים, ולא היא. השיטה הנפוצה ביותר ליצירת דיפ־פייק היא בתחום הווידאו (Sample, 2020). מספר הסרטונים המזויפים המופצים ברשת גדל באופן מעריכי (אקספוננציאלי) מאז 2018 ומכפיל עצמו כל חצי שנה. בדצמבר 2020 המספר כבר עלה על 85 אלף סרטונים (Sensity, 2021). היכולת לייצר תוכני שמע (אודיו) או וידאו סינתטיים בעלי מראה מציאותי באה לידי ביטוי בשלל יישומים חיוביים ממגוון תחומי חיים: החל בקולנוע וטלוויזיה דרך רפואה, פסיכולוגיה ועד אומנות, תרבות ושימור היסטורי. כך למשל, יכולות דיפ־פייק מצליחות לתת לחולי ALS קול, עצמאות ויכולת השפעה, "להשיב לחיים" דמויות היסטוריות במוזיאונים ואפילו לשפר את הרפואה באמצעות "חולים סינתטיים" (Jaiman, 2020). אולם חרף הפוטנציאל החיובי הטמון בתחום, הוא מביא עימו גם יישומים מסוכנים ומטרידים המתבססים על אותן יכולות ממש. בין היישומים הללו אפשר למנות למשל הונאות פליליות שעלולות לפגוע בפרטים ובארגונים, או יצירת פורנוגרפיה בהשתתפות לכאורה של דמויות מזויפות שלא נתנו לכך את הסכמתן, לעיתים אף למטרות סחיטה או שימושים פוגעניים אחרים. במקביל לאלו ניצבת סכנת פגיעה בחברה האזרחית או במדינות על ידי יצירת מצגים שקריים שתכליתם עיצוב תודעה, הלכי רוח ודעת קהל, על מנת להשפיע על המערכת הפוליטית, על מערכות בחירות ואף על יציבותן של מדינות (Sayler & Harris, 2021). הזיוף העמוק הוא למעשה יישום טכנולוגי המעצים את אתגרי התקופה הקיימים בתחום הביטחון הלאומי בעידן של פוסט־אמת, כאשר סרטונים מזויפים עלולים להטעות את דרג מקבלי ההחלטות ואולי אף להניע לפעולה, בהתאם לאינטרסים של יוצריהם (Schiff, 2019).

משום כך, השימוש ביכולות דיפ־פייק סומן על ידי ארגוני המודיעין האמריקאיים בשנת 2019 כאיום האסטרטגי הגדול ביותר על הביטחון הלאומי (Konkel, 2019; Ng, 2019). לא רק ארגוני מודיעין חוששים מהפצתו של תוכן מזויף. בסקר דעת קהל בינלאומי משנת 2020 טענו 66 אחוזים מן הנסקרים כי הם מוטרדים מכך שבשל הטכנולוגיה אי אפשר לדעת אם מה שאנשים רואים ושומעים הוא אמיתי (2020 Edelman Trust Barometer). המשמעות העיקרית של יכולות הזיוף העמוק היא אם כן ערעור על התוקף של האמת האובייקטיבית ובה בעת יצירת חשד בכל פיסת תוכן כשקרי, שכן אפשר לטעון כי היא מזויפת גם כאשר

היא אמיתית, ובכך להתנער מאחריות (אורפז, 2020). ליכולת זו עלולה להיות השפעה הרסנית על חברות בכלל ועל דמוקרטיות בפרט.

מחקר זה, **שבוצע במסגרת תוכנית ליפקין־שחק לפוסט אמת ופייק ניוז במכון למחקרי ביטחון לאומי**, סוקר ומציג את הטכנולוגיה המשמשת ליצירת זיופים עמוקים ואת מגוון השימושים החיוביים והשליליים המוכרים שלה, תוך דיון במקרי מבחן בולטים מן העולם. לאחר מכן ייבחנו ההתפתחויות בתחום ההיערכות וההתמודדות עם האתגר בעולם ובישראל. לבסוף, באמצעות מסקנות שעלו מסימולציה שנערכה במכון בהשתתפות מומחים, יוצגו המלצות למדיניות רצויה בתחום עבור ישראל על מנת להיערך כהלכה לקראת איום ההולך ומתגבש.

מחקר זה לא היה מתאפשר ללא עזרתם האדיבה של רבים, ועל כך תודתנו. לתת־אלוף (מיל") איתי ברון, סגן ראש המכון למחקר וראש תוכנית ליפקין־שחק במכון למחקרי ביטחון לאומי (INSS), על שזים את המחקר ועזר לעצב אותו ולצקת בו תוכן. לגב' ענבל אורפז – חוקרת בתוכנית ליפקין־שחק של המכון על סיוע בתובנות ובכל היבט נדרש אחר לאורך המחקר כולו. תודתנו הגדולה גם לגב' יובל קנפו, מתמחה בתוכנית טכנולוגיות מתקדמות וביטחון לאומי, שסייעה רבות בסקירה ובגיבוש של חומרים למזכר. אנו מבקשות להודות גם לכל חברי ועדת המומחים (הרשימה המלאה מופיעה בנספח א') שכונסה לצורך המחקר, אשר תרמו מידיעותיהם, מזמנם וממרחם למחקר זה; לד"ר ענת קורץ, עמיתת מחקר בכירה ומנהלת המחקר במכון למחקרי ביטחון לאומי, ולד"ר גליה לינדנשטראוס, עמיתת מחקר בכירה במכון למחקרי ביטחון לאומי, שעמלו על עריכתו ושיפורו של חיבור זה. לגב' מירה ילין על העריכה הלשונית, למר יגאל טליאנסקי על העיצוב הגרפי ולגב' נעם רן, מנהלת הפרסומים של המכון למחקרי ביטחון לאומי על הניצוח על האופרציה.

רבים לקחו חלק וסייעו לגיבושו ופרסומו של מחקר זה, אולם האחריות לכל טענה או טעות המופיעה כאן היא של המחברות בלבד.