9 בנובמבר 2020

Deepfake: מקרה בוחן להשפעת פייק ניוז על הביטחון הלאומי

ענבל אורפז

Deepfake הוא הלחם המילים *Deep Learning*, טכנולוגיות למידה עמוקה מתחום הבינה המלאכותית, ו-*Fake* (זיוף). טכנולוגיות *Deepfake*, המאפשרות יצירת וידאו או אודיו סינתטיים, למשל סרטונים שבהם נראים מנהיגים, הוכרזו ב-2019 על ידי ארגוני המודיעין האמריקאים כאיום אסטרטגי על הביטחון הלאומי. משמעות עיקרית של יכולות ה-*Deepfake* היא ערעור האמת וגרימת חשד בכל פיסת תוכן כשקרית: ניתן לטעון שהקלטה היא שקרית והופקה באמצעות *Deepfake* וכך להתנער מאחריות. הפלטפורמות החברתיות, ובראשן פייסבוק וטוויטר, ממלאות תפקיד מרכזי בהפצת *Deepfake*, ולקראת הבחירות לנשיאות בארצות הברית, שיתקיימו בנובמבר 2020, הרשתות נדרשו לקבוע מדיניות בתחום ואף לראשונה הסירו פוסטים שקריים של הנשיא טראמפ. עם זאת, למרות שהטכנולוגיה הגיעה לבשלות, האיום לא התממש וטרם התרחש אירוע בקנה מידה שהשפיע על דעת קהל או מקבלי ההחלטות. יתכן והסיבה היא שניתן לייצר קמפיינים של תודעה מבוססי שקרים באמצעים פשוטים יותר ואפקטיביים לא פחות.

בשנים האחרונות סומן השימוש ביכולות Deepfake, המאפשרות יצירה של וידאו או אודיו סינתטי בעל מראה מציאותי, כאיום על הביטחון הלאומי. זאת, בשל הפוטנציאל של מצגים שקריים להשפיע על דעת הקהל. בין היתר, ראשי ארגוני המודיעין האמריקאי התייחסו בשימוע השנתי בנושא האיומים העולמיים שנערך בפני ועדת המודיעין של הסנאט בראשית 2019 ל-Deepfake, לצד פיתוחים טכנולוגיים נוספים וביניהם קמפיינים ברשתות החברתיות שתכליתם להשפיע על התודעה, כאיום הגדול ביותר על הביטחון הלאומי. עם זאת, למרות הפופולריות שלה זוכה הנושא בשיח הציבורי ובמדורי הטכנולוגיה, נראה שעד כה מדובר באיום על הדמוקרטיה שלא התממש על אף שהטכנולוגיה בשלה ונגישה.

במאמר זה נבחן פוטנציאל הסכנה לדמוקרטיות המערביות הגלום בשימוש טכנולוגיות Deepfake במסגרת מאמצי השפעה על דעת קהל ומקבלי החלטות, ותידון השאלה מדוע עד כה לא נעשה בהן שימוש נרחב. אחת ההשערות היא שניתן להשפיע על הציבור ודרגי מקבלי ההחלטות על ידי הפצת פייק ניוז ושקרים באמצעים פשוטים יותר, ולכן אין צורך בטכנולוגיות מתקדמות ומורכבות לשם כך. עם זאת, נטען כי עצם קיום הטכנולוגיה בארסנל הכלים של משבשי התודעה מערער על תפיסת האמת ומעלה חשד כי כל פיסת תוכן המופצת עשויה להיות שקרית, ובכלל זאת הקלטות וסרטונים של פוליטיקאים. כלומר, האיום שבטכנולוגיה אינו בהכרח השימוש בה, אלא עצם הידיעה שהיא קיימת וישימה, וכן בהשפעה הפוטנציאלית שלה על השיח. בנוסף, נתייחס לסוגיית האחריות של הפלטפורמות הטכנולוגיות בהפצת Deepfake ובמאמצי תודעה.

צמיחת ה-Deepfake כאיום על הביטחון הלאומי

Deepfake הוא הלחם המילים Deep Learning, טכנולוגיות למידה עמוקה מתחום הבינה המלאכותית, והמילה Fake (זיוף). על פי הגדרת מילון מריאם וובסטר (Merriam-Webster) המונח מתייחס לרוב לווידאו או אודיו שנערך באמצעות אלגוריתם המשמש להחלפת פני אדם בסרטון המקורי במישהו אחר, לרוב דמות ציבורית מוכרת, או החלפת תוכן הדברים הנאמרים כך שהתוכן הערוך נראה אותנטי. באופן זה גורמים בווידאו שנראה אמיתי לדמות, למשל מנהיג של מדינה, להתבטא או לעשות דברים, שלא באמת נאמרו ונעשו. המונח "Deepfake" נטבע ב-2017 על ידי משתמש בקהילת המפתחים Reddit, שהשתמש בטכנולוגיות בתחום הלמידה העמוקה (Deep Learning) כדי להטמיע פרצופים של סלברטיאיות על גבי צילומי שחקנים בסרטונים פורנוגרפיים.

לפי דו"ח של חברת Deeptrace, המאפשרת זיהוי של סרטוני Deepfake, אשר פורסם באוקטובר 2019 (Patrini, G., 2019) בשבעת החודשים שקדמו לפרסומו מספר סרטוני ה-Deepfake בעולם כמעט שהוכפלה ועלתה למעל ל-14 אלף סרטונים. מתוכם 96 אחוזים הם פורנוגרפיים. אולם, בשוליים מתחילים להתפתח גם שימושים אחרים במדיה סינתטית, אשר לחלקם יש השלכות בתחום הביטחון הלאומי והפוליטיקה. לפי הדו"ח, מבין סרטוני ה-Deepfake שאינם פורנוגרפיים ואשר שותפו באתר יוטיוב, 12 אחוזים מהדמויות שהופיעו בהם היו של פוליטיקאים, אולם 81 אחוזים מסרטוני הווידאו היו זיופים של דמויות מעולם הבידור.

הצמיחה בקצב יצירת סרטוני ה-Deepfake מתאפשרת הודות לפיתוחים טכנולוגיים ושיפור בכוח העיבוד של מחשבים וביכולות אלגוריתמיות, וכבר כיום אפליקציות מובייל מאפשרות למשתמשי קצה לייצר בקלות ובלי השקעת משאבים סרטונים מזויפים. דו"ח (Samsung NEXT Ventures, 2020) שפרסמה באוגוסט 2020 Samsung NEXT מיפה 158 חברות בעולם שעוסקות ביצירת מדיה סינתטית ממגוון סוגים, בהם וידאו, קול ויצירת אוואטרים. 118 מהחברות הן סטארטאפים שמרוכזים בעיקר בארצות הברית, בריטניה וישראל - שבה זוהו עשרה סטארטאפים בתחום.

בשימוע השנתי בנושא האיומים העולמיים בפני ועדת המודיעין של הסנאט האמריקאי, שהתקיים בראשית 2019, ציינו ראשי ארגוני המודיעין האמריקאים את ה-Deepfake, לצד פיתוחים טכנולוגיים נוספים וביניהם קמפיינים להשפעה על התודעה ברשתות חברתיות, כאיום הגדול ביותר על הביטחון הלאומי (Ng, A., 2019 and Konkel, F., 2019). ביוני של אותה השנה נערך שימוע ייעודי בוועדת המודיעין של הסנאט בנושא איומי ה-Deepfake על מערכת הבחירות האמריקאית בפרט, ועל ביטחון לאומי ככלל (Tillett, E., & Gazis, O., 2019).

ה-Deepfake מצטרף אם כן לשורת כלים, המאפיינים את עידן הפוסט אמת והפייק ניוז ומשמשים ליצירת שקרים, שביכולתם להשפיע על דעת הציבור ומקבלי ההחלטות ושיש להם פוטנציאל נזק על הביטחון הלאומי. מאפייני עידן הפוסט אמת הם קושי הולך וגובר לברר את

המציאות, להבינה ולקבל על בסיסה החלטות נכונות גם בתחום הביטחון הלאומי (ברון ורויטמן, 2019). אמנם, מאז ומתמיד קיים קושי בבירור המציאות והבנתה, המשפיע על תהליכי קבלת ההחלטות. אך בעת הנוכחית, חלק מהבעיות המוכרות התעצמו בשל מהפיכת המידע ובעיקר עקב התופעה של התפוצצות המידע - כמויות אדירות של נתונים, מידע וידע, המגיעים בקצבים מהירים מאוד ובמבנים שונים, ובנוסף כלים המאפשרים יצירה והפצה של שקרים בקלות יחסית, דוגמת רשתות חברתיות. אתגרי התקופה בולטים על רקע החרפתם של מאבקים פוליטיים ואידאולוגיים בין קבוצות מתחרות - ימין ושמאל, שמרנים וליברלים, ועלייה לשלטון של מנהיגים פופוליסטיים, הרוחשים אמון מועט לדרגים מקצועיים בתחומים שונים.

ה-Deepfake מייצג יכולת טכנולוגית נוספת, המעצימה את הבעיה ומדגישה את אתגרי התקופה בתחום הביטחון הלאומי. סרטוני Deepfake מזויפים יכולים לייצר מצג שווא שעלול הן להטעות את דרג מקבלי ההחלטות, למשל, דרך סרטון מזויף של פעיל טרור או אירוע ביטחוני שיגרום להחלטות שגויות, והן להשפיע על דעת הציבור באזורים שיש להם חשיבות לביטחון הלאומי, למשל, על ידי הפצת מידע שגוי בתחום הבריאות או הטעייה בנוגע לנבחרי ציבור. יצירת והפצת סרטוני ה-Deepfake יכולה להיעשות על ידי גורמים שונים, מקומיים או חיצוניים למדינה, שיש להם עניין להשפיע על השיח.

ביולי 2019 שיגר אדם שיף, יו"ר וועדת המודיעין של בית הנבחרים האמריקאי, מכתבים למנהלי הפלטפורמות החברתיות הגדולות בעולם - פייסבוק, גוגל וטוויטר - בנוגע למדיניות הפצת Deepfake לקראת הבחירות לנשיאות ב-2020 והיערכותן בנושא (Congressman Schiff, 2019). "כשאנחנו מסתכלים קדימה לקראת בחירות 2020, אני מודאג מאוד שייתכן שהניסיון מ-2016 היה רק הפרולוג", כתב שיף בהתייחסו לניסיון הרוסי להשפיע על בוחרים אמריקאים בבחירות לנשיאות ארצות הברית שהתקיימו באותה שנה, באמצעות שימוש ברשתות חברתיות, והוסיף: "מדיה שעברה מניפולציה או תוכן מטעה הם חלק קבוע מהתוכן באינטרנט, אבל לטכנולוגיית Deepfake יש פוטנציאל להחריף את הבעיה... ההשלכות על הדמוקרטיה שלנו עשויות להיות הרסניות: סרטון Deepfake משכנע ומתוזמן היטב של מועמד שיהפוך ויראלי בפלטפורמה כמו טוויטר יכול להטות את הכף במרוץ ולשנות את מהלך ההיסטוריה". שיף דרש לברר כיצד נערכות הפלטפורמות מבחינת המדיניות שלהן בנוגע להפצת Deepfake, האם הן מתכוונות להסיר או לחסום תכני Deepfake והאם הן חוקרות טכנולוגיות המיועדות לזהות Deepfake.

הרקע לשליחת המכתבים היה, בין היתר, סרטון ערוך ששותף ברשתות החברתיות במאי 2019 אשר בו הופיעה ננסי פלוסי, יו"ר בית הנבחרים האמריקאי (CBS News, May 2019). סרטון זה הפך מקרה בוחן לסוגיות הקשורות לסרטונים מזויפים וערוכים, להשפעתם על הפוליטיקה ולמעורבות הפלטפורמות הטכנולוגיות בשיח הפוליטי. סרטון זה אפילו לא שילב טכנולוגיות Deepfake וסך הכול היה גרסה מואטת של וידאו אמיתי, שפלוסי נראתה בו בעקבות העריכה דמנטית או שיכורה. פייסבוק סירבה להסיר את הסרטון בטענה שהוא אינו מפר את כללי המדיניות של החברה. לעומת זאת, פלטפורמת שיתוף הווידאו יוטיוב, שנמצאת בבעלות גוגל,

נקטה עמדה שונה והודיעה שהיא מסירה את הסרטון בנימוק שהוא מפר את מדיניות האתר (Kelly, M., 2019).

למרות שהסרטון הערוך של פלוסי היה נטול תחכום טכנולוגי ונוצר באמצעים פרימיטיביים (שזכו לתואר Cheap Fake), הוא הפך אבן בוחן לגבי להשפעת סרטונים מזויפים על השיח ועל הביטחון הלאומי. בחודשים הבאים, סוגיית ה-Deepfake והאופן שבו מתמודדות איתה הפלטפורמות הטכנולוגיות המשיך להעסיק את מערכת הבחירות האמריקאית. בסוף חודש ינואר 2020 פרסמה ועדת האתיקה של בית הנבחרים מזכר (U.S. House of Representative, 2020) שלפיו חברי הבית ועובדיו שישתפו תכני Deepfake או קטעים אודיו-ויזואלים מעוותים, שמטרתם להטעות את הציבור, עשויים להפר את הקוד האתי של הבית. במזכר נכתב שמניפולציות בווידאו ותמונות שמטרתן להטעות את הציבור יכולות להזיק לשיח הציבורי ולפגוע באמינות בית הנבחרים, ולכן לפני שיתוף של תמונות, וידאו או אודיו ברשתות החברתיות ובערוצי תקשורת אלקטרונית, מצופה מחברי הבית וצוותיהם לנקוט מאמצים סבירים כדי לברר האם מדובר ב- Deepfake או שהם נוצרו בכוונה להטעות את הציבור.

שימושים ב-Deepfake בהקשר של ביטחון לאומי

ה-Deepfake אמנם הוכרז כאיום על הביטחון הלאומי, אך הדוגמאות לשימושים שכבר נעשו בו מצביעות על כי עד כה, למעשה, לא התרחש אירוע חמור שמימש את מלוא פוטנציאל האיום. אחד הסרטונים המוקדמים שהמחיש את הרלבנטיות של Deepfake לענייני ביטחון לאומי היה וידאו שפורסם ב-2018 באתר BuzzFeed, המדמה את נשיא ארצות הברית לשעבר, ברק אובמה (Mack, D., 2018). בסרטון נראה אובמה אומר "אנחנו נכנסים לעידן שבו האויבים שלנו יכולים לגרום לכך שזה יראה כאילו מישהו אומר דבר מסוים בכל נקודת זמן - אפילו אם הם לעולם לא היו אומרים את הדברים הללו". בהמשך הדברים, הווידאו חושף שלא מדובר באובמה, אלא ב-Deepfake, שנעשה באמצעות שימוש בצילום של הקולנוען ג'ורדן פיל. הסרטון מסתיים באזהרה מפיו של בן דמותו של אובמה: "זו תקופה מסוכנת. עם הפנים קדימה, עלינו להיות ערניים יותר לגבי הדברים באינטרנט שעליהם אנו סומכים. זה זמן שבו עלינו להסתמך על מקורות חדשותיים מהימנים". האזהרה מפיו של אובמה תקפה בעידן ה-Deepfake עבור הציבור הרחב שנחשף לתכנים הללו ועשוי להיות מושפע מהם, וכן לדרג מקבלי ההחלטות.

עם זאת, מספר אירועים שנחשד כי נעשה בהם שימוש ב-Deepfake המחישו את פוטנציאל ההשפעה על הביטחון הלאומי. במקרה אחד שימוש בווידאו שנחשד כמזויף, כמעט הוביל להפיכה צבאית במדינת גבון. נשיא המדינה, עלי בונגו, לא ערך הופעות ציבוריות במשך תקופה ממושכת לאחר שקיבל טיפול רפואי בערב הסעודית באוקטובר 2018 ולאחר מכן במרוקו. בגבון נפוצו שמועות וספקולציות בנוגע למצבו הבריאותי ואף נטען שהוא מת. בסופו של דבר, בערב השנה החדשה נשא בונגו נאום לאומה ששודר בערוצי המדיה החברתית באינטרנט (Gabon 24, 2018). בסרטון נראה בונגו קפוא בתנועתו, מעט מכני, עם תנועות עיניים מוגבלות. ההופעה גרמה לכך

שעלו חשדות שמדובר ב-Deepfake, על אף שבאותו זמן כבר היה ידוע שבונגו עבר שבץ. בחלוף שבוע בוצע במדינה ניסיון הפיכה צבאי שכשל. עד כה לא נמצאו הוכחות שהווידאו אכן היה מזויף (Breland, A., 2019).

במקרה אחר, במלזיה בשנת 2019, הופץ סרטון מין הומוסקסואלי שבו הופיע השר המקומי לענייני כלכלה. הסרטון עורר שערורייה במדינה, שבה קיום יחסים הומוסקסואליים אסור על פי חוק. תומכי הפוליטיקאי, כולל ראש הממשלה, טענו כי מדובר בסרטון Deepfake שמטרתו לחסל את הקריירה הפוליטית של השר. זאת, על אף שגבר אחר המופיע בסרטון טען כי הוא אמיתי (Lin, J., 2019). גם במקרה זה, לא נמצאו הוכחות לכך שמדובר בסרטון מזויף.

המקרים הללו ממחישים בעיה נוספת שאופיינית לעידן הפוסט אמת - לא רק שהטכנולוגיות גורמות לתוכן מזויף להראות אמיתי, אלא נוצר מצב שבו אמת נתפסת כשקר. במילים אחרות, בעידן הנוכחי תיעוד אינו עוד ראיה המוכיחה שמשהו הוא אמיתי, וה-Deepfake גורם פקפוק בעדויות מוחשיות באשר הן, כגון סרטון וידאו או שיחת טלפון מאדם שמדבר בקולו של אדם מוכר. כלומר, עצם הצגת הטכנולוגיה מאפשרת לשקרנים להכחיש את אמיתותן של עובדות וראיות אמיתיות. זו אולי אחת ההשפעות המרכזיות של ה-Deepfake - לא עצם השימוש בו, אלא הערעור והחתירה נגד מושג האמת. מכאן התהייה, מה למשל היו ההשלכות של פרסום הקלטות מפי הנשיא טראמפ, שהשפיעו על התגלגלות אירועים היסטוריים, לצד טענה שהן Deepfake.

גם בישראל נעשה שימוש בסרטוני וידאו הערוכים באופן מניפולטיבי, שנועדו להשפיע על דעת הקהל. בדומה למצב בארצות הברית, גם במקרים הללו בישראל נעשה שימוש בעריכה פשוטה בלי טכנולוגיות Deepfake. אך גם מקרים אלה חושפים את קצה הקרחון של היכולות לייצר מאמצי השפעה. דוגמה לכך היא סרטון שהפיץ ברשתות החברתיות ראש הממשלה בנימין נתניהו בצהרי יום הבחירות בישראל במארס 2020. בסרטון נשמע בני גנץ, מועמד 55״כחול לבן״ לראשות הממשלה, אומר: ״לא לוותר, עד הרגע האחרון, לא נשים ׳כחול לבן׳ בקלפיות״ (המשרוקית; גלובס, 2020). בסרטון ששיתף נתניהו הושמט המשך המשפט שהופיע בהקלטה המקורית: ״לא נשים ׳כחול לבן׳ בקלפיות - נקבל בחירות רביעיות״. מפלגת הליכוד נקנסה בשל כך על ידי ועדת הבחירות ב-7,500 ש״ח, ונדרש להסירו (שחק, ע. 2020), אך מספר כלי תקשורת הספיקו להדהד את המסר והציגוהו כ״בלבול״ של גנץ.

דוגמה נוספת מישראל ממחישה את פוטנציאל הנזק שעלול לגרום שימוש בטכנולוגיות Deepfake באופן שמסכן את בריאות הציבור. במהלך התפשטות מגפת הקורונה בחודש מארס 2020 הופצה בקבוצות וואטסאפ הקלטה ובה ביקורת על התנהלות מנכ״ל משרד הבריאות משה בר סימן טוב וראש הממשלה נתניהו, ונשמע כביכול קולו של פרופ׳ יצחק קרייס, מנהל המרכז הרפואי שיבא (MedNew, 2020). דוברת שיבא מסרה בתגובה להפצת ההקלטה שמדובר בפייק ניוז וקראה לציבור לא להפיץ ידיעות מוטעות ומטעות. גם בעתיד עלול להיעשות שימוש בקולותיהם של בכירים במערכת הבריאות באמצעות יכולות Deepfake על מנת לזייף הקלטות עם הנחיות בריאותיות שיש בהן כדי לגרום לנזק ניכר לבריאות הציבור.

אחריות הפלטפורמות החברתיות בהפצת Deepfake

הפלטפורמות החברתיות הן ערוצי ההפצה העיקריים של סרטונים לציבור הרחב - והן יכולות להכריע אילו מהם יופצו ואילו לא, ואיזה מידע יגיע לידיעת הכלל. על כן ניתן לטעון כי הפלטפורמות הטכנולוגיות הפכו ל-Policy Makers או צנזורים, המחליטים אילו תכנים ישותפו ואילו לא. מהבחינה הזו, הפלטפורמות הטכנולוגיות, שלמיליארדי אנשים ברחבי העולם יש גישה אליהן, הפכו גורם מרכזי בהשפעה על התודעה הציבורית, ובתוך כך גם על הביטחון הלאומי. ההחלטה שלהן לאפשר שיתוף או אי-שיתוף של תכנים שקריים או מוטעים, דוגמת Deepfake, מהווה התערבות חיצונית והשפעה זרה על השיח המקומי בזירה הפוליטית

ההתנהלות של הפלטפורמות עד כה מוכיחה שהן דואגות לאינטרסים של עצמן ולאו דווקא רואות חשיבות בהגנה על הדמוקרטיה או הביטחון הלאומי. כך למשל, בתגובה להחלטה של פייסבוק לא להסיר את הסרטון של פלוסי, צמד אמנים בריטיים הפיצו באינסטגרם, רשת חברתית הנמצאת בבעלות פייסבוק, סרטון ובו נראה מנכ״ל החברה, מארק צוקרברג, מדבר בגנות חברה הצוברת בידיה כמויות גדולות ורגישות של מידע על משתמשים. מטרת הסרטון, שנוצר באמצעות טכנולוגיה של חברת CannyAI הישראלית, הייתה לבחון את מדיניותה של פייסבוק בנוגע ל-Deepfake במקרה שנעשה בו שימוש כדי לבקר את החברה עצמה (Cole, S., 2019). סרטון זה המחיש את הקונפליקט שבו נמצאת חברת פייסבוק עצמה בהקשר זה.

כמענה להתמודדות עם בעיית ה-Deepfake והשפעתו על השיח, הובילה פייסבוק במהלך 2020 אתגר, כלומר תחרות, לזיהוי סרטוני Deepfake, בשיתוף חברות הטכנולוגיה אמזון ומיקרוסופט. במסגרת האתגר, שהשתתפו בו מעל 2000 אנשים מעולם הטכנולוגיה והאקדמיה, פייסבוק סיפקה למשתתפים 100 אלף סרטוני Deepfake שנוצרו על ידה בהשתתפות שחקנים. המשתתפים התחרו על פרסים בשווי כולל של מיליון דולר ונוצרו במסגרתה מעל ל-35 אלף מודלים לזיהוי Deepfake. עם זאת, שיעור הדיוק בזיהוי זיופים על ידי האלגוריתמים נמצא מאכזב: המודל המנצח בתחרות הצליח לזהות כ-65 אחוזים בלבד מסרטוני ה-Deepfake שנבחנו (Knight, W., 2020). רמת דיוק זו אינה מספקת כאמצעי לזיהוי אוטומטי של סרטונים סינתטיים. כלומר, גם אם הפלטפורמות יצרו להטמיע יכולות זיהוי אוטומטיות, עדיין אין ברשותן אלגוריתמים שמאפשרים זאת. לאחר סיום התחרות, סמנכ״ל הטכנולוגיות של פייסבוק, מייק שרופפר, אמר שהחברה עצמה מפתחת אלגוריתם כזה ושתוצאות התחרות ישמשו כאבן דרך לחוקרים בתחום (Vincent, J., 2020). כן ציין כי Deepfake לא מהווה בעיה משמעותית כעת, ועדיין החברה מבקשת להיות ערוכה לזהות תכנים כאלו בעתיד.

ככל שהפלטפורמות יהפכו לצנזור שאחראי להחליט אילו תכני Deepfake יופצו ואילו לא, הן יעמדו בפני מערך חדש של דילמות ויידרשו, למשל, להכריע בין מצבים שבהם שותף סרטון Deepfake לצרכי בידור לגיטימיים לבין סרטונים שנועדו לקדם מאמצי תודעה של מדינות וארגונים המעוניינים להפיץ דיסאינפורמציה. כך לדוגמה, פייסבוק תידרש להכריע בשאלה האם סרטון של עמית פישר מסטודיו אמברלה (Fisher, A., 2020), שבו מוחלפים פניו של מל ברוקס

בסרט "ההיסטוריה המטורפת של העולם" בזה של בנימין נתניהו, הוא בידורי או הטעייה מכוונת של ציבור.

בפברואר 2020 הופץ סרטון ערוך נוסף שבו הופיעה פלוסי ואשר קודם ושותף ברשתות החברתיות על ידי הנשיא טראמפ (Trump, D. J., 2020). בסרטון הפרובוקטיבי נראית פלוסי קורעת את נאום "מצב האומה" שנשא טראמפ בתגובה לסיפורים של גיבורים אמריקאים שהוזכרו בנאום. העריכה המטעה מובילה את הצופים לחשוב שפלוסי מזלזלת באנשים המופיעים בווידאו. גם במקרה זה, הסרטון היה ערוך באופן מניפולטיבי ולא היה Deepfake. הסרטון פורסם לאחר שהפלטפורמות הטכנולוגיות פרסמו נהלים הנוגעים לשיתוף תכני Deepfake, ובכלל תכנים ערוכים ושקריים, והעמיד לראשונה את המדיניות של החברות למבחן (Frankland & Gorman, 2020). בינואר 2020 הודיעה פייסבוק על צעדי האכיפה שתנקוט כלפי מדיה שעברה מניפולציות (Bickert, M., 2020). בפוסט שפירסמה פייסבוק והתייחס לשימוש ב-Deepfake, נכתב: "בעוד קטעי הווידאו הללו עדיין נדירים באינטרנט, הם מהווים אתגר משמעותי לתעשייה ולחברה שלנו בשעה שהשימוש בהם עולה". פייסבוק הודיעה במסגרת זו שהיא תסיר מדיה מטעה וערוכה, לפי הקריטריונים הבאים: 1. מדיה שעברה עריכה, מעבר לשינויים באיכות, באופן שאינו גלוי לאדם הממוצע ושסביר שיגרום למישהו לחשוב בטעות שמושא הווידאו אמר דברים שלא נאמרו; 2. מדיה שהיא תוצר של בינה מלאכותית או למידת מכונה שמחברת, מחליפה או מוסיפה לווידאו כך שהוא נראה אמיתי. בפייסבוק ציינו שמדיניות זו אינה תקפה לתוכן שהוא פארודי או סאטירי, או וידאו שהעריכה היחידה בו היא שינוי סדר המילים.

בפברואר 2020 הודיעה טוויטר על אודות שורת צעדים שתנקוט לסימון תוכן סינתטי ולחסימת תוכן סינתטי מזיק (Roth, Y., & Achuthan, A., 2020). בפוסט שפרסמה טוויטר צוין כי קרוב ל-90 אחוזים מהמשתמשים ברשת, שהשתתפו בסקר בנוגע למדיניות של טוויטר, השיבו כי הם חושבים שצריך להצמיד תוויות אזהרה לתכנים שעברו שינויים ניכרים. שיעור דומה של משיבים ציין שצריך להזהיר אנשים לפני שהם משתפים מדיה מטעה שעברה שינויים ועריכה. המשתתפים בסקר תמכו פחות בקריאה להסיר לחלוטין מדיה ערוכה ושקרית מהפלטפורמה. טוויטר אפיינה מספר מצבים אפשריים לפרסום מדיה שקרית ומטעה, לפי חומרת הנזק שלהם, ובהתאם קבעה מדיניות לגבי סימונם בתווית מיוחדת או הסרתם: התכנים שיוסרו הם המהווים סכנה לבטיחות של אדם או קבוצה, גורמים סיכון לאלימות נרחבת או מאיימים על הפרטיות או חופש הביטוי של אדם או קבוצה. יתר התכנים יסומנו בתווית מיוחדת בלבד.

עם פרסום ההחלטות של הפלטפורמות החברתיות לגבי מדיה שקרית וערוכה, הובעה כלפיהן ביקורת שלפיה גם לאחר קביעת המדיניות ניתן יהיה להמשיך להפיץ סרטוני Deepfake שקריים ומטעים (Holmes, A., 2020). פרסום הסרטון השני של פלוסי, היה הזדמנות לבחון את יישום קווי המדיניות שפורסמו זמן קצר קודם לכן על ידי הפלטפורמות הטכנולוגיות לגבי הפצת ושיתוף תכנים ערוכים ושקריים שתכליתם הטעייה. על אף הפנייה של פלוסי לטוויטר ולפייסבוק, החברות סירבו להסיר את הסרטון. "העם האמריקאי יודע שהנשיא לא מהסס לשקר להם - אבל זו בושה לראות שטוויטר ופייסבוק, שהם מקורות לחדשות עבור מיליונים, עושות אותו הדבר", צייץ בטוויטר דרו המיל, ראש הסגל של פלוסי (Hammill, D., 2020). "הווידאו האחרון של

יושבת הראש פלוסי תוכנן באופן מכוון כדי להטעות ולשקר לאמריקאים, ובכל יום בו הפלטפורמות מסרבות להסיר אותו, זו עוד תזכורת שאכפת להן יותר מהאינטרסים של בעלי המניות שלהן מאשר אלו של הציבור״. אנדי סטון, דובר חברת פייסבוק, השיב לביקורת וצייץ ״סלח לי, האם אתה רוצה להגיד שהנשיא לא נשא את הדברים הללו ושיושבת הראש לא קרעה את הנאום?״. (Stone, A., 2020)

חילופי הדברים הללו ממחישים שמדובר במאבק בין חברות הטכנולוגיה והפוליטיקאים, שמתנהל במתח שבין חופש הביטוי, הגנה על הדמוקרטיה ורצונן של חברות מסחריות להשיא רווחים לבעלי המניות שלהן. פרסום הווידאו השני של פלוסי, העלה כמה דילמות שהתחדדו בעקבותיו, בהן: האם הפצת סרטונים מזויפים ומטעים היא לגיטימית, כחלק מתעמולת בחירות? מי האחראי לסמן וידאו כערוך או שקרי - הפלטפורמות הטכנולוגיות או המשתפים שלו, כולל יריבים פוליטיים? מה האחריות של גופי התקשורת שממשיכים לתת תהודה לסרטונים המזויפים על ידי שידורם? בנוסף, הסרטון המחיש שלא נדרשות יכולות זיוף מתוחכמות כדי להשפיע על דעת הקהל.

במאי 2020 טוויטר תייגה לראשונה ציוץ של טראמפ עם תווית ״לא מבוסס״ (BBC News, 2020). באוגוסט אותה השנה הסירה פייסבוק לראשונה פוסט של הנשיא, שהכיל ריאיון וידאו איתו ובו הוא ציטט מידע שקרי לגבי חסינות ילדים לנגיף הקורונה (Ghaffary, S., 2020), בטענה להפרת המדיניות שלה בנוגע להפצת דיסאינפורמציה מזיקה על הנגיף. לא הייתה זו הפעם הראשונה שבה טראמפ אמר דברים שקריים או חסרי ביסוס, אך הפעם לדבריו הייתה משמעות לחיי אדם. טוויטר מצידה חסמה את הגישה לרשת לעיתונאים ששיתפו וידאו זה למטרות בדיקת עובדות (Haltiwanger, J., 2020). המקרים הללו ממחישים את התמודדות של הפלטפורמות עם הפצת מידע שקרי שאינו Deepfake ומעלים סוגיות נוספות - מי אחראי על הפצת השקרים? במקרה זה, טראמפ ציטט מידע שגוי בריאיון ששודר ברשת פוקס ולאחר מכן שותף ברשתות החברתיות. היכן יש לעצור את שרשרת הפצת השקרים, ומי ייתן דין וחשבון בגין הפצה?

בינתיים נמשך הדיון הציבורי בנוגע לאחריותן של הפלטפורמות להפצת דיסאינפורמציה. בהצהרת הפתיחה של מרק צוקרברג בשימוע שנערך בקונגרס ביולי 2020 בנושא ההגבלים העסקיים של חברות הטכנולוגיה הגדולות בעולם, הוא הכיר בכך שפייסבוק צריכה להשתפר בתחום (DailyMail, 2020). לדבריו, ״כפלטפורמה לרעיונות תמיד נהיה במרכז דיונים חשובים על חברה וטכנולוגיות״ [..] ״אלה זמנים מאתגרים בצורה יוצאת דופן, ולכן זה חשוב יותר מתמיד שאנשים יוכלו לנהל שיחות בפלטפורמות שלנו על הנושאים שחשובים להם - בין אם זה נגיף הקורונה, חוסר צדק חברתי וגזעי, דאגות כלכליות או משפחתיות, או הבחירות הקרבות. אנחנו מכירים בכך שיש לנו אחריות לעצור שחקנים רעים מהתערבות או חתירה תחת השיח הזה באמצעות מיסאינפורמציה, ניסיונות לדכא מצביעים או שיח שנאה או הסתה. אני מבין את הדאגות של אנשים בתחומים הללו, ואנחנו עובדים כדי לטפל בהן. בעוד אנחנו מתקדמים [..], אני מכיר בכך שאנחנו צריכים לעשות יותר״. ההצהרה של צוקרברג נמסרה מספר שבועות לאחר פרסום דו״ח לסיכום חקירה (audit) חיצונית לפייסבוק, שבו נטען כי החלטות הרשת החברתית אפשרו לשיח שנאה ודיסאינפורמציה לשגשג ושהרשת לא עשתה מספיק כדי להגן על המשתמשים

בה (Isaac, M., 2020; Facebook's Civil Rights Audit, 2020). בשימוע שנערך בקונגרס ביולי נשאל צוקרברג על ידי חבר בית הנבחרים הדמוקרטי, דיוויד סיסלין, יו"ר ועדת בית הנבחרים להגבלים עסקיים: "האם זה לא אומר שהפלטפורמה שלך כל כך גדולה, שאפילו עם המדיניות הנכונה במקום, אינכם יכולים להכיל תוכן קטלני?" (Wise, J., 2020). באוגוסט נשלח מכתב מטעם כ-20 פרקליטי מדינה אמריקאים להנהלת פייסבוק ובו הם ביקשו מהרשת החברתית, בין היתר, לנקוט צעדים נוספים כדי למנוע שימוש בה להפצת דיסאינפורמציה ושנאה ולאפשר אפליה (Letter from 20 state attorneys general to Facebook's Mark Zuckerberg, 2020).

לא רק הפלטפורמות הטכנולוגיות, גם גופי התקשורת המסורתית טרם יצרו פרקטיקות להתמודדות עם זיהוי סרטונים מזויפים והפצתם. ברוב מערכות התקשורת לא קיימים כיום אמצעים לברר האם סרטונים המתקבלים מגורמים חיצוניים הם אמיתיים או מזויפים - אף על פי שבעולם כבר קיימים פתרונות המאפשרים לזהות Deepfake. בנוסף, גופי התקשורת משדרים סרטונים מזויפים - גם במקרים שהם יודעים שמדובר בזיוף - וכך מסייעים להדהוד השקר ולזריעת חרדה והטעייה של הציבור. דוגמה לכך היא סרטון שהופץ על ידי חמאס במאי 2019 ובו נראה ירי טיל נ"ט נגד רכב ישראלי, רגעים ספורים לאחר מעבר רכבת סמוך למקום הפגיעה. ואולם, עקב מתיחות ביטחונית הופסקה באותו היום תנועת הרכבות באזור עוטף עזה (לביא א., 2019). למרות שהסרטון היה שקרי ועבר עריכה מניפולטיבית, הוא שודר בערוצי התקשורת הישראליים וקיבל תהודה (זיני, א., הלר, א., רביד, ב., 2019). סרטונים מסוג זה עלולים להיות חלק ממאמצי השפעה תודעתיים באמצעות יצירת פחד.

האם בהלת ה-Deepfake מוגזמת?

למרות שה-Deepfake הוכרז כאיום על הביטחון הלאומי, עד כה לא התממש האיום, כלומר: לא נרשמה השפעה רחבת היקף של סרטונים מעובדים או מזויפים אשר גרמו פגיעה בביטחון הלאומי. זאת, למרות הבשלות הטכנולוגית המאפשרת את יצירת ה-Deepfake וגם עלויות ההפקה הנמוכות יחסית של הטכנולוגיה. מכאן השאלה: מדוע לא נעשה שימוש נרחב ב-Deepfake במסגרת מאמצי השפעה? גם הדוגמאות שהובאו במסמך זה ממחישות שימוש בסיסי בלבד בעריכה מניפולטיבית, לא שימוש מתוחכם במדיה סינתטית.

אחד ההסברים לכך הוא שקיימות דרכים קלות ונגישות יותר להפצת דיסאינפורמציה ויצירת השפעה על התודעה הציבורית. דוגמה לכך היא קמפיין שעלה לאחרונה בעמוד פייסבוק בשם "האביב הציוני", בו הופיעו פוסטים, לכאורה של תומכים באגף השמאל של הזירה הפוליטית בישראל, אשר "התפכחו" ומאסו בהסתה של השמאל והפכו לתומכי בנימין נתניהו, המוביל את האגף הימני של הזירה (ב"ז א., 2020). הקמפיין שותף גם בעמוד "דף התמיכה בחדשות 20", שלפי כתב העת "העין השביעית" אינו קשור לערוץ 20. אלא שהתמונות של המשתמשים ששיתפו את הפוסטים בקמפיין הן תמונות סינתטיות, שנוצרו על ידי מחשב, ולא של אנשים אמיתיים. השימוש בתמונות הסינתטיות היה פרימיטיבי במיוחד ולאחת המשתמשות, תחת הכינוי "שרון אפשטיין", הוצמדה התמונה המופיעה בערך הויקיפדיה Human image synthesis (Wikipedia, n.d). כך גם נחשף קמפיין רשת של 19 כותבים מזויפים, שפרסמו יותר מ-90 מאמרי דעה שקידמו

אינטרסים של איחוד האמירויות וקראו לגישה נוקשה יותר כלפי קטר, איראן וטורקיה ב-46 גופי תוכן (Rawnsley, A., 2020). במילים אחרות – מדוע יטרחו יוצרי הקמפיינים להפיק סרטונים מורכבים ולהיעזר ברשתות בוטים גדולות, כשאפשר להשתמש בתמונות קיימות ובפרופילים מזויפים בפייסבוק להפצת שקרים?

ה-Deepfake התווסף לארסנל הכלים הקיימים, המאפשרים הפצת שקרים ויצירת מאמצי השפעה. הוא מייצג של רוח התקופה, בה גבר הקושי להבחין בין אמת ושקר, אך ניתן להפיץ באפקטיביות שקרים וליצור השפעה לא פחותה באמצעים פשוטים וזולים יותר. עם זאת, יש השפעה לעצם קיום הטכנולוגיה, ההופך את הציבור לחשדן יותר וגורמת פיקפוק באמינות של כל תוכן שאליו הוא נחשף. כלומר, הטכנולוגיה משבשת את מאזן הכוחות הקיים, שלפיו מה שאנו רואים הוא הקיים, ומערערת את האמינות של כל סרטון או הקלטה.

לפני סיום, נציין כי יש גם ניצנים לשימושים חיוביים ל-Deepfake, למשל לשם הנגשת מידע בשפות וניבים מקומיים. כך למשל, בפברואר הפיץ מועמד לבחירות בדלהי (הודו) שני סרטונים מבוססי Deepfake למיליוני מצביעים פוטנציאליים - אחד באנגלית והשני בניב הינדי מקומי (Christopher, N., 2020). זו דוגמה לשימוש המאפשר ניהול קמפיין פוליטי המגיע למגוון קהלי יעד. יתכן, שההשפעות החיוביות יבואו לידי ביטוי גם בשימושים בתחום הביטחון הלאומי, כדי להעצים אזרחים או להנגיש להם מידע מנבחרי הציבור - ולא רק יהוו איום.

בנוסף, יודגש שקיימות טכנולוגיות המאפשרות לזהות סרטונים מזויפים. למעשה מתנהל מרוץ בין הטכנולוגיות ליצירת Deepfake לבין טכנולוגיות המיועדות לזהותו. אולם, כפי שהראה מחקר קודם (אורפז, ע., 2019), נראה שהטכנולוגיה לא תיתן (גם במקרה זה) מענה לבעיה שאותה היא יצרה. במידה רבה, הפלטפורמות הטכנולוגיות מאיצות את הבעיה ומציבות אתגרים גם בפני מקבלי ההחלטות. לכן, במסגרת מאמצי ההגנה מפני ה-Deepfake, יש חשיבות להכשרות בתחום האוריינות הדיגיטלית, שיספקו לציבור ולמקבלי ההחלטות כלים משופרים יותר להבחין בין מידע אמיתי ושקרי וכן לפתח כלים לעצירת הסרטונים וההקלטות המזויפים עוד לפני שהם הופכים וויראליים ומשפיעים על השיח והביטחון הלאומי.

יתכן שבעתיד נראה אירוע בקנה מידה נרחב של השפעה על דעת קהל או דרג מקבלי ההחלטות על ידי סרטונים מזויפים. אם אכן כך יקרה, יתממשו האזהרות בנוגע לסכנות הטמונות ב-Deepfake מבחינת השפעותיו מרחיקות הלכת על התודעה הציבורית והביטחון הלאומי. עד אז, נמשיך להתייחס אליו כאל יכולת יצירת פייק ניוז משוכללת, ואיום פוטנציאלי.

ביבליוגרפיה

BBC News. (2020, May 27). Twitter tags Trump tweet with fact-checking warning. *BBC News*. Retrieved from https://www.bbc.com/news/technology-52815552

Bickert, M. (2020, January 6). Enforcing Against Manipulated Media. Facebook. Retrieved from https://about.fb.com/news/2020/01/enforcing-against-manipulated-media/

Breland, A. (2019, March 15). The Bizarre and Terrifying Case of the "Deepfake" Video that Helped Bring an African Nation to the Brink. *Mother Jones*. Retrieved from https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/

CBS News. (2019, May 24). Doctored videos of Nancy Pelosi spread online. Retrieved from https://www.youtube.com/watch?time_continue=31&v=utwZgn1o5LY&feature=emb_title

Christopher, N. (2020, February 18). We've Just Seen the First Use of Deepfakes in an Indian Election Campaign. *Vice*. Retrieved from https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp

Cole, S. (2019, June 11). This Deepfake of Mark Zuckerberg Tests Facebook's Fake Video Policies. *Vice*. Retrieved from https://www.vice.com/en_us/article/ywyxex/deepfake-of-mark-zuckerberg-facebook-fake-video-policy

Congressman Adam Schiff. (2019, July 15). Press Release: Schiff Presses Facebook, Google and Twitter for Policies on Deepfakes Ahead of 2020 Election. Retrieved from https://schiff.house.gov/news/press-releases/schiff-presses-facebook-google-and-twitter-for-policies-on-deepfakes-ahead-of-2020-election

DailyMail. (2020, July 29). MARK ZUCKERBERG: FACEBOOK NEEDS TO BE THIS SIZE - WE SHOULD WORRY ABOUT CHINA'S RIVAL INTERNET MORE. *DailyMail*. Retrieved from https://www.dailymail.co.uk/news/fb-8572427/MARK-ZUCKERBERG-STATEMENT-CONGRESS.html

Facebook. (2020, June 25). Deepfake Detection Challenge Dataset. Retrieved from https://ai.facebook.com/datasets/dfdc/

Facebook's Civil Rights Audit – Final Report. (2020, July 8). Retrieved from https://about.fb.com/wp-content/uploads/2020/07/Civil-Rights-Audit-Final-Report.pdf

Fisher, A. (2020, July 16). Retrieved from https://www.facebook.com/amit.fisher1/videos/10156371916682185

Frankland, A. & Gorman, L. (2020, January 13). Combating the Latest Technological Threat to Democracy: A Comparison of Facebook and Twitter's Deepfake Policies. *Alliance for Securing Democracy*. Retrieved from https://securingdemocracy.gmfus.org/combating-the-latest-technological-threat-to-democracy-a-comparison-of-facebooks-and-twitters-deepfake-policies/

Gabon 24. (2018, December 31). DISCOURS À LA NATION DU PRÉSIDENT ALI BONGO ONDIMBA. Retrieved from https://www.facebook.com/watch/?v=324528215059254

Ghaffary, S. (2020, August 6). Facebook took down a Trump post for the first time. *Recode*. Retrieved from https://www.vox.com/recode/2020/8/5/21356348/trump-children-immune-covid-19-facebook-post-taken-down-misinformation-twitter

Haltiwanger, J. (2020, August 6). Journalists are getting locked out of Twitter after fact-checking videos of Trump that push false information on COVID-19. *Business Insider*. Retrieved from https://www.businessinsider.com/journalists-get-locked-out-twitter-after-fact-checking-trump-videos-2020-8

Hammill, D. (2020, February 7). *Twitter*. Retrieved from https://twitter.com/Drew_Hammill/status/1225830200627298305

Holmes, A. (2020, January 7). Facebook just banned deepfakes, but the policy has loopholes — and a widely circulated deepfake of Mark Zuckerberg is allowed to stay up. *Business Insider*. Retrieved from https://www.businessinsider.com/facebook-just-banned-deepfakes-but-the-policy-has-loopholes-2020-1

Isaac, M. (2020, July 8). Facebook's Decisions Were "Setbacks for Civil Rights," Audit Finds. *New York Times*. Retrieved from https://www.nytimes.com/2020/07/08/technology/facebook-civil-rights-audit.html

Kaggle. (2020, March). Deepfake Detection Challenge. Retrieved from https://www.kaggle.com/c/deepfake-detection-challenge/overview/prizes

Kelly, M. (2019, May 24). Distorted Nancy Pelosi videos show platforms aren't ready to fight dirty campaign tricks. *The Verge.* Retrieved from https://www.theverge.com/2019/5/24/18637771/nancy-pelosi-congress-deepfake-video-facebook-twitter-youtube

Knight, W. (2020, June 12). Deepfakes Aren't Very Good. Nor Are the Tools to Detect Them. *Wired*. Retrieved from https://www.wired.com/story/deepfakes-not-very-good-nor-tools-detect/

Konkel, F. (2019, January 29). AI, Deepfakes and the Other Tech Threats That Vex Intel Leaders. *Nextgov*. Retrieved from https://www.nextgov.com/emerging-tech/2019/01/ai-deepfakes-and-other-tech-threats-vex-intel-leaders/154498/

Letter from 20 state attorneys general to Facebook's Mark Zuckerberg. (2020, July 29). *The Washington Post*. Retrieved from https://www.washingtonpost.com/context/letter-from-20-state-attorneys-general-to-facebook-s-mark-zuckerberg/8fa9e319-8a2c-408f-be5d-1766e9404b46/?itid=lk_inline_manual_2

Lin, J. (2019, July 17). Someone just uploaded more videos implicating Azmin in a sex scandal, including a recording of two men agreeing to meet in a hotel. *Business Insider Malaysia*. Retrieved from https://www.businessinsider.my/someone-just-uploaded-more-videos-implicating-azmin-in-a-sex-scandal-including-a-recording-of-two-men-agreeing-to-meet-in-a-hotel

Mack, D. (2018, April 17). This PSA About Fake News From Barack Obama Is Not What It Appears. *BuzzFeed*. Retrieved from https://www.buzzfeednews.com/article/davidmack/obama-fake-news-jordan-peele-psa-video-buzzfeed

Merriam-Webster. (n.d.). Words We're Watching: "Deepfake." Retrieved from https://www.merriam-webster.com/words-at-play/deepfake-slang-definition-examples

Ng, A. (2019, January 29). Deepfakes, disinformation among global threats cited at Senate hearing. *Cnet.* Retrieved from https://www.cnet.com/news/deepfakes-disinformation-among-global-threats-cited-at-senate-hearing/

Patrini, G. (2019, October 7). Mapping the Deepfake Landscape. Retrieved from https://sensity.ai/mapping-the-deepfake-landscape/

Rawnsley, A. (2020, July 7). Right-Wing Media Outlets Duped by a Middle East Propaganda Campaign. *Daily Beast*. Retrieved from https://www.thedailybeast.com/right-wing-media-outlets-duped-by-a-middle-east-pr opaganda-campaign

Roth, Y., & Achuthan, A. (2020, February 4). Building rules in public: Our approach to synthetic & manipulated media. *Twitter*. Retrieved from https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media.html

Samsung NEXT Ventures. (2020, August). Synthetic Media Landscape 2020. Retrieved from https://www.syntheticmedialandscape.com/

Stone, A. (2020, February 7). Retrieved from https://twitter.com/andymstone/status/1225837038190170112

Tillett, E., & Gazis, O. (2019, June 13). House holds hearing on "deepfakes" and artificial intelligence amid national security concerns. *CBS News*. Retrieved from https://www.cbsnews.com/news/house-holds-hearing-on-deepfakes-and-artificial-intelligence-amid-national-security-concerns-live-stream/

Trump, D. J. (2020, February 7). Retrieved from https://twitter.com/realDonaldTrump/status/1225553117929988097?s=20

U.S. House of Representative. (2020, January 28). Intentional Use of Audio-Visual Distortions & Deep Fakes. Retrieved from

https://ethics.house.gov/sites/ethics.house.gov/files/wysiwyg_uploaded/Deep%20Fakes%20Pink%20Sheet%20Guidance-Final.pdf

Vincent, J. (2020, June 12). Facebook contest reveals deepfake detection is still an "unsolved problem." *The Verge*. Retrieved from https://www.theverge.com/21289164/facebook-deepfake-detection-challenge-unsolved-problem-ai

Wikipedia. (n.d.). Human image synthesis. *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Human_image_synthesis

Wise, J. (2020, July 29). Cicilline grills Zuckerberg on coronavirus misinformation: This is "about Facebook"s business model'. *The Hill*. Retrieved from https://thehill.com/policy/technology/509686-cicilline-grills-zuckerberg-on-coronavirus-misinformation-this-is-about

אורפז, ע. (2019). טכנולוגיות נגד פייק ניוז: כשל שוק או הזדמנות מפוספסת? מבט על. (1205), Retrieved from https://www.inss.org.il/he/publication/טכנולוגיות-נגד-פייק-ניוז-כשל-שוק-או-הזד/

ב"ז, א. (2020, אוגוסט 8). קמפיין רשת מפיץ פוסטים שקריים של "שמאלנים שהתפכחו" והפכו ל"ביביסטים." העין השביעית https://www.the7eye.org.il/383761?s=08 . Retrieved from

ברון, א. & רויטמן, מ. (2019). ביטחון לאומי בעידן של פוסט־אמת ופייק ניוז. פרסום מיוחד. Retrieved from https://www.inss.org.il/he/publication/national-security-in-the-era-of-post-truth-and-fake-news/

המשרוקית של גלובס. (2020, מרץ 2). מפלס ההטעיות ביום בחירות היה נמוך במפתיע, ואז נתניהו שוב הפיץ סרטון מסולף של גנץ. גלובס Retrieved from . https://www.globes.co.il/news/article.aspx?did=1001320369

זיני, א., הלר, א., & רביד, ב. (2019, מאי 5). חמאס פרסם תיעוד: ירי טיל הנ"ט לעבר הרכב הישראלי. חדשות 13 Retrieved from . https://13news.co.il/item/news/politics/security/attack-south-236927/

לביא, א. (2019, מאי 6). חמאס הפיץ סרטון פיקטיבי של רגע הריגת משה פדר. זמן ישראל . Retrieved from https://www.zman.co.il/4132/

מתחזה למנכ״ל שיבא. (2020, מרץ 26). *MedNews*. Retrieved from https://mednews.co.il/impersonating-the-ceo-of-sheba/

שחק, ע. (2020, מרץ 3). ועדת הבחירות חייבה את נתניהו להסיר את סרטון הפייק ניוז. *Ice*. Retrieved from http://www.ice.co.il/digital-140/news/article/778114