"אבטחת אמת" – ניתוח טכנולוגי של תופעת הפוסט-אמת

רון הלד

כיצד השפיעה הטכנולוגיה על האופן בו מידע מיוצר, מופץ ונצרך? האם יש קשר בין השפעה זו לתופעת הפוסט-אמת? וכיצד ניתן להתגונן מפני נזקיה של תופעה זאת? טענת מאמר זה היא כי בשנים האחרונות חל שינוי דרמטי בעולם המידע, וכי יש לנתח את תופעת הפוסט-אמת על רקע שינוי זה, ולא על-סמך הבנת עולם-המידע המסורתי. שינוי זה אף מסביר את הקושי בהתמודדות אפקטיבית עם נזקי התופעה, שכן הכלים והמנגנונים הקיימים נבנו להתמודדות עם אתגרי עולם המידע המסורתי. לאור זאת, מציע המאמר לפתח מתודולוגיית הגנה חדשה, "אבטחת-אמת" (Truth Security), בהשראת מתודולוגיות ההגנה הקיימות בעולם-המידע המודרני.

'ודאי שהעולם הזה הוא עולם של שקר, אבל כפסע בינו לבין עולם-האמת.'

יצחק בשביס-זינגר

לתופעת הפוסט־אמת הוצעו מגוון הגדרות,[1] שכולן עוסקות במערכת היחסים המשתנה בין הבנת המציאות והעובדות לבין הרגשות, האמונות והדעות האישיות. לענייננו, משקפת התופעה קושי הולך וגובר לברר את המציאות ולהבינה, הן בקרב הציבור הרחב והן בקרב מקבלי ההחלטות בתחום הביטחון הלאומי בדמוקרטיות מערביות ליברליות.[2] מטרתו של מאמר זה, שהוא חלק ממחקר העוסק בתופעת הפוסט־אמת על שלל היבטיה, היא לנתח את תופעת הפוסט־אמת מההיבט הטכנולוגי. הוא אינו מתאר את התופעה ואת נזקיה, אלא מנסה לענות על שתי שאלות עיקריות: כיצד השפיעו השינויים שחלו בעולם המידע בשנים האחרונות על תופעת הפוסט־אמת והעצימו את נזקיה? כיצד אפשר לפתח מתודולוגיית הגנה אפקטיבית נגד נזקיה של התופעה?

טענתו הראשונה של המאמר היא שאי אפשר להבין את צמיחתה של תופעת הפוסט-אמת בלי להבין את השינויים שהתרחשו בעולם המידע במעבר מהדור הראשון של אתרי האינטרנט (Web 1.0) לדור השני (Web 2.0). שינויים אלה יתוארו בחלקו הראשון של המאמר, ובפרט יתוארו הפיתוחים והחידושים הטכנולוגיים והטכנולוגיים־עסקיים (המודלים העסקיים של חברות הטכנולוגיה), העומדים בבסיסו של עולם המידע המודרני ומבדילים בינו לעולם המידע המסורתי. השינויים הללו הולידו שני קשיים עיקריים: קושי גובר לברר ולהבין את המציאות באמצעים המסורתיים; אי-יכולת של מנגנוני ההגנה המסורתיים להתמודד עם בעיית הדיסאינפורמציה. שילובם של שני הקשיים הללו יצר מצע פורה להתעצמות מגמות שהחלו עוד קודם לשינוי.[3] השינויים בעולם המידע יתוארו בחלוקה לארבעה תהליכים מרכזיים, שכל אחד מהם ינותח בנפרד: יצירת המידע, הפצת המידע, צריכת המידע וביקורת המידע. כן יכלול ניתוח התהליכים הסבר לכמה מהתופעות החדשות בעולם המידע המודרני – תיבות תהודה (echo chambers),

בועות סינון (filter bubbles) וחללים ריקים ממידע (data voids). בסופו של החלק הראשון תוצג תופעת הפייק ניוז (fake news) וייבחנו מגבלותיהם של כמה מהפתרונות המוצעים להתמודדות איתה.

בשני החלקים הבאים של המאמר יוצג מענה לשאלה "מה עושים?" וייטען כי כדי להתמודד עם התופעות החדשות יש לפתח מתודולוגיה, אסטרטגיות וכלים חדשים המותאמים במיוחד להתמודדות עימן, וכי את אלה יש לפתח בהשראתן של המתודולוגיה, האסטרטגיות והכלים להגנה שפותחו וכבר הוכיחו את יכולתם להתמודד עם אתגריו של עולם המידע המודרני – אבטחת סייבר (Cybersecurity) ואבטחת מידע (Information Security). להגנה מפני תופעת הפוסט־אמת יוצע לקרוא "אבטחת אמת" (Truth Security).

ההתבוננות בתופעת הפוסט־אמת במאמר זה היא מנקודת מבטן של הדמוקרטיות המערביות הליברליות. במדינות שבהן שולטים משטרים מסוג אחר, ובפרט באוטוקרטיות, יש לעולם המידע מאפיינים אחרים וארגזי כלים אחרים להתמודדות עם תופעות הפוסט-אמת המתרחשות בו ועם נזקיהן.[4]

עולם המידע המודרני

עולם המידע הוא המרחב שבו המידע הדיגיטלי מיוצר, מופץ, נצרך ומאוחסן.[5] מרחב זה מורכב ממידע גולמי וממידע מעובד (הניתוח במאמר זה אינו נזקק להבחנה ביניהם ולכן שני סוגי המידע נדונים יחד בשם "מידע"), מתשתיות מידע, כגון שרתי אחסון ואתרי מדיה, ומתשתיות תקשורת ומחשוב. חשוב לציין כי גם במקרים שבהם יצירת המידע אינה דיגיטלית, הפצתו וצריכתו תהיינה לרוב דיגיטליות. לכן, כמעט אין אירוע מֵידָעי שאינו מתרחש גם, אם לא בעיקר, בעולם המידע.

כדי לתאר את השינויים שחלו במעבר מעולם המידע המסורתי לעולם המידע המודרני, יובחנו ארבעה תהליכים מרכזיים: יצירת המידע, הפצת המידע, צריכת המידע וביקורת המידע. יתוארו מאפייניו של כל אחד מהתהליכים[6] ויוסבר כיצד הפיתוחים הטכנולוגיים והטכנולוגיים־עסקיים שחלו בעשור האחרון השפיעו על התהליכים ושינו את אופיים.[7] לבסוף, תוצג תופעת הפייק ניוז וייבחנו מגבלותיהם של כמה מהפתרונות המוצעים להתמודדות איתה.

ארבעת התהליכים המרכזיים של עולם המידע

1. **יצירת המידע**

תהליך יצירת המידע עבר שינוי גדול מאוד בעשור האחרון. שינוי זה נגרם הן על ידי המעבר של הפרט מצרכן מידע פאסיבי למשתמש אקטיבי והן בעקבות השפע והעוצמה של האפשרויות ליצירת מידע דיגיטלי, שהגורמים העיקריים להם הם: (1) זמינותם של הכלים הטכנולוגיים, המאפשרים ייצור מגוון ומהיר של מידע, כגון המצלמות ומצלמות הווידאו במכשירי הטלפון הנייד, ובמיוחד במכשירי הטלפון החכם; (2) זמינותם של משאבי אחסון זולים – האחסון במכשירים האלקטרוניים עצמם ובשירותי אחסון ייעודיים אחרים, כגון מחשוב ענן (Cloud Computing)[8] – שמעודדת ייצור יתר; (3) פיתוחם של כלים אוטומטיים ליצירת מידע מעובד

ממידע גולמי. יתרה מכך, הקלות שבה מידע מועתק ונערך מאפשרת מחזור וסילוף של מידע בקנה מידה רחב.[9]

שינוי חשוב נוסף הוא השיפור הדרמטי ביכולות עיבוד מידע מתקדמות של מגוון כלים לעריכת תמונות, טקסט, קול ו-ווידאו, המאפשרים לכל הרוצה בכך לזייף מידע. השימוש בפיתוחים אלה נרחב, והם מוטמעים כבר באמצעי הייצור עצמם, לאחר שעברו תהליכי הפשטה והנגשה למשתמש ההדיוט. לאחרונה התפרסמו כמה מקרים של שימוש בכלים של זיוף עמוק (deepfake) לעריכת וידאו.[10] כלים אלה משתמשים בשיטות למידה עמוקה (deep learning) ומאפשרים לשבץ בתמונה או בפס הקול המקוריים תמונות וקולות שלא היו שם קודם לכן.[11] אמצעי עיבוד וזיוף היו קיימים תמיד, אך בעידן הנוכחי הם נגישים לכול, ורמת העיבודים והזיופים היא כה גבוהה עד כי במקרים רבים אין לאדם הממוצע יכולת להבחין בין זיוף לאמת, ולעיתים אין יכולת כזאת אף לאנשי מקצוע. כדי להתמודד עם האיום הזה, הוקמו גופים ייעודיים לפיתוח דיסציפלינות מקצועיות חדשות לזיהוי זיופים עמוקים.[12]

2. הפצת המידע

תהליך הפצת המידע עבר מהפכה של ממש בשנים האחרונות. הקישוריות הגבוהה, זמינותם של אמצעי התקשורת בכלל ואמצעי התקשורת האישית בפרט (במיוחד מכשירי הטלפון החכם) ורוחבי הפס הגדולים מאפשרים הפצת מידע בקצב ובהיקף שלא נראו מעולם.

שתי תכונות עיקריות מבחינות בין מגוון הסגנונות להפצת מידע. האחת היא היחס מוסר–מקבל: אחד–לאחד (unicast), אחד–לרבים (סלקטיבי, multicast), אחד–לכולם (broadcast) ורבים–לרבים (many-to-many). השנייה היא סוג השיח: שיח הדדי (דיאלוג), שבו כל צד מוסר ומקבל מידע, ושיח חד־צדדי (מונולוג), שבו צד אחד מוסר מידע והשני מקבל. לשיטות ההפצה יש כמובן תכונות נוספות, אך השתיים האלה הן החשובות לענייננו.

סביב כל אחד מסגנונות הפצת המידע התפתחו פלטפורמות טכנולוגיות־עסקיות ייעודיות, שמטרתן להקל על תהליך ההפצה. אתרי האינטרנט, שעד לא מכבר הפיצו את המידע שהם עצמם ייצרו, משמשים היום פלטפורמה להפצת מידע שמייצרים המשתמשים (Internet as a Platform). עבור סגנון התקשורת אחד–לאחד התפתחו פלטפורמות המסרים המידיים (IM – Instant Messaging), כגון WhatsApp, Facebook Messenger, ו־Telegram. עבור סגנון התקשורת אחד–לרבים התפתחו הרשתות החברתיות, כגון Facebook, Instagram, ו־Reddit, וכן התקשורת הקבוצתית בפלטפורמות המסרים המידיים, כגון קבוצות WhatsApp וערוצי טלגרם. עבור סגנון התקשורת אחד–לכולם התפתחו פלטפורמת הציוצים Twitter, שהיא גם רשת חברתית, ופלטפורמת שיתוף הסרטים YouTube. עבור סגנון התקשורת רבים-לרבים התפתחו פלטפורמות הבלוגים, כגון Blogger, WordPress, ו-Tumblr, וכן פלטפורמות ה-Wikis כגון Wikipedia. רשימת הפלטפורמות הזאת היא כמובן חלקית מאוד. שילובן של כל הפלטפורמות האלה, ריכוזן באמצעי תקשורת אחד (הטלפון החכם או המחשב האישי) ופיתוחם של ממשקים נוחים להפצת מידע ביניהן – כל אלה מאפשרים הפצת מידע יעילה מאוד לכל פרט ופרט, ללא כל צורך באמצעים מיוחדים.

חשיבות כבירה יש להתפתחות המודלים העסקיים סביב הפלטפורמות הטכנולוגיות. השוק של פלטפורמות המשרתות שני סוגי משתמשים בכל רגע, כגון מפרסמים–צרכנים וספקים–לקוחות,

מייצר דינמיקה המכונה "המנצח לוקח הכול" (Winner-Takes-All Market).[13] בשוק זה כל יתרון קטן, בביצועים או בנתח שוק, מתורגם ליתרון עסקי גדול, שמזין את עצמו במהירות וגורם להשתלטות של מיעוט שחקנים על השוק ולשחיקה מהירה של התחרותיות. כך הפכו הפלטפורמות של הפצת המידע לחברות ענק, שהשתלטו על שווקים שבהם פעלו קודם לכן גורמים אחרים, כגון העיתונות המודפסת, רשתות הטלוויזיה וענקיות המדיה. רווחיותן של חברות אלה אפשרה להן לרכוש מתחרות או חברות המייצרות עבורן דריסת רגל בשווקים משיקים (למשל הרכישה של Facebook את Instagram ואת WhatsApp). כך קרה, שחרף היכולת הטכנולוגית ליצור דרכים מגוונות להפצת מידע, הצטמצמה הפצת המידע למספר קטן מאוד של פלטפורמות, שבהן משתמשים מרבית האזרחים בדמוקרטיות המערביות הליברליות.

להלן כמה תופעות שהתפתחו בפלטפורמות הטכנולוגיות:

(1) נוסף על הפצת המידע בידי בני אדם, בעולם המידע המודרני מפיצים מידע דרך קבע יותר ויותר שחקנים וירטואליים, כגון בוטים (Bots) והמלצות תוכן אוטומטיות. השחקנים הווירטואליים הם לעיתים אוטונומיים לחלוטין ולעיתים משרתים אינטרסים עסקיים. הפצת המידע באמצעות השחקנים הווירטואליים מתאפיינת בנטייה להסתה[14] ובאמינות נמוכה במיוחד.[15]

(2) היכולת לשתף תוכן או להפיצו מחדש ייצרה תופעה חברתית־מֵידָעית, שבמסגרתה מידע חדש ומרעיש זוכה להפצת יתר. מכיוון שקל כל כך להפוך סיפור שקרי לחדשה מרעישה, המידע השקרי מופץ לעיתים אף יותר ממידע אמיתי.[16] לכך יש להוסיף את המניעים הפסיכולוגיים והסוציולוגיים, כגון חרדות ההחמצה (FOMO – Fear of Missing Out & FOBM – Fear of Being Missed), המעודדים כל פעילות שהיא בפלטפורמות המידע, ללא שימת לב לאיכותן ולאמינותן.

(3) למידע האישי המצוי בפלטפורמות הטכנולוגיות יש ערך אדיר למפרסמים, המשתמשים בו כדי להתאים ביתר דיוק את מאמצי הפרסום למשתמש. כך נוצר ה־microtargeting – פילוח שוק לסגמנטים על בסיס כריית מידע (data mining) והתאמת הקמפיין לכל פלח שוק בנפרד. כלי הפצת המידע רב העוצמה הזה משרת ביעילות רבה מגוון אינטרסים, ובראשם מסחריים ופוליטיים.

(4) סביב השימוש בפלטפורמות הטכנולוגיות נוצרה תעשייה משלימה, במיוחד בתחומי השיווק והמכירות, המעודדת אף היא הפצת מידע שאמינותו לא נבדקה ושכל תכליתו היא למשוך את תשומת ליבו של המשתמש: חברות קליקבייט (clickbait), המשתמשות בכותרות מרעישות או בתמונות מסקרנות כדי לפתות את המשתמש ללחוץ עליהן ולפתוח את התוכן המקושר אליהן; וחברות תשלום לפי לחיצה (PPC – Pay Per Click) או עלות לחשיפה (CPI – Cost Per Impression), המקבלות תשלום מהמפרסם לפי מספר הלחיצות על הקישור או לפי מספר החשיפות לפרסומת. חברות אלה תרות כל הזמן אחר מידע מסקרן שימשוך את תשומת ליבו של הצרכן, ובדרך זו מידע שקרי ומרעיש עובר פעמים רבות הדהוד נוסף. כך ניסח זאת בחריפות ג'ף המרבכר, מייסד חברת התוכנה וניתוח הנתונים Cloudera ולשעבר ראש צוות ה־data בחברת Facebook: "The best minds of my generation are thinking about how to make people click ads".[17]

נוסף על אלה התפתחו שלוש תופעות חשובות, הממחישות היטב את האתגרים החדשים של עולם המידע המודרני. תופעות אלה מראות כיצד נוצר בו כר פורה ועשיר להפצה ויראלית של מידע שקרי ושל תיאוריות קונספירציה, וכיצד הוא מטפח חוסר ביקורתיות והקצנה של דעות ואמונות.

(1) **תיבות תהודה** (echo chamber) – כאמור, לפרט הבודד קל היום יותר מאי פעם להפיץ מידע. נוסף על כך, פלטפורמות הפצת המידע מאפשרות לפרט להשתייך לקהילות סגורות המתאגדות סביב דעות ואמונות מסוימות. באופן טבעי, יפיץ הפרט ביתר נוחות מידע התומך בדעות ובאמונות מסוימות לקהילות המאוגדות סביבן, ויחזק עוד יותר את אמונותיהם של חברי הקהילה, שאף הם בתורם יפיצו את המידע התומך.[18] כך נוצר אפקט העדפת המוכר (mere exposure effect), המזין את עצמו: הפרט יצרוך את אותו המידע שוב ושוב, אמונתו באמיתות המידע תלך ותתחזק, וחוזר חלילה. תהליך זה מעצים ומקצין את הדעות והאמונות הרווחות בקהילות חברתיות סגורות בפלטפורמות השונות. יתרה מכך, הפעילות בקהילות אלה מפחיתה את הרצון של הפרט לשתף את דעותיו ואת אמונותיו בסביבות ידידותיות ותומכות פחות.[19] זוהי דוגמה לביטוי חדש של נטייה אנושית ידועה ומוכרת.

(2) **בועות פילטר** (filter bubble) – המודלים העסקיים של הפלטפורמות הטכנולוגיות מתבססים על רמת חיבור (engagement) גבוהה של המשתמשים לפלטפורמות. כדי לייצר את החיבור ולהעצים אותו הפלטפורמות מנסות להתאים את חוויית השימוש בהן לכל משתמש על פי צרכיו האישיים ועל פי הרגלי הצריכה שלו. לשם כך הן משתמשות במידע האישי של המשתמש, שאותו הן אוספות ומנתחות תוך כדי השימוש שלו בפלטפורמה. האלגוריתמים של הפצת המידע בכל פלטפורמה לומדים את דפוסי השימוש והצריכה של המשתמש ושל משתמשים דומים לו. הם לומדים, למשל, מהחיפושים שהמשתמשים עושים, מהזמן שהם מבלים בצריכת מידע כזה או אחר ומה־like-ים שהם נותנים למידע שמפיצים משתמשים אחרים. על בסיס הלמידה הזאת הם מפיצים אל המשתמש מידע הדומה למידע שצרך בעבר ושכנראה ירצה לראות גם בעתיד, ומציעים לו לעקוב אחר משתמשים שכנראה יעניינו אותו.[20] כמו כן, הם מפיצים אליו מידע מרעיש שיש לו כוח משיכה גדול, וכך נמנע ממנו להיחשף למגוון של דעות ורעיונות ולמידע הסותר את מערכת הדעות והאמונות שלו. כך, כשהיא מבודדת את הפרט בתוך בועה עם מסננת המותאמת לו אישית, הפלטפורמה מעצימה ומעודדת את החיבור של הפרט אליה.

(3) **"חללים ריקים ממידע"** (Data Voids) – שלא כמו השימוש ברשתות החברתיות, השימוש במנועי החיפוש מתחיל בשאלה שמזין צרכן המידע. חללים ריקים ממידע נוצרים כאשר המידע הנותן מענה לשאלה מצומצם, לא מדויק או לא קיים. גורמים בעלי עניין יכולים לזהות את החללים הריקים ולהזרים לצרכן בתשובה לשאלתו תוכן המקדם את עניינם. הבעיה מחריפה כשכלים מסחריים שתכליתם קידום מסחרי במנועי החיפוש (SEO – Search Engine Optimization) מנצלים את החללים הריקים כדי להגביר את החשיפה לתוכן שהם מקדמים. לעיתים מייצרים הגורמים הללו באזז מלאכותי, שתכליתו לעודד חיפושים שהם דאגו מראש שיניבו את התוצאות הרצויות להם. במקרים אחרים הם מנצלים היווצרות ספונטנית של באזז תקשורתי ואת העלייה במספר החיפושים שנוצרים בעקבותיו.[21]

3. צריכת המידע

כמו תהליכי היצירה וההפצה של המידע, גם תהליך צריכת המידע עבר שינוי קיצוני בשנים האחרונות. אחת התכונות החשובות של צריכת המידע היא סינכרוניות ואסינכרוניות. צריכה סינכרונית היא צריכה של פרטים רבים ברגע נתון. היא תוצר של הפצת מידע בסגנון אחד–לכולם, כגון מהדורות החדשות של השעה שמונה בערב, שאפיינה את עולם המידע המסורתי. לעומתה, צריכה אסינכרונית היא צריכה של פרטים שונים בזמנים שונים, לרוב בעקבות יוזמה של הפרט עצמו. בעוד שגופי המדיה של עולם המידע המסורתי גרמו להתנהגות סינכרונית של הצרכנים, הרי שבעולם המידע המודרני צריכת המידע היא ספוראדית וספונטנית יותר ולכן אסינכרונית בדרך כלל. המשמעות היא, שצרכני המידע בעולם המידע המודרני מחזיקים בכל רגע נתון בתמונת מציאות מעט שונה, ומידע מוטעה או שקרי יכול בקלות להרחיב את הפערים ובכך להעמיק את הקוטביות והפילוג בחברה. בסוגיות בעלות חשיבות לאומית רב הצורך בהסכמה ציבורית רחבה, ועל כן לתכונת האסינכרוניות, המאפיינת את עולם המידע המודרני, יש חשיבות רבה לענייננו.[22]

נוסף על כך, הפלטפורמות הטכנולוגיות הפכו להיות הפלטפורמות המרכזיות לצריכת מידע על חשבון פלטפורמות המדיה המסורתיות, כגון הטלוויזיה והעיתונות המודפסת.[23] בעקבות זאת, רווחי הפרסום של פלטפורמות המדיה המסורתיות, שהיו במשך שנים ארוכות בסיס המודל העסקי שלהן, נשחקו מאוד עוד בתחילת עידן האינטרנט, טרם עליית הפלטפורמות הטכנולוגיות, והועתקו מאז יותר ויותר לפלטפורמות הטכנולוגיות.

שחיקה עסקית זו יצרה מעגל קסמים (vicious circle), שבו יצירת מידע איכותי ויקר, שהולידה בעבר רייטינג ותחזקה מודל כלכלי בריא, מתוגמלת פחות מיצירת התוכן באיכות נמוכה יותר שמייצרות הפלטפורמות החדשות. ככל שיצירת תוכן איכותי מתוגמלת פחות, כך ההצדקה להמשיך ולתחזק את המערכות המסורתיות פוחתת והולכת, ואכן ניכר צמצום רב במספר המערכות מסוג זה.

עניין חשוב נוסף: בעבר, יצירת המידע והפצתו נעשו באותה פלטפורמה, היא פלטפורמת המדיה המסורתית, וכך האחריות על אמינותו של המידע ועל איכותו הייתה של מפיץ המידע עצמו. כיום התפצלו שני התהליכים, ויצירת המידע נעשית בנפרד מהפצת המידע,[24] ובפרט אין הפלטפורמות הטכנולוגיות אחראיות, או רוצות להיות אחראיות, על המידע שהן מפיצות,[25] פרט למקרים של עבריינות קיצונית, כגון הפצת תוכן פדופילי, שבהם נוקטות הפלטפורמות צעדים אקטיביים להסרת המידע. עניין האחריות על התכנים המופצים נמצא במרכז הלחץ המופעל על הפלטפורמות, שבעקבותיו החלו הפלטפורמות לשנות את גישתן.[26]

כך נוצר מעגל צרכני חסר רגולציה: פרט או ארגון מייצרים מידע ומפיצים אותו בעזרת פלטפורמות הפצת המידע; את המידע הזה צורכים פרט או ארגון אחרים (במיוחד אלה השייכים לתיבות התהודה הרלוונטיות או אלה שהאלגוריתמים להפצת המידע מצאו אותם מתאימים לקבלו), שמפיצים אותו בתורם באמצעות הפלטפורמות; וחוזר חלילה. וכל זה בקצב מהיר מאוד וחסר תקדים. מעגל צרכני זה הוא הזדמנות פז עבור פרטים או ארגונים בעלי עניין – מסחרי, פוליטי או אחר – ליצירת תוכן שמקדם את עניינם ולהפצתו בהיקף נרחב. הוא מאפשר לבחון את הלך הרוח בציבור, לזהות את המידע המשפיע ביותר, באמצעות ניסוי וטעייה או באמצעות בדיקת A/B (A/B Testing), להקצין את מסריו ולהעצים את תפוצתו. כל זאת, בזמן קצר ובעלות נמוכה,

המאפשרים קבלת משוב, הפקת לקחים ויישום המסקנות בזמן אמת. מעגלי המשוב (feedback loops) הקצרים והזולים של עולם המידע המודרני טובים במידה ניכרת מאלה של עולם המידע המסורתי ומאפשרים לבעלי העניין לפתח כלים אפקטיביים יותר מאי פעם להפצת המידע המשרת את עניינם.[27]

4. ביקורת המידע

בתהליך ביקורת המידע בודקים האם המידע נכון ומדויק, מזהים חריגות וחוקרים את מקורותיהן, ואם נמצא שהמידע כוזב וְיָצַר או הופץ בזדון, אוכפים את המחיר המשפטי והחברתי על יוצרי המידע ועל מפיציו. בשנים האחרונות, בעקבות השינויים בעולם המידע, חלה הפחתה ניכרת ביכולת לבקר את המידע ובאפקטיביות של הביקורת. האנונימיות המאפיינת את הפעילות בעולם המידע המודרני (anonymity & untraceability), המחסור ברגולציה[28] ובחקיקה נגד עבירות במרחב זה (no policy), ההיקף העצום של המידע ושל הפעולות בעולם המידע, שלא מאפשר לבקר הכול בזמן אמת (cacophony), וכן היעדרה של מערכת איזונים ובלמים חברתית־ערכית, שמטרתה לרסן ולפקח על הפלטפורמות הטכנולוגיות החדשות ועל פעילות הגולשים (uncontrolled space) – כל אלה החלישו מאוד את תהליכי הביקורת המסורתיים, שהתבססו על הצלבת מקורות מידע שונים, על היכרות אישית עם מפיץ המידע ועל המוניטין המוכח שלו, על חוש הביקורת האישי של צרכן המידע בנוגע לסוגיות מקומיות וקרובות אליו, על איום בתביעות דיבה ועל מומחים שיוודאו את אמינות המידע.

ההיחלשות העסקית של פלטפורמות המדיה המסורתיות והצורך שלהן להתחרות בדפוסי הצריכה של מידע באינטרנט, הובילו לשחיקה מואצת של האתיקה העיתונאית. אינטרסים פוליטיים וחולשה מקצועית בהתמודדות עם אתגרי העידן הנוכחי העצימו מאוד את תהליך היחלשותם של שומרי הסף המסורתיים, כגון מומחי תוכן, מוסדות ביקורת ממלכתיים ומערכת המשפט, ובשל כך גם גרמו לירידה באמון הציבור בהם.

תופעת הפייק ניוז

תופעת הפייק ניוז היא חלק מתופעת הפוסט־אמת הרחבה יותר וממחישה את מאפייני עולם המידע המודרני שתוארו לעיל תחת ארבעת התהליכים: יצירת מידע שקרי בקלות, בקצב מהיר מאוד ובאיכות גבוהה; הפצה נרחבת של המידע השקרי והקצנתו באמצעות הדהודים ושיתופים בפלטפורמות הפצת המידע; צריכה מוגברת ולא ביקורתית של הציבור הרחב את המידע השקרי; וחוסר ברגולציה, בחקיקה ובאכיפתן, כדי למנוע את יצירתו ואת הפצתו של המידע השקרי או כדי למחוק אותו לאחר מעשה.

כדי להתמודד עם תופעת הפייק ניוז משתמשים כיום במנגנוני הגנה שהיו אפקטיביים בעולם המידע המסורתי, אך אינם מותאמים לעולם המידע המודרני:

1. כלים לבדיקת עובדות (fact check), אנושיים, טכנולוגיים והיברידיים, שתכליתם לנתח את העובדות המופיעות בפריט המידע ולקבוע האם הן אמיתיות או שקריות. בפועל, כלים אלה מגנים בעיקר מפני צריכה לא מודעת של מידע שקרי[29] ומסייעים ביצירת תרבות ביקורתית. אך ללא מנגנונים משלימים של אכיפה נגד יצירה והפצה של מידע שקרי, הרי שכוחם של הכלים הקיימים מוגבל מאוד.[30]

2. חשיפת שחקנים וירטואליים ייעודיים (רשתות בוטים) המפיצים מידע שקרי. מטרותיה של החשיפה הן הסרתם של השחקנים הווירטואליים מפלטפורמות הפצת המידע והפסקת השימוש בהם. לעיתים הפלטפורמות עצמן הן שחושפות את השחקנים הווירטואליים[31], אך רבים המקרים שבהם השחקנים הווירטואליים מזוהים ונחשפים רק לאחר שעיקר הנזק כבר נעשה, דהיינו לאחר שכבר הספיקו להפיץ מידע שקרי רב, והוא מוסיף להיות מופץ בידי מפיצים אנושיים לגיטימיים.

כלי ההגנה הקיימים, אם כן, אינם יכולים לשמש אסטרטגיית הגנה אפקטיבית מפני תופעת הפייק ניוז, וכל ניסיון להשתמש בהם נידון לכישלון, כפי שאכן אירע בשנים האחרונות. מצב זה הוא כר פורה של הזדמנויות חדשות ליצירה ולהפצה של מידע שקרי, במזיד או בשוגג, ולערעור של העובדות כמסד לידע על המציאות ושל גבולות השיח הלגיטימי בדמוקרטיות המערביות הליברליות.

מתודולוגיה, אסטרטגיות וכלים להגנה בעולם המידע המודרני

כדי להתמודד עם מאפייניו החדשים של עולם המידע המודרני פותחה מתודולוגיה חדשה להגנה – 'הגנת מידע',[32] וכדי ליישמה פותח ארגז כלים, טכנולוגי ואנושי, חדש. המתודולוגיה וארגז הכלים החדשים שפותחו הם טובים ויעילים לצורכי הגנה שונים בעולם המידע המודרני – הגנה על מידע של פרט בודד, הגנה על מידע של ארגון או עסק, הגנה על תשתיות מחשוב, הגנה על תשתיות קריטיות או הגנה על מידע של מוסדות מדינתיים – אך הם חסרים עדיין את היכולת להתמודד ביעילות עם תופעת הפוסט־אמת. הטענה המרכזית במאמר זה היא כי המתודולוגיה וארגז הכלים האלה צריכים לשמש בסיס והשראה לפיתוח מתודולוגיה וארגז כלים המתאימים במיוחד להתמודדות עם תופעת הפוסט־אמת. להלן יתוארו עיקרי מתודולוגיית ההגנה, יוצגו דוגמאות לאסטרטגיות הגנה שפותחו על פיה ויודגם השימוש בכלים שפותחו כדי ליישם אסטרטגיות אלה. בהשראתם יוצעו בחלק המסכם של המאמר קווים מנחים לפיתוחם של מתודולוגיה, אסטרטגיה וכלים להתמודדות עם תופעת הפוסט־אמת.

הגנת המידע, שהיא גישת האבטחה וההגנה בעולם המידע, מקבילה, בכפוף להתאמות הנדרשות, לגישת האבטחה וההגנה במרחב הפיזי, שאותה הגדיר יהושפט הרכבי:

"בהגנה המטרה היא שמירה (preserve) על טריטוריה, על נכסים ועל כוחות [...] הדבר מתממש בסיכול מטרת המתקיף על-ידי עצירתו, בלימתו, אמנעתו והדיפתו (ward off, parry, interdict, repulse), וההגנה היא מבחינה זו מאמץ למנוע שינוי. בכל אלה המטרה היא **שלילית**, שכן פירושה למנוע הישג מן התוקף (denial) [...] (ראוי לשים לב שהתרגום של שני הפעלים באנגלית "defend" ו"protect" לעברית הוא "להגן"; הפועל protect קרוב לפועל לשמור או לאבטח.)"[33]

לפי הרכבי, מטרת ההגנה היא לשמור על נכסים, דהיינו למנוע מהתוקף פגיעה בהם או השגתם. ובהשאלה – מטרתה של הגנת המידע היא לשמור על הנכסים בעולם המידע. להלן דוגמאות לנכסים מרכזיים של עולם המידע, שעליהם יש לשמור:[34]

1. **נכס הפרטיות והסיווג**: הנכס הוא המידע גופא, והשמירה עליו מתבטאת בכך שרק לגורמים המורשים יש גישה למידע.

2. **נכס האמינות**: הנכס הוא פער מינימלי בין המידע כפי שהיה בעת יצירתו ובין המידע בעת צריכתו. גם אם אין כל פער, אך צרכן המידע חושש שיש פער, הרי שיש פגיעה באמינות ובעקבותיה פגיעה בשימושיות המידע. נכס זה שקול לקבילותה של ראיה בבית המשפט.

3. **נכס הרציפות והזמינות**: הנכס הוא יכולתו של צרכן המידע לגשת למידע המאוחסן בעולם המידע. בהגדרה זו נכללת גם הרציפות התפקודית של מערכות המשתמשות במידע לצורך תפקודן השוטף, כגון מערכות לניהול פס ייצור המתבססות על הזנת מידע מחיישנים הפזורים במפעל.

4. **נכס יכולת הייחוס**: הנכס הוא היכולת של צרכן המידע ושל הרשויות לייחס פעילות הקשורה למידע, כגון יצירה והפצה, לפרט או לארגון. נכס זה יכול להיות חשוב לקביעת אמינותו של מידע בתהליך ביקורת המידע.

מטבע הדברים, לארגונים שונים יש נכסים וסדרי עדיפויות שונים. לכן, אסטרטגיות הגנת המידע צריכות להתבסס, בראש ובראשונה, על מיפוי נכסי המידע של הארגון ועל קביעת סדר העדיפויות שלהם. משהוגדרו הנכסים, נשאלת השאלה כיצד יכול התוקף לפגוע בכל אחד מהם ומהו איום הייחוס. באופן מסורתי מקובלת החלוקה לשלוש תכליות התקפיות (Offensive Cyber):[35]

1. **תקיפה לצורך פגיעה, הרס ומחיקה** (CNA – Cyber Network Attack): מטרתה גרימת נזק למידע או לתשתיות היצירה, ההפצה והאחסון שלו. תקיפה מסוג זה יכולה להיות בעלת אופי של "כפתור אדום" (kill switch), כלומר, פעילות חשאית המאפשרת יכולת ביום פקודה, או בעלת אופי טקטי וגלוי יותר.

2. **תקיפה לצורך הפקת מידע או ריגול** (CNE – Cyber Network Exploitation): חשאית או גלויה, שמטרתה איסוף מידע.

3. **תקיפה לצורך השפעה** (CNI – Cyber Network Influence): מטרתה לוחמה פסיכולוגית, פגיעה בריבונות או פגיעה באורח החיים הנורמטיבי, באמצעות טְפָלוּל (מניפולציה) המידע או טפלול תשתיות המידע והתקשורת.

ככל פעילות התקפית, גם בעולם המידע יש לתוקף אסטרטגיות וטקטיקות מגוונות מאוד, שלהן מנעד רחב של יכולות מבצעיות ומודיעיניות משלימות (לא כאן המקום לפרטן). חשוב להדגיש כי לעיתים קרובות מנצל התוקף גורם פנים בארגון או בגוף הנתקף, ביודעין (משתף פעולה) או שלא ביודעין (מְשֻׁטֶּה). שימוש כזה בגורם פנים מוסיף מורכבות וקושי להגנה על המידע.

אם כן, כדי לנסח אסטרטגיית הגנה אפקטיבית יש לאפיין את איום הייחוס: מהם הנכסים? מהו הסיכון לכל נכס (CNA, CNE, CNI) ומהי חומרתו? ומהו סדר העדיפויות להגנה על הנכסים? נוסף על אלה, יש לאפיין את הגישה ההגנתית ואת הכלים למימוש האסטרטגיה.

לכל פעילות התקפית שרוצים להתגונן מפניה אפשר לבחור מבין מגוון גישות הגנה, ואלה המרכזיות שבהן: **גישת מניעה**, שמטרתה למנוע מהתוקף הישג כלשהו, יהא אשר יהא; **גישת הכלה**, שמטרתה למנוע מהתוקף לממש את ההישג שאליו כבר הגיע; **גישה פסיבית**, המתגוננת בעולם המידע של המגן בלבד; ו**גישה פרואקטיבית**, המתמודדת עם התקיפה גם בעולמות

נוספים, כגון עולם המידע של התוקף או העולם הפיזי. אפשר לשלב גישות שונות, אך תהיה לכך השלכה על המשאבים המוקצים למשימת ההגנה.

הדוגמאות להלן ממחישות כיצד התפתחו ארגזי כלים שונים להגנה על שניים מהנכסים שהוגדרו:

1. המידע הארגוני הרגיש והמסווג ביותר (crown jewels) הוא נכס שמקובל ומתבקש להגן עליו מפני העתקה, גניבה או פגיעה. הכלים ליישום ההגנה על נכס זה מגוונים:

   (1) הגבלת הגישה למידע באמצעות מנגנון הרשאות (מניעת גישה שאינה מורשית).

   (2) בידול המידע הרגיש מתשתיות שרגישותן נמוכה מרגישותו, כדי לאפשר הכלה של פגיעה בתשתיות אלה ומניעת פגיעה בתשתיות הרגישות ביותר.

   (3) שמירה על שערי הכניסה הישירים אל המידע באמצעות חומות אש (firewall), שתכליתן למנוע חדירה של גורמים זדוניים לרכיב או לרשת תקשורת, ובאמצעות אזורים מפורזים (Demilitarized Zone), שתכליתם לחצוץ בין הרשת הפנימית של הארגון ובין רשתות חיצוניות ולאפשר הכלה של חדירה מסוימת.

   (4) התקנת מלכודת דבש (Honeypot) ברשת – רכיב שכל מטרתו היא לכוון את התקיפה אליו במקום אל המידע הרגיש, וכך להתריע מפני התקיפה, לחשוף את שיטות הפעולה של התוקף או להכיל את התקיפה.

   (5) יצירת קווי הגנה וביצור במעגלים מתרחבים סביב המידע למניעת גישה לקרבתו של המידע הרגיש.

   (6) פיזור סנסורים והקמת מרכזי ניטור (SOC – Security Operation Center)[36] לניטור הגישה למידע הרגיש ולזיהוי חריגות בגישה אליו ובתעבורה שלו.

   (7) כדי להגביל הישג ולמנוע נזק היקפי נדרש מערך אופרטיבי שיוכל להתערב ברגע שתזוהה פעילות התקפית מעבר לקווי ההגנה. מערך אופרטיבי כזה מורכב למשל מ־CERT (Computer Emergency Response Team)[37] ומ־IRT (Incident Response Team)[38].

2. גם זמינותו הרציפה של המידע היא נכס שמקובל ומתבקש להגן עליו. שלא כמו בדוגמה הראשונה, כשמטרת ההגנה היא למנוע מהתוקף להשיג את המידע, בדוגמה זו המטרה היא לאפשר למותקף להמשיך בפעילותו האורגנית, כלומר, לאפשר לו נגישות רציפה למידע ושימוש רציף בתשתיותיו. מגוון האיומים שיש להתגונן מפניהם הוא רחב:

   (1) נגד איום של מחיקת המידע יש לנקוט צעדים שימנעו את הפגיעה בו או שיכילו אותה ויאפשרו את שרידותו באמצעות כלי גיבוי ושחזור (DRP – Disaster Recovery Plan).

   (2) נגד תקיפה שמטרתה למנוע ממשתמש מורשה את הגישה לשירות או לרשת (DoS – Denial of Service) יש לנטר את הפעילות התקשורתית, ברשת בכלל ובסביבת השירות בפרט, לזהות בה חריגות ולסווג את התעבורה הזדונית. לכל סוג של תעבורה זדונית יש להשתמש בכלי המניעה והבלימה המתאימים לו (IPS[39], DDS[40], WAF[41]) וכן בשירותי ההגנה של ספקית התקשורת, כדי לעצור את ההתקפה טרם הגעתה לרף המונע מהשירות

לתפקד (מניעה של מניעת השירות) או כדי ליצור תשתיות תקשורת העוקפות את ההתקפה ומאפשרות גישה חרף ההתקפה (הכלה של מניעת השירות).

(3) נגד איום שמטרתו השבתה של רכיב קריטי ברשת תקשורת צבאית, שתמנע מהרשת תפקוד בזמן מלחמה, יש לדאוג לגיבוי חם, המופעל אוטומטית ומידית עם נפילת התקשורת, או לגיבוי קר, המופעל ידנית עם נפילת התקשורת, או לשניהם.

לכל נכס, אם כן, נדרש מגוון רחב של כלים הרלוונטיים להגנה עליו והמשלימים זה את זה. אולם גישה הגנתית המפעילה בו־זמנית כלי הגנה רבים עלולה לגרום לפגיעה חמורה בזמינות של המידע ובעקבותיה לפגיעה בתפקוד האופרטיבי של הארגון. יתרה מכך, לעיתים דווקא ריבוי של כלי הגנה מייצר התאבכות הורסת (destructive interference) ופוגע בהגנה על הנכס. על כן, אופי ההגנה על כל נכס ייקבע גם לפי אופי השימוש בו, ואופן השימוש בכלים השונים ושילובם מחייב מחשבה ותכנון קפדניים.

כל הכלים שהוצעו לעיל בשתי הדוגמאות פועלים בעולם המידע של המגן בלבד. חשיפה של מאמץ התקפי, של זהות התוקף ושל שיטות הפעולה שלו, כמו גם נקיטת אמצעים משפטיים נגד התוקף ונגד המעורבים בבנייתן של יכולות התקיפה, יוצאות מתחומי עולם המידע של המגן ומיישמות גישת הגנה פרואקטיבית בשירות תכליות ההגנה. גישה פרואקטיבית להגנה יכולה, למשל, לשלול את החשאיות של פעילות התוקף (שלילה חשאית) – אם התוקף ידע שתקיפתו נחשפה, הוא עשוי להפסיק אותה. ניתן גם לסכל לגמרי את מערכי התקיפה שלו באמצעות פרסום פומבי (שלילה רועשת) של התקיפה ושל כלי התקיפה, יחד עם ניתוח יכולותיהם ותיאור הדרכים לזיהויים ולמניעתם. פעולות הגנה פרואקטיבית כגון אלה עשויות לגבות מהתוקף מחיר כלכלי – המשאבים שהושקעו בפיתוח כלי התקיפה ואובדן היכולות שפותחו – והן מחייבות איסוף מידע מקדים על פעולותיו של התוקף. כלים התקפיים כגון אלה, המשרתים את תכליות ההגנה, הם חלק מלוחמת מידע נגד התוקף ומטרתם להשפיע על התוקף ועל קבלת ההחלטות שלו.

לאחר שגובשה אסטרטגיית ההגנה והותאמו הכלים ההגנתיים וההתקפיים לכל נכס, נדרש לבחון ולוודא את האפקטיביות שלה באמצעות שיטות בקרה המדמות את פעילות התוקף: (1) בדיקות יזומות שבהן צוות מומחים מנסה לחדור לרשת התקשורת של ארגון ולפגוע בנכסיו (penetration testing); (2) צוות אדום (Red Teaming) של מומחים, שמטרתו לבחון ולבקר הנחות יסוד, תפיסות ותוכניות של ארגון;[42] (3) ומשחקי מלחמה. נוסף על הכלים הטכנולוגיים ועל השגרות ונהלי העבודה לשימוש בהם, יש להכשיר ולעודד את העובדים בארגון לאמץ תרבות הגנתית – שמירה על ערנות, דיווח בצירים שהוגדרו, שימוש בהיגיון בריא ובגישה ביקורתית. תכליתה של התרבות ההגנתית היא לצמצם את יכולתו של התוקף לנצל את החולשות האנושיות של עובדי הארגון כדי לגרום להם לפעול בהתאם לצרכיו (social engineering).

אבטחת אמת (Truth Security) – מתודולוגיה, אסטרטגיות וכלים להגנה מפני תופעת הפוסט־אמת

בהשראת המתודולוגיה, האסטרטגיות וכלי ההגנה שתוארו לעיל, יוצגו בחלק זה של המאמר השלבים הנדרשים לפיתוחם של מתודולוגיה, אסטרטגיות וכלים להגנה מפני תופעת הפוסט־אמת:

1. בשלב ראשון, יש להגדיר מהם הנכסים שיש להגן עליהם. אלה הם הנכסים שפגיעה בהם תפגע ביכולתה של החֶבְרָה (society) לקבל החלטות על בסיס עובדות ועל בסיס הבנה נכונה של המציאות.

בסוגיית החיסונים נגד נגיפים ומחלות, למשל, הנכסים להגנה יכולים להיות: (1) זמינות המידע על החיסונים, על האפקטיביות שלהם ועל הסכנות שבהם, בכל פלטפורמה שהמידע נצרך בה, כגון רשתות חברתיות, מנועי חיפוש ומדיה מסורתית; (2) אמינות המידע על החיסונים, על האפקטיביות שלהם ועל הסכנות שבהם; (3) שמה הטוב של מערכת הבריאות ושל העומדים בראשה, ובמקרה של חיסונים שמה הטוב של מערכת הבריאות לגיל הרך בפרט; (4) מערכות המידע האישיות והארגוניות של ארגונים אלה; (5) היכולת לייחס פגיעה בכל אחד מהנכסים לפרט או לארגון.

2. בשלב השני, יש להגדיר מהם הנזקים שאפשר להסב לכל נכס ומהי גישת ההגנה הנבחרת: מניעת הנזק, הכלתו או שילוב של השתיים; הגנה פאסיבית, הגנה פרואקטיבית או שילוב של השתיים.

בתהליך הבחירות (לפרלמנט, לנשיאות, לראשות הממשלה), למשל, הנכסים להגנה והנזקים האפשריים להם יכולים להיות:[43] (1) מערכות ההצבעה האלקטרוניות וספירת הקולות – הגנה מפני השבתת המערכות או שיבוש ההצבעות והספירה לטובת מועמד זה או אחר (איום מסוג CNA), והגנה מפני התקפה פסיכולוגית שמטרתה לערער את אמון הציבור במערכות הללו (איום מסוג CNI); (2) שמם הטוב של מועמדים – הגנה על המידע האישי של מועמד מפני תוקף שרוצה לסחוט אותו, להפעיל עליו לחץ או להשפיע על שותפיו הפוטנציאליים (איום מסוג CNE), והגנה מפני הפצת מידע שקרי ויצירת באזז תקשורתי סביב סוגיות בעייתיות שמטרתם לערער את אמינותו של המועמד בעיני הציבור או להסיט דיון ציבורי מסוגיות חשובות אחרות (איום מסוג CNI); (3) אמון הציבור במוסדות האמונים על תהליך הבחירות התקין – הגנה מפני הפצת מידע שקרי המייחס למוסדות הללו הטיה פוליטית כזו או אחרת (איום מסוג CNI).[44]

3. בשלב השלישי יש לפתח את הכלים הנדרשים להגנה על הנכסים מפני הנזקים האפשריים. יש להניח, כי נכסים שונים יזדקקו לכלי הגנה שונים וכי רבים מהכלים הללו חסרים היום ויהיה צורך לייצר את התנאים המאפשרים ומעודדים את פיתוחם. נגד חלק מהאיומים צריך יהיה לפתח כלים התקפיים שישמשו את תכלית ההגנה. נוסף על כלים טכנולוגיים, נדרש לפתח מנגנונים אנושיים – מיומנויות ומקצועות חדשים – שיאפשרו את יישומם ואת הטמעתם של הכלים הטכנולוגיים.

בסוגיית החיסונים, למשל, כלי ההגנה יכולים להיות: (1) כלים להפצת מידע אמיתי ולמניעת הפצה של מידע שקרי על חיסונים בפלטפורמות השונות;[45] (2) כלים לזיהוי חללים ריקים ממידע המתקבלים בתשובה לחיפושים אחר מידע על נושאים הנוגעים לסוגיית החיסונים, כגון אוטיזם (אחד החששות הנפוצים והלא-מוצדקים מחיסון ילדים

הוא אוטיזם) או חשיבות השירותים הניתנים בסניפי טיפת חלב, ומילויים במידע אמיתי; (3) כלים להגנה על שמם הטוב של אנשי מערכת הבריאות לגיל הרך, כגון כלים לניטור ההתייחסויות אליהם ברשתות החברתיות וכלים להתערבות במקרי משבר; (4) כלים להגנה על מערכות המידע האישיות של אישים וארגונים רלוונטיים; (5) הכשרתם של המותקפים הפוטנציאליים לפעילות זהירה ומודעת במערכות המידע האישיות שלהם ובפלטפורמות המידע; (6) כלים לחשיפת זהותם של תוקפים במקרי תקיפה מסוכנים במיוחד.

4. בשלב הרביעי, לאחר יישום אסטרטגיית ההגנה בכלים שפותחו, יהיה צורך לבקר ולוודא את האפקטיביות שלה ולפתח מענה לפערים שיאותרו. בין השאר, נדרש יהיה לערוך משחקי מלחמה ולהפעיל צוותים אדומים, שתכליתם תיקוף הנחות היסוד העומדות בבסיס האסטרטגיה החדשה. כמו כן, צריך יהיה להכשיר את הגורמים האנושיים הרלוונטיים לתרבות הגנתית – מהכשרות ייעודיות לפוליטיקאים, למשל, ועד להכשרות של כלל צרכני המידע במדינה באמצעות מערכת החינוך. הכשרות מסוג זה תעסוקנה, בין השאר, בצריכה ביקורתית של מידע, בהפצה אחראית של מידע ובדרכים להתמודדות עם אירועי פוסט־אמת. את ההכשרות צריך יהיה להתאים לקהל היעד על פי מאפייניו החברתיים־תרבותיים ועל פי היכולת שלו להתמודד עם עולם המידע המודרני. סביר להניח כי ייווצר מתח חריף בין רבים מכלי ההגנה ובין הערכים הדמוקרטיים והליברליים, כגון חופש הביטוי, הזכות לפרטיות ואף חופש העיתונות והעיסוק. לכן נדרשים לפיתוח כלים אלה מומחיות מקצועית, ליווי משפטי ומערכת רגולטורית תומכת.

לסיום, במאמר זה נותחו מאפייניו הייחודיים של עולם המידע המודרני והשפעתם על תופעת הפוסט־אמת, והוצעו קווים מנחים לפיתוח מתודולוגיה, אסטרטגיה וכלים להגנה מפני נזקיה של התופעה. יש להמשיך ולהעמיק את הניתוח ולהגדיר באמצעותו את כלל הנכסים המדינתיים והחברתיים הזקוקים להגנה. מיפוי הנכסים הוא בסיס הכרחי לפיתוח מלא של אסטרטגיות וכלים שיגנו עליהם מפני נזקי התופעה. בנוסף, יש לעצב מערכת של איזונים ובלמים, שתבטיח בקרה ופיקוח הדדיים של השחקנים בעולם הזה – האזרחים, מוסדות המדינה ובעיקר פלטפורמות הטכנולוגיה. ללא מערכת כזאת, שתְּגֻבָּה ברגולציה ובאמצעי אכיפה מדינתיים, וכן תישען על תרבות של אחריות ושל ביקורת אזרחית, ייוותר עולם המידע המודרני מערב פרוע.

הערות

[1] ראו למשל: Matthew d'Ancona, *Post-Truth: The New War on Truth and How to Fight Back*, London, Ebury Press, 2017; Lee McIntyre, *Post-Truth*, London and Cambridge, MIT Press, 2018

[2] יתי ברון ומיכל רויטמן, "ביטחון לאומי בעידן של פוסט אמת ופייק ניוז", המכון למחקרי ביטחון לאומי, אפריל 2019, https://www.inss.org.il/he/publication/national-security-in-the-era-of-post-truth-and-fake-news/

[3] יעל ברהמס, "פילוסופיה של פוסט-אמת", תל אביב, המכון למחקרי ביטחון לאומי, אפריל 2019, https://www.inss.org.il/he/publication/post-truth-philosophy/

[4] ההשוואה בין עולמות המידע של הדמוקרטיות המערביות הליברליות ובין אלו של המשטרים האחרים היא חשובה ומעניינת, אך אין לה מקום במאמר זה. ראו למשל:
Henry Farrell and Bruce Schneier, "Common-Knowledge Attacks on Democracy", Berkman Klein Center for Internet & Societ at Harvard University, Research Publication No. 2018-7, October 2018, https://cyber.harvard.edu/story/2018-10/common-knowledge-attacks-democracy

[5] להגדרה חלופית ראו למשל: משרד ראש הממשלה, "החלטה מספר 3611 של הממשלה מיום 07.08.2011", 7 באוגוסט 2011, https://www.gov.il/he/departments/policies/2011_des3611

[6] הניתוח אינו עוסק במידע מסווג ובמידע פנים־ארגוני מכל סוג שהוא.

[7] ראו למשל ניתוח דומה:
Sarah Dooley and Emma Moore with Alexander Averin, "Route 5 – Change and 21st Century Media", in: *Fake News: A Roadmap*, eds. Jente Althuis and Leonie Haiden, Riga, NATO Strategic Communications Centre of Excellence and King's Centre for Strategic Communications (KCSC), January 2018, pp. 34–40.

[8] מחשוב ענן: שירותים של מערכות ממוחשבות מרוחקות, כגון אחסון או עיבוד, שהמשתמש אינו מנהל אותם בעצמו אלא מתחבר אליהם באמצעות האינטרנט או באמצעות קו תקשורת ייעודי.

[9] ראו למשל: Staffan Truvé, "The Discovery of Fishwrap: A New Social Media Information Operation Methodology", Recorded Future, June 11, 2019, https://www.recordedfuture.com/fishwrap-influence-operation/; Guy Rosen, "Protecting Facebook Live From Abuse and Investing in Manipulated Media Research", Facebook Newsroom, May 14, 2019, https://newsroom.fb.com/news/2019/05/protecting-live-from-abuse

[10] ראו למשל: Aja Romano, "Jordan Peele's simulated Obama PSA is a double-edged warning against fake news", Vox, April 18, 2018, https://www.vox.com/2018/4/18/17252410/jordan-peele-obama-deepfake-buzzfeed

[11] לסקירה על איום הזיוף העמוק, על הטכנולוגיה העומדת בבסיסו, ועל כיוונים אפשריים להתמודדות עימו, ראו למשל:
Jeremy Hsu, "Experts Bet on First Deepfakes Political Scandal", IEEE Spectrum, June 22, 2018, https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/experts-bet-on-first-deepfakes-political-scandal; Robert Chesney and Danielle Keats Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 California Law Review (Forthcoming), https://ssrn.com/abstract=3213954
לדוגמאות לשימוש בזיוף עמוק ראו למשל:
Tero Karras, Samuli Laine and Timo Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks", Cornell University, March 29, 2019, https://arxiv.org/pdf/1812.04948.pdf; Egor Zakharov, Aliaksandra Shysheya, Egor Burkov and Victor Lempitsky, "Few-Shot Adversarial Learning of Realistic Neural Talking Head Models", Cornell University, May 20, 2019, https://arxiv.org/pdf/1905.08233.pdf;
Ohad Fried, Ayush Tewari, Michael Zollhöfer, Adam Finkelstein, Eli Shechtman, Dan B. Goldman, Kyle Genova, Zeyu Jin, Christian Theobalt and Maneesh Agrawala, "Text-Based Editing of Talking-Head Video", Cornell University, June 4, 2019, https://arxiv.org/abs/1906.01524;
OpenAI Team, "Better Language Models and Their Implications", OpenAI blog, February 14, 2019, https://openai.com/blog/better-language-models

[12] ראו למשל תוכניות של הסוכנות הצבאית האמריקאית לפרויקטים מחקריים מתקדמים (DARPA) לפורנזיקת מדיה (MediFor): https://www.darpa.mil/program/media-forensics ולפורנזיקת סמנטיקה (SemaFor): https://www.darpa.mil/news-events/semantic-forensics-proposers-day

[13] ראו למשל: Farhad Manjoo, "Abolish Billionaires", *The New York Times*, February 6, 2019, https://www.nytimes.com/2019/02/06/opinion/abolish-billionaires-tax.html;
גיא רולניק, "שלוש הערות על פשיטת הרגל של התחרות בכלכלה המודרנית", דה-מרקר, 20 ביולי 2018, https://www.themarker.com/markerweek/thisweek/.premium-1.6292392

[14] ראו למשל: Massimo Stellaa, Emilio Ferrarab and Manlio De Domenicoa, "Bots Increase Exposure to Negative and Inflammatory Content in Online Social Systems", *PNAS*, Vol. 115, No. 49, pp. 12435–12440, November 20, 2018, https://www.pnas.org/content/115/49/12435

[15] לטענת האמינות הנמוכה ראו למשל: Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini and Filippo Menczer, "The Spread of Low-Credibility Content by Social Bots", *Nature Communications*, Vol. 9, No. 4789, November 20, 2018, https://www.nature.com/articles/s41467-018-06930-7

[16] ראו למשל: Soroush Vosoughi, Deb Roy and Sinan Aral, "The Spread of True and False News Online", *Science*, Vol. 359, No. 6380, pp. 1146–1151, March 9, 2018, https://science.sciencemag.org/content/359/6380/1146
לטענה ההפוכה ראו למשל: שתי הגישות מסכימות כי מידע שקרי מופץ ביתר הרחבה מאשר מידע אמיתי, והמחלוקת נסבה רק סביב המקור האנושי או הווירטואלי של ההפצה.
שם; Nahema Marchal, Bence Kollanyi, Lisa-Maria Neudert and Philip N. Howard, "Junk News During the EU Parliamentary Elections: Lessons from a Seven-Language Study of Twitter and Facebook", The Computational Propaganda Project, Oxford Internet Institute, University of Oxford, May 21, 2019, https://comprop.oii.ox.ac.uk/research/eu-elections-memo
תהליכי הפצת מידע שקרי ברשתות החברתיות שונים במידה ניכרת מקמפיין אחד למשנהו, ממדינה למדינה ומשנה לשנה.

[17] Ashlee Vance, "This Tech Bubble is Different", Bloomberg Businessweek, April 11, 2015, https://www.bloomberg.com/news/articles/2011-04-14/this-tech-bubble-is-different

[18] ראו למשל את קהילת ה-Facebook "חיסונים בחירה מושכלת", העוסקת בשאלה האם לחסן ילדים: https://www.facebook.com/groups/VaccineChoiceIL/

[19] ראו למשל: Keith N. Hampotn, Inyoung Shin and Weixu Lu, "Social Media and Political Discussion: When Online Presence Silences Offline Conversation", *Information, Communication & Society*, Vol. 20, No. 7, pp. 1090–1107, August 16, 2016, https://www.tandfonline.com/doi/full/10.1080/1369118X.2016.1218526
לטענה ההפוכה ראו למשל: Elizabeth Dubois and Grant Blank, "The Echo Chamber is Overstated: The Moderating Effect of Political Interest and Diverse Media", *Information, Communication & Society*,

Vol. 21, No. 5, pp. 729–745 January 29, 2018, https://www.tandfonline.com/doi/full/10.1080/1369118X.2018.1428656
גם אם עוצמת ההשפעה של התופעה איננה גבוהה כפי שלעיתים נהוג לומר, הרי שעצם קיומה ממחיש היטב את החידוש שבעולם המידע המודרני. הערה זו כוחה יפה גם עבור תופעת בועות הפילטר.

[20] ראו למשל:
Joseph Cox, "It Took 10 Seconds for Instagram to Push me Into an Anti-Vaxx Rabbit Hole", Motherboard, March 21, 2019, https://www.vice.com/en_us/article/vbwkvm/10-seconds-instagram-recommends-anti-vaxx-vaccine-accounts;
Mark Bergen, "YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant", Bloomberg, April 2, 2019, https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant;
Max Fisher and Amanda Taub, "How YouTube Radicalized Brazil", *The New York Times*, August 11, 2019, https://www.nytimes.com/2019/08/11/world/americas/youtube-brazil.html

[21] Michael Golebiewski and Danah Boyd, "Data Voids: Where Missing Data Can Easily Be Exploited", Data & Society Research Institute, May 11, 2018, https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Data_Voids_Final_3.pdf

[22] Zane Pokorny interview with The grugq, "The grugq Illuminates Influence Operations", Recorded Future, March 25, 2019, https://www.recordedfuture.com/podcast-episode-100

[23] במדינות שונות מאפייני צריכת המידע מעט שונים וגם האיזון בין צרכים חברתיים, מקצועיים או אינטלקטואליים משתנה ממדינה למדינה. עם זאת, בכל הדמוקרטיות המערביות הליברליות שיעור החדירה של הפלטפורמות הטכנולוגיות הדומיננטיות ומידת השימוש בהן לעומת פלטפורמות אחרות בולטים מאוד.
לניתוח צריכת המידע בישראל ראו למשל:
בזק, דו"ח החיים הדיגיטליים 2018,
https://www.bezeq.co.il/media/minisites/doh/%D7%93%D7%95_%D7%97-%D7%91%D7%96%D7%A7-2018-%D7%90%D7%97%D7%A8%D7%95%D7%9F.pdf;
לניתוח צריכת המידע בארה"ב ראו למשל:
Aaron Smith, Skye Toor and Patrick van Kessel, "Many Turn to YouTube for Children's Content, News, How-To Lessons", Pew Research Center: Internet & Technology, November 7, 2018, https://www.pewinternet.org/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons;
Digital TV Research, North America Pay TV Forecasts, April 2019, https://www.digitaltvresearch.com/products/product?id=231

[24] יש כמובן יוצאים מן הכלל, כגון חברת Netflix, המייצרת חלק ניכר מהתוכן שהיא מפיצה, או השקת IGTV של הפלטפורמה Instagram ביוני 2018. ייתכן כי חריגים אלה מעידים כי לאחרונה החלה מגמה שתכליתה לתקן במעט את התופעה הזאת, אך אלה היוצאים מן הכלל המעידים על הכלל.

[25] Legal Information Institute, "Protection for Private Blocking and Screening of Offensive Material", *U.S. Code*, Title 47, Chapter 5, Subchapter II, Part I, Section 230, 1996, https://www.law.cornell.edu/uscode/text/47/230

[26] ראו למשל: Monika Bickert, "Combatting Vaccine Misinformation", Facebook Newsroom, March 7, 2019, https://newsroom.fb.com/news/2019/03/combatting-vaccine-misinformation/; Robert McMillan and Daniela Hernandez, "Pinterest Blocks Vaccination Searches in Move to Control the Conversation", *The Wall Street Journal*, February 20, 2019, https://www.wsj.com/articles/next-front-in-tech-firms-war-on-misinformation-bad-medical-advice-11550658601?mod=hp_lead_pos2; Joseph Cox, "Twitter and YouTube Won't Commit to Ban White Nationalism After Facebook Makes Policy Switch", Motherboard, April 2, 2019, https://www.vice.com/en_us/article/mbzz8x/twitter-youtube-wont-ban-white-nationalism-facebook

[27] thaddeus t. grugq, "Opaque at Both Ends", Medium, April 13, 2019, https://medium.com/@thegrugq/opaque-at-both-ends-bb3e2d6e0d58; Rebecca Lewis, "Alternative Influence", Data & Society, September 18, 2018, https://datasociety.net/output/alternative-influence

[28] המחסור ברגולציה החל לאחרונה לקבל מענה. להלן רשימה חלקית של דוגמאות מהשנתיים האחרונות:
תקנות ה־GDPR (General Data Protection Regulation) של האיחוד האירופי, שנאכפות מאז מאי 2018;
European Commission Press Release Database, "Statement on the Code of Practice Against Disinformation: Commission Asks Online Platforms to Provide More Details on Progress Made", Brussels, February 28, 2019, https://europa.eu/rapid/press-release_STATEMENT-19-1379_en.htm;
תוכנית המכון הלאומי האמריקאי לתקנים ולטכנולוגיה (NIST) ליצירת מסגרת עבודה בתחום הפרטיות https://www.nist.gov/privacy-framework;
European Union, "Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019: On Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC", Official Journal of the European Union, May 17, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.130.01.0092.01.ENG;
EU Parliament Press Release, "Terrorist Content Online Should Be Removed Within One Hour, says EP", April 17, 2019, http://www.europarl.europa.eu/news/en/press-room/20190410IPR37571/terrorist-content-online-should-be-removed-within-one-hour-says-ep;
Singapore Government Bill No. 10/2019, "Protection from Online Falsehoods and Manipulation Bill", April 1, 2019, https://sso.agc.gov.sg/Bills-Supp/10-2019/Published/20190401?DocDate=20190401;
UK Department for Digital, Culture, Media & Sport and Home Office, "Online Harms White Paper", June 26, 2019, https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper;

Parliament of Australia, "Criminal Code Amendment (Sharing of Abhorrent Violent Material) Bill 2019", April 4, 2019, https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bId=s1201;
Government of Canada, "Canada's Digital Charter: Trust in a digital world", https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00108.html;
Mark Zuckerberg, "Four Ideas to Regulate the Internet", Facebook Newsroom, March 30, 2019, https://newsroom.fb.com/news/2019/03/four-ideas-regulate-internet;
Alejandro de la Garza, "'Technology Needs to Be Regulated.' Apple CEO Tim Cook Says No Oversight Has Led To Great Damage To Society", *Time*, April 23, 2019, https://time.com/5574143/technology-needs-to-be-regulated-says-apples-tim-cook/

[29] ראו למשל: התוסף NewsGuard לבדיקת עובדות בדפדפן: https://www.newsguardtech.com/ ; אתר בדיקת העובדות הוותיק Snopes: https://www.snopes.com/
גישה דומה נקטה חברת WhatsApp: Manish Singh, "WhatsApp Explores Using Google to Fight Misinformation", VentureBeat, March 15, 2019, https://venturebeat.com/2019/03/15/whatsapp-explores-using-google-to-fight-misinformation;
Caroline Haskins, "WhatsApp Launches a Tip Line for Misinformation in India Ahead of Elections", Motherboard, April 2, 2019, https://www.vice.com/en_us/article/eveeyk/whatsapp-launches-a-tip-line-for-misinformation-in-india-ahead-of-elections;
Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner and Yejin Choi, "Defending Against Neural Fake News", Cornell University, May 29, 2019, https://arxiv.org/pdf/1905.12616.pdf;
המודל זמין: https://grover.allenai.org/

[30] ראו למשל: Jerry Lu, "The Fake News Arms Race", Hackermoon, July 10, 2018, https://hackernoon.com/the-fake-news-arms-race-448675592803

[31] ראו למשל: Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior From Israel", Facebook Newsroom, May 16, 2019, https://newsroom.fb.com/news/2019/05/removing-coordinated-inauthentic-behavior-from-israel
להדגמת המתודולוגיות הננקטות ראו למשל:
Aric Toler, "The Unintelligent Design of SureFire Intelligence", bellingcat, October 30, 2018, https://www.bellingcat.com/news/americas/2018/10/30/unintelligent-design-surefire-intelligence;
Staffan Truvé, "The Discovery of Fishwrap: A New Social Media Information Operation Methodology", Recorded Future, June 11, 2019, https://www.recordedfuture.com/fishwrap-influence-operation

[32] במאמר זה המושג 'הגנת מידע' כולל הן את אבטחת המידע (Information Security) והן את אבטחת הסייבר (Cybersecurity).

[33] יהושפט הרכבי, **מלחמה ואסטרטגיה**, תל אביב, הוצאת "מערכות"/משרד הביטחון, 1992, עמ' 407.

[34] לניתוח מקביל ומעט שונה ראו: דגנית פייקובסקי ואביתר מתניה, "מבצעי השפעה בסייבר – מאפיינים ותובנות", **המערכה על התודעה: היבטים אסטרטגיים ומודיעיניים**, מזכר 191, המכון למחקרי ביטחון לאומי, מאי 2019, https://www.inss.org.il/he/publication/מבצעי-השפעה-בסייבר-מאפיינים-ותובנות/

[35] גבי סיבוני ועופר אסף, **קווים מנחים לאסטרטגיה לאומית במרחב הסייבר**, מזכר 149, המכון למחקרי ביטחון לאומי, אוקטובר 2015, עמ' 18, https://www.inss.org.il/he/wp-content/uploads/sites/2/systemfiles/memo149.pdf

[36] תפקידו של מרכז הניטור להעריך את מצב האבטחה של מערכות המידע בארגון. מרוכזים בו הטכנולוגיות והאנשים הנדרשים להפעלת צעדי האבטחה הראשונים במקרה של פגיעה. אחד הרכיבים החשובים של מרכז ניטור הוא ה־SIEM (Security Information and Event Management), האוסף, מרכז ומנתח מידע ממערכות ומחיישנים המותקנים ברשתות התקשורת של הארגון.

[37] צוות מומחים המנהל אירועי אבטחה עבור ארגונים ומדינות.

[38] צוות מומחים המוכן לתגובה מיידית במקרה של אירוע אבטחה, ופועל להכלת ההתקפה ולהגבלת הישגיה.

[39] IPS - Intrusion Prevention System: מערכת שמטרתה לזהות פעילות עוינת ולפעול מיידית להגבלתה או לעצירתה.

[40] DDS – Denial of Service Defense System: מערכת הגנה מפני תקיפות DoS, מעין IPS ייעודי לתקיפה זו.

[41] WAF - Web Application Firewall: מערכת שמטרתה לסנן, לנטר ולחסום תעבורה אל יישומי רשת (Web Applications) ומהם.

[42] ראו למשל: אסף חזני, אורי גולדברג ויונתן רוזן, "האין פה נביא להי עוד? על הפעלת "צוות אדום"", **מערכות**, גיליון 442, אפריל 2012, עמ' 67-60, http://maarachot.idf.il/PDF/FILES/7/113077.pdf

[43] סוגיית ההגנה על תהליך הבחירות הדמוקרטי מעלה את השאלה: מהו הגבול הלגיטימי של ההגנה ומה הגבול הלגיטימי של המשחק הפוליטי? לכן היא דוגמה טובה למורכבות הפיתוח של אסטרטגיות הגנה ולחשיבות הגבוהה של פיתוח מושכל שלהן.

[44] ברוח זו פיתחה חברת Microsoft את תוכנית ההגנה על הדמוקרטיה (Defending Democracy Program): https://blogs.microsoft.com/on-the-issues/2018/04/13/announcing-the-defending-democracy-program/
התוכנית כוללת מרכיבים של בדיקת עובדות, הגנה על קמפיינים פוליטיים ועל ארגונים התומכים בדמוקרטיה, שימוש בטוח ושקוף במערכות הצבעה וכדומה.

[45] ראו למשל: Monika Bickert, "Combatting Vaccine Misinformation", Facebook Newsroom, March 7, 2019, https://newsroom.fb.com/news/2019/03/combatting-vaccine-misinformation