# קובץ פרסומים בנושא סייבר ובינה מלאכותית
# דצמבר 2020

**תרגום ועריכה: עמרי וקסלר ואופיר בראל**

סדנת יובל נאמן למדע, טכנולוגיה וביטחון
אוניברסיטת תל-אביב

ICRC Blavatnik Interdisciplinary Cyber Research Center
TEL AVIV UNIVERSITY אוניברסיטת תל אביב
Yuval Ne'eman Workshop for Science, Technology and Security Tel Aviv University

# תוכן העניינים


**דבר ראש הסדנה..... 5**

**א. העבר, ההווה והעתיד של האסטרטגיה וכוחות הסייבר של רוסיה ..... 6**

הקדמה ..... 6

מבוא ..... 6

הדוקטרינה והאסטרטגיה של רוסיה בתחום אבטחת סייבר ..... 7

שינויים במשמעות מושג הלוחמה בתפיסה הרוסית ..... 7

השקפותיה הרשמיות של רוסיה על לוחמת מידע ..... 8

איומים מרכזיים במרחב המידע ..... 9

תגובתה של רוסיה לאיומים במרחב המידע: עמדה הגנתית ומשותפת ..... 10

אבטחת הסייבר מחוץ לדוקטרינה של רוסיה: הערך של נשק סייבר ..... 10

התפתחות מבצעי הסייבר והמידע של גופי המודיעין של רוסיה ..... 11

השנים הראשונות למבצעי הסייבר של רוסיה: ה-FSB וגורמים א-מדינתיים ..... 11

חשיפתו של ה-GRU וכניסתו לתחום מבצעי לוחמת המידע ..... 12

התרבות הארגונית של ה-GRU וניהול מבצעי סייבר טכניים ..... 14

השלכות מבצעי הסייבר והמידע הצבאיים של רוסיה על פעילות עתידית
בחסות המדינה ..... 15

מסקנות ..... 16

סיכום ..... 17

**ב. השימוש בבינה מלאכותית בקרב צבא סין ..... 18**

הקדמה ..... 18

רקע ..... 18

בינה מלאכותית צפויה לשפר את האסטרטגיה הצבאית הקיימת של סין ..... 19

עקרון ה-Intelligentized Warfare ..... 19

שישה עקרונות של עקרון ה-Intelligentized Warfare ..... 20

הגישה הסינית לפיתוח יכולות בינה מלאכותית ..... 22

תובנות ומבט לעתיד ..... 23

סיכום ..... 23

**ג. דיפ-פייק ומדיה מלאכותית במערכת הפיננסית: הערכת תרחישי האיום ..... 24**

הקדמה ..... 24

מבוא ..... 24

מתודולוגיה ..... 25

סקירת תרחישים ..... 25

מדיה מלאכותית בהיקף מצומצם ובהיקף רחב ..... 25

שלוש טכניקות מרכזיות המשמשות למטרות זדוניות ..... 26

תרחישים המתמקדים באינדיבידואלים ..... 27

תרחישים המתמקדים בחברות ..... 30

תרחישים המתמקדים במגזר הפיננסי ..... 34

תרחישים המתמקדים בבנקים מרכזיים ובמוסדות רגולציה פיננסיים ..... 34
יישומי מדיניות ..... 36
האיום הכולל ליציבות הפיננסית ולמערכת המאקרו הכלכלית ..... 36
שימוש במדיה מלאכותית לעומת השימוש באמצעים זדוניים אחרים ..... 36
הפצת נרטיבים והעצמת משברים קיימים ..... 37
השוואה לקטעי Deepfake פוליטיים ..... 38
תגובות מגוונות מצד בעלי העניין ..... 39
מסקנות ..... 40
סיכום ..... 41

**ד. מדיניות החוץ המפוזרת והבלתי מתואמת של קנדה בתחום הסייבר:**

**קריאה להבהרה ..... 42**
הקדמה ..... 42
מבוא ..... 42
מדיניות חוץ בתחום סייבר: מדוע היא חשובה? ..... 43
התפתחויות אחרונות בתחומי המדיניות והחקיקה ..... 43
פעילות בין-לאומית, דיפלומטיה ומוקד מגדרי ..... 44
אסטרטגיית הגנת סייבר צבאית ומעורפלת ..... 45
פערים בעקביות, בבהירות ובתיאום ..... 46
הצורך המתמשך בשורה של עקרונות מוצהרים בתחום הסייבר ..... 46
סיכום ..... 47



# דבר ראש הסדנה

קובץ פרסומים זה מכיל ארבעה מחקרים מטעם מכוני מחקר, אוניברסיטאות זרות, ומומחים הדנים בממשק שבין הטכנולוגיה והביטחון, ובדגש על תחומי הסייבר והבינה המלאכותית. תחום הסייבר והשימוש לרעה בטכנולוגיות תקשורת ומידע הפכו בעשור האחרון לאיום משמעותי על כלכלתן וביטחונן הלאומי של מדינות, ושילובו בממד התודעתי משפיע גם על תחומי הפרט והחברה, כגון דעת קהל, תודעת ההמון, העצמת מחלוקות וערעור אמון הציבור ברשויות. לאיומים אלו, נוסף האיום הנשקף משימוש לרעה בטכנולוגיות הבינה המלאכותית ולמידת המכונה.

קובץ זה כולל 4 מחקרים בנושאים הבאים:

התפתחותה של לוחמת הסייבר והמידע של רוסיה;

דיונים בנושא תפיסת ההפעלה והשימוש בבינה מלאכותית בקרב צבא סין;

ניתוח איומי המדיה המלאכותית, ובראשם טכניקת הזיוף "דיפ-פייק", על המגזר הפיננסי.

תחום הסייבר במדיניות החוץ של קנדה.

מחקרים אלו מדגימים את החשיבות של הבנת הפן הטכני של הסייבר ושל טכנולוגיות חדשות, לצד הבנת הפן האסטרטגי והמדיני, קרי הבנת תפיסת ההפעלה ומדיניות החוץ והביטחון של מדינות, בראשן המעצמות הגדולות. חשיבות זו נובעת מהיכולת להבין את המתרחש בעולם ואת המגמות החדשות ומכך לגזור רעיונות חדשים למדיניות, לדרכי פעולה אפקטיביות, למענים לאתגרים הנובעים מהשימוש לרעה בטכנולוגיה, לשיתופי פעולה, להסדרים בין-לאומיים, וכיוצא בזה.

סדנת יובל נאמן למדע, טכנולוגיה וביטחון שמה לה למטרה לחקור ולהנגיש לציבור ולמקבלי ההחלטות את נושא הסייבר על כל היבטיו, ובכלל זאת, את הממשק שבין הסייבר לממד התודעתי ואת הממשק שבין הסייבר לבין טכנולוגיות חדשות, בראשן הבינה המלאכותית. המחקרים סוכמו ותורגמו לעברית על ידי צוות הסדנה.

קריאה נעימה,

**פרופסור יצחק בן ישראל**
ראש סדנת יובל נאמן למדע, טכנולוגיה וביטחון
ראש המרכז הרב-תחומי לחקר הסייבר ע"ש בלווטניק

# א. העבר, ההווה והעתיד של האסטרטגיה וכוחות הסייבר של רוסיה

**Bilyana Lilly & Joe Cheravitch**[1]

## הקדמה

מאמר זה נועד לתאר את התפתחותה של לוחמת המידע ולוחמת הסייבר הרוסיות, בעיקר לאחר תום המלחמה הקרה. זאת, באמצעות תיאור התפתחויות בגישות התיאורטיות ובאמצעות תיאור התפתחות הכוחות העיקריים האחראיים לביצוען. לטענת המחברים, מקבלי ההחלטות במערב צריכים להכיר התפתחויות אלו, על מנת להתמודד באופן יעיל עם לוחמת המידע ולוחמת הסייבר מצד רוסיה.

בחלקו הראשון של המסמך, מתארים המחברים את מקורותיה של התיאוריה הרוסית סביב לוחמת המידע ומתארים את ההתפתחויות שחלו בה. חלקו השני דן בתיאור הגופים האחראים לביצוע מבצעי מידע וסייבר וכן מתארים את התפתחות יכולותיהם ומסבירים את השינויים שנעשו בחלוקת האחריויות ביניהם, החל מראשית שנות ה-2000. בחלק האחרון, המחברים מציגים מסקנות לגבי המשך פעילותה של רוסיה במרחב הסייבר.

## מבוא

מבצעי סייבר שמיוחסים לרוסיה מתקיימים בהשפעת מספר גורמים, הכוללים שיקולים גיאו-פוליטיים רחבים; התרבות הארגונית של צבא רוסיה, המודיעין וההנהגה הפוליטית; והתפתחויות בגישתה של רוסיה לעימותים א-סימטריים בין מדינות. על מנת להבין את המניעים והמגבלות שבבסיס מבצעי הסייבר והמידע של רוסיה, מקבלי ההחלטות צריכים ללמוד באופן יסודי את המדיניות והדוקטרינה הקיימות ולהבין כיצד הן התפתחו מהשנים הראשונות של העידן הפוסט-סובייטי ועד לימינו. כמו כן, עליהם להבין טוב יותר את השחקנים האחראיים לביצוע מתקפות סייבר ומבצעי השפעה דיגיטליים. הדבר מצריך מחקר של מקורות ומסמכים רשמיים שפורסמו וחקירה מדוקדקת של הגורמים המעורבים במתקפות ובמבצעי סייבר. חקירה מסוג זה אפשרית כעת בעקבות חשיפתם של מבצעים שונים, כגון המעורבות בבחירות לנשיאות ארה"ב ב-2016. מבצעים אלו שנחשפו לציבור הובילו להיווצרותו של מידע ציבורי בהיקף חסר תקדים על אישים ויחידות בצבא ובגופי המודיעין של רוסיה. חקירה מסוג זה יכולה ללכד את הקהילה הבין-לאומית לטובת התמודדות משותפת כנגד מבצעי סייבר דומים עתידיים, וכן לסייע למנהיגים לקבל החלטות בנוגע לדיפלומטיית סייבר ולרעיון ההרתעה במרחב הסייבר.

1 Bilyana Lilly and Joe Cheratovitch. The Past, Present, and Future of Russia's Cyber Strategy and Forces. 12th International Conference on Cyber Conflict. NATO CCDCOE Publications. https://ccdcoe.org/uploads/2020/05/CyCon_2020_8_Lilly_Cheravitch.pdf



## הדוקטרינה והאסטרטגיה של רוסיה בתחום אבטחת סייבר

שינויים במשמעות מושג הלוחמה בתפיסה הרוסית

במהלך שני העשורים האחרונים, ההנהגה הצבאית והפוליטית של רוסיה ביצעה התאמות ושינויים יסודיים בתפיסתה את מושג הלוחמה ואת תפקידם של מבצעי סייבר. חוקרים שונים הציגו פרסומים חשובים, בהם הם ניתחו ניואנסים שונים בהתפתחויות הללו בצד הרוסי. חלק זה במאמר מוקדש לתיאור אותה ספרות ובאמצעותו ניתן להבין את כיוונה של דוקטרינת הסייבר של רוסיה ושל המחקר בנושא וכן את ההנחות העומדות בבסיסם. הדבר מניח את היסוד לניתוח האבולוציה של יחידות הסייבר של רוסיה ומדגיש את ההשוואה בין הדוקטרינה הקיימת והספרות המדעית הצבאית ברוסיה מצד אחד, לבין התרבות הארגונית של גופי הסייבר המרכזיים וטבעם של מבצעי הסייבר של רוסיה מצד שני.

הגדרת מושג המלחמה על ידי רוסיה עברה שינוי במרוצת הזמן: מהסכמה על כך שהמלחמה היא אלימות חמושה להסכמה שהמלחמה כוללת גם מיזוג מדויק בין אלימות חמושה לאמצעים לא צבאיים. הבנת הניואנסים המתפתחים בנקודת המבט הצבאית הרוסית היא חיונית עבור מקבלי החלטות במערב, מאחר וההבדלים בין התפיסה הרוסית למערבית בנושא הלוחמה כוללים גם הבדלים בהבנת אלמנטים שונים במדיניות חוץ כגון איתותים. הבדלים כאלו עשויים להכיל משמעויות והשלכות רחבות להרתעת רוסיה ולהבנת הקווים האדומים שלה. הבדלים אלו חיוניים לטובת ניסוח אסטרטגיה ארוכת טווח, שמתייחסת למניעים של ההתנהגות הרוסית.

חוקרים מערביים ורוסיים שניסו לתאר את אופייה המשתנה של תפיסת המלחמה של רוסיה השתמשו במונחים כגון 'לוחמה היברידית', 'לוחמה מהדור הבא', ' לוחמה פוליטית', 'כפייה בין-ממדית' (cross-domain coercion) ו-'טקטיקות מהתחום האפור' (grey-zone tactics). קיימים הבדלים במונחים השונים, אך כולם מייצגים את התפיסה הרוסית, לפיה לוחמה כוללת גם אמצעים לא צבאיים, שניתן להשתמש בהם לפני השימוש בכוחות צבא גלויים או במקומם.

הדיונים בנושא השימוש באמצעים לא צבאיים במסגרת הלוחמה הרוסית אינם תופעה חדשה, אולם דיונים אלו אומצו על ידי חלקים נרחבים בממסד הצבאי הרוסי רק בשנים האחרונות. תיאורטיקנים צבאיים רוסים דנו בתועלת שבהפעלת אמצעים מסוג זה עוד לפני המהפכה הבולשביקית. אמצעי לוחמה פסיכולוגית הופעלו כבר במהלך פלישת נפוליאון לרוסיה וגם הצבא האדום ראה בשנותיו הראשונות את התועלת שבהפעלת לוחמה פסיכולוגית ביצירת לחץ על האוכלוסייה שמאחורי החזית. התיאורטיקן הצבאי יבגני מסנר (Evgeny Messner), כתב בשנים שלפני המהפכה הבולשביקית על טשטוש הגבולות בין מלחמה לשלום ועל השימוש במבצעי מידע לשם השפעה על הלכידות החברתית. תובנות מסוג זה השתקפו גם במחקרים של חוקרים רוסיים מאוחרים יותר, שהציגו את תפיסתם על אופייה המשתנה של הלחימה החל משנות ה-90. למרות ההבדלים באמצעים, הבסיס האסטרטגי למתקפות הסייבר הצבאיות של רוסיה הונח לפני יותר ממאה שנים.

למרות מספר המאמרים הגדול על השימוש באמצעים לא צבאיים משנות ה-90 וה-2000, חשיבתה של הצמרת הצבאית הרוסית השתנתה בעיקר בין תחילת שנות ה-2000 לבין הפלישה לאוקראינה ב-2014. בתקופה זו נוצר קונצנזוס בין ההנהגה הרוסית הבכירה לבין תיאורטיקנים

צבאיים, לפיו הגבול שבין מלחמה לשלום היטשטש ואמצעי לוחמה לא אלימים יכולים להיות יעילים מספיק על מנת שהפעלתם תוגדר כלוחמה הלכה למעשה.

ראש המטה הכללי של רוסיה, גנרל ולרי גרסימוב (Valery Gerasimov), כתב בעבר שכללי המלחמה השתנו ושמרידות כגון אלו שהודגמו באביב הערבי מציגות מודל אפשרי של מלחמות העתיד, בהן ייעשה שימוש רב באמצעים פוליטיים, כלכליים ואמצעים לא אלימים אחרים. חוקרים צבאיים שונים הרחיבו את הטיעון הזה וטענו שהצד התוקפני בעימות יהיה הראשון להשתמש באמצעים לא צבאיים, כגון טכנולוגיות מידע שיכוונו כלפי מוסדות ציבוריים במדינת היעד, כגון גופי התקשורת, מוסדות תרבות, מוסדות דת, ארגונים א-ממשלתיים ותנועות בעלות מימון חיצוני זר. בשנת 2019, גרסימוב הדגיש את השימוש המתפתח והמתואם באמצעים צבאיים ולא צבאיים והסביר על עליונותם של אמצעים לא צבאיים על פני אלו הצבאיים, בהם יש להשתמש רק כאשר הצעדים הלא צבאיים לא משיגים את המטרה.

עדכונים אחרונים במסמכים האסטרטגיים החשובים ברוסיה משקפים גם הם התפתחויות בתפיסת המלחמה. הדוקטרינה הצבאית של רוסיה משנת 2010 הצהירה ששילוב בין אמצעים צבאיים ללא צבאיים אופייני לעימותים הצבאיים המודרניים. הדוקטרינה המעודכנת משנת 2014 חיזקה את ההנחה הזו ואף הציגה אותה כמאפיין הראשון של הלוחמה הצבאית המודרנית. המסמך המנחה את מדיניות החוץ משנת 2013, מנה אמצעים מדעיים, כלכליים וטכנולוגיים, כגון אמצעים בתחום ה-IT כבעלי חשיבות להשפעה על מדיניותה ועל המצב הפוליטי של מדינה יריבה. מסמכים אלו מציגים את השינויים התפיסתיים בהגדרתה של רוסיה את מושג המלחמה המודרנית.

## השקפותיה הרשמיות של רוסיה על לוחמת מידע

הצגת המתווה לתפיסה הרוסית את הלוחמה המודרנית חשובה להבנת אסטרטגיית הסייבר של רוסיה, מאחר ותפיסתה של רוסיה על אבטחת סייבר נעוצה בהבנתה את האופי המשתנה של המלחמה ומעוצבת על ידי הגדרתה את לוחמת המידע. אבטחת סייבר נתפסת בדיונים ברוסיה כמונח מערבי, בעוד שהמונח הרוסי המקביל לה הוא אבטחת מידע. חוקרים ומסמכים רשמיים מציגים הגדרות שונות במעט של לוחמת מידע ואבטחת מידע, אולם התפיסה הבסיסית היא שאבטחת מידע היא מרכיב של לוחמת מידע, מונח הכולל היבט טכני והיבט פסיכולוגי או קוגניטיבי. לוחמת מידע היא חלק אינטגרלי מעימותים בין מדינות ומטרתה לייצר עליונות בתחום המידע על פני היריב, באמצעות שימוש באמצעים טכניים ופסיכולוגיים. לפי תפיסה זו, מבצעי סייבר הם האמצעי בו משתמשים על מנת לשלוט במרחב המידע, שנחשב למרחב לחימה בפני עצמו. מסמך שפורסם על ידי משרד ההגנה של רוסיה משנת 2011 בנושא פעילותו של צבא רוסיה במרחב המידע, מספק הגדרה ברורה ללוחמת מידע. על פי מסמך זה, לוחמת מידע היא:

*"עימות במרחב המידע בין שתי מדינות או יותר, שמטרתו להזיק למערכות מידע, תהליכי מידע ומשאבי מידע במטרה לערער את המערכות הפוליטיות, החברתיות והכלכליות ולגרום למניפולציה פסיכולוגית נרחבת שמטרתה לערער את החברה והמדינה ולאלץ את המדינה לקבל החלטות שלטובת היריב."*

הגדרה זו מדגישה את שני האלמנטים העיקריים של לוחמת מידע: האלמנט הטכני, הנוגע



לתשתיות מידע, והאלמנט הפסיכולוגי, שכולל השפעה קוגניטיבית על הציבור ומקבלי ההחלטות, כדי להשפיע על החלטותיהם ועל רצונם להילחם.

לדברי גרסימוב, ללא גבולות לאומיים ברורים, מרחב המידע מאפשר השפעה חשאית לא רק על תשתיות חיוניות במדינה, אלא גם על האוכלוסייה שלה, מה שיכול להשפיע באופן ישיר על ביטחונה הלאומי. לדבריו, מאפיינים אלו הופכים את הלימוד של ניהול וביצוע פעולות במרחב המידע למשימה החשובה ביותר של המדע הצבאי. בהתחשב באופיים הרב-ממדי והלא שגרתי, לוחמת מידע ומבצעי סייבר יכולים להיות מיושמים לפני ההכרזה הרשמית על מלחמה וניתן להוציאם לפועל כדי להשיג מטרות פוליטיות, ללא הצורך להשתמש בכוח צבאי.

## איומים מרכזיים במרחב המידע

איומים הנובעים ממידע קיבלו חשיבות רבה בדוקטרינה הרוסית מאז תחילת המאה ה-21. בהתאם למסורת הסובייטית, המתארת את רוסיה כמבצר מכותר המתמודד עם איומים חיצוניים ופנימיים כאחד, רוסיה רואה את המאבק במרחב המידע כמאבק אינסופי. מסמך הביטחון הלאומי משנת 2000 (National Security Concept) הדגיש שהביטחון הלאומי של רוסיה מאוים במרחב המידע על ידי מדינות שמנסות לשלוט בו, בזמן שהן מפתחות את עקרונותיהן למלחמות מידע. המסמך הציג הבנה הוליסטית של לוחמת מידע באמצעות התמקדות באיומים, הנוגעים להיבטים הטכניים והפסיכולוגיים של לוחמת מידע. הדוקטרינה הצבאית משנת 2010 מייצגת שינוי בהבנת האיומים הלאומיים ומציינת את התפקיד ההולך וגובר של לוחמת מידע כמאפיין של הלוחמה המודרנית והחשיבות עבור צבא רוסיה לפתח אמצעים וכוחות המוקדשים ללוחמת מידע.

מסמכי דוקטרינת אבטחת המידע של רוסיה מהשנים 2000 ו-2016 הגדירו גם את השקפתה של רוסיה על תפקידם של איומי מידע בלוחמה המודרנית. מסמך הדוקטרינה משנת 2000 סיפק הגדרה רחבה של מרחב המידע, לפיה מדובר בשילוב של מידע, של תשתיות מידע ושל גורמים המעורבים באיסוף וביצירת מידע, בהפצתו ובשימוש בו. כמו כן, מדובר במערכת להסדרת עניינים הקשורים ליחסי ציבור. על בסיס הגדרה זו, המבטאת את המרכיבים הטכניים והקוגניטיביים של מרחב המידע, ניתן לתאר טווח רחב של איומים על אבטחת מידע הכוללים איומים טכניים, כגון איומים על שירותי מידע ותקשורת ואיומים על הלכידות החברתית, כגון הפגיעה במה שהוגדר כ-"הפוטנציאל הרוחני והיצירתי ובמורל של העם הרוסי".

בשנת 2013, פרסמה מועצת הביטחון של רוסיה את מסמך העקרונות הבסיסיים לאבטחת מידע בין-לאומית. המסמך הדגיש את האיומים שהודגשו לעיל ותיאר את טכנולוגיות המידע ככלי נשק, שניתן להשתמש בו למטרות פוליטיות וצבאיות ובמטרה לפגוע בריבונותה של מדינה ובשלמותה הטריטוריאלית. דוקטרינת אבטחת המידע המעודכנת משנת 2016 הלכה בעקבות קודמותיה והדגישה מחדש את האיומים ההולכים וגוברים לרוסיה במרחב המידע מצד יריבים שונים. הדוקטרינה תיארה את האיומים הקוגניטיביים במרחב המידע, שמקורם בגורמים זרים, ואת השפעתם על הערכים החברתיים ועל היציבות החברתית. מסמכים אלו ממחישים את האמונה שהיערכותה של רוסיה במרחב המידע היא תגובה לאיומים המאלצים אותה להגן על עצמה.

## תגובתה של רוסיה לאיומים במרחב המידע: עמדה הגנתית ומשותפת

האסטרטגיה הרשמית הגלויה של רוסיה לניהול איומים במרחב המידע היא רבת פנים ורחבה כמו האיומים עצמם. עם זאת, האסטרטגיה משמיטה כל אזכור ליכולות ואמצעים התקפיים או לפעולות תוקפניות. הממשלה מציינת במסמכים רשמיים רשימת יעדים בתחום המדיניות, המתארים עמדה הגנתית הנעזרת בשיתופי פעולה, ומטרתם למנוע או לבלום תוקפנות במרחב הסייבר באמצעות מסגרות משפטיות ורגולטיביות ובאמצעות שיתופי פעולה. קווי מדיניות אלו כוללים פיתוח ויישום של אמצעים רגולטוריים משפטיים, המבססים את אחריותם של יחידים וארגונים לגישה, להעברה או לשימוש בלתי חוקי במידע, וכן צעדים לחיזוק האבטחה של מערכות מידע חיוניות. ההמלצות בנושא המדיניות הבין-לאומית של רוסיה קוראות להקמת מערכת בין-לאומית לאבטחת מידע ולהקמת מנגנונים לשיתוף פעולה בין-לאומי במאבק נגד השימוש הזדוני במידע ובטכנולוגיות תקשורת למטרות טרור.

## אבטחת הסייבר מחוץ לדוקטרינה של רוסיה: הערך של נשק סייבר

לרוסיה אין דוקטרינה מפורשת בנושא אבטחת הסייבר, והמסמכים הרשמיים, העוסקים בעמדתה של רוסיה במרחב המידע, מציגים בעיקר עמדה הגנתית. עם זאת, התיאוריה הצבאית מספקת תובנות שימושיות נוספות על תפקידן של יכולות סייבר, בדגש על יכולות התקפיות, בתפיסתה של רוסיה בנושא עימותים. חוקרים צבאיים דנו בגמישות ההפעלה והיעילות של אמצעי סייבר התקפיים וכן בעלותם הנמוכה יחסית לאמצעים צבאיים קונבנציונליים. יכולות סייבר התקפיות הולמות את הרעיון של לוחמת מידע, מאחר והגבולות שבין מלחמה ושלום מטשטשים במרחב הסייבר. הטשטוש בא לידי ביטוי ביכולת לגרום נזק ליריב מבלי לחצות את סף העימות המזוין או להכריז על מלחמה. בהיעדר מסגרת משפטית ברורה המאפשרת להעמיד לדין גורמים שביצעו מתקפות סייבר, כל גורם מסוגל להוציא לפועל מתקפות סייבר הרסניות מכל מקום ובכך להחליש את יכולתו של האויב להגן על עצמו ולהגיב.

יתרון צבאי נוסף של נשק סייבר הוא היכולת להשיג עליונות בתחום המידע מבלי לחצות גבולות בין מדינות או לבסס נוכחות פיזית בטריטוריה של האויב. מאפיין נוסף, החיוני ביותר עבור רוסיה, הוא השימוש ביכולות סייבר התקפיות במסגרת לוחמה א-סימטרית, שיכולה לסייע למדינה, בעלת נחיתות טכנולוגית וכלכלית לנטרל יריב החזק ממנה.

חוקרים צבאיים רוסים הדגישו את היכולות ההרסניות ואת הוורסטיליות של יכולות סייבר התקפיות, ואת היכולת להשתמש בהם כנגד מטרות צבאיות, אזרחיות וממשלתיות. בהתאם להבנתה של רוסיה את נושא לוחמת המידע, חוקרים טוענים שיכולות סייבר התקפיות יכולות לפגוע בתשתיות היריב וכן להשפיע עליו פסיכולוגית. חוקרים נוספים, שעבדו עבור משרד ההגנה של רוסיה, הדגישו את יכולתן של יכולות סייבר התקפיות לגרום לנזק בתשתיות אנרגיה ותחבורה ואף לגרום למשבר פיננסי. מאפיין נוסף של יכולות סייבר התקפיות הוא עלותן הנמוכה יחסית לפיתוח ולשימוש בנשק קונבנציונלי.

על פי מחקר שפורסם ברוסיה ב-2012, יחידה המונה עד ל-600 "לוחמי מידע" תוכל להסב נזק נרחב ומתמשך לתשתיות המידע של מעצמות כגון ארה"ב. על פי חישובי החוקרים, אימון היחידה יארך כשנתיים ועלות ההכשרות לצד ביצוע המתקפה תעמוד על לא יותר מ-100 מיליון דולר.



על אף תיאור גישתה של רוסיה ללוחמת מידע כהגנתית בלבד במסמכים הרשמיים, הספרות הצבאית ברוסיה מתארת דיון ער על הפיתוח והיישום של יכולות סייבר התקפיות והגנתיות כאחד. לא מן הנמנע כי אימוץ אסטרטגיות סייבר פרו-אקטיביות מצד מדינות המערב, כגון אסטרטגיית ה-Persistent Engagement של פיקוד הסייבר האמריקני, יספק לרוסיה הצדקה להוסיף את ממד הסייבר ההתקפי לדוקטרינת לוחמת המידע שלה. מצד שני, השמטת ממד הסייבר ההתקפי ממסמכים רשמיים מספק לרוסיה מרחב הכחשה המשרת את הנרטיב לפיו פעילותה של רוסיה במרחב הסייבר נובעת מהגנה עצמית ומספק לרוסיה הצדקה להמשיך ולהשקיע בחידוש כוחותיה המזוינים.

# התפתחות מבצעי הסייבר והמידע של גופי המודיעין של רוסיה

## השנים הראשונות למבצעי הסייבר של רוסיה: ה-FSB וגורמים א-מדינתיים

בעידן שלאחר התמוטטות הגוש הסובייטי, שירות הביטחון הפדראלי (FSB) היה הגורם החשוב ביותר בכל הקשור למבצעי סייבר שכוונו כלפי חוץ. בשנות ה-90 ובתחילת שנות ה-2000, ה-FSB פיתח קשרים שאפשרו לו לשתף פעולה, או לכפות על מומחים וחוקרי סייבר פרטיים להוציא לפועל מבצעי סייבר עבורו. השימוש בהאקרים שלא גויסו לשירות המדינה סייע לרוסיה להתמודד עם פערים בכוח אדם מיומן בתחום הסייבר. לדברי גורם פנימי במרכז לאבטחת המידע (CIS)[2] של ה-FSB, המרכז העסיק עברייני סייבר כדי למלא מכסות גיוס עבור מומחי סייבר.

גורם מרכזי שסייע ל-FSB בפיתוח יכולות סייבר התקפיות בשנותיו הראשונות היה התפרקות הסוכנות הפדראלית לתקשורת ולמידע ממשלתיים (FAPSI)[3] בשנת 2003, וסיפוח יכולותיה ואנשיה ל-FSB. גורם נוסף היה סיוע נרחב במחקר טכנולוגי שקיבל הארגון מטעם מרכז ה-Kvant למחקר מדעי.[4]

לדברי העיתונאי העוסק בענייני אבטחת סייבר מהמגזין הטכנולוגי Wired, אנדי גרינברג (Andy Greenberg), בשנים האלו דָבַק המודיעין הצבאי של רוסיה (GRU) בעבודת איסוף מודיעין מסורתית לתמיכה במבצעי צבא רוסיה, והיווה גורם משני בלבד בהשוואה לפעילותו של ה-FSB בניהול מבצעי הסייבר של רוסיה כנגד גיאורגיה ואסטוניה.

למשך תקופה, מבצעי הסייבר ההתקפיים של רוסיה בוצעו במסגרת מבנה גמיש הכולל מגוון רחב של שחקנים מדינתיים וא-מדינתיים. דוגמה לכך ניתן לראות בחטיבת הרשת הסיבירית,[5] שהוקמה על ידי סטודנטים באוניברסיטת טומסק (Tomsk) שקיבלה חסות חוקית מה-FSB לבצע מתקפות DDoS כנגד אתרים צ'צ'ניים בתחילת שנות ה-2000. סדרת מתקפות הסייבר על אסטוניה בשנת 2007 בוצעה על ידי קבוצות שונות ולא מוכרות של האקרים שפעלו בחסות ממשלת רוסיה. שנה לאחר מכן, נתגלתה נוזקה שהשימוש בה יוחס ל-FSB ברשתות מחלקת ההגנה האמריקנית.

עם זאת, לא מן הנמנע כי דווקא מבצעי הסייבר המוקדמים של ה-FSB הובילו לבסוף להתפתחות

2 Center for Information Security
3 Federal Agency of Government Communications and Information
4 Kvant Scientific Research Institute
5 Siberian Network Brigade

תכנית הסייבר של צבא רוסיה, שנזנחה בעידן הפוסט סובייטי כתוצאה מפערי כוח אדם ומחסור בתקציב. הצלחתן של מתקפות הסייבר על אסטוניה וגיאורגיה והפריצה לרשתות מחלקת ההגנה האמריקנית על ידי רוסיה ומדינות נוספות הובילו את ארה"ב להרחיב את תכנית הסייבר הצבאית של מחלקת ההגנה, שגולת הכותרת שלה הייתה הקמת פיקוד הסייבר בשנת 2009. אירועים נוספים שהתרחשו במקביל להקמת הפיקוד, כגון מתקפת ה-Stuxnet שפגעה במתקני הגרעין של איראן, הובילו לחשש בקרב גורמי צבא וביטחון ברוסיה מפני הדומיננטיות האמריקנית במרחב הסייבר. ההשקעה האמריקנית הגוברת ביכולות הסייבר הצבאיות של ארה"ב הובילה את ממשלת רוסיה להכפיל את השקעתה ביכולות הסייבר של הצבא.

תהליכי משא ומתן בין ארה"ב ורוסיה בנושא הרגולציה על פיתוח יכולות סייבר מתקדמות נתפסו במרכז חוסר ההסכמה על נושאים כללים יותר, כגון משילות האינטרנט העולמית וכשלו. כישלון זה פגע בסיכוי לקיים בקרת נשק סייבר בין רוסיה ליריבותיה. בעוד שבשנים הראשונות, קואליציות אד-הוק של גורמים א-מדינתיים שפעלו במרחב הסייבר מחוץ לשליטתה הישירה של הממשלה, מילאו את שאיפותיה המוקדמות של רוסיה בתחום הסייבר, הפערים המתרחבים בין רוסיה ליריבותיה, ביניהן ברית נאט"ו, החמירו את החששות הקיימים בנוגע לאי-מוכנותה למלחמת מידע בלתי נמנעת עם המערב.

## חשיפתו של ה-GRU וכניסתו לתחום מבצעי לוחמת המידע

באמצע שנת 2013 השיק שר ההגנה של רוסיה, סרגיי שויגו (Sergey Shoygu), קמפיין גיוס נרחב עבור מתכנתים למילוי שורות היחידות הטכנולוגיות בצבא, שיקדמו את המחקר והפיתוח הצבאיים בתחומי מבצעי הסייבר, מודיעין האותות (סיגינט) והלוחמה האלקטרונית.

במאי 2014, משרד ההגנה של רוסיה הכריז על הקמת כוח מבצעי המידע (Information Operations Force), שעל פי התקשורת ברוסיה, התבסס בחלקו על הגידול במספר היחידות העוסקות במדע ובטכנולוגיה והקמתו הואצה על רקע פרשת הדלפות אדוארד סנודן. כמו כן, על פי הדוקטרינה הצבאית משנת 2014, משימת בניין הכוח ופיתוח האמצעים הדרושים לעימות במרחב המידע היא המטרה העיקרית בתהליך המודרניזציה וההצטיידות של הכוחות המזוינים של רוסיה.

בין השנים 2013 ל-2017, יוחסו רוב מתקפות הסייבר המשמעותיות של רוסיה ל-GRU והוא סומן כגוף המוביל בביצוע מבצעי סייבר מתוחכמים.

התרבות הארגונית של ה-GRU, הכוללת מדיניות תוקפנית, פזיזות וסף לקיחת סיכונים גבוה, נתפסת כלא הולמת את תחום מבצעי הסייבר, המורכב יותר ממבצעי ריגול שקטים ופחות ממתקפות רחבות היקף. חוסר התאמה זה משתקף בדבריו של קצין FSB לשעבר, שנעצר לאחר שניסה ככל הנראה להדליף את שמותיהם של ההאקרים השייכים ל-GRU, לפיו אופן הביצוע התוקפני והפזיז של ה-GRU מקל על מדינות המערב וחברות אבטחת סייבר פרטיות לייחס אליו את המתקפות.

עם זאת, הייחוס המתמשך של מבצעי סייבר ולוחמת מידע ל-GRU מעיד על כך, שהארגון זוכה לתמיכתו של הנשיא פוטין וימשיך ככל הנראה להוביל את המבצעים הללו. החשיבות שגורמי ביטחון רשמיים ברוסיה מייחסים לעבודת ה-GRU מחזקת אווירה של דחיפות והרפתקנות.



## התרבות הארגונית של ה-GRU וניהול מבצעי מידע

היבט נוסף בתרבות הארגונית של ה-GRU, שטרם נחקר ויכול להסביר את תחילת עיסוקו במבצעי סייבר היא מעורבותו ההיסטורית בהוצאה לפועל של מבצעי מידע. בניגוד למרבית היחידות האחראיות לתחום הסייבר ב-GRU, יחידות מבצעי המידע שלו הן בעלות היסטוריה ארוכה: הצבא האדום הקים יחידות להפצת תעמולה (spetsprop)[6] מעט לפני מלחמת העולם השנייה. יחידות ה-spetsprop עסקו בשידור מסרים והפצת עלוני תעמולה שנועדו לפגוע במורל הלחימה של כוחות האויב ואף פעלו להשפיע על האוכלוסייה האזרחית שמאחורי קווי האויב. לאחר 1991, שמותיהן של היחידות הללו שונו והן הוצבו באופן בלעדי תחת ה-GRU. ה-GRU ביצע ארגון מחדש וחילק את מומחי היחידות הללו לשמונה קבוצות למבצעי לוחמה פסיכולוגית במהלך מלחמת צ'צ'ניה הראשונה ופיזר אותן ברחבי צבא רוסיה.

עם זאת, המלחמה בגיאורגיה ב-2008 והאכזבה שנרשמה בקרב צבא רוסיה בגין חוסר היכולת להתמודד עם מה שנתפס כמבצעי מידע של המערב נגד רוסיה, הניעה גורמים בצבא להקים מחדש את יחידות ה-spetsprop. גורמי ביטחון רוסים הבינו, שהתעמולה המודרנית, כמו זו שהגיעה מצד ברית נאט"ו, חייבת להיות דיגיטלית. כמו כן, היבטים נוספים של לוחמת מידע, כגון מתקפות DDoS, נוספו לקורסי ההכשרה בפקולטה למבצעי מידע באוניברסיטה הצבאית של רוסיה, הפועלת תחת משרד ההגנה, לצד אמצעים ישנים כגון תפוצת מידע כוזב.

באותה מידה שמתקפות סייבר שימשו אמצעי חדש עבור טקטיקות א-סימטריות, כך טכנולוגיות ICT סיפקו ל-GRU מרחב חדש לניסוי ולהפעלה של טכניקות תעמולה שמקורן עוד בפעילותן של יחידות ה-spetsprop. פעילות הפצת התעמולה והמידע הכוזב של ה-GRU במאה ה-21 לא שונה מהותית מפעילותן של יחידות ה-spetsprop במהלך מלחמת העולם השנייה.

פעילותן של יחידות ה-spetsprop לייצר מתיחות ולסכסך בין פולין לאוקראינה, במטרה להקל על כיבוש מזרח פולין בתחילת מלחמת העולם השנייה, זהה לפעילות ה-GRU במאה ה-21, לסכסך בין שתי המדינות תוך שימוש ברשתות החברתיות.

פעילות היחידות הללו מאז ראשית שנות ה-2000 ממחישה את המעבר שלהן למרחב הדיגיטלי, שינוי שבסופו החל ה-GRU להוציא לפועל מתקפות סייבר. במהלך מלחמת צ'צ'ניה השנייה, ה-GRU השיק אתר חדשות במטרה להפיץ את הנרטיב הרוסי במלחמה. מבצעי ההשפעה המקוונים של ה-GRU החלו לעבור אבולוציה מסוימת בתחילת המשבר באוקראינה ב-2014, עם הפצתו של מדריך לשימוש ברשת החברתית פייסבוק, דבר שמלמד על חוסר הניסיון של פעילי ה-GRU במבצעי השפעה ברשתות החברתיות. שנה לאחר מכן, ה-GRU שילב מתקפות סייבר ומבצעי השפעה ברשת כנגד תחנת הטלוויזיה הצרפתית TV5 Le Monde, תוך התחזות לפעילי דאע"ש במסגרת קמפיין ה-CyberCaliphate. המעורבות של מרכז השירות המיוחד ה-72,[7] המכונה כיום יחידה 54777, ועוסק בלוחמה פסיכולוגית, ממחישה את שיתוף הפעולה בין מומחים למבצעי מידע לבין יחידות הסייבר של ה-GRU.

לדברי גורמי מודיעין במערב, יחידה 54777 פועלת בשיתוף פעולה עם האקרים של ה-GRU לפחות משנת 2014, ומשתמשת במתקפות סייבר כגורם משלים למבצעי מידע באמצעות ארגוני proxy וחברות קש. לפני המשבר באוקראינה, יחידה 54777 כללה 80 מומחים ששובצו בחמש

6 המושג מתייחס לתעמולה "מיוחדת".

7 72nd Special Service Center

מחלקות: המרכז למידע צבאי זר; מחלקה לניהול מבצעי מידע ולוחמה פסיכולוגית; מחלקה להפצת תעמולה ברדיו; מחלקה לניהול מבצעי מידע בגופי מדיה המונית; ומחלקה לעריכת תכנים. היחידה שלחה מומחים ויועצים ליחידות שונות בצבא במטרה לשלבם בדרגי פיקוד שונים. שילובן של יחידות מקומיות למבצעי מידע השייכות ל-GRU בתחומי לוחמת המידע והלוחמה האלקטרונית, ממחיש את יכולתם של מפקדים מקומיים לנהל מבצעי מידע בדרג הזוטר.

## התרבות הארגונית של ה-GRU וניהול מבצעי סייבר טכניים

בזמן שמתקפות הסייבר של ה-GRU משכו תשומת לב מחקרית רבה מאז שנת 2014, תשומת לב מעטה הוקדשה להבנה כיצד ההיסטוריה של הארגון משפיעה על פעילותו העכשווית. מקורם של מבצעי הסייבר של ה-GRU טמון בהיסטוריה של המודיעין הטכני של צבא רוסיה, שהחל להתקיים בתקופה טרום מלחמת העולם הראשונה. מודיעין טכני, ובעיקר הצפנה וסיגינט, ראה התפתחויות משמעותיות בעידן הסובייטי. ההנהגה הצבאית הסובייטית המוקדמת הכירה בחשיבותו והשימוש בסיגינט מילא תפקיד משמעותי בעימות בין סין לבריה"מ ב-1929. יכולות איסוף הסיגינט הסובייטיות הצליחו להתעלות מעל יכולותיה של בריטניה בתחום ואף להשתוות ליכולותיה של ארה"ב בשנת 1939. עם זאת, יכולות אלו שימשו את רוסיה בעיקר לצרכי ביטחון פנים.

התפתחות משמעותית נוספת ביכולות המודיעין הטכני הצבאי של רוסיה נרשמה במהלך מלחמת העולם השנייה, וב-1942 הצליחו מומחי הצפנה רוסים לפענח את מכונת ההצפנה הגרמנית, האניגמה. ה-GRU המשיך לפתח את יכולותיו בתחום הסיגינט במהלך המלחמה הקרה ובתקופת גורב'צוב, הצבא החזיק ב-40 ברג'ימנטים, כ-170 גדודים ובמעל ל-700 פלוגות שעסקו בסיגינט.

אחת ההתפתחויות המשמעותיות ביותר במודיעין הסיגינט הסובייטי במלחמה הקרה הייתה הקמתו של המרכז ה-85 לשירותים מיוחדים (יחידה 26165),[8] שהיה אחראי לשימוש באמצעי IT חדשים לטובת הצפנה. המרכז הועבר ממִנְהֶלת הסיגינט של ה-GRU והחל לפעול תחת מפקדת ה-GRU, דבר המלמד על חשיבותו. עם זאת, סביר להניח שגם המרכז סבל מקיצוצים שנערכו ברחבי הצבא בכלל וביכולות המודיעין בפרט, לאחר תום המלחמה הקרה.

בשנת 2017, נחתם הסכם בין מי ששימש ראש המרכז בשנת 2016, ויקטור נטישקו (Viktor Netyshko) לבין ה-FSB, בנושא שיתוף פעולה בגיוס ובהכשרת כוח אדם במרכז ההצפנה של ה-FSB לקראת העברתם ליחידות ה-GRU. כמו כן, מפקדי המרכז עסקו בקידום המחקר האקדמי והמדעי בתחומי מדעי המחשב, שהיו חיוניים לקידום מבצעי סייבר.

המרכז ה-85 לשירותים מיוחדים מייצג רק חלק ממנגנון מבצעי הסייבר ההתקפיים של ה-GRU. המרכז הראשי לטכנולוגיות מיוחדות (יחידה 74455) משך תשומת לב רבה בשל מעורבותו בבחירות לנשיאות ארה"ב ב-2016 ובביצוע מתקפת הסייבר העולמית NotPetya בשנת 2017. הקמתה של יחידה 74455 שיקפה את החשיבות שההנהגה הצבאית ברוסיה ייחסה למבצעי מחשבים ומפקדיה עוסקים גם במחקר של מדעי המחשב לצרכים צבאיים.



בנוסף, מנהלת יחידה 74455 קשרים עם המכון המרכזי הרביעי למחקר מדעי,[9] השייך ליחידות מערך הטילים האסטרטגיים של רוסיה. מטרתם של קשרים אלה היא ככל הנראה לספק להאקרים של ה-GRU מחקרים בנושא התפתחותה של התיאוריה הצבאית ובנושאי אסטרטגיה השייכים לתחום מבצעי הסייבר. בשנים 2008-2018, פרסם המרכז מאמרים שונים בנושא יכולות סייבר, דבר שהצביע על ההתעניינות הגוברת של הארגון בעניינים הנוגעים לסייבר.

כמו כן, מבצעי סייבר שהתמקדו במטרות באוקראינה, באירופה ובמערב ויוחסו ליחידה 74455, הדגימו את יכולותיה המתקדמות ומצביעים על השקעה ניכרת בתקציב ובכוח האדם המגויס ליחידה. לדברי המומחית למערכות בקרה תעשייתיות, מרינה קוטופיל (Marina Kotofil), קיימים הבדלים משמעותיים בין שתי מתקפות הסייבר ששיתקו את רשת החשמל של אוקראינה ב-2015 וב-2016 ויוחסו ל-GRU, המלמדים על פיתוח יכולותיו.

## השלכות מבצעי הסייבר והמידע הצבאיים של רוסיה על פעילות עתידית בחסות המדינה

תקיפות הסייבר שביצע ה-GRU נגד גיאורגיה בסוף 2019 הדגימו את החיבור שבין אמצעים טכניים לבין אלמנטים מעולם המידע והתודעה בלוחמת מידע מודרנית. המתקפות כללו שימוש בנוזקה מתוחכמת במטרה לחסום שידורי טלוויזיה ולהשבית אתרי רשת, תוך הפצת תמונתו של נשיא גיאורגיה לשעבר, שהורשע בשחיתות ב-2013, לצד הכיתוב שהוא עתיד לחזור לפעילות פוליטית במדינה. השילוב הזה צפוי להימשך גם במבצעים עתידיים, כגון אלו המתמקדים במערכות בחירות שצפויות להתקיים בהמשך שנת 2020 ובשנים הבאות, מבצעים הצפויים להמשיך ולהעמיק את השסעים הפוליטיים והחברתיים, ולערער את יריביה של רוסיה באמצעים דיגיטליים.

ככל שגדלה החשיפה של מדינות המערב למבצעי מידע וסייבר, כך המשיכה רוסיה לשדרג ולהרחיב את יכולות הסייבר שלה ולנצלן. על פי דו"ח של Check Point מסוף שנת 2019, רוסיה ביצעה במחצית הראשונה של 2019 השקעה חסרת תקדים ביכולת תקיפה וריגול סייבר. התפתחויות אלו ממחישות את הצורך להכיר את יכולותיה אלו של רוסיה. לימוד התרבות הארגונית וההיסטוריה של הגורמים המבצעים מבצעי סייבר ומידע, מאפשר לקבל תובנות על המניעים, האסטרטגיה ושיטות הפעולה המיוחדים להם.

בהתחשב בהיקפים ובהשלכות של מבצעי המידע והסייבר של ה-GRU, היכרות עם הגורמים המבצעים, ברמה פרטנית יותר חיונית לחיזוי פעולות עתידיות צפויות ולהבנה כיצד ניתן להתמודד איתן. היכרות זו מחייבת גם קיום מחקר היסטורי על גופי המודיעין של רוסיה.

בזמן שמומחים מערביים רבים התמקדו בדבריו של גרסימוב מ-2013 ותפיסתו את לוחמת המידע, מעטים בלבד שמו לב לדבריהם של מומחים רוסים לביטחון ולאסטרטגיה בדרג הבינוני, שהזהירו מפני עימות מתקרב במרחב המידע בין רוסיה למערב. הבנה מעמיקה של המניעים של הגורמים הרוסיים האחראיים למבצעי סייבר ומידע, תאפשר למנהיגי המערב להתכונן לאיום לוחמת המידע של רוסיה.

## מסקנות

בשנים האחרונות חלו התפתחויות בתפיסת מושג הלחימה של רוסיה, שהובילו לשילוב הגובר בין אמצעים לא צבאיים לבין אמצעים צבאיים קונבנציונליים. שילוב זה בא לידי ביטוי בחשיבות הגדלה של לוחמת המידע בדוקטרינה של רוסיה, לפיה לוחמת המידע הכוללת מבצעי מידע וסייבר והיא חלק בלתי נפרד מהמלחמה המודרנית. מסמך הדוקטרינה מתאר את רוסיה כמדינה המגנה על עצמה במרחב המידע אל מול יריביה התוקפניים. מחקרים ומאמרים של מומחים צבאיים רוסים משקפים את העניין ההולך וגובר בפיתוח יכולות סייבר התקפיות בזכות היעילות, הזמינות ויכולת ההתאמה שלהן לעימותים המודרניים. הניתוח הזה של יכולות סייבר התקפיות תואם לביצוע בפועל של מבצעי סייבר ומידע, שתפיסת הפעלתם פותחה במקביל לגיבוש תפיסתה של רוסיה על המלחמה המודרנית.

הגופים האחראים לתחום לוחמת הסייבר של רוסיה התפתחו במקביל לגיבוש התפיסה הרוסית ללוחמה המודרנית ולהופעת האיום הנשקף משימוש המערב בטכנולוגיות מידע לקידום מטרותיו הצבאיות והמדיניות. בשני העשורים הראשונים שלאחר סוף המלחמה הקרה, ה-FSB נטל חלק מרכזי בניהול מבצעי סייבר, לצד קבלת תמיכה מהאקרים פרטיים. באותה תקופה, גיבשה האליטה הרוסית קונצנזוס סביב התפיסה שלוחמה כוללת שימוש באמצעים צבאיים ולא צבאיים, בזמני שלום ומלחמה. תפיסה זו הובילה את משרד ההגנה הרוסי להגדיל את השקעותיו במטרה לייסד גוף ריכוזי שיהיה אחראי למבצעי סייבר. שינויים אלה, לצד ההזדמנויות המבצעיות שנקרו בפני רוסיה עקב פלישתה לאוקראינה, אפשרו ל-GRU לתפוס עמדת מובילה בתכנון ובביצוע מבצעי סייבר התקפיים, שבהתאם למורשת ההיסטורית של הארגון, התאפיינו בלקיחת סיכונים ובביצוע אגרסיבי. בנוסף, תפקידו ההיסטורי של ה-GRU כגוף האחראי למבצעי מידע ותפוצת תעמולה סיפק בסיס טבעי לקיומם של מבצעי סייבר לצד מבצעי מידע, שני מרכיבי הליבה של לוחמת המידע. כל אלו אפשרו את התפתחותם של מבצעי הסייבר האסטרטגיים של רוסיה מפעולות אד-הוק לקמפיינים מאורגנים המבוצעים באופן ריכוזי, ומשקפים את תפיסת הלוחמה המודרנית שלה.

תפיסת לוחמת המידע של רוסיה והגופים האחראיים למבצעי מידע, עשויים להשפיע על המדיניות והאסטרטגיה העתידיות של רוסיה במרחב הסייבר. התפיסה ההגנתית, לפיה רוסיה מתמודדת מול גורמים תוקפניים, שמנצלים את טכנולוגיות המידע במטרה לערער את הפוטנציאל הצבאי ואת והחברה ברוסיה, צפויה להמשיך ולהתקיים בעתיד הנראה לעין.

במקביל, הרעיון לפיו יריבותיה של רוסיה פגיעות במרחב המידע, יבטיח את מקומם של מבצעי המידע והסייבר בדוקטרינה של רוסיה ואת מעמדם של הגופים האחראיים לביצועם בשנים הקרובות. למרות שצבא רוסיה ימשיך להשקיע ביכולות לחימה מודרניות, חשיבותם הגדולה של האמצעים הלא-קונבנציונליים, בייחוד הדיגיטליים שבהם, תמשיך להוות מוקד לתשומת לב בדוקטרינה הצבאית, במחקר הצבאי ובמדיניות. ייתכן ומקבלי ההחלטות ברוסיה יחליטו לכלול את המחקר, הפיתוח והשימוש באמצעי סייבר התקפיים באופן רשמי בדוקטרינת לוחמת המידע של צבא רוסיה, אולם תרחיש זה נראה לא סביר, בהינתן אופייה ההגנתי של הדוקטרינה, המאפשר לרוסיה מרחב הכחשה מסוים שמטרתו להדוף את ההאשמות בנושא מעורבותה במבצעי סייבר ומידע.



## סיכום

לוחמת המידע תפסה חלק חשוב באסטרטגיה הצבאית של רוסיה עוד מתחילת המאה ה-20. במהלך שני העשורים האחרונים, ההנהגה הצבאית והפוליטית של רוסיה ביצעה התאמות יסודיות של תפיסת מושג המלחמה, שהפכה להיות מסגרת המשלבת אמצעים צבאיים ולא-צבאיים כאחד. השימוש באמצעים לא-צבאיים מטשטש את הגבול שבין מלחמה לשלום ועשוי לייתר לעיתים את השימוש בכוחות צבאיים.

לצד השינויים התפיסתיים, ניתן להצביע על שינויים ארגוניים, במסגרתם הועברה האחריות על מבצעי סייבר בין גופי המודיעין השונים. בעוד שבתחילת שנות ה-2000, הגוף העיקרי שהיה אחראי לביצוע לוחמת סייבר ומידע היה ה-FSB, החל מ-2009 המוקד עבר אל ה-GRU. שינויים אלה נבעו משורה של אירועים היסטוריים שהדגישו את הצורך בעדכון התפיסות הקיימות ובהם המלחמות בצ'צ'ניה ובגיאורגיה, הקמת פיקוד הסייבר האמריקני ופרוץ אירועי האביב הערבי.

השימוש בלוחמת מידע ובלוחמת סייבר נובע מהתפיסה הגנתית, שמדגישה את הצורך של רוסיה להתגונן מפני תוקפנותו של המערב. אולם, ניתוח הפעילות הרוסית מגלה שימוש נרחב במתקפות סייבר ובמבצעי מידע והשפעה עבור יעדים שאינם הגנתיים גרידא. השימוש באמצעים אלו מאפשר לרוסיה להתמודד עם הפער הטכנולוגי והכלכלי מול ארה"ב ולערער את יריביה מבלי להכריז מלחמה. גורמים אלו מובילים לכך שלוחמת סייבר ומידע תמשיך להיות מרכיב מרכזי באסטרטגיה הרוסית.

# ב. השימוש בבינה מלאכותית בקרב צבא סין

**Michael Dahm**[10]

## הקדמה

מטרתו המרכזית של המאמר היא לתאר כיצד השימוש בבינה מלאכותית בקרב צבא סין עשוי לקדם את תפיסות הלוחמה שלו. לטענת המחבר, השימוש בבינה מלאכותית יכול לסייע לצבא סין בהשגת שאיפתו האולטימטיבית: השגת יתרון קוגניטיבי בשדה הקרב, על מנת להכריע את האויב.

שני החלקים הראשונים במאמר משלימים זה את זה: חלקו הראשון מציג את מרכזיותה של הבינה המלאכותית באסטרטגיה הצבאית הסינית, הממוקדת במידע. חלקו השני מציג את הטענה לפיה השימוש בבינה מלאכותית יסייע בקידום אותה האסטרטגיה. חלקו השלישי של המאמר מציג את עקרון ה-"intelligentized warfare", מושג המייצג בעיני צבא סין את אופי לוחמת המידע העתידית. החלק הרביעי דן בשש דרכים ייחודיות שבהן בינה מלאכותית עשויה לסייע ליישום תפיסת לוחמת המידע הסינית. החלק החמישי מציג כיצד סין מחזקת את יכולותיה בתחום הבינה המלאכותית ומצמצת את הפערים מול ארה"ב. המחבר חותם את המאמר עם מסקנות כלליות על תפיסת לוחמת המידע הסינית ועל דרכו של המערב לאתגר אותה.

## רקע

גורמים בצבא סין מאמינים כי סין מפסידה לרוסיה ולארה"ב במרוץ להובלה העולמית בתחום הבינה המלאכותית. אותם גורמים פרסמו מאמרים הכוללים ממצאים וציטוטים של בכירים רוסים ואמריקנים, הקוראים לממשלותיהם להגביר את ההשקעה בטכנולוגיות בינה מלאכותית כחלק משאיפותיהן הלאומיות בתחום הבינה המלאכותית.

בשנת 2017, הייתה סין מהמדינות הראשונות לקדם אסטרטגיה לאומית לפיתוח טכנולוגיות בינה מלאכותית, שעסקה רבות בתרומתה של הטכנולוגיה לפיתוח הכלכלי של סין. לעומת זאת, כוונותיו של צבא סין והאסטרטגיה שלו בתחום הבינה המלאכותית נותרו עמומות. ללא ספק, גורמים בצבא סין מבינים שעליהם להתחרות בארה"ב באמצעות הסתגלות מהירה לשינויים שבינה מלאכותית ומערכות אוטונומיות צפויות לעורר בתחום הלוחמה. בחינה של הדיון המתקיים בקרב בכירים בצבא סין סביב השינויים הללו חושפת כי אופן הלחימה החדש, המכונה בצבא סין intelligentized warfare, הוא למעשה המשך של התפיסות האסטרטגיות והמבצעיות הקיימות בסין.

אסטרטגים סינים נוטים לטעון שטכנולוגיות בינה מלאכותית צריכות להיות מנוצלות במסגרת נשק קינטי ולא קינטי כאחד במטרה להשיג שליטה ודומיננטיות ברשתות ובמערכות מידע, על מנת לשתק את כוחותיו של האויב. לוחמת מידע ושליטה במידע נמצאות בלב גישתו של צבא

10 Michael Dahm. Chinese Debates on the Military Utility of Artificial Intelligence. *War on the Rocks*. June 5, 2020. https://warontherocks.com/2020/06/chinese-debates-on-the-military-utility-of-artificial-intelligence/



סין ללחימה ולבינה מלאכותית. התמודדות עם האסטרטגיה הצבאית של סין יחייב שימוש הגנתי והתקפי בטכנולוגיות בינה מלאכותית חדשות. בעימות עתידי, צבא ארה"ב יצטרך להשתמש בבינה מלאכותית וביכולות אוטונומיות לטובת שמירה על מערכות המידע שלו ולצורך שיבוש האסטרטגיה והיכולות של סין.

## בינה מלאכותית צפויה לשפר את האסטרטגיה הצבאית הקיימת של סין

האסטרטגיה המקיפה של צבא סין להבסת כוח צבאי זר מתמקדת בהשגת דומיננטיות בעימותים בין מערכות-על (system-of-systems).[11] שיטת הלחימה הזו מתמקדת בשיבוש או שיתוק של מערכות-העל של האויב.

ראשית, צבא סין ינסה להרוס את מערכות המידע של האויב באמצעות שימוש באמצעים קינטיים ולא קינטיים. לאחר מכן, בכוונת הצבא לנטרל מרכיבים מסוימים בצבא האויב, באמצעות שימוש במתקפות מדויקות ארוכות טווח. דוקטרינה זו תוארה בשם עימות מערכות (systems confrontation) אך מונח זה אינו מתאר בדיוק את ההשפעות המצטברות הפוטנציאליות המתחוללות במערכת-העל, המורכבת ממספר מערכות השלובות אחת בשנייה. בינה מלאכותית יכולה לסייע בהשגת מטרה זו.

השימוש בבינה מלאכותית בעימות בין מערכות-על עולה בקנה אחד עם האסטרטגיה הסינית הקיימת של "informationized warfare". על פי הדוקטרינה של צבא סין, מרכז הכובד של המבצעים הצבאיים המודרניים עבר מריכוז כוחות למערכות-על בתחום המידע, הכוללות את מערך התקשורת, הפיקוד והשליטה, זיהוי המטרות ועיבוד מידע. יש לכך השפעה על כלל ההיבטים של המלחמה המודרנית. מערכות-על צבאיות הן נרחבות ומורכבות ובעתיד ינוהלו כנראה באמצעות בינה מלאכותית. על כן, קיימת הערכה כי הדרך לנתח ולהבין אותן בזמן אמת ולתקוף אותן, היא באמצעות שימוש בבינה מלאכותית.

מטרתו של צבא סין היא להשתמש באלגוריתמים של בינה מלאכותית, בלמידת מכונה, בצוותים מעורבים של מכונות ומפעילים אנושיים ובמערכות אוטונומיות, כדי לשתק את יריביו. המטרה העליונה של הצבא היא להשיג יתרון קוגניטיבי, כלומר את היכולת להבין ולהסתגל למערכות-העל של היריב באופן מהיר יותר משהיריב יכול להבין ולהסתגל למערכת-העל של צבא סין. צבא סין שואף להשתמש בבינה מלאכותית כדי לשתק את האויב במהירות ובאופן מדויק, תוך הגנה על מערכות-העל שלו. כל מי שעתיד לקרוא תיגר על צבא סין יצטרך להבין כיצד יכולות עתידיות בבינה מלאכותית יוכלו לסייע בהבנת מטרותיה של סין בעימות בין מערכות-על.

## עקרון ה-Intelligentized Warfare

הגרסה האנגלית של אסטרטגיית ההגנה הסינית משנת 2019 מתארת שינוי בלוחמה המודרנית: המלחמה מתפתחת לכיוון המכונֶה informationized warfare והלוחמה הנבונה (intelligent warfare) נמצאת באופק. עם זאת, תרגום לאנגלית של הגרסה הסינית של האסטרטגיה מתארת

11 המושג מתייחס לאוסף של מערכות ייעודיות המקיימות ביחד מערכת משותפת המפיקה ביצועים ומאפשרת יכולות גדולות יותר מכלל המערכות המרכיבות אותה. מושג זה מתייחס לרוב למערכות C4I ו-ISR (Computers, Communications and Information; Intelligence, Surveillance and Reconnaissance).

את השינוי כאבולוציה מתמשכת שצפויה להוביל לעלייתה של הלוחמה הנבונה.

המושג הסיני של צורת המלחמה מתייחס לאופייה המשתנה של המלחמה. עידן המידע הניב את עיקרון ה- informationized warfare, המשפיע ומכוון את התפתחות צבא סין מתחילת שנות ה-2000. כעת, בכירי הצבא סבורים שצורת הלוחמה תתפתח לכיוון הלוחמה הנבונה.

ההערכה משנת 2019 בדבר האבולוציה לקראת לוחמה נבונה זהה לאסטרטגיה הצבאית משנת 2002, לפיה צורת המלחמה מתפתחת לכיוון של לוחמת מידע (informationization). דיונים שנערכו בקרב בכירי צבא סין בנושא המעבר ללוחמת מידע ולטכנולוגיות מידע מודרניות גיבשו, למעשה, את היסודות לתיאוריית ודוקטרינת ה-informationized warfare. כעת, מתמקדים הדיונים בעיקרון ה- intelligentized warfare ובשינויים שהוא צפוי להוביל בצבא.

פרסומים המופיעים בגופי תקשורת צבאיים רשמים, כגון עיתון ה-PLA Daily והאתר הרשמי של צבא סין מספקים במה לדיונים בנושאים צבאיים שונים, בהם עקרון ה-intelligentized warfare. מאמרים, הצעות למחקר ומאמרי דעה אלו מפורסמים לצד תמיכה מרומזת מצד הממסד הצבאי הסיני ויכולים לספק תובנות נוספות בנושא הקונצנזוס סביב עקרון ה-intelligentized warfare.

מאז פרסום האסטרטגיה בשנת 2019, התפרסמו פרשנויות שונות בגופי התקשורת הרשמית של הצבא, שעסקו ב-informationized warfare, intelligentized warfare, מערכות בלתי מאוישות, קבלת החלטות אוטונומית ולוחמה קוגניטיבית. פרסומים אלו מייצגים מגמה רעיונית של שילוב טכנולוגיות הבינה המלאכותית עם רעיון ה- informationized warfare והאסטרטגיה הצבאית המתמקדת במידע.

## שישה עקרונות של עקרון ה-Intelligentized Warfare

מאמרים סיניים הצביעו על כך שעל אף שבינה מלאכותית תשנה את מאפייני המלחמה, המלחמה תישאר פעולה אלימה המיועדת להשגת מטרות פוליטיות ובני אדם ימשיכו למלא בה חלק מרכזי. המחברים אף ציינו שבני אדם צפויים להמשיך לתכנן, לארגן וליזום מלחמות גם בעתיד, בעוד שמכונות יוכלו לסייע לבני אדם בהיבטים קוגניטיביים ובביצוע פעולות. עם זאת, בני אדם עדיין יהיו המרכיב הדומיננטי במלחמות.

עם זאת, בדיון הסיני על בינה מלאכותית ישנם מספר היבטים ייחודיים שיכולים למשוך את תשומת ליבם של אנליסטים מערביים. מרביתם משקפים את ההנחה שבינה מלאכותית תאפשר לפתח את התיאוריה והדוקטרינה הצבאיות.

**1.** על פי אסטרטגים בצבא סין, מושג ה-intelligentized warfare מוגדר כהתפתחות וכאבולוציה של מושג ה-informationized warfare. מדובר בעימות בין מערכות-על המסתמך על מידע העובר דרך מערכות ורשתות דיגיטליות. עם זאת, ישנם מאמרים בהם תהליך התבוּניות (intelligentization) של הלוחמה תואר בתור שלב עצמאי בהתפתחות המלחמה העתידית.

**2.** תיאוריטקנים צבאיים סינים טוענים שהשימוש ברשתות תקשורת נרחבות ובטכנולוגיות אלחוטיות (ubiquitous networks) יאפשרו לוחמה בין מערכות-על ויקצרו את הזמן הדרוש לקבלת החלטות וביצוע פעולות. התפיסה הסינית של רשתות מסוג הזה דומה לרעיון לוחמת הפסיפס (Mosaic Warfare) של דארפ"א, המתאר רשתות גמישות המשמשות לחיבור פלטפורמות



בלתי מאוישות. הצבא הסיני מתמקד בפיתוח מערכות-על צבאיות מזה כ-15 שנים. השימוש ברשתות גמישות ונפוצות אלו לצד השימוש בטכנולוגיות בינה מלאכותית עשויים לייצר גרסה סינית של לוחמת הפסיפס.

**3.** אסטרטגים סיניים טוענים שבינה מלאכותית משפיעה על הפיקוד והתכנון המבצעי. התפיסה הרווחת הקיימת, לפיה צבא סין מחזיק בתפיסות סובייטיות של מבני פיקוד ושליטה ריכוזיים, הינה שגויה. אסטרטגים בצבאות המערב מניחים כי ניתוק של כוחות צבא סין ממוקד השליטה והבקרה המרכזי, יוביל לחוסר תפקוד או לקבלת החלטות אקראית, אולם צבא סין סבור שפונקציות הפיקוד והשליטה יכולות להיות מוּבְנות מראש בעיצוב המערכות ובתהליכי התכנון המבצעי, במטרה למזער טעויות מצד מכונות או בני אדם בזמן לחימה.

בהתאם למורשת האינטלקטואלית שלו, המאופיינת בסגנון מרקסיסטי-לנינסטי, הצבא מתאר את המלחמה כתהליך מדעי, שניתן לנתח אותו במטרה לקבל תוצאות מחושבות מראש. טכנולוגיות בינה מלאכותית ולמידת מכונה עשויות להעניק לצבא סין את האלגוריתמים והכלים, שהוא סבור שנחוצים לו על מנת להגדיר את התוצאות הסופיות של המלחמה ולפתח מערכות-על חסינות לצד יכולות מבצעיות ותכניות צבאיות.

**4.** ישנה ציפייה שבינה מלאכותית תאפשר פיתוח ושימוש בקונספטים מבצעיים חדשים. מאמר בשם "Intelligentized Warfare, Where are the Changes?" שפורסם בינואר 2020 מתאר דפוסי מבצעים עתידיים, כגון לוחמת התשה המבוססת על נחילים אוטונומיים,[12] לוחמת מתקפות רדומות אוטונומיות,[13] ולוחמה ניידת רב-ממדית אוטונומית.[14] דפוס מבצעי נוסף שעולה ממאמרים נוספים הוא לוחמת השליטה האוטונומית בתודעה של האויב.

מחברי המאמרים הסינים ניתחו את עקרון קבלת ההחלטות OODA (observe, orient, decide, act) האמריקני, והגיעו למסקנה ששלב קבלת ההחלטות היא צוואר הבקבוק של התהליך. לטענתם, מערכות אוטונומיות עתידיות יציגו יתרון בקבלת החלטות ויאפשרו להניע מהר יותר את פעילות הצבא.

**5.** בינה מלאכותית יכולה לאפשר לצבא סין לבצע מתקפות קינטיות מדויקות ולשתק את מערכות-העל של היריב. מאמר של סוכנות הידיעות הסינית Xinhua שפורסם באתר משרד הביטחון הלאומי של סין בספטמבר 2019 חזה, שבינה מלאכותית ומערכות אוטונומיות יאפשרו שחרור מדויק של אנרגיה קינטית שתפוזר ברחבי מערכת-העל או תתמקד בחלקים מרכזיים וחיוניים בה. כל זאת, על מנת לשתק את האויב לזמן ממושן. היתרון בשדה הקרב ייטה לטובת הצד שיידע להשיג שליטה במרחב הקוגניטיבי. כלומר, הצד המנצח יהיה זה שיידע לתכנן, להתאים את פעולותיו ולפעול מהר יותר מהצד השני, על מנת לשבש ולשתק את מערכות-העל של היריב או על מנת לבטל את השפעת היריב על מערכות-העל שלו.

12 Autonomous swarm attrition warfare ; המושג מתייחס לשימוש בנחילים של פלטפורמות אוטונומיות לצורך מתקפות מתואמות ונרחבות.

13 Autonomous dormant assault warfare ; המושג מתייחס למתקפות פתע על מטרות מפתח ויכולות חיוניות של האויב תוך שימוש בפלטפורמות אוטונומיות המתוכנתות להישאר "רדומות" עד להפעלתן בעת זיהוי המטרה

14 Autonomous cross-domain mobile warfare ; המושג מתייחס לשימוש בפלטפורמות אוטונומיות בעלות ניידות גבוהה המסוגלות לפגוע במטרות בטווח רחוק ובאופן נרחב.

**6.** אנליסטים סיניים מכירים בכך שצבא סין חייב להאיץ את התקדמותו בבינה מלאכותית, מאחר וסין מפגרת אחרי ארה"ב בפיתוח טכנולוגיות בינה מלאכותית לצרכים צבאיים. לשם המחשת הפיגור, חוקרים צבאיים סיניים מדגישים התפתחויות אסטרטגיות שונות בצבאות ארה"ב ורוסיה, תוך קריאה להגביר את ההשקעה במחקר ופיתוח טכנולוגיות בינה מלאכותית צבאיות.

## הגישה הסינית לפיתוח יכולות בינה מלאכותית

ברור כי בינה מלאכותית היא תחום שמקבל עדיפות גבוהה בצבא סין. בשנים האחרונות אימץ צבא סין אסטרטגיה המכונה "עקיפה בסיבוב" (Overtaking on the Curve) שמשמעותה ניצול מגמות חדשות בתחומי המדע והטכנולוגיה במטרה לצמצם את הפער ובסוף לעקוף ארה"ב ורוסיה. שינויים, פיתוחים וחידושים פוטנציאליים עתידיים בתחום הבינה המלאכותית, יוכלו לפיכך להוות הזדמנות עבור סין למזער ואף לסגור את הפער הטכנולוגי מול ארה"ב ולהבין לעומק את האסטרטגיה הצבאית שלה.

על פי מחקרים מערביים עדכניים בנושא תעשיית הבינה המלאכותית, ארה"ב מובילה במחקר ובפיתוח של בינה מלאכותית אולם סין סוגרת את הפער מולה במהירות. עם זאת, הערכת מצב התחרות הצבאית בתחום הבינה המלאכותית קשה יותר. בקשת התקציב של מחלקת ההגנה האמריקנית לשנה הפיסקאלית 2021 כוללת בקשה לכ-841 מיליון דולר להשקעה ישירה בבינה מלאכותית, אך נתון זה אינו כולל פירוט על מיזוג טכנולוגיות הבינה המלאכותית והשפעתו על תקציבי הפיתוח של מערכות נשק אחרות. תקציב הביטחון של סין המוקדש לטכנולוגיות בינה מלאכותית צבאיות מעורפל אף יותר.

ההשוואה בין ארה"ב וסין בתחום טכנולוגיות הבינה המלאכותית הצבאיות נתקלת בקשיים גם בשל העובדה שהחדשנות וההתקדמות הטכנולוגית בתחום הבינה המלאכותית מוּנָעות על ידי התעשייה ולצרכי יישומים אזרחיים. הכוחות המזויינים ימשיכו להפיק תועלת מהשימוש הכפול בעיבוד נתוני-עתק (big data) ובאלגוריתמים של בינה מלאכותית, המגבירים את היעילות התעשייתית ומאפשרים לפתח מערכות אוטונומיות מסחריות.

במרוץ להובלה בפיתוח ובהטמעה של טכנולוגיות בינה מלאכותית נהנית סין מיתרון הגודל. ממשלת סין מקדמת פיתוח בינה מלאכותית באמצעות שימוש בערים שלמות כמעבדות. בסין ישנן מספר ערים המוגדרות כ"ערים חכמות", המשמשות את הממשלה לבחון את השימוש בבינה מלאכותית לצרכי ניהול העיר וביטחון. בסופו של דבר, צבא סין יפיק תועלת מהחדשנות האזרחית והמסחרית בבינה מלאכותית, רובה מבוססת על שיתופי פעולה עם תעשיות זרות, ביניהן האמריקנית. למודלים אלו, של שליטה ממשלתית ריכוזית או הובלה של התעשייה האזרחית, צפויה להיות השפעה על יכולת החיזוי של מרוץ החימוש בטכנולוגיות בינה מלאכותית צבאיות.



## תובנות ומבט לעתיד

על פי תפיסתו העצמית, צבא סין מפגר אחרי צבא ארה"ב המתקדם מבחינה טכנולוגית ושפריסתו ברחבי העולם מספקת לו השפעה רבה ויכולת להפעיל עוצמה. נשיא סין, שי ג'ינפינג, הורה לצבא סין צריך להשלים את תהליך המודרניזציה שלו בשנת 2035 ולהיות שווה ערך מבחינה טכנולוגית לצבא ארה"ב עד שנת 2050.

האסטרטגיה הסינית לשימוש בבינה מלאכותית היא הרחבה של עיקרון ה-"informationized warfare" של צבא סין ונובעת מהגדרתו את מושג המלחמה. אף על פי שיש נקודות חפיפה בין החשיבה הצבאית האמריקנית לסינית בנושא השימוש בבינה מלאכותית, התיאוריה הסינית מציגה מסקנות שונות מארה"ב. בעוד שהחשיבה האמריקנית נוטה להדגיש את תפקידה של הבינה המלאכותית בהגברת כוח האש ובאסטרטגיות העוסקות בתמרונים, צבא סין מקדם עקרונות העוסקים בקידום אסטרטגיות המתבססות על מידע.

טכנולוגיות חדשות בתחומי הבינה המלאכותית, למידת מכונה ומערכות אוטונומיות עשויות להעניק לצבא סין את הכלים הנחוצים לו לשם הבנה וזיהוי מטרותיו ארוכות הטווח לשליטה במרחב המידע, ביצוע מניפולציות בתפיסה המחשבתית ושיתוק תהליכי קבלת ההחלטות של היריב. בבואה לפתח אסטרטגיות להתמודדות עם יכולותיה הצבאיות של סין, על ארה"ב להקדיש תשומת לב לעקרונות הלחימה המתפתחים בצבא סין ולהשקפותיו המתפתחות על תפקיד הבינה המלאכותית בלוחמה העתידית.

## סיכום

בשלושת העשורים האחרונים ראתה התפיסה האסטרטגית הסינית במלחמה תהליך מתפתח שמושפע במידה רבה מהשימוש במידע. תפיסה זו מיוצגת על ידי השימוש במונחים "informationized warfare"", המייצג את מצב הלוחמה הנוכחי, ו-"intelligentized warfare", המתאר את השלב הבא בהתפתחות הלוחמה, המסתייע בבינה מלאכותית. בהתאם לכך, השימוש בבינה מלאכותית נועד לסייע במימוש אסטרטגיית לחימה שמעניקה משקל רב לשימוש במידע. הבינה המלאכותית, לטענת המחבר, יכולה לסייע לסין במשימה זו בשלביה השונים, הן בשלב תכנון המערכה והן בשלב הפגיעה באויב.

המיקוד הרב בספרות הצבאית בנושא בינה מלאכותית נובע מחשיבותה של לוחמת המידע באסטרטגיה הסינית, ומרצונה של ההנהגה הפוליטית בסין לצמצם את הפער בפיתוח ובשימוש בטכנולוגיות בינה מלאכותית צבאיות מול ארה"ב.

המחבר מצביע על כך שהתפיסה הסינית לגבי בינה מלאכותית שונה מהתפיסה האמריקנית, הרואה בבינה מלאכותית פוטנציאל להגברת כוח האש. בהתאם לכך, הוא סבור שאסטרטגיות מערביות שמיועדות להתמודד עם התפיסה הסינית, צריכות לתת משקל לתפיסה זו ולהתפתחותה בספרות הצבאית הסינית.

# ג. דיפ-פייק ומדיה מלאכותית במערכת הפיננסית: הערכת תרחישי האיום

**Jon Bateman**[15]

## הקדמה

מסמך זה מנתח את איומי קטעי ה-Deepfake והמדיה המלאכותית על המגזר הפיננסי. לשם כך, מחבר המסמך מציג עשרה תרחישים פוטנציאליים, המהווים איום כנגד ארבעה סוגים של מטרות: יחידים, חברות (במגזר הפיננסי ומחוצה לו), המגזר הפיננסי בכללותו ומוסדות פיננסיים. לכל תרחיש המופיע במסמך צורפו שלוש שאלות רלוונטיות: כיצד ניתן לבצע את ההונאה; כיצד שימוש במדיה מלאכותית עשוי לשנות את ביצוע ההונאה; אילו התפתחויות צפויות בשימוש העתידי במדיה מלאכותית.

בחלקו השני של המסמך, המחבר פורש תובנות הנוגעות לתחומים רחבים יותר: השפעת השימוש בקטעי Deepfake על היציבות המאקרו-כלכלית, היקף השימוש העתידי בהם לעומת השימוש באמצעי ההונאה המסורתיים והמלצות לדרכי התמודדות אפשריות עם האיומים הפוטנציאליים.

## מבוא

הופעתם של קטעי Deepfake וקטעי מדיה אחרים, שנוצרו בסיוע בינה מלאכותית וזכו למושג מדיה סינטטית (synthetic media) או מדיה מלאכותית, העלתה חששות מפני השימוש בהם לטובת הפצת מידע כוזב בסביבה הפוליטית. לעומת זאת, שאלת השימוש בטכנולוגיות אלו לגרימת נזק פיננסי זכתה לתשומת לב מועטה בלבד.

רמאות, הונאה, זיוף וביצוע מניפולציות בשוק הם אתגרים ייחודיים בכל כלכלה. גורמים זדוניים משלבים לעיתים קרובות טכנולוגיות חדשות בפעילותם, דבר המחייב מחקר נוסף במטרה להבין כיצד אמצעי הונאה חדשים, כגון Deepfake, משמשים לביצוע עבירות כלכליות ונזקים נוספים.

קיים חוסר הסכמה בקרב מומחים בנושא היקף איום ה-Deepfakes על המערכת הפיננסית. ישנם מעט מאוד מקרים מתועדים, דבר שמקשה על חיזוי מגמות עתידיות. חלק מהמומחים במגזר הפיננסי מגדירים את ה-Deepfakes כאתגר טכנולוגי ראשון במעלה ודורשים שינויי מדיניות משמעותיים. מנגד, מומחים אחרים סבורים שקיימת הערכת יתר של איום ה-Deepfakes וכי המערכות הקיימות לאימות וליצירת אמון יכולות להסתגל לטכנולוגיה החדשה.

מסמך זה מקדם את הדיון באמצעות: זיהוי הדרכים הספציפיות בהן איומי ה-Deepfakes והמדיה המלאכותית עלולים להוביל לפגיעה כלכלית; הערכת ההשפעה של האיומים השונים; והצגת תובנות למקבלי החלטות. המסמך גם מתאר סוגים שונים של תרחישי איומים, המופנים כלפי ארבע סוגי מטרות: יחידים (אינדיבידואלים), חברות, שווקים וגופים רגולטוריים. התיאור מתבסס על הטכנולוגיה הקיימת כיום ליצירת מדיה מלאכותית ועל המציאות של פשיעה כלכלית. מחבר המסמך מנסה לתאר את האתגרים באופן רחב, זאת על אף שאין מערך תרחישים מוגדר ומקיף.

15 Jon Bateman. Deepfakes and Synthetic Media in the Financial System: Assessing Threat Scenarios. Carnegie Endowment for International Peace. July 8, 2020. https://carnegieendowment.org/2020/07/08/deepfakes-and-synthetic-media-in-financial-system-assessing-threat-scenarios-pub-82237



## מתודולוגיה

מדיה מלאכותית יכולה לשמש לצורך הסבת נזק פיננסי למגוון רחב של מטרות פוטנציאליות. המטרות הבולטות כוללות מוסדות פיננסיים כגון בנקים, בורסות לניירות ערך, מסלקות וברוקרים, מטרות המסתמכות על מידע מהימן כדי לבצע עסקאות. מטרות נוספות הן רגולטורים פיננסיים ובנקים מרכזיים, המפקחים על השוק ונאבקים בתפוצת מידע כוזב בעל השפעה מזיקה. עם זאת, גם אינדיבידואלים וחברות שמחוץ למגזר הפיננסי עלולים להוות מטרות לתקיפה.

בהתאם לנאמר לעיל, המסמך מציג עשרה תרחישי איומים, הניצבים בפני ארבע קבוצות של מטרות פוטנציאליות, המדורגים באופן עולה בהתאם לחשיבותם המערכתית: אינדיבידואלים; חברות (פיננסיות ולא פיננסיות); שווקים (לאומיים ועולמיים); וגופים רגולטוריים בענף הפיננסים. מטעמי פשטות, המסמך משייך כל תרחיש לקבוצה מסוימת, אולם מכיר בכך שחלק מהתרחישים תקפים עבור כמה מטרות מסוגים שונים. כך למשל גניבת זהות פוגעת באינדיבידואלים שזהותם נגנבה אולם יכולה לפגוע גם בחברות, כגון בנקים שמנפיקים כרטיסי אשראי למתחזים וקמעונאים שמעבדים רכישות שבוצעו דרכם. במילים אחרות, הצטברות של פגיעות בהיקף קטן יכולה, מבחינה תיאורטית, להיות בעלת השפעות גדולות יותר בהיקפן ופגיעות בהיקף נרחב יכולות, באופן בלתי נמנע, להשפיע על אינדיבידואלים.

גורמים זדוניים מחזיקים ברשותם טכניקות רבות מוכחות ויעילות, ביניהן גם טכנולוגיות דיגיטליות. לכן, מסמך זה משווה מדיה מלאכותית לטכנולוגיות הנמצאות בשימוש נרחב למטרות פשיעה פיננסית. כל תרחיש מתאר פעילות לא חוקית, מעריך כיצד השימוש במדיה מלאכותית יכול לקדם ולהרחיב אותה ומעריך את רמת האיום של הפגיעה הפוטנציאלית.

השוואה ישירה בין מדיה מלאכותית לבין אמצעים טכנולוגיים קיימים שבשימוש עבריינים מסייעת להתמקד בתרחישים המדאיגים ביותר, שיכולים לחזק במידה רבה גורמים עברייניים ודורשים פתרונות מדיניות חדשים. באותה מידה, היא מסייעת לזהות את האיומים המדאיגים פחות, שאינם מסוכנים יותר מהאיומים הקיימים כיום ושאינם דורשים ככל הנראה שינוי בהיערכות.

כל עשרת התרחישים המתוארים במסמך הם תרחישים אפשריים הכוללים שימוש בטכנולוגיות זמינות ליצירת מדיה מלאכותית. תרחישים אחרים, כוללים שימוש בטכנולוגיות קיימות שניתן לשלוט בהתפשטותן ולהגביל את השימוש בהן לרעה.

## סקירת תרחישים

<u>מדיה מלאכותית בהיקף מצומצם ובהיקף רחב</u>

ניתן לחלק את עשרת תרחישי האיומים לשתי קטגוריות של מדיה מלאכותית. תרחישים 1-4 כוללים שימוש במה שמכונה מדיה מלאכותית בהיקף צר,[16] לצורך התמקדות במטרות יחידניות, כגון האחראי על השכר בעסק מסוים, ומועברת אל המטרה באופן ישיר בערוצים פרטיים, כגון שיחת טלפון. תרחישים 5-10 כוללים שימוש במדיה מלאכותית בהיקף רחב,[17] שמיועדת נגד מטרות נרחבות, כגון קהילת המשקיעים ומופצת באופן נרחב בערוצים ציבוריים, כגון רשתות חברתיות.

16 Narrowcast synthetic media
17 Broadcast synthetic media

כל קטגוריה מציגה הזדמנויות ואתגרים שונים למקבלי החלטות. לדוגמה, מדיה מלאכותית בהיקף רחב נוטה להתפשט בערוצי תקשורת כמו רשתות חברתיות או דיווחי חדשות, שהתכנים בהם מנוטרים ומספקים אפשרות לחשיפה ולבדיקת עובדות. בניגוד לכך, מדיה מלאכותית בהיקף צר נוטה לעבור בערוצי מדיה לא מנוטרים, כמו דוא"ל והודעות SMS. מצב שכזה מחייב שימוש בצעדי-נגד אחרים כגון שימוש בטכנולוגיות לסינון ספאם, המסייעות לבצע אימות זהויות ולסמן או לסנן גורמים חשודים.

מדיה מלאכותית בהיקף רחב זכתה לתשומת לב רבה ממקבלי החלטות, המודאגים מהשפעה והתערבות בבחירות ובתהליכים דמוקרטיים. לעומת זאת, מדיה מלאכותית בהיקף צר זכתה עד כה לתשומת לב מועטה בלבד. על המגזר הפיננסי לתקן את חוסר האיזון הזה ולעורר מודעות ציבורית גדולה יותר לאיומי מדיה מלאכותית בהיקף צר, שעלולים להוביל לנזק פיננסי במסגרת תרחישים רבים.

<u>שלוש טכניקות מרכזיות המשמשות למטרות זדוניות</u>

במסגרת עשרת תרחישי האיומים הנדונים במסמך, ניתן לזהות שלוש טכניקות מרכזיות: **(1)** Deepfake voice phishing (Vishing); **(2)** מסרים פרטיים מזויפים;[18] **(3)** רשתות בוטים.

**Deepfake Vishing:** איום מדיה מלאכותית בהיקף צר, העושה שימוש בקולות שהועתקו לצורכי הנדסה חברתית באמצעות שיחות טלפון. טכניקה זו משמשת לגניבת זהות, התחזות לצרכי הונאה והונאות תשלום. פעולת ה-Vishing, הכוללת שימוש ב-Deepfake קולי, מנצלת את אמון הנמען בדמות שאליה מתחזה המוען. קורבנות רבים של פעולות כאלו עשויים לייחס שיבושים בקול המלאכותי לשיבושים בקו או בקליטה ועלולים להיכנע למניפולציות רגשיות מצד התוקפים.

**מסרים פרטיים מזויפים:** איום מדיה מלאכותית בהיקף רחב, המשתמש בקטעי Deepfake קוליים או חזותיים (קטעי ווידאו), המתארים דמויות ציבוריות המשמיעות מסרים מזיקים באופן פרטי. טכניקה זו עלולה לשמש לביצוע הונאות שונות, בהן מניפולציות על שוק המניות, משיכה המונית של כספים מבנקים על ידי לקוחות מודאגים על מנת להוביל להתרסקות הבנק[19] וזיוף כביכול של פעולות ממשלתיות.

**רשתות בוטים המבוססות על בינה מלאכותית:** איום מדיה מלאכותית בהיקף רחב. בינה מלאכותית עלולה לשמש להקמת חשבונות מזויפים ברשתות החברתיות תוך שימוש בתמונות ובטקסטים מלאכותיים ולצורך הפעלתם. השימוש ברשתות בוטים עלול לשמש את התוקפים לצורך פגיעה בחברות, בשווקים ובגופים רגולטוריים. בוטים מלאכותיים יהיו יעילים יותר וקשים יותר לזיהוי מהבוטים הקיימים היום. סביר להניח שבוטים ברשתות החברתיות יעשו שימוש נרחב יותר בבינה מלאכותית בעתיד, דבר שיחריף את התחרות הטכנולוגית בין חברות המדיה החברתית לבין הגורמים הזדוניים.

18 Fabricated private remarks

19 Bank run and flash crash



## תרחישים המתמקדים באינדיבידואלים

### 1. גניבת זהות

**האיום הקיים**: גניבת זהות היא התלונה הנפוצה ביותר מבין התלונות המגיעות לנציבות הסחר הפדראלית של ארה"ב (FTC).[20] במקרה השכיח, עבריין פותח חשבון כרטיס אשראי חדש על שמו של הקורבן. פתיחת החשבון מצריכה שימוש במידע אישי על הקורבן, אותו האקרים יכולים להשיג באמצעות פריצה למאגרי מידע מסחריים. פריצות רחבות היקף למאגרי מידע הזינו את התפתחות האקו-סיסטם העברייני והרחיבו את מגמת גניבת הזהויות. נתונים שנגנבו ממאגרי מידע נסחרים על ידי עבריינים ברשת האפילה.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: ישנן לפחות שתי דרכים בהן קטעי Deepfake עלולים לסייע לגניבת זהות:

1. שימוש ב-Deepfakes באופן ממוקד במטרה לגנוב זהות או לבצע העברות כספים. דוגמה לכך היא שיחת טלפון הכוללת שימוש בקולו של הקורבן במטרה לשכנע את היועץ הפיננסי שלו לבצע העברות כספים בין חשבונות. סוג נוסף של גניבת זהות ממוקדת כולל שימוש בקטעי Deepfake לפתיחת חשבונות בנק תחת זהות בדויה לטובת הלבנת כספים.
2. עבריינים יכולים להשתמש ב-Deepfakes במסגרת קמפיינים של הנדסה חברתית על מנת להשיג באופן לא מורשה גישה למאגרי מידע אישי. במסגרת זו, ניתן להשתמש ב-Deepfake על מנת לזייף שיחת טלפון מטעם גורם אחראי או מנהל IT על מנת לשכנע עובדי חברות מסחר אלקטרוני למסור את הרשאותיהם. בתרחיש הזה, קטעי ה-Deepfake תופסים את מקומה של מתקפת הדיוג במטרה לקבל גישה לפרטי הרשאות, כשגניבת הזהות עצמה מתרחשת בשלב מאוחר יותר.

**הערכה**: השימוש ב-Deepfake לשם vishing אפשרי מבחינה טכנית. הטכנולוגיה הקיימת היום מאפשרת העתקה של קולות, הניתנים לשליטה בזמן אמת. טכנולוגיה להעתקת קולות באיכות גבוהה יכולה להשתמש בהקלטה באורך של חמש דקות בלבד. אלגוריתמים אחרים מסוגלים לחולל קולות מלאכותיים על סמך דגימת קול של שלוש שניות.

משמעות היכולת לזייף קולות על בסיס הקלטות קצרות בלבד, היא שניתן להשתמש גם בהקלטות של דמויות זוטרות שאינן מוקלטות בדרך כלל, על מנת לבצע הונאות שונות. ההתפתחות הטכנולוגית צפויה אף להקטין את משכי ההקלטות הדרושות במטרה להפיק קול מלאכותי באיכות גבוהה יותר.

העתקת קול יכולה להסתמך גם על מענה קולי או על קטעי ווידאו שהקורבן פרסם ברשתות החברתיות. גורמים זדוניים עלולים גם להתקשר לקורבן בתואנה כלשהי ולהקליט בחשאי את קולו, על מנת להפיק ממנו קול מלאכותי, שישמש עבור שיחות טלפון עם אינדיבידואלים אחרים. גורמים פנימיים בארגון עלולים לנצל את קרבתם לדמויות בכירות במטרה להקליט את קולן ולדעת מה בדיוק הקול המלאכותי צריך לומר, על מנת לעורר אמון ולהשיג את המטרה.

מצד שני, תרחישים אלו כוללים עבודה מאומצת יותר והם מורכבים יותר, בהשוואה לטכניקות המשמשות היום עבריינים. מתחזים מיומנים יכולים לחקות קולות ללא שימוש בבינה מלאכותית

20 Federal Trade Commission

ולחפות על סדקים אפשריים בקול באמצעות יצירת תחושת דחיפות בעת קיום השיחה.

בנוסף, גניבת זהות כפי שהיא מבוצעת כיום, כוללת שימוש בטכניקות סייבר שניתנות לשימוש בהיקף נרחב בקלות רבה יותר מאשר באמצעות Deepfakes. דוגמה לכך הן מתקפות הדיוג, באמצעותן ניתן להתמקד בעובדים רבים בחברה מסוימת או במספר חברות. בהשוואה לכך, שיחת טלפון המבוססת על Deepfake חייבת להיות מתוכננת בקפידה ולהיות נשלטת באופן אישי על ידי התוקף. בתיאוריה, ניתן לבצע שיחות טלפון אוטומטיות המבוססות על Deepfake, תוך שימוש בתוכנות מסוג interactive voice response, אך השילוב בין שתי הטכנולוגיות צפוי להיות מורכב.

מאחר וגניבת זהות מתבצעת כיום בהיקף נרחב ובמידת יעילות גבוהה, קטעי Deepfake לא צפויים להחליף באופן מלא את האמצעים הקיימים. עם זאת, קטעי Deepfake עשויים לשמש כאמצעים נלווים ועוצמתיים לביצוע פשעים פיננסיים מתוחכמים ביותר, בייחוד אלו שמבוצעים על ידי גורמים פנימיים.

### 2. הונאת התחזות

**האיום הקיים**: התלונות על הונאות מסוג זה הן סוג התלונות השני בשכיחותו, מבין התלונות המגיעות ל-FTC. במסגרת הונאות אלו, התוקפים מתחזים לגורם ממשלתי, לקרוב משפחה שנמצא במצוקה, לעסק ידוע או לאיש תמיכה טכנית במטרה להפעיל לחץ על הקורבן לשלם או לבצע העברת כספים. בדרך כלל, התוקפים יוצרים קשר באמצעות הטלפון, תוך שימוש במספרי טלפון מזויפים או בשירותי שיחות דיגיטליות כגון Skype, על מנת להציג את עצמם כמתקשרים מקומיים או כדמויות מֻכָּרות. במקרים אחרים, האקרים משתלטים על חשבון הדוא"ל של המותקף במטרה ליצור קשר עם בני משפחתו וחבריו.

לאחר מכן, מבצעים התוקפים מניפולציות על הקורבנות באמצעות איומים שונים, במטרה לסחוט אותם. בשנת 2019 דיווחו אזרחים אמריקנים על הפסדים כספיים בסך 667 מיליון דולר כתוצאה מהונאות התחזות.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: Deepfakes יכולים להקנות להונאות התחזות אמינות. עבריינים עלולים להעתיק קולות של קרובי משפחתו של הקורבן או של גורם ממשלתי המוכר לקורבן. תוקף מתוחכם יוכל גם להעתיק קולות של מספר דמויות ציבוריות ולבצע שיחות אוטומטיות למספר רב של קורבנות בו-זמנית. עבריינים מסוימים אף עשויים לשלב בקול המועתק מבטא מסוים ואף לחקות את קולותיהם של דמויות ידועות ובכך להגביר את האמון מצד הקורבן.

**הערכה**: Deepfakes עלולים לסייע לתוקפים מיומנים, המבצעים עבודת מחקר מקדימה נרחבת כדי למפות מערכות יחסים משפחתיות לטובת ביצוע חיקויים משכנעים. כבר כיום ישנם דיווחים לא מאומתים מצד קורבנות שונים, המאמינים שקולו של בן משפחתם הועתק באמצעות שימוש בבינה מלאכותית. עם זאת, סוג זה של הונאות מחייב עבודה רבה וצפוי להיות פחות נפוץ מהונאות נרחבות ובלתי מובחנות.

הונאות המבוססות על Deepfakes הכוללות חיקוי של אנשי ממשל, עשויות להיות קלות יותר



לביצוע, בהתחשב בנגישות של הקלטות קיימות, שיכולות לשמש לאימון המודל. אף על פי כן, משמעות השימוש בטכניקה הזו תהיה סטייה מאסטרטגיות הונאה שהוכחו כיעילות. תוקפים רבים מתחזים כיום לגורמי ממשל בדרג נמוך, כגון אנשי אכיפת חוק, שהקשר שלהם לקורבן סביר בלבד ושלא ניתן לאמת את זהויותיהם באופן מידי. שיחות מטעם דמויות ציבוריות בכירות עלולות להיראות חשודות בעיניי חלק מהקורבנות. עם זאת, הונאות אלו לא חייבות להיות משכנעות לחלוטין. ביצוע מניפולציות על הקורבן ויצירת תחושת דחיפות עשויים לחפות על בעיות בהעתקת הקול. כמו כן, במרבית המקרים, הונאות אלו מתמקדות בקורבנות רגישים יותר, כגון האוכלוסייה המבוגרת ובני משפחה של אנשי צבא.

### 3. סחטנות סייבר

**האיום הקיים**: סחטנות סייבר מתרחשת כאשר התוקף טוען שיש ברשותו מידע מביך על הקורבן ודורש תשלום או מידע רגיש, תוך איום לפרסמו אם לא תיענה דרישתו. לעיתים קרובות, המידע המביך הוא מידע בעל אופי מיני.

בחלק מהמקרים התוקף מקבל גישה למידע המשמש לסחיטה באמצעות מתקפת סייבר. לעיתים קרובות יותר, מדובר באיום סָרָק והמידע המביך אינו קיים. לעיתים קרובות, הסוחטים מציינים את מספר הטלפון או סיסמה של הקורבן במהלך ההתכתבות עמו, במטרה להקנות אמינות לאיום, נתונים הנאספים ככל הנראה במסגרת דליפת מידע ממאגרי נתונים. על פי ה-FBI, בשנת 2019 דיווחו אזרחים אמריקנים על הפסדים בסך 107 מיליון דולר עקב פעולות של סחטנות סייבר, שאינן כוללות תשלום כופר במסגרת מתקפות כופרה.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: סחטני סייבר יכולים להשתמש ב-Deepfakes על מנת לייצר מידע מלאכותי ומזויף שישמש לסחיטה, כך שייראה אמין ואותנטי יותר. סחטנים עלולים לשלוח לקורבנות תמונות או קטעי וידאו, בהם שולבו פניהם באופן מלאכותי במדיה פורנוגרפית, כהוכחה לכך שבידיהם גישה למידע אישי רגיש.

כפי שסחטני סייבר מחלצים היום מידע המשמש לסחיטה ממאגרי מידע מודלף, כך סחטנים המשתמשים ב-Deepfake יכולים לסרוק את הרשתות החברתיות, לאסוף תמונות וקטעי ווידאו של הקורבנות ולייצר חומרים שישמשו לצרכי סחיטה. הסחטנים יכולים להשתמש באותן הרשתות החברתיות כדי ליצור קשר עם הקורבן ועל מנת לפרסם את המידע המביך.

**הערכה**: היבטים מסוימים של תרחיש השימוש ב-Deepfakes לצרכי סחיטה כבר קיימים כיום, אך לא נמצאים בשימוש בהיקף נרחב לשם סחיטה. מאז הופעת טכנולוגיית ה-Deepfake ב-2017, הטכנולוגיה שימשה בעיקר ליצירת פורנוגרפיה מלאכותית. על פי מחקר שנערך ב-2019, 96% מקטעי ה-Deepfake שפורסמו באינטרנט היו פורנוגרפיים. השימוש בקטעי Deepfake בעלי אופי פורנוגרפי משמש בעיקר לשם פגיעה ממוקדת, בייחוד בנשים, לשם השגת מטרות פוליטיות ואישיות.

בתחילה, קטעי Deepfake פורנוגרפיים כוונו בעיקר כנגד ידוענים, מאחר ומיפוי תווי פנים דרש מספר גדול מאוד של קטעי וידאו של הקורבן. כיום, הטכנולוגיה הקיימת מאפשרת להפיק קטעי Deepfake על בסיס תמונה אחת של הקורבן. המשמעות היא שניתן לייצר פורנוגרפיה

מלאכותית באופן אוטומטית ובהיקף נרחב, תוך שימוש בפריט מדיה אחד בלבד.

קל לשער כיצד קטעי Deepfake יכולים לקדם סחטנות סייבר רחבת היקף לשם יצירת רווח. בעוד שחלק מהקורבנות יזהו את החומר המשמש לסחיטה כמזויף ולכן יסרבו לשלם, אחרים עלולים לציית לדרישת הסוחט. קורבנות עשויים לבחור לשלם מתוך חשש שאנשים אחרים, כגון בני משפחה, חברים או עמיתים לעבודה, יאמינו שהמידע אמתי. באפריל 2020, רשויות מקומיות בהודו הודיעו כי התקבלו דיווחים על סחיטות סייבר תוך שימוש ב-Deepfakes.

אף על פי כן, נתונים מצביעים על כך שהשיטות המסורתיות לסחטנות סייבר רווחיות ודורשות כישורים טכניים מעטים, כך ששימוש רחב היקף ב-Deepfake לא נראה סביר. יצירה של Deepfake מאתגרת יותר מבחינה טכנית ודורשת תוכנה שיכולה למזג חלקי תמונות במקטעי מדיה, וכן יצירת קשר עם הקורבן. סביר להניח שהשימוש ב-Deepfake לצרכי סחיטה יהיה יעיל יותר בפנייה אישית.

## תרחישים המתמקדים בחברות

### 4. הונאות תשלום

**האיום הקיים**: ניצול דוא"ל עסקי (Business email compromise) הוא מונח המתאר סוגים של הונאות תשלום המתמקדות בחברות. לעיתים קרובות, התוקפים פורצים לחשבון דוא"ל או מזייפים כתובת דוא"ל של מנכ"ל החברה ויוצרים קשר עם דמות בעלת סמכויות פיננסיות בחברה במטרה לבצע העברת כספים דחופה. עבריינים גם עשויים להתחזות לספקים מוכרים או לעובדים. מקרים מורכבים יכולים אף לכלול התחזות לתקופות של שבועות וחודשים במטרה לייצר מערכת יחסים עם הקורבנות ולהפעילם, לרוב באמצעות שיחות טלפון ודוא"ל.

על פי ה-FBI, בשנת 2019 עסקים בארה"ב דיווחו על הפסדים של יותר מ-1.7 מיליארד דולר עקב הונאות תשלום, קרוב למחצית מכלל ההפסדים מפשעי סייבר.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: תוקפים יכולים להוציא לפועל שיחות טלפון מבוססות Deepfake במטרה להקנות אותנטיות. תוקפים יכולים להשתמש ב-Vishing המבוסס על Deepfake ולהימנע מהצורך לפרוץ או לזייף חשבון דוא"ל. חלק מהנמענים עלולים להאמין לשיחות וידאו מבוססות Deepfake אף יותר מלשיחות קוליות.

**הערכה**: קיים תיעוד לשימוש ב-Deepfake לביצוע הונאות בהיקפים קטנים. בשנת 2019, עבריינים השתמשו בטכנולוגיה להעתקת קול על מנת להתחזות למנכ"ל בחברה גרמנית והצליחו לשכנע עובד בריטי הכפוף לו להעביר לידיהם 243,000 דולר. חברת הביטוח של הקורבן הודיעה כי סביר להניח שהעבריינים השתמשו בתוכנה לשימוש מסחרי.

תוקף שאפתני יכול לשלב קטעי Deepfake בשיחות וידאו, דבר שעשוי להפוך אותן למשכנעות יותר. הטכנולוגיה הקיימת מאפשרת לבצע החלפות פנים במהלך שיחת וידאו, ומאחר ושיחות וידאו הן לעיתים בעלות איכות ירודה, משתתפי השיחה ככל הנראה יתעלמו משיבושים באיכותן, שמקורם הוא בקטעי ה-Deepfake.



### 5. מניפולציה על מחירי מניות באמצעות מידע כוזב

**האיום הקיים**: האינטרנט מספק אמצעים רבים להפצת מידע כוזב המיועד להשפיע על מחירי המניות. לעיתים קרובות, גורמים זדוניים אנונימיים מפיצים מידע כוזב על מניה מסוימת בבלוגים, בפורומים וברשתות החברתיות באמצעות בוטים, ספאם ואמצעים נוספים. פעולות אלו נועדו לייקר או להוזיל באופן מלאכותי את מחיר המניה, לשם רווח מהיר. חברות קטנות הן המטרה הנפוצה ביותר, מאחר וקל יותר להשפיע על מחירי המניות שלהן. עם זאת, גם חברות גדולות נפלו לעיתים קורבן למזימות אלו, המונעות ממניעים פוליטיים וכלכליים.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: קטעי Deepfake עלולים לפגוע בשווי המניה באמצעות יצירת מידע ונרטיבים כוזבים שנראים מהימנים. דוגמה לכך היא יכולתו של גורם זדוני לפרסם קטע Deepfake, המציג כביכול מנכ"ל של חברה מודיע על פשיטת רגל, מודה בביצוע עבירות או משמיע הערות או דעות פוגעניות.

לחלופין, ניתן לייצר קטע Deepfake המציג אירועים חיוביים פיקטיביים במטרה להעלות את ערך המניות. דוגמה לכך היא היכולת לייצר קטע Deepfake המציג דמות מוכרת או פוליטיקאי המשתמשים במוצר או בשירות מסוים וממליצים עליו.

**הערכה**: Deepfake ברמת איכות גבוהה המופץ ברשתות החברתיות או כספאם בדוא"ל, יכול להוות כלי יעיל לביצוע מניפולציות במניות של חברות קטנות. לעיתים קרובות, לחברות קטנות אין מספיק משאבים ויכולת להתגונן בפני הונאות. גם אם קטע ה-Deepfake נחשף במהרה, עבריינים עדיין יכולים להפיק רווחים ממסחר בטווח קצר.

Deepfakes עשויים להוות איום גם עבור חברות גדולות, שמחירי המניות שלהן עמידים יותר בפני מניפולציות. ההופעות של מנהלי החברות הגדולות בתקשורת יוצרת מאגר של תמונות, קטעי ווידאו והקלטות, המאפשרים לגורמים זדוניים ליצור Deepfake באיכות גבוהה.

דוגמה לתרחיש מזיק במיוחד היא השימוש בקטעי Deepfake המציגים מנכ"ל או בכיר בחברה מציג דעות או הערות אישיות פוגעניות, מיזוגניות או גזעניות. בהינתן העובדה שלא ניתן להוכיח שהערה או שיחה פוגענית לא התקיימה, על המנכ"ל או הבכיר להסתמך על מוניטין חיובי במטרה לנהל את המשבר. תרחיש זה עלול גם להוביל לערבוב בין האמת לשקר. Deepfake המציג בכיר בחברה אומר או מבצע מעשה פוגעני עלול להוביל לקבלת תלונות על מקרים אמתיים של אפליה או השפלה.

גם אם קטע ה-Deepfake יוכח כמזויף, סביר להניח שהוא יוביל להשלכות שליליות ארוכות טווח על מוניטין החברה. כמו סוגים אחרים של מידע כוזב, Deepfakes עלולים ליצור השפעה פסיכולוגית מתמשכת על חלק מהציבור, גם לאחר שהוכח כי הם מזויפים. ניסויים שנערכו במסגרת מחקרים הראו שמיעוט לא מבוטל של אנשים יאמינו באמינותם של קטעי ה-Deepfake גם לאחר שכבר הוכח כי הם מזויפים.

### 6. מניפולציה על מחירי מניות באמצעות בוטים

**האיום הקיים**: השפעה על מחירי מניות עלולה גם להתבצע גם באמצעות הצגת פעילות כוזבת המונית כנגד חברה. דוגמה לכך היא הצגת רושם של מחאה המונית כנגד חברה או תאגיד מסוים ברשתות החברתיות.

השימוש בבוטים ברשתות חברתיות לשם השפעה על הרגש הציבורי כלפי חברות מסוימות אינו חדש. גורמים זדוניים מנהלים במקביל מספר גדול של חשבונות ברשתות החברתיות, המפרסמים באופן מתואם מסרים שנועדו לקדם או לפגוע בחברה כלשהי. עליות בפעילות הבוטים ברשתות החברתיות נמצאו כקשורות לשינויים קטנים וזמניים במחירי מניות.

הרשתות החברתיות עוקבות אחר רמזים המסייעים בזיהוי בוטים. רמזים אלו כוללים הקמת מספר רב של חשבונות בנקודת זמן מסוימת; שימוש בתמונות פרופיל גנובות או בתמונות ממאגרי מידע; פרסום פוסטים בתדירות גבוהה שלא ניתן להסביר אותה ועוד. הרשתות החברתיות משתמשות גם בלמידת מכונה על מנת לזהות דפוסי התנהגות וחשבונות חשודים.

למרות זאת, השימוש בבוטים ברשתות חברתיות נותר גבוה. ראשית, לא כל החשבונות האוטומטיים מפרים את המדיניות של הרשתות החברתיות וכן קשה להבחין בין בוטים מזיקים המפרים את מדיניות הרשתות החברתיות לבין בוטים בלתי מזיקים. שנית, ההתנהגות של בוטים זדוניים מתרחשת בהיקף נרחב ומתפתחת במהירות, דבר המקשה על היכולת לעקוב אחריה. שלישית, השימוש באלגוריתמים לזיהוי בוטים אינו מושלם ולא מן הנמנע שבוטים יזוהו כמשתמשים אנושיים או להיפך. לבסוף, פעולות נרחבות לניקוי הרשת החברתית מחשבונות מזויפים עלולות להרתיע משקיעים, שרוצים לראות גידול במספר המשתמשים ברשתות, או את המשתמשים הקבועים, המעוניינים במספר רב של עוקבים ברשת. זיהוי והסרה של בוטים היא משימה קשה, גם כאשר לרשתות יש גישה לאמצעי בינה מלאכותית מתקדמים, שאין ברשות גורמים זדוניים.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: גורמים זדוניים עלולים להשתמש בטכניקת למידה עמוקה במטרה ליצור רשתות בוטים מבוססות בינה מלאכותית, המתחזות למשתמשים אנושיים ומתחמקות מזיהוי. המטרה הסופית עשויה להישאר דומה: הצגת רגשות ציבוריים המוניים ביחס לחברה כלשהי במטרה להשפיע על מחיר המניה שלה. גורם זדוני עלול להמציא מגמה חדשה או להעצים מגמה מתפתחת.

גורמים זדוניים משתמשים כבר היום בבינה מלאכותית על מנת לייצר תמונות פרופיל מזויפות, דבר המקשה על זיהוי שימוש חוזר בהן. השלב הבא יכלול ככל הנראה ניסוח פוסטים מלאכותיים על ידי אלגוריתמים.

בעוד שבוטים קיימים משכפלים פוסטים או יוצרים פוסטים אקראיים באופן גס, בוטים המבוססים על בינה מלאכותית יכולו לפרסם פוסטים חדשים ומותאמים אישית למשתמשים.

בוטים אלו יהיו מודעים לפוסטים הקודמים שהם פרסמו, ויוכלו לשמר מאפיינים קבועים לאורך זמן, כגון מאפיינים אישיותיים, סגנון התבטאות, תחומי עניין וביוגרפיה. הבוטים המשכנעים ביותר יצברו עוקבים אנושיים, יעצימו את ההשפעה של המסרים שלהם ויקשו אף יותר על גילוים. לעיתים קרובות, בוטים קיימים עוקבים אחד אחר השני במטרה להגביר את הנראות שלהם, דבר המוביל לזיהוי דפוסי התנהגות קבוצתיים ומֻקָל על זיהוים.

בוטים המבוססים על בינה מלאכותית יכולים לפעול עם השגחה מינימלית במשך חודשים ואפילו שנים, במטרה ליצור אמינות ולבסס השפעה. כל בוט ברשת יוכל להתחיל לפרסם פוסטים על חברה כלשהי, באמצעות סגנון המיוחד לו ובהתאם לאישיות שנבנתה עבורו. הקמפיין עשוי לשקף דעת קהל של צרכנים מהשורה ובכך להשפיע על מחיר המניה.



**הערכה**: ישנם אמצעים כיום המאפשרים לזהות תמונות וטקסטים שנוצרו באמצעות בינה מלאכותית, שני מאפיינים מרכזיים המרכיבים בוטים המבוססים על בינה מלאכותית. בנוסף, הרשתות החברתיות יכולות לעיתים לחשוף בוטים על סמך התנהגותם ויחסי הגומלין שהם מנהלים ולא רק על סמך התוכן שהם מייצרים. עם זאת, המנגנון לזיהוי בוטים לא מושלם, אפילו ביחס לבוטים קיימים, הפועלים באופן גס יחסית. אף רשת חברתית לא פתרה בהצלחה את סוגיית הבוטים, וכן הופעתם של בוטים בעלי יכולות מתקדמות יותר תאלץ אותן לשכלל את טכניקות הזיהוי שלהן.

עבור גורמים המבצעים מניפולציות במניות, בוטים מבוססי בינה מלאכותית יכולים לנצל את רצונם של המשקיעים לבחון מגמות בקרב צרכנים ברשתות החברתיות. מספר הולך וגדל של חברות בתחום הטכנולוגיה הפיננסית (Fintech) משווקות אמצעים לניתוח מגמות ציבוריות ואת רגשות משתמשי הרשתות החברתיות כלפי חברות. מידע בנושא מגמות ורגשות המשתמשים משמש לטובת מסחר אלגוריתמי, המתנהל ללא התערבות אנושית. שימוש נרחב יותר בניתוח נתונים על מגמות ורגשות משתמשי הרשתות החברתיות ושילובם במסחר אלגוריתמי, עלולים לשפר את יכולתם של בוטים מבוססי בינה מלאכותית לבצע מניפולציות על מחירי המניות.

### 7. משיכות כספים המוניות מבנקים

**האיום הקיים**: בנקים חוו בעבר אירועי משיכת כספים המונית,[21] שהושפעו במידה מסוימת משמועות בנושא מצב פיננסי רעוע שהופצו ברשתות החברתיות. הרשתות החברתיות כשלעצמן אינן גורם ראשי למשיכת כספים המונית. שמועות על מצב פיננסי רעוע או שלילי הן בדרך כלל תגובה לבעיות במגזר הבנקאי, זאת למרות שלעיתים השמועות מקדימות את המציאות. הרשתות החברתיות, לצד התקשורת המסורתית והעברת מידע מפה לאוזן, מספקות מרחב בו ניתן להפיץ שמועות. לפיכך, יכולות הרשתות החברתיות להגביר חששות ציבוריים קיימים בנושא מצבו הפיננסי של בנק מסוים.

קשה לקבוע את מקורן ואמיתותן של שמועות בנושא הבנקים. בשנת 2014, ממשלת בולגריה האשימה גורמים, שאת זהותם לא חשפה, בפגיעה מתואמת במוניטין של מספר בנקים שסבלו ממשיכות כספים המוניות. הפצת השמועות כללה שימוש במסרונים, פוסטים והדלפות ברשתות החברתיות. תקריות אלו מציגות דפוס פעולה שגורמים זדוניים יכולים לאמץ בעתיד.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: רשתות בוטים המבוססים על בינה מלאכותית יכולות לעורר או להעצים שמועות שעלולות להוביל למשיכות כספים המוניות. כמו כן, גורמים זדוניים עלולים לפרסם קטע Deepfake ברשתות החברתיות, המציג כביכול גורם בכיר בבנק או גורם ממשלתי, מתארים בעיות נְזִילוּת חמורות בבנק. ניתן לשער כי השימוש בקטעי Deepfake במטרה להוביל למשיכות כספים המוניות מהבנקים, יתרחשו בעתות משבר במערכת הפיננסית במדינה.

**הערכה**: גורמים זדוניים עשויים דווקא להימנע מהשימוש במדיה מלאכותית וברשתות בוטים מבוססי בינה מלאכותית, זאת על אף יעילותם להפצת שמועות ברשתות החברתיות. הסיבה לכך היא שישנן דרכים יעילות ופשוטות הרבה יותר, כגון השימוש בתמונות ובקטעי וידאו תוך

21 Bank runs

הוצאתם מהקשרם המקורי. דוגמה לכך היא השימוש בתמונות מהעבר, המתארות משבר פיננסי בבנק במדינה אחרת והצגתן כתמונות המתארות משבר נוכחי בבנק אחר.

## תרחישים המתמקדים במגזר הפיננסי

### 8. יצירת קריסה בבורסה ופגיעה במסחר

**האיום הקיים:** ב-23 באפריל 2013, קבוצת ההאקרים המכונה הצבא האלקטרוני הסורי השתלטה על חשבון הטוויטר של סוכנות הידיעות Associated Press ו"צייצה" מידע כוזב, לפיו אירעו שני פיצוצים בבית הלבן וכי נשיא ארה"ב דאז, ברק אובמה, נפצע. הפרסום הוביל למפולת מידית וחדה במסחר. בשלוש דקות בלבד, מדד ה-S&P 500 איבד 136 מיליארד דולר מערכו, וכן מחירי הנפט ותשואות אגרות החוב הממשלתיות צנחו אף הם.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: גורמים שפועלים ממניעים פוליטיים או פיננסיים עלולים לפגוע במסחר תוך שימוש ב-Deepfakes. דוגמה לכך היא פרסום הקלטה מלאכותית של שר הנפט של ערב הסעודית או של רוסיה, המתמקחים ביניהם על מכסות ייצור הנפט, מתוך כוונה לבצע מניפולציות על מחירו.

חלק מפרסומים אלו ניתנים להפרכה בקלות יחסית. אישיות בעלת מוניטין של אמינות המופיעה בקטע Deepfake תוכל להכחיש את הפרסום, בגיבוי עדויות משכנעות. עם זאת, התפוגגות ההשפעה של פרסום הקטע תיקח זמן רב יותר כשמדובר באישיות בעלת מוניטין רעוע ובעיית אמינות.

**הערכה**: קטע Deepfake משכנע עשוי להיות בעל השפעה ניכרת יותר מהפצת מידע כוזב באמצעות השתלטות על חשבונות ברשתות החברתיות. קטעי וידאו כאלו מנצלים נטייה פסיכולוגית של בני אדם להאמין ולזכור מידע חזותי טוב יותר מסוגים אחרים של מידע. ניתן להשתמש ב-Deepfake גם לשם הפצת מידע כוזב ברשתות החברתיות ובתקשורת המסורתית, במקום לפרוץ לעמוד טוויטר של סוכנת ידיעות חשובה, כפי שנעשה ב-2013. אף על פי כן, יצירת Deepfake יעיל מחייבת כישורים טכניים ותחכום פוליטי וכן יצירת מידע כוזב שישטה בדעת הקהל וישפיע על השווקים אינה משימה קלה לביצוע.

## תרחישים המתמקדים בבנקים מרכזיים ובמוסדות רגולציה פיננסיים

### 9. הפצת מידע כוזב בנושא מהלכים רגולטוריים כלכליים או פיננסיים

**האיום הקיים**: בנקים מרכזיים ורגולטורים פיננסיים ברחבי העולם, נאבקים בשמועות על מהלכים כלכליים ופיננסיים עתידיים. ב-2019 הכחישו הבנקים המרכזיים של הודו ומיאנמר שמועות לפיהן הרשויות צפויות לסגור מספר בנקים.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: גורמים זדוניים עלולים להשתמש ב-Deepfakes על מנת להפיץ הקלטות המציגות גורמים בנקאיים מרכזיים, שדנים בבעיות פיננסיות כגון בעיות נְזִילוּת. ניתן גם למקד קטעי Deepfakes נגד גורמים בבנקים מרכזיים או במוסדות רגולציה פיננסיים למטרות רווח פוליטי. דוגמה לכך היא הפצת קטע Deepfake המציג גורם פיננסי



ממשלתי בכיר, מודה בקבלת שוחד מגוף עסקי בתמורה להפסקת החקירה נגדו.

**הערכה**: לקטעי Deepfake כאלו עשויה להיות השפעה גדולה יותר במדינות בהן הרגולטורים הפיננסיים סובלים מראש ממשבר אמון. אמון הוא אמצעי חשוב המסייע להפריך את אמינותם של קטעי ה-Deepfake, בייחוד כאשר הגוף הנפגע צריך להוכיח שמה שהוצג בקטע לא התרחש. בעתות משבר כלכלי, קטעי Deepfake עלולים לנצל חששות כלכליים ולהעצים אותם.

גופי רגולציה עלולים לסבול מרמת אמון ציבורי נמוכה גם במדינות גדולות. תגובה ממשלתית כושלת להפצת קטעי Deepfake המופנים נגדה עשויה להאריך את פרק הזמן בו עלולים הגורמים המפיצים לגרום לנזק ולהפיק רווחים מהמצב שנוצר.

### 10. יצירת רושם מוטעה של תמיכה ציבורית במהלכים בקרב מקבלי ההחלטות

**האיום הקיים:** גופי רגולציה צריכים להתמודד עם תופעת האסטרוטורפינג (Astroturfing), תופעה בה גורמים זדוניים מנסים בחשאי להשפיע על מקבלי החלטות, באמצעות יצירת רושם כוזב של תמיכה ציבורית בעמדות ובמהלכים מסוימים והצגתם כפופולריים. הנציבות האמריקנית לניירות ערך (SEC)[22] והלשכה האמריקנית להגנת הצרכן,[23] חוו שימוש לרעה במערכות דיגיטליות, במסגרתן נשלחו תגובות רבות מטעם הציבור בנושא הצעות רגולציה. במקרים רבים, שתדלנים ופעילים פוליטיים הגישו תגובות רבות בשמם של אנשים פיקטיביים, תוך שימוש בשמם של אנשים שנפטרו, או בצירוף כתובות דוא"ל מזויפות.

**כיצד מדיה מלאכותית עלולה להגביר את האיום**: מדיה שנוצרה באמצעות בינה מלאכותית עשויה להגביר את התופעה ולהקנות לה אמינות. אלגוריתמים המייצרים טקסטים מלאכותיים יכולים לכתוב על כל נושא ובכל היקף. גורמים המעורבים באסטרוטורפינג עלולים להשתמש באלגוריתמים לניסוח טקסטים על מנת לפרסם מספר רב של תגובות מזויפות, בעד או נגד הצעה לרגולציה פיננסית. תגובות אלו יכולות להיות מגוונות בתוכן ובשפתן וכן השימוש באלגוריתמים אלו יכול להקשות על טכניקות קיימות המשמשות לזיהוי קמפיינים לאסטרוטורפינג.

**הערכה**: כיום, טקסטים המיוצרים באמצעות בינה מלאכותית יכולים להיות שונים באיכותם וברמת האמינות שהם מעוררים. גורמים המעורבים באסטרוטורפינג עלולים להשתמש באלגוריתמים שאומנו לייצר טקסטים בסגנון מסוים במטרה להקנות להם אמינות.

אולם, ישנן דרכים להבדיל בין טקסטים שנכתבו על ידי בני אדם לטקסטים שנוצרו באמצעות בינה מלאכותית. כתיבה אנושית נוטה לכלול בחירת מילים יצירתית ובלתי צפויה, בעוד שטקסטים מלאכותיים נוטים להתבסס על דפוס כתיבה וניתן לאמן אלגוריתמים שיזהו אותם. טקסטים ארוכים יותר יוכלו לסייע לאלגוריתמים לזיהוי טקסטים בקבלת החלטות. עם זאת, אלגוריתמים אלו אינם חסינים מפני טעויות וניתן לשטות בהם. אלגוריתמים לניסוח טקסטים מלאכותיים יוכלו לפרסם טקסטים יותר לא-שגרתיים, אך כאלו שיכללו שימוש בסגנון אנושי יותר.

על פי דו"ח של הסנאט האמריקני משנת 2019, אף אחת מ-14 הסוכנויות הפדראליות שנבדקו, לא הפעילה אמצעים כדי לבדוק את אמינותן של תגובות המגיעות מהציבור. בניסוי שנערך בהרווארד, כל התגובות המלאכותיות שנוצרו נשלחו בהצלחה לסוכנות פדראלית, בטרם הוסרו

22 Securities and Exchange Commission

23 Consumer Financial Protection Bureau

באופן וולונטרי. תגובות מלאכותיות אלו היוו את מרבית התגובות שהסוכנות קיבלה באותו זמן. עם זאת, ההשפעה של האסטרוטורפינג נתונה במחלוקת. למרות הממצאים שתוארו בניסויים ובבדיקות השונות, סוכנויות ממשלתיות מעניקות משקל רב יותר לתגובות שמגיעות מעסקים וארגונים ידועים, מאשר מאלו שמקורן באזרחים אלמונים. בכל זאת, האסטרוטורפינג מהווה בעיה רגולטורית מתפתחת שעלולה לערער את אמון הציבור בתהליך החקיקה. בסיכומו של דבר, מדיה מלאכותית עלולה להגביר ולחזק את השפעת תופעת האסטרוטורפינג.

## יישומי מדיניות

האיום הכולל ליציבות הפיננסית ולמערכת המאקרו הכלכלית

עשרת התרחישים שהוצגו אינם מתארים איום רציני על יציבות המערכת הפיננסית העולמית או על השווקים הלאומיים בכלכלות חזקות. ככלל, נראה שכלכלות חזקות חסינות בפני השפעות של הפצת מידע כוזב, ללא קשר לטכניקת תפוצתו. עד להמצאת טכניקת ה-Deepfake, הפצת מידע כוזב או מוטעה הייתה לעיתים בעלת השפעת רוחב בשוק, אך בעשורים האחרונים מידע כוזב הוביל להשפעות מוגבלות בלבד ולטווח קצר. על מנת לערער את יציבות השוק, על ההשפעה של המדיה המלאכותית להיות חזקה יותר מהשיטות הקיימות להפצת מידע כוזב. בעת כתיבת שורות אלו, אין סיבה לצפות לכך.

סביר להניח שמדיה מלאכותית תוביל לפגיעה כלכלית באינדיבידואלים או בעסקים ספציפיים. פגיעה שכזו, יכולה לבוא לידי ביטוי באמצעות הונאה או מניפולציה על מחירי מניות ויכולה להיות משמעותית מנקודת מבטו של הקורבן. עם זאת, לא סביר שהמדיה המלאכותית תוביל לנזק מצטבר, שיהיה בעל השפעה ברמה הרחבה. הונאות כלכליות הן דבר שבשגרה, אפילו בכלכלות מפותחות. על מנת לחולל השפעות מאקרו-כלכליות, הנזק המצטבר ממדיה מלאכותית יצטרך לעלות בהיקפו על הנזק שנגרם ממרבית הפעילויות הלא-חוקיות הקיימות כיום. זהו מצב שכמעט בוודאות לא צפוי להתרחש.

אף על פי כן, שווקים מתפתחים ניצבים בפני איומים גדולים יותר מצד המדיה המלאכותית. כבר היום, מדינות בעלות כלכלות פחות יציבות, בהן המוסדות, הכוללים גם את ענף הפיננסים, זוכים לפחות אמון, נאבקות יותר עם מידע כוזב בעל אופי פיננסי. השימוש במדיה מלאכותית עלול להחמיר את המצב. גם מדינות מפותחות, שנמצאות במשבר כלכלי, עלולות להיות פגיעות יותר בפני שימוש זדוני במדיה מלאכותית. דוגמה לכך היא התרחיש השביעי שהוצג במסמך זה ותיאר כיצד מדיה מלאכותית עלולה לתרום למשיכות כספים המוניות מהבנקים. אם המערכת הבנקאית גם כך לא נהנית מאמון ציבורי גבוה, השימוש ב-Deepfakes עלול לעורר בהלה ציבורית ולהוביל לקריסת בנקים, הנמצאים כבר במשבר. הפרכת האמינות של קטעי ה-Deepfakes עלולה להיות קשה יותר במצב שבו הבנקים וגופי הרגולציה הממשלתית לא נהנים מאמון ציבורי גבוה.

שימוש במדיה מלאכותית לעומת השימוש באמצעים זדוניים אחרים

שימוש זדוני בקטעי מדיה מלאכותית עלול להוביל לנזקים, אולם הפוטנציאל של שימוש זדוני זה טרם נקבע וכן הוא מורכב יותר לביצוע בהשוואה לשיטות פליליות אחרות. היקף השימוש הזדוני במדיה מלאכותית ייקבע בעיקר על בסיס שיקולי עלות-תועלת. עברריינים



יעריכו את התועלת ואת העלות שבשימוש ב-Deepfake לעומת השימוש באמצעים הקיימים כגון Cheapfake או Shallowfake.[24]

קטעי Cheapfake הוכחו כיעילים, והובילו לנזקים גם במישור הפיננסי. כל התרחישים שהוצגו, מתארים כיצד ניתן כבר היום לבצע הונאות מסוגים שונים, ללא שימוש בבינה מלאכותית. בחלק מהמקרים, האמצעים הקיימים היום הם יעילים ביותר, כגון במקרה של גניבת זהות. השכיחות הגבוהה של מניפולציות שונות במדיה מעידה, שמומחיות טכנית היא לא תנאי הכרחי לביצוע הונאות. אפילו במצב שבו בינה מלאכותית תתפתח ותשתכלל יותר, גורמים זדוניים עדיין עשויים למצוא את השיטות המסורתיות כיעילות ופשוטות יותר.

אף על פי כן, למדיה מלאכותית יש מספר יתרונות חשובים: ראשית, היא יכולה להיות מציאותית ברמה גבוהה מאוד. בטכניקות אחרות כגון החלפת פנים, שאינה מבוססת על Deepfake, יש צורך ביכולות עריכה ברמה גבוהה ביותר. שנית, כמויות קטעי המדיה המלאכותית שניתן לייצר יכולה להיות אינסופית. שלישית, ניתן לייצר מדיה מלאכותית בהתאמה אישית. אלגוריתמים לניסוח טקסטים יכולים ללמוד במהירות סגנונות כתיבה וז'אנרים חדשים, בעוד שלבני אדם יקחו מספר שנים לעשות כן.

הבחירה של גורם כלשהו במדיה מלאכותית לעומת מדיה שעברה מניפולציה תלויה באסטרטגיה שלו וביכולותיו. הבחירה בטכניקה מסוימת נובעת משיקולי עלות ותועלת וכן רבות מההונאות אינן מתוחכמות, זולות לביצוע וכן מניבות רווחים קטנים. לפיכך, השימוש ב-Deepfake יהיה פחות ישים ורווחי בהשוואה לשיטות הקיימות. השימוש ב-Deepfakes יוכל להיות רלוונטי, עובר הונאות מתוחכמות, במסגרתן מבצעים העבריינים תחקירי עומק ומתאימים את ההונאה לאופי הקורבן, מתוך תקווה לחלץ ממנו תשלומים גבוהים יותר.

במקרים מסוימים, יכולותיו של העבריין יהוו גורם מכריע. עבריין בעל כישורים חברתיים וכישורי שפה חזקים, יעדיף להשתמש בטכניקות מסורתיות של vishing, כדי להונות עסקים בטלפון. עברייינים פחות כריזמטיים או בעלי חוש טכני יותר, יעדיפו להשתמש בקול מועתק מאשר בקולותיהם שלהם.

הפצת נרטיבים והעצמת משברים קיימים

הפצת מדיה מלאכותית בהיקף נרחב צפויה להוביל להשלכות נרחבות יותר, כשהמטרה היא העצמת נרטיבים קיימים. השימוש במדיה מלאכותית יכול להעצים את הנראות של מחאה, אמתית או מזויפת כנגד חברה או תאגיד במסגרת תרחיש בו מתקיים כבר חרם צרכנים. במצב שכזה, האמינות המעורערת של החברה תקשה על המאבק בהפצת המידע הכוזב.

חברות, מוסדות פיננסיים ורגולטוריים ממשלתיים, שניצבים בפני משבר ביחסי הציבור שלהם, פגיעים במיוחד לשימוש במדיה מלאכותית. על ארגונים להיערך להתמודדות מול השימוש באמצעים אלו כחלק מתכנית לניהול משברים. נקיטת צעדים שונים, כגון יצירת אמון עם קהלי מפתח של הארגון, עשויים לצמצם את הנזק.

24 אמצעים מסורתיים יותר לזיוף וביצוע מניפולציה על קטעי מדיה שאינם משתמשים בבינה מלאכותית.

## השוואה לקטעי Deepfake פוליטיים

הפרשנות הנרחבת על Deepfakes פוליטיים סייעה לעצב את התפיסה הציבורית הכללית על מדיה מלאכותית. ראשית, במצב שבו טוהר הבחירות הוא ערך מרכזי, Deepfakes נתפסים בדרך כלל כאמצעי להשפיע על דעת הקהל באופן נרחב וכאמצעי המתפשט ברשתות החברתיות ובמדיה המסורתית. שנית, אמצעי התגובה המשפטיים להתמודדות עם Deepfakes אינם יעילים, זאת בשל ההגנה על חופש הביטוי הפוליטי ובשל העובדה, שתהליכים משפטיים מתקדמים לאט. שלישית, השימוש ב-Deepfakes מעלה שאלות כבדות משקל: דוגמה לכך היא השאלה כיצד ניתן להבדיל בין סאטירה לגיטימית להטעיה בלתי חוקית? שאלה נוספת נוגעת לזהות הגורם שצריך להגדיר מה מותר ומה אסור ולאכוף את החוקים.

הניתוח של השימוש ב-Deepfakes כנגד מטרות בענף הפיננסים, הופך את התפיסה הרווחת למסובכת יותר. כמו במישור הפוליטי, Deepfakes בענף הפיננסים ישמשו לצורך מניפולציה על קהלים רחבים, אך תרחישים רבים שהוצגו במסמך זה כוללים הונאה של יעד מוגדר וספציפי. מענים משפטיים לסוגיית ה-Deepfakes בענף הפיננסים צפויים להיתקל בקשיים, כגון זיהוי ונקיטת צעדים כנגד עבריינים במדינות זרות. עם זאת, ניתן להשיב אובדן פיננסי, דבר שכמעט אינו קיים במישור הפוליטי, היכן שתרחיש של ביטול תוצאות הבחירות נדיר למדי. לבסוף, חברות צריכות למצוא את הבסיס המשותף למאבק בשימוש במדיה מלאכותית לטובת עבריינות פיננסית.



## תגובות מגוונות מצד בעלי העניין

כל עשרת התרחישים מציגים רצף אירועים משותף, בו גורם זדוני מייצר מדיה מלאכותית ומפיץ אותה לקהל יעד. חלקים מקהל היעד מייחסים למדיה המלאכותית אמינות וכך נוצר הצורך של הקורבן להגיב.

למרות שמסמך הזה לא נועד לספק הנחיות מפורטות בתחום המדיניות, רצף האירועים המוצג בתרחישים יכול לספק מסגרת בסיסית לגיבוש תגובה ברמת המדיניות (ראו טבלה 1). כל אירוע מערב גורמים ובעלי עניין שונים ולכן כל אירוע מציע הזדמנויות שונות בתחום המדיניות. חלק מהמענים המפורטים בטבלה 1 כבר נחקרו בעבר וכן חלק מהמענים יכולים להתבסס על מענים קיימים להתמודדות עם מידע כוזב ופגיעה כלכלית.

| | יצירת מדיה מלאכותית | הפצת מדיה מלאכותית | תגובת קהל היעד | תגובת הקורבן |
|---|---|---|---|---|
| **דרכי התמודדות אפשריות** | - הקמת מנגנוני בקרה וקוד אתי למחקר ופיתוח של בינה מלאכותית<br>- הקמת מנגנוני בקרה וקוד אתי עבור הפצה של טכנולוגיות בינה מלאכותית | - זיהוי ומעקב אחר מדיה מלאכותית<br>- בקרת תכנים<br>- מגנונים לאימות זהות<br>- מנגנונים למניעת זיופים<br>- מנגנונים לאנטי-ספאם<br>- אכיפת מדיניות נגד בוטים<br>- בדיקת עובדות<br>- שיתוף מידע ומודיעין איומים<br>- צעדי נגד משפטיים | - חינוך הציבור<br>- הכשרת עובדי החברה/ הארגון<br>- הכשרת עובדי המוסדות הפיננסיים להתמודדות עם האיום | - צעדים מקדימים לבניית אמון<br>- תכנון מקדים של התמודדות<br>- איסוף ראיות להפרכת נרטיבים כוזבים<br>- קמפיין יחסי ציבור |
| **בעלי עניין רלוונטיים למענים בתחום המדיניות** | - חוקרי בינה מלאכותית<br>- מפתחי טכנולוגיות בינה מלאכותית<br>- משקיעים בטכנולוגיות בינה מלאכותית | - חברות המדיה החברתיות<br>- גופי המדיה והתקשורת המסורתית<br>- ספקי שירותי טלפוניה ו-VoIP<br>- ספקי שירותי שיחות ווידאו<br>- ספקי שירותי דוא"ל<br>- סוכנויות המודיעין<br>- גורמי אכיפת חוק ומוסדות רגולציה פיננסיים | - הציבור הרחב<br>- עסקים וחברות<br>- מוסדות פיננסיים<br>- ממשלות<br>- ארגוני החברה האזרחית | - עסקים וחברות<br>- מוסדות פיננסיים<br>- בנקים מרכזיים ומוסדות רגולציה פיננסיים<br>- עיתונאים |

מחקרים קודמים שנערכו בנושא מדיה מלאכותית וטוהר הבחירות, מצאו כי אף גורם בעל עניין לא מסוגל להתמודד לבדו עם האתגרים. מענים מוצלחים יחייבו שימוש בטכנולוגיה חדשה,

שינויים בהתנהלות הארגונית ושינויים חברתיים. כמו כן, על המגזר הפיננסי להעריך את תפקידו בגיבוש מדיניות נגד מדיה מלאכותית.

במטרה לנסח מדיניות כללית, על המוסדות הפיננסיים והרגולטורים לחלק את פעילותם לשלושה חלקים:

**(1)** פעולות פנימיות, הכשרות ומגנוני בקרה, כגון הערכה מחדש של אמצעי אימות זהות לקוחות;
**(2)** פעולות בקרב התעשייה, כגון שיתוף מידע והרחבת מנגנונים לשיתוף מודיעין איומי סייבר כך שיכלול גם את איום המדיה המלאכותית;

**(3)** הרחבת שיתופי הפעולה בין בעלי העניין לבין גורמים מבחוץ, כגון חוקרי בינה מלאכותית, חברות הטכנולוגיה, גופי הממשלה וארגוני החברה האזרחית.

על המגזר הפיננסי למצוא דרך לבטא את הסוגיות המיוחדות לו בקרב שותפים פוטנציאליים, להתעדכן בפעולות ומענים המיושמים בתעשיות אחרות ולקדם שיתופי פעולה.

## מסקנות

מדיה מלאכותית לא צפויה לאיים על היציבות הפיננסית העולמית או על הכלכלה ברמת המאקרו. לעומת זאת, אינדיבידואלים, חברות ומוסדות ממשלתיים הם מטרות פגיעות יותר, כמו גם כלכלות מתפתחות ומערכות פיננסיות הנמצאות במשבר. מדיה מלאכותית מצטרפת לרשימה ארוכה של אמצעים מזיקים שקיימים כבר היום. גורמים בעלי יכולות טכניות מתקדמות ישלבו ככל הנראה בין אמצעים מסורתיים ולבין מדיה מלאכותית, בעוד שגורמים בעלי יכולות בסיסיות בלבד צפויים לבחור באמצעים המסורתיים. איומי המדיה המלאכותית על המגזר הפיננסי דומים לאלו הקיימים במישור הפוליטי, אך גם שונים מהם. ניתן להפיק תובנות לגורמים רלוונטיים בשני התחומים.

מקבלי החלטות המודאגים מאיום המדיה המלאכותית ניצבים בפני אי וודאות. מצד אחד, הטכנולוגיה הזו מעוררת דאגות מובנות. מדיה מלאכותית יכולה להיראות יותר מציאותית ולהיווצר בהתאמה אישית ברמת איכות גבוהה יותר מאשר האמצעים הקיימים היום. עבריינים צפויים להרחיב יכולותיהם על מנת להשתמש במדיה מלאכותית. למעשה, היבטים שצוינו בתרחישים שהוצגו במסמך כבר התממשו. אם המערכת הפיננסית לא תפעל כעת, היא עלולה להפסיד זמן יקר ערך במרוץ נגד גורמים זדוניים. פיתוחים טכנולוגיים, שותפויות בין המגזר הציבורי לפרטי ואמצעי מדיניות אחרים צפויים להגיע לכדי הבשלה רק בשנים הבאות.

מצד שני, לא ניתן להצדיק הקצאת משאבים חיוניים לבעיה שהיא ברובה תיאורטית. השימוש במדיה מלאכותית טרם הוביל לנזק כלכלי נרחב ואף אחד מהתרחישים שהוצגו לא התממש במלואו. מנקודת מבטם של מקבלי ההחלטות, מדיה מלאכותית היא סיכון אחד מני רבים במגזר הפיננסי ואף ייתכן כי מדובר באיום שלא יתממש במלואו. כמו כן, איומים רבים קשורים להפסדים כספיים שניתן למדוד, זאת לעומת הנזק שעלול להיגרם כתוצאה משימוש זדוני ב-Deepfakes שנותר עדיין בגדר ספקולציה.

מדובר כאן בדילמה קלאסית של ניהול סיכונים. בעולם אידיאלי, החלטות בתחום המדיניות מבוססות על מידול סיכונים מהימן והערכה על החזר ההשקעה. אבל, כאשר מדובר בטכנולוגיות



כמו מדיה מלאכותית, אין כיום מספיק מידע רלוונטי. על מנת להתמודד עם הדילמה הזו, גורמים בעלי עניין במגזר הפיננסי צריכים תחילה לנקוט בצעדי תגובה ראשוניים, תוך מעקב אחר התפתחות האיום לאורך זמן.

הניתוח שנעשה במסמך זה יכול לסייע למקבלי החלטות בכל הקשור לביצוע צעדי תגובה ראשוניים אלו. בהיעדר מידע על השימוש לרעה במדיה מלאכותית, מסמך זה מתמודד עם הבעיה בניהול סיכונים באמצעות התבססות על תרחישים קיימים ותוך הערכה כיצד מדיה מלאכותית צפויה להחריף את האיום. בהינתן משאבים מוגבלים, על המוסדות הפיננסיים להתמקד בניסוח מענה להתמודדות עם שלוש הטכניקות הנפוצות ביותר, כגון Deepfake vishing, זיוף מסרים פרטיים ושימוש ברשתות בוטים מבוססים על בינה מלאכותית. על המוסדות לשקול כיצד להתמודד עם הפצת מדיה מלאכותית בהיקף נרחב או מצומצם. עבור הפצה נרחבת, יש להתמקד במדיה מלאכותית המשמשת להפצת נרטיבים ולהעצמת משברים קיימים.

לצד ההתמקדות במענים אלו, חשוב למקמם במאמץ משותף נרחב יותר. התמודדות עם איום המדיה המלאכותית במגזר הפיננסי מחייבת שימוש בטכנולוגיות חדשות, בתהליכים ארגוניים חדשים ובחינוך, במגזר הפיננסי ובקרב הציבור.

## סיכום

מדיה מלאכותית עלולה לשמש לצורך ביצוע שורה של הונאות פיננסיות, במסגרת מספר תרחישים. עם זאת, ניתן להעריך שגורמים זדוניים יבחרו לעיתים דווקא באמצעים קיימים על פני שימוש במדיה מלאכותית, זאת מאחר והאמצעים הקיימים היום יעילים ומניבים תוצאות ומאחר ויש צורך ביכולות מתקדמות ובתחכום על מנת להשתמש במדיה מלאכותית.

על פי המחקר, השימוש ב-Deepfakes לשם הונאות כלכליות עלול לגבות מחיר כבד מחברות ומאינדיבידואלים. אף על פי כן, מעצבי המדיניות לא תמיד ירצו להקצות את המשאבים הנחוצים להתמודדות עם האיום, מפני שאיום השימוש ב-Deepfakes עודנו איום תיאורטי במרבית המקרים. קיים חשש שמא מקבלי ההחלטות יפעלו לאט מדי, דבר שעלול לעכב ולהכשיל מציאת מענה יעיל לשימוש זדוני במדיה מלאכותית. ההתמודדות עם איום ה- Deepfakes מחייבת שיתוף פעולה של מספר בעלי עניין הכוללים את המגזר הפיננסי, גורמים ממשלתיים, גורמים מסחריים מחוץ למגזר הפיננסי והציבור הרחב. מחבר המסמך קורא למגזר הפיננסי לפעול בשלושה מישורים הכוללים פעולות פנים-ארגוניות, פעולות בקרב המגזר הפיננסי ושיתוף פעולה עם בעלי עניין נוספים.

# ד. מדיניות החוץ המפוזרת והבלתי מתואמת של קנדה בתחום הסייבר: קריאה להבהרה

Josh Gold, Christopher Parsons & Irene Poetranto[25]

## הקדמה

מחברי המסמך סבורים שעל קנדה לנסח מדיניות חוץ מקיפה בתחום הסייבר. טענתם המרכזית היא שבעוד שקנדה השיקה בשנים האחרונות מספר יוזמות לקידום אבטחת סייבר ברמה הבין-לאומית, היא טרם ניסחה מדיניות מקיפה ומתואמת בנושא. מדיניות שכזו תוכל לסייע לקנדה לקדם את הערכים והאינטרסים החשובים לה בזירה הבינלאומית.

## מבוא

ביולי 2020 הצטרפה קנדה לארה"ב ולבריטניה בייחוס מתקפות סייבר שמטרתן לגנוב מידע בנושא חיסונים לנגיף הקורונה לממשלת רוסיה. שר ההגנה של קנדה, הרג'יט סג'אן (Harjit Sajjan), קרא לקהילה הבין-לאומית לחזק את ההבנה המשותפת בנושא נורמות התנהגות במרחב הסייבר ולחזק את ההרתעה כנגד גורמים זדוניים זרים. עם זאת, למרות ניסיונותיה של קנדה למלא תפקיד מפתח בשמירה על השלום והביטחון העולמיים, דבר שבא לידי ביטוי ברצונה להתקבל כחברה במועצת הביטחון של האו"ם, לקנדה אין אסטרטגיית סייבר בין-לאומית ברורה ומקיפה.

בשנת 2010, הכירה ממשלת קנדה לראשונה בצורך לפתח מדיניות חוץ בתחום הסייבר, במטרה להבטיח שפעילותה במרחב הסייבר תעלה בקנה אחד עם יעדים נרחבים יותר בתחומי החוץ והביטחון. מחברי אסטרטגיית אבטחת הסייבר הלאומית של קנדה משנת 2018 (NCSS),[26] קבעו כי האסטרטגיה תואמת למדיניות החוץ בתחום הסייבר של קנדה. שנתיים לאחר פרסום האסטרטגיה, קנדה טרם ניסחה מדיניות חוץ בתחום הסייבר, זאת בניגוד לבעלות בריתה וליריבותיה, שהציגו את יעדיהן, האינטרסים שלהן והערכים שלהן במרחב הסייבר ואת תכניותיהן לקדם אותם ולהגן עליהם.

מדיניות חוץ מקיפה בתחום הסייבר צפויה להחליף את גישתה הנוכחית של ממשלת קנדה, המבוססת על פעולות אד-הוק. ממשלת קנדה זקוקה להבעת עמדה עקבית במסגרת מדיניות חוץ במרחב הסייבר על מנת לקדם את האינטרסים שלה ולהגן עליהם ביעילות. בנוסף, עליה לנסח מדיניות חוץ בתחום הסייבר תוך שמירה על שקיפות, מדיניות שתשקף את ערכיה, כגון שמירה על זכויות אדם ועקרונות דמוקרטיים אחרים.

25 Josh Gold, Christopher Parsons and Irene Poetranto. Canada's Scattered and Uncoordinated Cyber Foreign Policy: A Call for Clarity. *Just Security*. August 4, 2020. https://www.justsecurity.org/71817/canadas-scattered-and-uncoordinated-cyber-foreign-policy-a-/call-for-clarity

26 The National Cyber Security Strategy



## מדיניות חוץ בתחום סייבר: מדוע היא חשובה?

בכל הנוגע לאבטחת סייבר, ובייחוד במישור הבין-לאומי, ממשלת קנדה לא הציגה באופן ברור את העקרונות שאותם ברצונה לקדם ולהגן ומדוע. הצגתם של עקרונות אלו חשובה, מפני שמרחב הסייבר מגלם בתוכו דיונים בעלי אופי פילוסופי-פוליטי. לא כל השחקנים הפעילים במרחב הסייבר חולקים את אותן ההבנות בנושא מהות הביטחון ומטרתו ואין הבנה משותפת בנושא הגדרתו של איום במרחב הסייבר.

בתור דמוקרטיה ליברלית, המבוססת על שמירה על זכויות אדם ועקרון שלטון החוק, קנדה תלויה בשמירתם בתוך גבולותיה ובזירה הבין-לאומית. עם זאת, מספר מגמות במרחב הסייבר, כגון השימוש בצנזורה, מרוץ חימוש והצטיידות באמצעי סייבר התקפיים, הקמת פיקודי סייבר צבאיים והסתמכות על מבצעי סייבר התקפיים, מהוות אתגר עבור קנדה והשגת יעדיה לקידום ערכים דמוקרטיים במרחב הסייבר. הצגת עקרונותיה של קנדה בהקשר של אבטחת סייבר תעזור להבהיר מהם האינטרסים של קנדה ומהי משמעותם בעולם דיגיטלי. רק לאחר הגדרת האינטרסים שלה, תוכל ממשלת קנדה להתמקד בהגנה עליהם ובמציאת הדרכים הטובות ביותר לעשות זאת, ולקדמן בזירה הבין-לאומית. על ממשלת קנדה להגדיר את ערכיה ומטרותיה באופן מקיף, על מנת שתוכל לנסח אסטרטגיה.

רבות מבעלות בריתה הקרובות של קנדה, כגון ארה"ב, בריטניה, אוסטרליה והולנד, פרסמו אסטרטגיות שמטרתן להבהיר את מטרותיהן הספציפיות במדיניות החוץ בנושא טכנולוגיות דיגיטליות והשימוש בהן, הן בהקשר הביטחוני והן בהקשר של שמירה על זכויות אדם.

קנדה ניצבת בפני אתגר. חברותה בברית ה-Five Eyes מעניקה לה ערך ביטחוני חיוני, אך טומנת בחובה גם התחייבויות, מגבלות וסתירות אפשריות לערכים קנדיים מסוימים, כגון שמירה על זכויות אדם. מעורבותה של קנדה במעקב המוני עלולה להתפרש כהפרה של זכויות של אזרחי מדינות שאינן שותפות לברית ה-Five Eyes. על מדיניות החוץ של קנדה בתחום הסייבר לספק מענה לשאלות כיצד תתמודד קנדה עם הפרות זכויות האדם הקשורות לביצוע מעקב המוני וכיצד ניתן לבצע מעקב המוני, מבלי לסכן את רווחתה הכלכלית של המדינה.

## התפתחויות אחרונות בתחומי המדיניות והחקיקה

ממשלת קנדה ניסחה מדיניות אבטחת סייבר, אולם אין מדובר באסטרטגיה מקיפה כנדרש. מסמך ה-NCSS משנת 2018 עדכן את מסמך האסטרטגיה משנת 2010, אולם נותר מעורפל. המסמך אינו מזכיר מונחים כגון 'דמוקרטיה' ו-'זכויות אדם', זאת על אף היותם ערכים מרכזיים של קנדה. תכנית הפעולה הלאומית בתחום אבטחת הסייבר,[27] שפורסמה בשנת 2019, מגדירה יוזמות ספציפיות שמטרתן ליישם את מטרות מסמך ה-NCSS ומדגישות את הצורך לקדם את האינטרסים של קנדה במרחב הסייבר בזירה הבין-לאומית.

תכנית הפעולה מכירה בכך שהממד הבין-לאומי של אבטחת הסייבר אינו מוקד הפעילות העיקרי של קנדה, על אף שאבטחת סייבר היא סוגיה בין-לאומית במהותה. כמו כן, תכנית הפעולה קוראת לממשלת קנדה לקחת על עצמה תפקיד מוביל בקידום אבטחת הסייבר ובתיאום עם בעלות בריתה במטרה לעצב את מרחב הסייבר הבין-לאומי לטובת האינטרסים הלאומיים של קנדה.

27 The 2019 National Cyber Security Action Plan

בהקשר זה, צוות מדיניות הסייבר של המשרד לענייני עולמיים של קנדה (GAC)[28] עוסק בניסוח אסטרטגיית סייבר בין-לאומית. נכון לכתיבת שורות אלו, אסטרטגיה זו טרם פורסמה על אף שתאריך היעד לפרסומה נקבע לשנת 2019. פרטים מעטים בלבד מהאסטרטגיה הצפויה פורסמו, ביניהם כיצד היא תיראה, היקף חשיפתה לציבור והאם בעלי עניין שונים, כגון ארגוני החברה האזרחית, השתתפו בהתייעצויות לקראת תהליך הניסוח.

בנוסף לניסוח מסמכי מדיניות ואסטרטגיה, בשנת 2019 אושרה חקיקה, שנועדה לאפשר לגורמים מדינתיים להתמודד טוב יותר עם איומים על ביטחונה הלאומי של קנדה. חוק הביטחון הלאומי שנחקק באותה השנה, נחשב לעדכון חשוב ומקיף של החקיקה הקנדית בתחום הביטחון הלאומי. חוק סוכנות הסיגינט של קנדה (CSE Act),[29] הכלול כסעיף בחוק הביטחון הלאומי, מאפשר לסוכנות, המשמשת כסוכנות הלאומית למודיעין ולאבטחת סייבר, להוציא לפועל מבצעי סייבר הגנתיים ומבצעי הגנת סייבר אקטיבית מחוץ לשטחה של קנדה, תוך הגדלת סמכויותיה וטווח פעילותה. לסמכויות חדשות אלו השלכות על זכויות אדם, שקיפות פוליטית וביטחון עולמי והן בולטות על רקע התנגדותו של האיחוד האירופי לביצוע מעקבים המוניים ועל רקע היעדר מנגנון לפיצוי אזרחים אירופיים שזכויותיהם נפגעו במסגרת פעילות מדינות ברית ה-Five Eyes.

## פעילות בין-לאומית, דיפלומטיה ומוקד מגדרי

קנדה משתתפת בפורומים אזוריים ובינלאומיים בתחום אבטחת הסייבר, תוך שיתוף פעולה עם מדינות בעלות ברית וקיום שיח מול מדינות פחות ידידותיות, במטרה להגיע עמן להסכמים בנושא אינטרסים משותפים ובמטרה לנסות להבין את עמדותיהן.

מאז 2015 הקצתה קנדה מעל ל-13 מיליון דולר קנדי (כ-9.8 מיליון דולר) עבור פרויקטים לפיתוח יכולות סייבר ברחבי העולם, במטרה לאמן ולהכשיר גורמים מקומיים בתחומי המדיניות, הטכנולוגיה והמשפט. על פי ה-GAC, מאמצים אלו מהווים חלק חיוני מהאסטרטגיה של קנדה להפיץ את חזונה לשימור אינטרנט פתוח, בטוח, המנוהל על ידי בעלי העניין השונים.

עם זאת, בהינתן העובדה שלקנדה אין מדיניות חוץ פומבית בתחום הסייבר, לא ברור כיצד צעדים אלו תואמים ליעדים רחבים יותר שלה, כגון קידום אבטחת סייבר המבוססת על דמוקרטיה וזכויות אדם, ושיטות להתמודדות עם נרטיב הצנזורה והשליטה הממשלתית באינטרנט, המקודם על ידי מדינות אוטוריטריות.

תחת ראש ממשלת קנדה הנוכחי, ג'סטין טרודו (Justin Trudeau), ענייני חוץ נוהלו על בסיס עקרונות פמיניסטיים, המכוונים להשגת שוויון מגדרי ולהעצמת נשים. בתחום אבטחת הסייבר, ה-GAC מימן מחקרים בנושא ההיבט המגדרי של אבטחת סייבר וארגן אירועים בנושא. במהלך המושב השני של קבוצת העבודה הפתוחה של האו"ם (OEWG),[30] שהתקיים בפברואר 2020, ה-GAC הצטרף למדינות אחרות במתן חסות לנציגי ממשל וארגוני חברה אזרחית ממדינות מתפתחות, שהגיעו לפגישה בחסות תכנית המלגות "נשים בסייבר".

28 Global Affairs Canada

29 The Communications Security Establishment Act

30 Open-Ended Working Group; הקבוצה הוקמה ביוזמתה של רוסיה בפורום נרחב לקביעת נורמות התנהגות במרחב הסייבר במקביל לקבוצת המומחים של האו"ם (GGE), המורכבת ממספר מצומצם יותר של מדינות.



החשיבות שקנדה מעניקה לענייני מגדר במדיניות החוץ שלה בתחום הסייבר ראויה להערכה, אולם פעילותה בתחום מתקיימת כסדרה של פעולות אד-הוק המהוות אלמנטים חלקיים בלבד מתוך מכלול שלם שאינו בא לידי ביטוי. על מדיניות עקבית להיות חלק מאג'נדה רחבה יותר. לפיכך, אם ברצונה של קנדה לקדם היבטים מגדריים בתחום אבטחת הסייבר, על מדיניות החוץ שלה בתחום הסייבר להדגיש את חשיבות המגדר וזכויות האדם בתחום האבטחה.

## אסטרטגיית הגנת סייבר צבאית ומעורפלת

קנדה הכירה בכך שבמטרה להגן על האינטרסים והערכים שלה במרחב הסייבר יש צורך ביותר מדיפלומטיה בלבד, עמדה המשתקפת בחוק הביטחון הלאומי מ-2019. בנוסף, מסמך מדיניות ההגנה של צבא קנדה משנת 2017 קרא לצבא לקבל על עצמו עמדה אסרטיבית יותר במרחב הסייבר, באמצעות ניהול מבצעי סייבר אקטיביים נגד יריבים פוטנציאליים. ההחלטה בנושא תפקידו וסמכויותיו של צבא קנדה לא לוותה בדיון ציבורי. הטלת האחריות על סוכנויות המודיעין ועל צבא קנדה, להוביל ולנהל את מדיניות הסייבר, עלולה לערער ולפגוע במאמציה הדיפלומטיים של קנדה, הכוללים את פעילותה במסגרת האו"ם, הממוקדת בחיזוק השלום והביטחון במרחב הסייבר.

קנדה פועלת להגברת נוכחותה ופעילותה הצבאית בתחום הסייבר. כחלק מכך, הכריזה קנדה כי תצטרף למרכז למצוינות בהגנת סייבר משותפת של ברית נאט"ו (CCDCOE).[31] ברית נאט"ו עוסקת בניסוח הדוקטרינה המבצעית שלה במרחב הסייבר וחנכה את המרכז למבצעים במרחב הסייבר שלה בבלגיה.[32] תשע מדינות החברות בברית, שאינן כוללות את קנדה, הודיעו רשמית כי יעמידו את יכולות הסייבר שלהן לרשותה של הברית על מנת להגיב למתקפות סייבר.

בספטמבר 2019 קנדה הצטרפה ל-26 המדינות החתומות על ההצהרה המשותפת לקידום התנהגות מדינתית אחראית במרחב הסייבר.[33] במסגרת ההצהרה הכריזו המדינות החתומות כי ישתפו פעולה בהטלת האחריות ויגבו מחיר ממדינות המעורבות בפעילות סייבר זדונית. החתימה על הצהרה זו מדגישה את החשיבות שקנדה מייחסת לגביית מחיר על התנהגות זדונית במרחב הסייבר.

חשיבות זו באה לידי ביטוי גם בתזכיר שנשלח לראש ממשלת קנדה באוקטובר 2019 לפיו קנדה מכירה בחשיבותן של נורמות התנהגות במרחב הסייבר, אולם יש צורך לנסח דרכי תגובה וצעדים לגביית מחיר מגורמים זדוניים. מחברי המסמך הוסיפו כי אחד מעמודי התווך באסטרטגיית אבטחת הסייבר קורא לקנדה לשתף פעולה עם בעלות בריתה במטרה לפתח מנגנון לגביית מחיר והטלת השלכות על גורמים זדוניים במרחב הסייבר. עם זאת, מהמסמך לא ברור סוג המחיר ועד כמה הוא יעמוד בחוק הבין-לאומי. על כן, מסמך אסטרטגיית אבטחת הסייבר של קנדה שולח מסר לא ברור לאזרחיה, לבעלות בריתה וליריביה.

כל 27 המדינות שחתמו על ההצהרה המשותפת מספטמבר 2019 הן בעלות בריתה של ארה"ב וחברות בברית נאט"ו, עובדה שמעלה חשד פוטנציאלי להטיה (bias), בהגדרת פעולותיהן של מדינות אחרות כזדוניות, בהשוואה להגדרת פעולותיהן של המדינות החתומות. אם קנדה

31 NATO Cooperative Cyber Defence Centre of Excellence

32 The Cyberspace Operations Centre

33 Joint Statement on Advancing Responsible State Behavior in Cyberspace

מעוניינת לקדם שיתוף פעולה בין-לאומי בתחום אבטחת הסייבר העולמית, עליה להתקדם מעבר לפתרונות המבוססים על מדינות בעלות ברית השותפות לעמדתה.

## פערים בעקביות, בבהירות ובתיאום

במסגרת מדיניות החוץ שלה, אימצה קנדה עמדות התומכות בזכויות האדם והאזרח, אך חוותה אתגרים ביישומן. דוגמה לכך היא חברת Netsweeper הקנדית, העוסקת בסינון תכנים פוגעניים ובניהול איומי רשת, הזוכה לתמיכה ממשלתית פדראלית ומחוזית, על אף מחקר מטעם מעבדת Citizen Lab באוניברסיטת טורונטו, לפיו טכנולוגיה מתוצרת החברה משמשת לצנזורה, חסימת תכנים והפרת זכויות אדם. כמו כן, חוק ה-CSE, שהרחיב את סמכויותיה של סוכנות הסיגינט של קנדה, כולל ניסוחים שיכולים להתפרש ככאלו המתירים התערבות חיצונית בהליכים משפטיים או בתוצאות בחירות. פרשנות זו, התומכת בהתערבות חיצונית בתהליכים מדינתיים פנימיים, עשויה לספק לגיטימציה למדינות יריבות, המתערבות בתהליכים פנימיים של מדינות זרות.

כמו כן, בניגוד לבעלות בריתה, קנדה לא התייחסה בפומבי לעמדתה בנושא החלת החוק הבין-לאומי על מרחב הסייבר ולשאלה כיצד יש להחילו, זאת על אף קריאתה למדינות אחרות לעשות כן. בניגוד לכך, מספר הולך וגדל של מדינות, בהן אוסטרליה והולנד, פרסמו מסמכי עמדה בנושא יישום החוק הבין-לאומי במרחב הסייבר. בהיעדר מסר ברור בנושא מדיניותה של קנדה ועמדתה בנושא החוק הבין-לאומי, בעלות בריתה ויריבותיה לא יעריכו את עמדתה, דבר שיקשה עליה לסייע בקביעת נורמות התנהגות אחראית במרחב הסייבר ולבסס הרתעה.

במונחים של הגנה וביטחון, מסמכים שפורסמו לציבור מעידים על כך שקנדה מיישרת קו עם גישתה של ארה"ב במרחב הסייבר, הכוללת גם את נושא יכולות הסייבר ההתקפיות. בשנת 2017, ניסח צבא קנדה את מסמך הדוקטרינה המשותפת העוסק במבצעי סייבר התקפיים. עם זאת, בניגוד למסמכי מדיניות הסייבר של ארה"ב ובריטניה, שחלקים מהם פורסמו לציבור, מסמך הדוקטרינה המשותפת נותר מסווג ולא פורסם לציבור.

על כן, כל החלטה של קנדה לאמץ גישה הכוללת שימוש במבצעי סייבר התקפיים, בדומה לאלו של ארה"ב, עלולה להתקבל ללא מעורבות ציבורית משמעותית. תהליך קבלת החלטות זה מעלה שאלות בנושא יכולת הציבור להשפיע על מדיניות, בנושא החשאיות סביב גישות אסטרטגיות ובנושא העמימות סביב מדיניות אבטחת הסייבר. כמו כן, חשאיות התהליך פוגעת ביכולתו של הציבור להטיל על הממשלה את האחריות להחלטותיה ומעלה שאלות על מידת הלגיטימציה הציבורית של ההחלטה שהתקבלה.

## הצורך המתמשך בשורה של עקרונות מוצהרים בתחום הסייבר

ניסוח מדיניות חוץ מגובשת בתחום הסייבר, הכוללת הצהרה על עקרונות, יוביל לקידום שיתופי פעולה בין-לאומיים, למיזוג שיקולים מגדריים במסגרת מדיניות הביטחון ולפיתוח יכולות סייבר התקפיות. עם זאת, על מדיניות הוליסטית להבטיח שהאינטרסים והעקרונות של קנדה, כגון דמוקרטיה, זכויות אדם ושמירה על שלטון החוק, מוגנים ומוקרנים כלפי חוץ.

בניגוד למצב הקיים, בו תהליכים לניסוח מדיניות חוץ בתחום הסייבר מתנהלים בחשאי, על



ממשלת קנדה לקיים התייעצויות עם בעלי העניין השונים, כגון ארגוני החברה האזרחית והמגזר הפרטי. נציבות ה-Cyberspace Solarium, שהוקמה על ידי הסנאט האמריקני במטרה לנסח אסטרטגיית סייבר לאומית, קיימה יותר מ-200 פגישות עם נציגי המגזר הפרטי ויותר מ-25 פגישות עם גורמי אקדמיה. ה-GAC לעומת זאת, לא קיים התייעצויות עם גורמים א-ממשלתיים במהלך גיבוש המסגרת האסטרטגית הבינלאומית למרחב הסייבר, שפורסמה כחלק מתכנית הפעולה לאבטחת סייבר לאומית ב-2019.[34]

קנדה יכולה להפוך למדינה מובילה בניסוח נורמות התנהגות במרחב הסייבר ובקידום עקרונות וערכים בנושא השימוש הגובר בטכנולוגיות דיגיטליות. במסגרת המצב הקיים, על בעלות בריתה ויריבותיה של קנדה לעקוב אחר מסמכי המדיניות הלא מגובשים שלה ולנסות להעריך כיצד יישום יעדיהם יענה על האינטרסים הרחבים של ממשלת קנדה בתחום מדיניות החוץ. המצב הקיים מקשה על קנדה להסביר את כוונותיה, רצונותיה והקווים האדומים שלה ומקשה עליה לשתף פעולה עם בעלות ברית, במטרה להשפיע על הזירה הבין-לאומית. קנדה זקוקה למדיניות חוץ הוליסטית בתחום הסייבר, אם ברצונה להפוך למעצמה בסדר גודל בינוני, המסוגלת להסביר כיצד ומתי תשתמש בכוח.

## סיכום

מחברי מסמך אסטרטגיית הסייבר הלאומית של קנדה מ-2018, טענו כי מסמך האסטרטגיה יותאם למדיניות החוץ של קנדה בתחום הסייבר. עם זאת, קנדה טרם ניסחה מדיניות חוץ בתחום הסייבר. ממשלת קנדה קידמה בשנים האחרונות מספר יוזמות במסגרת מדיניות החוץ שלה בתחום הסייבר. יוזמות אלו, הכוללות יוזמות לקידום שוויון מגדרי, לצד תמיכה בהחלת החוק הבין-לאומי במרחב הסייבר ובפיתוח יכולות סייבר התקפיות, אינן מהוות מדיניות מקיפה וברורה בנושא ופוגעות ביכולתה של קנדה להציג את עמדותיה, ערכיה ועקרונותיה במרחב הסייבר ולהגן עליהם. מדיניות מקיפה וברורה תסייע לקנדה להציג ולהגן על האינטרסים שלה תוך שיתוף פעולה עם בעלות בריתה ולייצר הרתעה מול יריבותיה. על ממשלת קנדה לקיים את תהליך ניסוח המדיניות בשיתוף פעולה עם בעלי עניין חוץ-ממשלתיים, כגון המגזר הפרטי וארגוני החברה האזרחית.

34 International Strategic Framework for Cyberspace