

נגזרות מודיעיניות של עולם הביג דאטה

סא"ל צ' - משרת באמ"ן

מבוא

בשנים האחרונות חלו התפתחויות מרחיקות לכת בתחום עיבוד המידע, בעיקר בשוק האזרחי. ההתפתחויות הללו מכוננות בשמות שונים, ובהם "מהפכת ה־Big Data" ו"מהפכת ה־Data Science". הן משפיעות על חיי היומיום של בני אדם בכל העולם ברמה האישית והציבורית, והן בעלות פוטנציאל השפעה דרמטי על האופן שבו מייצרים מודיעין. מטרת מאמר זה היא לסקור את ההזדמנויות שהשתנוות זו מייצרת עבור קהילת המודיעין, את הלקחים שנלמדו עד כה באמ"ן ואת האתגרים שעדיין עומדים לפתחנו.

השתנות בעולם

תחום המידע חווה בעשור האחרון שתי מגמות עולמיות הכרוכות זו בזו. מצד אחד, חל שינוי בדרך שבה העולם משתמש בטכנולוגיה שמייצרת מידע, שהוביל לגידול אקספוננציאלי בכמות המידע. מצד אחר, וכתגובה לכך, חלו כמה התפתחויות טכנולוגיות חשובות בתחום ניתוח המידע, שאפשרו ניתוח כמויות גדולות מאוד של מידע במהירות גדולה. כתוצאה מההתפתחויות האלה נוצרה תעשייה חדשה שמפתחת שיטות וכלים חדשים, תפקידים חדשים בארגונים, והכשרות ייחודיות לתחום עיבוד המידע.

בצד ייצור המידע, הגורמים המרכזיים לשינוי הם השימוש בטלפונים חכמים ומחשוב לביש ברמה האישית, וכן מגמת ה־Internet Of Things (IOT)¹, שבמסגרתה חל גידול אדיר במספר רכיבי המחשוב הרשתיים הנמצאים בבתים ("בתים חכמים") ובערים ("ערים חכמות"). גם ברמה הלאומית, מתקיימים פרויקטים גדולי ממדים, שבמסגרתם נאספות כמויות מידע גדולות מאוד, כמו פרויקטי מחקר בפזיקה, או במציאת תרופות מותאמות אישית באמצעות מיפוי גנטי.² מעבר לגידול במספר הרכיבים המייצרים מידע, חל גם שינוי באופן שבו בני אדם משתמשים ברכיבים אלה, בדגש על תופעת הרשתות החברתיות, שגורמת לבני אדם למסור יותר ויותר מידע על עצמם (מרצון או שלא מרצון).

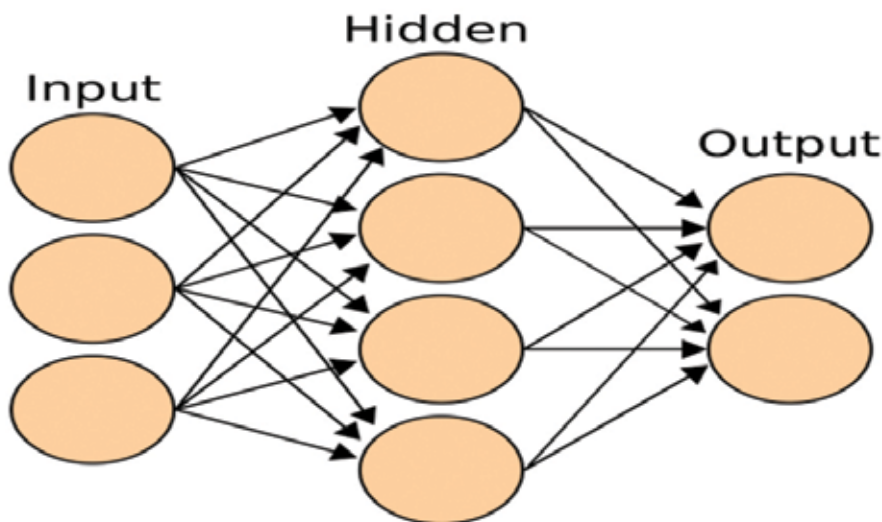
תחום המידע חווה בעשור האחרון שתי מגמות עולמיות - שינוי בדרך שבה העולם משתמש בטכנולוגיה שמייצרת מידע והתפתחויות טכנולוגיות בתחום ניתוח המידע, שמאפשרות ניתוח כמויות גדולות מאוד של מידע במהירות גדולה

1 קשה לאתר את מקור המונח; הוא הופיע במצגת של Kevin Ashton, או בחברת Procter & Gamble, בשנת 1999.
2 White House Special with DJ Patil, US Chief Data Scientist, Partially Derivative Podcast, Dec 12, 2016

בתחום ניתוח המידע, ניתן להצביע על כמה מגמות:

- **טכנולוגיות עיבוד מבוזר** - קיומם של מאגרי מידע אדירים בידי ענקיות הטכנולוגיה הגדולות (Google, Facebook, Twitter, Amazon), הוביל לקידום טכנולוגיות חדשות להתמודדות עם מידע באופן מבוזר. נקודת המפתח של טכנולוגיות אלו היא הבאת העיבוד המבוזר אל המידע, במקום הבאת המידע אל יחידת העיבוד. פרדיגמה חישובית זו היא הבסיס למשפחת טכנולוגיות ה-Hadoop, שהתפתחה במקור מטכנולוגיות שפותחו על ידי חברת Google, ונפתחו לשימוש כלל התעשייה.³ פרויקט ה-Hadoop הוא פרויקט קוד פתוח המגדיר מסגרת תוכנה ליישום מבוזר והמעבד כמויות גדולות של נתונים.
- **אלגוריתמיקה של למידת מכונה ו-deep learning** - למידת מכונה איננה מושג חדש, אבל היא חייבה בעבר כמויות גדולות של מידע מתויג ידנית לצורך אימון המכונה, וכוח מחשוב יקר. גם האלגוריתמיקה הייתה יחסית מוגבלת. הפרדיגמה החישובית המקבילית שנקראה "רשתות נוירונים מלאכותיות" (Artificial Neural Networks) שמדמות את הפעילות החישובית של רשתות נוירונים במוח האנושי, ומאפשרות למידת מכונה מקבילית מהירה, נחשבה לכזו שהפוטנציאל המוגבל שלה מיצה את עצמו.

תרשים 1: מבנה בסיסי של רשת נוירונים (מקור: Wikimedia commons, GFDL)



הירידה בעלויות החישוב הקטינה את עלות האימון, העלייה בכמות המידע אפשרה את האימון, ובתחום האלגוריתמיקה חלה פריצת דרך במחקר של רשתות נוירונים מלאכותיות עם הרעיון של שימוש ברשתות בעלות מספר רב של שכבות (Deep Neural Networks). מעלתה

3 .The Google File System, Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung, 2003
MapReduce; Simplified Data Processing on Large Clusters, Jeffrey Dean and Sanjay Ghemawat, 2004.

העיקרית של הטכנולוגיה החדשה היא שהיא מאפשרת למידת מכונה מדוגמאות, מבלי שבני אדם יגדירו מראש את התבניות המבחינות שיש לחפש במידע, באופן דומה ללמידה אנושית.⁴ השקעה מסחרית בטכנולוגיות לטיפול במידע לא מובנה - מידע לא מובנה (שמע, תמונה, וידיאו) הוא מידע קשה לאינדוקס, שליפה וחיפוש. אלגוריתמיקה לעיבוד שמע ותמונה היא תחום שהושקעו בו תשומות גדולות באקדמיה וביישומי טכנולוגיה נישתיים יחסית (צילומי לוויין, מעבדות קול). ההשתנות בתחום זה היא במעבר ממחקר שמבוצע באקדמיה ובארגונים ביטחוניים, למחקר שמתבצע על ידי ענקיות טכנולוגיות המכשיר (Google, Facebook, Microsoft, Apple, Amazon, IBM). ההשקעה שלהן מכוונת לבעיות החדשות שנובעות מקיום מאגרי ענק של תמונות וסרטוני וידיאו, ומהתעשייה המתפתחת של עזרים אלקטרוניים.⁵ התמריץ המסחרי האדיר, התחרות בין הארגונים, והשימוש באלגוריתמיקה החדשה, הובילו לשיפורים משמעותיים באלגוריתמיקה של מציאת פרטים בתמונות ובסרטי וידיאו, ובאלגוריתמיקה של עיבוד שמע. אלגוריתמיקה של זיהוי דובר בשמע, ומציאת מילות מפתח בשמע (Keyword spotting) היא כבר ברמה מסחרית, ולראשונה מדברים באופן גלוי על האפשרות המעשית של Speech to Text. זיהוי פנים בתמונות גם הוא הפך לטכנולוגיה מסחרית נגישה. בכל הטכנולוגיות הללו יש עדיין מרחק בין ה־hype הציבורי לבין מצבן המעשי, בפרט עבור בעיות המעניינות ארגוני מודיעין, אבל המרחק הזה הולך ומצטמצם.

- **טכנולוגיות האצה חומרית** - טכנולוגיה של מעבדים גרפיים, ששימשה בעבר לשיפור איכות התמונה במשחקי מחשב, נמצאה מתאימה מאוד להרצה של אלגוריתמים מקביליים של ביג דאטה. תחום הביג דאטה הפך ליעד מסחרי מרכזי של חברות שמייצרות חומרה כזו.

הזדמנויות לארגוני מודיעין

קיימות שתי משפחות של הזדמנויות שמאפשרת מהפכת הביג דאטה בארגוני מודיעין. האחת היא משפחה של הזדמנויות איסופיות - קיים היום מידע רב יותר, בתחומים רלוונטיים לשאלות מודיעיניות, שנמצא במאגרים ובמערכות מחשוב מרושתות בעולם באופן זמין יותר או פחות. הנכונות של בני אדם למסור על עצמם מידע רב מאוד במסגרת תופעת הרשתות החברתיות והאפליקציות בטלפונים החכמים, מאפשרת גם היא איסוף מידע שלא היה אפשרי בעבר, לרבות על אנשים שמנסים להסתתר. משפחה שנייה של הזדמנויות נובעת מזמינות הטכנולוגיות המסחריות לשימושים מודיעיניים. אם בעבר היו גופי המודיעין בחזית הטכנולוגיה, ונאלצו להתמודד בכוחות עצמם (יחד עם האקדמיה) עם אתגרים מחקרניים קשים, כמו עיבוד תמונה, עיבוד שמע, ומחקר של מאגרים ענקיים, כעת יכולים ארגוני המודיעין לרכב על גל, שבחזיתו נמצאת התעשייה המסחרית, בדגש על ענקיות המחשוב.

אתגר הביג דאטה בארגוני מודיעין

בטרם אתאר את האתגרים הייחודיים שמייצרת מהפכת הביג דאטה לארגוני המודיעין, אנסה לאפיין כמה מושגים הרלוונטיים לעולם המודיעיני המתארים את תופעת הביג דאטה. באופן כללי,

4 Hinton, G. Osindero, S. The, Y. "A fast learning for deep belief nets", 2006
5 חברת Apple מייצרת את Siri, Microsoft את Alexa, Cortana ו־Google Assistant את Google

נהוג לטעון שפרדיגמת הביג דאטה שונה מפרדיגמות מידע קודמות בשלושה מאפיינים, שמכונים שלושת ה-V: נפח המידע (Volume), מגוון המידע (Variety) והמהירות שבה יש לטפל במידע (Velocity).⁶

אפשר להשתמש במושגים הללו כדי להצביע על ההבדל בין אתגרי התעשייה המסחרית לאתגרי המודיעין. בכל הקשור לנפחי המידע, אין גוף מודיעיני שיכול להתחרות בכמויות המידע שברשות Google או Facebook. גם בתחום המהירות שבה יש לטפל במידע, האתגר המסחרי גדול מהאתגר המודיעיני - מרבית השאלות המודיעיניות על מידע הן כאלה שסביר לקבל עליהן תשובה בקבועי זמן של בין דקות לשעות, ובתחומים מוגבלים בקבועי זמן של שניות. אפילו במקרה האחרון

מדובר במה שמכונה בתעשייה Near Real Time ולא Real Time, שמתבטא בקבועי זמן של פחות משנייה. בעולם המסחרי, קביעה איזו פרסומת יקבל משתמש כאשר הוא גולש באתר, או החזרה של תשובה לשאלה במנוע החיפוש של Google, מחייבות פעולה בקבועי זמן של פחות משנייה. יוצא מהכלל הוא תחום המגוון. בניגוד לארגונים מסחריים, שהמידע שהם אוספים הוא לעיתים קרובות "שלהם" ולכן הם יכולים לפעול לתקנן אותו, והם גם לרוב ממוקדים במידע מסוג מסוים או במספר מצומצם של סוגי מידע (טרנסאקציות של כרטיסי אשראי, ציוצים ב-twitter

וכו'), ארגוני מודיעין מחזיקים במידע שנוצר במגוון רחב של מקורות, שעליהם אין לארגוני המודיעין השפעה, משום שהוא איננו בשליטת הארגון. בתחום המגוון, ארגוני מודיעין נמצאים בקצה הסקלה, והוא עבורם אתגר ענק.

האתגרים שעומדים בפני ארגוני המודיעין בכלל, ואמ"ן בפרט, הם אלו:

- **התמודדות עם נפחי מידע** - פרסומים שונים מדברים על גידול אקספוננציאלי של פי שניים בכמות המידע האגור בעולם מדי שנה. בשנים האחרונות, המדידה באמ"ן אכן הראתה גידול כזה, ולעיתים אף בקצב גבוה יותר. חלק מהגידול מוסבר בגידול בהצלחות האיסופיות, ובחלקו בגידול בכמות המידע אצל האויב ועל אודותיו. הגידול הזה חייב הקמת "מפעל מידע מודיעיני" המבוסס על טכנולוגיות ביג דאטה מסחריות. הדבר נתקל במגבלה הואיל ותקציב ארגוני המודיעין כמעט לא השתנה למול מחיר נפחי האחסון. אף שמחירי האחסון בעולם יורדים בקצב מהיר, הם יורדים בקצב נמוך יותר מקצב העלייה בכמות המידע. ומה שחשוב יותר, מגבלות הרכש הצבאי ומגבלות הסיווג מחייבות שימוש בחומרה קנויה ייעודית, במחירים שלעיתים אינם מגלמים את ירידת המחירים העולמית. כתוצאה מכך נמצא מפעל המידע המודיעיני בסד שחייב אותו להפסיק הכנסה של חומרים מודיעיניים מסוימים לאגירה,

6 השימוש הראשון בהבחנה הזו נעשה ככל הנראה בדוח של חברת גרטנר. מאז יש שמוספים קריטריונים נוספים, ואפשר למצוא התייחסות ל-Veracity (אמינות המידע), Value (תועלת עסקית), Variability (השתנות המשמעות של המידע) ו-Visualization (אופן הצגת המידע). הדוח המקורי של חברת גרטנר (או META):

Laney Douglas (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety, Gartner.

חרף המשאבים האדירים והסיכון שהושקעו בהשגתם, או להקטין את משך האגירה שלהם, מה שמקטין את היכולת לשאול שאלות מודיעיניות רבות שמתבססות על למידה לאורך זמן. בפועל, הבחירה כמעט תמיד נעשתה בהקטנת משך האגירה.

- **התמודדות עם מגוון המידע** - כאמור, בעיה זו אינה ייחודית לגופי מודיעין, אבל מוקצנת אצלם לעומת העולם האזרחי – מגוון אדיר של סוגי חומרים, לא מסודרים, לא מתוקננים, המכילים כפילויות ופערים, ובמידת אמינות משתנה. הבעיה הזו מחייבת בניית כלים ושיטות, והשקעה משמעותית של תשומות כוח אדם בתקנון ובסידור מידע.
- **התמודדות עם חומרים ייחודיים לשאלות המודיעיניות** - הגם שיש השקעה גדולה של ענקיות התעשייה האזרחית בתחום ניתוח שפה טבעית, ההשקעה הזו מכוונת לאזורים שבהם פוטנציאל הרווח הוא הגדול ביותר. מרבית ההשקעה נעשית בשפות ובלהגים שמדוברים על ידי אוכלוסיות גדולות ומתקדמות מבחינה טכנולוגית. אפילו השפה הערבית, שמדוברת על ידי אנשים רבים, אינה בראש סדר העדיפויות של התעשייה, ובמקרה של שפה מדוברת, שבה הערבית מתפצלת למספר רב של להגים, הבעיה הזו חמורה אף יותר. מעבר לסוגיית השפה, גם בתחומים אחרים המודיעין ממוקד בשאלות אחרות מאשר התעשייה. המודיעין מחפש מונחים שונים, באוצר מילים ובתחומי עניין שונים מאשר העולם האזרחי.
- **חוכמת ההמונים למול "ידיעת הזהב"** - ענקיות התעשייה מכוונות את מאמצייהן לנושאים שמעניינים את הקהל הרחב. כך, למשל, Google משתמשת במידע שמחפשים הגולשים כדי לדרג תוצאות למשתמשים אחרים. שימוש ב"חוכמת המונים" כזו במודיעין הוא מוגבל, שכן לעיתים קרובות המודיעין מחפש דווקא את הידיעות הייחודיות, ה"מסתרות". ידיעות שמישהו כבר צפה בהן הן לרוב פחות מעניינות ולא יותר מעניינות. קהילת המודיעין היא קטנה מדי ומפוצלת מדי לנושאי עניין מכדי שתוכל להוות "המון" במובן של האלגוריתמיקה המסחרית של "חוכמת המונים". ארגוני המודיעין צריכים למצוא פתרונות גם לשאלות שהתשובה עליהן היא חריגה סטטיסטית, או "ידיעת זהב" ייחודית. הבנה זו היא סוג של התפכחות מההתלהבות מרעיון הרשת החברתית וחוכמת ההמונים, בהקשר המודיעיני.⁷ כלומר, בעוד ההיגיון של הפעילות המסחרית הוא למצוא את המכנים המשותפים הנרחבים ולהפיק מהם את התנהגות ההמון, המודיעין, מעבר לתועלת שהוא יכול להפיק מכך, מחפש גם את החריג והייחודי.
- **למידה של תופעות פיזיקליות ותופעות אנושיות ותחרות למידה** - לצד ההתמודדות עם תופעות פיזיקליות (זיהוי פנים, לוחיות רישוי, חתימות דובר, צורת גופנים של טקסט), קיים קושי רב לתת למכונה ללמוד תופעות אנושיות, שיש בהן מרכיב של השתנות בציר הזמן, ואפילו "תחרות למידה", כמו דפוסי התנהגות של תאי טרור ומפגעים בודדים, או תו"ל של ארגון אויב. האויב שמתגלה, משנה את דפוסי פעילותו, והמכונה, שחייבת ללמוד מדוגמאות, לא בהכרח יכולה לעשות זאת. הקושי הזה מתחדד במעבר משגרה לחירום, שכן אויב שהתנהגותו

Calvin Andrus. "The Wiki and the Blog: Toward a Complex Adaptive Intelligence Community. *Studies in Intelligence*, Vol 49, No 3, September 2005

שונה בשגרה ובחירום, יכול לגרום ללמידה ממוכנת בשגרה להיות לא אפקטיבית בחירום, או מה שלא פחות גרוע, הלמידה במלחמה עלולה להיות איטית מדי מכדי להשפיע על תוצאות המלחמה.

- **אתגר מבני** - קהילת המודיעין, ואמ"ן בכלל זה, התעצבה מבחינה ארגונית לפני העידן הדיגיטלי, בעולם שבו סייגנט, ויזינט, יומינט ומידע גלוי היו דיסציפלינות זרות זו לזו. בעידן הנוכחי, שבו מרבית המידע הנאסף הוא דיגיטלי, ההבחנה הזו היטשטשה, והמבנה של קהילת המודיעין מייצר לה כפילויות וחוסר יעילות. ברגע שמתקיימים כמה מאגרי ביג דאטה שבנויים בטכנולוגיות שונות שאינן מאפשרות היתוך והצלבה דיגיטליים זה עם זה, המודיעין נפגע. נדרשת התארגנות חדשה שמרכזת את המידע שנאסף בארכיטקטורה אחת ואולי אפילו ב"מפעל מידע" אחד. קשה מאוד לייצר שינוי כזה בקהילת המודיעין, ללא שינוי מבני. שינוי כזה צפוי להיתקל בהתנגדויות קשות, שכן הוא מחייב שבירת הזהויות הדיסציפלינריות המסורתיות. בצד ארגוני הסינון, המחקר וההערכה, נדרשת גם כן התארגנות חדשה, שכן לא ניתן לענות על שאלות מודיעיניות בעידן הביג דאטה באותו אופן טורי שבו נוצר המודיעין בעבר.⁸

האתגר האנושי - בעידן הביג דאטה נדרשים מקצועות חדשים כגון: "מהנדסי מידע", "מדעני מידע" ו"אזרחי מידע"

האתגר האנושי - אתגר זה אינו ייחודי לקהילת המודיעין, אך מצדיק התייחסות. ההשתנות הזו נדרשת גם בארגוני המודיעין הטכנולוגיים, שזקוקים למומחים במגוון מקצועות חדשים. ממהנדסי מידע (Data Engineers), שתפקידם סידור ותקנון החומר, ניקוי, והכנסתו למאגרים בצורה שמאפשרת שליפה, ועד "מדעני מידע" (Data Scientists), שמפעילים אלגוריתמיות שונות על המידע לצורך מענה לשאלות מודיעיניות קשות.

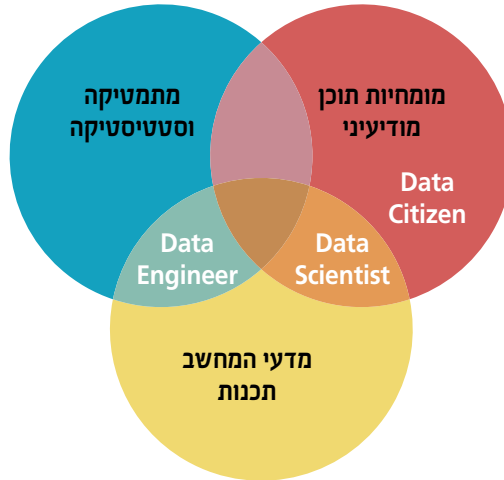
הכישורים הנדרשים למקצועות החדשים הללו נמצאים בשילוב של ידע מודיעיני מקצועי (Domain Experts), ידע מתמטי וסטטיסטי, וידע בתכנות. גם לארגוני המחקר והניתוח נדרשים כישורים חדשים. מרבית החוקרים העתידיים צריכים להיות "אזרחי מידע" (data citizens), כלומר בעלי "אוריינות או תודעת מידע". בתחום זה ארגוני המודיעין יכולים ליהנות מכך שהמקצועות הללו מתעצבים כעת באזרחות בכלל ובאקדמיה בפרט. (ראו איור 2).

- **ההשתנות הטכנולוגית ואתגר הענן** - מרבית הטכנולוגיות החדשות שנוצרות בתחום הביג דאטה מגיעות בראש ובראשונה לענני המחשוב הציבוריים. שוק ענני המחשוב הציבוריים הוא שוק ריכוזי יחסית, שבו פועלות מספר קטן של חברות ענק לא ישראליות.⁹ לקהילות המודיעין, ובפרט לקהילת המודיעין הישראלית, העוסקות במידע חשאי, יש מגבלות קשות

8 ראו: צוותי חקירה שילוביים כמענה לעידן הביג דאטה במבחן המעשה, א"א, מודיעין הלכה ומעשה, גיליון 1, דצמבר 2016, וכן קהילת ידע מודיעינית כמנגנון פעולה המספק גמישות אסטרטגית ומערכתית לאמ"ן, סא"ל ב', באותו גיליון.

9 בנקודת הזמן של כתיבת מאמר זה (2017), השוק נשלט על ידי AWS של חברת Amazon, ואחריה Azure של חברת מיקרוסופט, ו-GCP של חברת Google, בסדר הזה, לצד מספר רב של עננים עם נתח שוק קטן יותר. השוק הולך ומתמרכז בשנתיים האחרונות, וחברות "קטנות" (יחסית) נפלטות ממנו.

איור 2: מומחיות אנושיות נדרשות בעידן הביג דאטה



בשימוש בעננים ציבוריים. עם זאת, לא ניתן היום לשמר חדשנות בתחום הביג דאטה ללא קשר לקהילת הקוד הפתוח ולענני המחשוב הציבוריים. האתגר הזה יחייב מתן מענה טכנולוגי הגנתי ייחודי, וייתכן שיחייב גם מאמץ מדינתי. כיום, ענני המחשוב הגדולים קיימים במספר רב של מקומות בעולם. העננים הללו, שנבנו מתוך מיקוד בשוק האזרחי, מרחיבים את שירותיהם גם לעולם הממשלתי, אולם ההשתנות איטית, ומחייבת תשומת לב ברמה המדינית, בהקשרי רציפות תפקודית בחירום, הגנה בסייבר, וסוגיות מדיניות אחרות.

- **אתגר אתי** - הגם שנושא הביג דאטה מעלה שאלות אתיות רבות גם בסביבה האזרחית, בארגוני מודיעין וביטחון יש לו גוון אחר. פעולה אפקטיבית של ארגוני המודיעין מאפשרת ומחייבת שימוש בביג דאטה לחדירה לפרטיות בהיקפים רחבים. מעבר לכך, מכיוון שארגוני מודיעין משפיעים על פעולות מלחמתיות, מעבר מפעולה מלחמתית שבה ההחלטה אנושית לחלוטין, לפעולה שמבוססת על מכונה, בין אם מכונה שמחליטה, ממליצה או נותנת מידע רלוונטי להחלטה, מחייבת רמה גבוהה של מודעות להיבטים האתיים של פעולת המכונה, פעולה מודעת להבטחת איכות המידע, ומנגנוני הבטחת איכות משופרים.

סיכום: השינוי הנדרש מנקודת המבט המודיעינית

אמ"ן נמצא בתהליך כניסה לעידן הביג דאטה זה שנים מספר. התהליך הביא לכמה הצלחות מודיעיניות מרשימות לצד התמודדות עם קשיים לא מעטים. ניתן לסמן כמה שינויים נדרשים לאמ"ן בשנים הקרובות, כדי למצות את הפוטנציאל של מהפכת הביג דאטה. באופן כללי, השינויים הללו נגזרים משלוש תובנות מרכזיות:

1. מיצוי הפוטנציאל מחייב מעבר מהתמודדות לוקלית, "תוך כדי תנועה", במסגרת משאבית מצומצמת, להיערכות מדינית קהילתית וצבאית, למשל במסגרת התוכנית הרב-שנתית הצה"לית.
2. השינוי הוא גם פרדיגמטי בתוך ארגוני המודיעין, ובמרכזו, ההכרה בכך שתחום הביג דאטה

הוא תחום ליבה, לצד תחומי ליבה מסורתיים. למשמעות של הבנה זו השלכות נרחבות - ארגוניות, ניהוליות, ומשאביות.

3. בתחום הביג דאטה, קהילת המודיעין איננה יכולה ואינה צריכה להתחרות בארגוני הענק שמובילים את התחום בעולם אלא להתמקד בתחומים הנישתיים הנדרשים להצלחה שאינם מקודמים על ידי התעשייה, תוך שימוש בכלים, בטכנולוגיות ובשיטות שמתפתחים בתעשייה.

מהתמודדות לוקלית לתוכנית גלובלית

בשכבה הנמוכה ביותר של ההתמודדות עם עולם הביג דאטה נדרשים משאבי מחשוב גדולים מאוד גם לאחסון המידע, גם לעיבודו, וגם לאימון מכוונת לומדות. עד כה נקבע התקציב בארגוני מודיעין לנושא המחשוב, במקרה הטוב ביותר, בגידול ליניארי, ולעיתים קרובות גם פחות מכך. ההתייחסות אליו הייתה דומה לאזורים אחרים בארגון שמצריכים אחסון, ושאינם גדלים באופן אקספוננציאלי משנה לשנה. למול הגידול האקספוננציאלי בכמות המידע, הלחץ על מפעל המידע

מחייב יציאה מהמסגרת התפיסתית הרואה בציד המחשוב של אמ"ן הוצאה קבועה, כמו הוצאות בינוי והסעדה. ציוד מחשוב הוא יקר מאוד במונחים אמ"ניים, אבל פחות יקר במונחים צה"ליים, קהילתיים ולאומיים, ולכן נדרש להבין שזו יכולת ליבה צה"לית וקהילתית, ולממן אותה בהתאם.

אם מרחיבים את ההסתכלות אל מעבר להיבט הכמותי, להיבט של השימוש בטכנולוגיות ביג דאטה אזוריות, שברובן מתפתחות בעולם האזרחי, בדגש על העננים הציבוריים, ייתכן שגם התארגנות צה"לית או קהילתית היא קטנה מדי, ונדרשים התארגנות ומאמצים מדינתיים. התאמה של קהילת המודיעין וצה"ל לניצול העננים הציבוריים

**תחום המידע צריך
להפוך מתחום צדדי
לתחום ליבה במפעל
המודיעיני, הן בתחום
האיסוף והן בתחום
המחקר.**

**נדרש להשקיע משאבים
לצורך "אימון המכונה"
גם אם אין למשאבים
אלה, לכאורה, ערך
מודיעיני ישיר**

מחייבת פעולה מתואמת והיערכות הגנתית, משפטית, תקציבית וטכנולוגית. היא מחייבת שינוי בתהליכי רכש, שכן התשלום בעננים ציבוריים הוא תשלום לפי שימוש, ומחייבת גיבוש מדיניות משפטית וחוקית, שתקבע איזה מידע יכול לשבת באיזה ענן, ובאיזה אופן. שימוש בענני מחשוב ציבוריים מחייב שינוי במיקומם הפיזי, או במעטפת ההגנה שלהם. מדובר בעלויות גדולות מאוד, שארגוני הענן הגדולים יהיו מוכנים לקחת על עצמם רק אם יבינו שמדובר בעסקה כדאית, ומשום כך נדרשת פעולה ברמת מדינה, ולא ברמת צה"ל בלבד. כך למשל, קהילת המודיעין האמריקאית יצרה שיתוף פעולה עם חברת Amazon להקמת ענן מחשוב מסווג בעלות כוללת של 600 מיליון דולר.¹⁰ אפשר גם לדמיין, שבתחומים מסוימים, שמעניינים קהילות מודיעין של מדינות שונות, יהיו גם שיתופי פעולה שיאפשרו חיסכון משאבי ומיקוד מאמץ גם בתחום של מפעל המידע.

<http://www.businessinsider.com/amazon-web-services-launches-secret-region-2017-11> 10

מתחום צדדי לתחום ליבה אמ"ני

תחום המידע, או ה"IT" כפי שהוא עדיין מכונה לעיתים, נתפס כמשני לעומת תחומי הליבה האמ"ניים, האיסופיים, המבצעיים והמחקריים. הוא אינו נתפס כמרכזי בסדר היום המודיעיני. כך, למשל, ההזדמנויות האיסופיות והצי"ח מוכוון "ידעת הזהב" היו הגורם המכווין לפעולת יחידות האיסוף השונות, וההשפעה של מפעל המידע עליהן הייתה שולית יחסית. אלא שתפיסה חדשה של "מפעל מידע" וחשיבותו כפי שהוצגה במאמר זה, מחייבת השפעה משמעותית על הפרויקט המודיעיני, הן בתחום האיסוף והן בתחום המחקר. לצורכי אימון מכונה נדרש לפעמים להביא מידע שלפני כן נחשב לא חשוב בעיני האיסוף. אין הכוונה שלמפעל המידע יהיה מעמד מוביל למול הדיסציפלינות האיסופיות. לא צפוי ולא רצוי שינוי כזה בארגון, אך יש צורך בשינוי יחסי הכוחות, ובעלייה בהשפעת מפעל המידע על האיסוף.

האופן הצדדי שבו נתפס נושא המידע בארגוני המודיעין גרם לכך שמבנה ארגוני ההערכה והמחקר התעלם לגמרי מנושא המידע ומחשיבותו. ההצלחות בשנים האחרונות בניצול טכנולוגיות ביג דאטה נבעו בעיקר מ"החלפת הדיסקט" של מפקדים שהבינו את הפוטנציאל המהפכני של תחום הביג דאטה, הקדישו לנושא משאבי כוח אדם משמעותיים, ושינו את מבנה יחידותיהם כדי להתאימן לעידן החדש. עדיין, השינוי הזה היה ספורדי ולא מערכת.

היבט שלישי של התוצאות של מפעל המידע בעבר הוא שמפעל המידע פעל כתוצאה מהגעה של מקור כטריגר, או כמענה לבעיה שהציבו בפניו משתמשים מודיעיניים. לא התבצעה מחשבה עמוקה ולא התארגנות שתכליתה היא לשרת את מפעל המידע ככזה. המכונה עבדה "בשירות האדם" ולא להפך. כיום ישנם ניצני שינוי באמ"ן, ויש כיום יותר ויותר אנשים שפועלים "בשירות המכונה" - בתיוג מידע, באימון המכונה, ובהקמת תשתיות מידע וידע מקושרות בסיסיות בזירות השונות, בצמידות נמוכה יחסית למקור מסוים או לשאלה מודיעינית ספציפית. ההבנה שכשם שמפעל המידע משרת את החוקרים ואת יחידות האיסוף, כך אלו צריכים לשרת את מפעל המידע, היא הבנה הולכת ומתפתחת, אך עדיין רחוקה מלהיות מובנת מאליה באמ"ן.

מחזית הטכנולוגיה לטיפוס על כתפי ענקים

תחום הביג דאטה נולד ומקודם על ידי חברות ענק כמו פייסבוק, גוגל ואמזון. חלק מהטכנולוגיות בתחום מקודמות כיום על ידי פרויקטי קוד פתוח גדולים. ההבנה שאין סיכוי לארגוני מודיעין להתחרות במאמצי התעשייה האזרחית, מחייבת את המודיעין להיפתח יותר לתעשייה, ללמוד להשתמש בטכנולוגיה האזרחית, ולהיזהר מלנסות להתקדם על בסיס טכנולוגיה שונה מהטכנולוגיה האזרחית. הסגירות המסורתית של המודיעין, מנימוקים של חשאיות אך גם מחוסר רצון להתבסס על פיתוח חיצוני, עלולה להיות לו לרועץ בתחום הביג דאטה. אמ"ן חייב למצוא פתרונות וכלים שיאפשרו לגופי הטכנולוגיה שלו לעבוד עם התעשייה, וכך ייחנה מפירותיה. המשמעות המעשית היא מציאת פתרונות לחיבור אנשי המודיעין לסביבה אזרחית ומציאת מעטפת ההגנה המתאימה.