



המכון לחקר המתודולוגיה של המודיעין



המרכז למורשת המודיעין

## ”על המדף” – תוצר ביקורת ספרים

### מהפכת הביג דאטה

**Big Data: A Revolution That Will Transform How We Live, Work, and Think/ Viktor Mayer-Schönberger and Kenneth Cukier/ London: John Murray-Hachette, 2013**

#### אהוד ערני<sup>1</sup>



מתוך האתר של Amazon

זהו ספר קלאסי על תופעת הביג דאטה<sup>2</sup> ומשמעויותיה. הספר פורסם לפני שבע שנים (נצח במונחים טכנולוגיים) אך הסקירה המקיפה שהוא מעניק נותנת בסיס טוב להבנת התחום למי שאינם מומחים. הוא כתוב בשפה פשוטה וברורה ומספק דוגמאות שלמרות שחלקן מוכרות לחובבי הז'אנר, הן מבהירות בצורה נגישה את הטענות המרכזיות. הספר היה ברשימת רבי המכר של הניו יורק טיימס, תורגם ל 16 שפות ועדיין עולה במאי 2020 כבחירה הראשונה לאחר הקלדה (באנגלית) של המילים "ביג דאטה" בחיפוש הספרים של אתר אמזון.

המחברים מגדירים ביג דאטה כ"אותם דברים שניתן לעשות בהיקף גדול ושאי אפשר לעשותם בהיקף קטן, בכדי לכרות תובנות חדשות או בכדי ליצור אופנים חדשים של ערך [כלכלי] בדרכים שמשנות, בין היתר, שווקים, ארגונים, ואת היחסים בין אזרחים לבין ממשלות" (עמ' 6). הגדרה פשוטה יותר אולי תסתפק בקביעה שביג דאטה הוא כמות גדולה מאד של מידע מגוון שנוצר ומופץ במהירות רבה, לצד האפשרויות הטכנולוגיות לשימורו, עיבודו וניתוחו.

כך או כך, המציאות שהתניעה את עידן הביג-דאטה מעוגנת בהתחזקות והתרחבות האפשרויות לאיסוף מידע על ידי מגוון הולך וגדל של סנסורים, והתעצמות היכולות לשימור המידע (בייחוד נוכח התרחבות הדיגיטציה), ועיבודו בעיקר תוך שימוש באלגוריתמים של בינה מלאכותית. קשיי ההגדרה אינם מפתיעים. הכותבים מודים

<sup>1</sup> ד"ר אהוד (אודי) ערני: עמית מחקר בכיר במכון לחקר המתודולוגיה של המודיעין. מרצה בכיר, המחלקה ליחסים בינ"ל, אוניברסיטת חיפה, חוקר אורח, המחלקה למדע המדינה, אוניברסיטת סטנפורד.

<sup>2</sup> האקדמיה ללשון העברית קבעה ב-2016 כי המונח העברי הוא נתוני עתק. נדמה כי לפחות עתה המונח אינו מצוי בשימוש נרחב ובמאמר זה יעשה שימוש במונח ביג דאטה.

(עמ' 6) כי אין הגדרה מדויקת למונח. בספרות המקצועית והפופולרית יש הגדרות שונות לביג דאטה, כולל טענות כי מדובר במטבע לשון אופנתי ("באזורד") שנופח בייחוד על ידי חברות תוכנה כדי לקדם את מוצריהן.<sup>3</sup> במקומות אחרים נטען שהגבול בין ביג דאטה לבניה מלאכותית אינו ברור. מקורות רבים מתארים ביג דאטה ככזה שעונה לחמישה מאפיינים, המכונים<sup>4</sup> חמשת ה V :

**היקף (VOLUME):** עסקינן במידע בהיקפים גדולים מאד, מעבר ליכולות העיבוד של מערכות מחשוב מסורתיות. קושי ברור ברכיב זה של ההגדרה: "היקף גדול מאד" ו"יכולות עיבוד מסורתיות", הם מונחים יחסיים. הקושי הזה הציף גם קולות של חוקרים שטוענים כי אין לפנינו מהפכה של ממש.<sup>5</sup>

הפורום הכלכלי העולמי העריך ב 2019 כי עד שנת 2020, צפוי היקף המידע הדיגיטלי העולמי להגיע ל 44 זטה-בייט (בייט הוא יחידת הזיכרון הדיגיטלית המינימלית, זאטה-בייט אחד, מתאר מספר שבו יש 21 אפסים לאחר הספרה אחת). הפורום מעריך כי סביב 2025, ייצרו 463 אקסה בייטס מדי יום (אקסה-בייט אחד, מתאר מספר שבו יש 18 אפסים לאחר הספרה אחת). זוהי כמות מידע השווה למידע האגור על למעלה מ 212 מיליון CD. חלק מן המידע הנוצר כעת כולל, מדי יום: כ - 294 מיליארד אימליים, חצי מיליארד טוויטים, וכ - 65 מיליארד הודעות בוטסאפ.<sup>6</sup> המגמה של מחשוב של מוצרים רבים והחזון של אינטרנט של עצמים (THE INTERNET OF THINGS) צפוי להעצים עוד יותר את כמות המידע. על פי האקונומיסט, עד 2035, יהיו כטריליון מחשבים שיהיו מחוברים זה לזה.<sup>7</sup>

מחברי הספר מביאים מספר דוגמאות בכדי להדגים איך נראה BIG DATA בפרויקטים קונקרטיים: פרויקט למיפוי גרמי השמיים SLOAN DIGITAL SKY SURVEY שהונע בשנת 2000, אסף בתוך מספר שבועות יותר מידע ממה שנאסף בכלל ההיסטוריה של האסטרונומיה. הכותבים גם מזכירים כי לאחר המצאת מכונת הדפוס של גוטנברג במאה ה - 15, חלפו כחמישים שנה, בטרם כמות הספרים שהודפסה באירופה השתוותה לכמות הספרים שנוצרו טרם המצאה המכונה. לדברי הכותבים בעת פרסום הספר (ב - 2013), כמות המידע בעולם הוכפלה כל שלוש שנים. ב - 2018 העריכה חברת IBM כי בתחומים מסוימים כלל המידע מוכפל בתוך מספר ימים.<sup>8</sup>

**מהירות (VELOCITY):** המידע נוצר במהירות רבה, ומעובד במהירות רבה.

---

<sup>3</sup> Peter Zadrozny and Raghu Kodali, Big Data Analytics using Splunk, New-York: Apress, 2013, p. 1  
<sup>4</sup> Anil Jain, The 5 Vs of Big data, IBM, 17 September, 2016, see: <https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/>, Bernard Mar, Why only one of the five Vs of big Data really matters, IBM, March 19, 2015. See: <https://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters>

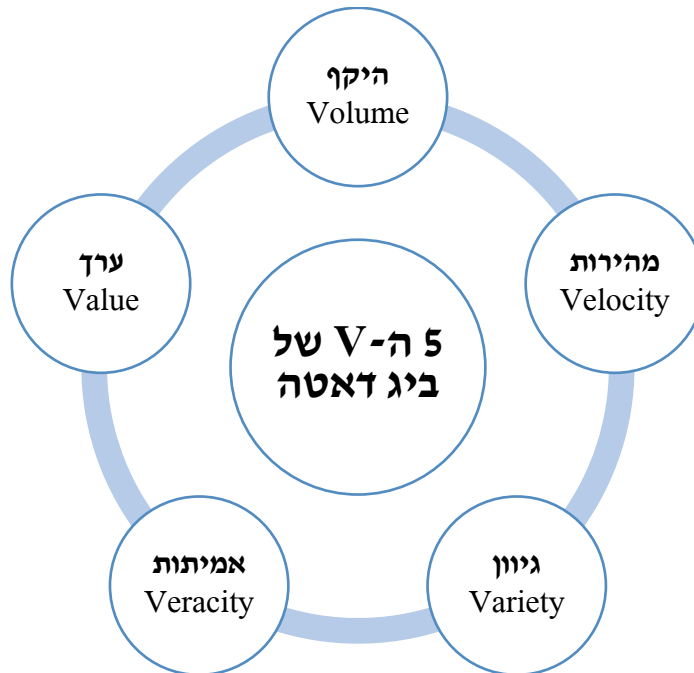
בגלגול הראשון, כללה ההגדרה רק את שלושת הראשוניות: היקף, מהירות ומגוון. מעבר לחמישה המוצגים למעלה יש המוסיפים גם: השתנות ערכיות המידע (VARIABILITY), והצגה גרפית של המידע (VISUALIZATION).  
<sup>5</sup> Stephn Pew, Big Data, Big Dupe: A Little Book about a bunch of non-sense. (Analytics press: El-Dorado Hills, CA, 2018).

<sup>6</sup> World Economic Forum, *How Much Data is Created Everyday*, See: <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/> accessed May 12, (2020).

<sup>7</sup> Economist Staff, Chips with Everything: How the world will change as computers spread into everyday objects, *The Economist*, 14 September, 2019

<sup>8</sup> IBM, IBM, How do we skill up the next generation of IT talent? IBM has a plan with New Collar <https://www.ibm.com/blogs/ibm-training/how-do-we-skill-up-the-next-generation-of-it-talent-ibm-has-a-plan-with-new-collar/> (accessed May 6, 2020)

**גיוון (VARIETY):** המידע הנצבר הוא מגוון מאד. החלוקה המסורתית היא בין מידע מובנה (STRUCTURED DATA) כלומר מידע שכבר הוטמע במאגר מידע (למשל: טבלת אקסל) ומידע שאינו מובנה (UN STRUCTURED DATA) כמו טקסט, תמונות, קבצי אודיו וכד'.  
**אמיתות (VERACITY):** המידע הוא בעל רמות שונות של אמת ואינו ודאי.  
**ערך (VALUE):** המידע הוא בעל רמות שונות של ערכיות למשתמש בו. טיעון מרכזי בהקשר זה הוא כי ביג דאטה מאפשר בגלל היקפו ובהינתן יכולות עבוד מתקדמות, להעניק תובנות בסדר גודל שונה באופן דרמטי, מאילו שניתן היה לגזור מהיקפי מידע קטנים יותר?



הספר בן שבע השנים שרד, כאמור, את התחלופה המהירה שהעידן הנוכחי מבשר; הוא עדיין נמכר בחנויות הספרים (כולל בעמק הסיליקון!) כמבוא כללי, לא טכני, לנושא. שני העמודים הראשונים של הספר מחדדים מיד את מידת האקטואליות שלו, לפחות לקורא בימי הקורונה. בכדי להדגים מהו ביג דאטה ומדוע הוא שימושי, הספר מראה כיצד אלגוריתמים של גוגל שניתחו את תפוצתן הגאוגרפית של מילים מסוימות במנוע החיפוש של החברה היו לכלי היעיל ביותר בבניית מפת התפוצה הגאוגרפית של הוירוס H1N1, שהביא למגפת השפעת ב-2009. הטקסט ממשיך לסיפור נוסף עם נגיעה רלוונטית, הפעם לקורא הישראלי. זאת, תוך תיאור הישגיו העסקיים של הפרופ' הישראלי-אמריקני אורן עציוני (בנו של לוחם הפלמ"ח והסוציולוג הנודע אמיתי עציוני). החברה הראשונה שלו, פיירקאסט, ייצרה כבר ב-2004 תחזיות לגבי עלות עתידית של כרטיסי טיסה בהתבסס על מיצוי מסדי נתונים של הזמנות כרטיסי טיסה בכדי להקל על צרכנים להחליט מתי יהיה מועד טוב לרכוש כרטיס.

הכותבים מייחסים את העלייה בחשיבות של הביג דאטה להתפתחויות טכנולוגיות, בייחוד לצמיחה הדרמטית בהיקף המידע, לצד הרחבה של האפשרויות הטכניות לטיפול בו. בה בעת, הם מראים כי בבסיס התופעה עומד

גם שינוי תודעתי ביחס לערך של מידע ולאופן בו הוא משרת יעדים כלכליים וחברתיים. בעוד שבעבר מטרתו של המידע הסתיימה בעת שהועברו הנתונים הנדרשים (למשל: הושלם החיפוש אותו ביצענו במנוע חיפוש), הרי שעתה צריך לחשוב, לשיטתם של הכותבים, על המידע כחומר גלם לגיבוש תובנות נוספות, כמו היכולת לבסס תמונת מצב לגבי תפוצת מגיפה לפי תדירות השימוש במונחים מסוימים במנועי חיפוש. בנסיבות אלה, לארגוציה של המידע יש ערך כלכלי בפני עצמו, מעבר לערך הגלום בפיסות המידע הספציפיות.

ליבת הספר מקדמת שלוש טענות מרכזיות. ראשית, היקפי הביג דאטה ויכולות החישוב מאפשרות לעבד את כלל המידע הנדרש לשאלה מסוימת. זאת, בשונה מן המציאות שהחלה במאה ה-19, בה רוב המידע שנאסף בכדי להתמודד עם צרכים חברתיים, התבסס על מדגם בלבד. איסוף כלל הנתונים בסוגיה מסוימת גם מעלה את רמת המהימנות וגם מאפשר את העמקת ההבנה שלנו לגבי המידע לרבות ראיית גוונים שונים בתוך המידע. שנית, איסוף כלל הנתונים מאפשר לוותר במידת מה על דיוק בפיסות המידע הקונקרטיים. לדברי הכותבים, ההישענות על מדגם בכדי להבין קבוצה גדולה (כמו, משאלי דעת קהל), דחפה למאמץ גדול לדיוק במדידה. באופן אירוני, אולי, תמונת מציאות המבוססת על כלל המידע, והמעניקה תמונה ברורה על המגמה, אינה מחייבת דיוק מלא היות שהמגמה ברורה מכלל הנתונים. לבסוף, היכולת לאסוף ולנתח את כלל המידע בנושא מסוים, מעבירה את הדגש בחקירה האנושית של המציאות, מהישענות על לוגיקה סיבתית (מדוע קוקטייל מסוים של תרופות מרפא סרטן?) ללוגיקה של קורלציה או מתאם.

הספר מסביר כי ההשפעה המצטברת של שלושת השינויים הללו היא שינוי מהותי של סדרי יסוד חברתיים בדגש על העולמות הכלכליים-חומריים אך לא רק על אלה. כך למשל, המעבר בין הבנת המציאות דרך סיבתיות להבנתה דרך מתאם (קורלציה) מאיים לערער נדבכים שלמים של האופן בו בני אנוש מבארים את המציאות והכותבים תוהים מה יהא על אינטואיציה? פעולה אנושית בניגוד לעובדות? לימוד מניסיון אנושי?

מאז שנכתב הספר דומה שדווקא התחזקו הספקות לגבי המענה המושלם שביג דאטה אמור לתת. זאת, בין היתר נוכח מה שהתברר כמגבלות של יכולות הניתוח. פרויקטים כמו "קוד המוטה" של החוקרת והפעילה החברתית ג'וי בולמוויני,<sup>10</sup> הראו כי אלגוריתמים מסוימים משכפלים הטיות חברתיות קיימות ובמקום לספק תמונה מלאה של המציאות, הם רק מקצינים קשיים קיימים בתפיסת המציאות. היבטים משמעותיים של משבר הקורונה גם הם מחזקים את התחושה כי המפץ הגדול עדיין רחוק. כך למשל, המספר המצומצם של בדיקות זמינות לנגיף בחודשים הראשונים של המשבר, חידד כי אנו עדיין חיים בעולם של מדגמים ולא הבנה מלאה של המציאות. עצם ההתפרצות של הנגיף והמגבלות חסרות התקדים על פעילות אנושית שבאו בעקבותיה, גם הם מזכירים כי בסופו של יום חיזוי הסתברותי על בסיס אירועי העבר הוא מוגבל. דווקא מרחבים בהם הדמיון האנושי דומיננטי, כמו ספרות או סרטים אפוקליפטיים, היו אפקטיביים אולי יותר בחיזוי.

שני פרקים נוספים עוסקים בהיבטים רחבים של מידוע (DATAFICATION). בראשון, ממקמים הכותבים את הביג דאטה על פני רצף היסטורי ארוך שנועד לחדד אצל הקורא את האבחנה בין דאטה לבין דיגיטציה. הסקירה שלהם מזכירה כי בני אדם החלו למדד את עצמם ואת סביבתם הרבה לפני עידן המחשבים. בהמשך, המחברים מדגימים את התהליך בו פיסות ידע על המציאות, הופכות – לאחר עיבוד בסיסי – לתמונת מידע רחבה של ביג דאטה המאפשרת תובנות חדשות כמו יומני מסע נשכחים של צי האמריקני, שעובדו במאה ה-19 על

---

<sup>10</sup>Algorithmic Justice League, *Coded Bias*, <https://www.ajlunited.org/> (accessed May 11, 2020)

ידי קצין צי בשם מתיו פונטיין מאורי, לטבלאות מומלצות לנתיבי שייט בהתבסס על דפוסים שאיתר בזרמים בים וברוחות השמיים.

המחברים מחדדים כי היכולת לתעד מידע היא בעיניהם אחד מן הקווים המבחינים בין חברות פרימיטיביות ומתקדמות. זאת, בייחוד בגלל שעצם התיעוד מאפשר חזרה על פעולות קודמות. הם מראים כיצד חברות אנושיות הרחיבו את הממדים שהן מדדו בהם את המציאות, ומראים את השינויים מן העשורים האחרונים בהם הפכו אותיות, מיקום גאוגרפי ואינטראקציות אנושיות לדאטה. בסיום הפרק המחברים תוהים האם הדורות הבאים יעברו למנטליות שהם מכנים "מודעות של ביג דאטה", כלומר ראיית המציאות רק דרך משקפיים של דאטה. המחברים מתייחסים גם לסוגיית צורת המדידה שהשתנתה ומראים כי שינוי צורני של שיטת מניה, סייע ליכולת לחשב את הנמדד.

המחברים: מימין, VIKTOR MAYER-SCHÖNBERGER; משמאל, KENNETH CUKIER



מתוך The Economist



מתוך Chartwell

בפרק השני העוסק ב"דאטפיקציה" המחברים מראים כי המגזר העסקי היה הראשון למצות את הערך הגלום בתופעה. הם מראים כיצד הפך מידע ממרכיב שתמך בעסקאות בשווקים שונים (למשל: מידע על המסגרת המשפטית של חוזים שנכרתו בין צדדים או מידע על שוויין של סחורות שונות), לטובין שעומדים במוקד המסחר. בנסיבות אלה, המודלים העסקיים של חברות משתנים, ויש צורך לחשב את הערך של מידע המצוי בידי חברות לא רק בשל ערכו העכשווי אלא גם, ככל שניתן, נוכח שוויו העתידי.

הכותבים טוענים כי ב"עידן המידע כל המידע חשוב...אפילו המידע הגולמי והזניח ביותר" (עמ' 100). לדבריהם, יש לחשוב על הערך של המידע כקרחון: רק חלק קטן מן הערך שהוא משקף בולט ורוב הערך עדיין נסתר. דוגמא עדכנית היא עסקת המכירה של חברת מוביט הישראלית לחברת אינטל העולמית במאי 2020, שהונעה במידה

רבה, מן הרצון של אינטל לשלב את המידע ש"מוביטי" אספה, אוספת ותאסוף על הרגלי השימוש בתחבורה ציבורית, והתנהלות מפעיליה.<sup>11</sup>

המחברים סוקרים את הנתונים העקרוניים העיקריים בהם מידע שנאסף יוצר ערך לחברות:

1. שימוש חוזר במידע (DATA REUSE) בכדי לחלץ תובנות שונות מאלו להן נועד המידע מלכתחילה. למשל, שימוש חוזר במונחי החיפוש לאחר שהסתיים תהליך החיפוש הבדיד, בכדי לאתר מגמות כלליות באוכלוסייה.
2. הארכת המידע (EXTENSIBLE DATA) בעת איסופו, כדי לאפשר שימוש משני במידע. דוגמה בולטת לגישה זו היא האיסוף הרב-ממדי של מידע שנעשה על-ידי המכונית שמפעילה חברת גוגל בשביל צילומי STREET-TV. המכוניות לא רק מצלמות את הרחובות בהן הן נוסעות, אלא גם שומרות מידע כמו מיקומי גי.פי.אס, ושמות של רשתות "וויפי" באזור.
3. היתוך של שני מאגרי נתונים נפרדים בכדי לחלץ תובנות חדשות (DATA FUSION) למשל, פרויקט דני שכלל היתוך מאגר המידע חולי סרטן בדנמרק המידע של בעלי טלפונים ניידים במדינה בכדי לנסות לאתר קורלציה בין שתי הקבוצות.
4. הנגשת מידע (OPEN DATA), בייחוד כזה שנאסף על ידי רשויות המדינה. הגם ש"מרבית, בהפרש גדול, של ההיקף הנרחב של המידע שנאסף, נערך ומנוצל בנוגע לפעולותיהם של אזרחים אמריקנים מתרחש המגזר הפרטי"<sup>12</sup>

הכותבים מצרפים את קולם למגמה (שהתחדדה בייחוד בימי ממשל אובמה) לפיה יש להנגיש לציבור את המידע הממשלתי בכדי למצות את הערך הכלכלי הגלום בו. אולם, בכדי למצות אותו נדרשת פתיחות וחשיבה יצירתית, לפי גישה זו על המדינה לחלוק את הביג דאטה עם המגזר העסקי בכדי למצות את הערך שלו. דרך נוספת למצות את הביג דאטה היא ניתוח מדויק של "שרשרת הערך" של המידע: הבנה ברורה מי מחזיק במאגרי המידע, מי מחזיק ביכולת החישובית למצות אותו, ומי במגזר העסקי פועל מתוך ראייה שמודרכת בראש ובראשונה מדאטה. הכותבים סבורים כי הקבוצה שמפיקה את מירב הערך בשרשרת הזו היא דווקא הקבוצה השלישית. משמעות אחת העולה ממצאות זו היא ירידת השיפוט האנושי בתהליכי מחקר וקבלת החלטות ועליית מומחיות בתחומי מדעי הנתונים.

המחברים חותמים את הספר בהתייחסות לסיכונים שמביא עמו עידן הביג דאטה ובכללם שאלות של אובדן פרטיות, האפקט החברתי של השענות על סטטיסטיקה והסתברות ובייחוד הסכנה שבלחזות פעילות עתידית (למשל, מועדות - רציסיביזם - של אסיר משוחרר) בהתבסס על דפוסי עבר. בנוסף, החוקרים מזהירים מן הפיתוי שבהישענות על נתונים מספריים בלבד. כל הנתונים, הם מזכירים, יכולים להיות מוטים, מנותחים בצורה לא נכונה והחמור מכל: עשויים שלא לתאר את הממד המהותי של התופעה אותה אנו מנסים לחקור.

---

Christine Fisher, Intel buys transit app maker Moovit to advance its mobility ambitions, *Engadget.com*<sup>11</sup>  
<https://www.engadget.com/intel-mobileye-acquires-moovit-mobility-185110875.html>

Paul R. Pillar, "Big Data: Public and Private", *Brookings Institution*, December 10, 2013,<sup>12</sup>  
<https://www.brookings.edu/opinions/big-data-public-and-private/>

בחלקו האחרון של הספר המחברים מציעים כמה כיוונים להתמודד עם הקשיים והאתגרים שהביג דאטה מעלה. הם מזכירים כי חשוב לשמור על הראיה לפיה לאדם יש רצון חופשי וכי בני אדם לא יפעלו רק לפי דפוסים קודמים שנותחו סטטיסטית. זה רלוונטי בייחוד במקרים שמערכת נדרשת לקבל החלטה בנוגע לאדם, על בסיס התנהגותו בעבר, כמו בבית משפט. עולם שבו אלגוריתמים יחזו הכל לא מותיר מרחב לרצון חופשי. בנוסף, הכותבים מחדדים כי יש למצוא דרך להפוך את האלגוריתמים שמעבדים את הביג דאטה לנהירים כדי להימנע ממצב בו "קופסה שחורה" תהיה המנגנון המרכזי לניתוח (ובהמשך, גם לקבלת החלטות). דרך אחת שהם מצביעים עליה היא יצירת מקצוע של בקרי אלגוריתמים, חוקרים חסרי פניות שיבחנו מעת לעת אלגוריתמים בכדי לוודא דיוק ותקפות שלהם. חשיבות הביג דאטה והיכולת לעבדו בבניה מלאכותית ברורה לארגוני המודיעין השונים. החשיפה סביב משבר הקורונה של היכולות של השב"כ לחשוף מגעים בין אנשים על פי נתוני האיכון של הטלפונים שלהם היא בוודאי רק קצה הקרחון של היכולות הקיימות.

המפגש בין התובנות של הספר לבין עבודת המודיעין מעלה שורה ארוכה של שאלות. אציע רק כמה כיוונים אפשריים:

ראשית, האם עלינו לשנות את האופן בו אנו תופסים את המקצוע המודיעיני? האם ראוי לאמץ כמסגרת מתודולוגית את העולם התיאורטי של מידע כדרך להבין את המקצוע. מדעי המידע הוא התחום האקדמי והמקצועי העוסק באיסוף, איכסון, איחזור ושימוש במידע.<sup>13</sup> היות שהגזרה המתקדמת ביותר של עולם זה מצויה כנראה בעולם העסקים, האם ראוי לחשוב על המעשה המודיעיני במונחים עסקיים של עולם הידע. זאת גם במובן של הטובין המוצעים: איזה מידע ספציפי או תובנות מביג דאטה הם הערכיים ביותר? אבל, גם במובן התהליכי: האם עלינו לחשוב על המעשה המודיעיני כשרשרת ערך עסקית, ובתוך כך לאתר את המקום הספציפי בו נוצר הערך הרב ביותר? או איתור צווארי בקבוק? אולי ההבנה כי המגזר הפרטי הוא המוביל בתחום הביג דאטה היא תובנה נושנה, אך השימוש שעשו מדינות שונות (כולל ישראל) בביג דאטה להתמודדות עם משבר הקורונה בשבועות האחרונים הזכיר לנו עד כמה המדינה עודנה מרכזית באיסוף, עיבוד ופעולה המבוססת על מידע. זאת, ולו משום שיש לה את הסמכות – בשונה מן החברות הפרטיות – להפעיל אמצעי כפיה. כך או כך, למרות חשיבותו של המגזר העסקי, לא נראה כי מערכות המודיעין הלאומיות – בייחוד במדינה כמו ישראל, בה המודיעין הצבאי מאד דומיננטי – הפכו את האתוס שלהן לאתוס עסקי. מערכות מודיעין לאומיות הן חלק ממנגנון השירות הציבורי. נוכח ההבדלים היסודיים בין השירות הציבורי למגזר הפרטי – יעדי הארגונים, השיטה שבה הם ממנים את עצמם, ובעלי העניין הרלוונטיים – לא בטוח שמערכות מודיעיניות יכולות או צריכות לעשות כן.

שנית, ברובד המתודולוגי נשאלת השאלה האם הביג דאטה הוא הכלי המתאים להוביל את הניתוח של כלל התחומים בהם מערכות המודיעין עוסקות? כאמור, הערך של ביג דאטה למודיעין מסכל הוא ברור, אבל האם הוא מתאים לציפייה של חלק ממקבלי ההחלטות שהמודיעין אסטרטגי<sup>14</sup> יחזה החלטות של תהליכי קבלת

Tefco Saracevic, *Information science*. In M. J. Bates (Ed.), *Encyclopedia of library and information sciences* (3rd ed.) New York: Taylor and Francis, 2009, pp. 2570-2585 as seen in: Association of information science and technology, <https://www.asist.org/about/what-is-information-science/> (accessed May 12, 2020)

<sup>14</sup> להגדרות שונות של מודיעין אסטרטגי וחלק מן הדיונים סביב הגדרות אלה ראה: United States Department of Defense, *Dictionary of Military and Associate Terms*, 2001, 2004, p. 505, John G. Heidenrich, *The State of Strategic Intelligence: The Intelligence Community's Neglect of Strategic Intelligence*, *Studies in Intelligence studies Vol 51(2)*, 2007; Itai Shapira, *Strategic intelligence as an art and a science: creating and*

החלטות באליטות מצומצמות של היריב? אם זה המצב, יתכן מאוד שאין מספיק מידע בכדי לנתח את דפוסי קבלת ההחלטות בכדי להגיע לתובנה משמעותית. אפילו היה מספיק מידע, עסקינן במידע שנודע לאירועי עבר ולא יסייע לנו בהכרח להבין אירועים שאין להם תקדים. למעשה, לרוב ההפתעות המודיעיניות לא היה תקדים ובוודאי שלא היו אלפי תקדימים שמאפשרים חילוץ תובנה משמעותית.

שלישית, ברובד הארגוני, הספר מעלה את שאלת עתידם של מומחי התוכן מול מומחי המידע. לכאורה, בעולם של ביג דאטה, בעל המקצוע המרכזי שנדרש לסייע למקבלי החלטות הוא איש דאטה, שידע כיצד למצות את המידע בכדי להבין את תמונת המציאות ואף לחזות אותה. המעבר האירגוני מדגש על מומחי תוכן למומחי דאטה, משקף שינוי מתודולוגי של מעבר מניתוח סיבתי שנעשה על ידי מומחה התוכן כמקובל היום (הסבר כמו: "שורת לחצים פנימיים וחיצוניים, שנפגשת עם מפקד מקומי אגרסיבי, יוצרת סבירות גבוהה לפיגוע") להבנת המציאות דרך קורלציות עליהן אמון מומחה הדאטה ( ניתוח כמו: "אם חמישה חברי חוליה נפגשים, יתרחש פיגוע. המכניזם המדויק לכך, אינו ידוע"). זהו שינוי מן עמדה המסורתית של רוב מערכות המחקר המודיעיני שהדגישו לאורך שנים דווקא את המאמץ "להיכנס לראשו של היריב" ועודן מקדישות את מירב ההכנה של חוקרי המודיעין ללימוד התוכן המהותי.

בסופו של יום, ככל שהעולם, ובכלל זה העולם המודיעיני, הופך לכזה שממוקד בנתונים, כימותם וניתוחם, ראוי אולי לחזור לדבריו של הוגה בלתי-כמותני בעליל, אלן בלום. בספר על אמנות הקריאה ( HOW TO READ AND WHY) תהה בלום כבר בשנת 2000: "מידע עומד לרשותנו באופן אינסופי. אך היכן תימצא התבונה?". הספר שנסקר אינו נותן מענה מלא לתהייתו של בלום, אך בוודאי משרטט חלק מקווי המתאר של התשובה.

---

using conceptual frameworks, *Intelligence and National Security*, Published online 22 October 2019; David T. Moore, ; Kevjn 1. *Sensemaking: A structure for an intelligence revolution*, Washington DC: Defense Intelligence College, 201  
Sherman Kent, ; Lim, Big Data and Strategic Intelligence, *Intelligence and National Security* 31(4) 2016, pp. 619-635  
]Strategic Intelligence for American Foreign Policy, Princeton: Princeton University Press, 1949 [1956