

# "קמ"ן המידע" - תפיסה חדשה של המחקר המודיעיני

רס"ן (מיל') ד"ג - לשעבר מפקד המעבדה לחדשנות מתודולוגית בחטיבת המחקר

## רקע

עוד מתקופת היישוב והש"י (שירות הידיעות) הושפע המודיעין בארץ מתפיסות האקדמיה בישראל.<sup>22</sup> קווי הדמיון בין המחקר האקדמי לבין המחקר המודיעיני מאפשרים למודיעין להתבסס על הידע התיאורטי המתודולוגי שנרכש באקדמיה, כמו גם על שיטות מחקר מדעיות כדי לשכלל את שיטות פיתוח הידע שלו. זאת, כמובן, תוך מציאת ההתאמות הדרושות להקשר הצבאי-מודיעיני של פיתוח הידע. בעשורים האחרונים בולטת בהקשר זה תרומתו של יצחק בן-ישראל לשיח התפיסתי המודיעיני.<sup>23</sup> אחת ההתפתחויות המתודולוגיות המעניינות במדעי החברה והרוח בשנים האחרונות היא הניסיון לרתום את תחום המחשוב לטובת תחומי מחקר איכותניים. מגמה זו באה לידי ביטוי באסכולות מחשבתיות שונות: מדעי חברה ממוחשבים,<sup>24</sup> מדעי רוח דיגיטליים,<sup>25</sup> ניתוח תרבות,<sup>26</sup> קריאה רחוקה.<sup>27</sup>

התפתחות מתודולוגית-אקדמית זו מתאימה ככפפה ליד לרבים מהאתגרים שמלווים את קהילת המודיעין בעת הנוכחית, וזאת בשל הדמיון בתהליכי השינוי של מספר מעצבים בסביבת הפעולה של תחומי מחקר אלה ושל ארגוני המודיעין, בעיקר בתחום היצף המידע הזמין באינטרנט, בליווינות המסחרית ועוד. בסביבה המודיעינית יש להוסיף גם את הפוטנציאל האיסופי חסר התקדים של ממד הסייבר על היצף המידע שזה מביא עמו.

מעצב נוסף משותף לאקדמיה ולמודיעין בהקשר זה הוא הרצון והיכולת ללמוד מהתפתחויות טכנולוגיות ומתודולוגיות במגזר העסקי. שני המגזרים שואפים "לייבא" כלים ושיטות מתחום מדע הנתונים (Data Science), כלים שפותחו, למשל, לצורך ייעול הקצאת משאבים בתחום השיווק המקוון הממוקד – "מגויסים" ומותאמים לצרכים מחקריים.

מטבע הדברים, השינויים בסביבת המידע ובכלים העומדים לרשות החוקרים משפיעים על האקדמיה ועל קהילת המודיעין לא רק בישראל. מכון RAND האמריקאי, למשל, פרסם לאחרונה מחקר שהוזמן על ידי קהילת המודיעין האמריקנית, ובו כיווני מענה עבור שילוב יכולות Data Science במשרד ההגנה האמריקאי.<sup>28</sup> התפתחויות אלו הובילו לאחרונה את אמ"ן לגבש תפיסה חדשה במחקר המודיעיני: תפיסת "קמ"ן המידע", המשפיעה הן על תפיסת הפעלה והן על בניין

22 גיל אייל. הסרת הקסם מן המזרח. (ואן ליר, 2008).

23 יצחק בן-ישראל. דיאלוגים על מדע ומודיעין (תל-אביב: משרד הביטחון. יצחק בן-ישראל, 1988). הפילוסופיה של המודיעין. (תל-אביב: משרד הביטחון, 1999).

24 Lazer, D', Pentland, A', Adamic, L', Aral, S', Barabasi, A' L', & Brewer, D'. (6 Feb 2009). "Life in the network: the coming age of computational social science". *Science*, 721-723

Berry, D'. (2012). "Understanding Digital Humanities". *Palgrave Macmillan*.

26 "Manovich, L'. (2017). "Cultural Analytics, Social Computing. Amsterdam University

27 .Moretti, F'. (2013). Distant Reading. Verso Books

28 Knopp, B', Beaghley, S', Frank, A', Orrie, R', & Watson, M'. (2016). "Defining the Roles, Responsibilities, and Functions for Data Science Within the Defense Intelligence Agency

הכוח באמ"ן. במאמר זה אנסה להמחיש את הכיוונים, שאליהם מכוונת תפיסה זו, הן מבחינת התועלת המודיעינית והן מבחינת הדרישות בתחום בניין הכוח.



## **על תחום ה-Data Science כיום ועל השלכות מודיעיניות מתבקשות**

בחלק זה אסקור בקצרה שתי טיפולוגיות שתעזורנה להבין את הפוטנציאל הקיים באימוץ כלים אלו כבר היום לצרכים מודיעיניים: סוגי שאלות, שלגביהן קיימים כיום כלים מתחום מדע הנתונים, וסוגי נתונים, שעליהם ניתן להפעיל מניפולציות שכאלו.

### **סוגי שאלות**

מאמר שפורסם לאחרונה באתר של חברת מיקרוסופט מציע חלוקה לחמישה סוגי שאלות, שלגביהן קיימים כיום כלים מתחום מדע הנתונים:

- **סיווג** ("האם מדובר ב-X או ב-Y?") - בעולם העסקי אפשר לחשוב על שאלות, כגון "האם מייל מסוים הוא דואר זבל?". בעולם מדעי הרוח אפשר למצוא דוגמאות לשאלות, כגון "האם החיבור נכתב על ידי מחבר א' או ב'<sup>29</sup>? בעולם המודיעיני אפשר לחשוב על שאלות, כגון "האם דיווח X אותנטי, או שמא מדובר בניסיון זיוף?", "האם יום X הוא יום שבו צפוי אימון של יחידה מסוימת, או שמא מדובר ביום שגרתי ביחידה?" וכו'.
- **זיהוי חריגות** ("האם X יוצא דופן?") - בעולם העסקי אפשר לחשוב על שאלות, כגון "האם פעולה מסוימת בכרטיס האשראי מעידה על ניסיון זיוף?" בעולם המודיעיני ההשלכות בתחום ההתרעה ברורות ("האם התנהלות יחידה X שגרתי או חריגה?", "האם התנהלות אדם מסוים

Khmelev, D', & Tweedie, F'. (2001). "Using Markov Chains for Identification of Writers. *Literary and Linguistic Computing*", 299-307

- שגרתית או חריגה?", ולמעשה מיושמות מזה מספר שנים במספר ארגונים.<sup>30</sup>
  - **הערכת כמות** ("כמה X צפויים להיות?") – בעולם העסקי אפשר לחשוב על שאלות, כגון "מה צפוי להיות הביקוש למוצר מסוים ברבעון הבא?" בעולם מדעי החברה אפשר למצוא דוגמאות לניסיונות בעלי חשיבות מודיעינית לכשעצמם לחזות מספר אירועי משברים פנימיים במדינות.<sup>31</sup> בעולם המודיעיני ניתן להצביע על שאלות מודיעיניות קלאסיות מתחום הערכת כשירות אויב, העוסקות אינהרנטית בהערכת כמויות (אמל"ח במלאים, למשל).
  - **זיהוי מקבצים** ("לפי אילו קבוצות אפשר לארגן את המידע?") – בעולם העסקי אפשר לחשוב על שאלות, כגון "אילו קבוצות משתמשים קיימות בקרב לקוחות החברה?" בעולם מדעי החברה ובמודיעין אפשר למצוא דוגמאות, כגון חלוקה ממוכנת לקבוצות של משתמשי טוויטר, המצייצים אודות הלחימה בסוריה.<sup>32</sup>
  - **סיוע בקבלת החלטות ממוכנת** ("מה כדאי שאעשה עכשיו?") – בעולם העסקי קיימות מערכות אוטונומיות שצריכות לקבל הרבה החלטות קטנות באופן עצמאי ("האם כדאי ששואב האבק הרבובטי ימשיך לשאוב את החדר, או יחזור לנקודת הטעינה?"). אומנם קשה לחשוב על דוגמאות אקדמיות לסוג כזה של שימוש בנתונים, אולם דווקא בהקשר המודיעיני-מבצעי ישנן מספר דוגמאות, כמו כטמ"מ האיסוף האוטונומי Perdix שבפיתוח הפנטגון, שעל פי הפרסומים הרשמיים, מקבל החלטות באשר לאופן ביצוע משימת האיסוף.<sup>33</sup> ללא מעורבות אדם.
- נקודה מעניינת בהקשר לחלוקה זו היא שבניגוד לאינטואיציה, השאלות, שבהן אפשר להשתמש ביכולות מחשוב, נוגעות הן לתהליכי מודיעין דדוקטיביים המבוססים על הגדרת השאלה מראש, והן לתהליכי מודיעין אינדוקטיביים המבוססים על זיהוי תופעות בנתונים עצמם (כגון שאלת זיהוי המקבצים). הנקודה המכרעת הופכת להיות שאלת סוג הנתונים שלגביהם קיימות פרקטיקות מוכרות לתפלול (מניפולציה) לטובת שאלות מעין אלה.

## סוגי נתונים

- קיימים שלושה סוגים עיקריים של נתונים אשר להם "הוכחת היתכנות" בתחום מדע הנתונים ורלוונטיות מודיעינית ברורה. עם זאת, מובן שאפשר לבצע מניפולציות מהסוגים שנוכרו לעיל גם על סוגי נתונים נוספים. שלושת הסוגים הם:
- **מידע מובנה** – מידע שנמצא ברמת ארגון גבוהה כדוגמת מאגרי נתונים טבלאיים, שבהם מיקום הנתון (באיזו טבלה נמצא, באיזו עמודה, ובאיזו שורה) מעיד על "ההקשר" שלו. דוגמא לנתונים מסוג זה אפשר למצוא באתרים רבים (רשתות חברתיות, אתרים ממשלתיים ועוד)

Marr. **Big Data In Practice**. 2016. Wiley. Bernard 30  
 Keneshloo, Y', Korkmaz, G', Cadena, J', & Ramakrishnan, N'. (2014). "Detecting and forecasting domestic political 31  
 (crises: a graph-based approach". Proceedings of the 2014 ACM conference on Web science, (192-196  
 Lynch, M', Freelon, D', & Aday, S'. (2014). **Syria's Socially Mediated Civil War**. Washington DC: United States Institute 32  
 Of Peace  
 .Perdix Fact Sheet .(2017) אתר הפנטגון 33  
<https://www.defense.gov/Portals/1/Documents/pubs/Perdix%20Fact%20Sheet.pdf>

המאפשרים "לתשאל" את מאגר הנתונים, שעליו מתבסס האתר באמצעות שירות המכונה API ולקבל כפלט מידע מובנה. מידע מובנה מוגדר כמידע המאפשר בקלות יחסית מניפולציות כגון אלו שפורטו לעיל, כאשר קיימת כמות גדולה דיה של מידע אמין, מלא ובעל "רעש" מועט.

- **מידע טקסטואלי** – ישנם מחקרים רבים הנוקטים את גישת "Data as Text" שבמסגרתה החוקר אינו קורא את הטקסט ומפרש אותו במוחו, אלא משתמש במחשב לביצוע מניפולציות המבוססות על שכיחות מילים בחיבורים שונים. בדומה לכך, ישנם מחקרים העוסקים בזיהוי נימת הדובר על בסיס אוצר המילים שבו הוא משתמש, בזיהוי מאפיינים המתייחסים לדמות מחבר הטקסט, ובמציאת דמיון בין טקסטים שונים, כגון מציאת מפלגות בעלות עמדה פוליטית דומה על בסיס אוצר המילים שבמצעיהן וכו'.<sup>34</sup>
- **מידע ויזואלי** – ישנם מחקרים רבים מתחום "ראיית המכונה" (Vision Computer) העוסקים בזיהוי עצמים בתמונות, בזיהוי פנים וסמלים, בחלוקת תמונות למקצבים וכד'.



סוגי נתונים אלו זמינים לאיש המודיעין בכל רמות הניתוח ולכלל מושאי המחקר, בין אם על בסיס חומר גלוי, חומר מסחרי או חומר שהושג באופן אחר.

## חזון קמ"ן המידע

התיאור התיאורטי של התחום עשוי להישמע מעורפל לקורא שאינו מעורה בהתפתחויות שחלו בו במגזר האזרחי או בקהילת המודיעין. אנסה להדגים את כיווני המחשבה שאליהם מוביל התחום. אעשה זאת דווקא בתחום, שנתפס אולי כפחות אינטואיטיבי לכך, תחום המחקר המדיני. המחקר המדיני מבוסס במידה רבה על ניתוח פעולות המתרחשות במרחב הציבורי, ולכן יש בו חשיבות רבה לאיסוף הגלוי. תפיסת המודיעין המסורתית מדגישה מאוד ניתוח איכותני (של טקסטים, תמונות וכו') ו"קריאה מעמיקה" של כל פיסת מידע. אולם בעידן "התפוצצות המידע",

Grimmer, J., & Brandon, S' M'. (2013). "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Text". *Political Analysis*, 1-31 34

המודיעין הגלוי "מוצף" בשלל פיסות מידע: כתבות בתקשורת הכתובה, בטלוויזיה, ברדיו, מאמרי דעה, ראיונות, נאומים, תמלילי ועידות, פרסומים רשמיים (למשל נוסחי חקיקה, פרסומי לשכות מרכזיות לסטטיסטיקה) וכמובן, התבטאויות ברשתות החברתיות.

"קריאה מעמיקה" של כל אחת מפיסות המידע הללו דורשת לא רק ידע בשפה הרלוונטית, אלא גם היכרות עם הזירה הפוליטית הרלוונטית, עם מודלים רלוונטיים מתחום מדע המדינה, עם היחסים הבינלאומיים ועוד. על כן, בגישת המודיעין המסורתית כדי לבצע תהליך מודיעיני מיטבי יש לחלק את הכוח בין "גורמי איסוף והפקה", שתפקידם לקרוא כמה שיותר פיסות מידע, ולסנן מתוכן רק את "החשובות ביותר", לבין "גורמי מחקר", שתפקידם להכיר באופן מעמיק את הזירה ואת המודלים התיאורטיים ולהרכיב מתוכם "תמונת מצב".

גישת "קמ"ן המידע" גורסת, לעומת זאת, כי אפשר להפיק ידע מודיעיני רלוונטי גם מרבבות פיסות המידע "החשובות פחות". לפי גרסה זו, יש להשתמש לשם כך בכלים סטטיסטיים ומחשוביים (שהרי משאבי כוח האדם אינם מאפשרים להתעמק בכל פיסות המידע). כך אפשר לחשוב על תהליך מודיעיני, שמהותו אינה לסנן רק את מעט הטקסטים החשובים ולקרוא אותם לעומק, אלא, לזהות, למשל, באמצעות שכיחות המילים בכלל הטקסטים, מה עומד על סדר היום הציבורי במדינה כלשהי, או מה התנודות בנוגע למסגור של ישראל בהתבטאויות רשמיות של גורמי משטר זה או אחר או בקרב ציבורים שלמים.

לכך אפשר להוסיף גם ניתוח של מגמות המתייחסות לביקורי דיפלומטים במדינות העולם. גם אם לא כל ביקור חשוב בפני עצמו, אפשר לזהות מגמות של "התקרבות" או "התרחקות" בין מדינות על בסיס מאגר גדול של פרטים מעין אלו, לנתח דפוסי הצבעה משותפים במוסדות בינלאומיים שונים, התייחסויות משותפות בעיתונות העולמית, נתוני סחר, ועוד. המחקר המדיני הוא דוגמה אחת (נוחה), אך כיוונים מתודולוגיים דומים אפשריים

**גישת "קמ"ן המידע"  
גורסת, כי יש להפיק  
ידע מודיעיני רלוונטי גם  
מרבבות פיסות המידע  
"החשובות פחות", ולשם  
כך יש להשתמש בכלים  
סטטיסטיים ומחשוביים**

גם בתחום המחקר הצבאי, מחקר הטרור, המחקר הכלכלי והחברתי, המחקר הטכנולוגי וללא ספק גם בתחום הסייבר (תחום רווי נתונים במהותו).

אסכם ואומר, שבמקום תהליך מודיעיני המעוצב לאור הצורך "לסנן" מעט פיסות מידע "חשובות" מתוך "בליל הרעש", מקדמת תפיסת "קמ"ן המידע" תהליך מודיעיני המעוצב לאור ניסיון ליצור ערך מודיעיני על ידי שימוש בכלים סטטיסטיים ומחשוביים לקיבוץ של מירב המידע, כולל פיסות מידע "זניחות", לכאורה.

## השלכות על בניין הכוח המודיעיני

**בתחום ההכשרות** – בדומה לתהליכים המתרחשים בפקולטות למדעי החברה והרוח כיום, התאמת המודיעין לתפיסה החדשה מחייבת שינוי במיומנויות המוקנות לחוקרים. תפיסת "קמ"ן המידע" שואפת להשפיע על ה-DNA של איש המחקר המודיעיני ועל התרבות הארגונית

המודיעינית העוסקת בשאלה מהו מחקר מודיעיני? באופן זה ובהתבסס על מודל המיומנויות הדרושות לצוות מדע נתונים של חברת בוז-אלן-המילטון<sup>35</sup> במסגרת תפיסת "קמ"ן המידע", אנו ממקדים את הכשרותינו בשלושה צירים:

- **בניית סיפור** – במסגרת הכשרתו של חוקר המודיעין נלמדים כלים לחשיבה ביקורתית ולניתוח איכותני באופן המדגיש יכולת בניית נראטיב קוהרנטי ומשכנע המאפשר למקבל ההחלטות (שלא היה חלק מתהליך המחקר עצמו) להבין ולהפנים את התובנות העולות ממנו.
- **חשיבה כמותית וסטטיסטיקה** – בתחום זה רוכשים כלים לבדיקת תוקפן ומהימנותן של

השערות המתייחסות ל"התנהגות" הנתונים. זאת על בסיס התפיסה שאין די לחוקר במודיעין, כמו גם באקדמיה,<sup>36</sup> ללמוד להשתמש ב"כלי מדף" המאפשרים, למשל, מניפולציות מתחום ניתוח הרשתות החברתיות (SNA) או סימולציה מבוססת סוכן (ABM). הוא אינו יכול לקבל את "גזר הדין" של אלגוריתם מסוים של חלוקה למקבץ כ"קופסא שחורה", אלא עליו להבין לעומק את המסד התיאורטי העומד מאחורי כלים אלו, כדי להיות מסוגל לעשות

בחירה מושכלת באיזה כלי עליו להשתמש לצורך איזו בעיה, ומתי עליו להודות כי אין נתונים או כלים רלוונטיים, וכי יש לנקוט בגישה אחרת.

- **מיומנויות מחשב** – להקנות לחוקרים כלים בסיסיים להשגת ניקוי, ארגון, וביצוע מניפולציות על הנתונים. כך, מושם דגש על תשאול מאגרי נתונים מובנים באמצעות שפת SQL, או תכנות בסיסי באמצעות שפת Python, שימושים מתקדמים באקסל וכד'.

### ברובד הארגוני והתפיסתי

פרט לשינויים הנדרשים ברמת מיומנויות איש המודיעין, הרי ששינוי תפיסתי זה דורש לבחון שינוי ביחסים שבין ארגוני המודיעין. כלומר, לצד שימור דפוסים קיימים, המדגישים עקרונות "מעגל המודיעין" וחיפוש צר, יש לפתח איסוף רחב של נתונים רבים אודות מושא המחקר, גם אם כל אחד בפני עצמו "חסר ערך". באופן זה, ייתכן שבעתיד יידרש גם עיצוב מחודש גם של יחסי איש המחקר עם איש האיסוף – נקודה המהווה בעת הנוכחית אתגר נוסף מיני רבים ל"מעגל המודיעין" כרעיון המסדר המרכזי של אמ"ן.<sup>37</sup>

## תפיסת "קמ"ן המידע" שואפת להשפיע על ה-"DNA" של איש המחקר המודיעיני ועל התרבות הארגונית המודיעינית העוסקת בשאלה "מהו מחקר מודיעיני?"

<sup>35</sup> Booz Allen Hamilton. (2015). "The Field Guide To Data Science".  
<sup>36</sup> Rieder, B', & Röhle, T'. (2017). "Digital Methods". Amsterdam University Press  
<sup>37</sup> דודי סימון-טוב וסא"ל עופר. "מודיעין 2.0 - גישה חדשה לעשיית מודיעין". INSS, צבא ואסטרטגיה, 2013.

## סיכום

לאורך ההיסטוריה השתנה המודיעין פעמים רבות באופן ששיקף שינויים חברתיים, טכנולוגיים ואחרים ביחס לשאלה "מהו ידע?" או "מהו הידע הרלוונטי למקבל החלטות?" בעידן הנוכחי נראה, כי "מפנה המחשוב" מהווה את אחת הבשורות המשמעותיות לשאלה זו, גם בתחומים "רכים", כגון מדעי החברה והרוח, אשר התבססו פעמים רבות על מיומנויות ניתוח איכותניות בלבד. על "קמ"ני המידע" כבעלי מקצוע מחקרי חדש ובעלי כישורים שיוקנו בעתיד לכלל החוקרים, מוטלת המשימה להתאים את המודיעין לסביבה האסטרטגית הנוכחית.