09

ביקורת המדינה על שימוש בבינה מלאכותית בגופים ציבוריים: עקרונות לבחינה מוסדית ונורמטיבית

יובל ריינפלד וענבל בן ארצי

ד"ר יובל ריינפלד הוא מרצה באוניברסיטת רייכמן, יו"ר (משותף) ועדת בינה מלאכותית בלשכת עורכי הדין וחבר בפורום המומחים הלאומי לבינה מלאכותית.

עו"ד ענבל בן ארצי היא מנכ"לית HighLaw ועוסקת בפיתוח, בהדרכה ובהטמעה של כלי AI.

המאמר בקצרה

- שילוב בינה מלאכותית לצורך קבלת החלטות על ידי גופים ציבוריים טומן בחובו סיכונים לפגיעה בזכויות אזרח.
- בישראל לא קיים חוק רוחבי בעניין השימוש בבינה המלאכותית, והרגולציה היא מבוזרת וסקטוריאלית.
- הדין האירופאי והאמריקאי עשויים להוות השראה לאימוץ עקרונות לפיקוח על השימוש בבינה מלאכותית.
- בהיעדר מסגרת חקיקתית אופקית, ביקורת המדינה עשויה לתרום לא רק לחשיפת ליקויים, אלא גם לעיצוב סטנדרט ציבורי של שימוש אחראי בבינה מלאכותית.
- לשם כך נדרשת הקמת תשתית ביקורתית חדשה, הכוללת שפה מקצועית עדכנית, מומחיות טכנולוגית, שיתוף פעולה בין-מוסדי, כלי הנחיה מוקדמים ויכולת ללמוד מן הנעשה בעולם בנושא.

מבוא

בשנים האחרונות החלו גופים ציבוריים ברחבי העולם לבחון שימוש במערכות בינה מלאכותית (להלן גם - AI) לצורכי דירוג סיכון, תיעדוף טיפול ותמיכה בקבלת החלטות. במערכות רווחה, למשל, משתמשים לעיתים בכלים המעבדים מגוון רחב של נתונים, כגון היסטוריית דיווחים, מצב כלכלי, מבנה משפחתי ושינויים במקום המגורים, כדי להפיק "ציון סיכון" המשמש בסיס לקביעת סדרי עדיפויות ולעיתים אף להחלטות רגישות הנוגעות להתערבות בחיי משפחה. ניתן למצוא דוגמאות דומות לשימוש באלגוריתמים לצורך זיהוי צרכים עתידיים ותכנון תקציבי במגזר הציבורי, כגון הקצאת משאבים במערכת הבריאות, לרבות תכנון מיטות, כוח אדם וציוד[1]. מציאות זו מעלה חשש כי כאשר אנשי מקצוע יפעלו תחת עומס כבד תגבר ההסתמכות על פלט המערכת, ואילו היכולת להסביר את ההחלטה, לנמקה באופן עצמאי ולאפשר ביקורת אפקטיבית עליה תיחלש[2].

אין מדובר בחשש תיאורטי בלבד. מערכות דומות כבר פועלות ברשויות רווחה, במחלקות הטבות סוציאליות ובמוקדי אכיפה במדינות שונות, והדיון בשילובן במגזר הציבורי הולך ומתרחב גם בישראל[3]. על רקע זה מתחדדת שאלה יסודית: כאשר גוף ציבורי משלב בינה מלאכותית בקבלת החלטות הנוגעות לזכויות האזרח, מהם הכלים הנורמטיביים והמוסדיים שעל ביקורת המדינה להפעיל כדי להבטיח שלטון חוק, אחריותיות ושמירה על זכויות? שאלה זו מקבלת משנה תוקף נוכח פערי ההסדרה הקיימים.

באיחוד האירופי נכנסה לתוקף מסגרת רגולטורית ייעודית, ה-EU AI Act, הקובעת דרישות מחמירות ביחס למערכות בינה מלאכותית, ובפרט למערכות המוגדרות כבעלות סיכון גבוה[4]. בישראל, לעומת זאת, טרם עוצבה מסגרת חקיקתית ייעודית ומחייבת בתחום זה, ונותרו פערי פיקוח ניכרים. ביקורת המדינה עשויה למלא תפקיד מרכזי בחלל זה.

מאמר זה טוען כי ביקורת המדינה היא אחד הגורמים המוסדיים המתאימים ביותר למילוי התפקיד. שלושה טעמים עיקריים תומכים בכך: ראשית, עצמאותו המוסדית של מבקר המדינה מאפשרת לו לבחון את פעולת הרשות המבצעת; שנית, סמכויותיו מקנות לו גישה למסמכים, לנתונים ולתהליכים פנימיים שאינם נגישים בדרך כלל לגורמים חיצוניים; ושלישית, ניסיונו המצטבר בבחינת חוקיות, תקינות ויעילות של גופים ציבוריים מספק בסיס מוסדי מתאים להרחבת הביקורת גם לתחום השימוש בבינה מלאכותית.

גם במבט השוואתי ניכרת מגמה ברורה של התעוררות מוסדות ביקורת עליונה לסוגיה זו. משרד הביקורת הלאומי של בריטניה (UK National Audit Office; NAO) עסק בשימוש בבינה מלאכותית בממשלה הבריטית, משרד הביקורת הממשלתי של ארצות הברית (US Government Accountability Office) פרסם מסגרת עבודה לבחינת השימוש הממשלתי ב-AI, ובמדינות אירופיות שונות התחדד הקשר בין שימוש במערכות בינה מלאכותית לבין שאלות של שקיפות, זכויות ופיקוח. ישראל יכולה ללמוד מניסיונם של גופים אלה ולהתאים את כלי הביקורת לצרכיה.

חלק א' של המאמר עוסק במהות הסיכון הייחודי שמציבות מערכות בינה מלאכותית בשירות הציבורי; חלק ב' סוקר את המסגרת הנורמטיבית הקיימת ומצביע על פערי ההסדרה בישראל, בין היתר בהשוואה לדין האירופי; חלק ג' מציע שישה עקרונות לבחינה מוסדית של שימוש בבינה מלאכותית; וחלק ד' מתרגם עקרונות אלה לכלים ביקורתיים מעשיים ומציע המלצות מדיניות. המאמר מסתיים בסיכום שמחזיר את הדיון מן המישור הנורמטיבי אל משמעותו המעשית עבור הפרט אשר עשוי להיפגע מהחלטה מינהלית הנתמכת במערכת בינה מלאכותית.

מהות הסיכון: מערכות בינה מלאכותית ציבוריות אינן כלי ניטרלי

בשירות הציבורי פועלות מגוון מערכות בינה מלאכותית: מערכות ניבוי המזהות "פרופיל סיכון", מנועי דירוג לצורכי תיעדוף טיפול, מסנני בקשות אוטומטיים ועוזרי החלטה לפקידים ולבעלי סמכות. חרף השונות ביניהן, למערכות הללו מאפיין יסודי משותף: אין מדובר בכלים מינהליים ניטרליים במובן הרגיל. אופן פעולתן, מקורות הנתונים שעליהם הן נשענות והמשקל שניתן לפלט שהן מייצרות

1 European Commission, *Artificial Intelligence in Healthcare* (n.d.), https://health.ec.europa.eu/ehealth-digital-health-and-care/artificial-intelligence-healthcare_en

2 UK Government, *Artificial Intelligence Playbook for the UK Government* (2025), https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html

3 Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, And Punish The Poor (2018).

4 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, OJ L, 12.7.2024.

עשויים להשפיע באופן ממשי על זכויות, על נגישות לשירותים ועל אופן הפעלת שיקול הדעת המינהלי. שלושה מאפיינים מרכזיים מבדילים אפוא את המערכות האלה מכלים מינהליים רגילים ומצדיקים פיקוח ייחודי עליהן, כפי שיפורט להלן.

אי-שקיפות אלגוריתמית

מודלים מבוססי למידת מכונה, ובפרט מודלים מורכבים, עשויים להפיק תוצרים שקשה לפרשם באופן אינטואיטיבי. פקיד המקבל ציון סיכון, דירוג או המלצה אופרטיבית לא תמיד יכול להסביר לאזרח כיצד התקבלה התוצאה, ולעיתים גם מפתחי המערכת עצמם אינם מסוגלים להציג הסבר מלא וברור להיגיון הפנימי של המודל. אי-שקיפות זו, הנדונה תדיר בספרות כבעיית הקופסה השחורה (Black box), מאיימת במישרין על עקרונות יסוד של המשפט הציבורי, ובהם זכות הטיעון, חובת ההנמקה ועקרון שלטון החוק[5].

כאשר אזרח מבקש לדעת מדוע בקשתו נדחתה, מדוע סווג באופן מסוים או מדוע עניינו תועדף באופן מסוים, ותשובת הרשות נשענת הלכה למעשה על פלט אלגוריתמי בלתי

5. Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information 1 - 18 (2015).

מוסבר, נפגעת יכולתו להבין את ההחלטה, לתקוף אותה ולהעמידה לביקורת[6].

הספרות האקדמית מצביעה על כמה מקורות לאי-שקיפות בפעולתן של מערכות אלגוריתמיות, ובהם מורכבות טכנית המקשה על הבנת אופן הפעולה של המודל, סודיות מסחרית או קניינית המגבילה את היכולת לבחון את המערכת וכן היעדר גילוי מספק מצד הרשות בעניין האופן שבו נעשה בה שימוש[7]. גורמים שונים אלה לאי-שקיפות עשויים להתקיים גם בהקשרים ציבוריים בישראל, ומשום כך עליהם לעמוד במוקד בחינתה של ביקורת המדינה.

הטיות מובנות

מערכות בינה מלאכותית מבוססות, במישרין או בעקיפין, על נתוני עבר. כאשר נתונים אלה משקפים דפוסים היסטוריים של אפליה, הקצאה בלתי שוויונית של משאבים או סטיגמטיזציה של קבוצות אוכלוסייה מסוימות, קיים חשש ממשי כי המודל לא רק ישחזר דפוסים אלה, אלא אף יעצים אותם.

הקושי המרכזי טמון בכך שההטיה אינה נובעת בהכרח מאפליה מכוונת במובנה הישיר, אלא מן העובדה שהמערכת מזהה קורלציות סטטיסטיות בתוך מציאות חברתית ומוסדית שכבר הייתה מוטה מלכתחילה ופועלת על בסיסן.

כך, למשל, אם בשל נגישות נמוכה, דפוסי טיפול בלתי שוויוניים או הטיות מוסדיות אחרות סווגה בעבר קבוצה מסוימת כבעלת "סיכון" גבוה יותר בהקשרים של זכאות לשירותי רווחה, של אכיפה או של גישה לשירותים ציבוריים, המודל עשוי לפרש נתונים אלה כאינדיקציה אובייקטיבית לכאורה לסיכון גם בעתיד. במובן זה, הטיה שיטתית אינה תוצר של כוונה זדונית בהכרח, אלא של מבנה הנתונים, של תנאי יצירתם ושל האופן שבו המציאות החברתית מתורגמת למשתנים חישוביים.

פוטנציאל לכשל אלגוריתמי בעל השפעה רוחבית

כאשר גוף ציבורי מפעיל מערכת בינה מלאכותית בקנה מידה רחב, הסיכון הטמון בכך אינו מתמצה בהחלטה שגויה אחת, אלא ביכולתה של אותה שגיאה להשתכפל במהירות ובהיקף נרחב. בניגוד לטעות אנוש "רגילה", אשר לרוב מוגבלת למקרה מסוים, לבעל תפקיד מסוים או להקשר עובדתי נקודתי, טעות אלגוריתמית שיטתית עלולה להשפיע בעת ובעונה אחת על ציבור רחב מאוד של פרטים.

יתרונה התפעולי של המערכת - יכולתה לפעול באופן אחיד, רציף ומהיר - הוא גם מקור הסיכון הייחודי שלה: ככל שהמערכת יעילה יותר בהפקת החלטות או המלצות בקנה מידה רחב, כך גדל גם פוטנציאל ההפצה של כשל מובנה, הטיה שיטתית או הנחת יסוד שגויה.

הקושי מתחדד עוד יותר משום שפגיעה אלגוריתמית רוחבית לא תמיד נחשפת מייד. כל אזרח או תושב חווה את ההחלטה כעניינו הפרטי בלבד, ואינו יכול לדעת אם מדובר בשגיאה חד-פעמית או בתוצאה של דפוס שיטתי המשפיע באופן דומה גם על אחרים.

במילים אחרות, אף שהפגיעה נחווית באופן אינדיבידואלי, הכשל הוא לעיתים קרובות מבני וקולקטיבי. דווקא בשל כך, מנגנוני הביקורת הרגילים, הנשענים על תלונות פרטניות או על בחינה של מקרים בודדים, עלולים להתקשות בזיהוי הבעיה בזמן אמת. כאן מתחדדת חשיבותו של מוסד ביקורת בעל פרספקטיבה מערכתית.

גוף כגון מבקר המדינה אינו נדרש לבחון רק אם החלטה מסוימת הייתה סבירה בנסיבותיו של אדם יחיד, אלא אם קיים דפוס רוחבי של קבלת החלטות לקויה, בלתי שוויונית או בלתי מוסברת. יכולתו לעיין בנתונים מצטברים, להשוות בין מקרים, לבחון נהלים, לזהות סטיות חוזרות ולחבר בין מופעים נפרדים של אותו כשל היא שמקנה לו יתרון מוסדי מובהק בבחינת מערכות AI ציבוריות. יתרון זה מתחדד בישראל גם נוכח תפקידו של מבקר המדינה כנציב תלונות הציבור, שכן בירור של תלונה פרטנית עשוי לחשוף ליקוי מערכתי רחב יותר הדורש תיקון כללי ולא מענה נקודתי בלבד. במובן זה, הפיקוח על שימוש ציבורי בבינה מלאכותית אינו נוגע רק לשאלת תקינותה של החלטה מסוימת, אלא גם, ואולי בעיקר, ליכולת לזהות בזמן כשל מערכתי בטרם יעמיק את פגיעתו בציבור רחב.

הפסיקה ברחבי העולם והספרות ההשוואתית כבר מספקות דוגמאות מוחשיות לכך שהסיכונים הכרוכים בשימוש במערכות אלגוריתמיות אינם תיאורטיים בלבד. אחת הדוגמאות הבולטות היא פרשת SyRI (System Risk Indication) בהולנד: רשויות הרווחה החלו להשתמש במערכת ממשלתית שנועדה לזהות סיכון להונאה באמצעות הצלבת מידע ממאגרי מידע ממשלתיים שונים, לרבות מידע שאספו גופים ציבוריים נוספים. בית המשפט בהאג קבע כי ההסדר הנורמטיבי שאפשר את השימוש במערכת אינו עומד בדרישות סעיף 8 לאמנה האירופית לזכויות אדם, בין היתר בשל הפגיעה הבלתי מידתית בפרטיות והיעדר שקיפות מספקת ביחס

---

6. Danielle Keats Citron, *Technological Due Process,* 85 Washington Univ. Law Rev. 1249, 1268 (2008).

7. Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms,* Big Data & Society 3 (2016).

לאופן פעולת המערכת[8]. במוקד הקושי עמדה העובדה שהמערכת פעלה על בסיס הצלבת מידע ממקורות ממשלתיים רבים, והפרט לא היה יכול לדעת בוודאות מהם הנתונים והשקלולים שהובילו לסימונו כמי שנמצא ב"סיכון". בכך המחיש המקרה כיצד שימוש שלטוני במערכת אלגוריתמית עלול לפגוע לא רק בפרטיות, אלא גם ביכולת להבין את בסיס ההחלטה ולהעמידה לביקורת אפקטיבית.

המקרה ממחיש כי כאשר המדינה מפעילה מערכת אלגוריתמית בעלת השפעה ממשית על זכויות הפרט, אין די בתכלית מינהלית ראויה - נדרשים גם מנגנוני הגנה אפקטיביים, שקיפות מספקת ואיזון נורמטיבי הולם.

דוגמה שונה אך משלימה עולה מן המחלוקת האמריקאית סביב מערכת COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), כלי להערכת סיכון פלילי המשמש, בין היתר, בשלבי המעצר, השחרור טרם המשפט, הפיקוח והענישה, לצורך הערכת הסיכון להישנות עבירה או לאי-התייצבות. תחקיר ProPublica הצביע על כך שהמערכת יצרה פערים מובהקים בשיעורי הטעות בין נאשמים שחורים לבין נאשמים לבנים. בפרט נטתה המערכת לסווג נאשמים שחורים כבעלי סיכון גבוה בשיעור גבוה לעומת כלל האוכלוסייה, גם כאשר בפועל לא שבו לבצע עבירות. הדיון המתודולוגי סביב ממצאים אלה עוד לא הסתיים, אך המקרה הפך לסמל מרכזי של החשש מפני הטיות מובנות במערכות חיזוי אלגוריתמיות, בעיקר כאשר הן פועלות על בסיס נתוני עבר המשקפים פערים חברתיים ומוסדיים קיימים[9].

לצד זאת, בפסק דין מוכר שעסק בתחום, בעניין ערעורו של State v. Loomis על פסיקה שהסתמכה בין השאר על ציון שהופק ממערכת COMPAS קבע בית המשפט העליון של ויסקונסין כי השימוש בהערכת הסיכון של המערכת בנסיבות המקרה לא הפר את זכותו של הנאשם להליך הוגן, אך הדגיש כי אין להשתמש בציון כדי לקבוע אם הנאשם ייאסר או מה תהיה חומרת העונש, וכי ההסתמכות על הכלי חייבת להיות מלווה באזהרות ברורות בדבר אופיו הקנייני, הישענותו על נתונים קבוצתיים, היעדר תיקוף מספק לאוכלוסייה המקומית והחשש מהטיה כלפי קבוצות מיעוט. פסק דין זה ממחיש כי גם כאשר בית משפט אינו פוסל את עצם השימוש בכלי אלגוריתמי, אין פירוש הדבר הכשרה מלאה ובלתי מסויגת שלו[10].

בכך מבטא פסק הדין עמדה מורכבת יותר: לא של פסילה מוחלטת, אך גם לא של קבלה חסרת תנאים. דווקא המתח הזה, בין הסתייעות טכנולוגית לבין שימור שיקול דעת אנושי ואחריות משפטית, הוא שעומד בלב שאלת הפיקוח על מערכות AI ציבוריות. מן המקרים הללו עולה דפוס משותף: בכל אחד מהם, הדיון במערכת האלגוריתמית לא התמקד רק בדיוק הטכני שלה, אלא גם בשאלות רחבות יותר של שקיפות, זכויות, אחריותיות ומידת ההסתמכות הראויה על פלט ממוכן. במובן זה, הדוגמאות אינן רק המחשה נקודתית לסיכונים, אלא אינדיקציה לכך שהשאלה המרכזית איננה אם המדינה רשאית להשתמש בכלים המבוססים על בינה מלאכותית, אלא באילו תנאים מוסדיים, דיוניים ונורמטיביים שימוש כזה יכול להיחשב לגיטימי.

מסגרת נורמטיבית: הדין הקיים ופערי הרגולציה

הדיון בשימוש בבינה מלאכותית בגופים ציבוריים בישראל אינו מתנהל במצב של ריק נורמטיבי מוחלט. בשנים האחרונות החלה להתעצב בישראל תפיסת מדיניות ממשלתית ביחס לאסדרת התחום, אשר אינה מבוססת, לפחות בשלב זה, על רגולציה רוחבית אחידה בסגנון האירופי, אלא על גישה סקטוריאלית, הדרגתית ומבוססת סיכון.

גישה זו מבקשת להטיל על כל רגולטור ענפי את האחריות למפות את השימושים במערכות בינה מלאכותית בתחום שעליו הוא מופקד, לזהות את הסיכונים המתעוררים ולבחון את המענים הרגולטוריים המתאימים להם[11].

בכך מבקשת המדיניות הישראלית לאזן בין הרצון שלא לבלום חדשנות טכנולוגית באמצעות חקיקה רוחבית מוקדמת מדי, לבין ההכרה בכך ששימוש בבינה מלאכותית עלול לעורר סיכונים מהותיים לזכויות יסוד, לשוויון, לפרטיות, לאחריותיות ולשקיפות. נקודת המוצא המוסדית לגישה זו קיבלה ביטוי ברור במסמך "עקרונות, מדיניות,

8 Stichting Privacy First v. State of the Netherlands, ECLI:NL:RBDHA:2020:1878 (Dist. Ct. The Hague Feb. 5, 2020).

9 Jeff Larson et al., *How We Analyzed the COMPAS Recidivism Algorithm,* ProPublica (23.5.2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

10 State v. Loomis, 881 N.W.2d 749, 769 (Wis. 2016).

11 משרד החדשנות, המדע והטכנולוגיה ומחלקת ייעוץ וחקיקה (משפט כלכלי) **עקרונות, מדיניות, רגולציה ואתיקה בתחום הבינה המלאכותית 2023** (2023), https://www.gov.il/he/pages/ai_23.

רגולציה ואתיקה בתחום הבינה המלאכותית 2023", שפרסמו משרד החדשנות, המדע והטכנולוגיה ומחלקת ייעוץ וחקיקה (משפט כלכלי) במשרד המשפטים.

המסמך קובע במפורש כי בשלב זה יש לפעול במסגרת רגולציה ענפית ולא במסגרת רגולציה רוחבית, וכי על הרגולטורים לפעול, ככל האפשר, בהתאמה לנעשה במדינות מפותחות ובארגונים בין-לאומיים ובהתבסס על ניהול סיכונים והתאמת הרגולציה לרמת הסיכון[12]. לצד זאת, המסמך ממליץ על שימוש בכלי רגולציה גמישים ומדורגים כגון עקרונות אתיים, תקינה, רגולציה עצמית, מודולריות, פיילוטים וניסויים מבוקרים וכן על פיתוח רגולציה בשיתוף מומחים, גורמים בתעשייה, אנשי אקדמיה והציבור.

ואכן, ישראל לא אימצה בשלב זה מודל של חוק רוחבי אחד בעניין הבינה המלאכותית, אלא מודל של רגולציה מבוזרת, סקטוריאלית ומתואמת. המסמך האמור אינו מסתפק בקביעה מוסדית בדבר אופי הרגולציה, אלא גם משרטט את ציר הסיכונים והעקרונות שאמורים להנחות את הרגולטורים בפעולתם.

בין האתגרים המנויים בו נזכרים, בין היתר, אפליה, מעורבות אנושית, הסברתיות, גילוי, אמינות, עמידות, אבטחה, בטיחות, אחריותיות ופרטיות. כן מודגש בו כי כל רגולטור נדרש לזיהוי ולמיפוי של השימושים בבינה מלאכותית בגופים מפוקחים ואחרים בתחום שעליו הוא אמון, של הסיכונים הכרוכים בכך ושל דרכי ההתמודדות האפשריות. חשיבות הדברים לעניינו כפולה: ראשית, הם מלמדים כי קיימת בישראל תפיסת מדיניות המזהה את הסיכונים הכרוכים בשימוש ב-AI; שנית, הם מלמדים כי המדינה עצמה בחרה, לפחות לעת הזו, להפקיד את עיקר מלאכת הרגולציה והפיקוח בידי הרגולטורים הענפיים.

גישה זו קיבלה ביטוי יישומי מאוחר יותר גם במהלכים ענפיים קונקרטיים. כך למשל, בדוח הסופי שפורסם בדצמבר 2025 בעניין השימוש בבינה מלאכותית בסקטור הפיננסי, הוצגה עבודתו של צוות בין-משרדי שכלל את רשות שוק ההון, הפיקוח על הבנקים, הרשות לניירות ערך, רשות התחרות, משרד המשפטים ומשרד האוצר[13].

עצם הקמתו של צוות כזה, בהרכב בין-רגולטורי ובשביל סקטור מוגדר, ממחישה היטב כי הרגולציה הישראלית המתהווה מבוססת על תפיסה של טיפול ענפי מתואם, ולא על רגולציה רוחבית אחת החלה באופן אחיד על כל המשק.

הדוגמה הפיננסית אף מחזקת את ההבנה כי המדינה רואה ב-AI נושא המחייב מאמץ מערכתי, אך כזה שמתורגם בפועל לזירה הסקטוריאלית. עם זאת, העובדה שקיימת בישראל תפיסת מדיניות סקטוריאלית אינה מבטלת את קיומם של פערים רגולטוריים של ממש.

ההפך הוא הנכון. דווקא משום שהמדינה בחרה במודל מבוזר, שבו על כל רגולטור לעצב את המענה בתחום שעליו הוא מופקד, מתחדדת השאלה כיצד יובטח היקף ההגנה על זכויות באופן אחיד ואם קיימים מנגנוני פיקוח אפקטיביים שיכולים לזהות פערים, אי-עקביות או חוסר ממשי בהסדרה. במילים אחרות, האתגר הישראלי אינו רק היעדרו של חוק רוחבי, אלא גם הסיכון שפיצול רגולטורי ייצור אי-אחידות, פערי הגנה או עיכוב בטיפול בתחומים שבהם הרגולטור הענפי טרם גיבש מסגרת מספקת. על רקע זה מתחדדת חשיבות ההסתמכות על עקרונות הדין הישראלי הכללי, ובפרט דיני הפרטיות, חופש המידע והמשפט המינהלי. דינים אלה לא נועדו במקורם להסדיר שימוש בבינה מלאכותית, אך הם מספקים מסגרת נורמטיבית בסיסית החלה גם כאשר גוף ציבורי נשען על מערכות אלגוריתמיות.

חוק הגנת הפרטיות, התשמ"א-1981, מעגן את ההגנה על פרטיות ועל מידע אישי והוא נקודת מוצא מרכזית לכל בחינה של עיבוד מידע בידי רשויות[14]. עם זאת, החוק אינו כולל כיום הסדר מפורש בדבר החלטות המבוססות על עיבוד אוטומטי ואינו יוצר מנגנון דומה להסדר האירופי בתחום זה.

חוק חופש המידע, התשנ"ח-1998, עשוי לשמש בסיס עקרוני לדרישת גילוי גם ביחס לשימוש במערכות המשפיעות על הפרט. עם זאת, כאשר מתבקש גילוי מפורט יותר בדבר אופן פעולת המערכת, הנתונים שעליהם היא נשענת או מאפייני המודל שבו משתמשים, עשויות להתעורר טענות מצד הרשות או הספק בדבר סוד מסחרי, קניין רוחני או מגבלות אחרות על חשיפת המידע[15]. במקביל, עקרונות היסוד של המשפט המינהלי, ובהם חובת ההנמקה, החובה להפעיל שיקול דעת עצמאי, זכות הטיעון ואיסור השרירות, מוסיפים לחול גם כאשר הרשות נעזרת בכלים טכנולוגיים.

ואולם, הפסיקה הישראלית טרם פיתחה דוקטרינה ייעודית וסדורה ביחס להיקף ההסבר, הגילוי והבקרה הנדרשים כאשר החלטה מינהלית נשענת על מערכת AI.

---

12 שם.

13 רשות שוק ההון, ביטוח וחיסכון **בינה מלאכותית בסקטור הפיננסי, דוח סופי** (2025).

14 חוק הגנת הפרטיות, התשמ"א-1981, ס"ח 128.

15 חוק חופש המידע, התשנ"ח-1998, ס"ח 227.

לשם השוואה, הדין האירופי מציע כיום מסגרת מפורטת ומתקדמת יותר. הרגולציה האירופית להגנת המידע (General Data Protection Regulation; GDPR) כוללת הגנות ביחס להחלטות המבוססות אך ורק על עיבוד אוטומטי, לצד חובות שקיפות ואמצעי הגנה דיוניים[16], ואילו חוק הבינה המלאכותית של האיחוד האירופי (EU AI Act) מאמץ משטר רגולטורי מבוסס סיכון הקובע דרישות מחמירות ביחס לשימוש בבינה מלאכותית בתחומים מסוימים המוגדרים כבעלי סיכון גבוה, כגון תעסוקה, חינוך, שירותים ציבוריים חיוניים, אכיפת חוק, הגירה וניהול הליכים שיפוטיים או מינהליים[17].

ההשוואה לדין האירופי לא נועדה בהכרח להכתיב אימוץ מלא של המודל האירופי בישראל, אך היא ממחישה עד כמה ההסדרה הישראלית נותרה בשלב זה חלקית, מבוזרת ותלויה ביכולתם של רגולטורים ענפיים לפתח מענה עצמאי. מכאן נגזרת גם חשיבותה האפשרית של ביקורת המדינה. במצב שבו הרגולציה הישראלית מבוססת על חלוקה סקטוריאלית, על עקרונות מדיניות כלליים ועל דינים כלליים שאינם ייעודיים לשימוש בבינה מלאכותית, גובר הצורך בגוף ביקורת בעל ראייה מערכתית, המסוגל לבחון לא רק אם כל רגולטור פועל בתחומו, אלא גם אם המבנה הכולל יוצר הגנה מספקת, עקבית ואפקטיבית על הציבור. תפקידה של ביקורת המדינה בהקשר זה אינו להחליף את הרגולטורים הענפיים, אלא לבדוק אם המדיניות שהותוותה אכן מתורגמת הלכה למעשה למנגנוני פיקוח, תיעוד, בקרה והגנה על זכויות.

שישה עקרונות לבחינה מוסדית

הפרק הנוכחי מבקש לעבור מן המישור הנורמטיבי אל המישור היישומי. הפרקים הקודמים עסקו בשאלה מדוע השימוש הציבורי בבינה מלאכותית מעורר קושי משפטי ומוסדי, ופרק זה ידון בשאלה מה צריך מבקר המדינה לבדוק כאשר גוף ציבורי משתמש במערכת AI. התשובה המוצעת אינה נשענת על עיקרון אחד בודד, אלא על מסגרת מוסדית משולבת, המיועדת לבחינת מערכות בינה מלאכותית לאורך מחזור חייהן המלא - משלב הרכש וההטמעה, דרך שלב ההפעלה השוטפת ועד לשלב העדכון, ההשעיה או הפסקת השימוש.

המסגרת המוצעת כאן שואבת השראה משלושה מקורות מרכזיים: מסגרת ניהול הסיכונים של המכון הלאומי לתקנים ולטכנולוגיה של ארצות הברית (National Institute of Standards and Technology; NIST) המציעה גישה מחזורית ורב-שלבית לניהול סיכוני AI; הדין האירופי, ובפרט הוראות ה-EU AI Act ביחס למערכות בסיכון גבוה; והספרות והפסיקה שעמדו על הסכנות שבאטימות, בהטיה, בפיקוח פורמלי בלבד ובפגיעה בזכויות דיוניות.

מסגרת NIST AI RMF 1.0 מדגישה כי ניהול סיכוני השימוש בבינה מלאכותית אינו פעולה חד-פעמית אלא תהליך מתמשך הנשען על ארבע פונקציות ליבה: ממשל ברור, מיפוי שימושים וסיכונים, מדידה והערכה של השפעות וניהול שוטף של הסיכון לאורך חיי המערכת (Govern, Map, Measure, Manage)[18]. מסגרת זו אומנם אינה מחייבת משפטית בישראל, אך היא מספקת היגיון ארגוני שימושי במיוחד לביקורת מוסדית.

ה-EU AI Act דורש כי בעת שימוש במערכות בסיכון גבוה יקפידו על תיעוד, ניהול סיכונים, לוגים, פיקוח אנושי, דיוק, עמידות ואבטחה לאורך כל מחזור החיים של המערכת.

להלן שישה עקרונות מוצעים לבחינה מוסדית. אין לראות בהם רשימת בדיקה טכנית בלבד, אלא מערכת משולבת של תנאים הכרחיים לשימוש אחראי בבינה מלאכותית במרחב הציבורי.

עיקרון 1: שקיפות ותיעוד

נקודת המוצא לביקורת אפקטיבית היא קיומו של תיעוד. גוף ציבורי אינו יכול לעשות שימוש אחראי במערכת בינה מלאכותית אם אינו מחזיק בתיעוד מספק של מאפייניה, ייעודה, מגבלותיה, מקורות הנתונים שעליהם היא נשענת, זהות הספק, גרסת המודל, אופן העדכון, תוצאות הביצועים הידועות והגורם המוסדי האחראי להפעלה.

בהיעדר תיעוד כזה, לא ניתן לקיים פיקוח אמיתי, לא ניתן לעקוב אחר שינויים בגרסאות ולא ניתן לייחס אחריות כאשר מתגלה כשל. למעשה נדרש Model Card - מסמך תיאורי שיטתי של המערכת. אין הכרח לאמץ דווקא את המונח הזה, אך חשוב לאמץ את ההיגיון: כל מערכת צריכה להיות ניתנת לזיהוי, להסבר מוסדי בסיסי ולהערכה בדיעבד. על כן, יש לשאול לא רק אם המערכת קיימת, אלא אם קיימת עבורה תשתית תיעודית נאותה.

16 Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation), arts. 13, 14, 15 & 22, OJ L 119/1 (4.5.2016).

17 Regulation (EU) 2024/1689 , לעיל ה"ש 4, סעיף 6.

18 National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (2023), chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf .

הוראות ה-EU AI Act כוללות דרישות דומות לתיעוד טכני, למידע למפעילים, ללוגים ולשמירת תיעוד עבור הרשויות המוסמכות[19]. עם זאת, לצורך ביקורת מדינה אין הכרח להמתין להסדר ישראלי זהה. די להבין שתיעוד הוא תנאי ראשוני לכל פיקוח. מובן כי תיתכן שאלת ביקורת מעשית: האם לכל מערכת AI פעילה בגוף המבוקר יש מסמך תיאורי עדכני הכולל את מטרת השימוש, הגרסה הפעילה, זהות הספק, סוגי הנתונים הרלוונטיים, ביצועים ידועים, מגבלות ידועות והגורם האחראי בארגון.

עיקרון 2: בקרה אנושית ממשית

המונח "פיקוח אנושי" (Human oversight) הפך לביטוי כמעט שגרתי בכל מסמך רגולטורי העוסק בבינה מלאכותית, אך לא כל פיקוח אנושי הוא פיקוח ממשי. לא פעם, תהליך קבלת ההחלטות מציג כלפי חוץ מעורבות של פקיד או בעל סמכות, אך בפועל האדם רק מאשר את פלט המערכת בלי להבין לעומק את משמעותו, בלי לבחון אפשרות לסטות ממנו ובלי לתעד שיקול דעת עצמאי.

במצב כזה, הנוכחות האנושית היא פורמלית בלבד, ואין בה כדי לספק את ההגנה המוסדית הנדרשת. סיטרון (Citron) עמדה כבר לפני שנים על הסכנה שבמנגנוני procedural safeguards הנחזים כלפי חוץ כהגנה דיונית מספקת, אך בפועל אינם אלא טקס ריק של מקבלי החלטות הנוטים לייחס למערכת הממוחשבת אמינות יתר או שאינם מצוידים בכלים המאפשרים להבין את הפלט שלה[20].

ה-EU AI Act מתייחס לכך במפורש ודורש שתכנונן של מערכות בסיכון גבוה יאפשר פיקוח אנושי אפקטיבי, שמטרתו לצמצם או למנוע סיכונים לבריאות, לבטיחות ולזכויות יסוד. המשמעות בעניין ביקורת המדינה ברורה: אין די בכך שהחלטה "עוברת דרך אדם". יש לבחון אם האדם באמת מסוגל להפעיל שיקול דעת עצמאי, להבין את ההמלצה, להטיל בה ספק ולחרוג ממנה בעת הצורך.

עיקרון 3: הערכת השפעת המערכת לפני הטמעתה

אחד הכשלים המרכזיים בהטמעת מערכות AI במגזר הציבורי הוא המעבר המהיר מדי משלב ההתלהבות הטכנולוגית לשלב ההפעלה בפועל, בלי שנערכה בחינה מוקדמת ומסודרת של הסיכונים. כאן נדרשת הערכה של השפעת המערכת טרם הפעלתה. יש להעריך את ההשפעה על הפרטיות וכן על זכויות, שוויון, שקיפות, תלות טכנולוגית, אחריותיות וסיכוני הטיות.

המשפט הקיים כבר מכיר אומנם בצורך בכלים כמו DPIA (Data Protection Impact Assessment) בעת עיבוד הכרוך בסיכון גבוה, ויש לכך התייחסות ב-GDPR, אך הכוונה היא בעיקר למצבים שבהם העיבוד עשוי להשפיע על זכויות וחירויות, ובפרט כאשר מדובר בהערכה שיטתית ואוטומטית של מאפיינים אישיים, בקבלת החלטות בעלות השפעה ניכרת או בעיבוד רחב היקף של מידע רגיש. הערכת ההשפעה של AI במגזר הציבורי צריכה להיות רחבה יותר[21]. יש לשאול מי עלול להיפגע, אילו קבוצות מושפעות במיוחד, האם קיימת חלופה פחות פוגענית, מהי רמת ההסתמכות המתוכננת על המערכת, מהן ההשלכות של כשל אפשרי, מי אישר את השימוש במערכת ואיזו תשתית תיעודית היא תכלול.

עיקרון 4: אחריותיות ומנגנוני השגה

החלטה מינהלית הנשענת על מערכת אלגוריתמית אינה חדלה להיות החלטה מינהלית. משום כך, עליה לעמוד בדרישות היסוד של אחריותיות: יש לוודא כי יהיה ניתן להבין מה עמד בבסיסה, לזהות מי נושא באחריות לה ולאפשר לפרט שנפגע ממנה לבקש בחינה חוזרת או להשיג השגות. אם ההחלטה מוצגת כקביעה של מחשב, ההנמקה נפגעת ושרשרת האחריות מתערפלת.

לעיתים הרשות מפנה לספק, הספק מפנה למודל, והפרט נותר ללא כתובת ברורה. דווקא כאן יש חשיבות מיוחדת לביקורת מוסדית. מבקר המדינה צריך לבחון אם קיימים מסלולי השגה ברורים, אם יש גורם אנושי מוסמך האחראי להחלטה, אם ניתן לשחזר את הנתונים והשלבים שהובילו לתוצאה ואם האחריות אינה מתפזרת על פני שרשרת האספקה באופן המסכל ביקורת. דרישה זו נובעת הן מעקרונות המשפט המינהלי והן מהדיון הרחב יותר על אחריותיות בשרשראות טכנולוגיות מורכבות.

עיקרון 5: הפחתת הטיה ובחינת נתוני אימון

העובדה שמערכת AI אינה "רואה" משתנה של גזע, מגדר או שיוך קבוצתי באופן מפורש, אינה מבטיחה היעדר אפליה. זהו אולי אחד הלקחים המרכזיים שעלו מן הספרות ומהמקרים האמפיריים הבולטים. מערכות חיזוי ולמידת מכונה עלולות לשעתק דפוסים היסטוריים של אי-שוויון גם כאשר הן נשענות על משתנים הנחזים כניטרליים. משום כך, בחינת הטיה אינה יכולה להסתפק בהצהרה שהאלגוריתם "עיוור" למאפיין מוגן. נדרשת בדיקה שיטתית

19 Regulation (EU) 2024/1689, לעיל ה"ש 4, סעיפים 9 - 15, 18.

20 Citron, Technological Due Process, לעיל ה"ש 6, 149, 1287 - 1295.

21 Regulation (EU) 2016/679 , לעיל ה"ש 16, סעיף 35.

של ייצוג הנתונים, איכותם, מקורותיהם וההשפעה בפועל של המערכת על קבוצות שונות.

בכלל זה יש לבחון אם נתוני האימון משקפים באופן מאוזן את האוכלוסייה הרלוונטית, אם הם כוללים קבוצות שאינן מיוצגות די הצורך, ואם הם נושאים עימם הטיות היסטוריות או מוסדיות שהמערכת עלולה לשעתק. בחינה זו חשובה במיוחד כאשר המודל אומן על בסיס נתוני עבר שנוצרו בתוך מערכות ציבוריות שכבר פעלו בתנאים של אי-שוויון או נגישות לא אחידה לשירותים.

תחקיר ProPublica על COMPAS ממחיש זאת בבירור[22]. אף שאין להתייחס אליו כאל פסק דין או אמת מדעית סופית, התחקיר מראה כיצד פערים בשיעורי הטעות בין קבוצות אוכלוסייה שונות עשויים להופיע גם במערכת שמתיימרת להיות מכשיר ניטרלי.

22 Julia Angwin et al., *Machine Bias*, ProPublica (23.5.2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

מבקר המדינה אינו צריך להפוך למעבדת מחקר אלגוריתמי, אך עליו לדרוש מן הגוף המבוקר להראות אם נבדקה הטיה, באילו שיטות, באיזו תדירות ומה נעשה כאשר מתגלים פערים.

עיקרון 6: אבטחה, עמידות והיערכות לכשל

מערכות AI ציבוריות חשופות לשאלות של הוגנות ושקיפות, אמינות, אבטחה, עמידות ותפקוד תחת תנאי שינוי. AI Act דורש ממערכות בסיכון גבוה רמה מתאימה של דיוק, חוסן ואבטחת סייבר, ומסגרת NIST מדגישה כי ניהול הסיכונים צריך להיות רציף ומחזורי וכי עליו להתעדכן לאורך חיי המערכת.

על מוסדות המשתמשים במערכות AI לבחון לא רק אם המערכת פעלה היטב ביום ההשקה, אלא גם כיצד היא מתמודדת עם סטיית מודל (drift), עם שינויים באוכלוסיית המשתמשים, עם עדכוני תוכנה, עם תקיפות סייבר או עם כשל של ספק. בהקשר זה עולה גם שאלת התלות בספק (vendor lock-in).

אם גוף ציבורי רוכש מערכת שאינו יכול לבדוק, שאינו יכול להחליף בקלות או שאינו מחזיק בגישה מספקת לנתוניה, ללוגים שלה או לתיעוד הרלוונטי, הוא עלול למצוא עצמו תלוי בספק באופן הפוגע בפיקוח, בבקרה ואף ברציפות השירות.

על פי ה-AI Act, בעת שימוש במערכות AI המוגדרות כמערכות בסיכון גבוה, קיימות דרישות ברורות ביחס ללוגים ולתיעוד מסמכים ושמירתם. גם ללא העתקה ישירה למשפט הישראלי, ניתן לגזור מה- AI Act אמת מידה מוסדית חשובה: שימוש אחראי ב-AI מחייב היערכות לכשל, לגיבוי תהליכים אנושיים ולניהול תלות בספק באופן שאינו מסכל ביקורת.

ששת העקרונות אינם רשימה אקראית של דרישות נפרדות. הם שלובים זה בזה: תיעוד הוא תנאי לפיקוח אנושי אפקטיבי, פיקוח אנושי הוא תנאי לאחריותיות, אחריותיות תלויה באפשרות להשיג על החלטה, בחינת הטיה תלויה באיכות התיעוד והמדידה ואבטחה ועמידות משלימות את כל אלה בכך שהן מוודאות שהמערכת אינה רק "הוגנת" בתיאוריה, אלא גם נשלטת, נבדקת ומבוקרת בפועל לאורך זמן. ביקורת מדינה המבקשת להתמודד ברצינות עם AI ציבורי תצטרך אפוא לבחון את מכלול העקרונות יחד ולא להסתפק בקיום פורמלי של אחד מהם.

תפקיד מבקר המדינה: לקראת ביקורת מותאמת AI

ששת העקרונות שהוצגו בפרק הקודם אינם יכולים להישאר בגדר מסגרת נורמטיבית מופשטת בלבד. כדי שיהפכו לאמות מידה אפקטיביות, עליהם להיות מתורגמים לכלים מוסדיים, מתודולוגיים וארגוניים המותאמים לאופי הייחודי של מערכות בינה מלאכותית.

כאן בדיוק מתחדדת שאלת תפקידו של מבקר המדינה, אשר בוחן חוקיות, תקינות, יעילות וחיסכון במובנם המסורתי, אך נדרש להתאים את ארגז הכלים הביקורתי שלו למציאות שבה החלטות ציבוריות מושפעות ממודלים, נתונים, ספקים חיצוניים, מערכות לומדות ותשתיות דיגיטליות מורכבות.

סמכותו הרחבה של מוסד מבקר המדינה ואחריותו כלפי הכנסת ולא כלפי הממשלה מציבות אותו בעמדה מוסדית ייחודית לבחינה מסוג זה[23]. עם זאת, עצם קיומה של סמכות רחבה אינו מבטיח כשלעצמו יכולת אפקטיבית לפעול בתחום חדש זה. במובן זה, האתגר אינו נובע מהיעדר סמכות פורמלית, אלא מהפער שבין כלי הביקורת המסורתיים לבין האובייקט החדש שיש לבדוק.

המבקר יודע היטב לשאול אם הליך רכש בוצע כדין, אם נשמרה הפרדה בין סמכויות, אם קיימים נהלים פנימיים ואם הוקצו תקציבים בהתאם להחלטות המוסמכות. אך כאשר מדובר במערכת AI, שאלות הביקורת צריכות להתרחב באופן ניכר: האם בוצעה הערכת סיכון טרם ההטמעה, האם נבחנו פערי הטיה בין קבוצות, האם קיימת בקרה אנושית ממשית שאינה אישור פורמלי בלבד, האם נשמרים לוגים ותיעוד טכני, האם קיימת תלות בספק יחיד והאם קיימים מסלולי השגה אפקטיביים לאזרח שנפגע מהחלטה הנתמכת במערכת.

מסגרות כמו NIST AI RMF 1.0, המדגישות ממשל, מיפוי, מדידה וניהול מתמשך של סיכוני AI, ממחישות היטב את הצורך לעבור מביקורת נקודתית לביקורת מחזורית ומבוססת סיכון[24]. דוח של משרד הביקורת הלאומי של בריטניה מ-2024 על השימוש ב-AI בממשלה מחזק אף הוא את ההבנה שהצלחה או כישלון בתחום זה תלויים לא רק בטכנולוגיה עצמה, אלא גם בשאלות כגון מי אמון על המערכת בארגון, כיצד נבדקת איכות הנתונים שעליהם היא נשענת, האם לעובדים המשתמשים בה יש ידע והכשרה מספקים והאם מתקיים תיאום בין הגורמים המינהליים,

23 חוק-יסוד: מבקר המדינה, סעיפים 2 ו-6.

24 NIST, *NIST*; 18 לעיל ה"ש, National Institute of Standards and Technology, *Artificial Intelligence Risk AI RMF Playbook* (2023) https://www.nist.gov/itl/ai-risk-management-framework/nist-ai-rmf-playbook.

המשפטיים והטכנולוגיים המעורבים בהפעלתה[25].

על רקע זה, ניתן להציע חמש המלצות מוסדיות מרכזיות:

ההמלצה הראשונה היא יצירת מסגרת בדיקה שיטתית ומובנית שתאפשר להפוך את עקרונות הביקורת על AI לכלי עבודה מעשי. נכון שמבקר המדינה יפתַח שאלון ביקורת ייעודי לשימוש בביקורות של גופים ציבוריים המשתמשים במערכות AI.

שאלון כזה לא צריך להיות מסמך טכני סגור וקשיח, אלא מסגרת מודולרית שתתאים את עצמה לאופי השימוש. נדרשת למשל הבחנה בין מערכות תומכות החלטה פנימיות, מערכות דירוג סיכון, מערכות המכוונות הקצאת משאבים או מערכות המתקשרות ישירות עם הציבור.

חשיבותו של שאלון כזה כפולה: ראשית, הוא מספק למבקרים שפה משותפת וסדרה עקבית של שאלות, ובכך מצמצם את הסיכון לביקורת חלקית, מקרית או אינטואיטיבית מדי. שנית, פרסומו עשוי לייצר אפקט מניעתי חשוב. גופים ציבוריים ידעו מראש מה צפוי להיבחן ויוכלו להיערך בהתאם באמצעות תיעוד מסודר, ניהול סיכונים והטמעת בקרה פנימית טובה יותר.

במובן זה, השאלון אינו רק מכשיר ביקורת בדיעבד, אלא גם כלי ממשל מקדים. מסגרות כמו NIST AI RMF Playbook וממצאי ה-NAO בדבר הצורך בתפקידי בעלות, אחריות ותיאום בממשלה מספקות בסיס טוב לפיתוח כלי כזה, גם אם נדרשת התאמה שלהן להקשר הישראלי.

ההמלצה השנייה נוגעת לפיתוח מומחיות טכנית פנימית. ביקורת אפקטיבית על מערכות AI מחייבת הבנה טכנולוגית מספקת. אין הכוונה לכך שכל מנהל ביקורת יהפוך למדען נתונים או למהנדס למידת מכונה, אלא לכך שמוסד הביקורת יפתח מוקד מומחיות פנימי - יחידה ייעודית או צוות רוחבי תומך - שיסייע לצוותי הביקורת הסקטוריאליים בבדיקות הנוגעות לאלגוריתמיקה, נתונים, מודלים, הטיה, אבטחה ותיעוד טכני.

ה-NAO מדגים כי מוסד ביקורת המדינה יכול להישען על תשתיות ניתוח נתונים ושירותי דאטה כדי לייעל עבודת ביקורת, לחבר מאגרי מידע, לזהות דפוסים ולבסס ממצאים באופן אמפירי יותר[26]. גם אם לא מדובר בדיוק ביחידת AI ייעודית במובן המצומצם, המודל הבריטי בהחלט תומך בהיגיון של שילוב יכולות אנליטיות וטכנולוגיות בעבודת הביקורת. המשמעות היישומית לישראל היא ברורה. בהיעדר מומחיות טכנולוגית מספקת, ביקורת AI עלולה להצטמצם לבדיקה פורמלית של נהלים, בלי לחדור אל לב הבעיה. גוף מבוקר עשוי להציג "מדיניות AI", "תיעוד", "אישור אנושי" או "בדיקות תקופתיות", אך רק צוות בעל הבנה מספקת יוכל לבחון אם מדובר במנגנון ממשי או בתחליף סמלי.

השקעה ביחידה מומחית כזו עשויה להיות יקרה בטווח הקצר, אך תשואתה הציבורית עשויה להיות גבוהה מאוד, משום שכשל מערכתי אחד במערכת AI ציבורית עשוי להשפיע על אוכלוסייה רחבה ביותר.

ההמלצה השלישית נוגעת לצורך בשיתוף פעולה מוסדי. השימוש הציבורי בבינה מלאכותית יושב על התפר שבין כמה תחומים רגולטוריים: פרטיות, הגנת מידע, שוויון, הליך מינהלי תקין, רגולציה ענפית, אבטחת מידע ועוד. משום כך, ביקורת המדינה אינה יכולה ואינה צריכה לפעול כאילו מדובר בטריטוריה מוסדית השייכת לה בלבד. בפרט, קיימת חפיפה מסוימת, גם אם לא זהות, בין עולמה של הרשות להגנת הפרטיות לבין תחום הביקורת על שימוש ציבורי ב-AI.

הרשות עוסקת בעיבוד מידע אישי ובהיבטי פרטיות, ואילו מבקר המדינה בוחן תקינות מוסדית רחבה יותר. דווקא בשל כך, קיים ערך משמעותי לשיתופי פעולה ממוסדים, לפחות ברמת החלפת ידע, מתודולוגיות, מוקדי סיכון ודפוסי אכיפה.

תפקידו של מבקר המדינה בהקשר זה אינו לבטל את המבנה הענפי, אלא לבחון אם הוא אכן פועל, אם קיימת בו עקביות ואם נוצרו פערים בין רגולטורים שונים או בין תחומים שונים של השירות הציבורי.

ההמלצה הרביעית עניינה לעודד היערכות מוקדמת של גופים ציבוריים באמצעות פרסום מסמך מנחה פומבי. הביקורת על AI אינה צריכה להתחיל רק לאחר שכבר נגרם נזק. לצד שאלון הביקורת הפנימי של מבקר המדינה, ניתן לגבש מסמך מנחה שאינו חקיקה ואינו רגולציה מחייבת במובן הפורמלי, אך יוכל לסייע לגופים ציבוריים להבין מראש מהן הציפיות הבסיסיות ביחס לשימוש אחראי בבינה מלאכותית.

כשם שגופים ממשלתיים מסתמכים על מדריכים לניהול סיכונים, לבקרה פנימית, לממשל תאגידי או לניהול מידע,

25 UK National Audit Office, *Use of Artificial Intelligence in Government* (2024), https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government/.

26 National Audit Office, *How Data Analytics Can Help with Audits* (2020), https://www.nao.org.uk/insights/how-data-analytics-can-help-with-audits/; National Audit Office, *Using Data Analytics to Tackle Fraud and Error* (2025), https://www.nao.org.uk/reports/using-data-analytics-to-tackle-fraud-and-error.

כך ניתן לגבש מדריך עקרונות לשימוש אחראי ב-AI בגופים ציבוריים[27]. מדריך כזה לא יצטרך להכריע בכל שאלה משפטית מורכבת, אך כן יוכל לנסח ציפיות ברורות ביחס לתיעוד, להערכת סיכון, לבקרה אנושית, לשקיפות, לניהול ספקים, לניטור הטיות ולהיערכות לכשל.

הרעיון הזה נתמך גם בניסיון הממשלתי הרחב שנצבר בישראל בתחום ניהול הסיכונים ברגולציה ובמדיניות ציבורית. מדריכים רוחביים אינם יוצרים לבדם משטר משפטי מלא, אך הם מסייעים ביצירת אחידות מושגית, שפה מוסדית משותפת וסטנדרט מקצועי שהמבוקר והמבקר גם יחד יכולים להתייחס אליו. לכן, מדריך כזה יכול לשמש בעת ובעונה אחת גם ככלי הכוונה מוקדם לגופים ציבוריים וגם כאמת מידה ביקורתית בבדיקות עתידיות.

ההמלצה החמישית היא שילוב של למידה ושיתוף פעולה בין-לאומיים, אך בזהירות מושגית. ביקורת הבודקת שימוש בבינה מלאכותית יכולה להסתייע הן בדוחות קונקרטיים של מוסדות ביקורת עליונים והן בעקרונות הבין-לאומיים הרחבים של ניהול סיכונים, בקרה פנימית ואחריותיות ציבורית. אין להתעלם מן הערך של למידה ממוסדות ביקורת עליונים וממסגרות מקצועיות בין-לאומיות: ה-NAO הבריטי כבר עסק ישירות בשימוש ב-AI בממשלה; מוסדות ביקורת אחרים בעולם מפתחים בהדרגה כלים אנליטיים, יכולות של ניתוח נתונים (data analytics) ומסגרות ביקורת מותאמות לעולם דיגיטלי; ו-INTOSAI בכללותה מדגישה זה שנים את הקשר בין בקרה פנימית, ניהול סיכונים, אחריותיות וממשל תקין.

עם זאת, כאן בדיוק דרושה זהירות: לא כל מסמך בין-לאומי שמדבר על ממשל תקין (good governance) או ניהול סיכונים (risk management) הוא מסמך ייעודי העוסק ב-AI, ולא נכון "לצבוע" בדיעבד מסמכים כלליים כסטנדרטים ייעודיים לביקורת בתחום.

המסקנה העולה מן האמור היא שהמעבר לביקורת מותאמת AI אינו מצריך מהפכה מוסדית מוחלטת, אך הוא בהחלט מחייב שינוי משמעותי של הכלים, הכשירויות והמתודולוגיה. סמכות מבקר המדינה לבדה אינה מספיקה. נדרשת גם בנייה מודעת של תשתית ביקורתית חדשה, הכוללת שפה מקצועית עדכנית, מומחיות טכנולוגית, שיתוף פעולה בין-מוסדי, כלי הנחיה מוקדמים ויכולת ללמוד מן הנעשה בעולם בלי לאמץ בצורה לא ביקורתית מונחים או מסמכים שאינם מתאימים במדויק למציאות הישראלית.

סיכום

המסגרת שהוצעה במאמר זה אינה מבקשת להרתיע גופים ציבוריים מפני שימוש בבינה מלאכותית ואינה נשענת על עמדה ריאקציונית שלפיה עצם הטמעת הטכנולוגיה היא בגדר תקלה מוסדית. ההפך הוא הנכון. במקרים רבים, מערכות בינה מלאכותית עשויות לשפר את איכות השירות הציבורי, לייעל תהליכים אדמיניסטרטיביים, לצמצם זמני טיפול, לסייע בזיהוי צרכים שלא זוהו קודם לכן ולאפשר הקצאה מדויקת יותר של משאבים ציבוריים מוגבלים. דווקא משום שהפוטנציאל הציבורי של טכנולוגיות אלה הוא ממשי, השאלה המרכזית אינה אם להתיר את השימוש בהן, אלא באילו תנאים משפטיים, מוסדיים וביקורתיים ניתן להבטיח שהשימוש יהיה אחראי, הוגן ונתון לפיקוח אפקטיבי.

מנקודת מבט זו, הדיון בבינה מלאכותית בשירות הציבורי אינו דיון טכנולוגי גרידא. הוא נוגע בליבה של תפיסת השלטון החוקתי והמינהלי. כאשר החלטה על זכאות, על גישה לשירות, על הקצאת משאב, על סימון "סיכון" או אף על התערבות עמוקה בחיי הפרט נשענת על פלט אלגוריתמי, אין מדובר רק בשאלה של יעילות, אלא בשאלה של כוח שלטוני: מי מקבל את ההחלטה, על סמך אילו נתונים, באיזו מידה ניתן להבין אותה, מי נושא באחריות לה וכיצד יכול הפרט להתמודד עימה.

במובן זה, השאלה המרכזית אינה אם האלגוריתם "חכם" דיו, אלא אם מוסדות המדינה נותרו נאמנים לחובות היסוד שלהם גם כאשר הם פועלים באמצעות מערכות חישוביות מורכבות. חזרה אל דוגמת הרווחה הפותחת ממחישה את הדברים היטב. אם בוחנים את המקרה דרך פריזמת ששת העקרונות שהוצעו במאמר, מתגלה רצף של כשלים מצטברים: היעדר תיעוד מספק של המערכת, היעדר בקרה אנושית ממשית עליה, היעדר בחינה מוקדמת של השפעותיה על זכויות ועל קבוצות פגיעות, היעדר מנגנון ברור להשגה או לערר, היעדר בחינה שיטתית של הטיה בתוצאות והיעדר בחינה של תלות בספק ושל היערכות לכשל. רצף כזה אינו רק כשל טכנולוגי או ניהולי, הוא כשל מוסדי, וכאשר כשל מוסדי כזה אינו מזוהה, אינו נבדק ואינו מתוקן בזמן, הוא עלול להפוך במהירות לכשל חוקתי-מינהלי המשפיע על ציבור רחב.

כאן בדיוק מצוי תפקידו האפשרי, ואף המתבקש, של מבקר המדינה. מבקר אינו אמור להפוך לרגולטור AI במובן המלא, ואינו מחליף את המחוקק, את הממשלה או את הרגולטורים הענפיים. תפקידו שונה: לחשוף פערים, לזהות כשלים מבניים, להצביע על היעדר מנגנוני הגנה ולהפוך בעיות

27 משרד ראש הממשלה, **מדריך לניהול סיכונים ברגולציה ובמדיניות ציבורית** (2018), chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/ https://www.gov.il/BlobFolder/generalpage/reg-risk-management-guide/he/file_file_Guide171018.pdf.

טכנולוגיות שנוטות להישאר מוסתרות בתוך שיח מקצועי או קבלני לבעיות ציבוריות ומשפטיות גלויות. במציאות של רגולציה סקטוריאלית, מבוזרת ומתפתחת, חשיבותו של גוף ביקורת בעל ראייה רוחבית אף גוברת, משום שהוא עשוי לבחון לא רק את התנהלותו של כל רגולטור בנפרד, אלא גם אם המכלול אכן יוצר הגנה מספקת, עקבית ואפקטיבית על הציבור.

במובן זה, התרומה האפשרית של ביקורת המדינה חורגת משאלת האכיפה במובנה הצר. ביקורת טובה אינה רק מצביעה על הפרה - היא גם מייצרת שפה מוסדית. אם מוסד הביקורת יבחן באופן עקבי תיעוד, בקרה אנושית, הערכת השפעה מוקדמת, מנגנוני השגה, בדיקות הטיה, אבטחה, עמידות ותלות בספקים, הוא יתרום לא רק לחשיפת ליקויים, אלא גם לעיצוב סטנדרט ציבורי חדש של שימוש אחראי בבינה מלאכותית. בכך, הביקורת עשויה למלא תפקיד כפול: תפקיד רטרוספקטיבי של איתור מחדלים ותפקיד פרוספקטיבי של הכוונת התנהגות מוסדית עתידית.

מן העבר השני, אין לייפות את התמונה. ביקורת המדינה לבדה לא תספיק. האחריות המוסדית לפיתוח מסגרת משטרית ראויה לשימוש ציבורי בבינה מלאכותית מוטלת גם על המחוקק, על הממשלה, על הרגולטורים הענפיים ועל ספקי המערכות עצמם. המחוקק נדרש להכריע אם, ובאיזו מידה, נחוצה בישראל מסגרת חקיקתית אופקית או לכל הפחות עיגון מפורש של מנגנוני הגנה בסיסיים. הממשלה נדרשת להבטיח שהגישה הסקטוריאלית שכבר הותוותה לא תישאר הצהרת מדיניות בלבד. הרגולטורים הענפיים נדרשים למלא בפועל את תפקיד המיפוי, הניתוח והפיקוח שהמדיניות הטילה עליהם, וספקי מערכות AI נדרשים להבין כי מכירה למגזר הציבורי אינה יכולה להתבסס על דרישה לאמון עיוור במערכת אטומה. ואולם, דווקא משום שכל אלה הם תהליכים איטיים, ולעיתים גם מפוצלים ומורכבים, בטווח הקצר מבקר המדינה עשוי להיות הגורם המעשי ביותר ליצירת שינוי. הוא אינו צריך להמתין להשלמת חקיקה ואינו חייב להמתין להסכמה מלאה בשאלה איזה מודל רגולטורי יאמץ המשפט הישראלי בעתיד.

די בכך שיבחן אם מוסדות המדינה, כאן ועכשיו, מפעילים מערכות המשפיעות על זכויות הפרט בלי שהוקמו סביבן אמצעי הבקרה המתבקשים. במובן זה, ביקורת המדינה יכולה לשמש מנגנון ביניים הכרחי בין מצב של אימוץ טכנולוגי מואץ לבין הסדרה משפטית שעדיין נמצאת בתנועה. הספרות הביקורתית על ממשל אלגוריתמי עומדת שוב ושוב על כך שהסכנה המרכזית אינה עצם קיומם של מודלים מורכבים, אלא האפשרות שכוח ניהולי, מסווג, מדרג ומכריע יופעל מאחורי מעטה של מורכבות טכנית, סודיות מסחרית או סמכות מקצועית לכאורה.

לפי יובנקס (Eubanks), הביקורת מתמקדת באופן שבו מערכות דיגיטליות מציגות הכרעות פוליטיות ומבניות כשאלות של יעילות אדמיניסטרטיבית, וכך נוצרים שחיקה של הליך הוגן (due process), שיעתוק אי-שוויון והעמקת פערי כוח[28]. על פי פסקוואלה (Pasquale), הדגש הוא על הסכנה של "קופסה שחורה", כלומר מציאות שבה מוסדות חזקים בוחנים, מדרגים ומעריכים אחרים בלי להיחשף בעצמם לביקורת שקולה.[29]

שתי התובנות האלה רלוונטיות במיוחד למרחב הציבורי. הן מלמדות כי לצד הבדיקה אם הטכנולוגיה מדויקת, חשוב לבדוק אם הכוח שהיא מפעילה נותר כפוף לביקורת, לחשיפה ולאחריותיות. בסופו של דבר, שאלת הבינה המלאכותית בשירות הציבורי אינה שאלה של חדשנות בלבד ואף לא שאלה של פרטיות בלבד. זו שאלה של ארכיטקטורת כוח.

כאשר רשות ציבורית נעזרת במערכת אלגוריתמית כדי לדעת יותר, להחליט מהר יותר ולפעול בהיקף רחב יותר, מתעצמת גם החובה להצדיק, להסביר, לבקר ולתקן. חברה דמוקרטית אינה נבחנת רק ביכולתה לאמץ טכנולוגיה מתקדמת, אלא גם ביכולתה להבטיח שהטכנולוגיה לא תשחרר את השלטון מן החובות הבסיסיות המוטלות עליו כלפי האזרח. אם בעבר ביקורת המדינה נועדה להגן על הציבור מפני שרירותיות שלטונית גלויה, הרי שבעידן הנוכחי עליה להסתגל גם לצורות חדשות, שקטות ומתוחכמות יותר של שרירותיות, הפועלות באמצעות נתונים, מודלים וציון סיכון.

לכן, הצעדים שהוצעו במאמר זה הם אינם "מהפכה" אלא תנאי סביר לשמירה על יסודות הממשל התקין בעידן אלגוריתמי, ומפני שהם צעדים סבירים קשה להצדיק את דחייתם. בעידן שבו קצב ההטמעה של מערכות AI מהיר בהרבה מקצב התגבשותם של כללי בקרה, האיפוק הביקורתי אינו ניטרליות אלא ויתור בפועל על יכולת ההגנה של המשפט הציבורי. אם מבקר המדינה יבקש למלא את תפקידו ההיסטורי גם במרחב הדיגיטלי, לא די שיבחן את תוצאות המדיניות, עליו לבחון גם את התשתיות החישוביות שמעצבות אותה.

---

28 Eubanks, Automating Inequality, לעיל ה"ש 3, במיוחד הדיון בהליך הוגן, בשיעתוק אי-שוויון ו"בבית המחסה הדיגיטלי לעניים" (Digital poorhouse).

29 Pasquale, The Black Box Society, לעיל ה"ש 5, דיון בכוחן של מערכות "הקופסה השחורה", בחוסר השקיפות שלהן ובצורך בפיקוח משמעותי על מערכות דירוג, ניקוד וקבלת החלטות.