מיכאל בירנהק*

פרטיות ובינה מלאכותית

טכנולוגיות של בינה מלאכותית צורכות מידע בהיקפים נרחבים, ובכלל זה מידע אישי שנוגע לבני אדם. גם העיבודים, השימושים והפלט של הטכנולוגיות מתייחסים לבני אדם, ומשפיעים על חייהם. מאמר זה בוחן את הממשק שבין הטכנולוגיה של הבינה המלאכותית לבין הפרטיות. הדיון משלב את פרדיגמת המחקר של משפט וטכנולוגיה, אשר עֵרה לקשרים המורכבים שבין שני הענפים ומוצאת בהם מכנה משותף, ערכי, עם הבסיס העיוני של הזכות לפרטיות. המאמר מתייחס לבינה מלאכותית כשם כולל לשיטות טכנולוגיות לעיבוד מידע, ובכלל זה מידע אישי, תוך שימוש בנתוני עתק בדרך של לימוד מכונה לשם איתור דפוסי פעולה כלליים חדשים ויישומם לאדם מסוים. כדי לברר האם הבינה המלאכותית פוגעת בפרטיות ואם כן כיצד, המאמר בוחן את שלבי החיים השונים של המידע ומצביע על שורת אתגרים בקשר לכללים משפטיים כמו חובת ההודעה ודרישת ההסכמה, עקרונות מזעור המידע וצמידות המטרה שבדיני הפרטיות, אתגר קצירת מידע פומבי מהרשת, אתגרים של זיהוי בשלב איסוף המידע אך לא בשלב השימוש בטכנולוגיה, ואתגרים שקשורים לעיבוד המידע, שימוש בו לצורכי הסקת מסקנות, וגיבוש תחזיות על האדם. המסקנה בעת הזו היא שבמידה לא מבוטלת לדיני הפרטיות המתקדמים יש כלים סבירים להתמודד עם אתגרי הבינה המלאכותית, הגם שיש בהם סוגיות שטעונות פתרון. אולם מעל לכול מרחף אתגר הכוח – פערי הכוח שבין הצדדים: אנחנו, נושאי המידע, מול תאגידי המידע הרעבים, ומול המכונה.

מבוא

ושוב, טכנולוגיה חדשה פורצת לחיינו, והיא משנה סדרי עולם.[1] בינה מלאכותית שהייתה בגדר מדע בדיוני במשך שנים רבות קורמת כעת רשתות וסיביות, בולעת נתונים בהיקפי

* פרופסור מן המניין, הפקולטה למשפטים, אוניברסיטת תל אביב. תודה לניבה אלקין־קורן, לעמית אשכנזי, לתמר מגידו, לניר גרסון, לחיים רביה, לעמרי רחום־טוויג, לתומר שדמי ולתהילה שוורץ אלטשולר על הערות מצוינות, ליובל אלי־עז על עזרה במחקר, ולמרכז הנשיא מאיר שמגר למשפט דיגיטלי וחדשנות על תמיכתו במחקר, וכן למענק (1838/24) של הקרן הלאומית למדע במסגרת תכנית מפ״צ.

ענק, מייצרת מידע וידע חדשים, ואפשרויות שמסחררות אותנו. זו טכנולוגיה רב־תכליתית (general purpose) ודו־שימושית (dual use), שניתן להשתמש בה לטוב או לרע, בדומה לטכנולוגיות קודמות:[2] טכנולוגיות בינה מלאכותית יכולות לפענח מידע רפואי,[3] לאתר התנהגות חריגה בצילומי מעקב,[4] לסייע בכתיבה,[5] ולשרת לקוחות באמצעות צ׳אט־בוט. אבל הטכנולוגיה משתמשת במידע אישי, מאפשרת מעקב תאגידי ומדינתי, והיא עשויה לשמש כלי בידי חורשי־רעה לפגוע באחרים, למשל לייצר תוכן מזויף או לבצע מתקפות סייבר. השלכות הבינה המלאכותית רק מתחילות להתברר.

מאמר זה בוחן את גורלה של הפרטיות אל מול הבינה המלאכותית. האתגר – ניכר, ומגיע הן מצד הטכנולוגיה והן מצד הפרטיות. הטכנולוגיה עדיין מתפתחת, ורק מתחילה לפלס את דרכה לשימוש ציבורי נרחב, אם כי במהירות. בשלהי 2022 חזינו בפליאה בכניסתן של טכנולוגיות בינה מלאכותית יצרניות (Generative AI), שמייצרות תוכן חדש, כמו ChatGPT ו־Midjourney. ״בינה מלאכותית״ היא שם כולל לשיטות רבות לאיסוף ועיבוד מידע. גיוון זה מוסיף לקושי להבין ולהכיל את השינויים המהירים. בהמשך אציע הגדרת עבודה לצורך הדיון. הבירור מאתגר, משום שגם הפרטיות היא מושג מורכב ועמום. נורמות חברתיות, שיכולות לעיתים לסייע בהכוונה מול טכנולוגיות חדשות, עדיין לא התגבשו. המפגש בין טכנולוגיה חדשה ודינמית למושג משפטי מופשט, בהיעדר נורמות ברורות ופרקטיקות חברתיות מבוססות, מייצר שאלות רבות, ולפי שעה פחות תשובות.

עוד לפני עשרים שנה עמד טל ז׳רסקי על המשמעויות של טכנולוגיות לכריית מידע, ואף הזכיר במפורש את הבינה המלאכותית[6] ואת השפעתה על הפרטיות. מאז השתכללו הטכנולוגיות וקנו אחיזה בשוק, והמשפט התפתח. המאמר מציע מבט עכשווי, ובוודאי לא אחרון, על הסוגיה. אציע מסגרת חשיבה שנשענת על הבנת מאפייניה הערכיים

1 המונח המקובל הוא ״טכנולוגיה משבשת״ (disruptive technology), ומקורו מיוחס לחוקרים בתחום מנהל העסקים: Joseph L. Bower & Clayton M. Christensen, *Disruptive Technologies: Catching the Wave*, 73 HARV. BUS. REV. 43 (1995).

2 לטכנולוגיה דו־שימושית, ראו John Forge, *A Note on the Definition of 'Dual Use'*, 16 SCI. & ENG. ETHICS 111 (2010).

3 ראו Pranav Rajpurkar & Matthew P. Lungren, *The Current and Future State of AI Interpretation of Medical Images*, 388 N. ENGL. J. MED. 1981 (2023).

4 ראו R. Madhavan et al.,*Violence Detection from CCTV Footage Using Optical Flow and Deep Learning in Inconsistent Weather and Lighting Conditions*, *in* ADVANCES IN COMPUTING AND DATA SCIENCES 638 (Mayank Singh et al. eds. 2021).

5 אפשרות זו מעוררת שאלות אתיות. ראו למשל Brady D. Lund et al., *ChatGPT and A New Academic Reality: Artificial Intelligence-Written Research Papers and the Ethics of the Large Language Models in Scholarly Publishing*, 74 J. ASS'N INFO. SCI. & TECH. 570 (2023).

6 Tal Z. Zarsky, *'Mine Your Own Business!': Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion*, 5 YALE J. L. & TECH. 1 (2003).

הרלוונטיים של הבינה המלאכותית במשולב עם מובנָהּ העיוני של הפרטיות, ואמפֶּה את הסוגיות העיקריות במפגש זה.

נקודות האחיזה לדיון – מעטות יחסית, ומתוך אלה, רובן כללי וראשוני. בישראל, חקיקה ייעודית שעוסקת בבינה מלאכותית – אין. בעולם, יש ניצנים לחקיקה, והאיחוד האירופי מוביל עם חוק הבינה המלאכותית (AI Act, להלן: AIA), שיש בו התייחסות להיבטי פרטיות מסוימים בלבד ולא הסדרה כוללת של הפרטיות.[7] חקיקת הפרטיות הקיימת אינה עוסקת במפורש בבינה מלאכותית (למעט החלטות אוטומטיות, ואדון בכך בהמשך), אולם אין חולק שהדין חל על היבטים שונים של הבינה המלאכותית. הדין שואף להיות ניטרלי לטכנולוגיה, משום שהמחוקק מבקש לחסן את החקיקה מפני שינויים עתידיים. התוצאה היא שהדין מורכב מעקרונות ברמת הפשטה גבוהה יחסית ובצידם כללים קונקרטיים יותר, אבל עדיין כלליים. אתגר זה מוכר מתחומי משפט רבים בנוגע לממשק בין משפט לטכנולוגיות מידע: כיצד להתאים דין ישן למציאות חדשה. חקיקת הפרטיות המובילה בעולם היא הרגולציה של האיחוד האירופי להגנת מידע אישי, ה-GDPR,[8] שיש לה השפעה גלובלית ניכרת.[9] בינתיים, מספר התביעות שהוגשו בקשר להיבטי פרטיות של בינה מלאכותית גדל, בעיקר בארצות הברית,[10] ובישראל

7 הצעת החוק אושרה במוסדות האיחוד והנוסח הסופי ייכנס לתוקף בהדרגה, בשנתיים שלאחר פרסומו. ראו Artificial Intelligence Act, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024 O.J. (L 304) 1, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689. היבטי הפרטיות העיקריים ב־AIA הם איסור על איסוף מידע ביומטרי באמצעות מצלמות מעקב (CCTV), ועיבוד מידע שעלול לשמש לאפליה. לסקירה תמציתית, ראו עמרי רחום־טוויג "המשפט לומד את המכונה: ה־AI Act האירופי כחקיקה טכנולוגית בזמן אמת" **פורום עיוני משפט** מח (2024).

8 General Data Protection Regulation, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC O.J. (L 119) (1) (להלן: GDPR).

9 ההשפעה החלה עוד לפני ה־GDPR, בדרך של "גלובליזציה משפטית רכה". ראו Michael Birnhack, *The EU Data Protection Directive: An Engine of a Global Regime*, 24 COMPUT. L. & SEC. REP. 508 (2008). ההשפעה הנוכחית הומשגה כ"אפקט בריסל": ANU BRADFORD, THE BRUSSELS EFFECT: HOW THE EUROPEAN UNION RULES THE WORLD 114 (2020).

10 בארצות הברית הוגשו תביעות ייצוגיות אחדות של משתמשים נגד חברות כמו OpenAI, גוגל ו־Clearview. בחלקן הושגה פשרה, וחלקן עדיין תלוי ועומד כעת. לפשרה האחרונה, ראו דיווח של מי שחשפה את השימוש של Clearview בצילומים שקצרה ברשת: Kashmir Hill, *Clearview AI Used Your Face. Now You May Get a Stake in the Company*, N.Y. TIMES (June 13, 2024)

הוגשה במאי 2024 תביעה ראשונה בעניין.[11] בשל השינויים התכופים בטכנולוגיה ובמשפט, הדיון המוצע אינו דוקטרינרי באופיו, אלא מבקש לברר שאלות עקרוניות ותיאורטיות.

בשנים האחרונות החל גל דוחות של ארגונים בין־לאומיים ומדינתיים שבחנו היבטים של הבינה המלאכותית, כמו מסמכים של ה־OECD,[12] דוח הממשל האמריקני,[13] ובישראל דוח ממשלתי מקיף.[14] רשויות להגנת פרטיות פרסמו מסמכים ייעודיים.[15] הדוחות ממפים סוגיות שונות שקשורות בבינה מלאכותית, ומציעים עקרונות כלליים (למשל, ב־OECD – "האדם במרכז"),[16] ולעניינו, הערות כלליות יחסית על הפרטיות. סוגיית הפרטיות עולה כמעט בכל המסמכים,[17] אולם אין בהם אחידות בנוגע לזיהוי הפגיעה בפרטיות

https://www.nytimes.com/2024/06/13/business/clearview-ai-facial-recognition-settlement.html. במקביל, פלטפורמות רשת תבעו חברות שקצרו מהן מידע, ובינתיים, נראה שיד החברות הקוצרות על העליונה. ראו למשל Meta Platforms, Inc. v. Bright Data Ltd. (N.D. Cal. Jan. 23, 2024). מֶטָה, מפעילת פייסבוק ואינסטגרם, תבעה את ברייט-דאטה הישראלית בגין קציר מידע פומבי משירותים אלה, כאשר עילת התביעה הייתה הפרת תנאים חוזיים באתרים אלה. בית המשפט דחה את הטענות.

11 ת"צ (מחוזי מר') 42145-05-24 **כהן נ' Microsoft Corporation** (נבו 19.5.2024).

12 OECD, Recommendations of the Council on Artificial Intelligence (2022) (להלן: המלצות ה־OECD), ולאחרונה: OECD, AI, Data Governance and Privacy: Synergies and Area of International Co-operation (2024).

13 White House Office of Science and Technology Policy, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People (Oct. 2022). **ראו** גם את דוח הרשות הלאומית לתקשורת ומידע: National Telecommunications and Information Administration (NTIA), AI Accountability Policy Report (Mar. 2024).

14 משרד המשפטים ומשרד החדשנות, המדע והטכנולוגיה **עקרונות מדיניות, רגולציה ואתיקה בתחום הבינה המלאכותית** (2023) (להלן: דוח הממשלה). הרשות להגנת הפרטיות הייתה מעורבת בכתיבת החלקים המתאימים בדוח זה. ראו גם ועדת המשנה של המיזם הלאומי בנושא בינה מלאכותית, אתיקה ורגולציה **דין וחשבון** (2019).

15 למשל **באיחוד האירופי**, European Data Protection Supervisor (EDPS), Artificial Intelligence, Robotics, Privacy and Data Protection (2016); **בבריטניה**, דוח **נציב הגנת הפרטיות**: Information Commissioner's Office, Big Data, Artificial Intelligence, Machine Learning and Data Protection (2017) (להלן: דוח ה־ICO).

16 **ראו** המלצות ה־OECD, לעיל ה"ש 12, בעמ' 7; ובדומה, דוח הממשלה, לעיל ה"ש 14, **בעמ' 104**.

17 מחקר שבחן 36 מסמכי מדיניות של ממשלות, ארגונים בין־לאומיים, ארגוני החברה האזרחית והמגזר הפרטי, מצא שב־97% מהם נזכרה הפרטיות – יותר מכל עיקרון אחר. **ראו** Jessica Fjeld et al., *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI* 21 (Berkman Klein Center for Internet & Society, 2020) https://dash.harvard.edu/handle/1/42160420. **גם**

שנגרמת על ידי בינה מלאכותית. גם בספרות יש ניצנים ראשונים,[18] אבל גם שם הדגשים מגוונים. גם כמה תאגידים פרסמו דוחות בנושא, ובאלו, מטבע זהות המפרסמים, יש לעיין בזהירות.[19] ניכר שאנחנו בראשיתו של הבירור.

קושי מרכזי שעולה מהשיח הראשוני הוא לדייק את הפגיעה בפרטיות שהבינה המלאכותית מביאה עימה. האם ChatGPT פוגע בפרטיות? מה באשר לפיענוח הדמיות רפואיות? האם טכנולוגיות ביתיות ("האינטרנט של הדברים") כמו Alexa פוגעות בפרטיות?[20] כדי להשיב, אמסגר את הדיון בפרדיגמת המחקר של משפט וטכנולוגיות מידע. הנחת היסוד היא שהטכנולוגיה איננה ישות נפרדת שמתפתחת מעצמה בחלל ריק, אלא היא עמוסה בערכים, כפי שאסביר בהמשך. גם הבינה המלאכותית מפותחת על ידי בני אדם במסגרת תאגידית או מדינתית, והיא עדיין אינה יוצרת את עצמה. כדי לעצב כללים מיטביים לטכנולוגיה משתנה, או ליתר דיוק כדי לעצב כללים להתנהגות האנושית והמוסדית בנוגע לטכנולוגיה החדשה, יש לברר את מאפייניה, בבירור ערכי-טכנולוגי.

אציע לבחון את שלבי החיים העיקריים של המידע: יצירתו ואיסופו והשימוש בו (כולל עיבוד המידע), או במונחים טכנולוגיים, את שלבי הקלט של המידע הגולמי והפלט של תוצרי העיבוד של הבינה המלאכותית. לעיתים הפגיעה בפרטיות סמויה: היא מתבצעת רק בשלבי האיסוף ונמוגה בשלבים מאוחרים יותר, של שימוש בתוצרים. לעיתים, האתגר הפוך: אנחנו ערים לפגיעה בפרטיות בשלב השימוש בטכנולוגיה, אבל איננו מבינים מה קורה בתוך "הקופסה השחורה".[21] בנוגע לשלב האיסוף, אבחן סוגיות

דוח ה־NTIA בארצות הברית, לעיל ה"ש 13, בעמ' 11, ציין שכמעט כל התגובות שהתקבלו דנו בפרטיות.

18 בספר הביאורים המקיף ביותר על ה־GDPR (שבו כ־1,400 עמודים) יש רק שלושה אזכורים של הבינה המלאכותית: The EU General Data Protection Regulation (GDPR): A Commentary (Christopher Kuner et al. eds., 2020), בעמ' 112 (בנוגע למאגרי מידע שכוללים מידע מעורבב על בני אדם ומידע אחר), 416 (בנוגע לעקרון השקיפות וחובת ההודעה. הבינה המלאכותית מוצגת כקופסה שחורה); 568 (אזכור כללי, בנוגע לעקרון האחריותיות).

19 למשל דוח של חברת אינטל: David Hoffman & Riccardo Masucci, Intel's AI Privacy Policy White Paper: Protecting Individuals' Privacy and Data in the Artificial Intelligence World (2018). לסקירת דוחות תאגידיים נוספים, ראו מוסד שמואל נאמן למחקר מדיניות לאומית **בינה מלאכותית, מדעי הנתונים ורובוטיקה חכמה: דו"ח בנושא אתיקה, משפט ופרטיות** (2018).

20 לדיון, ראו למשל Eldar Haber, *The Internet of Children: Protecting Children's Privacy in A Hyper-Connected World*, 2020 U. Ill. L. Rev. 1209 (2020).

21 המונח מתאר את חוסר הסימטריה בנוגע לאיסוף מידע והשימוש בו, ובנוגע להחלטות אלגוריתמיות: אוספי המידע יודעים הכול על נושאי המידע (data subjects), אבל אלה

של הודעה והסכמה, של מזעור המידע וצמידות המטרה, ואת הקושי שמתעורר בעת קציר מידע (scraping) ממקורות פומביים ברשת. אבחן גם את אתגר הזיהוי וההתממה. בשלב השימוש, אבחן את משמעות התוצר של הבינה המלאכותית – המידע המעובד, את השימוש בתוצר להפקת תחזיות על התנהגות אנשים, ואת אתגר המעקב הקפיטליסטי אחרי צרכנים,[22] והמעקב המדינתי אחרי אזרחים.[23] הסך הכול מצטרף לאתגר של כוח.

מתוך הדיון תעלה טענה מרכזית בקשר ליכולת של דיני הפרטיות להתמודד עם אתגרי הבינה המלאכותית: במידה רבה, אתגרי הבינה המלאכותית כבר מוכרים למשפט ממפגשים טכנולוגיים-משפטיים קודמים.[24] ברמה העיונית, הבנת הפרטיות כשליטה מספקת לנו מורה דרך לבחינת הממשק החדש, וגם מציעה כלים מסוימים למענה.[25] הדברים אמורים בדיני הפרטיות במישור העקרוני: בפועל, קיומו של סל כלים וניסיון משפטי אינם מבטיחים שהכלים יופעלו כראוי, וגם בדין הקיים יש אתגרים רבים שטרם נפתרו. בישראל, עוד יש כברת דרך עד לגיבושם של דיני פרטיות במתכונת מספקת.

חלק א מציג את יסודות הדיון: את הפרדיגמה של משפט וטכנולוגיה, ובקצרה, את הזכות לפרטיות. כמו בכתיבתי הקודמת, אצדד בעמדת ה"פרטיות כשליטה" כמורת דרך מתאימה,[26] שממנה נגזרת בחינת מחזור החיים של המידע. מתוך זה, אציג את הגדרת העבודה של "בינה מלאכותית" ואת הקשר שלה למושג של נתוני עתק (big data), שכבר נדון בספרות המשפטית. **חלק ב** יברר האם בינה מלאכותית פוגעת בפרטיות, וכיצד. אבחן את מחזור החיים של המידע במערכות בינה מלאכותית. אסכם בדיון באתגר הכוח, ויציע על אפיקים אפשריים להתמודדות עם האתגרים.

יודעים מעט מאוד על השימושים במידע. ראו FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015).

22 SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR THE FUTURE AT THE NEW FRONTIER OF POWER (2019).

23 לפרטיות אזרחים מול המדינה, ראו מיכאל בירנהק **פרטיות חוקתית** (2023).

24 למסקנה דומה בהקשר האמריקני, ראו Daniel J. Solove, *Artificial Intelligence and Privacy*, 77 FLA. L. REV. (Forthcoming 2025). סולוב מזהיר מפני AI exceptionalism, כלומר מפני בידוד ההתייחסות המשפטית ל־AI. האתגר המשפטי המרכזי בעיניו הוא שסוגיות נפרדות מתערבבות זו בזו.

25 סולוב חולק באופן נחרץ על עמדה זו, וסבור שהבינה המלאכותית חרצה לשבט את גורלה של "הפרטיות כשליטה", בשל חוסר היכולת של משתמשים להבין את השפעת הטכנולוגיה על פרטיותם. שם, בעמ' 21. אני מסכים שהקושי ניכר, אבל כמו משפטנים אמריקנים אחרים, סולוב מרדד את הפרטיות כשליטה למרכיב ההסכמה בלבד.

26 ראו מיכאל בירנהק **מרחב פרטי: הזכות לפרטיות בין משפט לטכנולוגיה** 89 (2010).

א. אבני הבניין: משפט וטכנולוגיה, פרטיות ובינה מלאכותית

הבינה המלאכותית היא עוד מקרה שבו טכנולוגיה חדשה פוגשת מציאות משפטית קיימת. הפיתוי לומר שהמשפט מדדה בעצלתיים אחרי הטכנולוגיה – גדול.[27] במובן טכני, האמירה נכונה. החקיקה הישראלית אינה מאזכרת כיום במפורש את הבינה המלאכותית. אבל במובן המהותי, האמירה שגויה: למשפט ולטכנולוגיה יש מכנה משותף ערכי. חלק זה מציג בתמציתיות את מסגרת הדיון, שהיא פרדיגמת המחקר של משפט וטכנולוגיה, את הזכות לפרטיות, ואת השילוב ביניהן, שמביא להגדרת עבודה של הבינה המלאכותית לצורך בחינת היבטי הפרטיות. הדיון ינחה אותנו בחלק ב, בבירור הפגיעה בפרטיות שהבינה המלאכותית יוצרת.

1. פרדיגמת המחקר של משפט וטכנולוגיה

בעת עיסוק משפטי בטכנולוגיות בכלל, ובטכנולוגיות מידע בפרט, ראוי לחשוף את הנחות היסוד בדבר הקשר שבין המשפט לטכנולוגיה. מי שסבור שהטכנולוגיה פועלת ומתפתחת כישות עצמאית לפי חוקים משלה – מחסן אותה בפועל מפני המשפט,[28] שמוצג כגורם חיצוני, ומזיק. מנגד, מי שסבור שאין ייחוד בטכנולוגיה יסיק שהיא דומה לכל פעילות אנושית אחרת שיש להסדיר לפי הצורך. ברוח זו, ידועה אימרתו של השופט האמריקני פרנק איסטרברוק (Easterbrook) שדיני הסייבר כמוהם כדיני הסוס: את הסוס אפשר לקנות ולמכור, הסוס יכול לבעוט ולהזיק, ובאלה, הסביר איסטרברוק, עוסקים דיני הקניין, החוזים, הנזיקין ועוד.[29] לפיכך, הסיק, אין צורך בענף משפטי מיוחד.

פרדיגמת המחקר של משפט וטכנולוגיה דוחה את שתי עמדות הקיצון האלה.[30] הפרדיגמה מדגישה שאנשים, תאגידים ומדינות הם שמפתחים טכנולוגיות חדשות, ובכך דוחה את הגישה שהטכנולוגיה מתפתחת בעצמה. הטכנולוגיה נטועה תמיד

27 ראו למשל עע"ם 3782/12 **מפקד מחוז תל אביב-יפו במשטרת ישראל נ' איגוד האינטרנט הישראלי**, פ"ד סו(2) 159, 178 (2013), שפירש את הביטויים "מקום משחקים אסורים" ו"חצרים" לפי ס' 224 ו־229(א)(1) לחוק העונשין, התשל"ז–1977, כדי לברר האם יש למשטרה סמכות להורות לספקים של שירותי אינטרנט לחסום גישה לאתרי הימורים מקוונים.

28 ראו למשל את גישתו של גורו הטכנולוגיה האמריקני: KEVIN KELLY, WHAT TECHNOLOGY WANTS (2010).

29 Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEG. F. 207, 207-8 (1996).

30 להרחבה, ראו ניבה אלקין־קורן ומיכאל בירנהק "הקדמה: משפט וטכנולוגיות מידע" **רשת משפטית: משפט וטכנולוגית מידע** 11 (ניבה אלקין־קורן ומיכאל בירנהק עורכים 2011).

בסביבה חברתית. לפיכך, מדובר בפעילות אנושית שבהסדרתה המשפט מורגל. בצד זה, הפרדיגמה מכירה בכך שלטכנולוגיה יש מאפיינים ייחודיים.

המאפיין המרכזי הוא שהטכנולוגיה אינה חפה מערכים.[31] לעיתים, אלה מהנדסות של הטכנולוגיה שהטמיעו בה ערכים מסוימים בכוונה תחילה. "מעלית שבת" היא טכנולוגיה שנועדה לאפשר לשומרי שבת להשתמש במעלית מבלי לחלל את השבת. זה הערך שהוטמע בכוונה בטכנולוגיה. לעיתים המהנדסים כיוונו לערך אחד, אבל התוצאה הייתה שונה. למשל חברת נטסקייפ, שפיתחה את אחד הדפדפנים הראשונים ברשת, הטמיעה את הקוקית (עוגייה, cookie) כדי לסייע לגולשים כך שלא יצטרכו להירשם מחדש בכל אתר (ערך של יעילות), וכדי להימנע מיצירת מספר מזהה אחד לכל משתמש בכל הרשת (ערך של פרטיות).[32] המטרות הושגו, אבל הקוקית הפכה לטכנולוגיה מרכזית למעקב ברשת. אופני יצירת הערכים שמוטמעים בטכנולוגיה מגוונים, אבל לענייננו, די בכך שהטכנולוגיה היא תוצר של תהליכים חברתיים וטומנת בחובה ערכים שונים. מעת שתובנה זו התבהרה, אפשר היה לגייס טכנולוגיות באופן יזום כדי להשיג מטרות ערכיות חברתיות.[33] לורנס לסיג (Lessig) בחן את הארכיטקטורה (הפיזית והטכנולוגית) והכריז: "הקוד הוא המשפט".[34] כלומר, הטכנולוגיה – שמיוצגת על ידי קוד המחשב – יכולה להיות גורם מאסדר עצמאי. כך גם לגבי הבינה המלאכותית.

מול הטכנולוגיה ניצב המשפט. אם אנו מבינים את המשפט כאוסף כללים טכניים שאין בהם ממד ערכי, או ששיטת הפרשנות שנאמץ מסרבת ליצוק תכנים חדשים מעבר למה שהמחוקק התכוון בשעתו, די לבחון את לוח הזמנים: חוק שנחקק לפני פריצתה של טכנולוגיה חדשה – מיושן. אולם המשפט הוא ביטוי ציבורי להכרעות ערכיות ואינו רק אוסף כללים שרירותי, והוא נתון לפרשנות דינמית. במשפט הישראלי מקובלת הפרשנות התכליתית שמניחה שלכל כלל יש תכלית, ברמת הפשטה כללית (למשל, ערכי המדינה כמדינה יהודית ודמוקרטית),[35] או באופן קונקרטי.[36]

התוצאה היא שלמשפט ולטכנולוגיה יש מכנה משותף, ערכי. אלה שתי דרכים שונות למימוש ערכים. במשפט – הדברים נעשים בדרך כלל בדיון פומבי ומשיקולים ציבוריים, ואילו במקרה הטכנולוגי, החלטות מתקבלות במסגרת פרטית ותאגידית, משיקולים עסקיים. פעמים רבות, מה שנראה כהתנגשות בין המשפט לטכנולוגיה אינו אלא התנגשות בין ערכים שונים.

---

31 HUMAN VALUES AND THE DESIGN OF COMPUTER TECHNOLOGY (Batya Friedman ed., 1997).

32 ראו Steven Johnson, *The Magic Cookie: How Lou Montulli Cured the Web's Amnesia*, HIDDEN HEROES (2022), https://hiddenheroes.netguru.com/lou-montulli.

33 Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553 (1998).

34 LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (1999).

35 ראו אהרן ברק **פרשנות תכליתית במשפט** (2003).

36 ראו למשל יישום בהקשר טכנולוגי, עניין **איגוד האינטרנט**, לעיל ה"ש 27, בעמ' 177–178.

פרדיגמת המחקר של משפט וטכנולוגיה חושפת את מורכבות היחסים בין שני הגורמים ומאפשרת לנו להימנע מאמירות כוללניות נוסח "זה גובר על זה". במקום להנגיד את שני הגורמים החברתיים, אפשר ליזום שיתופי פעולה ביניהם; למשל הרעיון של הנדסת פרטיות, שלפיו יש להטמיע את הפרטיות כבר בשלב תכנון הטכנולוגיה.[37]

הפרדיגמה מסייעת גם לחוקרי וחוקרות המשפט שבוחנים את הטכנולוגיה, ומנחה אותנו לברר מצד אחד את תכלית הדין, וזו פעולה פרשנית-משפטית, ומצד שני לברר מהם הערכים שמגולמים בטכנולוגיה. פעולה זו אינה מובנת מאליה למשפטנים, ובדרך כלל עוסקים בה חוקרי תקשורת שאמונים על פיענוח מאפייני המדיום, כמו גם ענף המחקר של Science, Technology, Society – STS. הניתוח המשפטי צריך להיעזר בתחומי ידע אלה כדי לפענח את המשמעות הערכית של הטכנולוגיה. המשימה היא לחלץ את מאפייני הטכנולוגיה, את המזמינויות שלה (affordances), כלומר את התכונות של הטכנולוגיה שמכתיבות כיצד אפשר להשתמש בה וכיצד אי אפשר להשתמש בה. מירייה הילדברנדט (Hildebrandt) כותבת בדומה: "Tracing potential normative impacts means investigating what types of behaviours are invited or inhibited, enforced or ruled out by a particular technological device or infrastructure".[38] את הערכים שבטכנולוגיה נציב מול תכלית הענף המשפטי הרלוונטי.

בהתאם, אפנה לבחינת היסוד המשפטי-ערכי של הפרטיות, ומשם נוכל לבחון את הבינה המלאכותית באמצעות משקפיים ערכיים.

2. הזכות לפרטיות

אך מובן שאי אפשר להקיף את הזכות לפרטיות במאמר זה, ואסתפק בהערות ממוקדות בלבד. לזכות לפרטיות יש שלושה מישורי פעולה מרכזיים: פרטיות קלאסית בין־אישית – שהמשפט מסדיר בדיני הנזיקין,[39] פרטיות חוקתית, שפועלת בין המדינה לאזרחיה והמשפט מסדיר במשפט הציבורי,[40] ופרטיות מידעית, שמתמקדת באיסוף מידע ועיבודו, בדרך כלל בין צרכנים לתאגידים. הכלי המשפטי למישור האחרון הוא דינים ייעודיים, באירופה הם מכונים data protection, בארצות הברית פרטיות מידעית (informational privacy), ובישראל יש הסדר שמתמקד במאגרי מידע, בפרק ב של חוק הגנת הפרטיות,

37 ראו Ann Cavoukian, Privacy by Design – Take the challenge (2009). הרעיון עוגן במפורש במשפט האירופי, בס׳ 25 ל־GDPR.

38 Mireille Hildebrandt, *Introduction: A Multifocal View of Human Agency in the Era of Autonomic Computing*, *in* Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology 1, 2 (Mireille Hildebrandt & Antoinette Rouvroy, eds., 2011).

39 Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 Harv. L. Rev. 193 (1890).

40 להרחבה ראו בירנהק **פרטיות חוקתית**, לעיל ה״ש 23.

התשמ״א–1981.[41] כיום, הסוגיות העיקריות שהבינה המלאכותית מעוררת הן במישור הפרטיות המידעית ובמישור הפרטיות החוקתית. ככל שיתרחב השימוש בטכנולוגיה תתעוררנה שאלות גם במישור הפרטיות הבין־אישית, למשל מי שנעזר בבינה מלאכותית לכתיבת מכתב המלצה ומזין נתונים אישיים של מושא ההמלצה, ואלה יופיעו בפלט שמקבל משתמש אחר.

דיני הפרטיות אינם מגדירים מהי הפרטיות אלא רק קובעים את קיומה באופן כללי, כדרכו של חוק־יסוד: כבוד האדם וחירותו (סעיף 7(א), שם הוא מכריז: ״כל אדם זכאי לפרטיות ולצנעת חייו״). חוק היסוד מתייחס גם למצבים מסוימים (למשל בנוגע ל״סוד שיחו״ של אדם (סעיף 7(ד)). חוק הגנת הפרטיות מתייחס לפרטיות על דרך השלילה, ומגדיר מצבים שהם פגיעה בפרטיות.[42] אף אחד מאלה אינו דן בבינה מלאכותית. החוק מנוסח ברמת הפשטה גבוהה ונמנע מהתייחסות טכנולוגית, ולא בכדי. הסיבה היא העמימות המובנית של הפרטיות, ותלותה בנורמות חברתיות משתנות ובטכנולוגיות משתנות.

במקום אחר תיארתי את הצדקות הפרטיות כמעגלים קונצנטריים הולכים ומתרחבים – מהאדם עצמו ליחסים בין־אישיים, לקהילה ולמדינה.[43] ההצדקות המגוונות, מישורי הפעולה השונים ומגוון ההקשרים של הפרטיות מקשים לראות בה מרקם משפטי אחד. בספרות יש ניסיונות בולטים להפגת העמימות. ניסיון אחד מדגיש את הגישה אל האדם ומתאר את הפרטיות ככוח למניעת גישה כזו באופנים שונים.[44] ניסיון שני מדגיש את שליטתו של האדם במידע שעל אודותיו.[45] גישה זו מנחה אותנו להגן על הפרט ולהפקיד בידיו שליטה במידע זה. שתי הגישות הן שני הצדדים של אותו מטבע – האחת מתחילה מחוץ לאדם ובוחנת את חדירותו (פרטיות כגישה), והאחרת מתחילה בתוך האדם ומרחיבה את שליטתו החוצה (פרטיות כשליטה). בכל מקרה, שתי הגישות אינן הצדקות כשלעצמן אלא תרגום של עקרונות מופשטים למושג ברמת הפשטה בינונית, שגם אותו יש להמשיך ולתרגם לכללים קונקרטיים. ואכן, שתי הגישות מתכנסות לעקרונות הגנת המידע האישי (FIPs – Fair Information Practices).

עקרונות הגנת המידע מופיעים בשורת חוקים, הנחיות והמלצות בעולם מאז שנות השבעים של המאה הקודמת. החקיקה המובילה היא כאמור רגולציית הגנת המידע האישי באיחוד האירופי, ה־GDPR. נמצא שם עקרונות של הודעה והסכמה, מגבלות שמוטלות על אוספי המידע – כמו עקרון מזעור המידע (איסוף מינימום המידע שנחוץ

41 באוגוסט 2024 תיקנה הכנסת את החוק בתיקון מקיף, שייכנס לתוקף באוגוסט 2025. ראו חוק הגנת הפרטיות (תיקון מס׳ 13), התשפ״ד–2024, ס״ח 1430 (להלן: תיקון 13). בצד ההסדר הכללי יש חקיקה ייעודית שעוסקת בסוגי מידע מסוימים, למשל חוק מידע גנטי, התשס״א–2000.

42 ס׳ 2, 2א לחוק הגנת הפרטיות.

43 בירנהק **פרטיות חוקתית**, לעיל ה״ש 23, בעמ׳ 68–86.

44 Ruth E. Gavison, *Privacy and the Limits of Law*, 89 YALE L.J. 421 (1980).

45 ALAN WESTIN, PRIVACY AND FREEDOM 7 (1967); לקריאה עדכנית, ראו: Lisa M. Austin, *Re-Reading Westin*, 20 THEORETICAL INQUIRIES L. 53 (2019).

להשגת תכלית ראויה), עקרון צמידות המטרה (יש להשתמש במידע רק למטרה שלשמה נאסף), חובות סודיות ואבטחת מידע, וחובות נלוות, כמו חובה לדווח על דלף מידע. לנושאי המידע (data subjects) יש זכויות שונות, כמו זכות לעיין במידע שעל אודותיהם, לדרוש את תיקונו, ובאיחוד האירופי גם הזכות להישכח. בדין האירופי נמצאות גם חובת מינוי ממונה הגנת פרטיות, חובה לעריכת תסקיר מקדים להגנת הפרטיות, וחובה לבצע תהליך של הנדסת פרטיות. הדין הישראלי, שהיה מהמובילים בעולם בשנות השמונים במאה הקודמת, נותר מאחור, וכיום מנסה להדביק את הפערים.

כפי שהסברתי במקומות אחרים, יש ציר משותף לכל הצדקות הפרטיות, בכל מעגליה, וזו גישת הפרטיות כשליטה.[46] משמעותה היא שלאדם יש כוח (משפטי) לשלוט במידע האישי שעל אודותיו: להחליט אם למסור אותו, למי ולאילו מטרות, מתי ובאילו תנאים. השליטה משקפת את הצורך שלנו במרחב פרטי – פיזי, וירטואלי ואינטלקטואלי. גישה זו קשורה קשר הדוק לאוטונומיה של האדם החופשי ולרעיון של כבוד האדם, ומתאימה היטב לסביבה המשפטית הישראלית.

הפרטיות כשליטה נתונה לביקורת נוקבת, בעיקר מצד חוקרי פרטיות אמריקניים שמצביעים על כישלונה של ההודעה וההסכמה (notice and choice, בלשונם), ומסיקים מכך שאין תוחלת לעמדת הפרטיות כשליטה.[47] הם חוששים שההסכמה מכשילה את המשתמשים, ואני מצטרף לביקורתם. אכן, בפועל, בסביבה הרשתית, "הסכמה" רוקנה מתוכן. ההודעה על השימושים המיועדים במידע מגולמת במדיניות פרטיות, שעימה הגולשים ("משתמשים") מתבקשים להסכים. ההודעות רבות, ארוכות, מנוסחות בשפה משפטית לא נגישה, תוך ניצול כשלים קוגניטיביים.[48] אבל מההבחנה התיאורית בדבר כישלון ההסכמה במתכונתה הנוכחית אין להסיק מסקנה נורמטיבית. המבקרים מרדדים ללא הצדקה את הפרטיות כשליטה לשלב ההודעה וההסכמה ומתעלמים מכללים נוספים להגנת המידע האישי; הם מתעלמים מיתרונות מסוימים שיש בהודעה כשלעצמה, והם מסיקים את הראוי מתוך המצוי, וזהו כשל לוגי.

אזכיר שתי גישות נוספות לפרטיות. המשפטן האמריקני דניאל סולוב (Solove) בחן בפירוט הצדקות שונות לפרטיות, ומצא שהן צרות מדי או רחבות מדי.[49] הוא תיאר את הפרטיות כמושג-על כללי, והשתמש במושג הוויטגנשטייני של "דמיון משפחתי".[50] אבל אז, במקום לאתר הצדקה משכנעת, סולוב הציע טקסונומיה של פגיעות בפרטיות.[51]

46 בירנהק **פרטיות חוקתית**, לעיל ה"ש 23, בעמ' 87.

47 ראו למשל Woodrow Hartzog, *The Case Against Idealising Control*, 4 EUR. DATA PROT. L. REV. 423, 426 (2018); Joshua A.T. Fairfield & Christoph Engel, *Privacy as a Public Good*, 65 DUKE L.J. 385, 408-9 (2015); ARI EZRA WALDMAN, PRIVACY AS TRUST: INFORMATION PRIVACY FOR AN INFORMATION AGE 30-31 (2018).

48 ראו למשל Jonathan A. Obar & Anne Oeldorf-Hirsch, *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services*, 23 INFO. COMMC'N & SOC'Y 128 (2020).

49 DANIEL J. SOLOVE, UNDERSTANDING PRIVACY (2008).

50 שם, בעמ' 40.

51 שם, בעמ' 101 ואילך.

גישתו נוחה לאיתור ראשוני של פגיעות בפרטיות כאשר אנו ניצבים נבוכים מול טכנולוגיה חדשה, אולם אין בה הנחיה נורמטיבית. הפרטיות כשליטה, לעומת זאת, נכללת בכל אחת מההצדקות, ויוצרת את הציר המחבר שאותו חיפש סולוב.

פילוסופית הטכנולוגיה הלן ניסנבאום (Nissenbaum) הציעה מסגרת חשיבה בשם Contextual Integrity.[52] היא ממקדת את המבט ל"הקשר" חברתי מסוים, למשל הקשר רפואי, מבררת את הנורמות שחלות בו ומתמקדת בזרימת מידע: מה מקובל לגבי איסוף מידע, העברתו וכדומה. לשיטתה, עלינו לבחון את השינוי הטכנולוגי-חברתי החדש, למשל מערכת מידע חדשה: האם יש שינוי בזרימת המידע באופן שאינו תואם לכללי זרימת המידע שאיתרה בהקשר המדובר? אם כן, זו פגיעה לכאורה בפרטיות. המתווה שלה הוא מדריך פעולה נוח ושובה לב, אולם הדברים מורכבים יותר.[53] לא תמיד ברור מהו ההקשר המתאים; וכאשר מדובר בהקשר חדש – אין עדיין נורמות מגובשות בדבר זרימת המידע, ולכן אין אמת מידה להערכת השינוי. בנוסף, הצעתה חסרה אמת־מידה ערכית. רבים מההקשרים חדשים לגמרי ואין בהם נורמות ברורות בנוגע למידע. החברות שמציעות מוצרי בינה מלאכותית הן שיוצרות את ההקשר, והן שקובעות את הכללים. הפלא ופלא, הכללים מתאימים לתאגידים, לאו דווקא לבני האדם. וכיצד נעריך האם השינוי ראוי או אינו ראוי?[54]

3. מאפייני הבינה המלאכותית

לפי פרדיגמת המחקר של "משפט וטכנולוגיה" אפנה כעת לטכנולוגיה, ואנסה לחלץ את מאפייניה. "בינה מלאכותית" היא שם גג לטכנולוגיות רבות. בספרות יש אפיונים מגוונים, כל אחד מדגיש פן אחר: את התהליך, את התוצאה, את הדמיון לפעולה האנושית, או את מידת העצמאות של הטכנולוגיה. נראה שהעניין וההתרגשות הנוכחיים סביב הטכנולוגיה נובעים מאפיונה כמחליפה את בני האדם – ביכולת להבין,

52 HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE (2010).

53 לביקורת, ראו Michael Birnhack, *A Quest for a Theory of Privacy: Context and Control: Review of Helen Nissenbaum's Privacy in Context*, 51 JURIMETRICS 447 (2011).

54 הקשיים בגישת ניסנבאום נחשפים במקרה של הבינה המלאכותית. מהו ההקשר הרלוונטי? הבינה המלאכותית היא רב־תכליתית, ונמצא אותה בהקשרים של חינוך ורפואה, ניווט, שכתוב טקסטים, יצירה והמצאה, ובכל תחומי החיים.

לשקול ולהחליט באופן עצמאי.[55] המיקוד כאן הוא בפרטיות, ובהתאם אתמקד בפן המידעי. אתייחס לגישה השלטת כיום לבינה מלאכותית.[56] לצורך מאמר זה:

> בינה מלאכותית היא שם כולל לשיטות טכנולוגיות לעיבוד מידע, ובכלל זה מידע אישי, תוך שימוש בנתוני עתק בדרך של לימוד מכונה, לשם איתור דפוסי פעולה כלליים שלא זוהו עד אז, ויישומם לאדם מסוים.

הגדרת העבודה מדגישה את היבטי המידע ואינה מתיימרת להיות ממצה. זו הגדרת עבודה זמנית, שכן הטכנולוגיה מוסיפה להשתנות ולהתפתח.[57] יש בה מרכיבים מספר. ראשית, הדגש הוא **תהליכי**. אבחן לא רק את התוצאה, אלא גם את השלבים המוקדמים שבהם מידע נאסף ומעובד. הגדרות מבוססות־תוצאה תחמצנה שימושים מוקדמים במידע. בצד הטכנולוגיה מקובל להבחין בשלבים של תכנון המערכת, איסוף המידע ועיבודו, אימות ווידוא הפעולה, הטמעה ופיקוח.[58] בצד הפרטיות – הדין עוסק בכל מחזור החיים של המידע.

שנית, אתמקד בעיבוד **מידע אישי**. לא כל מידע הוא אישי. מידע על צמחים, למשל, אינו כזה. טכנולוגיה לזיהוי צמח לפי צילום אוספת ומעבדת מידע, אולם אין בה היבטי פרטיות. "מידע אישי" הוא מונח מפתח בדיני הפרטיות. יש גישות שונות להגדרתו, והן משקפות תפיסות שונות לגבי הפרטיות. בדין הפדרלי האמריקני נמצא הגנה על סוגי מידע מסוימים לפי תוכן המידע.[59] יש שם שורת חוקים ספציפיים שעוסקים במידע

---

55 הדימוי הפופולרי של בינה מלאכותית הוא של מכונה אוטונומית לחלוטין. בפועל תיתכנה דרגות שונות של מעורבות אנושית, ונקודות שונות בממשק בין האדם למכונה. לסקירה תמציתית ראו דוח הממשלה, לעיל ה"ש 14, בעמ' 53–59. לעיון פילוסופי ראו Peter-Paul Verbeek, *Subject to Technology: On Autonomic Computing and Human Autonomy*, *in* LAW, HUMAN AGENCY AND AUTONOMIC COMPUTING 27, לעיל ה"ש 38, בעמ'.

56 להתפתחות הטכנולוגית של הבינה המלאכותית ראו: Sanjay Chawla et al., *Ten Years after ImageNet: A 360° Perspective on Artificial Intelligence*, 10 ROYAL SOC'Y OPEN SCI. 221414 (2023).

57 השוו להגדרת דוח הממשלה, לעיל ה"ש 14, בעמ' 19: "תחום הבינה המלאכותית הוא שם כולל להתפתחות בתחום טכנולוגיות המידע, התקשורת ומדעי הנתונים, המאפשרת קבלת החלטות, ייצור תחזיות, או ביצוע פעולות על ידי מחשב ברמת עצמאות גבוהה, באופן המדמה או מסוגל להחליף בינה אנושית". להגדרות נוספות, ראו עמיר כהנא ותהילה שוורץ אלטשולר **אדם, מכונה, מדינה: לקראת אסדרה של בינה מלאכותית** 28–31 (2023).

58 ראו בהמלצות ה־OECD, לעיל ה"ש 12, בעמ' 7: ["AI system lifecycle phases involve: i) 'design, data and models'; which is a context-dependent sequence encompassing planning and design, data collection and processing, as well as model building; ii) 'verification and validation'; iii) 'deployment'; and iv) 'operation and monitoring'"].

59 לדיון, ראו בירנהק **מרחב פרטי**, לעיל ה"ש 26, בעמ' 194.

מסוגים שונים, למשל מידע פיננסי, או מידע על קלטות וידאו שאדם שאל.[60] בכל חוק נמצא תנאים נוספים להגדרת מידע אישי, ובמקרים רבים אפשרות לזיהוי האדם.[61] בדין האיחוד האירופי, "מידע אישי" מוגדר לפי האפשרות לזהות את האדם שהמידע הוא על אודותיו.[62] בעיני האירופים כל מידע עשוי להיות רגיש, ולכן די שניתן לקשרו לאדם מסוים כדי להפעיל את הדין. בדין הישראלי, הגדרת "מידע" בחוק מתייחסת ל"נתונים על אישיותו של אדם, מעמדו האישי, צנעת אישותו, מצב בריאותו, מצבו הכלכלי, הכשרתו המקצועית, דעותיו ואמונתו".[63] בכך יש מרכיב תוכן בדומה לדין האמריקני, אך מרחיב ממנו, ועל גביו מרכיב אירופי של זיהוי האדם, לפי התיבה "של אדם", כפי שפורשה בפסיקה.[64] המידע האישי שמשמש את הבינה המלאכותית מגוון: למשל מידע רפואי, גנטי או ביומטרי; מידע על פעילות, תוכן תקשורת, ומידע על מידע (מֶטַה-דאטה), כמו מידע על מיקום או זמן הפעילות.

שלישית, הגדרת העבודה קושרת את הבינה המלאכותית ל**נתוני עתק**, אבל תוך שמירה על עצמאות המושגים. גישה דומה יש גם בדוחות רשמיים אחדים,[65] ובספרות,[66]

60 ראו Gramm Leach Bliley Act (GLBA), Pub. L. 102–106, 113 Stat. 1338 (1999) (מידע פיננסי); Video Privacy Protection Act, Pub. L. 100–618, 102 Stat. 3195 (1988) (קלטות וידאו).

61 ראו Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. Rev. 1814 (2011).

62 ס׳ 4(1) ל־GDPR. ס׳ 9 שם מגדיר גם מידע רגיש במיוחד שכולל למשל מידע על מוצא אתני, דעות פוליטיות, חברות באיגוד מקצועי, ועיבוד מידע גנטי ומידע ביומטרי.

63 ס׳ 7 לחוק הגנת הפרטיות. החוק מגדיר גם מידע רגיש, בהבדלים זניחים. תיקון 13 החליף את ההגדרה הוותיקה בהגדרה האירופית, ולפיו "מידע אישי – נתון הנוגע לאדם מזוהה, או לאדם הניתן לזיהוי [...] מי שניתן לזהותו במאמץ סביר, במישרין או בעקיפין...".

64 ע"א 1697/11 **א. גוטסמן אדריכלות בע"מ נ' ורדי**, פס׳ 18 לפסק הדין של השופט פוגלמן (נבו 23.1.2013). בית המשפט קבע כי כאשר ניתן לזהות אדם מתוך הדמיה ממוחשבת של ביתו, הגם שהוא עצמו וחפציו האישיים אינם נכללים בה, זו פגיעה בפרטיות. שם, פס׳ 24–25 לפסק הדין של השופט פוגלמן.

65 למשל בדוח הממשלה, לעיל ה"ש 14, בעמ׳ 78; דוח ה־ICO, לעיל ה"ש 15, בעמ׳ 6 ("The terms 'big data', 'AI' and 'machine learning' are often used interchangeably but there are subtle differences between the concepts").

66 ראו Adam Thierer et al., Artificial Intelligence and Public Policy, Mercatus Ctr. At George Mason Univ. (2017). לפי המחברים "נתוני עתק" מתייחסים למערך הנתונים, ו"בינה מלאכותית" מתייחסת לאופי הפעילות הטכנולוגי. שם, בעמ׳ 8. אבנר לוין מפריד: להגדרתו, "תחום נתוני העתק הוא תחום עיבוד נתונים העוסק ביכולת לנתח נתונים הנאספים בצורה לא מסודרת...", ואילו בתחום הבינה המלאכותית "מפותחים אלגוריתמים שאמורים לחקות ואף לשפר בסופו של דבר את ההבנה האנושית בקשר לשליטה על תהליכים שונים...". אבנר לוין "הפרטיות בעידן הטכנולוגיות המתפרצות" **טכנולוגיות מתפרצות: אתגרים בדין הישראלי** 215, 221 (ליאור זמר, דב גרינבאום ואביב גאון עורכים 2022).

אך לעיתים הכריכה נעשית תוך חפיפה בין המושגים.[67] הקשר הזה חשוב, משום שהספרות המשפטית בחנה את הממשק של נתוני עתק ופרטיות בעשור האחרון באופן אינטנסיבי.[68] מי שחופף את המושגים, צפוי להסיק שהיבטי הפרטיות שהבינה המלאכותית מעוררת אינם חדשים.[69]

המונח "נתוני עתק" מתייחס הן למאפיינים של המידע עצמו והן לאופן הניתוח של המידע. זהו שם כולל לדרכים לאיסוף של מידע רב (volume), ממקורות מגוונים (variety), שניתן לעבדם על־אתר (velocity). מכאן, מקובל לאפיין נתוני עתק כשלושה V.[70] בדומה, גם המונח "בינה מלאכותית" כולל את חומרי הגלם – את הנתונים, שנאספים לצורך למידת המכונה (Machine Learning),[71] ואת השיטה לניתוח הנתונים.[72] במילים אחרות, נתוני עתק ובינה מלאכותית הם מושגים נפרדים אך קשורים.[73]

רביעית, מרכיב נוסף בהגדרה נזכר כעת – **למידת מכונה**. זהו התהליך הראשוני שבו המכונה לומדת לפעול (השפה המאנישה היא חלק מההילה של הבינה המלאכותית). כדי להבחין בין פטרוזיליה לכוסברה יש ללמד את המכונה מהם ההבדלים בין השתיים, וכך, גם בהקשרים משמעותיים יותר, למשל לימוד המכונה כיצד להבחין בין גידול סרטני לאיבר תקין. הלימוד יכול להיעשות תוך מעורבות והכוונה אנושית, בדרך של למידה

---

67 למשל Rainer Mühlhoff, *Predictive Privacy: Collective Data Protection in the Context of Artificial Intelligence and Big Data*, BIG DATA & SOC'Y 1 (2023), שמתייחס במאוחד ל־"AI and big data technology".

68 ראו למשל דיון מוקדם, עוד לפני חקיקת ה־GDPR, אצל Omer Tene & Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 NW. J. TECH. & INTELL. PROP. 239 (2013).

69 Thierer ועמיתיו כותבים: "it may be the case that many of the concerns related to privacy, bias, and discrimination are already covered by existing laws and regulations that address human failings in this regard". לעיל ה"ש 66, בעמ' 37–38.

70 נהוג לייחס את האזכור הראשון ל־Douglas Laney, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, Gartner, Working Paper No. 949, (2001), https://studylib.net/doc/8647594/3d-data-management--controlling-data-volume--vel...ocity--an. ראו גם: VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK AND THINK (2013).

71 ובניסוחם של כהנא ושוורץ אלטשולר, לעיל ה"ש 57, בעמ' 53: "נתוני עתק הם הדלק המאפשר את פיתוחה ואת שכלולה [של הבינה המלאכותית]".

72 ראו נחום קרייתי ואח' "קניין גופני ולמידה עמוקה" **ספר יורם דנציגר** 567, 567–568 (לימור זר-גוטמן ועידו באום עורכים 2019); EDPS, לעיל ה"ש 15, בעמ' 4; Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?* (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914, בעמ' 17.

73 Mitrou, שם, מצטטת את דוח הרשות הצרפתית להגנת הפרטיות, שלפיה "AI and Big Data are indissociable".

מפוקחת (supervised), או ללא מגע יד אדם, בלמידה לא־מפוקחת (unsupervised).[74] לשם הדיון הנוכחי, די לנו להבחין כי כדי לשפר את הבינה המלאכותית נחוץ מידע איכותי בהיקף גדול: אם המערכת תלמד מידע שגוי או מוטֶה, התוצאה תשעתק את ההטיות.[75] למשל, מערכת שתנחה מה השכר הראוי לעבודה, ותתבסס לשם כך על המצב הקיים שבו שכרן של נשים נמוך משכרם של גברים, תשעתק את ההטיה.

חמישית, המרכיב האחרון בהגדרת העבודה הוא **התוצאה**. בינה מלאכותית מאתרת דפוסים על בסיס ניתוח המידע שלמדה. הדפוסים כלליים, ולאחר שנוצרו ניתן להשוות אליהם נתונים של אדם מסוים. לשם כך דרוש מידע נוסף על אותו אדם. התוצאה תהיה מסקנה לגבי האדם: זיהוי, סיווג, הערכה, או תחזית לעתיד. הפלט המוסק אינו בהכרח מידע שנמסר על ידי האדם במפורש או תוך כדי התנהגותו, וגם אינו חוות דעת על האדם. התחזית מוסקת מתוך המידע של האדם המסוים, תוך השוואה למידע שנאסף על אנשים רבים אחרים.[76] מאפיין זה ממחיש שהמידע שלנו כרוך במידע על אחרים. הפרטיות היא ביסודה זכות של היחיד. אולם, בנתוני העתק ובעיבודם באמצעות בינה מלאכותית, המידע של אנשים נפרדים נכרך יחדיו.[77]

להמחשה: טלו אלגוריתם שפותח על בסיס מידע רפואי של אנשים רבים, בהיקף ניכר וממגוון מקורות (בדיקות דם, הדמיות רפואיות, רישומים רפואיים), כלומר מדובר בנתוני עתק. האלגוריתם למד (או שלימדו אותו) "לקרוא" מידע רפואי ולנתחו, והתוצאה היא נוסחה כללית בנוגע למחלה מסוימת. כעת, נבחן נתונים של אדם מסוים מול האלגוריתם. התוצאה היא הערכה, בהסתברות מסוימת, האם אדם נמצא במצב רפואי מסוים. כך פותח אלגוריתם לזיהוי סרטן המעי הגס, במחקר שנערך בישראל.[78] בדומה, חברות אשראי לומדות את דפוסי השימוש בכרטיסיהן לגבי כלל המשתמשים ובאופן פרטני לכל מחזיק בכרטיס, וכך יודעות לזהות חריגה שעשויה ללמד על גנבת הכרטיס.[79]

*

74 ראו כהנא ושוורץ אלטשולר, לעיל ה"ש 57, בעמ' 38–40.

75 לדיון בהטיות הבינה המלאכותית ראו כהנא ושוורץ אלטשולר, שם, בעמ' 179; Anupam Chander, *The Racist Algorithm?*, 115 MICH. L. REV. 1023 (2017).

76 לדיון בהסקת מידע על אנשים אחרים ראו Alicia Solow-Niederman, *Information Privacy and the Inference Economy*, 117 NW. U.L. REV. 357 (2022).

77 ראו Solon Barocas & Karen Levy, *Privacy Dependencies*, 95 WASH. L. REV. 555 (2020).

78 Yaron Kinar et al., *Development and Validation of a Predictive Model for Detection of Colorectal Cancer in Primary Care by Analysis of Complete Blood Counts: A Binational Retrospective Study*, 23 J. AM. MED. INFORMATICS ASS'N 879 (2016).

79 ליישומי הבינה המלאכותית בתחום זה ואתגריה ראו Eleanor Mill, et al., *Opportunities in Real Time Fraud Detection: An Explainable Artificial Intelligence (XAI) Research Agenda*, 14(5) INT'L J. ADVANCED COMPUT. SCI. & APPLICATIONS 1172 (2023).

לסיכום חלק זה, מובן מדוע המחוקק לא הזכיר עד כה במפורש את הבינה המלאכותית בחקיקה: החוק קדם להופעת הטכנולוגיה הזו, וממילא אין זה נבון להתייחס לטכנולוגיה מסוימת במפורש, ובייחוד כאשר מדובר בשם כולל לטכנולוגיות מגוונות. היעדר ההתייחסות המפורשת אין פירושו שדיני הפרטיות אינם חלים על היבטים רלוונטיים של הטכנולוגיה. להפך. הם חלים גם חלים, והשאלה היא כיצד. לכן, נדרשת פרשנות. פעמים רבות, לנוכח טכנולוגיה משבשת, המשימה מורכבת, ועלינו לפנות לעקרונות יסוד, כלומר לרובד הערכי, לתכלית דיני הפרטיות. פרדיגמת המחקר של משפט וטכנולוגיה מנחה אותנו לברר את יסודות הזכות לפרטיות ואת מאפייני הבינה המלאכותית ולהציבם אלה מול אלה, כדי לברר מהי בעצם הפגיעה בפרטיות שהבינה המלאכותית מביאה.

ב. בירור הפגיעה בפרטיות

הזירה ערוכה: הטכנולוגיה של הבינה המלאכותית ניצבת בצד אחד, והיא ערכית ביסודה. אפיינו כמה מהיבטיה המידעיים. הטכנולוגיה מתנהלת לפי מחזור חיים של תכנון המערכת, איסוף מידע, עיבודו, ושימושים בו. התהליך מעגלי – המערכת מוזנת בתובנות שהפיקה, מעבדת אותן, וחוזר חלילה. התיאור הלינארי כאן הוא לצורכי בהירות ההסבר. מול הטכנולוגיה ניצבת הזכות לפרטיות, גם היא ערכית ביסודה ומבקשת לממש תכליות שונות. הצדקות הפרטיות מיתרגמות למושג של פרטיות כשליטה: היכולת של אדם לשלוט במידע שעל אודותיו. דיני הפרטיות המידעית מנחים אותנו לעקוב אחרי שלבים בחיי המידע. האם הבינה המלאכותית פוגעת בפרטיות, ומהי בדיוק הפגיעה? בספרות המחקר הראשונית ובדוחות יש הערות בנוגע לפגיעת הבינה המלאכותית בפרטיות, אולם בדרך כלל אלה אבחנות כלליות. חלק זה מתאר את הניסיונות האלה ובוחן את שלבי החיים העיקריים של המידע בבינה המלאכותית.[80]

1. ניסיונות בירור ראשונים

יש הסכמה רחבה בקרב חוקרים ומחברי הדוחות השונים שבינה מלאכותית פוגעת בפרטיות, אולם באשר לבירור הפגיעה עצמה – יש גיוון ופיזור ניכרים.

המשפטן ריין קאלו (Calo) היה מהראשונים שהצביעו על אתגרי הפרטיות בבינה המלאכותית.[81] עוד בשנת 2010 טען שהבינה המלאכותית מקילה את מלאכתם של מבצעי מעקב,[82] והצביע על היקפי המידע האישי בידי תאגידים.[83] במאמר נוסף הדגיש

80 גם חוק הבינה המלאכותית האירופי מדגיש שההגנה על הפרטיות בהקשר הזה צריכה להישמר במהלך כל מחזור החיים של המידע במערכת הבינה המלאכותית. ראו AIA, לעיל ה"ש 7, בהערת הפתיח 69.

81 Ryan Calo, *Peeping HALs: Making Sense of Artificial Intelligence and Privacy*, 2(3) EUR. J. LEGAL STUD. 168 (2010).

82 שם, בעמ' 168–169.

83 שם, בעמ' 184.

את יכולת הטכנולוגיה לזהות דפוסי פעולה אנושיים בלי שנושאי המידע מבינים מה נעשה במידע שעל אודותיהם, ואת חוסר השוויון בנוגע למידע.[84] המשפטנית ליליאן מיטרו (Mitrou) בחנה את התאמת ה־GDPR לבינה מלאכותית, והצביעה באופן כללי על מעקב, על פגיעה באוטונומיה של האדם, ועל פגיעה בשליטה של האדם במידע האישי.[85] היא הטילה ספק בהתאמת עקרון ההודעה והסכמה לבינה מלאכותית,[86] משום שתוצרי העיבוד אינם ידועים מראש.[87] המשפטנים קארל מנהיים (Manheim) וליריק קפלן (Kaplan) הצביעו על פגיעתה של הבינה המלאכותית בפרטיות חוקתית ובפרטיות מידעית. הם התייחסו לקטגוריה החוקתית האמריקנית של פרטיות בהחלטות,[88] לחשש ממעקב,[89] לאיסוף מידע למטרות שלא הודעו לנושאי המידע מראש,[90] ולחשש מזיהוי מחדש של מידע אנונימי.[91]

המשפטנית סנדרה ווכטר (Wachter) והפילוסוף ברנט מיטלשטאדט (Mittelstadt) הדגישו את יכולת הבינה המלאכותית להגיע לתובנות על בני אדם – אפיון ההעדפות שלהם, זיהוי נתונים רגישים כמו מוצא אתני, העדפה מינית ודעות, ותחזית על התנהגותם העתידית.[92] את החשש ניסחו כפגיעה בשליטה של אנשים באופן שבו המידע על אודותיהם משמש להסקת מסקנות לגביהם.[93] הם הציעו להכיר בזכות חדשה, ל-reasonable inferences, במקרים שבהם לתחזית שעל הפרק יש השלכות משמעותיות.[94] לפי הצעתם, הזכות מטילה חובה על מעבדי המידע להצדיק, לפני השימוש במידע, שהמסקנה העתידית תהיה סבירה. בדומה, פילוסוף הטכנולוגיה ראיינר מולהוף (Mühlhoff) הציע שהפרטיות תחול גם על תחזיות.[95] טענתו נשענת על האופן שבו הבינה המלאכותית פועלת: על בסיס לימוד ועיבוד של נתוני עתק, הטכנולוגיה מאתרת דפוס פעולה מסוים שלא היה נגיש קודם לעין האנושית, וכעת הדפוס הזה מיושם על אדם אחר שיש לו מאפיין דומה. דוגמה פשוטה היא מערכת המלצות: על בסיס ניתוח רכישות ספרים באמזון או הרגלי צפייה בנטפליקס, הבינה

84 Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 420–421 (2017).

85 ראו Mitrou, לעיל ה״ש 72, בעמ׳ 22–23.

86 שם, בעמ׳ 37.

87 שם, בעמ׳ 39.

88 Karl Manheim & Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J.L. & TECH. 106, 112 (2019).

89 שם, בעמ׳ 126.

90 שם, בעמ׳ 121.

91 שם, בעמ׳ 128.

92 Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2 COLUM. BUS. L. REV. 494 (2019).

93 שם, בעמ׳ 497.

94 שם, בעמ׳ 500–501.

95 Mühlhoff, לעיל ה״ש 67, בעמ׳ 4, 6–8.

המלאכותית מזהה דפוסי קריאה וצפייה, ובהתאם, כאשר היא מזהה אדם שהמידע עליו לא נכלל בנתוני העתק הגולמיים אבל יש לו מאפיין דומה, הטכנולוגיה משערת שהוא יתעניין במוצרים דומים.[96] דוגמה רגישה יותר היא זיהוי עמדה פוליטית של אדם או מצבו הרפואי.[97] המשפטנית אלישיה סולו-נידרמן (Solow-Niederman) מחדדת חשש זה בנוגע להסקת מידע על אנשים אחרים שלא מסרו את המידע שלהם בעצמם.[98]

בדוחות המדיניות יש התייחסויות מגוונות. חוקרים במכון ברקמן-קליין באוניברסיטת הרווארד סקרו בשיטתיות דוחות של גורמים ממשלתיים, פרטיים וארגוני חברה אזרחית שעסקו בהיבטים שונים של בינה מלאכותית, ומצאו שהפרטיות נזכרת כמעט בכולם, ויותר מכל עיקרון אחר.[99] עם זאת, הדוחות מזכירים היבטים שונים של הפרטיות: כמעט בכולם נמצאו אמירות כלליות על חשיבות הפרטיות;[100] חלקם הזכיר את הפרטיות כשליטה, אולם באופן כולל ותוך שונות בנוגע לתיאור השליטה הרצויה ועוצמתה;[101] ולאחר מכן, נזכרו עקרונות הגנת מידע שונים בשכיחות הולכת ופוחתת: הסכמה, הנדסת פרטיות, הגבלת עיבוד מידע, זכות לתיקון מידע, זכות למחיקת מידע.[102]

דוח הממשלה בישראל בחן את אתגרי הפרטיות שיש בבינה המלאכותית לפי הדין הקיים, כלומר לפי עקרונות הגנת המידע האישי. הדוח קובע שנתונים אישיים הם מידע אישי, וחוק הגנת הפרטיות חל על איסופם,[103] ומציין את הקשיים בחובת היידוע, הסכמה, שקיפות, עקרון מזעור המידע, עקרון צמידות המטרה, סיכוני אבטחת מידע, והחשש מהסקת מידע רגיש.[104] הדוח מצביע על פתרונות ארגוניים-טכנולוגיים ומשפטיים אפשריים.[105]

הסקירה מעלה כמה מסקנות. **ראשית**, היעדר האחידות בזיהוי סוגיות הפרטיות מצביע על צורך ממשי בבירור ודיוק הפגיעה בפרטיות. **שנית**, כדי לבצע זאת, יש צורך באמת מידה ברורה להערכה. יש שתי אמות מידה רלוונטיות. האחת, נורמטיבית, לפי עקרונות היסוד של הפרטיות, והאחרת, לפי כללים פוזיטיביים בדין. חשוב לשים לב שהדין המצוי אינו בהכרח הדין הראוי. בספרות ובדוחות יש מי שמתייחסים לעקרונות הגנת המידע האישי כמורה דרך ערכי.[106] **שלישית**, אנו רואים את קשרי הגומלין בין

96 להרחבה ראו מיקי זר "'אנשים שקראו מאמר זה התעניינו גם ב...': על הקשר בין פרטיות לפרופיילינג" **משפט, חברה ותרבות** ב 69 (מיכאל בירנהק עורך 2019).

97 מוכרת הדוגמה של רשת טארגט האמריקנית שזיהתה שנערה בהיריון, לפי שינויים בהרגלי הצריכה שלה. ראו שם, בעמ' 73.

98 Solow-Niederman, לעיל ה"ש 76.

99 Fjeld, לעיל ה"ש 17.

100 שם, בעמ' 26. מחברי המסמך מצאו אזכורים כאלה ב־92% מהמסמכים שבחנו (לפירוט השכיחות ראו שם, בעמ' 21).

101 שם, בעמ' 22–23. הפרטיות כשליטה אוזכרה ב־42% מהמסמכים (שם, בעמ' 21).

102 לפירוט השכיחויות ראו שם, בעמ' 21, ולסקירה עצמה בעמ' 22–27.

103 דוח הממשלה, לעיל ה"ש 14, בעמ' 77.

104 שם, בעמ' 12, 24, 78–80.

105 שם, בעמ' 89–90.

106 ראו את דוח ה־ICO, לעיל ה"ש 15.

המשפט לטכנולוגיה. לעיתים המשפט מוצג כמנחה את הטכנולוגיה, לעיתים כמפגר אחריה, אבל בה בעת חלק מהמחקרים מציע לפתח את המשפט, בתגובה לטכנולוגיה, למשל ההצעות להכיר בזכויות חדשות בנוגע לתחזיות התנהגותיות.

2. תחנות בחיי המידע

דיני הפרטיות בכלל, והפרטיות המידעית בפרט, צועדים עקב בצד אגודל אחרי שלבי החיים של המידע. בכך הם מבקשים להבטיח את שליטת האדם במידע שעל אודותיו, גם אחרי שהמידע יצא מחזקתו. אומנם, לאחר שהמידע נמסר (במפורש ובהסכמה מדעת) או נאסף (למשל אגב שימוש ביישומונים) קשה יותר לאדם לשלוט במידע האישי. לשם כך המשפט מתגייס, ומפקיד בידי נושא המידע כוח משפטי. המשפט מייצר נקודות מפגש בין נושא המידע לבין המידע האישי שעל אודותיו,[107] למשל הזכות לעיין במידע, לדרוש את תיקונו, ובדין האירופי הזכויות של נושא המידע לחזור בו מההסכמה הראשונית ולדרוש את מחיקת המידע. חובות שמוטלות על מעבדי המידע מצטרפות: עקרון מזעור המידע נועד להבטיח שלא ייאסף מידע עודף; עקרון צמידות המטרה נועד להבטיח שהמידע ישמש רק למטרה הראשונה. כלי אכיפה מצטרפים. כמובן, היישום בפועל נתקל באתגרים. דיני הפרטיות, רגולטורים ובתי משפט עוסקים בסוגיות כאלה באופן שוטף.

שני שלבי המידע המרכזיים שרלוונטיים לבינה המלאכותית הם איסוף המידע והשימושים בו. לעיתים יצטרף גם שלב של מות המידע, כלומר מחיקתו והפסקת השימוש בו. לעיתים השלבים כרוכים זה בזה, למשל תחושת המעקב בשלב האיסוף מתעוררת גם בשלב השימוש. הדיון להלן מונה אתגרים שונים בכל שלב. ההפרדה היא מכשירנית ופדגוגית, לשם הבירור, כדי להבין מהי בעצם הפגיעה בפרטיות שיש בבינה המלאכותית. בהתאם, שיוך האתגרים לשלב מסוים אינו בהכרח ממצה.

(א) איסוף המידע

בשלב איסוף המידע נמסר לאדם מידע על האיסוף הצפוי ומטרותיו ("הודעה"), ויש לו הזדמנות לבחון, לשקול ולהחליט האם למסור את המידע לשם שימוש בטכנולוגיות של בינה מלאכותית ("הסכמה"). זה כמובן התסריט הפשוט ביותר. בפועל, הפעולות מורכבות בהרבה: לרוב, נאסף מידע שלא בידיעת האדם, ממידע פומבי ברשת או ממקורות אחרים, ולא תמיד מי שאוסף את המידע הוא הגורם שמעבד אותו. לענייננו, המיקוד הוא באיסוף מידע ראשוני לשם למידת מכונה. בשלב זה, למפתחים אין בהכרח עניין בזיהוי אנשים מסוימים. המכונה זוללת מידע בהיקפי עתק.

אציג חמישה אתגרי פרטיות: **ראשית**, שלב ההודעה וההסכמה נתקל בקשיים רבים גם בלי קשר לבינה המלאכותית. זהו אתגר הקופסה השחורה. **שנית**, בשל מטרתה של

107 ראו Michael Birnhack & Niv Ahituv, *Privacy Implications of Emerging and Future Technologies* (PRACTIS Project, 2013), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2364396.

הבינה המלאכותית להציע עיבודי מידע שעין אנושית מתקשה לזהות, קשה לומר מראש מהי המטרה ומהו המידע שדרוש. זהו אתגר לעקרון האחריותיות, שמגולם בעקרונות ההודעה, ההסכמה, צמידות המטרה ומזעור המידע. **שלישית**, אתגר של קציר המידע הפומבי מהרשת. מי שמפתחים את האלגוריתמים של הבינה המלאכותית זקוקים למידע בהיקף עצום לצורך לימוד ואימון המכונה. חלק לא מבוטל מהחברות קוצר את המידע ממקורות מידע פומביים. האם היותו של המידע נגיש ברשת מתיר את השימוש בו למטרות הבינה המלאכותית? **רביעית**, העובדה שהמידע מזוהה בעת איסופו אבל לא בהכרח בשלב התוצר הסופי, מעוררת שאלה כללית: האם דרושה הסכמה לשימוש כזה? זהו אתגר הזיהוי. **חמישית**, המידע נאסף מקבוצה מסוימת של אנשים, ומוחל גם על מי שאינו נכלל בה. זהו אתגר הקבוצה.

אתגר הקופסה השחורה: הודעה והסכמה

במישור הפרטיות הבין־אישית, יש להסכמה חשיבות רבה. ההסכמה מאיינת את הפגיעה בפרטיות מלכתחילה.[108] ההסכמה צריכה להיות מדעת וחופשית. בלעדיה, מדובר בפגיעה בפרטיות. במעבר לפרטיות המידעית, היקפי המידע גדולים יותר, פערי הכוח בין הצדדים גדלים, והיכולת לנהל משא ומתן קטנה. לכן, החוק קובע חובת הודעה: הגורם שאוסף מידע חייב להודיע לנושא המידע על כוונתו לאסוף מידע ולפרט את מטרות השימוש.[109] בדינים שונים יש פירוט נוסף בדבר תוכן ההודעה.[110] ההסכמה ניתנת על בסיס ההודעה, ולכן ההודעה כובלת את אוסף המידע באשר לשימושים במידע שנאסף. אם נמסר לנושא המידע המיועד שהמידע ישמש למטרה אחת, אין להשתמש במידע למטרות שונות. זהו עקרון צמידות המטרה.[111] בסביבה הדיגיטלית, ההודעה מיתרגמת ל"מדיניות פרטיות", וכאן הקשיים מתחילים.

מחקרים אמפיריים מצאו שמדיניות הפרטיות אינה משיגה את יעדה ליידע את נושא המידע באופן מספק. יש אינספור מדיניויות כאלה באתרים וביישומונים, והן מתעדכנות לעיתים מזומנות. מדיניות הפרטיות הנפוצה ארוכה,[112] שפתה משפטית ואינה נגישה.[113]

---

108 ס' 1 לחוק הגנת הפרטיות קובע כי "לא יפגע אדם בפרטיות של זולתו ללא הסכמתו". ס' 3 מגדיר הסכמה – "הסכמה מדעת, במפורש או מכללא".

109 בישראל, ס' 11 לחוק הגנת הפרטיות; באיחוד האירופי, ס' 13 ל־GDPR.

110 ס' 13(2) ל־GDPR דורש פירוט של משך החזקת המידע, פירוט הזכויות לעיין במידע ולתקנו, לחזור מההסכמה ולהגיש תלונה.

111 בישראל, ראו את ס' 2(9) ואת ס' 8(ב) לחוק הגנת הפרטיות, שעוסק במטרות של מאגר מידע. באיחוד האירופי, ראו ס' 5(1)(b) ל־GDPR.

112 Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 J.L. & POL'Y INFO. SOC'Y, 543, 563 (2008). המחקר מצא עוד בשנת 2005 שקריאת מדיניות פרטיות באתרים שונים תארך בממוצע 244 שעות בשנה.

113 ראו מחקר שמצא ירידה נמשכת ברמת ההבנה של המדיניויות בארצות הברית: George R. Milne et al., *A Longitudinal Assessment of Online Privacy Notice Readability*, 25(2) J. PUBLIC POL'Y & MKTG. 238 (2006). ומחקר בנוגע לאתרים בריטיים: Shmuel I.

בארצות הברית, משתמשים רבים בטוחים שהכותרת "מדיניות פרטיות" משמעה שהאתר שהם גולשים בו מגן על פרטיותם, ולא היא:[114] המדיניות מסבירה כיצד האתר פוגע בפרטיות. המדיניות הנפוצה נוטה לנצל קשיים קוגניטיביים של המשתמשים,[115] ובכל מקרה, מדובר בחוזה אחיד. בפועל, עקרון ההודעה וההסכמה נכשל, ומכשיל את המשתמשים.[116] לנוכח מצב זה, חוקרים אמריקניים הציעו לוותר על הדרישה במקרים הרגילים ולהשאירה רק במקרים מיוחדים,[117] או לוותר עליה כליל.[118]

הקשיים מתעצמים בנוגע לבינה המלאכותית. *ראשית*, בשלב הזה מדובר בטכנולוגיה חדשה, ויש להניח שהבנת הציבור לגביה – מעטה. גם לגבי טכנולוגיות אחרות אין צורך שנדע מה בדיוק קורה בנבכי התוכנה, אבל יש לנו מושג כללי. לגבי בינה מלאכותית, נראה שאין כיום הבנה כזו. *שנית*, השימושים במידע אינם בהכרח ידועים בשלב האיסוף. הסכמה שתינתן (לשם למידת המכונה), בהכרח תהיה כוללנית. לעיתים מפתחי הטכנולוגיה עצמם ולקוחותיהם אינם יודעים בדיוק למה תשמש בהמשך.

כיצד מתמודדים תאגידי המידע עם האתגר הזה? מדיניות הפרטיות של גוגל ציינה עד לראשית מארס 2024: "we may collect information that's publicly available online or from other public sources to help train Google's AI models and build products and features like Google Translate, Gemini Apps, and Cloud AI capabilities".[119] בחנו את עצמכן: האם הבנת מה המשמעות? הנוסח הנוכחי (מיום 28.3.2024) מסתפק בכך שגוגל עשויה לאסוף מידע פומבי ברשת, מבלי לציין את מטרות השימוש. במדיניות הפרטיות של Meta, שמפעילה בין היתר את פייסבוק ואינסטגרם, מצוין שהמידע האישי שנאסף משמש בין היתר כדי "[to] understand and enable creation of content like text, audio, images and videos, including through artificial intelligence technology we provide".[120] באתר "אמזון" לא מצאתי התייחסות ישירה לבינה מלאכותית. נזכרות שם מטרות, שיש להניח שלשמן מופעלת בינה מלאכותית – המלצות, התאמה אישית, וזיהוי

Becher & Uri Benoliel, *Law in Books and Law in Action: The Readability of Privacy Policies and the GDPR*, *in* CONSUMER LAW & ECONOMICS 179 (Klaus Mathis & Avishalom Tor eds., 2021).

114 Joseph Turow et al., *Persistent Misperceptions: Americans' Misplaced Confidence in Privacy Policies, 2003–2015*, 62 J. BROAD. & ELEC. MEDIA 461 (2018).

115 Luiza Jarovsky, *Improving Consent in Information Privacy through Autonomy-Preserving Protective Measures (APPMs)*, 4 EUR. DATA PROT. L. REV. 447 (2018).

116 ראו למשל Elettra Bietti, *Consent as a Free Pass: Platform Power and the Limits of the Informational Turn*, 40 PACE L. REV. 310, 316 (2020).

117 Neil Richards & Woodrow Hartzog, *The Pathologies of Digital Consent*, 96 WASH. U.L. REV. 1461 (2019).

118 לוין, לעיל ה"ש 66, בעמ' 222.

119 הנוסח נמצא בהשוואה בין נוסחים קודמים של המדיניות: https://policies.google.com/privacy/archive/20240208-20240304.

120 מדיניות הפרטיות של Meta (26.6.2024): https://www.facebook.com/privacy/policy.

הונאות.[121] חברת זום מציינת במדיניות הפרטיות שלה שהיא אינה אוספת מידע לצורכי למידת מכונה של בינה מלאכותית, אלא אם המשתמש הפעיל שירות מיוחד (intelligent" feature"), ואז החברה עשויה להשתמש בבינה מלאכותית למטרה זו בלבד.[122] מהם השירותים האלה? מהו השימוש? לזום פתרונים. חברת הבינה המלאכותית OpenAI, שמפעילה את ChatGPT, מבהירה כיום למשתמשיה באיחוד האירופי שהיא אוספת מידע ללימוד מכונה לשם הפעלת המערכת. החברה מאפשרת לבקש את הפסקת השימוש במידע שלהם ללימוד המכונה, וזאת בעקבות דרישת הרשות האיטלקית להגנת הפרטיות.[123]

הדוגמאות ממחישות את הקושי לשקף למשתמשים את המשמעות של איסוף מידע לצורכי למידת מכונה. החברות מסתפקות בתיאור כללי מאוד. קשיי ההודעה וההסכמה הרגילים מתעצמים. זהו אתגר הקופסה השחורה.[124]

המענה המשפטי הוא חובת שקיפות, שהיא ערך מרכזי בסל הכלים שהמשפט מציע לנושאי המידע.[125] ה־GDPR קובע חובת הודעה כללית שמתייחסת למערכות אוטומטיות: אוסף המידע צריך לספק לנושאי המידע מידע משמעותי על הלוגיקה של המערכת האוטומטית.[126] עוד בדירקטיבה להגנת מידע אישי משנת 1995 נכללה הזכות שלא להיות כפוף להחלטה אוטומטית. הזכות מופיעה גם בסעיף 22 ל־GDPR, שקובע שלנושא המידע זכות שההחלטות לגביו לא תתקבלנה רק על בסיס מערכת אוטומטית. במילים אחרות, אנחנו זכאים שבני אדם יחליטו על זכויות וחובות שלנו, ולא מכונות. הניסוח בסעיף עצמו צר יחסית, ואינו דורש במפורש שקיפות לגבי אופן פעולת האלגוריתם. עם זאת, הערת המבוא 71 של ה־GDPR, שיש לה כוח פרשני, מסבירה שמעבד המידע צריך לגלות את העובדה שההחלטה התקבלה באופן אוטומטי, לציין את הזכות להחלטה אנושית, ואת הזכות לקבל הסבר לגבי אופן קבלת ההחלטה. שילוב ההוראות עורר ספק בדבר היקפה של הזכות בדין הפוזיטיבי באיחוד האירופי.[127]

סך ההוראות הוא זכות להסבריות (explainability). אבל, מה בדיוק יש להסביר, ואיך? ההסבר עשוי להיות מורכב מאוד לנושא המידע הרגיל, שממילא הוא עמוס,

121 מדיניות הפרטיות של Amazon (31.3.2024): https://www.nevo.co.il/shorturl/9556C2.

122 מדיניות הפרטיות של Zoom (19.7.2024): https://explore.zoom.us/en/privacy.

123 ראו Natasha Lomas, *ChatGPT Resumes Service in Italy after Adding Privacy Disclosures and Controls*, TECHCRUNCH (Apr. 28, 2023) https://techcrunch.com/2023/04/28/chatgpt-resumes-in-italy.

124 ראו Pasquale, לעיל ה"ש 21.

125 ס' 5(1)(b) ל־GDPR קובע שמידע אישי צריך להיות מעובד באופן חוקי, הוגן ובשקיפות.

126 ס' 13(2)(f) ל־GDPR.

127 ראו Bryan Casey et al., *Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise*, 34 BERKELEY TECH. L.J. 143 (2019). בית הדין האירופי בחן שאלה מוקדמת יותר לגבי אופייה של החלטה אוטומטית, וקבע ששירות של הערכת ההסתברות האם אדם יחזיר הלוואה הוא החלטה אוטומטית כזו. ראו Case C-634/21, OQ v. Schufa Holding, AG, ECLI: EU:C:2023:957 (Dec. 7, 2023).

ומתקשה בעיבוד מידע מורכב.[128] מעבד המידע מבקש לא לחשוף סודות מסחריים,[129] ובכלל, גם הוא יתקשה במקרים רבים להסביר את הלוגיקה של המערכת. הרי זה סודה של הבינה המלאכותית: מוֹתַר הטכנולוגיה מהאדם. נציב הגנת הפרטיות הבריטי סבור שדרוש הסבר לא־טכני על הסיבות שהובילו להחלטה מסוימת של מערכת טכנולוגית, והתייחסות לסוגיות מעטפת: מי אחראי למערכת ולמי אפשר לפנות, איזה מידע נשקל בקבלת ההחלטה, אילו אמצעים ננקטו להבטיח שהמערכת אובייקטיבית ואמינה.[130] מהנדסים מתחבטים בסוגיה גם הם, לפי אמת מידה של היקף ההסבר (האם צריך להסביר את כל המערכת, או רק החלטה מסוימת?), ולפי ממד הזמן (האם די בהסבר בדיעבד, או שיש להקדים הסבר מראש?), ובהתאם, מנסים לפתח שיטות מתקדמות להסברים.[131] וסרמן־רוזן, בכרך ואלקין־קורן הציעו לבחון את הטעמים לדרישת ההסבריות לפי השחקנים המעורבים: מקבל ההחלטה, מושא ההחלטה, והחברה בכללותה, והצביעו על השוני בין המשמעות של "הסבר" בהקשר של יחסים בין בני אדם, להקשר של המכונה.[132]

אתגר האחריותיות: צמידות המטרה ואתגר מזעור המידע

הזכרתי את עקרון צמידות המטרה, שמשקף את שליטת נושא המידע במידע שעל אודותיו. הצד המשלים הוא חובה שמוטלת על מעבד המידע: אם אדם הסכים למסירת המידע למטרה מסוימת, מותר להשתמש במידע רק לאותה מטרה, ובאיחוד האירופי – גם למטרה תואמת (compatible). עיקרון נלווה הוא שיש לאסוף רק את המידע שדרוש למימושה של המטרה הלגיטימית. זהו עקרון מזעור המידע (מינימיזציה).[133] אם לשאול מונח חוקתי, הדרישה היא למידתיות: אין לאסוף מידע במידה העולה על הנדרש.

יישום שני העקרונות לבינה המלאכותית מעורר קושי ניכר. גם מפעילי הבינה המלאכותית לא תמיד יודעים מה מטרתם באיסוף המידע ובעיבודו. למשל, במודלים לעיבוד שפה טבעית – כלומר מערכות שיודעות לקרוא טקסטים שנכתבו בידי בני אדם,

128 למשל Chun-Hua Tsai & John M. Carroll, *Logic and Pragmatics in AI Explanation*, *in* XXAI: BEYOND EXPLAINABLE AI 387 (Andreas Holzinger et al. eds., 2022).

129 בית משפט בגרמניה קבע שלאדם יש זכות לעיין במידע ששימש בסיס להחלטה בנוגע לאשראי שלו, אבל לא לנוסחה ששימשה את מדֶרֶג האשראי. ראו Philipp Hacker & Jan-Hendrik Passoth, *Varieties of AI Explanations Under the Law: From the GDPR to the AIA, and Beyond*, *in* XXAI, שם, בעמ׳ 348.

130 ראו The Alan Turing Institute & ICO, Explaining Decisions made with AI, 21 (2022).

131 ראו Hacker & Passoth, לעיל ה״ש 129.

132 Hofit Wasserman Rozen et al., *Lost in Translation: The Limits of Explainability in AI* (2023), https://ssrn.com/abstract=4531323.

133 ס׳ 5(1)(c) ל־GDPR קובע שמידע אישי צריך להיות "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')".

המפעילים יוכלו לומר שהמטרה היא "לימוד המכונה", אולם זה יהיה תיאור כוללני שמעיד על הטכניקה, לאו דווקא על מטרת העיבוד עצמו. לעיתים יוכלו לומר מעט יותר על המטרה – למשל, כפי שאמזון מציינת את מערכת ההמלצות שלה.[134] אבל כאשר מפעילי המערכת מבקשים ממנה לאתר דפוסים והקשרים שהעין האנושית אינה מזהה, קשה לומר יותר מאשר "אנחנו מבקשים מידע כדי ללמוד על הלא נודע". בלי לדעת מהי בדיוק המטרה, אי אפשר לבחון האם המידע שנאסף מספיק, או חורג מהדרוש.

במידה רבה האתגרים אינם חדשים, ומוכרים מהדיון בנתוני עתק. שם, נאסף מידע בהיקף ניכר מאנשים רבים בלי שהמטרה מוגדרת ובלי שהיקף המידע הדרוש ברור, כדי לגלות מידע חדש ולהשתמש בו.[135] במובן הזה, אתגר הבינה המלאכותית חופף לאתגרי הפרטיות של נתוני עתק.

בעיה נוספת היא של מידע עודף. החשש הוא שתוך כדי איסוף מידע למטרה אחת, לגיטימית, שניתנה הסכמה לגביה, יתגלה מידע חדש. גם בעיה זו אינה חדשה. היא מוכרת, למשל, מתחום המידע הביומטרי. מידע שנאסף בצילומי וידאו (במצלמות "אבטחה") ומשמש מזהה חד־ערכי לזיהוי, חושף מידע על מין, גזע ומוצא אתני, מבנה גוף, מצב רפואי ועוד.[136] האם מותר לשמור את המידע ולהשתמש בו למטרה שחורגת ממטרת הזיהוי הביומטרי?

אתגר קציר המידע הפומבי

לפי שעה, עילת התביעה העיקרית בתביעות שנדונות בעולם בקשר למפגש של הפרטיות והבינה המלאכותית הוא לגבי קציר מידע ממקורות נגישים ברשת.[137] קציר המידע נעשה באופן אוטומטי, ואוסף (גם) מידע אישי ומזהה ללא הסכמת המשתמשים, כמו טקסטים ותמונות שפרסמו ברשת. חברת Clearview האמריקנית, למשל, קצרה מאות מיליוני תמונות מרשתות חברתיות ליצירת אלגוריתם זיהוי ביומטרי.[138] חברת Bright Data הישראלית קצרה מידע מפייסבוק ומאינסטגרם.[139]

134 ראו מדיניות הפרטיות של Amazon, לעיל ה"ש 121.

135 ראו Tene & Polonetsky, לעיל ה"ש 68, בעמ' 259–260.

136 ראו למשל דיון ביכולות ביומטריות להשיג מידע עודף – שמוצג שם כיתרון: Antitza Dantcheva et al.,*What Else Does Your Biometric Data Reveal? A Survey on Soft Biometrics*, 11 IEEE TRANSACTIONS ON INFO. FORENSICS & SECURITY 441 (2016). לדיון ביקורתי, ראו MARK ANDREJEVIC & NEIL SELWYN, FACIAL RECOGNITION 40-42 (2022).

137 ראו את התביעות בארצות הברית, לעיל ה"ש 10, ואת התביעה בישראל, לעיל ה"ש 11.

138 בעקבות תביעות בגין הפרת חוק הפרטיות הביומטרי באילינוי, הוגבל השימוש במאגר המידע. ראו Ryan Mac & Kashmir Hill, *Clearview AI Settles Suit and Agrees to Limit Sales of Facial Recognition Database*, N.Y. TIMES (May 9, 2022), https://www.nytimes.com/2022/05/09/technology/clearview-ai-suit.html.

139 ראו לעיל ה"ש 10.

לקציר המידע שני מופעים משפטיים: הראשון, קוצרי המידע מול נושאי המידע, והשני, הפלטפורמות שמהן נקצר המידע מול קוצרי המידע. באשר לממשק הראשון, הפעם, הבעיה איננה רק הודעה עמומה שנמסרה לנושאי המידע, אלא שהם לא נתבקשו כלל למסור את המידע – הוא פשוט נלקח מהם. באיחוד האירופי ובמדינות שהדין בהן דומה לדין האירופי (ובכלל זה ישראל), התרגום המשפטי של הבעיה הוא עקרון צמידות המטרה.[140] כך למשל, חברת בינה מלאכותית בקוריאה (שדיני הפרטיות בה הוכרו כתואמים לדין האיחוד האירופי) קצרה כעשרה מיליארד הודעות משירות שיחות פופולרי, ליצירת צ׳אטבוט.[141] רשות הגנת הפרטיות שם קבעה שמדובר בהפרה של עקרון צמידות המטרה. בדין האמריקני, התרגום המשפטי הוא שנטילת המידע פוגעת בציפיות הסבירות של המשתמשים.[142] בתי משפט בארצות הברית קבעו שקציר מידע מאתרים פומביים אינו מפר את חוק המחשבים שם, אך עלול להפר תנאי שימוש חוזיים.[143]

במילים אחרות, גם כאן, הכלים המשפטיים קיימים ומוכרים: עקרון צמידות המטרה ומבחן הציפיות הסבירות. אלה עקרונות משפטיים ניטרליים לטכנולוגיה, ואין מניעה להחילם גם לבינה המלאכותית. אם כן, האתגר אינו חֶסר בדין הקיים, אלא אתגר פרשני: האם כאשר אדם פרסם מרצון מידע במסגרת אחת, הוא מסכים לקציר המידע ושימוש בו למטרה של לימוד מכונה של גורם אחר? אני סבור שהתשובה שלילית.[144] פרטיות אינה קניין, ולכן פרסום צילום אינו ״הפקרה״ של המידע האישי, ואין להסיק ממנו הסכמה רחבה לכל שימוש. אדם שמפרסם תמונה ברשת חברתית עושה זאת כדי להיות בקשר עם חבריו שם, ותו לא. יש מגוון ניכר של משתמשים, שיש להם עמדות מגוונות, אפשרויות הפצה רבות, ונורמות חברתיות משתנות. למשל, מי שיש לו רק עשרים ״חברים״ בחשבון פייסבוק פומבי, והוא מכיר את כל העשרים באופן אישי – כלומר הם חברים ללא מירכאות – אינו דמות ציבורית ופעילותו הרשתית רגילה לחלוטין: יש להניח שהמשתמש מצפה לפרטיות, יותר ממשתמש שיש לו 5,000

140 לדיון השוואתי של הגישות האירופית והאמריקנית, ראו Javier Torre de Silva & López de Letona, *The Right to Scrape Data on the Internet: From the US Case hiQLabs, Inc. v. LinkedIn Corp. to the ChatGPT Scraping Cases: Differences Between US and EU Law*, 5 GLOB. PRIV. L. REV. 5 (2024).

141 Jasmine Park, *South Korea: The First Case Where the Personal Information Protection Act was Applied to an AI System* (Future of Priv. F., May 21, 2021) https://fpf.org/blog/south-korea-the-first-case-where-the-personal-information-protection-act-was-applied-to-an-ai-system.

142 ראו למשל אצל Calo, *Artificial Intelligence Policy*, לעיל ה״ש 84, בעמ׳ 422. מבחן הציפיות הסבירות פותח בארצות הברית בהקשר החוקתי, ראו Katz v. United States, 389 U.S. 347, 361 (1967), וזלג למשפט הפרטי.

143 הדברים עלו בנוגע לקציר מידע מהרשת החברתית לינקדאין. ראו hiQ Labs v. LinkedIn Corp., 31 F.4th 1180 (9th Cir. 2022).

144 ראו Michael Zimmer, *"But the Data is Already Public": On the Ethics of Research in Facebook*, 12 ETHICS & INFO. TECH. 313, 323 (2010).

"חברים" וחשבונו נעול. בהקשר הזה, מתערבבים מישורי פרטיות שונים: המישור של פרטיות המשתמשים זה מול זה, והמישור של פרטיותם מול הפלטפורמה המארחת, וכעת נוסף מישור של גורמים חיצוניים. בדין האמריקני אפשר להתמודד עם הסוגיה לפי מבחן הציפיות הסבירות,[145] ובדין הישראלי לפי עקרון צמידות המטרה.

המישור השני של קציר המידע הוא של הפלטפורמות המארחות – פייסבוק למשל – מול קוצרי המידע. לפלטפורמה יש אינטרסים מסחריים שלא לאפשר את קציר המידע. ראשית, הקציר מעמיס על שרתיה ללא תועלת מבחינתה; שנית, הפלטפורמה משתמשת במידע בעצמה למטרותיה ואין לה עניין לשתף את המידע עם מתחרים; ושלישית, משתמשי הפלטפורמות עשויים לבוא בטענות נגד הפלטפורמה על שהיא מפקירה את המידע שלהם לגורמים חיצוניים. האינטרסים המסחריים מתלכדים עם האינטרס של המשתמשים בפרטיות. לרשות הפלטפורמה כלים טכנולוגיים לנסות למנוע קציר מידע (למשל הטמעת אמצעי זיהוי אנושיים – captcha – המערכת שמבקשת מאיתנו לזהות למשל תמונות של גשרים מתוך סדרת תמונות), ואמצעים משפטיים שונים, כלומר דיני חוזים. לפי שעה, ידם של קוצרי המידע על העליונה.[146]

אתגר הזיהוי

כאשר המידע אינו מזהה, אין חשש לפגיעה בשליטת האדם במידע שעל אודותיו, ואז הדין אינו מופעל כלל. למשל, מידע סטטיסטי כללי אינו מזהה אנשים מסוימים. העובדה שבישראל יש כך וכך ג'ינג'ים מתייחסת למידע אישי – אולם מאחר שהמידע אינו מזהה את הג'ינג'ים, אין פגיעה בפרטיותם. כך גם לגבי מידע על מיקום: העובדה שבמרפאה מסוימת ביקרו ביום מסוים מאה איש, אינה פוגעת בפרטיותם. כאשר המידע מפורט יותר, או שניתן להצליבו עם מאגר מידע אחר, וניתן לזהות את המבקרים במרפאה, נתייחס לכך כמידע מזהה, ודיני הפרטיות יחולו. בעבר, אוספי מידע רבים התבססו על התממה, כלומר מחיקת מידע מזהה, כדי להימנע מתחולת הדין מלכתחילה. כיום, ההבנה היא שגם מידע שנראה מותמם ניתן לזיהוי.[147] השאלה כיום איננה אם ניתן לזהות את האדם מתוך מידע אנונימי לכאורה, אלא מה דרוש כדי לזהותו. זו שאלה של יכולת טכנולוגית, זמן, כסף ומוטיבציה.[148]

מפעילי הבינה מלאכותית יאמרו שהיא מניבה דפוס פעולה כללי לאחר תהליך של לימוד מכונה. הלימוד התבסס על מידע אישי, אבל התוצר עצמו אינו מזהה. במילים

145 ראו את עמדתו של סולוב, בהקשר האמריקני, לעיל ה"ש 24, בעמ' 25–28. הוא כופר בהבחנה הבינארית בין "פרטי" ל"ציבורי".

146 ראו עניין Meta, לעיל ה"ש 10.

147 Paul Ohm, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, 57 UCLA L. Rev. 1701 (2010).

148 דן חי "פרטיות וטכנולוגיה בישראל" **טכנולוגיות מתפרצות**, לעיל ה"ש 66, בעמ' 241, 277, מציע מבחן של סבירות. להתממת מידע במגזר הציבורי בישראל ראו שרון בר־זיו וטל ז'רסקי "פרטיות במשבר זהות: אסטרטגיות הסדרה בעידן ההתממה" **משפט, חברה ותרבות** ב 125 (מיכאל בירנהק עורך 2019).

אחרות, הטענה אוחזת בממד הזמן: משום שבשלב השימוש בתוצרי הבינה המלאכותית אין זיהוי של האנשים שהמידע על אודותיהם שימש ללימוד המכונה, מסיקים הטוענים, אין תחולה לדיני הגנת הפרטיות. גם טענה זו התעוררה עוד לפני פרוץ הבינה המלאכותית, למשל בנוגע לשימוש במידע רפואי לצורכי מחקר. מידע נוצר אגב טיפול רפואי. כאשר אנחנו משוחחים עם הרופאה, נבדקים, מקבלים טיפול, צורכים תרופות וכדומה, המטרה ברורה: המידע נמסר ונוצר כדי לקבל טיפול מיטבי. אלא, שלמידע יש ערך עצום למחקר אקדמי ומסחרי. האם ניתן להשתמש במידע שנוצר ונמסר באופן מזוהה, לאחר שהותמם, והשימוש בו נעשה באופן אגרגטיבי? דרך אחת לפתרון היא בחקיקה מפורשת. זה הילוכו של ה־GDPR, שקובע במפורש ששימוש במידע למחקר מותר בתנאים מסוימים.[149] כלומר, הדין האירופי זיהה את הפגיעה בפרטיות, את היתרונות במחקר, ואיזן בין הזכות והאינטרס המתנגשים.

בהיעדר הכרעה ברורה בדין, ראוי לומר שההסכמה משתרעת רק על הפעולה הראשונה. כל שימוש למטרה אחרת, שאינה תואמת את המטרה הראשונה – אסור, ולכן יש לבקש את הסכמת האדם גם לשימושים אנונימיים במידע שנאסף ממנו.[150] לפי עמדה נוגדת, כפי שמציג אותה דן חי, ההתממה נועדה להגן על המידע ומיטיבה עם נושא המידע, ולכן אין לראות בה "שימוש" במידע, מונח שמופיע בחוק בנוגע לעקרון צמידות המטרה.[151] אינני סבור כך. חישבו על אישה שמתנגדת להפלות בשל אמונתה הדתית, והיא עצמה חוותה הפלה טבעית. האם תסכים שהמידע האנונימי שנוצר אגב הטיפול במצבה הרפואי ישמש לשיפור טכניקות להפלה מלאכותית? שליטתה במידע משקפת את כבודה. יש להביט על מכלול מעגל החיים של המידע. ההדגשה על שלב השימוש במידע אגרגטיבי אנונימי שבו אין פגיעה בפרטיות אינה מאיינת פגיעות בשלבים קודמים. עמדת הפרטיות כשליטה מדגישה שהשליטה היא לכל אורך חיי המידע.

קושי נוסף בקשר לזיהוי מתעורר בשלב השימוש במידע, וכאן אני מקדים את המאוחר: כבר כיום יש דרכים לזהות נושא מידע מתוך פלט של בינה מלאכותית.[152] במקרה כזה, טכנולוגיות בינה מלאכותית עלולות להביא במישרין לחשיפת מידע אישי שלא נועד לפרסום, מבדיקות רפואיות וציונים של סטודנטים ועד לשיחות אישיות שנקצרו לשם לימוד מכונה.

כך או כך, מטרתי הנוכחית צרה יותר, והיא להראות שאתגר זיהוי המידע שמתעורר בבינה מלאכותית אינו חדש.

149 ס׳ 89 ל־GDPR מתיר עיבוד מידע לצורכי ארכוב ציבורי, מחקר מדעי או היסטורי ולשימושים סטטיסטיים, בתנאים מסוימים הקבועים שם.

150 Michael Birnhack, *A Process-Based Approach to Informational Privacy and the Case of Big Medical Data*, 20 THEORETICAL INQUIRIES L. 257 (2019).

151 חי, לעיל ה״ש 148, בעמ׳ 247–248.

152 ראו Reza Shokri et al., *Membership Inference Attacks Against Machine Learning Models*, 2017 IEEE SYMP. ON SEC. & PRIV. 1 (2017).

אתגר הקבוצה

גם אם אדם אינו משתף מידע אישי לשם לימוד המכונה, יש אחרים שמסרו מידע והמידע שלהם משמש להפקת הדפוס הכללי, למשל תחזית התנהגותית, שתיושם גם על זה שבחר שלא להיכלל שם. במילים אחרות: קשה להתחמק מהבינה המלאכותית. אם כך, האם עדיין יש משמעות לשליטת האדם במידע שעל אודותיו?

זהו אתגר מורכב, אבל גם הוא אינו חדש. יש סוגי מידע שמזכירים לנו שחיי האחד כרוכים בחיי האחר. במישור הפרטיות הבין־אישית, ספר שאדם חיבר על רומן שהיה לו, חשף את האישה. בהתנגשות בין חופש הביטוי של המחבר לפרטיות האישה הכריע בית המשפט בעד הפרטיות, תוך הבחנה בין ליבת הזכויות המתנגשות למעטפת שלהן.[153] במישור הפרטיות המידעית, מידע גנטי הוא מקרה מובהק שבו מידע של האחד חושף מידע על האחר. אדם שמבצע בדיקה גנטית בחברה מסחרית, פועל לפי הסכמה מדעת. אבל המידע הגנטי של האחד הוא גם המידע הגנטי של קרוביו הביולוגיים.[154] בסביבה הדיגיטלית, הבעיה תוארה באופן רחב יותר כ־Privacy dependencies.[155] כך גם בבינה המלאכותית. אין פה פתרון פשוט, למעט אולי התערבות רגולטורית כללית. אבל שוב, הבעיה אינה חדשה.

*

בשלב איסוף המידע בנוגע לבינה מלאכותית עשויים להיות אתגרים נוספים. למשל, אנשים נוטים למסור מידע (לרבות מידע אישי) רב יותר לבני אדם מאשר למכונה. ככל שהטכנולוגיה שבאמצעותה ייאסף המידע תהיה "אנושית" יותר, למשל בצ'אטבוט, יהיו משתמשים שיטעו לחשוב שהם משוחחים עם אדם (ובכך ייכשלו במבחן טיורינג), וימסרו מידע רב יותר. הדיון נועד להצביע על סוגיות מרכזיות שהתעוררו עד כה בשיח. הרשימה אינה ממצה. האתגרים – קיימים, אבל כולם כבר התעוררו במידה זו או אחרת בהקשרים קודמים. לא לכל האתגרים יש תשובות מניחות את הדעת בדין הקיים. לכן, אין מנוס אלא לשוב לעקרונות היסוד של הפרטיות, ולחפש שם את התשובה.

(ב) שימושים במידע

משנאסף המידע, הוא משמש ללמידת מכונה ולעיבוד של נתוני העתק לקראת הפקה של תוצר: איתור סטייה מהנורמה (למשל, מערכת לזיהוי טעויות במרשמי תרופות),[156]

153 ע"א 8954/11 **פלוני נ' פלונית**, פ"ד סו(3) 691 (2014).

154 ראו Ellen Wright Clayton et al., *The Law of Genetic Privacy: Applications, Implications, and Limitations*, 6(1) J.L. & BIOSCIENCES 1 (2019).

155 ראו Barocas & Levy, *Privacy Dependencies*, לעיל ה"ש 77, שהצביעו על תלות מבוססת-קשרים (קשרים חברתיים של אדם חושפים מידע על קשרים חברתיים של אחר); תלות מחמת דמיון (הסקת מידע על אדם אחר מחמת דמיונו למידע על אנשים ידועים), תלות מחמת שוני (זיהוי התנהגות שחורגת מהנורמה).

156 ראו למשל את השירות של חברת MedAware: https://www.medaware.com.

תחזית על התנהגות (למשל, הערכת מסוכנות עבריינים;[157] הערכת סיכוני אשראי),[158] המלצות (צפייה בסרט) או סיווג. לשימוש בקשר לאדם מסוים יש צורך במידע נוסף עליו. לכך נוספת מורכבות שכבר ציינתי: פעמים רבות המטרה אינה ידועה מראש. אתגר אחד בשלב זה כרוך בשלב הקודם: כאשר השימוש חורג מהמטרה שנושא המידע הסכים לה, השימוש מפר את עקרון צמידות המטרה. אדון בשלושה אתגרים נוספים: **ראשית**, בשאלת מעמדו של תוצר העיבוד: האם מסקנה או תחזית הם "מידע" על האדם? זהו אתגר העיבוד; **שנית**, שימוש בבינה מלאכותית להסקת מסקנות על אדם כרוך בשימוש במידע של אחרים ועלול ללקות בהכללות ובטעויות. זהו אתגר התחזית; **שלישית**, כאשר המדינה משתמשת בבינה מלאכותית כלפי אזרחיה, יש חשש לפיקוח ומעקב שלטוניים, שאינם ראויים במדינה דמוקרטית. זהו אתגר המעקב.

אתגר העיבוד

האם תוצר הבינה המלאכותית הוא "מידע אישי"? אם כן, דיני הגנת המידע חלים גם לגביו, ובכלל זה זכויות גישה למידע, תיקונו במידת הצורך, מחיקתו (בדין האירופי), חובות של סודיות ואבטחת מידע, ועוד. אם התשובה שלילית, נוצר נתק בין שלב האיסוף לשלב השימוש במידע. ראינו גִרסה אחת של השאלה הזו בדיון באתגר הזיהוי בשלב איסוף המידע. השאלה הייתה, האם יש צורך בהסכמת נושא המידע להתממת המידע ושימושים לא מזהים בו. כעת, השאלה רחבה יותר.

תוצרי העיבוד של הבינה המלאכותית יכולים להיות אלגוריתם ש"יודע" לזהות התאמה או סטייה, וכך למשל לזהות גנבה של כרטיס אשראי או בעיה רפואית, אמצעי לתחזית התנהגותית, וכדומה. התוצרים הם שיטות כלליות, וככאלה, אינם מידע אישי. אולם כאשר מפעילים את האלגוריתם ומיישמים אותו על אדם מסוים, בין אם המידע שלו נכלל מלכתחילה בנתוני העתק שהיו חומר הגלם בתהליך למידת המכונה ובין אם לאו – האם העיבוד הוא "מידע אישי"?

השאלה התעוררה בסביבה האנלוגית, בנוגע לחשיפת תוצאות של מבחני מיון לעבודה. מועמדים ביקשו לעיין במידע על אודותיהם שבידי מכוני המיון, כדי להבין כיצד התקבלה ההערכה לגביהם וכדי לחסוך את הצורך במבחן נוסף למעסיק אחר. המכונים מעוניינים לצמצם את חשיפת המידע בפני המועמדים מטעמיהם המסחריים. בשנת 2012 פרסם רשם מאגרי המידע הנחיה שלפיה יש למסור לנבחן את חוות הדעת שנמסרה למעסיק.[159] בפשרה שהושגה במסגרת עתירה מנהלית בנושא נקבע שמהמידע שנמסר למועמד ניתן להשמיט את "ניתוח מידת התאמתם של תכונותיו ומאפייני

157 בית המשפט העליון בוויסקונסין אישר שימוש במערכת כזו: Loomis v. Wisconsin, 881 N.W.2d 749 (Wis. 2016), *cert. denied*, 137 S.Ct. 2290 (2017).

158 לסקירת טכנולוגיות כאלה, ראו Imane Rhzioual Berrada et al., *A Review of Artificial Intelligence Approach for Credit Risk Assessment*, 2ND INT'L CONF. ON AI & SIGNAL PROCESSING (AISP) 1 (2022).

159 הנחיית רשם מאגרי המידע 2/2012 "תחולת הוראות חוק הגנת הפרטיות על פעילות מכוני מיון לקבלה לעבודה" (28.2.2012).

אישיותו של הנבחן למשרה... ובלבד שהנתונים עצמם ופענוחם לגבי תכונותיו ואישיותו של הנבחן כלולים בחומר שנמסר לנבחן".[160] עיבוד מידע זה משלב מידע על העובד עם מידע עסקי של המעסיק על המשרה.

התוצאה מבחינה בין חומרי הגלם שמסר המועמד, שהם בחזקת מידע אישי, לבין הציון שניתן לו (שאמור לייצג באופן אובייקטיבי את נתוניו), וגם הוא נחשב למידע אישי, ולבין ההערכה הסובייקטיבית של מכון המיון לגבי המועמד, שאינו נחשב למידע אישי של המועמד. נראה שניתן להחיל את ההבחנה הזו גם בנוגע לבינה המלאכותית. למשל, חישבו על טכנולוגיה מבוססת בינה מלאכותית שנועדה להערכת הסיכוי שאדם יעמוד בהתחייבויות כספיות, ומסייעת לבנקים להחליט האם לתת הלוואה למי שמבקש אותה.[161] המידע הגולמי שנאסף הוא מידע אישי של האדם (נתונים פיננסיים, בדוגמה זו), ויש לאפשר לו לעיין במידע זה. בדרך זו יוכל האדם לראות אם המידע שגוי ולבקש לתקנו, ויוכל להעריך, במידה מסוימת לפחות, מה המידע ששימש בסיס להחלטה בעניינו. דרישת ההסבריות מצטרפת כאן, והיא משמעותית ליכולת לבחון את ההחלטה שהתקבלה, ולערער עליה במידת הצורך (מהן אמות המידה שלפיהן הטכנולוגיה העריכה שהאדם לא יעמוד בהתחייבות כספית). ומה באשר לתוצר המכונה, שמותאם לאדם מסוים (חיווי מספרי שמשקף את הסיכוי שאדם יחזיר הלוואה במועדה)? כאשר בינה מלאכותית מתיימרת להיות אובייקטיבית וניטרלית, ללא מגע יד אדם, הדבר דומה יותר לציון שאדם מקבל במבחני המיון לעבודה, כלומר ייצוג אובייקטיבי שלו, ולא הערכה (אנושית) לגביו, ולכן גם הוא בבחינת מידע אישי. אם התוצאה סובייקטיבית (מרכיב שיפוטי של אדם שמקבל את ההחלטה הסופית אם לתת הלוואה), תוצר העיבוד אינו בגדר "מידע אישי". הסוגיה טעונה בירור נוסף, והמטרה כאן אינה להכריע בה אלא לזהותה כאתגר שהבינה המלאכותית מציבה לפרטיות ולהצביע על כיוון פתרון שעלה בהקשר אנלוגי.

אתגר התחזית

כמה חוקרים הצביעו על החשש מיכולת החיזוי של הבינה המלאכותית, קושי שמתעצם בגלל אתגר הקבוצה, משום שניתן לחזות התנהגות גם בלי שיתוף פעולה של האדם עצמו, לפי השוואתו לקבוצה שדומה לו במאפיינים רלוונטיים.[162] למשל, תחזית לגבי אדם שמצבו הכלכלי ידוע, על בסיס השוואת האדם לתוצר עיבוד מידע לגבי אנשים בעלי מצב כלכלי דומה. האתגר מתגבש בשלב מתקדם יותר בחיי המידע, כאשר העוסקים במלאכה זיהו דפוס פעולה מסוים וגיבשו יכולת להסיק מסקנות. מתעורר חשש לסיווג שגוי, למסקנות שגויות, ולתחזית שתסליל את האדם להתנהגות שאחרת לא היה בוחר בה.

160 עת"מ (מנהלי ת"א) 4749-04-12 **אדם מילא בע"מ נ' רשם מאגרי המידע**, הודעת הצדדים: https://www.nevo.co.il/shorturl/4FC2CE. להסכמה ניתן תוקף של פסק דין.

161 נושא זה מוסדר כיום בחוק נתוני אשראי, התשע"ו–2016. הדיון פה עקרוני וכללי, ואינו ניתוח של חוק זה.

162 ראו לעיל, דיון באתגר הקבוצה.

בתרגום לשיח הפרטיות, החשש הוא מהוצאת אדם מהקשרו. זו אחת ההצדקות לפרטיות, שנולדה בהקשר הבין־אישי. ג׳פרי רוזן (Rosen) טען שכאשר יש בידי גורם אחד מידע אישי על אודות אחר, הראשון עלול לרדד את השני לאותו פרט מידע, ולהחמיץ את מלוא עולמו.[163] הפרטיות אמורה לשמש מגן בידי האדם, למנוע את העברת המידע ואת השימוש המרדד. הדברים נכתבו אומנם בקשר ליחסים בין בני אדם ובין אדם לקהילה, אבל גם בסביבה של בינה מלאכותית יש חשש דומה, ששימוש בטכנולוגיה יביא לתחזית שמרדדת את האדם לכדי נתון אחד, אקראי, וכך תפקיע התחזית את שליטת האדם במידע האישי שלו. פתרון טכנולוגי לבעיה הוא עוד מידע – וכך תהיה לגורם המעבד תמונת־אמת של האדם הנדון. אולם בהיעדר הסכמה, יצירת פרופיל עשיר ומלא יותר של האדם תחמיר את הפגיעה בפרטיותו. הפגיעה היא בכך שמפרקים אדם לגורמים מידעיים, כפי שהסבירה מיקי זר, ומרכיבים אותו מחדש.[164] המשפט אינו כולל הוראה מפורשת בעניין הזה, אולם ההגנה על זהותו של אדם נלמדת מהגרעין של הזכות לפרטיות שקשור קשר הדוק לכבוד האדם, ומהכלים הרבים שהמשפט מעניק לאדם לקבל החלטות אוטונומיות לגבי זהותו.

אתגר המעקב

ראינו את המעקב התאגידי בשלב איסוף המידע. אוסיף כעת גם מעקב מדינתי. הפרטיות היא זכות יסוד חוקתית, וחוק־יסוד: כבוד האדם וחירותו מגביל את המדינה שלא לפגוע בפרטיות, אלא לפי פסקת ההגבלה.[165] מדינה דמוקרטית יכולה לעקוב אחרי אזרחיה רק לפי סמכויות מוגדרות, למשל סמכויות של רשויות לאכיפת חוק. כאשר המשטרה מבקשת לעקוב אחרי חשוד היא זקוקה לצו חיפוש, לצו האזנת סתר, לצו נתוני תקשורת וכדומה. כאשר המדינה מבקשת לעקוב אחרי האזרחים באופן כללי, עליה להצביע על מקור הסמכה מפורש, שבלעדיו חוקתיות המעקב מוטלת בספק. עקרון החוקיות מעוגן בפסקת ההגבלה שבחוק היסוד.[166] כך, כאשר התברר שהמשטרה מפעילה מערכת מצלמות (״עין הנץ״) בכבישים שמתעדת נהגים בלי שיש חשד לגביהם, הורה בג״ץ למדינה לנמק את פעולתה. בתגובה, הכנסת חוקקה חוק מסמיך.[167]

163 JEFFREY ROSEN, THE UNWANTED GAZE: THE DESTRUCTION OF PRIVACY IN AMERICA (2000). להמחשה בדין הישראלי, ראו ת״א (שלום י-ם) 6023/07 **אפריאט נ׳ ידיעות אחרונות בע״מ** (נבו 5.10.2008). בית המשפט קבע ששימוש עיתונאי חוזר בתמונת התובע הפכה אותו לסמל בעל כורחו.

164 זר, לעיל ה״ש 96.

165 ס׳ 7-8 לחוק־יסוד: כבוד האדם וחירותו.

166 למשל בג״ץ 6824/07 **מנאע נ׳ רשות המסים**, פ״ד סד(2) 479 (2010), ודיון, בירנהק **פרטיות חוקתית**, לעיל ה״ש 23, בעמ׳ 391–400.

167 ראו בג״ץ 641/21 **האגודה לזכויות האזרח בישראל נ׳ משטרת ישראל** (נבו 11.1.2022); חוק לתיקון פקודת המשטרה (מס׳ 40), התשפ״ד–2024.

כאשר המדינה מבקשת להשתמש באמצעי מעקב באופן ישיר היא כפופה לעקרונות החוקתיים, שאתגריהם בצידם.[168] למשל, בשימוש במצלמות בעלות יכולות ביומטריות: מדובר במצלמות מעקב שמקשרים אותן למאגר ביומטרי.[169] מערכת כזו מאפשרת לזהות אנשים במרחב הציבורי בזמן אמת. לכך אפשר לצרף מערכות בינה מלאכותית, למשל מערכת שתדע לזהות אדם שנמצא בסביבה שאינה סביבתו הרגילה.[170] למערכות כאלה ייתכנו שימושים חשובים, למשל זיהוי מחבלים לפני שיבצעו את זממם. אבל מובן שהפגיעה בפרטיות רבה. הסוגיה מערבת שאלות משפטיות אחדות, שלחלקן כבר יש תשובות: מידע ביומטרי הוא מידע פרטי;[171] המסגרת החוקתית מנחה אותנו לאתר מקור סמכות מספק, לבחון שהתכלית ראויה, ולבחון שהאמצעי שנבחר מידתי.[172] יש פה שאלת רקע חשובה: האם יש פרטיות במרחב הציבורי? הספרות השיבה בחיוב,[173] בייחוד מול המדינה. מול המדינה יש לנו חופש תנועה, חופש הפגנה, חופש התאגדות, וחירות כללית ממעקב. כמה ערים בארצות הברית אסרו על מעקב ביומטרי בתחומן.[174] ה־AI Act באיחוד האירופי אוסר על שימוש במצלמות כאלה לזיהוי אדם בזמן אמת,

168 למשל, כאשר המדינה משתמשת בטכנולוגיות של תאגידים פרטיים, יש חשש שהיא תחמוק מחובותיה החוקתיות. זו "לחיצת היד הנעלמה" שהתעוררה שנים לפני פריצת הבינה המלאכותית. ראו Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA. J.L. & TECH. 6 (2003).

169 לדיון בפן האתי של מצלמות ביומטריות במרחב הציבורי ראו Marcus Smith & Seumas Miller, *The Ethical Application of Biometric Facial Recognition Technology*, 37 AI & SOC'Y 167 (2022).

170 חוקר התקשורת אוסקר גנדי ציין חשש זה לפני שנים רבות: OSCAR H. GANDY, THE PANOPTIC SORT: A POLITICAL ECONOMY OF PERSONAL INFORMATION (1993).

171 עס"ק (ארצי) 7541-04-14 **הסתדרות העובדים הכללית החדשה מרחב המשולש הדרומי – עיריית קלנסווה** (נבו 15.3.2017).

172 לדיון במתווה החוקתי בנוגע לפרטיות ראו בירנהק, **פרטיות חוקתית**, לעיל ה"ש 23.

173 ראו Joel R. Reidenberg, *Privacy in Public*, 69 U. MIA. L. REV. 141 (2014); Helen Nissenbaum, *Protecting Privacy in an Information Age: The Problem of Privacy in Public*, 17 LAW & PHILO. 559 (1998).

174 הראשונה הייתה סן פרנסיסקו. ראו Shannon Van Sant & Richard Gonzales, *San Francisco Approves Ban on Government's Use of Facial Recognition Technology*, NPR (May 14, 2019), https://n.pr/3wPwm5N.

בכפוף לחריגים מסוימים ומרובי תנאים בקשר לרשויות אכיפת חוק.[175] ומול אלה, בישראל, הממשלה דווקא מעוניינת לקדם שימוש במצלמות כאלה.[176]

במילים אחרות, השימוש המדינתי בבינה מלאכותית עלול לפגוע בפרטיות ובזכויות יסוד דמוקרטיות נוספות. הבעיה כאן אינה בטכנולוגיה עצמה, אלא בשימושים בה. אבל אפשרות המעקב טבועה בבינה המלאכותית (כלומר, זו אחת המזמינויות שלה). המדינה אינה אמורה לעקוב אחרי אזרחיה, אלא אם יש טעם חזק ומשכנע, שאותו יש לבחון במסגנת החוקתית. אעיר שגם כאשר המדינה רוכשת מערכות פרטיות-מסחריות של בינה מלאכותית אין הדבר פוטר אותה מחובותיה החוקתיות. היעדר השקיפות לגבי אופן פעולת הטכנולוגיה מגביר את הקושי. אבל אלה אינן בהכרח סוגיות חדשות, ויש בידינו סל כלים חוקתי לבחון את הדברים.

סיכום: אתגר הכוח

הצבעתי על האתגרים המרכזיים שהבינה המלאכותית מציבה לפרטיות בעת הזו. אמת המידה לזיהוי האתגרים הייתה בראש ובראשונה עיונית, לאור הבנת הפרטיות כשליטה, עמדה שמייצגת שורת הצדקות לפרטיות. אמת מידה שנייה שנשזרה בדיון היא הדין הפוזיטיבי. הגם שהדין מתיימר להיות ניטרלי לטכנולוגיה, הוא עוצב בסביבה שקדמה לבינה המלאכותית. הדיון העלה שורת אתגרים בשלב איסוף המידע והשימוש בו: אתגר הקופסה השחורה (ואתגרי ההודעה וההסכמה הנלווים), אתגרי צמידות המטרה ומזעור המידע, אתגר קציר המידע הפומבי, אתגר הזיהוי, אתגר הקבוצה, אתגר התחזית, אתגר העיבוד ואתגר המעקב. בסופו של דבר כולם מתכנסים לפערי הכוחות, וזהו אתגר הכוח.[177]

במישור הפרטיות המידעית ניצב גורם אחד (תאגיד המידע) מול אנשים רבים.[178] תאגיד המידע הוא בנק, חברת ביטוח, ספק תקשורת, מוסד רפואי, פלטפורמה דיגיטלית כמו רשת חברתית, ויישומונים אינסוף. במאגרי המידע התאגידיים יש מידע על משתמשים רבים – כשלושה מיליארד אצל פייסבוק (מֶטָה), והמידע מגוון, עשיר ואישי. התאגידים משתמשים במידע לצורכיהם. המודל העסקי שלהם פשוט ומוכר: הם נותנים שירות "חינמי", אבל אוספים מידע אישי, לפי "הסכמה" ל"מדיניות הפרטיות". המידע

175 ראו ס' 5(1)(h) לחוק הבינה המלאכותית, לעיל ה"ש 7. לפי החוק יותר שימוש במערכת ביומטרית בחקירות על חטיפת אדם, סחר בבני אדם, ניצול מיני של אדם ובחיפושים אחרי נעדרים, וכן כדי למנוע סכנה מהותית ומיידית לחיי אדם או רכוש, או חשש אמיתי לפיגוע טרור. ה־AIA מבהיר שהפעלת הסמכות אינה פוטרת את המדינה מציות ל־GDPR בקשר לשימוש במידע ביומטרי.

176 ראו תזכיר הצעת חוק לתיקון פקודת המשטרה (מס') (מערכת צילום ביומטרית), התשפ"ד–2023.

177 ראו גם אצל Calo, *Artificial Intelligence Policy*, לעיל ה"ש 84, בעמ' 424.

178 ראו Michael Birnhack, *S-M-L-XL Data: Big Data as a New Informational Privacy Paradigm*, *in* BIG DATA AND PRIVACY: MAKING ENDS MEET 7 (Future of Privacy Forum & Center for Internet & Society, Stanford Law School) (2013).

משמש להתאמה אישית של מוצרים, שירותים ופרסומות – שלהם ושל "שותפים".[179] חוקרת המידע האמריקנית שושנה זובוף (Zuboff) תיארה זאת כ"מעקב קפיטליסטי" (Surveillance Capitalism).[180] לעיתים המידע המסחרי הופך למידע פוליטי, למשל כפי שאירע בפרשת פייסבוק-קיימברידג' אנליטיקה.[181]

דיני הפרטיות מעניקים לנושא המידע (מעט) כוח מול תאגידי המידע. עקרונות המידע האישי מבקשים להבטיח את האפשרות לשלוט במידע. ועדיין, רב כוחם של תאגידי המידע. הם יוזמים את התהליך, הם שולטים בעיצוב המערכת,[182] הם מנסחים את ההודעות בשפה תאגידית ידידותית שמנסה להסוות מכשולים רבים.[183] ראינו שנתוני העתק הם חומר הגלם של הבינה המלאכותית, אבל הם יותר מכך. הבינה המלאכותית תובעת את ליטרת המידע שלה כל העת. ככל שהשימושים בבינה מלאכותית יתגברו, תיווצר תלות במידע הזה. המטפורה המתאימה היא אולי של התמכרות, לא רק של צריכת דלק.[184]

הבינה המלאכותית מעצימה פערי כוח. בחיינו הדיגיטליים אנחנו מייצרים שובל מידע, 24 שעות ביממה:[185] אינסוף השימושים ממכשיר המעקב האישי המכונה "טלפון חכם",[186] במחשבים, בתשלומים דיגיטליים, באביזרים רפואיים ואביזרי ספורט, בתקשורת בין־אישית, במיקום, בגוף שלנו. לתאגידים יש כעת יכולת להפיק תובנות חדשות שלא היו נגישות קודם. כשחיינו הופכים דיגיטליים, אין לנו ברירה אלא למסור להם מידע. הסך הכול הוא שאנחנו נתונים למעקב תאגידי רציף.

179 בשימוע בסנאט למנכ"ל פייסבוק מארק צוקרברג, בעקבות פרשת פייסבוק-קיימברידג' אנליטיקה, תהה הסנטור אורין האצ' מהו המודל העסקי של החברה:

"HATCH: Well, if so, how do you sustain a business model in which users don't pay for your service?

ZUCKERBERG: Senator, we run ads".

לתמליל, ראו WASH. POST (Apr. 10, 2018), https://www.nevo.co.il/shorturl/FDBAFC.

180 ZUBOFF, לעיל ה"ש 22.

181 ראו Federal Trade Commission (FTC), Opinion In the Matter of Cambridge Analytica, LL.C. (Nov. 25, 2019), https://www.nevo.co.il/shorturl/87442B.

182 WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES (2018).

183 לדיון ביקורתי ראו ARI EZRA WALDMAN, INDUSTRY UNBOUND: THE INSIDE STORY OF PRIVACY, DATA, AND CORPORATE POWER (2021).

184 מטפורת הדלק ראו כהנא ושוורץ אלטשולר, לעיל ה"ש 57. תשובה אפשרית לכך עשויה להימצא ב"דאטה סינטטי". לדיון, ראו Trivellore E. Raghunathan, *Synthetic Data*, 8 ANNU. REV. STAT. APPLICATION 129 (2021).

185 לשובל המידע ראו בירנהק **מרחב פרטי**, לעיל ה"ש 26, בעמ' 169; עת"ם (מנהלי ת"א) 24867-02-11 **איי.די.איי חברה לביטוח בע"מ נ' רשם מאגרי המידע, הרשות למשפט טכנולוגיה ומידע במשרד המשפטים**, פס' 5-6 (נבו 5.8.2012).

186 על היבטי הפרטיות בטלפון החכם ראו למשל, בהקשר הפלילי, את דעת המיעוט של השופט אלרון, דנ"פ 1062/21 **אוריך נ' מדינת ישראל**, פס' 4 (נבו 11.1.2022).

האם החשש אמיתי? האם תחושת המעקב שלנו מתעצמת כאשר מכונה עוקבת אחרינו, ולא אדם? התגובה של אנשים למעקבים מהסוגים השונים היא שאלה אמפירית. ניתן לשער שיש מי שחוששים יותר מהאדם העוקב מאשר מהמכונה. מי שחברים בקבוצה חברתית מוחלשת, למשל מהגרי עבודה ומבקשי מקלט, עשויים להעדיף אינטראקציה עם מכונה מאשר עם שוטר. אחרים יחששו יותר מהמכונה העוקבת: המכונה פועלת באופן עמום, מרוחק, ולעיתים סמוי, גם אם אנחנו יודעים שהיא שם. המכונה אינה מתעייפת, היא מדייקת, צוברת מידע, ונוכחת בשטח ובתודעה. המעקב, כפי שהסביר לנו הפילוסוף הצרפתי מישל פוקו לפני שנים רבות, ממשטר התנהגות.[187] הפעולה הסמויה מקשה כמובן גם על אכיפה.

המאמר ביקש למפות את אתגרי הפרטיות. כיצד צריך המשפט להגיב? קשה מאוד לזהות את נקודת הזמן המיטבית להסדרה. התערבות מוקדמת מדי תצנן פיתוחים רצויים ותגביל את מרחב הנשימה שדרוש לפיתוח טכנולוגיה חדשה; התערבות בשלב מתקדם עלולה להיות מאוחרת מדי, אחרי שכבר נקבעו עובדות בשטח, ואז היכולת לשנות מורכבת ויקרה.

בינה מלאכותית היא טכנולוגיה חדשה ומשבשת, אבל מרכיבים מידעיים רבים שלה מוכרים למעלה מעשור. שאלות רבות שמתעוררות כעת התעוררו כבר לגבי מידע ברשת בכלל, ולגבי נתוני עתק בפרט: הקשיים בנוגע להודעה והסכמה חופשית מדעת, היכולת לתאר מראש את השימושים במידע כדי למזער את איסוף המידע וכדי לציית לעקרון צמידות המטרה, היכולת לפצח את "הקופסה השחורה", ועוד. המשפט מנסה כל העת לאתר פתרונות, ואלה רלוונטיים גם לבינה מלאכותית. הניסיון בהתמודדות עם אתגרי הפרטיות בעשורים האחרונים מעלה שאין פתרון קסם אחד, והדרך הטובה ביותר היא שילוב כלים משפטיים, טכנולוגיים, ארגוניים וחברתיים.[188]

האפיק המשפטי, ובייחוד הדור המתקדם של עקרונות הגנת המידע האישי שמגולם ב-GDPR, מציע סל כלים שחלקו מתאים לבינה מלאכותית. את ההודעה לנושאי המידע יש לפשט ככל שניתן, מבחינת נגישות, עיצוב, שפה פשוטה, אורכה והמשלב הלשוני שלה. ראוי לציין במפורש האם מידע שנאסף ישמש ללימוד מכונה. ראוי לציין את מטרות איסוף המידע ועיבודו בבינה המלאכותית. לעיתים המטרה ידועה וברורה, למשל איתור חריגות בשימוש בכרטיס אשראי, לשם התראה על הונאה. כאשר המטרה כללית יותר ואינה ידועה בשלב האיסוף, יש לומר זאת במפורש. ניתן להשתמש בהסכמה מדורגת, בדומה למקובל בתחום הרפואי.[189] הסבריות חשובה גם היא.

אני סבור שאין לאסוף מידע בדרך של קציר מידע כללי ללא הסכמת נושאי המידע; אין מקום להיתלות בהתממת מידע בשלב האיסוף כנימוק לחמוק מתחולת דיני הגנת

187 מישל פוקו **לפקח ולהעניש: הולדת בית הסוהר** (דניאלה יואל מתרגמת 2015).

188 LESSIG, לעיל ה"ש 34, הצביע על המשפט, על נורמות חברתיות, על נורמות שוקיות ועל הטכנולוגיה כגורמים מאסדרים.

189 לדיון, ראו Brent Daniel Mittelstadt & Luciano Floridi, *The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*, 22 SCI. & ENG. ETHICS 303, 312 (2016).

הפרטיות בשלבי ההמשך של העיבוד והשימושים בתוצרי העיבוד. הדין הקיים אוסר איסוף מידע עודף, כלומר מידע שחורג מהדרוש למימוש המטרה המקורית. אם נאסף מידע כזה, אין להשתמש בו. כדי לצמצם את המעקב התאגידי והשלטוני ניתן לאסור שימושים מסוימים מראש, למשל איסור לעניין מצלמות ביומטריות במרחב הציבורי. את אלה יש להשלים באמצעים משפטיים כדי למתֵן את פערי הכוח בין הצדדים, כמו סמכויות אכיפה משמעותיות של הרשויות המתאימות, אפיקים של תביעה ייצוגית וכדומה.

בצד זה, יש כלים משפטיים-ארגוניים-טכנולוגיים שה־GDPR מצטיין לגביהם: חובה לערוך תסקיר מקדים בנוגע להיבטי הפרטיות לפני הפיתוח וההטמעה של מערכת חדשה, חובה למנות ממונה הגנת פרטיות שתהיה מעורבת בכל התהליך, מראשיתו, וחובה לבצע תהליך של הנדסת פרטיות. למשל, שילוב פתרון טכנולוגי בניתוח המידע – הוא אפיק שיש להוסיף ולפתח.[190]

הפרטיות היא ערך דמוקרטי יסודי. אין לוותר עליה רק משום ההתפעלות מהמכונה.

190 ראו למשל שיטות לעיבוד מידע רפואי בלי זיהוי המטופלים ובלי איבוד מידע: Ravit Geva et al., *Collaborative Privacy-Preserving Analysis of Oncological Data Using Multiparty Homomorphic Encryption*, 120(33) PROC. NAT'L ACAD. SCIS. 1 (2023).