

חקיקה אלגוריתמית ולקחי המודל הלוגריתמי: על מס דירה שלישית וחקיקה מבוססת נתונים

בשולי העניין האקדמי בפרשת קוונטינסקי, הדיון בתוכן החקיקה של מס ריבוי דירות נשכח. הנוסחה שלפיה חושב המס התבססה על שני פרמטרים והצריכה חישוב הכולל פעולות מתמטיות לא טריוויאליות, כגון שימוש בלוגריתם, ומשמעותה לא הייתה מובנת למשפטנים מן השורה. תהליך החקיקה של נוסחת מס ריבוי דירות הוא דוגמה לחקיקה מבוססת נתונים (data based law) המתבססת על מידול סטטיסטי "דיני", שממנה ניתן להפיק לקחים לקראת הדור הבא – חקיקה מונחית נתונים (data driven law), המתאימה באמצעות למידת מכונה ומודלי בינה מלאכותית את הוראות החוק באופן אישי למושאיו. המאמר מציג עקרונות מנחים לחקיקה מבוססת נתונים, המביאים בחשבון את הצורך בדילברציה אנושית במהלך תהליך החקיקה; את הצורך במנגנוני עדכון לחקיקה על מנת לתקף את המודל; בממשל נתונים ראוי; במנגנוני השגה על החלטות אוטומטיות; וכן במיפוי והתייחסות למקרי קצה.

מבוא

בעניין קוונטינסקי¹ פסל בית המשפט העליון את הפרק בחוק להתייעלות כלכלית (תיקוני חקיקה להשגת יעדי התקציב לשנות התקציב 2017 ו-2018)² שמכוחו הוטל מס ריבוי דירות, המוכר בשיח הציבורי כ"מס דירה שלישית". עניין קוונטינסקי, שבו נפסל לראשונה דבר חקיקה על בסיס פגמים היורדים לשורש הליך החקיקה שלו, הוא נדבך חשוב במשפט החוקתי הישראלי.³ ואולם, נוכח העילה הפרודצורלית שבגינה פסל בית

* עמיר כהנא הוא תלמיד מחקר לתואר שלישי בפקולטה למשפטים, האוניברסיטה העברית בירושלים, עמית מחקר במרכז פדרמן לחקר הסייבר, התוכנית לסייבר ומשפט, ועמית בקבוצת המחקר Three Generations of Digital Human Rights. תודתי לחברי המערכת, ולקוראים האנונימיים על הערותיהם המועילות.

1 בג"ץ 10042/16 קוונטינסקי נ' כנסת ישראל (נבו 6.8.2017).
2 פרק י"ב לחוק ההתייעלות הכלכלית (תיקוני חקיקה להשגת יעדי התקציב לשנות התקציב 2017 ו-2018), התשע"ז-2016 (להלן: חוק מס ריבוי נכסים).
3 איתי בר-סימן-טוב "החלטה מתבקשת או 'סטייה רבתי' בלתי מוצדקת מההלכה הפסוקה?" **מחקרי משפט** לב 889 (2020); סוזי נבות "הדיאלוג החוקתי: דו-שיח בכלים מוסדיים" **משפטים על אתר** יב 99 (2019); גילה שטופלר "בית המשפט העליון כבלם מפני השתלטות עוינת של הממשלה על סמכויות הכנסת" **ICON-S-IL Blog** (25.3.2018).

המשפט את חוק מס ריבוי הדירות, תשומת לב מועטה הוקדשה לתוכן החוק עצמו,⁴ והעיתירות בעניין תוכנו נדחו על הסף.⁵

מאמר זה אינו עוסק במשפט חוקתי ובעילות פרוצדורליות לפסילת חוקים – אלא בתהליך החקיקה של מס ריבוי דירות, ובתוכו בגלגוליה של הנוסחה לחישוב סכום המס שהושת לפיו בגין בעלות על דירה שלישית ומעלה. כפי שיתואר בהמשך, נוסחה זו החלה את דרכה כמס אחיד, עד שנתגלגלה לכדי ביטוי מתמטי סבוך שלתיאורו המילולי בגוף החקיקה נדרש כעמוד שלם.⁶

ביטויים מתמטיים מורכבים בחקיקה מעלים שאלות משל עצמם הנוגעות לכוח האקספרסיבי של החוק: האם בשעה שהחוק אינו מנוסח באופן שנהיר לחברי הקהילה יש ביכולתו להמשיך לשמש תזכורת לקהילה בנוגע לערכיה?⁷ תיתכנה גם שאלות הנוגעות לכוח ההסברי של החוק, ליכולתו המוגבלת להכווין התנהגות כאשר הוא משתמש במושגים משדה סמנטי אחר, ולתוקף הרטורי של מראית העין המדעית של הנוסחאות. בתוך כך אפשר גם לתהות בנוגע לשימוש בתיאור מילולי של נוסחה, המתרגם ביטוי מתמטי תמציתי ללשון המשפטית הסבוכה, ובתוך כך מגדיל את הסיכוי לטעות אגב המרתו חזרה לנוסחה המתמטית (בדרך כלל באמצעות משפטים שמוגבלים בהכשרתם המתמטית).

אולם מאמר זה מבקש להתמקד בהיבט אחר של הנוסחה לחישוב מס ריבוי דירות, והוא היותה דוגמה של חקיקה מבוססת נתונים (data-based law), ועל הקשר בין תהליכי חקיקה מבוססת נתונים לבין הסתמכות על בינה מלאכותית לצורכי חקיקה. מאמר זה מבקש לעשות זאת באמצעות הדרך שעשה מס ריבוי דירות מדיונים במשרד האוצר שאפשר שמצאו את דרכם לעיתונות כ"בלון ניסוי" ועד לעדכונו הראשון בצו, וממנו ללמוד על הבעייתיות בשימוש במודלים סטטיסטיים בחקיקה, להקיש מכך על הבעייתיות שבהסתמכות על בינה מלאכותית בחקיקה, ולהציע כמה דרכים למתן את הסיכונים הכרוכים בחקיקה מבוססת נתונים.

⁴ <https://www.nevo.co.il/shorturl/25A19CEF-8DC>; עידו פורת "המינהליזציה של המשפט החוקתי ואקטיביזם פרו-פרלמנטרי" **ICON-S-IL Blog** (1.4.2018) <https://www.nevo.co.il/shorturl/A59D7E20-A21>; מנגד, ראו גונן אילן "שלוש שנים לבג"ץ קוונטינסקי: רוב מהומה על לא מאומה" **ICON-S-IL Blog** (30.4.2020) <https://www.nevo.co.il/shorturl/2D49711F-18B>.

⁵ למעט שתי הפסקאות הראשונות לפסק הדין של השופט סולברג בעניין **קוונטינסקי**, לעיל ה"ש 1, המשמשות כמבוא לפני פרק התשתית העובדתית, העוסק בהשתלשלות הפרוצדורלית של חקיקת חוק מס ריבוי נכסים, ופס' 94 לפסק דינו, המתעכבת על מורכבות ההסדר המוצע בחוק על מנת להסביר מדוע אין להידרש אליו כלאחר יד, אין בעניין **קוונטינסקי** התייחסות של ממש לתוכן החוק.

⁶ עניין **קוונטינסקי**, שם, פס' 55 לפסק הדין של השופט מזוז.

⁷ תוספת לפרק י"ב לחוק מס ריבוי נכסים.

⁷ ראו למשל Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1391–92 (1971).

אין הגדרה מוסכמת, ולו משפטית, למערכות בינה מלאכותית,⁸ אולם נראה שהארגון לשיתוף פעולה ולפיתוח כלכלי (OECD) והאיחוד האירופי מתכנסים לנוסח מוסכם, שעשוי להיות הסטנדרט הבינלאומי המקובל. ה־OECD מגדיר בינה מלאכותית כ"מערכת מבוססת מכונה שלמטרות מפורשות או משתמעות מסיקה מהקלט שהיא מקבלת כיצד לחולל תכנים כגון תחזיות, תוכן, המלצות או החלטות שיש ביכולתן להשפיע על סביבות פיזיות או וירטואליות. למערכות בינה מלאכותית שונות ישנן רמות שונות של אוטונומיה ויכולת הסתגלות לסביבתן לאחר פרישתן".⁹

מאמר זה יוצא מנקודת הנחה שמערכות בינה מלאכותית הן בסופו של דבר אוסף של טכניקות אלגוריתמיות סטטיסטיות ולמידת מכונה.¹⁰ משכך, באמצעות בחינה של חקיקה מבוססת נתונים המסתייעת במודלים כמותיים "פשוטים", ניתן להפיק לקחים בנוגע לחקיקה המסתייעת בכלי בינה מלאכותית. ההיקף העצום של הנתונים שמודלי למידת מכונה ניזונים מהם, כמו גם היקף הפרמטרים שבמודלים אלו, מעלים את רמת הקושי בהתמודדות עם אתגרי החקיקה מבוססת הנתונים, ולעיתים מייצרים אתגרים יחודיים.

חלקו הראשון של המאמר מניח בפני הקוראים את היסודות המושגיים שבהם ישתמש המאמר; חלקו השני עוסק בחקיקה מבוססת נתונים, על יתרונותיה וחסרונותיה. חלקו השלישי מתאר את תהליך התגבשותה של הנוסחה להערכת שווי דירה שבחוק מס ריבוי נכסים. חלקו הרביעי בוחן את חוק מס ריבוי נכסים כדבר חקיקה מבוססת נתונים, וחלקו החמישי דן בחקיקה מונחית נתונים. לקראת סיום, חלקו השישי מציע עקרונות מנחים לחקיקה מבוססת נתונים.

א. חקיקה מבוססת ראיות, חקיקה מבוססת נתונים וחקיקה מונחית נתונים

חקיקה מבוססת נתונים (data-based law) היא חקיקה המסתמכת על מודלים כמותיים-סטטיסטיים בתהליך עיצובה, עד כדי הטמעת מודלים אלו בלשון החוק עצמו. היא מתקיימת בתוך ספקטרום אשר לו שני קצוות: חקיקה מבוססת ראיות (evidence based law), שעושה שימוש מסוים בראיות אמפיריות לקביעת הכלל המשפטי (למשל, מתן תימוכין לענישה מחמירה בהסתמך על מחקר אקדמי שמראה שיש בה כדי להרתיע), נמצאת בקצהו האחד. בקצהו השני של הספקטרום נמצאת חקיקה מונחית נתונים (data

8 עמיר כהנא ותהילה שוורץ אלטשולר אדם, מכונה, מדינה: לקראת אסדרה של בינה מלאכותית 28–31 (2023).

9 Luca Bertuzzi, *OECD Updates Definition of Artificial Intelligence 'to Inform EU's AI Act'*, EURACTIV (Nov. 9, 2023) <https://bit.ly/4deQIZQ>.

10 LUC JULIA, *THERE IS NO SUCH THING AS ARTIFICIAL INTELLIGENCE* (2020); Stefano Quintarelli, *Let's Forget the Term AI. Let's Call Them Systematic Approaches to Learning Algorithms and Machine Inferences (SALAMI)*, QUINTA'S WEBLOG (Nov. 24, 2019) <https://bit.ly/4cT8Fxp>.

driven law), ששואפת להסתייע בטכניקות של למידת מכונה כדי לייצר כללים משפטיים מותאמים לנסיבות בהסתמך על נתוני עתק (big data), ולעיתים – מותאמים אישית למושא החוק (למשל, חקיקה העושה שימוש במודל לחיזוי רצידיביזם על מנת לקבוע את מרחב הענישה הקונקרטי של כל נאשם ונאשמת). בין שני קצוות אלו של הספקטרום מצויה חקיקה מבוססת נתונים, השואפת לייצר כללים משפטיים מדויקים יותר על בסיס נתונים רחב (למשל, קביעת מדרגות ענישה בחתכים של סוגי עבירה, על בסיס מדידה מתמשכת של הרתעה).

בשלהי המאה שעברה קראה ממשלת בריטניה בראשותו של טוני בלייר לעיצוב מדיניות "צופה פני עתיד, המבוססת על ראיות אמיתיות ושאינה תגובה ללחצים קצרי טווח, המתמודדת עם סיבות ולא עם סימפטומים".¹¹ ואכן ניתן לראות את התחזקותה של גישת המדיניות מבוססת הראיות בבריטניה ובעולם.¹² גישה זו, על פניה, מבקשת לעצב מדיניות באופן טכני-רציונלי, באופן המבוסס על ראיות אמפיריות.¹³ מדיניות מבוססת נתונים מבקשת לבחור במדיניות ש"עובדת",¹⁴ שניתן להצביע על קשר סיבתי חזק בינה לבין התוצאות המבוקשות על ידי מקבלי ההחלטות.¹⁵ לפי גישה זו, ה"ראיות" הן עניין מדעי נתון, והמדיניות שהיו מקבלים הגורמים הפוליטיים אילו רק היו מודעים להן הייתה משופרת ועדיפה.¹⁶ הספרות נוטה שלא להבחין בין המונח "מדיניות מבוססת ראיות" (evidence-based policy) לבין המונח "מדיניות מבוססת נתונים" (data-based policy), זאת משום שמדיניות מבוססת נתונים נתפסת כסוג של מדיניות מבוססת ראיות.¹⁷

ביסוסו של דבר חקיקה על ראיות אמפיריות מאפשר לחוק לעטות חזות רציונלית ומדעית, ולייצר מראית עין אובייקטיבית לטעמים שמאחוריו. מנגד, הוא מאפשר גם בחינה ניסויית של החוק על בסיס פרמטרים מדידים,¹⁸ כדי לבחון אם מטרות החקיקה

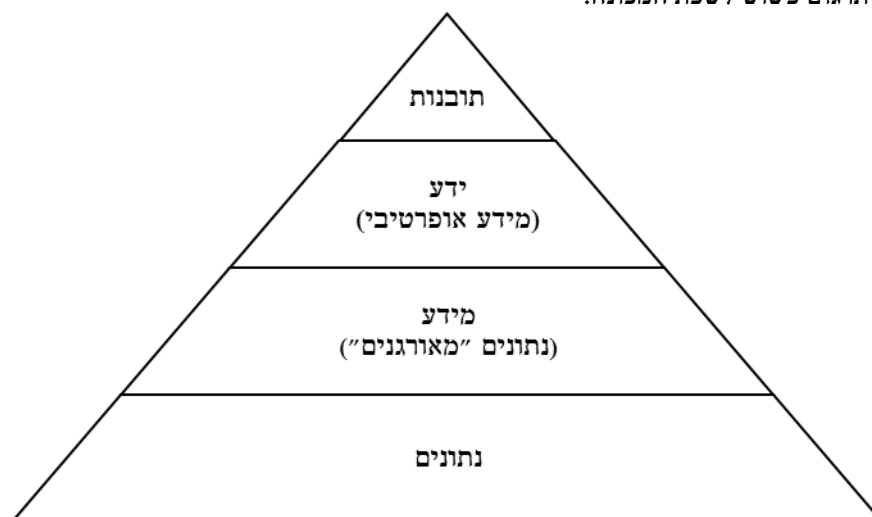
- 11 Cabinet Office, *Modernising Government*, 1998-9, HC 4310, at 14-22 (1999).
- 12 ראו למשל Adrian Pabst, *Rethinking Evidence-Based Policy*, 255 NAT'L INST. ECON. REV. 85 (2021).
- 13 Warren Pearce et al., *Evidence and Meaning in Policy Making*, 10 EVIDENCE & POL'Y. 161 (2014); Joshua Newman et al., *Policy Capacity and Evidence-Based Policy in the Public Service*, 19 PUB. MGMT. REV. 157, 158 (2017).
- 14 ואכן, בעקבות ההצהרה של ממשלת בלייר ב-1999 הוקמו בבריטניה עשרה מרכזי הערכת מדיניות על בסיס ראיות אמפיריות שנקראו What works centers ("מרכזי מה-עובד"). ראו Pabst, לעיל ה"ש 12.
- 15 Eileen Munro, *Evidence-Based Policy in PHILOSOPHY OF SOCIAL SCIENCE: A NEW INTRODUCTION* 48 (Nancy Cartwright & Eleonora Montuschi, eds., 2014).
- 16 לביקורת על גישת המדיניות מבוססת הראיות ראו PAUL CAIRNEY, *THE POLITICS OF EVIDENCE-BASED POLICY MAKING* (2016).
- 17 ראו למשל Taewoo Nam, *Do the Right Thing Right! Understanding the Hopes and Hypes of Data-Based Policy*, 37 GOV'T. INFO. Q. 1, 2-4 (2020).
- 18 Sofia Ranchordas, *The Whys and Woes of Experimental Legislation*, 1 THEORY & PRACT. LEGIS. 415 (2013).

הושגו.¹⁹ ראיות אמפיריות יכולות לספק – ולו בדיעבד ולצרכים רטוריים – הנמקה רציונלית למדיניות גם אם הצדקתה המקורית מבוססת על בחירה ערכית.²⁰ פעמים רבות ביסוסה של חקיקה על ראיות אמפיריות יכולה להקל על בחינת המידתיות החוקתית שלה, במסגרת מבחן הקשר הרציונלי שבין המטרה לבין האמצעי, שאותו ניתן לאשש על בסיס תשתית עובדתית.²¹ עם זאת, פעמים רבות מדיניות (או חקיקה) מבוססת ראיות אינה אלא ראיות מבוססות מדיניות (או חקיקה). היינו, החלטות פוליטיות שמוצדקות בדיעבד באמצעות עובדות, או פרשנות מחודשת של ראיות מדעיות באופן שמאשש את העמדה הפוליטית הראשונית של מקבליהן,²² ומצדיק את המדיניות.²³

תחת ההיוריסטיקה הנפוצה של נתונים-מידע-ידע-תובנות (DIKW, Data-Information-Knowledge-Wisdom),²⁴ נתונים הם הבסיס שממנו מופק (ובלעדיו אין) מידע. המידע בתורו מעובד לידע והופך לתובנות אופרטיביות. המעבר בין שלבי היררכיית DIKW כרוכים בעלייה בערך ובמשמעות של הנתונים (שעובדו בהדרגה לכדי תובנות),²⁵ ובה בעת בירידה ביכולתם של הנתונים המעובדים להזין אלגוריתמים.²⁶ נתונים גולמיים, גם כאשר הם מעובדים לכדי מידע מאורגן ומתויג, כשלעצמם הם בעלי משמעות פעוטה ונדרש עיבוד נוסף כדי לחלץ מהם ידע ותובנות. כך למשל רשומה בודדת של ערך נכס במקום גאוגרפי מסוים כשלעצמו עשוי להיות חסר משמעות, אך קל למערכות אוטומטיות לעבד את הנתון הזה. לשייך לו, למשל, את המודר

- 19 ראו למשל Anne Seidman & Robert B. Seidman, *ILTAM: Drafting Evidence-Based Legislation for Democratic Social Change*, 89 B.U. L. REV. 435 (2009).
- 20 Ken Young et al., *Social Science and the Evidence-Based Policy Movement*, 1 SOC. POL'Y & SOC'Y 215, 216–19 (2002).
- 21 אהרן ברק *מידתיות במשפט* 382–384 (2010); Rob van Gestel & Jurgen de Poorter, *Putting Evidence-Based Lawmaking to the Test: Judicial Review of Legislative Rationality*, 4 THEORY & PRACT. LEGIS. 155 (2016).
- 22 Sidney A. Shapiro, *Political Science: Regulatory Science After the Bush Administration*, 4 DUKE J. CONST. L. & PUB. POL'Y 31 (2009).
- 23 ראו למשל Michael G. Marmot, *Evidence Based Policy or Policy Based Evidence?*, 328 BRIT. MED. J. 906 (2004); Caitlin E. Hughes, *Evidence-Based Policy or Policy-Based Evidence? The Role of Evidence in the Development and Implementation of the Illicit Drug Diversion Initiative*, 26 DRUG & ALCOHOL REV. 363 (2007).
- 24 ראו למשל R.L. Ackoff, *From Data to Wisdom*, 16 J. APPLIED SYS. ANALYSIS 3 (1989); Jennifer Rowley, *The Wisdom Hierarchy: Representations of the DIKW Hierarchy*, 33 J. INFO. SCI. 163 (2007); Martin Frické, *The Knowledge Pyramid: the DIKW Hierarchy*, 46 KNOWLEDGE ORG. 33 (2019).
- 25 DAVE CHAFFEY & STEVE WOOD, *BUSINESS INFORMATION MANAGEMENT: IMPROVING PERFORMANCE USING INFORMATION SYSTEMS* 224 (2005).
- 26 ELIAS M. AWAD & HASSAN M. GHAZIRI, *KNOWLEDGE MANAGEMENT* 41 (2004).

הסוציאקונומי המתאים לו. עם צבר נתונים כזה, אפשר לנסות ולמצוא מודל סטטיסטי לפי טיב הקשר בין שני הפרמטרים המנבא את ערך הנכס בהסתמך על המדד הסוציאקונומי שלו. זהו ידע. היכולת האנושית לעשות שימוש בידע הזה – למשל, לגזור ממנו תובנות לגבי מדיניות – היא תולדה של המשמעות שלו, שאינה ניתנת לתרגום פשוט לשפת המכונה.



הבחירה בחקיקה מבוססת נתונים תחת חקיקה מבוססת ראיות היא מעבר לשלב נמוך יותר בהיררכיית DIKW, שבו המידע הראייתי מאבד מהמשמעויות שנטענו בו וחוזר להיות נתונים גולמיים. חקיקה מבוססת נתונים יכולה להיתפס כלגיטימית יותר מחקיקה מבוססת ראיות. הגם שלחקיקה מבוססת ראיות יש בסיס אמפירי מסוים, עדיין נותר החשש שראיות אלו עשויות להיות חלקיות או אנקדוטליות. שכן, המימרה המוכרת שלפיה צורת הרבים של "אנקדוטה" איננה "נתונים"²⁷ משקפת חשש מהיסק אינדוקטיבי מוטעה על בסיס תצפיות בודדות,²⁸ המגולמות בראיות, והשלכתו על הכלל. חשש אחר יכול להיות מביסוס מדיניות כללית על בסיס ראיות חלקיות באופן שמאשש (במודע או שלא במודע) הנחות קודמות. הגולמיות של נתונים בלתי מעובדים מאפשרת להם לכל הפחות להיחזות כאובייקטיביים, בכפוף לכך שמתודולוגיית האיסוף שלהם נכונה,²⁹ ובהיותם נוחים לעיבוד שוויוני ואוטומטי. העדפה של "נתונים"

27 הנוסח המקורי של הביטוי היה הפוך לחלוטין: "צורת הרבים של 'אנקדוטה' היא 'נתונים'", אך הנוסח ההפוך הוא השכיח יותר. לדיון קצר במקורות הביטוי ראו Dan Nguyen, *Don't forget: The plural of anecdote is data*, DAN NGUYEN'S BLOG (Sep. 10, 2015), <http://blog.danwin.com/don-t-forget-the-plural-of-anecdote-is-data>.
להשוואה ב-Google Ngram של שכיחות שני הביטויים ראו <https://did.li/2VSCN>.

28 ראו למשל Jeffrey J. Rachlinski, *Evidence-Based Law*, 96 CORNELL L. REV. 901, 919 (2011).

29 כאשר מתודולוגיית האיסוף שגויה הנתונים עשויים להיות "צורת הרבים של אנקדוטה", ולייצג באופן חסר את האוכלוסייה. איסוף שגוי מתודולוגית של נתוני אימון למודל בינה

על "ראיות" היא חזרה לאבן יסוד אפיסטמולוגית שבאמצעותה מבנים את הידע, ואפשר שהיא משקפת הנחה סמויה שלפיה הבניה זו נעשית באופן שיטתי ואובייקטיבי, אך בה בעת היא מתעלמת מהצורך להגדיר את המתודולוגיות שבאמצעותן ממירים נתונים גולמיים לידע אופרטיבי.

ניתן לראות מדיניות או חקיקה מבוססת נתונים ככאלו המגלמות מודל שנוצר באמצעות עיבוד כמותי של בסיס נתונים מקיף וסדור ככל הניתן.³⁰ מדיניות או חקיקה כזו יכולות שתהיינה "מידועות נתונים" (data informed), היינו הגם שהן מבוססות על נתונים ניתן בתהליך העיצוב של המדיניות או החקיקה מקום גם לשיקולים נורמטיביים ולהיבטים תלויי-הקשר.³¹

בהקשר זה, נחזור לספקטרום שבקצהו אחד נמצאת חקיקה מבוססת ראיות. ראיות אלו יכולות להיות על בסיס של הערכה מידועת של השפעת החקיקה (במסגרת הערכת השפעת אסדרה),³² או כאלו שנתקבלו בעקבות ניסויים רגולטוריים,³³ אך לעיתים אלו עשויות להיות ראיות אנקדוטליות שאינן מתארות את המצב העובדתי שהמחוקק מבקש לתקן במלואו.³⁴

בקצהו השני של הספקטרום מצויה חקיקה מונחית נתונים (data driven). חקיקה כזו נותנת לנתונים את הבכורה ואף את הבלעדיות בתהליך עיצובה, ולעיתים גם בתהליך אכיפתה. אנו יכולים לתאר לעצמנו דינים המעוצבים בהסתמך על אלגוריתמים של למידת מכונה, המייצרים מודלים מורכבים הניזונים מבסיסי נתוני עתק, המכילים רבבות פרמטרים. באופן זה ניתן לעצב את החוק באופן אישי לכל אזרח ובכפוף להיקף המערך הסנסורי של המדינה (היינו, האמצעים הטכנולוגיים שבאמצעותם המדינה אוספת

מלאכותית עלול להביא להטיות אלגוריתמיות. ראו למשל כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 186–198.

30 לדוגמה להכנת התשתית למדיניות מבוססת נתונים, ראו החוק האמריקאי ליסודות לקביעת מדיניות מבוססת נתונים, המורה לראשי סוכנויות פדרליות (המנויות ב- 31 U.S.C. § 901(b) לכלול באסטרטגיה השנתית שלהם סוגיות מדיניות שתיבחנה על בסיס ראיות, לפרט את הנתונים שיידרשו לשם כך, ואת הדרכים לאיסופם, כמו גם את המתודולוגיות שבאמצעותן תיבחנה הראיות. ראו An Act to Amend titles 5 and 44, United States Code, to require Federal evaluation activities, improve Federal data management, and for other purposes [Foundations for Evidence-Based Policymaking Act], 132 Stat. 5529 (2019); גיא מור "מדיניות מבוססת נתונים" הרגולטור (16.4.2019) <https://regulator.online/2019/04/16/evidencebasedpolicy>.

31 שילדקמפ ואח' טוענות שנתונים אינם יכולים לבדם להנחות החלטות, וכי תמיד יש בהם מרכיב הנוגע לשיקולים מקצועיים של מקבל ההחלטה: Kim Schildkamp et al., *How School Leaders Can Build Effective Data Teams: Five Building Blocks for a New Wave of Data-Informed Decision Making*, 20 J. EDUC. CHANGE 283, 284 (2019).

32 ראו למשל ס' 3 לחוק עקרונות האסדרה, התשפ"ב–2021.

33 לנסיינות רגולטורית באמצעות ארגוי חול ראו למשל כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 296–302.

34 ראו Rachlinski, לעיל ה"ש 28.

נתונים על המתרחש בה), לספק לכל אזרח "מיקרו-דירקטיבות" בתפירה אישית, המורות לו כיצד עליו להתנהג בכל עת בהתאם לפרופיל הייחודי שלו.³⁵ זהו החזון שמציעים בן-שחר ופורת של חוק בהתאמה אישית (personalized law), המתאים את המהירות המרבית בכביש לגילו של הנהג, ערנותו, השעה ביום, ולעשרות פרמטרים אחרים שעשויים להשפיע על ביצועיו בכביש.³⁶

כהערת הבהרה, יש לציין בשלב זה שמודלי שפה גדולים (LLM),³⁷ או בינה מלאכותית יוצרת (גנרטיבית),³⁸ אינם מצויים כלל במרחב שתואר לעיל הנמצא בין חקיקה מבוססת ראיות לבין חקיקה מונחית נתונים. פונקציית המטרה של מודלים אלה אינה מודל כמותי שמייצג מערכת נסיבתית תומכת החלטות מדיניות או חקיקה, אלא לחזות מה המענה הנכון ביותר לשאלתא (prompt) המוצגת להם. אם נתאר לעצמנו חקיקה באמצעות מודלי שפה גדולים, כלומר מצב שבו מנסח החוק פונה אל ChatGPT או דומיו על מנת שאלו יחוללו לו תזכיר חוק, ניתן לשער שהמודלים הזמינים בשוק בעת כתיבת שורות אלו אכן ידעו להחזיר פלט שיוצר מצג של חקיקה מנוסחת היטב, אך ללא תימוכין אמפיריים לרציונל המהותי שמאחוריה.

כדי לעמוד על היחס בין חקיקה מבוססת נתונים לחקיקה מונחית נתונים, יש לבחון את היחס בין סטטיסטיקה לאלגוריתמים. האב הגנאלוגי של הסובייקט האלגוריתמי הוא הסובייקט האקטוארי/סטטיסטי.³⁹ ההיגיון של הסובייקט האלגוריתמי הוא היגיון של התאמה, שמחפש בנתונים את "כפילי המידע" של הסובייקט (data או data double) ⁴⁰ (doppelgänger). זאת, להבדיל מרציונל אקטוארי של שיוך לקבוצה, או מהרציונל הסטטיסטי של מודל שמביא להתאמה חלקית.⁴¹ הגם שהסובייקט האלגוריתמי קורם

35 על כך ראו כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 81–83. בנוסף ראו OMRI BEN-SHAHAR & ARIEL PORAT, PERSONALIZED LAW: DIFFERENT RULES FOR DIFFERENT PEOPLE (2021); Anthony J. Casey & Anthony Niblett, *Self-Driving Laws*, 66 U. TORONTO L.J. 429 (2016).

36 ראו למשל BEN-SHAHAR & PORAT, שם, בעמ' 5–7.

37 מודלי שפה גדולים (Large Language Models) הם מודלי למידה עמוקה שאומנו על טקסט בכמות עצומה, ויש ביכולתם לבצע מטלות רבות בתחום עיבוד השפה הטבעית (NLP). בינה מלאכותית יוצרת מבוססת טקסט, כגון ChatGPT, Claude או Gemini מבוססת על מודלי שפה גדולים.

38 כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 41–45.

39 BERNARD E. HARCOURT, EXPOSED 156 (2015).

40 HARCOURT, שם, בעמ' 141–165. על "כפילי מידע" ראו גם Kevin D. Haggerty & Richard V. Ericson, *The Surveillant Assemblage*, 51 BRIT. J. SOCIO. 605 (2000); מיקי זר "אנשים שקראו מאמר זה התעניינו גם ב...": על הקשר בין פרטיות לפרופילינג" **משפט, חברה ותרבות** ב 69, 77–73 (2019).

41 HARCOURT, לעיל ה"ש 39, בעמ' 162.

עור וגידים באמצעות כלים סטטיסטיים, הוא מורכב מאסופה של מודלים שמכוננים ומתווכים באמצעות קופסה שחורה.⁴²

בהתאמה, חקיקה מבוססת נתונים היא הורתה של החקיקה מונחית הנתונים, והיא נמצאת בתוך שבציר שבין האחרונה לחקיקה מבוססת ראיות. חקיקה מבוססת נתונים נשענת על רציונל סטטיסטי או אקטוארי. תהליכי הייצור שלה הם ידניים: באמצעות אנליסט אנושי שמעצב את המודל ומכוון אותו תוך שימוש בשיקול דעת ובניסיון מקצועי, ובאמצעות המחוקק עצמו, שנדרש להבין את המודל כדי להחליט אם יש להטמיעו בחקיקה.

כפי שמדיניות מבוססת נתונים היא סוג של מדיניות מבוססת ראיות, כך חקיקה מבוססת נתונים היא סוג של חקיקה מבוססת ראיות. היא נשענת על מתודולוגיות כמותיות, לרוב סטטיסטיות, התומכות בעיצוב כללים נורמטיביים. המודל הסטטיסטי שמופק בתהליך יכול לשמש הצדקה לכלל הנורמטיבי, לסייע בבחינת האפקטיביות שלו לאחר מעשה, או להיות מוטמע בפועל בחקיקה עצמה.

נחשוב על מחוקק המבקש לנסח חזקה כלשהי שבעובדה⁴³ – בהתקיים אירוע X (בקבוק משקה אלכוהולי פתוח ברכב), ייראו הנסיכות כאילו אירע אירוע Y (הנהג צרך את המשקה מעבר לכמות המותרת), ויחול בהתאמה הכלל המשפטי הרלוונטי לאירוע Y . אם החזקה נתמכת במודל סטטיסטי המעיד על קיומו של קשר אמפירי בין X ו- Y (דגימות מצטברות של רמת האלכוהול בדמם של נהגים ושל הימצאות הבקבוקים הפתוחים ברכבם),⁴⁴ הרי שהחזקה היא הטמעה בפועל של המודל בחקיקה (איסור על החזקת בקבוקי אלכוהול פתוחים ברכב), וניתן גם לאחר מעשה לבחון את תקפות המודל באמצעות תצפיות מצטברות (המשך המדידה). ואולם, מחוקק הסבור כי מלכתחילה אין זה ראוי לצרוך אלכוהול בנהיגה, ואף לא בכמות המותרת (אם מטעמים ערכיים-פוריטניים, ואם מתוך שנאת סיכון), עשוי להשתמש בקשר הסטטיסטי שבין הימצאות בקבוק פתוח ברכב לרמת אלכוהול אסורה בדמם של הנוהגים בו כדי להצדיק את עמדתו הערכית.

42 FRANK PASQUALE, THE BLACK BOX SOCIETY (2015). עוד לפני שהשימוש במודלי למידה עמוקה שיש להם "עכירות אלגוריתמית" ניכרת נעשה שכית, התריעו מומחים מהתגבשותה של "חברת הקופסה השחורה". המורכבות של הטכנולוגיה יוצרת מעין קופסה שחורה שלא ניתן לפרקה ולהבין את הנעשה בה עד תום. אם מערכות מומחה ישנות מבוססות תכנות קלאסי ניתן היה במאמץ ניכר "להנדס לאחור" ולהבין כיצד הן פועלות וכיצד התקבלו בהן החלטות – במודלים של למידה עמוקה קשה עד בלתי אפשרי להסביר כיצד נתקבל פלט זה או אחר. על בעיית העכירות ראו גדי פרל ותהילה שוורץ אלטשולר מודל ליצירת שקיפות אלגוריתמית 14–15 (2022).

43 ראו דנ"א 1516/95 מרום נ' היועץ המשפטי לממשלה פ"ד נב(2) 813, 828 (1998), שם נאמר כי "הרציונל העיקרי לקביעת חזקה שבעובדה הוא הסתברות קיומה של העובדה המוחזקת".

44 בשלב זה נניח שמתודולוגיית הבדיקה האמפירית חפה מבעיות.

ב. חקיקה מבוססת נתונים: יתרונות ואתגרים

להסתייעות בחקיקה מבוססת נתונים ישנם יתרונות הקשורים לעצם ההסתמכות על נתונים, וכאלו הקשורים להיבטים הנוגעים לתוצאה של תהליך החקיקה, אם זו אכן באה לידי ביטוי במודל כמותי או בנוסחה מתמטית (להבדיל, למשל, ממודל שהוטמע בחקיקה בדרכים אחרות, דוגמת כלל הכרעה בינארי או חזקה שבעובדה).

היתרונות של ההסתמכות על נתונים כמותיים נוגעים להיותם אובייקטיביים ומדויקים למראה (אם בסיס הנתונים הרלוונטי הוא "נקי" (מדויק כלל הניתן, ללא נתונים שגויים שנתווספו לו אגב תהליך יצירתו), מקיף ומהווה ייצוג מלא או נאמן של המציאות).⁴⁵ בכפוף לכך שאיסוף הנתונים נעשה באופן מקצועי ונטול פניות, הנתונים הגולמיים נחזים כאובייקטיביים וניטרליים (גם אם הם עשויים להיות עמומים או בלתי מדויקים).⁴⁶ יש להם ארשת מדעית, שמעניקה לחוק המבוסס עליהם סמכות, ויוצר רושם של ניטרליות פוליטית, ובהתאמה של הצדקה שחפה משיקולים לא ענייניים. הדבר נכון גם בנוגע לתוצר הסופי של ההליך הסטטיסטי בחקיקה – המודל או הנוסחה – והאופן שבו הוא נתפס כמדעי, כ"אמיתי", וכמשקף את המציאות. הכוח הנורמטיבי, המצדיק, של מודלים סטטיסטיים, המגלמים אובייקטיביות מדעית ומתיימרים להציג אמת נטולת פניות, הוא בין היתרונות של ההסתייעות בחקיקה מבוססת נתונים.

חקיקה מבוססת נתונים תורמת להגשמת התכלית החברתית של החוק כמכשיר לתיאום וארגון.⁴⁷ בהתאם לרמת התחכום של המודלים הסטטיסטיים או הנוסחה שהחוק מתבסס עליה, ככל שהחוק "נתפר" ברזולוציה גבוהה יותר הוא מאפשר סטנדרטיזציה ברמה גבוהה יותר, הן במונחים של יישום שוויוני שלו והן במובנים של הקצאה יעילה יותר של זכויות וחובות משפטיות, בהתאם לקריטריונים מדידים וברורים.

יתרון נוסף הוא הקטנת מרחב שיקול הדעת האנושי של השחקנים השונים המיישמים את החוק. הקריטריונים האחידים והשטוחים מצמצמים את הרקמה הפתוחה של השפה המשפטית, ומקטינים את ההסתברות לאכיפה בררנית,⁴⁸ ולפרשנויות

45 ראו למשל כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 186–198; OSKAR MORGENSTERN, ON THE ACCURACY OF ECONOMIC OBSERVATIONS 13–61 (2nd Revised edition, 1965).

46 ראו למשל Paul Starr, *The Sociology of Official Statistics*, in THE POLITICS OF NUMBERS 51 (William Alonso & Paul Starr eds., 1980).

47 ראו הרמה הראשונה של הפונקציה המארגנת של החוק והמשפט לפי פוסתמה, Gerald J. Postema, *Coordination and Convention at the Foundations of Law*, 11 J. LEGAL STUD. 165 (1982); 4 JOHN FINNIS, PHILOSOPHY OF LAW: COLLECTED ESSAYS 66–73 (2011); RICHARD H. MCADAMS, THE EXPRESSIVE POWERS OF LAW: THEORIES AND LIMITS (2015).

48 וורבק, למשל, טוען שמערכות הדירוג החברתי הסיניות נועדו לצמצם את מרחב התגובה האפשרי של רשויות אכיפת החוק. Kevin Werbach, *Orwell that Ends Well? Social*

מרחיבות של החוק הסוטות ממטרות המקורית. הסתמכות רגולטורית על מודלים סטטיסטיים מאפשרת מענה רגולטורי אחיד ושוויוני לאזרחים, שאינם תלויים במזל, ברצון הטוב של הרגולטור או בשיקולים בלתי עניינים אחרים. הקטנת המרחב הפרשני יכולה לבוא לידי ביטוי גם באיונו המוחלט – כפי שנעשה בעניין מס ריבוי דירות (שם המודל הסטטיסטי נועד להחליף שיקול דעת שמאי אנושי).

ואולם, חלק מהיתרונות של חקיקה מבוססת נתונים שנמנו לעיל עשויים להיות תמונת מראה של חסרונותיה. רמת הפירוט הגבוהה של המודלים הסטטיסטיים, וביתר שאת זו של האלגוריתמים, סותמת את נקבוביות הרקמה הפתוחה של החוק, מקטינה את חופש התמרון האנושי, את זכותו של הפרט "לכתוב את סיפור חייו",⁴⁹ ושמה אותו בסד אקטוארי שמוכתב על ידי סיפורי חייהם של אחרים.

תהליך התרגום בין המציאות למודל המתמטי טומן בחובו בעיות הנובעות מהיחס שבין המפה לטריטוריה – מידע שנאבד בתהליך הקידוד, משתנים רכים שמושלכים הצידה,⁵⁰ רמת דיוק פחותה – חלקן אינהרנטיות לתהליך וחלקן אילוצים פרגמטיים של מתודולוגיה מחקרית. ההישענות האפריורית על מודל כמותי מחייבת מדידה של המציאות בכלים כמותיים, אשר במסגרתה יש לייצר לפעמים סטנדרטים כמותיים או קטגוריאליים אחידים. אגב כך, משתנים רכים עלולים להתעוות כדי להתאים לסד הקטגוריאלי הנוקשה או להיות מושלכים הצידה לחלוטין.⁵¹

ואולם, במסגרת חקיקה מבוססת נתונים, המודל לא רק מתרגם את המציאות האמפירית למערך של משתנים כמותיים, אלא גם מתרגם עולמות שווי שונים. בולטנסקי ות'בנו (Boltanski and Thévenot) מונים שישה "עולמות שווי" אנושיים, נפרדים בקואורדינטות, במוסכמות ובערכים שלהם, שבמסגרתם ניתן להגיע להסכמה פוליטית בנוגע לסוגיה מסוימת.⁵² חקיקה באשר היא מתרגמת בין "עולם השווי

Credit as Regulation for the Algorithmic Age, 2022 U. ILL. L. REV. 1417, 1425 (2022).

49 ע"א 10064/02 מגדל חברה לביטוח נ' אבו חנא, פ"ד ס (3) 13 (2005). ראו גם אליעזר ריבלין וגיא שני "תפיסה עשירה של עקרון השבת המצב לקדמותו בתורת הפיצויים הנזיקיים – המקרה של שימוש בסטטיסטיקה בפסיקת פיצוי לפגע קטין" **משפט ועסקים** י 499 (2009).

50 משתנים רכים הם כאלו שקשה עד בלתי אפשרי לכמת אותם. קל עד מאוד לבחון את כמות המתגייסים לצבא, אך מדידת המשתנה הסובייקטיבי, הבלתי מוגדר, של רמת המוטיבציה שלהם לשירות עשויה להיות מורכבת יותר. בעוד שייתכנו פתרונות מתודולוגיים לאתגרים של משתנים רכים, קיים הפיתוי בקרב חוקרים להתעלם מהם. אולם השמטת משתנים רכים בהיעדר נתונים מספריים על אודותיהם היא "הדרך המובטחת לגבולות מודל צרים, תוצאות מוטות, והתנגדות למדיניות". John D. Sterman, *All Models are Wrong: Reflections on Becoming a Systems Scientist*, 18 Sys. DYNAMICS REV. 501, 523 (2002).

51 ראו למשל Tribe, לעיל ה"ש 7, בעמ' 1361–1365.

52 LUC BOLTANSKI & LAURENT THÉVENOT, ON JUSTIFICATION: ECONOMIES OF WORTH 159–212 (2006).

האזרחי", המבטא את הערכים הקולקטיביים הנורמטיביים, לבין עולמות שווי אחרים (דוגמת "עולם השווי השוקי", "עולם השווי התעשייתי", "עולם השווי של המוניטין"). כאשר התיווך בין שתי מערכות קואורדינטות נפרדות של עולמות שווי שונים נעשה באמצעות חקיקה מבוססת נתונים, הרגישות של הקידוד ובחירת המשתנים גבוהה יותר, מאחר שנדרשת מידה מסוימת של "השוואה בין תפוזים ותפוחים". הגרנולריות שנדרשת ממודל סטטיסטי (קיטוע הרצף הרך של המציאות למקטעים כמותיים מדידים), שעליו לכסות את הטריטוריה כולה, עשויה להציף פערים באופן שבו ערכים אלו מכומתים ומושווים.

חיסרון נוסף של חקיקה מבוססת נתונים הוא הטיית השיח של מקבלי ההחלטות. הישענות נרחבת על כלים מתמטיים בחקיקה עשויה להשפיע על סוג השאלות שמקבלי החלטות, מעצבי מדיניות ואף שופטים שואלים בנוגע לסוגיות המובאות אל דלתותיהם, בבחינת "הב לילד פטיש, והכול ייראה לו בר־חבטה".⁵³ כך למשל חשיבה כמותית השפיעה על הגדרת יעדי הניצחון במלחמת וייטנאם באמצעות ספירת הרוגי האויב.⁵⁴ כפי שצוין לעיל, יש בכך כדי להגביר את סיכויי ההתעלמות ממשתנים רכים או הרדוקציה שלהם לקטגוריות נוקשות. נוסף על כך, אפשר שהדבר ימקד את תשומת הלב של מקבלי ההחלטות האנושיים במשתנים חליפיים (proxy) בני־כימות (למשל ההסתברות שפלוני נמצא בזירת הרצח), תחת הנתון שאותו הם אמורים לייצג (ההסתברות שפלוני הוא הרוצח).⁵⁵ יתר על כן, כאשר תוארו לעיל היתרונות של חקיקה מבוססת נתונים, אלו נמנו בהנחה שההליכים המדעיים – איסוף הנתונים, טיובם, והפרוצדורה הסטטיסטית עצמה – הם תקינים. כך למשל תיתכנה טעויות במתודולוגיה המחקרית. האם ההנחה כי הקורלציה בין משתנה המטרה למשתנים המסבירים מעידה על סיבתיות היא נכונה? ואם כן, האם היחס בין סיבה לתוצאה הוא זה שקיים במודל? בלבול בין סיבה לתוצאה עשוי להביא לאפקט פיגמליון בחקיקה, המסתמכת על מודל היוצר את המציאות שאותה הוא מנבא.

האם המודל הסטטיסטי מניח אי־תלות בין המשתנים, או תלות, והאם זו נבדקה?⁵⁶ אם המשתנים תלויים זה בזה – כלומר, יש ביניהם מתאם סטטיסטי – אך נבחר מודל שמתעלם מכך, הדבר עשוי להביא לתוצאות שגויות בחיזוי. בהקשר של מס ריבוי דירות, למשל, ניתן לתהות אם ישנה תלות פנימית בין המדדים שנלקחו מהלמ"ס (מדד הפריפריאליות ומדד סוציאקונומי). אפשר, כמובן, שהמודלים השונים מניחים שישנה תלות כזו, אך בהעדר שקיפות של תהליכי העבודה הפנימיים במשרד האוצר שבמסגרתם פותחו הנוסחאות של מס ריבוי דירות, אין לדעת זאת לבטח.

ABRAHAM KAPLAN, THE CONDUCT OF INQUIRY: METHODOLOGY FOR BEHAVIOURAL SCIENCE 28 (1998).

Kathrin Cresswell et al., *Evaluation of Digitalisation in Healthcare and the Quantification of the "Unmeasurable"*, 38 J. GEN. INTERNAL MED. 3610 (2023).

ראו למשל Tribe, לעיל ה"ש 7, בעמ' 1366–1366.

שם, בעמ' 1368–1366.

ודוק, גם כאשר המתודולוגיה הסטטיסטית חפה מכל בעיה, תיתכנה טעויות בהכנת הנתונים שעליהם היא מתבססת. מודלים עשויים להיות מוטעים בשל מידע חלקי, חסר או "מלוכלך" (היינו, בסיס נתונים שרמת אמינותו פחותה בשל פגמים באופן איסוף המידע).⁵⁷ הטיית המודל עשויה לאיין באחת את אחד היתרונות המובהקים של חקיקה מבוססת נתונים – שוויון עקרוני ביישום החוק, ומשכך, יש להקפיד ביתר תשומת לב על הליכי חקיקה המבטיחים שהמודל שעליהם היא מסתמכת אינו מוטעה. מראית העין המדעית והאובייקטיבית של דברי חקיקה מבוססי נתונים, גם כאשר אלו כוללים טעויות המביאות להטיות ביישומם של החוק, עשויה ראשית להקשות על קבוצות מוגנות להראות שנגרם להן עוול במקרים של הטיות ביישום. שנית, משנתגלתה טעות כזו לאחר מעשה, הדבר עשוי לערער את האמון בחקיקה מבוססת נתונים באשר היא, גם כאשר היא חפה משגיאות.

בזהירות מסוימת יש לציין חיסרון נוסף של חקיקה מבוססת נתונים: היא תוצר של הליך פרלמנטרי. תוקפו של המודל הסטטיסטי/כמותי עשוי להשתנות במהרה נוכח המציאות, לעיתים בשל האפקטים שיוצרת החקיקה עצמה. משתנים או נתונים חדשים עשויים להתווסף או להתעדכן, ואפשר שהדברים יעלו כדי צורך בעדכון החקיקה כדי לבצע התאמות במודל. משכך, ייתכנו מצבים שבהם עדכון המודל המוטמע בחוק לא יתאפשר בשל נסיבות פוליטיות שמונעות או מעכבות את הליך החקיקה. אומנם, הקצב האיטי שבו החוק מדביק את המציאות אינו תופעה ייחודית לחקיקה מבוססת נתונים. עם זאת, מאחר שזו אינה מושתתת על עקרונות כלליים שמהם ניתן להקיש פרשנות נוכח נסיבות חדשות, פערי הזמנים ב"עדכון הגרסה" של החוק עשויים להשפיע על האפקטיביות ואף על מידת ההוגנות שלו, אם הוא מבוסס על מודל מתמטי שהתיישן ומידת ההלימה שלו את המציאות פחותה במידה ניכרת משהייתה. הסטטיות של החוק עלולה להכשיל את הדינאמיות של הסטטיסטיקה.

מקל וחומר, חיסרון זה מרכזי גם בנוגע לחקיקה אלגוריתמית, מונחית נתונים. פרישה נרחבת של מיקרו דירקטיבות מותאמות אישית למושאייהן עשויה להשפיע על התחום שאותו הן מבקשות להסדיר באופן מהיר ובהיקף ניכר, ובהתאמה להשפיע על תוקף המודל. ככל שהמודל האמור מוסיף להתעדכן לאחר פרישתו באמצעות למידת מכונה, מתעורר קושי בעיצוב מנגנוני פיקוח פרלמנטריים אפקטיביים המבטיחים שכוונת המחוקק באה לידי ביטוי במודל שמתאים את הוראות החוק לכל אזור, וזאת בנוסף לאתגר העכירות האלגוריתמית,⁵⁸ שמקשה על היכולת להנדס לאחר את המודל ולהבין כיצד הוא מתאים את הוראותיו לכל אזור ואזור.

57 דוגמה שכיחה לשימוש במידע חלקי היא הסתמכות מחקרית על בסיסי נתונים שקל להשיג. הקלות בהשגתם נובעת מהבדל מבני בינם לבין הנתונים החסרים, שהיעדרם משפיע באופן מהותי על התוצאות. ראו למשל Mireille Hildebrandt, *Algorithmic Regulation and the Rule of Law*, 376 PHIL. TRANS. R. SOC. 1, 6–7 (2018).

58 עכירות אלגוריתמית היא תולדה של האופן שבו למידת מכונה פועלת. להבדיל מקוד מבוסס "אם-אז", שניתן להתחקות אחריו, מסובך ככל שיהיה, על הרציונל שבו, למידת מכונה מבוססת על נתונים. מטרתה היא לייצר מודל סטטיסטי שמבא בצורה הטובה

בהשלמה לדיון בחסרונותיה של חקיקה מבוססת נתונים בהיותה חקיקה, אפשר לציין חיסרון נוסף של חקיקה מבוססת נתונים והוא היותה מבוססת על נתונים. כדי שהמודל יוכל לייצג נאמנה את המציאות, אנו עשויים להידרש לכמה שיותר נתונים (הן מבחינת סוגי הנתונים והן מבחינת נפח הנתונים מכל סוג). כאשר נתונים אלו כוללים מידע רגיש או אישי,⁵⁹ עיבודם חוסה תחת עקרונות דיני הגנת הפרטיות, ובפרט תחת עקרון צמידות המטרה.⁶⁰ חוק הגנת הפרטיות, למשל, קובע ש"שימוש בדיעה על עניניו הפרטיים של אדם או מסירתה לאחר, שלא למטרה שלשמה נמסרה" הוא בגדר פגיעה בפרטיות.⁶¹ התקנות הכלליות בדבר הגנת מידע (GDPR) קובעות ש"מידע אישי [...] לא יעובד באופן שהוא בלתי תואם [עם המטרות שלשמן הוא נאסף]".⁶² לפיכך, בתהליך הכנת דברי חקיקה מבוססי נתונים יש לנקוט משנה זהירות בכל הנוגע לשימוש בנתונים אישיים, ולהקפיד על הלימה עם עקרון צמידות המטרה.⁶³ עם זאת, ככל שמדובר בחקיקה מבוססת נתונים ולא בחקיקה מונחית נתונים, אפשר שבחלק מהמקרים זו תסתמך על מדדים סטטיסטיים מצרפיים (שעשויים להיות תוצר של נתונים אישיים), שאין בהם משום פגיעה בפרטיות.

שאלה אחרונה שיש לשאול לגבי חקיקה מבוססת נתונים היא בנוגע לכוח הנורמטיבי שלה. האם נוסחה לוגריתמית יכולה להכווין התנהגות אנושית? האם הסובייקט המשפטי מסוגל להפנים את הנוסחה ולחלץ מתוכה את הערכים הכלליים של השיטה? לחוק יש תפקידים אקספרסיביים שעשויים להתערער אם לשונו אינה ברורה או טכנית מדי, לרבות אם הוא ביטוי באמצעות נוסחה.⁶⁴ האם לחוק שדרך פעולתו נשענת על נוסחה שתוארה בעיתונות הכלכלית כ"10 בחזקת שיגרון"

ביותר על בסיס נתוני אימון. המודל הזה אינו מוסבר על ידי המכונה באופן שיכול לתת לנו תובנות על השאלה כיצד הוא נבנה. ראו גם VIRGINIA DIGNUM, RESPONSIBLE ARTIFICIAL INTELLIGENCE: HOW TO DEVELOP AND USE AI IN A RESPONSIBLE WAY 59 (2019).

59 ראו למשל ס' 7(2) לחוק הגנת הפרטיות, התשמ"א-1981; רחל ארידור הרשקוביץ ותהילה שוורץ אלטשולר **הצעת חוק הגנת הפרטיות, התשע"ט-2019** 33–35 Regulation (EU) 2016/679 of the European Parliament and of the council of; (2019) 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 (General Data Protection Regulation), 2016 O.J. (L 119) (1) (EU) (להלן: GDPR).

60 מיכאל בירנהק **מרחב פרטי: הזכות לפרטיות בין משפט לטכנולוגיה** 52 (2010); ארידור הרשקוביץ ושוורץ אלטשולר, לעיל ה"ש 59, בעמ' 51–52; ס' 5(b) ל-GDPR.

61 ס' 2(9) לחוק הגנת הפרטיות.

62 ס' 5(b) ל-GDPR. המובאה מה-GDPR תורגמה על ידי המחבר.

63 ראו מיכאל בירנהק "פרטיות מותאמת אישית" **עיוני משפט** מב 381, 385–387 (2020).

64 ראו למשל Tribe, לעיל ה"ש 7, בעמ' 1374, 1391.

כאינדיקציה לחוסר נהירותה להדיוטות⁶⁵ ישנו כוח סימבולי או מחנך הראוי לדבר חקיקה? יתר על כן, טרייב הצביע על כך שכאשר שיטות לקביעת אשם מוסרי מסתייעות בראיות שמתקשות לשכנע הדיוטות בצדקתן (כגון ראיות סטטיסטיות), יש בכך כדי לצבוע את מערכת המשפט באור בלתי אנושי.⁶⁶ בהתאמה, חקיקה שמוצדקת באמצעות טיעונים מתמטיים המובנים רק לעיניו של מומחה מאבדת מהכוח ההצדקי שלה ותורמת לדהומניזציה של החוק. הגירעון ההסברתי של הוכחה ראייתית או של כלל מתמטי מקשה על היכולת לבטוח במערכת כולה.⁶⁷ עם זאת, השימוש בנוסחה לוגריתמית בהקשר של מיסוי נכסי דלא נידי, תחום שנתפס כ"טכני" יותר מהדין הפלילי וכנחלתם של מומחים, עשוי להפיג את החששות מהכרסום בכוחו המחנך של החוק.

ג. גלגוליה של נוסחה

בחודש יולי 2016 פורסם בעיתונות הכלכלית שבמסגרת דיוני ההכנה במשרד האוצר לחוק ההסדרים הועלתה הצעה למסות משקיעים בדירות ולהשית מס בסך 1,000 שקלים חדשים בחודש על בעלי דירה שלישית ומעלה, בגין כל דירה כזו, וזאת כדי לצנן את הביקושים.⁶⁸ ההצעה זכתה לביקורת בשל עצם הטלת המס, או בשל חששות מגלגולו על השוכרים,⁶⁹ אך גם בשל הרגרסיביות שבהטלת מס אחיד על כל דירה שלישית ומעלה, באופן המתעלם מהשונויות בערך הדירות.⁷⁰ נוכח הביקורות על אחידות המס, תזכיר חוק ההסדרים שהופץ כשלושה חודשים לאחר הפרחת "בלון הניסוי" בתקשורת הציע בסיס אחר לחישוב המס, לפי נוסחה.⁷¹ לפי דברי ההסבר בתזכיר, שיעור המס על דירה שלישית ומעלה יעמוד על אחוז מערך ממוצע של דירה למטר רבוע בישראל, כשהוא מוכפל במקדם המבוסס על מדד חברתי-

65 אריק מירובסקי "10 בחזקת שיגעון: הנוסחה החדשה והמפחידה של מס דירה שלישית" **TheMarker** (18.12.2016) <https://bit.ly/4cKGoxP>.

66 Tribe, לעיל ה"ש 7, בעמ' 1376.

67 Peter Tillers, *Trial by Mathematics – Reconsidered*, 10 L. PROBABILITY & RISK 167, 171 (2011).

68 מוטי בסוק ואריק מירובסקי "תוכנית באוצר: מס מיוחד על בעלי שלוש דירות ויותר" **TheMarker** (6.7.2016) <https://www.themarker.com/news/2016-07-06/ty-article/>; מוטי בסוק וצבי זרחיה "כחלון החליט: מס של 1,000 שקל בחודש לבעלי 3 דירות" **TheMarker** (22.7.2016) <https://www.themarker.com/realstate/2016-07-22/ty-article/0000017f-f1ed-d223-a97f-fdfdaa280000>.

69 בנצי גיספאן "המס יגולגל אליכם שוכרי הדירות" **Ynet** (10.8.2016) <https://www.ynet.co.il/articles/0,7340,L-4839830,00.html>.

70 ראו למשל עמרי מילמן "זינוק של 10% במספר בעלי שלוש דירות ויותר מאז 2015" **כלכליסט** (10.8.2016) <https://www.calcalist.co.il/sport/articles/0,7340,L-3695267,00.html>.

71 תזכיר חוק ההתייעלות הכלכלית (תיקוני חקיקה לשנים 2017 ו-2018), התשע"ז-2016, פרק 1: מס ריבוי נכסים (5.9.2016) (כך במקור. להלן: תזכיר חוק מס ריבוי נכסים).

כלכלי ומדד פריפריאלי שמפרסמת הלמ"ס.⁷² הנוסחה, שתיאורה נמסר באופן מילולי בלשון התזכיר, נראית כך, כאשר P הוא מדד הפריפריאליות של הלמ"ס, ו-S הוא המדד החברתי-כלכלי שלה:

$$15,148*(1.49-0.2362*P-0.01037*S+0.0238*P^2+0.0066*S^2+0.0042*P*S)$$

הערך הממוצע למטר רבוע של דירה בישראל – 15,148 ש"ח – צוין במפורש בלשון התזכיר, ונקבע שהוא יעודכן מדי שנה בהתאם לשינוי בפרמטר רביעי – מדד מחירי הדירות.⁷³ כמו כן נקבע חסם עליון ולפיו מס ריבוי הדירות לא יעלה על סך של 18,000 ש"ח בעד כל דירה.⁷⁴

הנוסחה שבתזכיר נועדה לאמוד את ערך הנכס לפי שלושה פרמטרים עיקריים (שטח, מדד פריפריאליות של הלמ"ס, והמדד החברתי-כלכלי של הלמ"ס) ולאורך זמן גם לפי פרמטר רביעי – מדד מחירי הדירות, שמעדכן את הערך הממוצע לדירה. כבר בשלב זה ניתן להעיר שלא נקבעו הוראות בחקיקה באשר להליכי העדכון של כל שלושת המדדים של הלמ"ס, שבהם תלויה נוסחה זו.

הצגת הנוסחה החדשה שבתזכיר לא שיכנעה את הביקורות האידאולוגיות על עצם הטלת המס, ואולם בחינת רגישות פשוטה של דני בן-שחר ורוני גולן ממכון אלרוב, שנועדה לבדוק עד כמה הנוסחה המוצעת רגישה לשינויים בפרמטרים שלה, ובדקה מהו גובה המס שיוטל לפי כל השילובים האפשריים שניתן להרכיב מאשכול סוציאקונומי ומדד פריפריאלי, מצאה שב-58 מתוך 100 מתוכם הנוסחה מחזירה ערך הגבוה מהחסם העליון.⁷⁵

ודוק, ידוע ששווייה של דירה אינו מסתמך אך ורק על שלושת הפרמטרים האלו. כבר באותו בניין ייתכנו הבדלים הנובעים משיקולים של קומה וכיווני אוויר.⁷⁶ בשכונה מסוימת ייתכנו פערים הנובעים ממרחק למוסדות לימוד, מגבלות או זכויות תכנוניות שונות, וכדומה. חוסר היכולת של המפה שהציע משרד האוצר לשקף במדויק את הטריטוריה היה ידוע מראש,⁷⁷ ואף על פי כן, בהצעת חוק ההסדרים שעלתה לשולחן ועדת הכספים הופיעה נוסחה אחרת לגמרי.

72 ס' 4 לתזכיר חוק מס ריבוי נכסים.

73 ס' 4(ג) לתזכיר חוק מס ריבוי נכסים.

74 ס' 4(ב) לתזכיר חוק מס ריבוי נכסים.

75 דני בן-שחר ורוני גולן "נוסחת המס על דירה שלישית מסובכת ומעוותת" *TheMarker* (23.9.2016). <https://bit.ly/3zWQzfx>. הגם שבדיקה זו מתעלמת מהשכיחות של שילובים אלו בפועל, במסגרת בדיקתם של בן-שחר וגולן נבחנו גם עסקאות הדיוור שנערכו בין ינואר 2015 למארס 2016, ונמצא שהנוסחה שבתזכיר מעריכה את שוויין של הדירות הערכת יתר בשיעור ממוצע של 49% למעלה ממחיר השוק.

76 ראו למשל דבריה של נציגת משרד האוצר, גב' מעין נשר, בפרוטוקול ישיבות 569–581 של ועדת הכספים, הכנסת ה-20, 371 (15.12.2016) (להלן: פרוטוקול הדיון בחוק).

77 ראו חורחה לואיס בורחס *דברי ימי תועבת העולם* 119 (רנה ליטוין מתרגמת, עם עובד 1987); ז'אן בודריאר *סימולקרות וסימולציה* 7–9 (אריאלה אזולאי מתרגמת, הקיבוץ המאוחד 2007).

ככל הנראה, הביקורת המקצועית שלעיל הביאה את האוצר להציע נוסחה מורכבת יותר ופחות אינטואיטיבית. המודל הפשוט יחסית למראה הומר בביטוי מתמטי מורכב הרבה יותר. הפעם הנוסחה חישה את המס שיוטל על כלל הנכס (ולא על מטר רבוע) באופן הבא, כאשר s ו- p הם אותם מדדי הלמ"ס ו- x הוא שטח הדירה:

$$10^{\left(4.421 - \frac{2.21ps + 79.7 \log_{10} x + 7.8p + 8s + 2.8p^2 + 2s^2}{100}\right)}$$

אומנם גם במקרה זה נעשה שימוש באותם פרמטרים, אך בשלב זה היכולת של מי שלא עבר הכשרה מתמטית על-תיכונית לתפוס באופן אינטואיטיבי את משמעות הנוסחה פחתה במידה ניכרת.⁷⁸ נוסחה זו, שתיאורה בהצעת החוק הקיף שני עמודים של מלל, היא זו שעמדה לפני ועדת הכספים בדיון נשוא עניין **קוונטינסקי**, והיא נותרה כשהייתה עם כניסתו של חוק ההסדרים לתוקף. הדיון בחוק מס ריבוי דירות בוועדת הכספים היה ארוך, ונוכח סד הזמנים הקצר שבו ניסו נציגי הממשלה להעבירו (במסגרת חוק ההסדרים) הוא היה אף רווי בניסיונות להאיצו מבלי להתעכב על פרטיו.⁷⁹ העיסוק של הוועדה במהות של הנוסחה עצמה, לא כל שכן בירידה לפרטיה, נקטע לבקשת יושב הראש שלה לצאת להפסקה.⁸⁰

סעיף נוסף בחוק מס ריבוי דירות אפשר לשר האוצר באישור ועדת הכספים של הכנסת לשנות בצו את התוספת לחוק שבה הוגדרה הנוסחה.⁸¹ הסיפא של הסעיף מורה ש"ובלבד שבשינוי הנוסחה שלפיה יחושב הסכום הקובע כאמור בסעיף קטן (א) או רכיב מרכיבה, יובא בחשבון הערך המקורב של דירת מגורים באזור שבו נמצאת הדירה, לפי אומדנים סטטיסטיים". לשון ההוראה מבקשת לצייר נוסחאות עתידיות ככאלו שיש בהן מרכיב שקושר בין שיעור המס לאומדן ערך הנכס בהתבסס על פרמטרים גאוגרפיים. ואולם ההוראה שותקת בנוגע לטיבו של האומדן הסטטיסטי (כלומר, השיטה המתמטית שבה ייעשה שימוש כדי להגיע לקירוב המבוקש), משקל הפרמטר הגאוגרפי בהשוואה לפרמטרים נוספים שעשויים להיות רלוונטיים, ושאלות הנוגעות לעדכון המודל, תיקופו, והנתונים שעליו הוא מסתמך.

יתר על כן, כבר בסמוך להעברת החוק מיהר משרד האוצר לעדכן את הנוסחה באמצעות המנגנון שהתווה בחוק.⁸² בדיון בוועדת הכספים נטען שהצורך בעדכון עלה לאחר ישיבת ועדת הכספים שבה נדון החוק עצמו, נוכח פרסום נתונים חדשים של הלמ"ס.⁸³ בדיון בוועדת הכספים הסבירו נציגי משרד האוצר שהעדכון בנוסחה נובע

78 ראו למשל הסיקור העיתונאי לנוסחה, לעיל ה"ש 65.

79 עניין **קוונטינסקי**, לעיל ה"ש 1, פס' 90 לפסק הדין של השופט סולברג.

80 פרוטוקול הדיון בחוק, לעיל ה"ש 76, בעמ' 372.

81 ס' 118 (ב) לחוק מס ריבוי נכסים.

82 צו ההתייעלות הכלכלית (תיקוני חקיקה להשגת יעדי התקציב לשנות התקציב 2017 ו-2018) (שינוי התוספת לפרק י"ב לחוק), התשע"ז-2017, ק"ת 7767.

83 ראו דבריה של עו"ד אפרת פרוקצ'יה בפרוטוקול ישיבה 604 של ועדת הכספים, הכנסת ה-20, 5.1.2017.

מעדכון מדדי הלמ"ס, אך גם משינוי בפרמטרים עצמם. המדד החברתי-כלכלי שונה באופן שאפשר להתייחס לאזורים גאוגרפיים ברמת דיוק גבוהה יותר.⁸⁴ הסבר נוסף לנוסחה החדשה היה בניסיון לדייק את האומדן לנכסים בפריפריה. להלן הנוסחה החדשה שבצו:

$$\frac{30 \cdot 10^{\left(4.194 - \frac{2.2(p+4)(s+4) - 5.6(4+4) + 79.7 \log_{10} x + 1.2(p+4) + 2.8(p^2 + 2s^2)}{100}\right)}}{30 + (7.482 - s + 4) + (6.728 - p + 4)}$$

במהלך הדיון בצו ביקשו חלק מחברי ועדת הכספים לראות כיצד הנוסחה פועלת בנוגע לנכסים המוכרים להם,⁸⁵ או לקבל פירוט של ההשפעה של השינוי בנוסחה על סך כל גביית המס המשוערת, ולראות מקרי קצה לכאן ולכאן. חלק הביעו תרעומת מהפערים שבין האומדן שבנוסחה לשווי בפועל של הנכסים שאותו היא מנסה לאמוד.⁸⁶ נציגי האוצר, שהתעקשו שהנוסחה החדשה מעדנת את גובה המס בפריפריה, האיצו בחברי הוועדה לאשר את הצו בהבטיחם שישבו בעתיד עם חברי הכנסת ויערכו סימולציות, ככל שיעלה הצורך.⁸⁷ יש לציין שלהבדיל מהנוסחה המקורית שבתזכיר, לא ניתן לציבור זמן הולם ללמוד אותה ולהגיב לה. גם לפני חברי הכנסת לא נראה שהוצגו מלוא המשמעויות שלה.

נוסף על כך, הגם שהשינוי בנוסחה הוסבר על ידי נציגי משרד האוצר בין השאר כתוצאה של הגדלת הרזולוציה הגאוגרפית של המדדים, נראה שעצם עדכון המדדים על ידי הלמ"ס (שנעשה אחת לכחמש שנים) העלה צורך בעדכון המקדמים במודלי הרגרסיה שבנוסחה. הדבר עשוי להצביע על חשש לתופעה סטטיסטית המכונה התאמת יתר (over fitting).⁸⁸ התאמת יתר היא מצב שבו המודל מתאים יתר על המידה לנתוני האימון, ויודע לנבא אותם בדיוק רב, אולם אם יוזנו למודל כזה נתונים אחרים מהנתונים שעליהם הוא אומן הוא עלול לטעות. זאת משום שהמודל זיהה רעשים אקראיים בנתוני האימון כפרמטרים מסבירים.⁸⁹ האם עדכון המדדים של הלמ"ס הביא לעיוות תוצאות המודל? אין ביכולתנו לדעת זאת ללא תיעוד של פיתוח המודל שעל בסיסו התקבלה הנוסחה שבחוק, ושל פיתוח המודל שלפיו עודכנה הנוסחה בצו. רמת

84 ראו דבריה של רחל קופר, שם, בעמ' 97.

85 ראו למשל דבריו של ח"כ מיקי מכלוף זוהר, שם, בעמ' 108.

86 ראו למשל דבריו של ח"כ דוד ביטן, שם, בעמ' 124, 127.

87 ראו דבריו של ערן ניצן, שם, בעמ' 129.

88 ראו למשל Tom Dietterich, *Overfitting and Undercomputing in Machine Learning*,

27 ACM COMPUT. SURVS. 326 (1995).

89 לדוגמה, מודל למידת מכונה שאומן לזהות ילדים בתמונות, ונתוני האימון שלו הורכבו מתמונות של ילדים שמשחקים בגן שעשועים, עשוי להסתמך על קיומן של מגלשות בתמונה כדי לזהות דמות כ"ילד". אם נציג למודל תמונה של ילד יושב בביתו וקורא ספר, המודל עשוי שלא לזהות אותו כילד.

הדיוק של המקדמים השונים, המגיעה לשלוש ספרות אחר הנקודה העשרונית, עשויה לתרום לחשש הזה.⁹⁰

ד. חוק מס ריבוי נכסים כחקיקה מבוססת נתונים

הנוסחה שעודכנה בצו האריכה ימים עד לביטולו של חוק מס ריבוי נכסים בעניין קוונטינסקי.⁹¹ פסק הדין, למעט אזכור אגבי של אותה "נוסחה מתמטית-סוציולוגית", התעלם מתוכנו המהותי של החוק והעמיד את ההנמקה שלו בפגמים שנפלו בהליך החקיקה. חלק זה של המאמר אינו נוקט עמדה כלפי תוכן החוק, אולם הוא מבקש להתעכב ביתר פירוט על האופן שבו מס ריבוי דירות הוא דוגמה לחקיקה מבוססת נתונים, כדי להמשיך ולמקם אותה בהקשר הרחב יותר של חקיקה מבוססת בינה מלאכותית, ולראות כיצד ניתן למתן חלק מהבעיות הייחודיות לסוג זה של חקיקה.

הנוסחה של חוק מס ריבוי דירות עוצבה תחת פרדיגמה של חקיקה מבוססת נתונים. הראיות שעליהן הסתמכו נציגי משרד האוצר לא נועדו לספק תימוכין למדיניות או לעצב אותה (זו הייתה ידועה מראש: מס בסדר גודל של כ-1,000 ש"ח לחודש לדירה, שאמור היה לשקף שנתי כ-1% משווייה), אלא לאפשר את הוצאתה לפועל באמצעות מודל. המדיניות שנבחרה היא להשית מס שנתי בשיעור של אחוז מערך הנכס, והמודל נועד לממש אותה. מבחינה זו, המודל אינו משמש ככלי ראייתי כדי לתמוך במדיניות או בתוצאות החקיקה. למען הסר ספק, הנוסחה שבמס ריבוי דירות אינה מקרה של חקיקה מונחית נתונים כפי שתואר בחלק ה' להלן – אין כאן הסתמכות על אלגוריתמים של למידת מכונה, וזהו אינו מודל שממשיך לעדכן ולטייב את עצמו באורח אדפטיבי, לאור נתונים חדשים.

יתר על כן, זהו מקרה קל יחסית של חקיקה מבוססת נתונים. פונקציית המטרה של הנוסחה (ערך הנכס) הוגדרה באופן חיצוני, והיא מדידה. מצב הדברים עשוי להיות מורכב יותר כאשר המודל המתמטי מנסה להגיע לתוצאה שאיכות המדידה שלה בעייתית או שנויה במחלוקת (למשל "מיקסום האושר המצרפי").

סיפור המעשה בגלגוליה של הנוסחה שבחוק מס ריבוי דירות מתאר תהליך ארוך של בחירת מודלים, שחלקים ממנו הוצגו לציבור. בתחילת התהליך, שיתוף הציבור הצליח לעורר שיח וביקורת. בתחילה הייתה זו ביקורת על הרגרסיבות שבמס אחיד של 1,000 ש"ח לחודש. בהמשך, לאחר הצגת הנוסחה הלינארית הראשונה, נשמעה ביקורת מקצועית על המודל.⁹² שיאו של התהליך, במעמד הדיון בוועדת הכספים נשוא עניין

90 מורגנשטרן העיר לגבי דיווחים סטטיסטיים של נתונים ברמת דיוק מופרזת כי "סוג זה של דיווחים יוצר את הרושם של איכות גבוהה במיוחד, אפילו 'מדעית', של העבודה [הסטטיסטית], ויש בה כדי להשרות אמינות גבוהה יותר על המסקנות": MORGENSTERN, לעיל ה"ש 45, בעמ' 63–64.

91 עניין קוונטינסקי, לעיל ה"ש 1, פס' 113 לפסק הדין של השופט סולברג.

92 לעיל ה"ש 75.

קוונטינסקי, נעשה ללא התעניינות פרלמנטרית בנוסחה עצמה (בין השאר, בשל הנסיבות שהביאו לפסילת החוק עצמו). גם כאשר תוקנה הנוסחה בצו, חברי הוועדה הובאו בפני עובדות מוגמרות, וחרף התעניינותם בהשלכות של המודל, נראה שלא ניתנה למחוקק שהות מספקת לבחון את המודל החדש.⁹³

בחירת המודלים השונים לאורך גלגולי הנוסחה לא הוסברה במלואה, וכך גם הבחירה בפרמטרים. לא שוחררו ניירות עבודה ולא ניתן פומבי לשיקולי הבחירה במודל, בפרמטרים וגם לא במשתנה המטרה. שאלה נוספת נוגעת לבסיסי הנתונים עצמם, ולנתונים שנכללו בהם. לעוסקים בתחום הנדל"ן בישראל מוכר הפער בין נתוני משרד האוצר לנתוני הלמ"ס בנוגע למספר הדירות החדשות הנמכרות, למשל.⁹⁴ לעיתים תיתכנה שגיאות במסדי הנתונים הממשלתיים, או במתודולוגיה של שליפת נתונים מתוכם.⁹⁵ היכולת לבחון באופן ביקורתי את דיוקם של המודלים השונים תלויה במידה לא מועטה בנתונים שעליהם הם מסתמכים. הדברים נכונים הן ביחס לבחינה מהותית של המודל המוצע והן לבחינת השפעות בלתי רצויות שלו, כגון הטיות.⁹⁶

ואולם מצבם של הנישומים, מושאי ההחלטה האקטוארית של הנוסחה שבחוק מס ריבוי דירות, עדיף על זה של הנישום בעניין **הר שמש** – שם ביקשה האגודה השיתופית הר שמש לקבל את הקוד שבאמצעותו חישובה רשות המיסים את השומה שלה, כדי לבדוק את נכונות החישוב.⁹⁷ במקרה של מס ריבוי דירות, רמת השקיפות טובה בהרבה: אין קוד מכונה שיש לחלץ מקרבי המערכת המנהלית, אלא יש להתמודד עם הנוסחה עצמה, הגלויה לעין כול. אולם האופן שבו הגיע המחוקק (או נציגיו של משרד האוצר) למודל הסטטיסטי של שווי הדירה לא היה ברור, ועל כן ההליך הדליברטיבי היה מוגבל. הנוסחה עצמה לא הוסברה מעולם למחוקק, אלא רק המוטיבציה לתיקון שלה בצו, והדבר פגע בהסברתיות (explainability) שלה,⁹⁸ וזאת בשים לב לכך

93 עם זאת, תיתכן עמדה ספקנית שלפיה שיקולים של תמריצים אלקטורליים ושל ניהול זמן, גם בהינתן שהות מספקת לחברי הוועדה להיערך לדיון, היו מונעים מחברי הוועדה להתעמק בחומר.

94 אורי חודי "אז כמה דירות באמת נמכרות בישראל בכל חודש? הגיע הזמן שהאוצר והלמ"ס יסכימו על נתון אחיד ואמין" **גלובס** (1.11.2018) <https://bit.ly/4cUZNr4>.

95 נמרוד בוסו "נתונים שגויים ממשרד השיכון הביאו לליקויים בנתוני הלמ"ס על מכירת דירות – הפרסומים הוקפאו" **TheMarker** (2.2.2017) <https://bit.ly/4deQOAG>. מהדיווח לא ברור אם מדובר בטעויות בבסיס הנתונים עצמו, או בליקויים בחילוף הנתונים האגרטיביים ממנו.

96 ראו למשל Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. 2218 (2019); כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 179–202.

97 עת"ם (מנהלי י-ם) 2663-05-19 **הר שמש מושב שיתופי בע"מ נ' הממונה על חופש המידע ברשות המיסים** (נבו 5.9.2019).

98 ראו פרל ושוורץ אלטשולר, לעיל ה"ש 42; כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 101–104.

שהנוסחאות (לרבות הנוסחה הליניארית הראשונה) מורכבות ולא אינטואיטיביות, עד שהן נדמות כשרירותיות.

לצד שאלות הנוגעות להסברתיות ולאיכות הנתונים והמודל, היבט נוסף שיש לתת עליו את הדעת בעניין חוק מס ריבוי דירות הוא עניין תיקונו של המודל ימים ספורים לאחר החקיקה. כאמור לעיל, רמת הדיוק של המקדמים בנוסחאות הסטטיסטיות השונות עשויה להעיד על התאמת יתר של הנתונים למודל, או על רצון להציג אותו כאובייקטיבי באמצעות מראית עין של דיוק. אולם המדדים שהמודל נסמך עליהם, כמו גם פונקציית המטרה שלו, מתעדכנים, אם באופן יזום באמצעות הלמ"ס ואם בשל טיבם של חיי הכלכלה, כך שערכי הדירות עשויים להשתנות מיום למחרתו. תיקונו של המודל בצו, לפי דברי נציגי משרד האוצר בדיון בוועדת הכספים, הגיב על שינוי ברמת דיוק המדידה של הלמ"ס. בהעדר מסמכי עבודה המתעדים את פיתוח הנוסחה, ובהעדר ביקורת פרלמנטרית מיוחדת של חברי כנסת – אין לדעת אם מדובר בהתאמת יתר.

יתר על כן, האם ראוי לעדכן או להחליף מודל באמצעות צו? הוראת סעיף 188(ב) לחוק ההסדרים אפשרה לשר האוצר להחליף את הנוסחה בכפוף לשמירה על אותה פונקציית מטרה (ערך דירה מקורב). הוראות החוק לא פירטו תנאים נוספים לשינוי הנוסחה, מדדים להשוואתה עם הנוסחה הקודמת, או קריטריונים המצדיקים החלפה זו. במשתמע, אלו נושאים שאמורים לעלות בדיוני ועדת הכספים, שאישורה נדרש לשם החלפת הנוסחה, אך ראוי היה לפתח מנגנון עדכון זהיר יותר, ואולי אף בחינה עיתית של תוקף הנוסחה. יתר על כן, אפשר שסעיף 188(ב) הוא הנתיב שבו חקיקה מבוססת נתונים אמורה לצעוד: הנחיה עקרונית-כללית למחוקק המשנה לבסס את חקיקת המשנה על מודל כמותי עם פונקציית מטרה מוגדרת. אך לצד הנחיה כזו יש להוסיף מנגנונים לבדיקת תוקפו של מודל כאמור ומידת ההתאמה שלו להנחיית המחוקק הראשי, ולכלול הוראות בנוגע לשקיפות המודל ולפיקוח פרלמנטרי על המודל ומימושו.

יש לציין שמטרת הנוסחה האפריורית הייתה לסכל הערכת שווי בידי שמאי אנושי, הפתוחה לערעורים ולפרשנויות. מתוך שמראש הוחלט לוותר על בחינה אנושית של הערכת השווי שבנוסחה, תיחומה בחסם עליון של מס בערך של 18,000 ש"ח לשנה היא דרך להימנע ממקרי קצה בעייתיים. אולם לא בכל סוגיה שחקיקה מבוססת נתונים באה לתת עליה מענה תהיה בהכרח אפשרות לפתור רעשים סטטיסטיים באמצעות ססתום מסוג זה, ואפשר שיהיה צורך במעורבות אנושית בבחינת יישום החקיקה⁹⁹. (Human in the loop)

99 על Human in the Loop ראו בקצרה High-Level Expert Grp. on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI* (Apr. 8, 2019), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, בנוסף ראו Fabio Massimo Zanzotto, *Viewpoint: Human-in-the-Loop Artificial Intelligence* 64 J. A.I. RES. 243 (2019).

ה. חקיקה מונחית נתונים: כשחקיקה מבוססת נתונים פוגשת בינה מלאכותית

כפי שצוין בחלק א', בקצהו השני של הספקטרום שתחילתו בחקיקה מבוססת ראיות מצויה חקיקה מונחית נתונים. חקיקה מונחית נתונים – בשלב זה בגדר רעיון תאורטי – נעשית בהסתמך על בסיס נתוני עתק עשירים במשתנים, באמצעות טכניקות של למידת מכונה, ותניב מודל מורכב בהרבה מתוצר של עיבוד סטטיסטי אנושי. מבחינות רבות, הפערים בין חקיקה מבוססת נתונים לבין חקיקה מונחית נתונים עשויים להצטייר כפערים של סדרי גודל, אך חקיקה מונחית נתונים מוסיפה אתגרים ייחודיים משל עצמה, והשינוי בסדרי הגודל – שאינו מוגבל רק לסדרי הגודל העצומים של המידע המזין את המודל אלא גם לכמות העצומה של הפרמטרים שהמודל משתמש בהם – עשוי להצריך התייחסות פרטנית משל עצמה.

הילדברנט מתארת שני סוגים של רגולציה אלגוריתמית.¹⁰⁰ הראשונה היא רגולציה מונחית קוד, על בסיס עצי החלטה ידועים מראש ודטרמיניסטיים, בבחינת "אם זה אז כך".¹⁰¹ המודלים של עצי ההחלטה עשויים להיות סבוכים ומורכבים, אך ניתן לחלץ מהם הסבר לגבי ההחלטות שמתקבלות באמצעותם. עם זאת, אפשר שלא יהיה די בהסברים אלו כדי להצדיק את ההחלטה. סוג שני של רגולציה אלגוריתמית היא רגולציה מונחית נתונים, שבה הקוד מושפע מנתוני האימון. אלו מודלים של למידת מכונה, שסובלים מעכירות אלגוריתמית.¹⁰²

אפשר שהערתו של סופר המדע הבדיוני קלארק כי לא ניתן להבחין בין טכנולוגיה מתקדמת דיה לבין קסם¹⁰³ תקפה גם בנוגע לטכנולוגיות סטטיסטיות.¹⁰⁴ אם, כפי שצוין בחלק הקודם שעסק בחקיקה מבוססת נתונים, היכולת האנושית להכיל ולהתמודד עם נוסחאות מורכבות בחקיקה היא מוגבלת, מקל וחומר שנוכח בעיית העכירות האלגוריתמית חקיקה מונחית נתונים עלולה להצטייר כשרירותית וסתומה שבעתים.

בינה מלאכותית מביאה עימה אופטימיזם טכנולוגי בלתי מבוטל. כך, למשל, אלארי ובקר מנסים לתאר דגם של אקוסיסטם משפטי עתידי מבוסס בינה מלאכותית שהולך ונעשה מבוזר יותר, מונחה נתונים יותר ומתוך כך, לטענתם, דינאמי ויעיל, צודק

100 Hildebrandt, לעיל ה"ש 57, בעמ' 2.

101 IFTTT – If This Then That.

102 Hildebrandt, לעיל ה"ש 57, בעמ' 3; לפירוט על בעיית העכירות האלגוריתמית ראו לעיל ה"ש 58.

103 Arthur C. Clarke, *Clarke's Third Law on UFO's*, 159 Sci. 255 (1968).

104 *People v. Collins*, 438 P.2d 33, 33 (Cal. 1968) דומה, ראו דברי בית המשפט בעניין *People v. Collins*, 438 P.2d 33, 33 (Cal. 1968): "בשעה שהמתמטיקה, שהיא קוסמת של ממש בחברתנו מבוססת המחשוב, מסייעת לבודקי העובדות בחיפושם אחר האמת, אל לה להטיל עליהם כישוף".

ודמוקרטי יותר.¹⁰⁵ אולם חזונות אופטימיים אלו, כמו זה שמציעים פורת ובן-שחר,¹⁰⁶ של חקיקה מונחית נתונים המאפשרת חוק בהתאמה אישית, או זה של קייסי וניבלט,¹⁰⁷ אודות "מיקרו דירקטיבות", שלשיטתם יהפכו לצורת האסדרה הדומיננטית בעתיד, עשויים גם להצטייר כסיוט דיסטופי. המעבר מ"כלליות החוק" ל"חקיקה אלגוריתמית-פרסונלית" מדאיג אחרים הטוענים שאין מדובר באבולוציה של צורות משפטיות אלא מעבר משלטון החוק לשלטון של טכנולוגיה ניהולית.¹⁰⁸

בעוד שמודלים סטטיסטיים המוטמעים בקוד עשויים כשלעצמם לחסות תחת הגדרת ה-OECD לבינה מלאכותית, המקבלת "רמות שונות של אוטונומיה", אפשר שהיסוד האדפטיבי הוא שעשוי להבחין בין חקיקה מונחית נתונים לבין חקיקה מבוססת נתונים גרידא.¹⁰⁹ מנגד, היסוד האדפטיבי הוא זה שמטשטש בין הכלל המשפטי לאכיפה שלו, מצמצם את מרחב הפעילות האנושית (human agency) והמוסדית, ובכך תורם לשחיקה בשלטון החוק שאנדיקוט ויונג מתריעים עליה,¹¹⁰ ולהחלפתו של הסובייקט המשפטי בסובייקט אלגוריתמי שאינו אלא צל-תחזית של מושאו הממשי.¹¹¹ המודל שבנוסחת המס אכן מהווה דוגמה לחקיקה מבוססת נתונים, אך אין מדובר בחקיקה מונחית נתונים. הוא אומנם מסיק מהקלט שהוא מקבל (גודל הדירה ומדדי הלמ"ס) כיצד לחולל תחזית (הערכת שווי הנכס), אך אלו גבולות האוטונומיה שלו. החקיקה לא הכילה מנגנון אוטומטי לעדכון המודל באופן מהותי נוכח שינויים מבניים באופן החישוב של מדדי הבסיס, למשל, אלא רק את המנגנון הפרוצדורלי לעדכון בחקיקה (כאשר המודל הסטטיסטי החליפי מוכנס על בסיס עבודת מטה ממשלתית).¹¹² בעוד שחקיקה מבוססת נתונים, שגם בהיותה מבוססת על עצי החלטה סבוכים נותרת חקיקה מבוססת קוד שניתן לתרגמה ללשון חקיקה אנושית של אס-אז, קשה לדמיין את מראיה של חקיקה מונחית נתונים בספר החוקים הלכה למעשה: האם תהיינה אלו שורות אינספור של קוד, או שמא הנחיות כלליות לחילול קוד (code generation) בהתבסס על הפניה לבסיס נתונים? בעיות הנוגעות לניסוח של חקיקה מבוססת נתונים, כגון שאלות של עדכון מודלים נוכח שינויים במדדים, עלולות להחריף.

סביר להניח כי מודלים מורכבים שנוצרו בלמידת מכונה קשה לתעתק ללשון אנושית ולכלול אותם בגוף החקיקה כפי שנעשה בחוק מס ריבוי נכסים. מבחינה

- Benjamin Alarie & Samuel Becher, *LexOptima: The Promise of AI-Enabled Legal Systems*, SSRN (21.11.2023) 105
- BEN-SHAHAR & PORAT, לעיל ה"ש 35. 106
- Casey & Niblett, לעיל ה"ש 35. 107
- Timothy Endicott & Karen Yeung, *The Death of Law? Computationally Personalized Norms and the Rule of Law*, 72(4) U. TORONTO L.J. 373, 395 (2022) 108
- כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 123. 109
- Endicott & Yeung, לעיל ה"ש 108. 110
- Katrina Geddes, *The Death of the Legal Subject: How Predictive Algorithms Are (Re)constructing Legal Subjectivity*, 25 VAND. J. ENT. & TECH. L. 1 26–46 (2022) 111
- ס' 118 (ב) לחוק מס ריבוי נכסים. 112

מעשית, חקיקה מונחית נתונים תצטרך להסתמך על הנחיות כלליות למידול בחקיקה ראשית, שאפשר שיישומו ייעשה באמצעות חקיקת משנה. עם זאת, יש לפתח סטנדרטים לפירוט המינימלי הנדרש מחקיקה ראשית במקרה כזה, לרבות הגדרת בסיס נתוני האימון, הגדרת פונקציית המטרה של המודל, קביעת קריטריונים לבדיקת תקפות המודל (לפני, אחרי ותוך כדי פרישתו) ואפשר גם הגדרה של טווחי תוצאות סבירים באופן קשיח, כדי להימנע מהתמודדות עם מקרי קצה.

מעקב אחר גלגוליה של הנוסחה לחישוב מס ריבוי נכסים, המתואר בחלק ג' לעיל, מראה כי כאשר אנו מצויים בעולמות של חקיקה מבוססת נתונים, למומחים חיצוניים יש גישה נוחה יותר למודל כדי לבחון אותו. ההיקף המצומצם יחסית של הפרמטרים המסבירים במודל כזה יכול לאפשר גם למחוקק הדיוט לשאול שאלות ולאתגר את הרציונל שמאחורי המודל. חקיקה מונחית נתונים, מבוססת למידת מכונה, שלה היקף פרמטרים עצום, מקשה על היכולת לבחון את המודל מבפנים ומותרת את מקבלי ההחלטות להתמודד עם אתגרי העכירות האלגוריתמית. ההליך הדליברטיבי של המחוקק מוגבל אפריורית, ולכן יש חשיבות ליצירת מנגנונים המפצים על כך: כלים פרוצדורליים שמבטיחים יכולת לתחקר את המודל לפני מעשה, לאחר מעשה ותוך כדי פעולתו.

ניטור מערכות בינה מלאכותית לאורך כל מעגל החיים שלהן הוא קריטי להבטחת תפקודן היעיל והבטיחותי. גם לאחר הפרישה (post deployment) של מודלי בינה מלאכותית שניזונים מנתונים "חיים", קלט של משתמשים, או בסיסי נתונים שאליהם מתווסף מידע באופן מקוון, חשוב לנטר אותן.¹¹³ האם חקיקה מונחית נתונים העושה שימוש, למשל, במודל המתבסס על חישוב ערכים של נכסים, מסוגלת להגיב די הצורך לשינויים אינפלציוניים חדים בערכם הריאלי? כיצד המחוקק יכול לבצע ניטור שלאחר פרישה לחקיקה מונחית נתונים באופן אפקטיבי, כך שבעיות וסיכונים יזוהו בטרם יסבו נזק, ובאופן שיאפשר את תיקונם במהרה?

מחוקק שמותח את הקפיץ של המכונה הלומדת ומאפשר לה לשכלל ולפתח הוראות חוק בהתאמה אישית, המגיבות באופן אדפטיבי לשינויים בנסיבות, נדרש למידה רבה של זהירות. ביישום של כל חקיקה ייתכנו טעויות, הטיות או מקרי קצה. אולם להבדיל מחקיקה מונחית נתונים הנעזרת במודלים סטטיסטיים פשוטים יחסית, שבהם ניתן לשער מהם מקרי הקצה מבעוד מועד, סדרי הגודל הכרוכים בשימוש בבינה מלאכותית מציבים אתגר אחר של התמודדות עם שגיאות, ומצריכים חשיבה על יצירת תשתית המאפשרת ערעור מעשי על תוצאות שגויות של המודל: תיעוד מלא של נתוני האימון, גישה לתחקור מערכי הקלט של הנתונים בזמן אמת, וכדומה.

עניין נוסף הנוגע לחקיקה המבוססת על נתוני עתק וטכניקות של למידת מכונה הוא הפגיעה הפוטנציאלית בפרטיות.¹¹⁴ אם חקיקה מונחית נתונים מוגבלת למספר פרמטרים בודדים, החזון של בן-שחר ופורת על פריסה נרחבת של סנסורי קלט המשפיעים בכל רגע

113 כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 261.

114 ראו מיכאל בירנהק "פרטיות ובינה מלאכותית" משפט חברה ותרבות ח 139 (ניבה אלקין-קורן ומעיין פרל עורכות 2025).

נתון על החובות המשפטיות של מושא המידע מעלה, בנוסף לחששות שהוזכרו מעיבוד שניוני של מידע ללא הסמכה לצורך פיתוח המודל, גם איסוף בלתי פוסק של מידע, במתווה פנאופטיקוני, כדי ליישם את המודל על מושאיו הקונקרטיים.

1. חקיקה מבוססת נתונים – עקרונות מנחים

מקרה הבוחן של הנוסחה בחוק מס ריבוי דירות אפשר לנו להציץ (מבלי להיפגע) לתהליכי העיצוב של חקיקה מבוססת נתונים, ולהעלות שאלות לגבי עצם הכדאיות של חקיקה כזו, ועל האופן שבו רצוי ליישם אותה.

ראשית, יש לפיכך לבחון את התכלית הקונקרטית של החקיקה ולראות אם עצם השימוש במודל מתמטי-סטטיסטי מתאים לה. הדבר תלוי גם באופי המודל הסטטיסטי שנבחר. כך, למשל, מושאיו של חוק מס ריבוי דירות הינם כנראה בעלי אמצעים כדי להיעזר במומחים מקצועיים כדי להתמודד עם מורכבות החוק. יתר על כן, ההיבט האקספרסיבי של המס – שנע בין הבעת הסתייגות מרכישת דירות להשקעה לבין יצירת תמריצים למשקיעים בנדל"ן לצמצם את מלאי הנכסים שברשותם – אינו מושפע מהותית מהמנגנון שבו מוטל המס.

מדרג כזה של תכליות יכול שיהיה לפי המודל של חוק הבינה המלאכותית האירופי,¹¹⁵ המגדיר רמות סיכון שונות הנשקפות מטכנולוגיות בינה מלאכותית, ובהתאמה את הדינים החלים עליהם. כך גם ניתן לחשוב על רציות השימוש במודלים סטטיסטיים בהתאם לתוחלת הסיכון שלהם בתחומי משפט שונים – אין דין הטיות אלגוריתמיות בחקיקה פלילית כדינו של מודל קשה לפיענוח בחקיקת מס.

יש להביא בחשבון את הגירעון ההסברתי של ביטויים מתמטיים סבוכים בחקיקה, ואת הקושי להצדיק אותם להדיוטות באמצעות השכל הישר. טילרס, למשל, מציע לעשות שימוש בכלים מתמטיים פורמליים באופן שהם לא יגרעו מהכוח ההסברתי של החוק אלא יתרמו לו.¹¹⁶ משכך, המחוקק נדרש לשקול אם המחירים של שימוש בנוסחה מורכבת ובלתי נהירה עדיפים על הסתייעות במנגנון פשוט יותר (כגון שיעור מס אחיד) או על שיקול דעת אנושי (כגון הערכה שמאית, במקרה של חוק מס ריבוי דירות). שיקול הדעת האנושי יכול שיהיה מגולם בעקיפין – ככל שניתן לסוכנים האנושיים מושאי החוק מרחב תמרון פרשני רחב ידיים דיו.

הליך חקיקה, לרבות חקיקה מבוססת נתונים, כולל יסוד של **דליברציה** אנושית, התדיינות בין שחקנים אנושיים רציונליים המנסים לשכנע זה את זה באשר לנכונות עמדתם, ולהגיע תוך דיון להסדר מוסכם ואופטימלי, שהרי המחוקק אינו מאציל את סמכותו למודלים של למידת מכונה או למודלי שפה מתקדמים (LLM). הילדברנט ממליצה לכלול בו יסוד אגוניסטי: קבלת החלטות שנעשית על בסיס לעומתי, באופן שיאפשר להגיע לתוצאות מושכלות.¹¹⁷ בהליך דליברטיבי שכזה נדרש שיקחו חלק

115 כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 120–134, 138.

116 Tillers, לעיל ה"ש 67, בעמ' 172.

117 Hildebrandt, לעיל ה"ש 57, בעמ' 7.

מומחים משפטיים, מדעני נתונים והמחוקקים (בייצגם את מושאי החקיקה מבוססת הנתונים).

מבחינה טכנית, יש לתת את הדעת על כך שהיסוד הדליברטיבי בהליכי חקיקה מבוססת נתונים מחייב גם אינטראקציה בין המחוקק ובין המודל: כאשר המחוקק האנושי בוחר להחריג מקרים מסוימים מהמודל (למשל, לפטור סוגי דירות מסוימים מהמס), יש לוודא כי הם מוחרגים גם מהמודל, או שהכללתם צריכה להיות מנומקת. לאחר שהמודל הותאם מחדש נוכח השינויים בבסיס הנתונים, הוא נדרש להיבחן מחדש באמצעות המחוקק באותו הליך דליברטיבי.

יש לוודא שההליך הדליברטיבי הוא מיועד. בשלב ראשון, על המחוקק עצמו להבין את מהות המודל הסטטיסטי, להכיר את מקורות הנתונים שלפיהם הוא נבנה, ולדעת להכווין את מדעני הנתונים לתוצאה הנורמטיבית הרצויה בעיניו. לשם כך, יש להבטיח רמה נאותה של שקיפות והסברות – עקרונות בסיס המצויים במרבית מסמכי האתיקה של בינה מלאכותית.¹¹⁸ ייתכן שמקום שבו עכירות אלגוריתמית אינה מאפשרת רמה הסברתית נאותה, ניתן לשקול במקרים המתאימים בחינה של איתנות ההליך המדעי, כלומר שייצור המודל נעשה באופן מדעי מוכח (תוך בחינת בסיס הנתונים שהמודל מסתמך עליו, אופן עיבודו, ויישומו בחקיקה).¹¹⁹

אולם ההליך הדליברטיבי אינו מוגבל לשיג ושיח פנימי של המחוקק עם מומחים טכניים של הרשות, אלא עליו להיות פתוח לציבור. אותם עקרונות של שקיפות והסברות כלפי המחוקק צריכים לחול גם כלפי הציבור. אומנם, כאשר המודל פשוט ניתן לנסות לשחזר אותו בתנאי מעבדה ביתיים, כפי שעשו בן-שחר וגולן עם המודל הלינארי הראשוני של נוסחת מס דירה שלישית.¹²⁰ אולם, ככל שהמודל מורכב יותר, והוא מסתמך על בסיסי נתונים עשירים יותר, ורמת העכירות האלגוריתמית שלו גדלה, חשיבות התייעוד של המודל, פיתוחו ובסיסי הנתונים שהוא נעזר בהם עולה, כדי שיתאפשר לציבור ולארגוני החברה האזרחית לקחת חלק מיועד ומשמעותי בהליך החקיקה.

בדברי חקיקה מבוססי נתונים, כמו גם בדברי חקיקה מונחי נתונים, חשוב לכלול סעיף מטרה המנוסח באורח שאינו מתמטי-פורמלי, שלאורו ייבדק המודל ותפורשנה תוצאותיו. רצוי לכלול בדברי ההסבר אמות מידה כמותיות לבדיקת פעולת החוק לאחר מעשה, או לבחינת הבסיס הכמותי שהצדיק את חיקוקו מלכתחילה.

מעריך נוסף של המלצות כרוך במיתון סיכונים, לרבות סיכונים אלגוריתמיים, בחקיקה מבוססת נתונים. כאמור לעיל, חסרונה של חקיקה מבוססת נתונים הוא הישענותה על הליך פרלמנטרי, וזה עשוי להתארך ולהתעכב בהתאם לנסיבות פוליטיות, בעוד שהמודל הסטטיסטי עשוי לאבד מתוקפו. על חקיקה מבוססת נתונים

118 כהנא ושוורץ אלטשולר, לעיל ה"ש 8, בעמ' 101–104; משרד החדשנות, המדע והטכנולוגיה עקרונות מדיניות, רגולציה ואתיקה בתחום הבינה המלאכותית, 106–107 (2023).

119 פרל ושוורץ אלטשולר, לעיל ה"ש 42, בעמ' 38, 42–44.

120 ראו בן-שחר וגולן, לעיל ה"ש 75.

לכלול בתוכה את **מנגנוני העדכון** של המודל, באופן הכולל קריטריונים למועד העדכון ולדרך שבה יבוצע. השארת פתח להחלפת המודל כליל באמצעות צו, כפי שנעשה בחוק מס ריבוי דירות, מבלי להבטיח דיון פרלמנטרי ראוי (שאפשר שמקומו לא רק בוועדת הכנסת הרלוונטית אלא במליאה), עלולה לאפשר למחוקק המשנה להציג מודלים חדשים שאינם תואמים את כוונת המחוקק. מנגנון חלופי יכול להיות הגבלת חקיקה מבוססת נתונים בהוראת שעה, שתבטיח דיון מחדש במודל עם פקיעת תוקפה.¹²¹

חקיקה מבוססת נתונים נדרשת גם לתת דעתה על **הנתונים**. אם הנתונים שעליהם התבסס דבר החקיקה מגיעים ממקור חיצוני (למשל הלמ"ס) ואינם נאספים או מחושבים לצורך המודל הסטטיסטי של אותו דבר חקיקה, יש לכלול הוראות המבטיחות **קישוריות** פנים-חקיקתית: להבטיח שהחשוב של אותם פרמטרים יישאר קבוע, שהאיסוף שלהם יתמשך כל עוד דבר החקיקה מבוסס הנתונים עומד בתוקפו, וכי שינויים בטכניקות מדידה של מדדים סינתטיים לא ייעשו בטרם יעודכן המודל. יש להבטיח ממשל נתונים (data governance)¹²² נאות, ככל שיש בהם צורך – האם הנתונים זמינים? האם הם מגובים? האם ניתן לשחזר את המודל באמצעותם? האם ישנו תיעוד ראוי ביחס לאיסופם? אילו פרוצדורות של ניקוי המידע נערכו?

על חקיקה מבוססת נתונים להתייחס למקרי הקצה של המודל. על פי רוב, התוצאה של המודל תהיה רציפה, כבדוגמת חוק מס ריבוי דירות: ערך הדירה. הגבלת המס בתקרה קבועה מנעה מקרי קצה שבהם תחזיות יתר של המודל יכלו להביא לסכומי מס גבוהים. גם בהקשר זה יש לתת את הדעת על היבטים של אינטראקציה במידול, שעליו להיעשות תוך הכרה בתקרות כאלו.

יש לבחון כיצד להכניס לחקיקה מבוססת נתונים (ומקל וחומר – לחקיקה מונחית נתונים) **מנגנוני השגה** על תוצאותיהן. מודל סטטיסטי מציע לנו אמירה על המציאות. גם חזקות משפטיות מציעות לנו אמירה כזו, והן לעיתים סמוכות ונשענות על הערכה הסתברותית של המציאות.¹²³ כפי שהדין מאפשר להפריך חזקות מסוימות כמכשיר של העברת נטלי הוכחה, ניתן לראות במודל סטטיסטי המוטמע בחקיקה מבוססת נתונים מעין חזקה – ועל המחוקק לבחון מתי לאפשר להפריכה ומתי לעשותה חלוטה. המנגנון של חוק מס ריבוי דירות הציג חזקה חלוטה בנוגע לשווי המוערך של הדירה.

121 איתי בר-סימן-טוב וגאיה הררי-הייט "שעתן היפה של הוראות-השעה? עלייתה של החקיקה הזמנית בישראל ועקרונות לטיובה" **עיוני משפט** מא 539, 564–566, 589–591 (2019). בר-סימן-טוב והררי-הייט ממליצים על שימוש בחקיקה זמנית ככלי לחקיקה אקספירימנטלית, בתחומים שיש בהם אי-ודאות או חוסר במידע, או ככלי גמיש להתמודדות עם תחומים דינאמיים. חקיקה מבוססת נתונים אינה אקספירימנטלית בהכרח, אך הישענותה על נתונים שעשויים להיות דינמיים מצריכה תיקוף עיתי שלה.

122 ממשל נתונים הוא מסגרת לניהול נתונים כנכס אסטרטגי, באופן המגדיר סמכויות ואחריות ארגוניות בנוגע לנתונים. ממשל נתונים מגדיר מדיניות, סטנדרטים ונהלים ומנטר ציות בנוגע לנתונים.

123 ראו למשל **עניין מרום**, לעיל ה"ש 43, בעמ' 832–833; Charles T. McCormick, *What Shall the Trial Judge Tell the Jury about Presumptions*, 13 WASH. L. REV. & ST. B. J. 185, 190 (1938).

מנגד, ניתן לראות במנגנוני ההשגה על איכוני השב"כ בתקופת הקורונה כלי המאפשר הפרכת חזקה בנוגע להיותם של המאוכזבים בקרבת נשאי קורונה (ואכן, שיעורים גבוהים של ערעורים על החלטות איכון אוטומטיות התקבלו לאחר בדיקה אנושית).¹²⁴ מנגנוני השגה על החלטות אוטומטיות שומרות על מעורבות אנושית בתהליך (Human in the loop), ומקטינות את הסיכוי להחלטה שרירותית.

סיכום

להבדיל מחקיקה מבוססת ראיות, חקיקה מבוססת נתונים אינה תופעה שכיחה עדיין, ולא כל שכן חקיקה מונחית נתונים. חוק מס ריבוי נכסים, שבו הוצג מודל סטטיסטי בחקיקה ראשית, שקצב העדכון שלה איטי ובלתי סדור ושנדרשת להליך חקיקה מיוחד ושקול (שלא כבעניין קוונטינזציה), מאפשר לנו לבחון כעומדים מהצד את תהליך הפיתוח של המודל הסטטיסטי שנועד לנבא את ערכה של דירה בישראל (אך התכחש לכך רשמית). ראינו כיצד שקיפות מאפשרת לציבור לבחון את המודל באופן מקצועי, וכיצד היעדר דיון מעמיק במודל מונע מהמחוקק להבינו עד תום, באופן המאפשר למחוקק המשנה להחליף את המודל באמצעות צו ללא כל התנגדות פרלמנטרית מהותית.

לצד יתרונותיה של חקיקה מבוססת נתונים, יש להיות מודעים לסיכונים שהיא טומנת בחובה, ולצורך בהליך חקיקה מיוחד, זהיר ושקוף ככל הניתן. בטרם משסים בציבור ובנציגיו "שיגעונות" בדמות נוסחה לוגריתמית מורכבת בחקיקה, יש להבטיח שניתן להסבירה למושאי החוק ולמנסחיו, ולוודא שיש היגיון ב"עשר בחזקת שיגעון". מקל וחומר, כאשר עסקינן בחקיקה מונחית נתונים הנסמכת על מודלים מורכבים של למידת מכונה, שעשויים להתעדכן בזמן אמת – שאלות הנוגעות לשקיפות, הסברתיות ובקרה על יישום המודל מקבלות משנה תוקף. הכנסת מערכות בינה מלאכותית לחקיקה מאלצת אותנו להתמודד עם מודלים מורכבים שלשון אנושית אינה יכולה לתאר בעל פה, ולא כל שכן להבינם עד תום. במקום לפענח את הרציונל של מודל שנוצר בלמידת מכונה, מחוקק של חקיקה מונחית נתונים נדרש ראשית להגדיר היטב את פונקציית המטרה של החוק, ובה בעת להבטיח את קיומם של מנגנונים שיבטיחו את תקינות המודל והנתונים שעליהם הוא מסתמך טרם פרישתו, שיאפשרו תחקור בדיעבד של המודל, שינטרו אותו תוך כדי פעולתו, ושיאפשרו השגה על החלטותיו.

124 עמיר כהנא "עידן הזוחלים: ההסמכה ההדרגתית של השב"כ לאיתור מגעים של חולי קורונה" **משפטים על אתר** כ 22 (צפוי להתפרסם ב-2024).