



ח' בטבת, התשפ"ג

1 בינואר, 2023

לכבוד

משרד החדשנות, המדע והטכנולוגיה

משרד המשפטים – ייעוץ וחקיקה

בדוא"ל AI@most.gov.il

התייחסות מכון תכלית ומרכז פדרמן לעקרונות מדיניות, רגולציה ואתיקה בבניה מלאכותית

שלום רב,

אנו מתכבדים להגיש את עמדת תכלית-המכון למדיניות ישראלית ומרכז פדרמן לחקר הסייבר באוניברסיטה העברית לקול הקורא אשר הופץ ביום 30.10.2022 ביחס למסמך עקרונות מדיניות רגולציה ואתיקה בתחום הבינה המלאכותית ("מסמך העקרונות"). אנו מברכים אתכם על העבודה המקיפה והמקצועית המשתקפת מבעד למסמך. התייחסותנו תתמקד בשני נושאים עיקריים אשר ראויים בעינינו להמשך העמקה ובחינה במסגרת קידום המתווה הרגולטורי המופיע במסמך העקרונות, כדלקמן.

⬅ **הקשיים העולים מהבחירה שלא לאמץ רגולציה רוחבית צריכים לקבל מענה בגיבוש מהודק יותר של מדיניות רגולציה הכוללת מאסדר לאומי ומתודולוגיית חשיבה אחידה שתחול כלל הרגולטורים המגזריים**

הבחירה העולה ממסמך העקרונות שלא לאמץ גישה של רגולציה רוחבית (בדומה לחקיקה המקודמת באירופה) מהווה מדיניות המבכרת אינטרסים של חדשנות טכנולוגית ויזמות בתחום הבינה המלאכותית על ערכים אחרים. זוהי בחירה רציונאלית ומובנת, אולם אין לחדד כי היא מעצימה סיכונים חברתיים שונים הכרוכים מהטמעה לא זהירה של בינה מלאכותית ועלולה להביא למדיניות עמומה ולא קוהרנטית בתחומים שונים ומשיקים זה לזה. מסמך העקרונות מבקש לתת מענה לבעיה הזו על ידי גיבוש של "מדיניות רגולטורית", אולם מדיניות זו, מלבד להציע תפריט של כלים (רגולציה רכה, הרמוניה עם הנעשה בעולם וכו') שכל רגולטור מגזרי יוכל לבחור ולהשתמש בהם לפי שיקול דעתו, לא מספקת הנחיה מהודקת ולא מבטיחה קידום של יעדי מדיניות אחידים ברמה הלאומית. נוסף, שמתן מקום גדול יותר לתפיסת האדם במרכז, מובילה, הלכה למעשה גם לקידום חדשנות ומניעת חסמים בשווקים בינלאומיים. לכן, חשוב להבהיר שבאזון זה אין זכויות האדם מעכבות את החדשנות, אלא הולך עמה יד ביד. החשיבות של הסדר נוקשה

יותר, מגדילה את הציות לזכויות אדם במיוחד אצל תאגידים קטנים יותר – וחברות הזנק – אשר לא קיבלו על עצמן מחויבות וולנטרית לזכויות אדם, למשל בדומה למסגרת של עקרונות האו"ם לזכויות אדם.¹

אנו סבורים כי במטרה לסגור את הפער הזה, יש לאמץ שני צעדים משלימים.

1. **מיסוד של מאסדר לאומי** אשר תפקידו יהיה להנחות ולייצר הרמוניה בין הרגולטורים המגזריים (בדומה לתפקידו של מערך הסייבר הלאומי באסדרת הגנת הסייבר).

2. **אימוץ של מתודולוגיה אחידה לבחינת המענה המשפטי והרגולטורי לאתגרי הבינה המלאכותית** במגזרים השונים.

קידום האחריות של מפתחי ומפעילי מערכות בינה מלאכותית המשפיעה על זכויות יסוד



הסתפקות בכלי רגולציה רכה ומחויבות אתית ביחס למערכות בינה מלאכותית אשר יש להן השפעה יסודית על זכויות אדם עלולה לייצר גרעון דמוקרטי הנובע מהעדר אחריות משפטית מספקת מצד מפעילי בינה מלאכותית כלפי משתמשים. בכך, תיפגע מטרה מרכזית אותה מסמך הרגולציה מציב – תפיסת האדם במרכז. בהתאם, אנו מציעים להסדיר את סוגיית האחריות ביחס ליישומי בינה מלאכותית בעלי השפעה על זכויות יסוד באופן מפורש, וממחישים את עמדתנו תוך פניה אל מקרה הבוחן של סינון תכנים באמצעות בינה מלאכותית ברשתות חברתיות. כפי שנראה, דוגמה זו מלמדת כי לעיתים אימוץ גישה דה-רגולטורית בשלבים הראשונים עלולה להטמיע בשטח מציאות בעייתית שרגולציה מאוחרת תתקשה לתת לה מענה.

א. הידוק המדיניות הרגולטורית: מיסוד מאסדר לאומי ואימוץ מתודולוגיה אחידה לבחינת המענה הרגולטורי והמשפטי

1. החלטת מדיניות מרכזית העולה ממסמך העקרונות היא הבחירה שלא לכלול בשלב הזה הצעה לקידום חקיקה כוללת שתסדיר בצורה רוחבית את הפיתוח והשימוש בכלי בינה מלאכותית. בהחלטה זו טמונה ההכרעה לבכר ערכים של עידוד חדשנות ויזמות, אשר הם בעלי חשיבות עליונה לאינטרס הלאומי לקדם את המגזר הטכנולוגי, תוך נטילה מודעת של סיכונים אחרים: החשש ליצירת אי-בהירות ואי-אחידות, והאפשרות הממשית ליצירה של אזורי חפיפה רגולטורית ולצידם "חורים שחורים" שלא זוכים לכיסוי על אף הסיכונים העולים מהפעילות והפיתוח של בינה מלאכותית.
2. המסמך מעודד פרקטיקה של יישום עצמי והפנמה של העקרונות, על ידי מפעילי הבינה המלאכותית והמשתמשים, תוך חידוד הצורך בהפנמת הידע לאורך זמן, ולא קידום של רגולציה מקיפה ומחייבת כלפי מפעילי

הבינה המלאכותית. עוד מציע המסמך כי הרגולציה שתקודם תכלול התאמה לרגולציה בינלאומית על מנת למנוע הצבה של חסמים רגולטוריים ייחודיים לישראל.²

3. המענה המוצע במסמך העקרונות לסיכונים האמורים כולל אימוץ של מדיניות רגולציה, אשר מיועדת לשמש כמסגרת-על לקידום מדיניות כלל משקית אחידה, והצעה למיסוד "מוקד ידע ותיאום ממשלתי להסדרת בינה מלאכותית". **אנו סבורים כי אין בפתרונות אלה מענה מספק לסיכונים האמורים.** מדיניות הרגולציה המתוארת בפרק 6.1 היא כללית מדי ואינה מייצרת מסגרת למדיניות אחידה. מדובר למעשה בתפריט אקלקטי של הצעות וכלים (ביניהם, מתן חשיבות להרמוניה עם הרגולציה בעולם, אימוץ גישה של ניהול סיכונים, קריאה להטמעת רגולציה רכה ועידוד נסיינות ושיתוף ציבור) אותם רגולטורים שונים צפויים לקרוא, לפרש ולתעדף בצורות שונות.

4. עובדה זו בולטת במיוחד ביחס למערכות בינה מלאכותית כלליות (general-purpose AI systems), שלא צפויות לקבל מענה אחיד ולעתים מענה כלל על ידי רגולטורים מגזריים, למרות שמדובר בתחום הזוכה לפיתוח מואץ ועשויות להיות לו השלכות משמעותיות על בטיחות וזכויות.³ המענה המוצע בדמות מוקד ידע ותיאום ממשלתי (פרק 6.3) צפוי להיות, במבנה הנוכחי, בעל השפעה מוגבלת בשל העובדה כי לא מתוארים במסמך העקרונות סמכויות חוקיות מחייבות אשר יינתנו לגוף הזה, לכשיוקם. קשה לתאר כי גוף ממשלתי שתפקידו יהיה לייעץ ולהנגיש מידע יתפקד כגוף בעל שיניים שיוכל להנהיג מדיניות אחידה וליצור האחדה של מדיניות רגולטורית בין רגולטורים מגזריים אשר לכל אחד מהם זווית ראייה צרה הנוגעת למגזר שלו. מציאות של תת-רגולציה ביחס ליישומים מסויימים ורגולציה חופפת ואולי אף סותרת ועודפת ביישומים אחרים היא אפשרות סבירה במצב עניינים זה.

5. מצב דברים זה מעורר קושי לא רק ביחס לפיתוח של רגולציה מגזרית שאינה מסונכרנת למדיניות לאומית כוללת, אלא גם ביחס להתפתחות הדין על ענפיו השונים. התפתחות תעשיית הבינה המלאכותית והפריסה הרחבה של השימוש בה צפויים להציף אתגרים לאופן פרשנות והתאמת הדין הקיים בענפים שונים – נזיקין, הגנת הצרכן, תחרות, איסור הפליה ועוד. בהיעדר מדיניות משפטית אחידה המכוונת על ידי גורם הבוחן את השאלות במישור הכלל-משקי, קיימת סכנה שבמגזרים שונים יאומצו פרשנויות סותרות ותוטמע מדיניות משפטית לא-אחידה. הפתרונות המוצעים במסמך העקרונות לא מספקים מענה לבעיה זו.

6. על רקע זה, אנו ממליצים לשקול שני פתרונות משלימים כחלק מפיתוח המדיניות הרגולטורית. בנוסף, נטען שהצורך במדיניות רגולטוריות נוקשה יותר, נובע, הן מעקרונות החדשנות והן מעקרונות האדם במרכז, בהינתן העיקרון לפיו יש למנוע חסימה מכניסה לשווקים אחרים. נטען, שרגולציה רכה שאיננה עולה בקנה אחד עם המדיניות האירופאית, עלולה לייצר בעתיד חסמים לחברות ישראליות להיכנס לשוק האירופאי. על מנת להוכיח את הסכנה הזו, נדון במקרה הבוחן של תאימות דיני הפרטיות הישראליים ביחס לדיני הפרטיות באיחוד האירופאי.

7. ההסדר האירופאי והסכנות שבאי-תאימות

- 6.1 בטייטה המוצעת מטעם האיחוד האירופי, הועדפה גישה אסדרה לפיה יש צורך בחקיקה אירופאית אופקית נוקשה ביחס לבינה מלאכותית עם סיכון גבוה יותר, על פני של רגולציה סקטוריאלית, תוך יצירת קוד התנהגות מחייב פחות כלפי בינה מלאכותית עם סיכון נמוך.⁴ גישה זו מבקשת לקדם באופן מיטבי הגנה על זכויות אדם (כבוד האדם, פרטיות, שוויון ומניעת אפליה, חופש ביטוי, התארגנות, הזכות להליך הוגן ועקרונות כלליים של משפט מנהלי),⁵ להגביר את הוודאות המשפטית, לשפר את אפקטיביות האכיפה, וליצור סביבה המאפשרת פיתוח של מערכת בינה מלאכותית בטוחה למשתמשים.⁶
- 6.2 בהתאם, החקיקה האירופאית המוצעת מטילה חובות על מערכות בינה מלאכותית שסווגו בתור בעלות סיכון גבוה, לרבות יצרנים, יבואנים, מפיצים, משתמשים וצדדים שלישיים.⁷ בין החובות שהוגדרו, ניתן לציין פיקוח אנושי, בקרת איכות, שמירת מידע, הסברתיות ועוד. כך למשל, מוצעת חובת דיווח על תאונות ותפקוד לקוי של המערכת⁸ וחידוד הצורך באכיפה של חוקי האיחוד על מערכת בינה מלאכותית.⁹
- 6.3 גם המועצה האירופאית מקדמת במישור הבין-מדינתי הבינלאומי ניסיון לגיבוש מסמך משפטי מחייב להסדרת בינה מלאכותית, המתמקד בהגנה על זכויות אדם, דמוקרטיה ושלטון החוק,¹⁰ תוך מתן שימת לב מיוחדת לאוכלוסיות רגישות לפגיעה דוגמת קטינים.¹¹ בעוד ההתחשבות באוכלוסיות רגישות קיבלה ביטוי בהנחיות רגולציה ממשלתיות אחרות,¹² הדבר לא נעשה במסמך דנן.
- 6.4 נבקש ללמוד ממקרה הבוחן של תאימות דיני הפרטיות הישראלים עם האירופאיים. דרישת התאימות מגיעה מן האיחוד האירופי והכרה בתאימות הדינים מקלה מבחינה רגולטורית על קבלת מידע אודות נושאי מידע אירופאיים. מתן סטטוס זה הוא קריטי לחברות ישראליות רבות, ויש לו חשיבות כלכלית רבה. לכן, אנו מציעים לכתחילה, לקדם רגולציה ישראלית שעולה בקנה אחד עם המדיניות האירופאית, כדי למנוע מצב בעתיד בו לא תוכר ההתאמה. במצב כזה, עלולה להיווצר בעיה עבור חברות ישראליות לפעול בשוק האירופי בתחום הבינה המלאכותית, לצד פגיעה אפשרית בזכויות אדם ובפרט בזכות לפרטיות, ובכך ייפגע ערך האדם אותו מבקש מסמך הרגולציה לקדם. ערך נוסף אותו מבקש לקדם מסמך הרגולציה דנן אך עלול להיפגע הוא ערך החדשנות, במידה וחברות ישראליות יוגבלו בפעילותן בשוק האירופי. הפסיקה של מוסדות השיפוט של האיחוד האירופי, לאורך השנים, דורשת להרחיב את ההגנה על זכויות האדם כחלק מקבלת הכרה בתאימות – מדיניות שניתן לקדמה כבר במסגרת מסמך זה.¹³
- 6.5 מהאמור עולה, כי הטלת חובות מעבר לרגולציה רכה הכרחית על מנת להגדיל את ההגנה על זכויות אדם, את האמון הציבורי במערכות הבינה המלאכותית ולהגברת המשילות. ככל שלא יוטלו כאלו ברגולציה הישראלית, ייווצרו פערים נורמטיביים אשר עלולים לפגוע במרחב הפעולה של חברות ישראליות בשוק האירופי, ובכך לסכל את קידום ערך החדשנות כמו גם ערך האדם במרכז.

8. מיסוד מאסדר לאומי

8.1. חלף הכוונה למסד מוקד יידע בעל סמכויות מוגבלות ובלתי מחייבות, אנו ממליצים על מיסוד מאסדר לאומי אשר ימלא שתי פונקציות מרכזיות. האחת, **המאסדר הלאומי ישמש כ-“רגולטור של הרגולטורים”**. תפקידו יהיה להבטיח הרמוניזציה וחתירה משותפת של כלל המגזרים במשך להגשמת יעדי המדיניות הלאומית בתחום הבינה המלאכותית. במסגרת זו, המאסדר יהיה בעל הסמכות המחייבת לקבוע מדיניות רגולטורית אחידה ביחס ליישומי בינה מלאכותית חוצי ענפים, לנפק הנחיות והערות פרשניות לרגולטורים המגזריים ויבטיח תיאום לשמירה על אחידות של הרגולציה הענפית. השנייה, **מאסדר העל יפתח את המדיניות הרגולטורית וייעץ לממשלה** על הצעדים הנדרשים להמשך פיתוחה וקידומה. בין היתר, מאסדר העל יקבל אחריות על הובלת המדיניות בתחומים המוסדרים כעת בפרקים 6.4-6.7 במסמך העקרונות.

8.2. מיסוד אורגן של מאסדר על מהווה מודל ביניים רצוי בין החלופה של רגולציה רוחבית לבין החלופה של רגולציה מגזרית ומאפשר להנות מחלק ניכר מהיתרונות של כל מודל. מאסדר העל פועל מעמדה מוסדית המאפשרת לו לגבש מדיניות לאומית וכלל-משקית אחידה, להוות גורם לאימוץ וסינון מקצועי של תקנים וסטנדרטים בינלאומיים ולהבטיח רמה נאותה של גמישות אצל הרגולטורים המגזריים באופן שאינו חותר תחת המדיניות הלאומית.

8.3. מודל דומה של מאסדר לאומי הפועל כגורם המסדיר את הרגולציה המגזרית אומץ בישראל במסגרת תזכיר חוק הגנת הסייבר ומערך הסייבר הלאומי (התשע"ח – 2018). אמנם קיימים הבדלים מהותיים בין מישור הגנת הסייבר הלאומית לבין פיתוח ושימוש בבינה מלאכותית, אך בשני המקרים ענייננו בנושא בעל השלכות רחב לאומיות אשר לצד קיומם של אתגרים מגזריים ורצון לשמר על גמישות ושונות בין מגזרים שונים נדרשת בו מדיניות לאומית כוללת להבטחת אחידות וקוהרנטיות.

8.4. מוצע כי המאסדר הלאומי יהיה בעל סמכויות להנחות את הרשויות המאסדרות במימוש מדיניות הרגולציה לאומית ובהסדרה של סוגיות משפטיות פרשניות להתאמת הדין הקיים. בנוסף, מי שרואה עצמו נפגע מהחלטות של רגולטורים מגזריים יוכל לפנות למאסדר הלאומי לבחינה מחדשת של החלטות הנוגעות למדיניות הרגולטורית, ולהחלטותיו יהיה מעמד מחייב.

9. הטמעת מתודולוגיה אחידה להתמודדות עם אתגרים רגולטוריים ומשפטיים

9.1. להרחבת הפיתוח והשימוש של בינה מלאכותית צפויה השפעה משבשת (disruptive) על שווקים וחברות, וכנגזרת לכך, על הענפים המשפטיים המסדירים את פעילותם. חלק ניכר מאותן השפעות יהיו חוצות מגזרים, וראוי כי ייבחנו ויקבלו מענה באופן הרמוני. כאמור, איננו סבורים כי הדבר מחייב רגולציה רוחבית (אשר לה עלויות משמעותיות על היכולת של התעשייה המקומית למצות את הפוטנציאל שלה). אולם

מסמך העקרונות לא מספק כרגע לקהילה הרגולטורית והמשפטית כלים מתודולוגיים להתמודדות אחידה, קוהרנטית ורציונאלית עם האתגרים הצפויים.

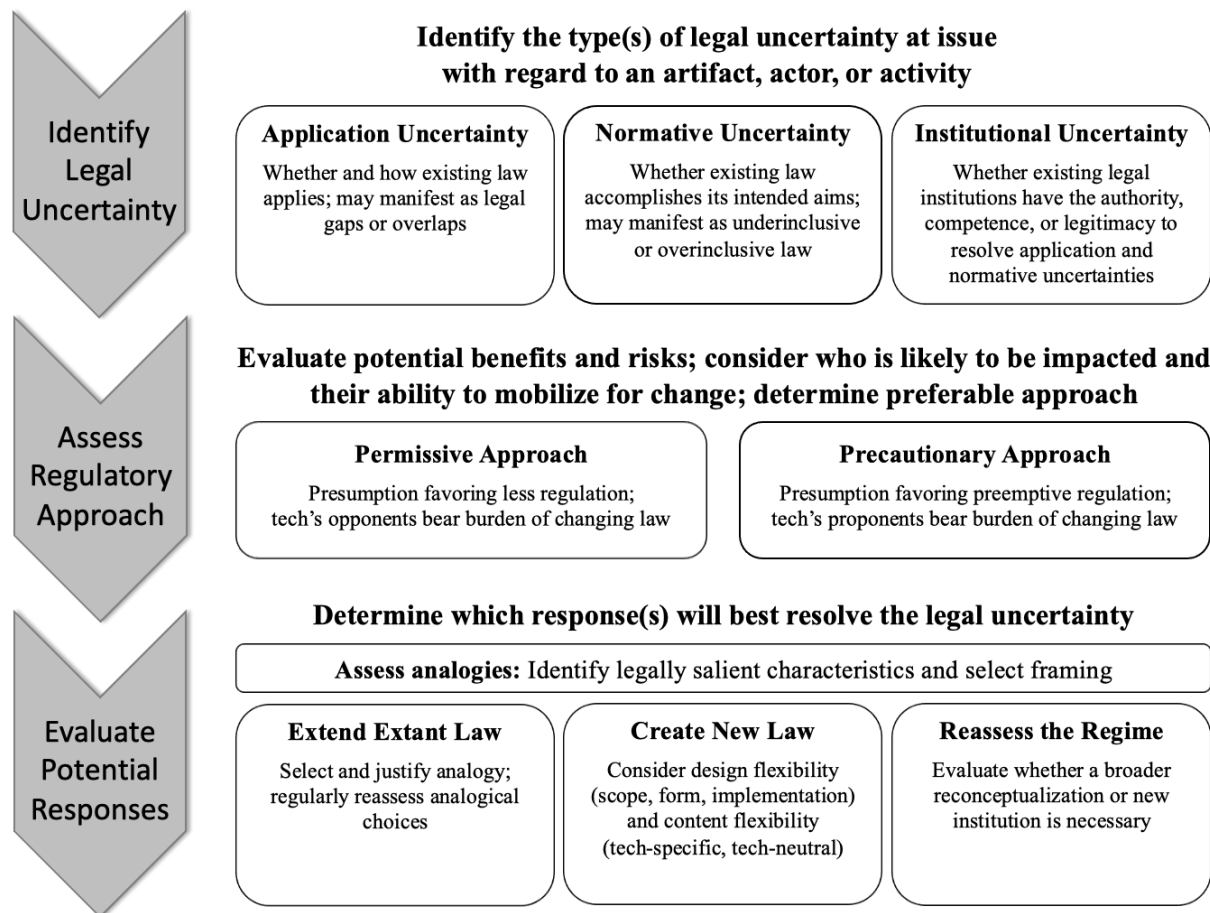
9.2. לדוגמה, מבנים תעסוקתיים יעברו שינוי דרמטי עם הופעתן של טכנולוגיות מתקדמות, אשר בתורן ייצרו סיכוני בטיחות חדשים. התפתחויות אלה ישפיעו על המגזר הפיננסי, הרפואי, האקדמי-חינוכי ואחרים, ולמרות שההשפעות הקונקרטיות ישתנו ממגזר למגזר, האתגרים הרגולטוריים והמשפטיים בתחומי דיני העבודה, דיני הבטיחות והגהות, דיני הנזיקין ואחרים, הם חוצי מגזרים ונדרשים לקבל מענה קוהרנטי.

9.3. על מנת לייצר שפה משותפת ועמדות קוהרנטיות בין רגולטורים שונים ויועצים משפטיים בין המשרדים השונים ובתוך משרד המשפטים, חיוני להטמיע מתודולוגיית חשיבה אחידה להתמודדות עם הבעיות המשותפות. בסיוע מתודולוגיה כזו, יהיה ניתן להגשים מדיניות רגולטורית ומשפטית הנשענת על רעיון מארגן אחיד ועל קוהרנטיות במקום איים מבודדים של מענים רגולטוריים ומשפטיים, שבאופן טבעי יהיו לא אחידים ולעיתים אף סותרים.

9.4. אנו סבורים כי נושא זה ראוי שיהיה באחריות מחלקת ייעוץ וחקיקה במשרד המשפטים אשר ינחה את יתר היחידות הממשלתיות. במסמך זה אנו מבקשים לשים זרקור על מתודולוגיה שפורסמה על ידי החוקרים קרוטוף וארד במאמר שפורסם ב- *Harvard Journal of Law & Technology* בשנת 2021.¹⁴ המאמר מציע מסגרת אנליטית לבחינת ההשפעה המשבשת של הטכנולוגיה על ענפי משפט מושפעים ועל הגישה הרגולטורית והמשטרים הרגולטורים המסדירים אותה, אשר ניתן ליישם ולהפעיל ביחס ליישומי בינה מלאכותית. החידוש במסגרת המוצעת שהפך אותה למרכזית בשיח המחקרי טמון בכך שאינה לוקחת כמובן מאליו את ההנחה שטכנולוגיה חדשה מייצרת צורך קבוע לעדכון הדין. במקום זאת, היא מציעה להתמקד בעמימות המשפטית, נורמטיבית או מוסדית שהטכנולוגיה מייצרת ומתוכה להבין מהי הגישה המטא-רגולטורית הנכונה וכיצד הדין והמשטר הרגולטורי צריכים להגיב.

9.5. להלן נציג את עיקרי המתודולוגיה.

תרשים 1. מתודולוגיית קרוטוף-ארד להערכת ההשלכות המשפטיות של טכנולוגיות משבשבות



9.6. המתודולוגיה כוללת שלושה שלבים.

9.6.1. בשלב הראשון שחקנים רגולטוריים ומשפטיים נדרשים לסווג את אופי העמימות המשפטית

שמייצרת הטכנולוגיה מבין שלוש קטגוריות: עמימות ביישום הדין, עמימות נורמטיבית ועמימות מוסדית. לדוגמה, הטמעה של מכשור לדיאגנוזה רפואית בקופות החולים המיועדת לייעל את מהירות ואיכות הניתוח של הדמיות ונתונים של המטופל עשויה לייצר רמות שונות של עמימות (שמובילה לרגולציית יתר או חורים בכיסוי הרגולטורי) ביחס להחלה של הרגולציה על איברים ומכשירים רפואיים: עמימות ביישום הדין תשאל האם הרגולציה הקיימת חלה על מכשור מבוסס בינה מלאכותית, עמימות נורמטיבית תעסוק בשאלה האם החלה של הרגולציה הקיימת מובילה לתוצאות התואמות את תכלית הרגולציה, ועמימות מוסדית תבחן האם הגופים הרגולטוריים הרלוונטיים (משרד הבריאות, הרשות להגנת הפרטיות וכו') מחזיקים ביכולת המוסדית ובסמכות החוקית להשיב על שתי השאלות הראשונות.

9.6.2. בשלב השני, בהנחה שקיימת עמימות משפטית כלשהי שמצדיקה תגובה משפטית, נדרשים

השחקנים הרגולטוריים והמשפטיים לבחון מה ראוי שתהיה התגובה המטא-רגולטורית

הראשונית. אפשרות אחת שלרוב תהווה ברירת המחדל היא גישה מתירנית המבוססת על חזקה

שניתן להשתמש בטכנולוגיה ללא רגולציה מגבילה או על פי הרגולציה הקיימת עד שיתברר מידע

נוסף באשר להחצנות העלולות להיגרם ממנה. אולם, במקרים מסויימים בהם מתגלה כי הסיכון הידוע

הוא משמעותי, ייתכן כי כבר בשלב הזה יעדיף הרגולטור להטיל הגבלות ואף איסור כולל על השימוש

בטכנולוגיה. התגובה הראשונית תיעשה בדרך כלל במציאות של מידע חסר, אך ניתוח נורמטיבי של

אופי העמימות המשפטית יאפשר במרבית המקרים תגובה רציונאלית.¹⁵

9.6.3. בשלב השלישי מתגבש הפתרון הכולל לעמימות המשפטית והסטטוס-קוו המשפטי/רגולטורי מושב

על כנו. הפתרון יפול לרוב לאחת או יותר מבין שלוש קטגוריות אפשריות: "מתיחה" של הדין הקיים

כך שיחול על הטכנולוגיה החדשה, פיתוח של כללים משפטיים חדשים, או פיתוח של משטר

משפטי/רגולטורי חדש.

9.7. אנו סבורים כי הטמעה רוחבית של מתודולוגיית קרוטוף-ארד תאפשר התפתחות רציונאלית וזהירה של

הדין תוך מודעות לאיזונים הנדרשים בין ערכים המושפעים מטכנולוגיות מבוססות בינה מלאכותית, באופן

שיאפשר יישום קוהרנטי ואחיד בין ענפים משקיים שונים ובין שחקנים רגולטוריים ומשפטיים שונים.

ב. קידום האחריות של מפתחי ומפעילי מערכות בינה מלאכותית המשפיעה על זכויות יסוד

10. מסמך העקרונות מבקש לקדם אחריותות בתור ערך אתי מרכזי: "מפתחי בינה מלאכותית, מפעיליה או

המשתמשים בה, יגלו אחריות לתפקודה התקין, ולקיום העקרונות האתיים האחרים בפעילותם, בין היתר בשים

לב לתפיסות ניהול סיכונים מקובלות ולאפשרויות הטכנולוגיות הזמינות".¹⁶ לאורך המסמך, האחריותות

מתוארת בתור אחד מהאתגרים המשמעותיים ביותר של הבינה המלאכותית, ובהתאם מוצעות פרקטיקות

רצויות כגון הצורך בהסברות ובמעורבות אנושית.

11. המסמך כולל מספר הצעות להתמודדות עם אתגר זה בתחום הרגולציה הרכה והאתיקה, אך ניכר כי נעשה

מאמץ להימנע מרגולציית יתר אשר עלולה להוות חסם לחדשנות ויזמות טכנולוגית. אנו סבורים שכלל, זוהי

גישה ראויה. יחד עם זאת, קיימת הצדקה לחרוג ממנה בהתקיים שני תנאים.

12. הראשון, כאשר שימוש במערכת בינה מלאכותית הוא בעל השפעה מיידית ומשמעותית על זכויות אדם; השני,

כאשר קיים חשש שהימנעות מרגולציה בשלב המוקדם תוביל למציאות בה רגולציה מאוחרת תתקשה לתת

מענה להגנה על זכויות. חשוב לתת את הדעת כי לעיתים "the social consequences of a technology"

cannot be predicted early in the life of a technology. By the time undesirable consequences are discovered, however, **the technology is often so much part of the whole economic and social fabric that its control is extremely difficult**¹⁷.

13. השימוש בבינה מלאכותית משפיע באופן עמוק על החברה הדמוקרטית בה אנו חיים – בין אם ביחס לאגירת מידע על משתמשים והשפעה על דעת הקהל, בין אם מדובר על השפעה ושליטה על במות ביטוי עיקריות, וכן, כמובן, מערבותה של בינה מלאכותית בהליכי קבלת החלטות, אשר יכולים להוביל לפגיעה בזכויות יסוד. שילוב בין רגולציה רכה לבין השפעה יסודית של בינה המלאכותית על חיינו, עלולה לייצר גרעון דמוקרטי במסגרתו למרות ההשפעה הגדולה על חיי המשתמשים והאזרחים, אין מפעילי הבינה המלאכותית חבים באחריות משפטית מספקת.

על מנת להמחיש את עמדתנו, ניקח כמקרה בוחן את שאלת סינון תכנים באמצעות בינה מלאכותית ברשתות חברתיות.

14. מניעת גרעון דמוקרטי באמצעות אחריותיות - מקרה בוחן

14.1. ההשפעה של מפעילי בינה מלאכותית על החברה כוללת תחומים רחבים מאוד, וביניהם יצירת ואיגוד קבוצות חברתיות שונות, באופן שמאפשר קידום אג'נדות חברתיות.¹⁸ בכך נכללים בין היתר, הפצה של חדשות כזב (fake-news), שלאורך העולם מתחילה הסדרת הפצתם, וקידום של תאוריות שונות, כפי שניתן היה לראות בתקופת נגיף הקורונה ביחס לקבוצות של מתנגדי חיסונים. כך, בינה מלאכותית, ומקרה הבוחן של הרשתות החברתיות מייצר echo-chambers – פרופילים שמאפשרים הדווד והקצנה של עמדות, באופן שמגדיל קיטוב חברתי, ומשפיע על המארג החברתי.¹⁹ השפעות אלו של הבינה המלאכותית, מעניקות דוגמה נוספת ביחס להשפעתה על המארג הדמוקרטי של החברה המודרנית ואת הצורך בקידום שקילות שיקולים דמוקרטיים בעת אסדרת בינה מלאכותית.

14.2. מניעה של גרעון דמוקרטי תקדם מטרה מרכזית אותה מסמך הרגולציה דן מציב לעצמו – תפיסת האדם במרכז. בנוסף לשיפר מאזן ההגנה על זכויות האדם של משתמשי בינה מלאכותית, הדבר יגביר את אמון הציבור במערכות בינה מלאכותית. לשם ההשוואה, המסגרת המקודמת באיחוד האירופי לאסדרת בינה מלאכותית שוקלת שיקולים חברתיים רחבים – לעומת שיקולים סקטוריאליים כפי שמוצע במסמך דן.

14.3. הגישה המטא-רגולטורית הראשונית לפרקטיקות של סינון תכנים בפלטפורמות היתה גישה דה-רגולטורית מתירנית. בארצות הברית חל כידוע סעיף 230 לחקיקה הפדרלית האמריקאית ה- (CDA) Communications Decency Act, אשר מעניק חסינות רחבה לפלטפורמות מאחריות לתוכן שמעלים

משתמשים תוך שמירת זכותן לסנן תכנים מבלי לשאת באחריות על כך. במשך הזמן הפלטפורמות הגדולות פיתחו כלים אלגוריתמיים שעושים חלק ניכר מעבודת הניטור והסינון של תכנים פוגעניים. בד בבד, כלים אלה מעצבים בפועל את היקף זכויותיהם של מיליארדי משתמשים לביטוי, לפרטיות ולבטיחות ברשת.

14.4. בשנים האחרונות, הגישה הדה-רגולטורית מתחילה להיסדק ונזקיה צפים על פניה שטח. בארצות הברית אנו רואים ניצני חקיקה מדינתית המבקשת להגביל את כוחן כך, לאחרונה אושר חוק רשתות חברתיות חדש שאומץ במדינת טקסס,²⁰ אשר אוסר על רשתות חברתיות גדולות לצנזר, לחסום, להוריד או לצמצם חשיפה של פוסטים על בסיס עמדה פוליטית. מעבר לאיסור זה, דורש החוק לספק מידע על הסרת תכנים והשעיית חשבונות פעמיים בשנה.²¹ תיק מעניין נוסף, אשר יידון בזמן הקרוב, הוא פרשת גונזלס נ' גוגל.²² מדובר על תביעה נגד גוגל, פייסבוק וטוויטר, שאפשרו להציג בפיד של גונזלס פוסטים התומכים בדאע"ש, וכעת על בית המשפט העליון להחליט האם לפלטפורמות עומדת אותה חסינות מסורתית או שלא. כלומר, אנו נמצאים בנקודה קריטית אשר עלולה לסמן על כרסום בהיקף החסינות המוענקת לפלטפורמות חברתיות בארצות הברית.²³

14.5. גם באירופה חלות התפתחויות משמעותיות. החקיקה של ה- Digital Services Act (DSA), מטילה על הפלטפורמות שורה של חובות רגולטוריות להבטחת טיפול הולם במניעה וטיפול בנזקי המשתמשים. לדוגמה, תחת ה-DSA, הפלטפורמות מחויבות לפרסם מידע בנוגע למדיניות הסרת התכנים שלה, תוך פילוח מידע רב אודות הביטוי שהוסר. הגברת שקיפות זו מציידת את האזרחים ביותר מידע אודות האופן שבו הרשת החברתית עובדת, ומעניק מידע רב יותר לציבור אודות ההשפעה של הרשתות החברתיות במדינה. לכך מצטרף ההסדר במסגרת תקנות הפרטיות האירופאיות (GDPR), המטיל חובות על פלטפורמות תוכן דיגיטליות ובפרט לגבי השימוש שלהן בבינה מלאכותית. כך, למשל, נהוגים הסדרים רחבים ביחס לזכות להישכח, גם ביחס לסידור התוצאות בחיפושים במנועי חיפוש ומנגנוני ההשלמה האוטומטית מהווים מושא לזכות.

14.6. בנוסף, ישנה התעוררות בינלאומית לגבי הצורך בקידום אחריות של רשתות חברתיות בהקשר של השימוש לרעה בהן לביצוע פשעים בינלאומיים. כך, למשל, המעורבות של הרשת החברתית פייסבוק בפרסום והדהוד מסרים של צבא מיאנמר,²⁴ תרם לרצח העם של מיעוט הרוהינגה במדינה, כעולה מעדותו של ראש צוות החקירה העצמאי לחקר רצח העם במיאנמר.²⁵ הדבר הוביל לביקורת ציבורית רבה, וכן למהלכים משפטיים שונים נגד החברה, הכוללים תביעות ייצוגיות וכן תביעה של מדינת גמביה.²⁶

14.7. ההתעוררות לגבי חשיבות קידום אחריות של מפעילי בינה מלאכותית, ובפרט בהקשרי סינון תכנים, לא פסחה על ישראל. כך, למשל, במסגרת תביעות בתחום לשון הרע, יוחסה אחריות למנועי

חיפוש, אודות פעולת האלגוריתם ביחס להשלמה אוטומטית.²⁷ בפסק הדין בעניין סביר נ' בר נוי נקבע, בפרט, כי כאשר המעוול אינו משתף פעולה או מסרב להסיר את הפרסום השגוי, על מנוע החיפוש לפעול להסרת הפרסום כאשר תוצאת החיפוש שגויה.²⁸ יישום מאוחר של פסק הדין, בעניין ברבי, מהווה אינדיקציה לכך שישנם ניצני פסיקה להכרה במנגנון מצומצם של "הודעה והסרה" ביחס לגוגל מכוח דיני הנזיקין.²⁹

14.8. לאחרונה, פורסם דו"ח ועדת הנדל, אשר נועד להציע מודלים נוספים להתמודדות עם סיכון תכנים ברשתות חברתיות. מן הדו"ח עולה העמדה שניתן לראות בפלטפורמות קרובים לגופים דו-מהותיים.³⁰ קביעה זו מאפשרת להטיל חובות מכוח המשפט הציבורי, מתוך ההנחה שמדובר על חברה פרטית שממלאת פונקציה ציבורית.³¹ מגמות אלו מאפשרות הרחבות נוספות של קידום אחריותיות לרשתות חברתיות, גם בהקשרן הישראלי.

14.9. הנקודה החשובה ביחס לכל ההתפתחויות הרגולטוריות הללו היא שהן מהוות תגובה מאוחרת, כנראה מאוחרת מדי. נוהגי השימוש וההסדרים על פיהן פועלות פלטפורמות תוכן התעצבו בעולם ובישראל בסביבה נטולת רגולציה ויצרו לדעת חוקרים שורה של נזקים חברתיים שרגולציה עכשווית תתקשה לעקור מן השורש. תעשיית הפלטפורמות צמחה והפכה לחלק משמעותי מהכלכלה בעידן הדיגיטלי, ויש כוחות עצומים המקשים על פיתוח ואימוץ של רגולציה שתיצור שינוי משמעותי. הפעילות של כלי בינה מלאכותית בעיצוב ומימוש זכויות משתמשים צמחה תחת מציאות שבה לפלטפורמות היו מנגנוני תמריצים מעוותים, והדבר ניכר בהתפשטות של תכנים פוגעניים מחד גיסא, וצנזורת יתר של משתמשים ללא הליך הוגן מאידך גיסא.

15. אנו סבורים כי השתלשלות זו מלמדת על לקח חשוב בהקשר הנוכחי: כלי בינה מלאכותית המופעלים על ידי שחקנים פרטיים עשויים להיות בעלי השפעה משמעותית ומיידית על זכויות אדם. רגולציה מאוחרת שלהם עלולה לא להספיק. מסיבה זו, ראוי כי בבחינת המשטר הרגולטורי שיחול בישראל, ייתן הרגולטור את הדעת על פיתוח ויישום מידי יותר של כלי רגולציה לעידוד האחריותיות של מפעילי מערכות בינה מלאכותית מסוג זה.

לסיכום, המשטר הרגולטורי שיאומץ בישראל צפוי להשפיע בצורה ניכרת על התפתחות התעשייה מזה והשימושים החברתיים בטכנולוגיה מזה. אנו סבורים כי המסקנות שאומצו במסמך העקרונות מהוות התפתחות חיובית לצמיחת רגולציה מטיבה. בנייר זה ביקשנו לעמוד על שני תחומים עיקריים בהם נדרשת תשומת לב נוספת להשפעות של המדיניות שנבחרה ולאיומן תיקונים שיבטיחו כי לצד עידוד היזמות והחדשנות, תתפתח בישראל סביבה רגולטורית קוהרנטית ואחידה שתדע להבטיח שמירה על זכויות אדם.



בפרט, טיפול בגירעון בדמוקרטי העולה משימוש במערכות בינה מלאכותית בעלות השפעה על זכויות תקדם מטרה מרכזית אותה מסמך הרגולציה דן מציב לעצמו – תפיסת האדם במרכז. בנוסף לשיפור מאזן ההגנה על זכויות האדם של משתמשי בינה מלאכותית, הדבר יגביר את אמון הציבור במערכות בינה מלאכותית.

נשמח להרחיב בכל עניין הנוגע לקול הקורא ולנייר עמדה זה, וכן להציג את עמדתנו בעל-פה בפניכם. אנו עומדים לרשותכם בכל שאלה או עניין.

ד"ר אלעד גיל

עמית בכיר וראש המחקר, תכלית- המכון
למדיניות ישראלית
חוקר, מרכז פדרמן לחקר הסייבר,
האוניברסיטה העברית

ד"ר טל מימרן

ראש תוכנית האמנה החברתית בעידן
הדיגיטלי, תכלית- המכון למדיניות
ישראלית
מנהל מחקר, מרכז פדרמן לחקר הסייבר,
האוניברסיטה העברית

¹ Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework (United Nations Human Rights Office of the High Commissioner) (2011) ניתן לראות תאגיד טכנולוגיה ענקיים שהצהירו על קבלת העקרונות: כמו גוגל, (<https://about.google/human-rights/>) מייקרוסופט (https://www.microsoft.com/en-us/corporate-responsibility/human-rights-statement?activetab=pivot_1%3aprimar5/responsibility/pdf/Apple-Supplier-Code-of-Conduct-and-Supplier-Responsibility-Standards.pdf) ואפל (<https://www.apple.com/supplier-responsibility/pdf/Apple-Supplier-Code-of-Conduct-and-Supplier-Responsibility-Standards.pdf>).

² עקרונות מדיניות, רגולציה ואתיקה בתחום הבינה המלאכותית (משרד החדשנות, המדע והטכנולוגיה, משרד המשפטים), 17-16 (2022) (להלן: "טיוטת עקרונות מדיניות").

³ יודגש כי אין בכוונתנו לטעון כי חקיקה כוללת היא המענה הנכון לבעיה זו. הסדרה רוחבית של בינה מלאכותית כללית, במיוחד יישומים המבוססים על קוד פתוח, היא בעלת תועלת מוגבלת והשפעות של אפקט מצנן על התעשייה. ראו, Alex Engler, [The EU's attempt to regulate open-source AI is counterproductive](https://www.brookings.edu/articles/the-eus-attempt-to-regulate-open-source-ai-is-counterproductive/), Brookings (Aug. 24, 2022).

⁴ EU Commission, Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts, Brussels 21.04.21, COM(2022) 206 Final, p.9 (<https://eur-lex.europa.eu>) ("Draft AI Regulation").

⁵ שם, 11.

⁶ שם, 3.

⁷ Articles 8-29.

⁸ Article 62.

⁹ Article 63.

Council of Europe, Ad Hoc Committee on Artificial Intelligence (CAHAI), Possible elements of a legal framework on artificial intelligence based on the Council of Europe's standards on human rights, democracy and the rule of law, pp.2-4 Strasbourg 03.12.21, <https://rm.coe.int/cahai-2021-09>

שם, 6.

מדריך ממשלתי: הערכת השפעת רגולציה מדריך RIA (משרד ראש הממשלה) עמ' 66, 73-71 (2015); **מדריך: הפחתת הנטל הרגולטורי: הנחיות לניהול תהליך עבודה** (משרד ראש הממשלה) עמ' 11 (2015).

13 ר' בעניין זה *Judgment of the Court (Grand Chamber) Data Protection Commissioner v Facebook Ireland* European Commission, *European Limited and Maximilian Schrems* (C-311/18; 16 July, 2020) *Commission adopts adequacy decision on Japan, creating the world's largest area of safe data flow* (press release, 23 Jan., 2019)

14 Rebecca Crootof & BJ Ard, *Structuring Tachlaw*, 34 HARV. J. L. TECH. 348 (2021).

15 לדוגמה לאימוץ גישה מטא-רגולטורית מתירנית בשלב החדירה של טכנולוגיות בינה מלאכותית חדשות ראו משרד המשפטים, גילוי דעת: שימושים בתכנים מוגנים בזכויות יוצרים לצורך למידת מכונה (2022).

16 טיוטת עקרונות מדיניות, 18.

17 DAVID COLLINGRIDGE, *THE SOCIAL CONTROL OF TECHNOLOGY* 11 (1981).

18 מיכל לביא **אחריות מתוכי תוכן לעוולות ביטוי: הקשר חברתי, משפט וטכנולוגיה** 76-73, 80 (2018).

19 Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi and Michele Starnini, *The echo chamber effect on social media* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES Vol. 118(9) 1 (2021).

20 *NetChoice v. Paxton US*, no. 21-51178 (2022).

21 מדובר בהחלטה השנייה במחלוקת (לאור הרקע של החוק, שנועד למנוע צנזור של ביטוי רפובליקני לאור מאורעות הקפיטול שהובילו למדיניות סינון תכנים קפדנית יותר מטעם הפלטרמות). ראו, למשל Alan Z. Rozenshtein, *The Fifth Circuit's Social Media Decision: A Dangerous Example of First Amendment Absolutism* lawfare (20.9.2022) <https://www.lawfareblog.com/fifth-circuits-social-media-decision-dangerous-example-first-amendment-absolutism>

22 *Gonzalez v. Google, LLC*, US No. 18-16700 (9th Cir. 2021).

23 ר' טל מימרן, "הרשתות החברתיות: פיקוח מפחיתשנאה" **ישראל היום** (18.10.2022).

24 The Social Atrocity Meta and the Right to Remedy for the Rohingya (Amnesty international) pp.26-61(2022).

25 Statement to the Human Rights Council by Nicholas Koumjian, Head of the Independent Investigative Mechanism for Myanmar, on the 51st Regular Session of the Human Rights Council <https://iimm.un.org/statement-to-the-human-rights-council-by-nicholas-koumjian-head-of-the-independent-investigative-mechanism-for-myanmar-on-the-51st-regular-session-of-the-human-rights-council/>

26 Laurie Kim, Tech The Republic of Gambia. v. Facebook, Inc., 567 F. Supp. 3d 291 (D.D.C. 2021); Accountability in Face of Genocide: Gambia v. Facebook, 36 Emory Int'l L. Rev. 165 (2022); Michael A Becker, *The Gambia v Facebook: Obtaining Evidence for Use at the International Court of Justice (Part I)* EJIL: TALK! (5.10.2021); <https://www.ejiltalk.org/the-gambia-v-facebook-obtaining-evidence-for-use-at-the-international-court-of-justice-part-i/>; Priya Pillai, The Republic of The Gambia v. Facebook, Inc.: Domestic Proceedings, International Implications, OPINIO JURIS (Aug. 8, 2020), <https://opiniojuris.org/2020/08/08/the-republic-of-the-gambia-vfacebook-inc-domestic-proceedings-international-implications>.

27 בהקשר הזה, העניין נסגר בפשרה ר' שירות גלובס, "גוגל כבר לא תציע השלמת חיפוש ל'צחי רוזנטל'" **גלובס** (24.11.2013) <https://www.globes.co.il/news/article.aspx?did=1000896242>

28 ע"א (מחוזי ת"א) 39345-11-14 **עמי סביר נ' שאול בר נוי** פס' 11-13 (נבו) 22.06.2015; רע"א 4673/15 **שאול בר נוי נגד עמי סביר עו"ד ואח'** פסקות 9-8 לפסק הדין של השופט זילברטל (2016).

29 ת"א (שלום ת"א) 67847-03-18 **ברבי נ' Google Inc** פס' 31-26 לפסק דינו של השופט הימן (נבו) 5.5.2021.

30 הצוות המייעץ לשר התקשורת לבחינת האסדרה על פלטפורמות תוכן דיגיטליות **דוח המלצות הוועדה** 10-9 (2022).

31 אסף הראל **גופים ונושאי משרה דו-מהותיים** (מהדורה שנייה) 78-72 (2019).