מדריך ליישום טכנולוגיות מגבירות-פרטיות במערכות בינה מלאכותית

דצמבר 2025

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

מדריך ליישום טכנולוגיות מגבירות-פרטיות במערכות בינה מלאכותית

הקדמה

מערכת בינה מלאכותית (Artificial Intelligence - AI) פירושה "מערכת ממוכנת אשר מסיקה מהקלט המוזן לה כיצד להפיק תחזיות, תוכן, המלצות או החלטות שיכולות להשפיע על הפרט או פעילותו של בעל השליטה או המחזיק במאגר, הפועלת ברמות משתנות של עצמאות והסתגלות".[1]

השימושים במערכות בינה מלאכותית מציבים אתגרים רבים להגנה על הפרטיות.[2] מערכות בינה מלאכותית מבוססות על עיבוד[3] כמויות גדולות של מידע, שחלקו עשוי להיות גם מידע אישי.[4] מידע זה נדרש לכל אורך מחזור החיים של מערכות בינה מלאכותית – בשלב הבניה של מערכת חדשה (כולל פיתוח ואימון), ובשלב השימוש במערכת (כולל שיפור ודיוק המערכת לאורך זמן).[5]

טכנולוגיות מגבירות פרטיות (Privacy-Enhancing Technologies - PETs) הן אוסף של גישות ופתרונות דיגיטליים המסייעים להגנה על מידע אישי.[6] טכנולוגיות אלו מאפשרות לסייע ולאזן בין היכולות של מערכות הבינה המלאכותית והשימושים בהן לבין הגנה על פרטיות המשתמשים. באמצעות טכנולוגיות מגבירות פרטיות ניתן לצמצם את סיכוני הפגיעה בפרטיות לכל אורך מחזור החיים של מערכות בינה מלאכותית.

מטרות המסמך

1. להציג גישות לצמצום הסיכונים לפרטיות הכרוכים בפיתוח של מערכות בינה מלאכותית ובשימוש בהן באמצעות טכנולוגיות מגבירות-פרטיות.
2. להציג דוגמאות ליישום גישות אלו בפרויקטים בעולמות תוכן שונים.

---

[1] ראו הגדרת "מערכת בינה מלאכותית" בטיוטת ההנחיה של הרשות להגנת הפרטיות בנושא "תחולת הוראות חוק הגנת הפרטיות על מערכות בינה מלאכותית" (אפריל 2025), אשר פורסמה להערות הציבור. טיוטת ההנחיה זמינה כאן. השוו גם להגדרה העדכנית בעקרונות ארגון ה-OECD ממאי 2024:

An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment. OECD Council Recommendation on Artificial Intelligence (amended 3 May 2024).

[2] ראו דוח הביניים של הצוות הבינמשרדי בנושא בינה מלאכותית בסקטור הפיננסי (2024), עמ' 93-72. זמין כאן. ראו גם מיכאל בירנהק, פרטיות ובינה מלאכותית, משפט חברה ותרבות, ח' (2024).

[3] עיבוד כהגדרתו בסעיף 3 לתיקון מס' 13 לחוק הגנת הפרטיות, התשמ"א-1981 – "'עיבוד', 'שימוש' – כל פעולה שמבוצעת על מידע אישי, לרבות קבלתו, איסופו, אחסונו, העתקתו, עיון בו, גילויו, חשיפתו, העברתו, מסירתו או מתן גישה אליו".

[4] מידע אישי כהגדרתו בסעיף 3 לתיקון מס' 13 לחוק הגנת הפרטיות, התשמ"א-1981 – "'מידע אישי' – נתון הנוגע לאדם מזוהה או לאדם הניתן לזיהוי; לעניין הגדרה זו, 'אדם הניתן לזיהוי' - מי שניתן לזהותו במאמץ סביר, במישרין או בעקיפין, בכלל זה באמצעות פרט מזהה, כגון שם, מספר זהות, מזהה ביומטרי, נתוני מיקום, מזהה מקוון, או נתון אחד או יותר הנוגע למצבו הפיזי, הבריאותי, הכלכלי, החברתי או התרבותי".

[5] Explanatory Memorandum on The Updated OECD Definition of an AI System, OECD Artificial Intelligence Papers, No. 8 (March 2024).

[6] לפירוט נוסף: מדריך לטכנולוגיות מגבירות פרטיות, הרשות להגנת הפרטיות (2025).

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

אוכלוסיית יעד

מסמך זה נועד לשמש בעלי תפקידים האחראים על הערכת סיכוני פרטיות ויישום פתרונות מתאימים בפרויקטים לפיתוח מערכות ושירותים דיגיטליים הכוללים רכיבי בינה מלאכותית. המסמך עשוי לשמש גם עבור אנשי פיתוח בתחום הבינה המלאכותית, ככלי מסייע להטמעת טכנולוגיות מגבירות-פרטיות כבר מהשלבים המוקדמים של ייזום הפרויקט ולאורך כל מחזור חייו. בפרט, המסמך רלוונטי לבעלי התפקידים הבאים:

- ממוני הגנת הפרטיות (Data Protection Officers – DPOs)[7] ויועצים משפטיים העוסקים בהיבטי פרטיות בפרויקטים המשלבים בינה מלאכותית.
- מנהלי מוצר ומנהלי פרויקטים המעורבים בתחום הפיתוח, ההטמעה והתפעול של מערכות מבוססות בינה מלאכותית בשירותים ומוצרים דיגיטליים.

השימוש במסמך אינו דורש ידע טכני או רקע טכנולוגי. בהתאם לכך, רמת הפירוט הניתנת עבור כל טכנולוגיה נועדה לאפשר היכרות עם מהותה, ולסייע בהערכה עקרונית של מידת התאמתה למשימות או לתחומים מסוימים. יש לציין כי המסמך אינו מיועד לתאר ולפרט את מלוא המורכבות של מערכות בינה מלאכותית, אלא מתמקד בהיבטים הנדרשים להבנת היישום של טכנולוגיות מגבירות-פרטיות בהקשר זה.

ביישום מעשי, לרבות כחלק מהתמודדות עם אתגרים וסיכונים אפשריים, יש להיעזר במקורות נוספים. לאורך המסמך מופיעים מקורות וקישורים להרחבה. יש לציין שלצורך בחינת הארכיטקטורה המתאימה לפרויקט או מוצר טכנולוגי, אין במסמך זה כדי להחליף התייעצות עם המומחים הרלוונטיים.

אופן הצגת הטכנולוגיות ויישומן

מסמך זה מציג טכנולוגיות מגבירות-פרטיות ויישומן לעולמות הבינה המלאכותית, תוך הסבר של עקרון פעולתן לשם השגת יעדי הגנת הפרטיות. כל טכנולוגיה מלווה, לפי העניין, בדוגמאות או איורים שממחישים את המנגנון המרכזי שלה.

בהמשך המסמך מובאים שיקולים מרכזיים ליישום מעשי של כל טכנולוגיה, לצד דוגמאות ליישומים ככל האפשר ממגוון תחומים ועולמות תוכן בהקשר של בינה מלאכותית. בנוסף, מוזכרים אתגרים, מגבלות וסיכונים הנלווים ליישום, תוך הדגשת סוגיות שדורשות תשומת לב מיוחדת.

הדוגמאות המובאות במסמך נועדו להמחיש בצורה פשוטה ותמציתית את עקרונות הפעולה של הטכנולוגיות בהקשרים של בינה מלאכותית. יש לראות בדוגמאות אלו המחשה בלבד – הן אינן מהוות המלצה לדרך פעולה מסוימת או הגבלה על שימושים אפשריים אחרים.

[7] ראו טיוטת גילוי הדעת של הרשות להגנת הפרטיות בנושא "מינוי ממונה על הגנת הפרטיות בארגון לפי דרישות תיקון 13 לחוק הגנת הפרטיות", אשר פורסם להערות הציבור. גילוי הדעת זמין כאן.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

יישום מעשי של טכנולוגיות מגבירות-פרטיות בעולם הבינה המלאכותית מחייב בחינה מעמיקה של מגוון שיקולים, לרבות מאפיינים ייחודיים של כל מקרה ופרויקט, הקשרים טכנולוגיים, מסגרת משפטית ושיקולים ארגוניים.

מקורות מידע למסמך

המסמך והדוגמאות שבסופו מתבססים בחלקם על מספר מקורות מרכזיים בנושא יישום טכנולוגיות מגבירות-פרטיות במערכות בינה מלאכותית, וביניהם:

1. טיוטת הנחיית הרשות להגנת הפרטיות בנושא תחולת הוראות חוק הגנת הפרטיות על מערכות בינה מלאכותית, טיוטה להערות הציבור, הרשות להגנת הפרטיות, אפריל 2025.
2. מדריך לטכנולוגיות מגבירות-פרטיות - PETs, הרשות להגנת הפרטיות, פברואר 2025.
3. מסמך בנושא שיתוף מודלים אמינים של בינה מלאכותית באמצעות טכנולוגיות מגבירות-פרטיות של ה-OECD, יוני 2025.[8]
4. מדריך לניהול סיכונים ושימוש אחראי בכלי בינה מלאכותית (AI) במגזר הציבורי, גרסה להערות ציבור, יוני 2025.
5. מאגר מקרי שימוש בטכנולוגיות מגבירות פרטיות של משרד המדע, החדשנות והטכנולוגיה של ממשלת בריטניה.[9]
6. סקירה בנושא טכנולוגיות מגבירות-פרטיות וטכנולוגיות משמרות-פרטיות של CIPL[10] - Center for Information Policy Leadership.
7. מאגר מקרי שימוש בטכנולוגיות מגבירות-פרטיות של המרכז לאתיקה של מידע וחדשנות של ממשלת בריטניה (כיום תחת משרד המדע, החדשנות והטכנולוגיה).[11]
8. מאגר מחקרי מקרה בשימוש בטכנולוגיות מגבירות-פרטיות ליישומי סטטיסטיקה רשמית של צוות המשימה הגלובלי של קבוצת העבודה של האו״ם לטכנולוגיות מגבירות-פרטיות.[12]

[8] OECD (2025), "Sharing trustworthy AI models with privacy-enhancing technologies", OECD Artificial Intelligence Papers, No. 38, OECD Publishing, Paris, https://doi.org/10.1787/a266160b-en.
[9] Repository of Privacy Enhancing Technologies (PETs) Use Cases, Updated: November 7, 2024.
[10] Privacy-Enhancing and Privacy Preserving Technologies: Understanding the Role of PETs and PPTs in the Digital Age, December 2023.
[11] Centre for Data Ethics and Innovation Repository of Use Cases, Updated June 2023.
[12] United Nations Global Working Group Task Team on Privacy Preserving Techniques: Case Study Repository, last modified on Apr 11, 2024.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

תוכן העניינים

*3103 | 073-3928529 | ppa@justice.gov.il
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

מידע אישי במערכות בינה מלאכותית

מבוא למערכות בינה מלאכותית

מערכת בינה מלאכותית מורכבת מאלגוריתמים ומודלים מתמטיים שמאפשרים לה לפעול ברמות שונות של אוטונומיה. המערכת מבוססת בין היתר על אלגוריתמים של למידת מכונה (Machine Learning) שמעבדים ומנתחים נתונים (קלט – Input) המאוחסנים במאגרי נתונים, אותם היא מעבדת ומנתחת, ועל בסיסם מייצרת תוצרים, קרי, פלט (Output)[13].

ייחודה של מערכת זו נעוץ בכך שהיא מחקה תהליכים קוגניטיביים אנושיים שנחשבו בעבר כבלעדיים לבני אדם. לרוב, המערכת פועלת לאור הנחיות טקסטואליות או קוליות (פרומפטים – Prompts) שניתנות על ידי משתמשים אנושיים. קיימים מספר סוגים של בינה מלאכותית, כל אחד מתמקד בתחום יישום שונה ומספק תוצאות שונות. בין הסוגים, בולטת בשנים האחרונות הבינה המלאכותית היוצרת (Generative AI), שמאפשר לייצר תוצרים מקוריים לכאורה, כמו טקסטים, סרטונים, ותמונות.

מחזור החיים של מערכות בינה מלאכותית[14]

מערכת בינה מלאכותית לרוב כוללת מודל[15] אחד או יותר המפותח על בסיס מידע (קלט) או הוראות המפעיל. מודלים של בינה מלאכותית יכולים לכלול ייצוגים סטטיסטיים של הקלט המאפשרים הפקה של פלט רצוי בתנאים ובסיטואציות מתאימות. מודל בינה מלאכותית יכול לאפשר למערכת בינה מלאכותית לבחור פעולה מועדפת על ידי בחינת ההשלכות החזויות שלה שהוסקו ממידע הקלט. מודלים של בינה מלאכותית יכולים להיבנות באופן ידני על ידי מתכנת אנושי או באופן אוטומטי, למשל באמצעות אלגוריתמים של למידת מכונה ותהליכי קבלת החלטות.

מחזור החיים של מערכת בינה מלאכותית כולל מספר שלבים שנועדו לבנות, לאמן ולאפשר את השימוש המיטבי במודל. במסמך זה נתמקד בשני שלבים מרכזיים:[16]

1. **בניה**: זהו שלב הפיתוח והאימון של מערכת חדשה. במהלך הבניה קובעים את המטרות והפונקציות של המערכת. הבחירה במודלים מתמטיים מתבצעת בהתאם לסוג המשימות שהמערכת צפויה לבצע, כאשר כל מודל מותאם לאופי הנתונים והפעולות הנדרשות.
התהליך כולל בניית מאגר נתונים אשר ישמש כבסיס לאימון המודל. המאגר צריך לכלול היקף נתונים רחב, שיכלול את התרחישים שבהם המערכת תפעל או התוצרים אותם

[13] משרד החדשנות, המדע והטכנולוגיה, עקרונות מדיניות, רגולציה ואתיקה בתחום הבינה המלאכותית (2023), עמ' 21-22.
[14] הפרק מבוסס על מסמך דברי ההסבר של ה-OECD להגדרה המעודכנת של בינה מלאכותית, לעיל ה"ש 5.
[15] מודל מוגדר כ"ייצוג פיזי, מתמטי או לוגי אחר של מערכת, ישות, תופעה, תהליך או נתונים" בתקן ISO/IEC 22989:2022 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology.
בשפת המקור: Model: physical, mathematical or otherwise logical representation of a system, entity, phenomenon, process or data
[16] מבוסס על ההבחנה המובאת במקור בה"ש 5.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

המערכת תצטרך לייצר. המערכת מתאמנת על פי הנתונים המוזנים, מנתחת אותם ולומדת לזהות דפוסים, מגמות וקשרים שבין גורמים שונים באמצעות למידה והסקה מתוך הקלט, שיכול לכלול נתונים הרלוונטיים למשימה שתתבצע, בקשות משתמש או שאילתות חיפוש. כדי להבטיח תוצאות איכותיות נדרש שהקלט יהיה איכותי, מאוזן, מייצג ורלוונטי למטרה המבוקשת. במהלך ולאחר הבניה, ביצועי המודל נבדקים כדי לוודא שהוא מפיק תוצאות שעומדות ביעדים שנקבעו. המפתחים מנתחים את הפלטים שהמודל מייצר, בודקים את רמת הדיוק והאמינות שלו, ומבצעים התאמות נדרשות לפי צורך. שלב זה חיוני כדי להבטיח שהמודל יכול לבצע את המשימה בצורה מיטבית ויעילה.

2. **שימוש**: בשלב זה המערכת זמינה לשימוש המשתמשים, אשר מזינים פרומפטים והנחיות ספציפיות כדי להכווין את המערכת להפקת פלטים מותאמים אישית (המלצות, תחזיות והחלטות). המערכת מקבלת קלט חדש – לדוגמה, נתוני משתמש, טקסטים או תמונות, ומפעילה את המודלים שאומנו מראש לצורך ניתוח, חיזוי, הפקת המלצות, קבלת החלטות או יצירת תוכן חדש. ביצועי המערכת בשלב זה מושפעים מהאופן שבו תוכננה ונבנתה, וכן מאיכות ודיוק הקלט המוזן לה.
לדוגמה, מערכת בינה מלאכותית תומכת החלטה קלינית עשויה לזהות שממצא מסוים בסריקת דימות הוא גידול חשוד; היא עשויה להמליץ על בדיקת המשך או ביופסיה, או שהיא עשויה להתריע בפני הצוות הרפואי באופן אוטומטי.

שלבי הבניה והשימוש עשויים להשתלב ביניהם: בשלב הבניה של מערכת בינה מלאכותית בחלק מהמקרים נדרש לעשות שימוש חוזר בנתוני קלט לצורך שיפור הביצועים והדיוק באמצעות התאמה ועדכון של המודל. במקרים אחרים, במהלך השימוש במערכת, נתוני קלט חדשים (כגון שאילתות משתמשים או תוצאות בפועל) משמשים לבניה (עדכון ושיפור המודל) – בין אם באופן שוטף (ובין אם בפרקי זמן מוגדרים.

מידע אישי בבניה ושימוש במערכת בינה מלאכותית

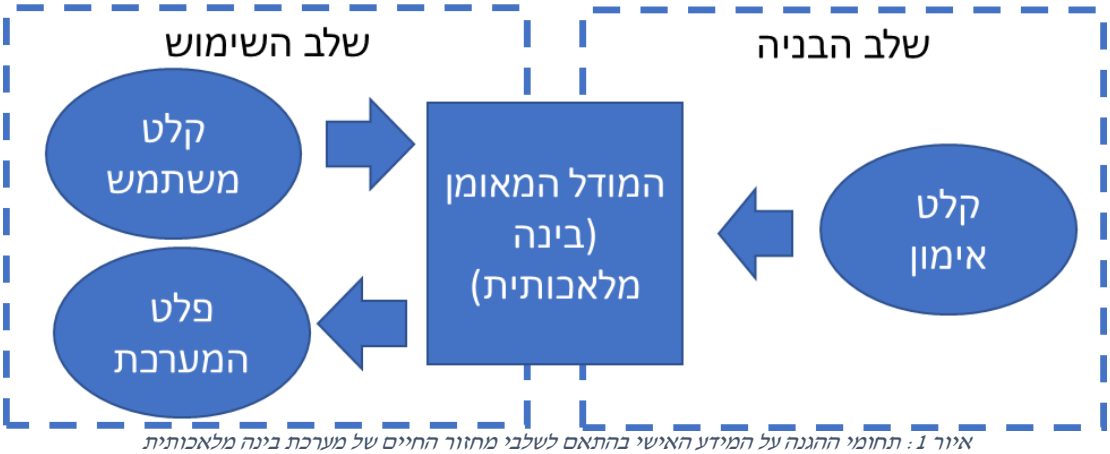
במהלך הבניה והשימוש במערכת בינה מלאכותית, הקלט למערכת עשוי לכלול מידע אישי. מידע אישי זה עשוי להשפיע על התנהגות המודל ואף להיטמע בו, דבר שעשוי להביא לחשיפה של נתונים אישיים דרך הפלט. מכאן, הגנה על מידע אישי במערכת בינה מלאכותית עוסקת בתחומים הבאים:

1. **קלט אימון (Training Data)**: מידע אישי המשמש לאימון המערכת. לדוגמה – רשומות רפואיות, נתוני משתמשים או עסקאות פיננסיות.
2. **המודל המאומן (Trained Model)**: מידע אישי המוטמע במודל לאחר האימון. בתנאים מסוימים, שחזור מידע אישי מתוך המודל עשוי להיות אפשרי באמצעות הסקת נתוני קלט אופייניים לפי תגובות המודל (Model Inversion) ושיטות נוספות.
3. **קלט משתמש (User Input)**: מידע אישי המוזן בשימוש במערכת. לדוגמה – שאילתה חופשית, בקשת שירות, תמונה או טקסט אישי. קלט משתמש עובר עיבוד על ידי המודל ועשוי לשמש לאימון ועדכון המודל לצורך שיפור ביצועים עתידיים.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

4. **פלט המערכת (Model Output)**: מידע אישי המופק על ידי המערכת, בין אם נוצר מהקלט, נלמד ממידע קודם, או נחשף בטעות. לדוגמה – מענה מותאם אישית, המלצה על בסיס פרופיל אישי, או מידע שנשלף בטעות מהמודל.

איור 1 מתאר את תחומי ההגנה על המידע האישי בהתאם לשלבי מחזור החיים של מערכת בינה מלאכותית:



איור 1: תחומי ההגנה על המידע האישי בהתאם לשלבי מחזור החיים של מערכת בינה מלאכותית

הגנה על מידע אישי במערכות בינה מלאכותית

טכנולוגיות מגבירות-פרטיות

טכנולוגיות מגבירות-פרטיות הן אוסף של שיטות, תהליכים וכלים דיגיטליים המסייעים בהגנה על מידע אישי. טכנולוגיות מגבירות-פרטיות מאפשרות לערפל את המידע האישי ולצמצם את רמת הפירוט שלו, להקטין את הסיכון לחשיפת מידע אישי במהלך העיבוד ולהגביר את השליטה בשימוש במידע אישי.[17]

העקרונות, הפרקטיקות והפירוט של סוגי טכנולוגיות מגבירות-פרטיות הובאו בהרחבה במדריך שפרסמה הרשות להגנת הפרטיות בנוגע לטכנולוגיות אלו.[18] במסמך זה נתמקד בטכנולוגיות שהן בעלות התאמה ורלוונטיות למערכות בינה מלאכותית בחלוקה לשלוש קבוצות מרכזיות, כאשר כל אחת מהן מציעה עקרון פעולה מרכזי ומובחן לצמצום הסיכון לפגיעה בפרטיות בהקשרי פיתוח ושימוש במערכות בינה מלאכותית.

הקבוצה הראשונה של טכנולוגיות מגבירות-פרטיות פועלת לפי עקרון **שינוי המידע**. קבוצה זו כוללת מגוון טכניקות שנועדו להסוות, להסיר או להחליף מאפיינים מזהים מתוך הנתונים הגולמיים. דוגמאות נפוצות כוללות הסרת מזהים ישירים כגון שם או מספר תעודת זהות, טשטוש

[17] הגדרה מתוך המדריך לטכנולוגיות מגבירות-פרטיות, הרשות להגנת הפרטיות, 2025.
[18] ראו ה"ש קודמת לקישור למסמך.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

מזהים עקיפים (כגון מיקוד גיאוגרפי, תאריך לידה מדויק או מאפיינים ייחודיים נוספים), עיגול ערכים, הוספת רעש אקראי לנתונים, הפקת נתונים סינתטיים הדומים במבנה הסטטיסטי לנתונים המקוריים אך מבלי לשמר רשומות אמיתיות, ולעיתים גם הכללה של ערכים לצמצום רזולוציה (כגון קיבוץ גילאים ספציפיים לטווחים של מספר שנים).

העיקרון המרכזי בקבוצה זו הוא שינוי ישיר של הנתונים הגולמיים, כך שהם לא ישמרו על הזיהוי המקורי של נושאי המידע הנוגעים לנתונים, ועם זאת ישמרו במידה מסוימת על שמישותם. זהו איזון עדין: מצד אחד, אנו רוצים להפחית את הסיכון לפגיעה בפרטיות על ידי מזעור הסיכוי לזיהוי של נושאי המידע, ומצד שני, לשמר את יכולת הנתונים לתרום לאימון מערכת הבינה המלאכותית.

בין השיטות בקבוצה זו ניתן למנות:

א. **התממה – Anonymization**: התממה היא הסרת מאפיינים או שינוי ערכים על מנת לצמצם או למנוע זיהוי של נושא המידע או יכולת שחזור מידע אישי מתוך קלט האימון או מתוך המודל המאומן. הפעלת התממה על קלט האימון מקשה על קישור בין דוגמה מסוימת במידע לבין אדם מסוים, ובהתאם מקטינה את היכולת לזיהוי חוזר מתוך המודל המאומן. אחת הדרכים הנפוצות לביצוע התממה היא עיגול ערכים או קיבוץ טווחים – לדוגמה, במקום גיל מדויק (כגון 47), המידע יוצג בטווחים (40–49), או במקום הכנסה חודשית מדויקת – טווחי הכנסה. הדבר מצמצם את הסיכון לזיהוי אך מפחית גם את רזולוציית הנתונים.

ב. **מידע סינתטי – Synthetic Data**: מידע סינתטי הוא מידע מלאכותי שנוצר באופן אוטומטי על ידי אלגוריתמים, במטרה לדמות מידע "אמיתי" – כלומר, מידע שמבוסס על מאפיינים סטטיסטיים של נתוני אמת, אך אינו אמור לשמר את הפרטים המקוריים של אף אדם או אירוע. במילים אחרות, זהו מידע בדוי אך מדויק מבחינה מבנית, שנועד לשמש כתחליף לנתונים רגישים – למשל בעת פיתוח, בדיקה או אימון של מערכות בינה מלאכותית. מידע סינתטי מאפשר לאמן מודלי בינה מלאכותית תוך צמצום הסיכון לחשיפה למידע אישי מזוהה.

ג. **פרטיות דיפרנציאלית – Differential Privacy**: שיטה זו מגבילה את ההשפעה של כל נתון בודד על תהליך הלמידה, ומונעת את האפשרות לזהות או לשחזר מידע אישי של משתמשים מתוך המודל המאומן, גם אם לתוקף יש גישה לפלט של המודל או למידע חיצוני. באמצעות הפעלת פרטיות דיפרנציאלית, המודל מצמצם עד מאוד את התלות בפרטים מזהים של דוגמה אחת.

היתרון המרכזי בקבוצה זו הוא הפשטות היחסית והיכולת ליישם את השיטות גם כאשר אין צורך במשאבי חישוב כבדים או בשיתוף בין מספר גורמים. אולם שינוי הנתונים עלול לפגוע באיכות המודל. כאשר מוסיפים רעש או מבצעים הכללות, המידע מאבד דיוק או רזולוציה – מה שעלול

*3103 | 073-3928529 | ppa@justice.gov.il
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

להשפיע על הביצועים של מודלים מבוססי בינה מלאכותית. לדוגמה, מודל שמבוסס על טווחי גיל בלבד עשוי להיות פחות מדויק לעומת מודל שאומן על מידע הכולל גיל מדויק.[19]

הקבוצה השנייה של טכנולוגיות מגבירות-פרטיות פועלת לפי עקרון **החישוב המבוזר**. גישה זו נועדה לאפשר למערכות לבצע חישובים על נתונים רגישים מבלי לחשוף את המידע המעובד, תוך שימוש בשיטות קריפטוגרפיות מתקדמות וטכניקות עיבוד מבוזר. בניגוד לעקרון שינוי המידע, לפיו הנתונים עצמם משתנים או מעובדים מראש לצורך הגנה עליהם, כאן הדגש הוא על אופן השימוש והעיבוד של הנתונים ללא שינוי הנתונים עצמם, ובהתאם לכך מבלי לוותר על דיוק או שימושיות.

העיקרון המרכזי בגישה זו הוא פיזור הנתונים או החישוב בין מספר ישויות או מקורות, כך שאף גורם יחיד אינו מחזיק בכל המידע הרגיש. במקום להעביר את כל הנתונים לשרת מרכזי, הם נשארים במקומם – למשל, על מכשיר המשתמש, בשרתים שונים, או אצל שותפים בארגון – והחישוב מתבצע באופן מבוזר או מוצפן. לעיתים, רק התוצאה הסופית של החישוב נאספת, ללא חשיפה ולעיתים אף ללא העברה של הנתונים עצמם.

בין השיטות בקבוצה זו ניתן למנות:

א. **למידה מבוזרת** – Federated Learning: אימון מודלי בינה מלאכותית ללא איסוף של כל נתוני הקלט במקום אחד. במקום להעביר את הנתונים עצמם לשרת מרכזי, המודל נשלח אל המכשירים או השרתים המקומיים, שם הוא מתעדכן על בסיס הנתונים המקומיים ורק העדכונים של המודל (ולא הנתונים עצמם) נשלחים בחזרה לשרת המרכזי, שם הם משולבים לגרסה משופרת של המודל.
למידה מבוזרת מאפשרת לדוגמה לבצע עיבוד מקומי על המכשירים של המשתמשים (כגון טלפונים חכמים, מחשבים ניידים או מכשירים חכמים אחרים). כלומר, הנתונים האישיים לא עוזבים את המכשיר, ואינם חשופים לשרת המרכזי או למשתמשים אחרים. לדוגמה, במערכת השלמה אוטומטית של שאילתות חיפוש על בסיס הקלדות המשתמש, במקום לשלוח את הטקסטים המוקלדים עצמם, המכשיר שולח לשרת רק עדכונים למודל החיזוי ולא את מידע המשתמשים. יש לציין כי המידע האישי הגולמי של נושא המידע אמנם לא נחשף עצמו בפני צדדים שלישיים, אבל נעשה בו שימוש לאימון המודל והמודל מעודכן באופן שהוא יוכל להשפיע על תוצאות החיזוי ואף להפיק את אותו מידע אישי בפלט.

ב. **חישוב רב-משתתפים** – Multi-Party Computation: ביזור מידע אישי באופן המאפשר למספר צדדים לחשב תוצאה משותפת מבלי לחשוף את המידע האישי או הרגיש לכל אחד מהצדדים. כל צד מחזיק בנתונים משלו, ומבצע עליהם חישובים מבלי לגלות אותם לכל צד אחר. בעזרת טכניקות של חישוב רב-משתתפים, כל צד תורם לעדכון המודל ללא חשיפת

[19] עשויים להיות גם מקרים הפוכים, ובהם שימוש במידע השונה מהמידע האישי המדויק יביא דווקא לשיפור בביצועים. לדוגמה, אימון מערכת בינה מלאכותית על מספר רב של דגימות מידע סינתטי (שאינו מייצג מידע אישי של נושאי מידע אמיתיים) עשוי להביא ליכולת הסקה והכללה טובה יותר לעומת אימון המערכת על היקף מצומצם של מידע אישי אמיתי.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

הנתונים האישיים שלו. לדוגמה, מספר ארגונים יכולים לאמן מודל בינה מלאכותית על נתונים רגישים מבלי לחשוף את המידע האישי לכל צד, ובכך לשמור על פרטיות המידע.
חישוב רב-משתתפים מאפשר שיתוף פעולה בין ארגונים שונים (כגון בין חברות, מוסדות פיננסיים או גופי בריאות) לצורך פיתוח מודלים מדויקים יותר מבלי לחשוף את המידע הרגיש של הלקוחות או המשתמשים, או לעבוד על מידע שלא ניתן להעביר אותו בצורה ישירה ממקום למקום. לדוגמה, מספר בתי חולים יכולים לחשב יחד מדד רפואי על כלל המטופלים מבלי לשתף זה עם זה את המידע הרפואי של החולים.

ג. **הצלבת מערכי מידע – Private Set Intersection**: השוואה בין מערכי נתונים מבלי לחשוף את המידע האישי שמצוי רק אצל אחד מהצדדים. כל צד יכול ללמוד אילו פריטים קיימים אצל הצד השני (למשל, רשימות של פריטים, משתמשים או נתונים) מבלי לחשוף פריטים שאינם משותפים ביניהם.
לדוגמה, נניח שבמסגרת סמכויותיו משרד הרווחה מעוניין לפנות למשרד התיירות כדי לדעת על שהיה ארוכה בחו"ל של אנשים הזכאים לקצבה – הוא יכול להעביר תעודות זהות של הזכאים לקצבה למשרד התיירות לבדיקה. במקרה כזה מידע על כלל זכאי הקצבאות יועבר שלא לצורך, כאשר באמצעות הצלבת מערכי מידע ניתן לצמצם את העברת המידע רק לאלו מהם שאכן שהו בחו"ל תקופה ארוכה.

השימוש בחישוב מבוזר מציע יתרונות מהותיים בהקשרים של הגנה על הפרטיות, ובוודאי כאשר מדובר במידע רגיש (כמו מידע רפואי, ביומטרי או פיננסי) או בעבודה מול מספר שותפים (כגון מוסדות ציבוריים, גופים רגולטוריים או ספקי שירותים שונים). הוא מאפשר למצות את הערך מהנתונים תוך מזעור החשיפה, צמצום סיכוני הדלפה, ומניעת ריכוז מידע במקום אחד. עם זאת, קיימים גם אתגרים: מערכות אלו לרוב דורשות תיאום מורכב יותר ועלויות חישוב גבוהות יותר. בנוסף, יש צורך בתשתיות מתקדמות ובסטנדרטים משותפים בין הארגונים המשתתפים.

הקבוצה השלישית של טכנולוגיות מגבירות-פרטיות פועלת לפי עקרון **הפרדה והצפנה**. בקבוצה זו נעשה שימוש בטכניקות מתקדמות שמטרתן הגנה על מידע אישי תוך כדי עיבודו, ולא רק לפני או אחרי העיבוד. זאת, באמצעות הצפנה חזקה או בידוד פיזי/לוגי של תהליכים רגישים, באופן שמאפשר עיבוד נתונים באופן מאובטח.

בעוד שבשיטות שינוי מידע המידע עובר שינוי לפני שמגיע למערכת, ובעוד שבשיטות חישוב מבוזר חלקים שונים של המידע נשארים אצל גורמים שונים – בשיטות הפרדה והצפנה המידע יכול להגיע לעיבוד ריכוזי, תוך הפעלת מנגנונים שמוודאים כי גם במהלך העיבוד עצמו המידע אינו גלוי לגורמים בלתי מוסמכים. שיטות אלו מאפשרות הגנה על המידע בזמן השימוש, מתוך ההבנה שבינה מלאכותית לרוב מצריכה כוח חישוב מרכזי (כגון שרתים בענן), ולעיתים אין שליטה מלאה של בעל המידע על סביבת ההרצה.

בין השיטות בקבוצה זו ניתן למנות:

*3103 | 073-3928529 | ppa@justice.gov.il
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

א. **הצפנה הומומורפית – Homomorphic Encryption**: טכניקת הצפנה שמאפשרת לבצע חישובים על נתונים מוצפנים מבלי לפענח אותם. שרת יכול לבצע פעולות מתמטיות על קלט מוצפן, והתוצאה המוצפנת יכולה להתפענח מאוחר יותר, ולהכיל את התוצאה הנכונה כאילו החישוב בוצע על המידע הגלוי. ניתן לבצע אימון מערכת בינה מלאכותית על נתונים מוצפנים, כך שהמידע נשאר מוגן במהלך התהליך. בדרך זו, גם המודל המאומן וגם פלט המודל יהיו מוצפנים ונגישים רק על ידי משתמשים מורשים. לדוגמה, חברת ביטוח יכולה לשלוח נתונים רפואיים מוצפנים לספק ענן לצורך אימון מודל חיזוי – מבלי לחשוף את המידע עצמו. הספק מבצע את החישובים על המידע המוצפן, והתוצאה הסופית ניתנת לפיענוח רק על ידי החברה.

ב. **סביבת ביצוע מהימנה – Trusted Execution Environment**: עיבוד נתונים בתוך סביבה מבודדת ומאובטחת, שמיועדת למנוע גישה לא מורשית או שינוי של הנתונים. המודל המאומן והנתונים שהוזנו אליו יכולים להישאר בתוך סביבת ביצוע מהימנה לאחר סיום תהליך האימון, כדי לשמור על הגנת המידע. במקרה כזה, המודל נשאר מוגן מפני גישה חיצונית, והמידע האישי מוגן מפני זליגה או חשיפה. לדוגמה, חברת בריאות יכולה לאמן מודל זיהוי מחלות בתוך סביבת ביצוע מהימנה, כך שהנתונים הרגישים לא יהיו נגישים מחוץ לסביבה זו. גם לאחר האימון, המודל יכול להישאר בסביבה המהימנה ונגיש רק לשימושים מאושרים, תוך שמירה על פרטיות המטופלים.

יתרונן הגדול של הטכנולוגיות מקבוצה זו טמון בכך שהן מאפשרת עיבוד מלא של מידע אישי תוך הגנה על פרטיותו. בפרט, טכנולוגיות אלה מאפשרות לאמן ולהפעיל מערכות בינה מלאכותית רגישות על גבי נתונים חסויים גם על גבי שרתים שאינם מהימנים. עם זאת, טכנולוגיות כמו הצפנה הומומורפית דורשות זמן עיבוד וזיכרון רבים יותר, והן אינן נתמכות בכל סביבות הפיתוח. סביבות ביצוע מהימנות דורשות רכיבי חומרה ותוכנה ייעודיים, ואין בהן כדי למנוע לחלוטין אפשרות של פריצה באמצעים הזמינים היום או בעתיד.

ההחלטה על הפעלת טכנולוגיות מגבירות פרטיות בתחומים של קלט אימון, המודל המאומן, קלט משתמש ופלט תלויה במגוון שיקולים, ובהם רגישות המידע, היקפו או אופן הבניה והשימוש במערכת הספציפית. כמו כן, חלק מהטכנולוגיות ניתנות להפעלה מול אתגרים רחבים בבניה ושימוש במערכת בינה מלאכותית, וחלקם מתאים להיבטים ספציפיים בהגנה על מידע אישי במערכת בינה מלאכותית. עקרונות הפעולה המגוונים של טכנולוגיות מגבירות-פרטיות מייצרים השפעה שונה על קלט אימון, המודל המאומן, קלט משתמש ופלט המערכת והבחירה בטכנולוגיה מסוימת צריכה להביא זאת בחשבון.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

כדי לסייע בהתאמה של עקרונות הפעולה של טכנולוגיות מגבירות-פרטיות לתחומים ספציפיים מוצע המיפוי הבא לפי קבוצות ועיקרי פעולתן במערכות בינה מלאכותית בטבלה 1:

טבלה 1: מיפוי של טכנולוגיות מגבירות-פרטיות לפי קבוצות ועיקרי פעולתן במערכות בינה מלאכותית

קבוצה	טכנולוגיה מגבירת-פרטיות	קלט אימון	המודל המאומן	קלט משתמש	פלט המערכת
שינוי המידע	**התממה:** הסרה וטשטוש של מאפיינים מזהים במידע	ערפול המידע האישי וצמצום רמת הפירוט שלו	נגזר מההגנה על קלט האימון	ישימות מוגבלת	ערפול המידע האישי וצמצום רמת הפירוט שלו
	מידע סינתטי: שימוש במידע מלאכותי כתחליף לנתונים רגישים				
	פרטיות דיפרנציאלית: שילוב רעש שמטשטש את המידע האישי				
חישוב מבוזר	**חישוב רב-משתתפים:** ביזור מידע אישי למספר צדדים מבלי לחשוף את המידע האישי או הרגיש	צמצום חשיפת המידע האישי במהלך השימוש	ישימות מוגבלת	ישימות מוגבלת	ישימות מוגבלת
	הצלבת מערכי מידע: השוואה בין מערכי נתונים מבלי לחשוף את המידע שמצוי רק אצל צד אחד				
	למידה מבוזרת: אימון מקומי של מודלי בינה מלאכותית ללא ריכוז של כלל נתוני הקלט				
הפרדה והצפנה	**הצפנה הומומורפית:** ביצוע חישובים על מידע בהיותו מוצפן	עיבוד מידע בהיותו מוצפן **בכל השלבים** לאורך שרשרת הבניה והשימוש במערכת בינה מלאכותית			
	סביבת ביצוע מהימנה: עיבוד נתונים בחלק מבודד ומאובטח של מערכת המחשב	עיבוד מידע מאובטח בסביבה מוצפנת **בשלב אחד או יותר** בשרשרת הבניה והשימוש במערכת בינה מלאכותית			

שילובים של טכנולוגיות מגבירות-פרטיות

עקרונות הפעולה המגוונים של טכנולוגיות מגבירות-פרטיות, והשוני בין הטכנולוגיות האפשריות למימוש כל עקרון פעולה, מאפשרות שילובים רבים ביניהן להשגת רמה גבוהה של הגנת מידע אישי ויצירת מעטפת הגנת מידע רב-שכבתית, גמישה ומבוססת הקשר. הפרק הבא של המסמך מציע דוגמאות ליישום טכנולוגיות מגבירות-פרטיות במערכות בינה מלאכותית. בדוגמאות אלו ניתן לזהות מספר שילובים נפוצים של טכנולוגיות המשלימות זו את זו. שילובים אלו מופעלים כחלק מתהליך משולב בהתאם לסוגי מידע אישי ולמאפייני העיבוד של מידע זה. שילובים אלו נבנים לרוב באופן מודולרי בהתאם למאפייני המידע, רמת הרגישות, הקשרים רגולטוריים וסוגי העיבוד הצפויים. בין דוגמאות אלו:

1. **פרטיות דיפרנציאלית + סביבת ביצוע מהימנה (Differential Privacy + Trusted Execution Environment):** בפרטיות דיפרנציאלית נוסף רעש אקראי למידע, כדי להבטיח שלא ניתן להסיק מידע על פרט בודד מתוך התוצאה. סביבת ביצוע מהימנה מספקת "בועה" מבודדת בתוך התוכנה או החומרה, שמונעת מגורמים בלתי מורשים –

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

כולל מערכת ההפעלה – לגשת למידע או לקוד שרץ בתוך הסביבה. כתוצאה מהשילוב בין שתי הטכנולוגיות, הוספת רעש אקראי נעשית בתוך סביבת הביצוע המהימנה, כך שלא רק הנתונים הרגישים, אלא גם פעולת הוספת הרעש האקראי עצמה מוגנת. החישוב מבוצע כולו בתוך סביבה מבודדת, והתוצאה המופקת מחוץ לסביבה היא כבר תוצר של מנגנון פרטיות דיפרנציאלית.

שילוב זה מאפשר הפחתת סיכון לזליגת מידע במהלך תהליך החישוב או הלמידה, ומספק הגנה כפולה גם מפני חשיפת נתונים, וגם מפני שיבוש או התקפות תוך כדי הוספת הרעש.

דוגמה לשימוש המאפשר הפקת מידע סטטיסטי שימושי תוך שמירה על הפרטיות: נניח שמשרד הבריאות רוצה לנתח מגמות צריכת תרופות באוכלוסייה, אך הנתונים מגיעים מקופות חולים שונות ואינם יכולים להישלח בפורמט גלוי. כל קופת חולים מבצעת חישוב בתוך סביבת הביצוע המהימנה, מוסיפה לנתונים רעש (פרטיות דיפרנציאלית), ושולחת רק את הסיכום. כך ניתן לאחד את המידע באופן מאובטח ולבצע הערכה סטטיסטית מבלי לחשוף מידע אישי של אף מטופל.

2. מידע סינתטי + הצלבת מערכי מידע (Synthetic Data + Private Set Intersection)

מידע סינתטי הוא מידע שנוצר באופן מלאכותי אך מדמה את מבנה והתפלגות הנתונים האמיתיים, מבלי להיות שייך לאף אדם מזוהה. הצלבת מערכי מידע מאפשרת לשני צדדים או יותר לבדוק אילו ערכים משותפים יש בין מאגריהם – למשל, מספרי תעודת זהות – מבלי לחשוף את יתר הערכים שאינם חופפים. השילוב של הטכנולוגיות מאפשר לזהות רק את התיקים הרלוונטיים או החופפים בין גופים שונים, ולאחר מכן להמיר את המידע האמיתי במידע סינתטי עבורם – כך שמודל בינה מלאכותית יאומן רק על תצפיות שאושרו לשימוש, וגם הן אינן חושפות פרטים מזהים של אנשים אמיתיים.

שילוב זה מאפשר מניעת חשיפת זהויות הן בשלב זיהוי התיקים המשותפים והן בשלב השימוש במידע, כלומר גמישות רבה בביצוע ניתוחים על אוכלוסיות מסוימות מבלי לחשוף מידע אישי, לצד האפשרות לשיתוף פעולה בין גופים ללא צורך בהעברת מאגרי מידע.

דוגמה: רשות מפקחת בתחום הפיננסי מעוניינת לבדוק האם קיימות קבוצות סיכון מסוימות שמקבלות הלוואות ממספר בנקים. כל בנק משתתף במנגנון הצלבת מערכי מידע כדי לחשוף רק את הלקוחות המשותפים. לאחר מכן, מייצרים מידע סינתטי על אותם לקוחות כדי לאמן מודל ניתוח סיכונים – מבלי לחשוף או לשתף מידע אישי אמיתי.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

דוגמאות ליישום טכנולוגיות מגבירות-פרטיות במערכות בינה מלאכותית

ריכוז דוגמאות לפי מדינות ותחומים

ארגון	מדינה	מידע	יישום	טכנולוגיה	תחום	שנה	מקור
Statistics Canada[20]	קנדה	מידע צרכני	פיתוח מודל לסיווג טקסט	הצפנה הומומורפית	**כת**	2021	קישור
Secretarium[20]	דנמרק	מידע פיננסי	שיתוף מידע פיננסי ללא חשיפתו בין משתתפים	חישוב רב-משתתפים, סביבת ביצוע מהימנה	**כת**	2023	קישור
Statice[21]	גרמניה	מידע פיננסי	שימוש במידע סינתטי לאימון מודל	מידע סינתטי	**קא**	2023	קישור
Statistics Netherlands[20]	הולנד	בריאות	פיתוח מודל לחיזוי סיכון קרדיווסקולרי מנתונים מבוזרים	חישוב רב-משתתפים, למידה מבוזרת, הצפנה הומומורפית	**כת**	2020	קישור
Frontier Development Lab / Intel[21]	ארה"ב	בריאות	פיתוח מודל לניתוח הקשר בין חשיפה לקרינה בחלל וסרטן	למידה מבוזרת	**קא**	2021	קישור
United Nations Economic Commission for Europe[20]	או"ם	בריאות	פיתוח מודל על בסיס מידע על אורח חיים הנאסף ממכשירים ניידים	למידה מבוזרת, פרטיות דיפרנציאלית, הצפנה הומומורפית	**קא**	2023	קישור
Twitter and OpenMined[20]	ארה"ב	מידע צרכני	מחקר על מידע משתמשים ללא חשיפה למידע עצמו	למידה מבוזרת, פרטיות דיפרנציאלית, חישוב רב-משתתפים	**קא** **פמ**	2023	קישור

מקרא: **קא:** קלט אימון **קמ:** קלט משתמש **ממ:** מודל מאומן **פמ:** פלט המערכת **כת:** כלל התחומים

[20] הפניה מאתר האו"ם - UN GWG Task Team on Privacy Preserving Case study repository Techniques Case Study Repository

[21] הפניה מאתר ICO - Repository of Privacy Enhancing Technologies (PETs) Use Cases

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

פירוט לגבי הדוגמאות מתוך מקורות המידע הרלוונטיים

יש לציין ביחס לכלל הדוגמאות כי הפירוט נסמך על התיאור שמופיע באתרי האו"ם, ICO ובמקורות הרלוונטיים שמהם לקוחות הדוגמאות ואין בתיאור זה כדי לחוות דעה או המלצה על אופן ההגנה על המידע האישי.

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
Statistics Canada	קנדה	מידע צרכני	פיתוח מודל לסיווג טקסט	הצפנה הומומורפית	2021	קישור

כללי

לשכת הסטטיסטיקה של קנדה ביצעה הוכחת היתכנות לאימון מודל למידת מכונה לסיווג טקסט בענן, תוך הבטחת הגנה על פרטיות המידע באמצעות הצפנה הומומורפית.

תהליך

1. נתוני הקלט לאימון מערכת בינה מלאכותית הוצפנו באמצעות הצפנה הומומורפית.
2. מערכת בינה מלאכותית פותחה בשרת מרוחק ואומנה על הנתונים המוצפנים שהועברו אליה, כך שנתוני המודל (משקלים של הרשת העצבית) היו מוצפנים לאורך התהליך.
3. השימוש במערכת הבינה המלאכותית נעשה על קלט משתמש שהיה מוצפן גם הוא.

תוצאות

1. נתוני הקלט ומידע משתמש שהוכנסו למערכת הבינה המלאכותית לא היו חשופים בשום שלב במהלך העיבוד.
2. נתוני מערכת הבינה המלאכותית (משקלים של הרשת העצבית) לא היו חשופים בשום שלב במהלך העיבוד.
3. דיוק התוצאות שהושגו, למרות אי דיוקים הנובעים מחישובי ההצפנה, היה דומה לדיוק באימון ושימוש ללא שימוש בהצפנה.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
Secretarium	דנמרק	מידע פיננסי	שיתוף מידע פיננסי ללא חשיפתו בין משתתפים	חישוב רב-משתתפים, סביבת ביצוע מהימנה	2023	קישור

כללי

הקונסורציום DANIE מאגד בנקים וספקי נתונים הפועלים יחד על פלטפורמה משותפת, אליה מועלים נתונים בנקאיים לצורך ניתוח. בין המטרות העיקריות של הקונסורציום: (1) שיפור איכות נתוני הלקוחות, (2) מניעת הלבנת הון, ו-(3) איתור הונאות. הפלטפורמה, שהושקה ב-2020, עושה שימוש בטכנולוגיות הצפנה מתקדמות ובסביבות ביצוע מהימנות, כך שהנתונים הנמצאים בעיבוד נותרים חסויים גם מפני המשתמשים עצמם.

DANIE נשענת על פתרונות הגנה על פרטיות המידע של חברת Secretarium, כאשר שתי היוזמות – גם Secretarium וגם DANIE – נולדו במסגרת תוכנית חממת החדשנות של Société Générale בלונדון. השתתפות ביוזמה זו מעניקה לארגונים יתרונות משמעותיים, כולל: עמידה ב-GDPR (רגולציית הגנת המידע של האיחוד האירופי) ומניעת קנסות; חיסכון במשאבים בזכות הפחתת הצורך בתיקוף ותיקון נתונים; ושיפור הביצועים והאפקטיביות של תהליכי ניתוח נתונים הודות למערכת עיבוד מרכזית ויעילה.

תהליך

1. המאגד משתמש בטכנולוגיות מחשוב מאובטח וקריפטוגרפיה כדי לאפשר שיתוף פעולה בין מוסדות פיננסיים ללא חשיפת נתונים רגישים.
2. מאז 2018, בנקים בינלאומיים משתמשים בטכנולוגיות הזמינות דרך הקונסורציום כדי לבצע עיבוד והתאמה של מיליוני רשומות.
3. במסגרת המאגד, השימוש בסביבת ביצוע מהימנה נעשה לעיבוד והתאמות של נתונים רגישים תוך שמירה על הפרטיות בין ארגונים ברשת מחשוב מאובטחת.

תוצאות

1. שמירה על שליטה בנתונים: כל צד שומר על שליטה מלאה בנתוניו, ללא חשיפת המידע למשתתפים אחרים.
2. הצפנה מלאה ואמינות מוכחת: הנתונים מוצפנים בכל עת, עם הוכחות אמינות ואפשרות לביקורת מלאה.
3. יעילות ויכולת הרחבה לקנה מידה גדול: המערכת תומכת במיליארדי עסקאות וכוללת ממשק משתמש ידידותי ויכולת פתרון שגיאות בצורה יעילה לשיפור איכות הנתונים.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
Statice	גרמניה	מידע פיננסי	שימוש במידע סינתטי לאימון מודל	מידע סינתטי	2023	קישור

כללי

חברת שירותי הביטוח הגרמנית Provinzial שיתפה פעולה עם חברת שירותי פרטיות הנתונים Statice, והשתמשה במידע סינתטי כדי לאמן מודלים של למידת מכונה במטרה לשפר את יכולות האנליטיקה החזויה שלהם (במיוחד מנוע ההמלצות "ההצעה הבאה הטובה ביותר"). הפעילות הביאה לחיסכון של למעלה משלושה חודשים בהערכת סיכוני פרטיות המידע שנמנעו עקב השימוש במידע הסינתטי.

תהליך

1. הפקת מידע סינתטי מתוך מאגרי המידע הקיימים של חברת הביטוח, תוך שמירה על המאפיינים הסטטיסטיים של הנתונים המקוריים.
2. אימון מודל אנליטיקה חזויה על המידע הסינתטי שנוצר, במטרה לזהות תבניות ולבצע תחזיות מדויקות.
3. השוואת ביצועי המודל שאומן על מידע סינתטי לביצועי מודל שאומן על נתונים מקוריים, לצורך הערכת האפקטיביות.

תוצאות

1. המידע הסינתטי התאים למאפיינים הסטטיסטיים של הנתונים המקוריים, ואפשר לאמן מודלים מבלי לחשוף מידע אישי רגיש.
2. המודל שאומן על המידע הסינתטי הציג ביצועים דומים למודל שאומן על נתונים מקוריים, מה שמעיד על איכות גבוהה של המידע הסינתטי.
3. השימוש במידע סינתטי אפשר ל-Provinzial לשפר את תהליכי האנליטיקה החזויה, תוך עמידה בדרישות הרגולציה.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
Statistics Netherlands	הולנד	בריאות	פיתוח מודל לחיזוי סיכון קרדיווסקולרי מנתונים מבוזרים	חישוב רב-משתתפים, למידה מבוזרת, הצפנה הומומורפית	2020	קישור

כללי

הלשכה המרכזית לסטטיסטיקה של הולנד (Statistics Netherlands - CBS) שיתפה פעולה מספר ארגונים בפרויקט CARRIER (קיצור של Coronary ARtery disease: Risk estimations and Interventions for prevention and EaRly detection). מטרת הפרויקט הייתה לפתח מודלים לחיזוי סיכון למחלות לב וכלי דם, תוך שמירה על פרטיות המידע הרפואי הרגיש ששימש לצורך כך.

תהליך

1. איסוף נתונים: נעשה שימוש במקורות נתונים שונים – נתוני טיפול ראשוני (מרפאות), נתוני טיפול שניוני (בתי חולים) ומידע חברתי-כלכלי.
2. טכנולוגיות מגבירות-פרטיות: כדי להבטיח שמירה על פרטיות המידע, נעשה שימוש בשיטות מתקדמות כגון חישוב רב-משתתפים, הצפנה הומומורפית, ולמידה מבוזרת. שיטות אלו אפשרו לנתח את הנתונים מבלי לחשוף מידע אישי מזהה.
3. פיתוח המודל: אלגוריתמים של למידת מכונה יושמו על המידע המוצפן לצורך בניית מודלים מדויקים לחיזוי סיכון קרדיווסקולרי.

תוצאות

1. שילוב נתונים מבוזרים: הפרויקט הצליח לשלב מקורות נתונים ממספר ארגונים שונים מבלי לחשוף את המידע האישי, ובכך הדגים את היכולת לשיתוף פעולה בין גופים תוך הגברת ההגנה על הפרטיות.
2. דיוק תחזיות: המודלים שפותחו הציגו רמת דיוק גבוהה, דומה לזו של מודלים שאומנו על מידע גלוי, והראו כי השימוש בטכנולוגיות מגבירות-פרטיות מאפשר להשיג תובנות איכותיות ללא פגיעה בסודיות הנתונים.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
Frontier Development Lab / Intel	ארה"ב	בריאות	פיתוח מודל לניתוח הקשר בין חשיפה לקרינה בחלל וסרטן	למידה מבוזרת	2021	קישור

כללי

חוקרים ממעבדת Frontier Development Lab (FDL) בשיתוף עם מנטורים מחברת Intel ערכו מחקר חדשני במטרה להבין טוב יותר את ההשפעות הפיזיולוגיות של חשיפה לקרינה על אסטרונאוטים. במחקר שבחן את הקשר בין חשיפה לקרינה בחלל לבין התפתחות מחלות סרטן, השתמשו החוקרים בגישה של למידה מבוזרת כדי לאפשר גישה לנתונים רגישים ומוגנים על אסטרונאוטים, מבלי לחשוף את המידע עצמו.

היתרון המרכזי של הגישה היה הפחתת עלויות משמעותית: בדרך מסורתית, הגישה לנתונים רפואיים כאלו מחייבת השקעה גדולה בתשתיות אבטחת מידע ובתהליכים בירוקרטיים. השימוש בלמידה מבוזרת אפשר לחסוך משאבים מבלי להתפשר על רמה גבוהה של פרטיות המידע.

תהליך

1. שימוש בלמידה מבוזרת: בפרויקט נעשה שימוש במסגרת OpenFL של אינטל, שאפשרה למודלים להתאמן על נתונים מקומיים מבלי להעביר נתונים רגישים בין מוסדות.
2. שילוב נתונים ממקורות שונים: הנתונים הגיעו מנאס"א, מאיו קליניק ומעבדת הגנים של נאס"א, כאשר כל הנתונים נשמרו במקומם תוך הגנה על המידע האישי.
3. אימון ואיחוד המודלים: מודלים אומנו ואוחדו כך שהנתונים האישיים נותרו חסויים לאורך כל שלבי הלמידה והעיבוד.

תוצאות

1. שמירה על פרטיות המשתתפים: הנתונים הרפואיים לא עזבו את המוסדות שהחזיקו בהם, ובכך נשמרה פרטיותם של אסטרונאוטים ומטופלים.
2. הקלה משמעותית על חסמים משפטיים ואתיים: הגישה סייעה להתגבר על אתגרים משפטיים ואתיים משמעותיים, דבר שאפשר שיתוף פעולה רחב יותר בין מוסדות ציבוריים ופרטיים.
3. אימון מוצלח של מודלים מדויקים: הפרויקט הדגים שניתן לקדם מחקר רפואי מתקדם תוך שימוש בטכנולוגיות מגבירות-פרטיות, מבלי לוותר על רמה גבוהה של הגנה על הפרטיות.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
United Nations Economic Commission for Europe	או״ם	בריאות	פיתוח מודל על בסיס מידע על אורח חיים הנאסף ממכשירים ניידים	למידה מבוזרת, פרטיות דיפרנציאלית, הצפנה הומומורפית	2023	קישור

כללי

נציבות האו״ם לכלכלה באירופה (UNECE) יזמה פרויקט ניסיוני לבחינת יישום של למידה מבוזרת, במטרה לאפשר שיתופי פעולה בין לשכות סטטיסטיקה לאומיות במדינות שונות מבלי לחשוף נתונים אישיים. הפרויקט פעל על אוסף נתונים פומבי של פעילות אנושית של קריאות מד התאוצה והגירוסקופ ממכשירים חכמים. הנתונים חולקו לארבע קבוצות משנה, אחת לכל לשכת סטטיסטיקה שהשתתפה בניסוי.

תהליך

1. בוצע אימון של מודל בינה מלאכותית באמצעות למידה מבוזרת על פני ארבע מאגרי המידע בלשכות הסטטיסטיקה.
2. נבנתה סביבה סימולטיבית על ידי שימוש בכלים וספריות של קוד פתוח במטרה לזהות ולסווג פעילויות אנושיות לקטגוריות מרובות, בהתבסס על נתוני מד תאוצה שנאספו ממכשירים חכמים ולבישים.
3. הסביבה הסימולטיבית שימשה לבדיקה של רמת ההגנה על הפרטיות וגם של תוצאות האימון.

תוצאות

1. הגנה על פרטיות: הפרויקט הצליח להמחיש שניתן לאמן מודלי בינה מלאכותית מדויקים מבלי להעביר נתונים אישיים.
2. היתכנות טכנית: הפרויקט הדגים את ההיתכנות של שימוש בטכנולוגיות למידה מבוזרת בעולם הסטטיסטיקה הרשמית.
3. תיאום ואישורים: הפרויקט הראה שבתרחישים אמיתיים יידרשו תיאום והסדרים בין הגופים המשתתפים.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa

ארגון	מדינה	תחום	נושא	טכנולוגיה	שנה	מקור
Twitter and OpenMined	ארה"ב	מידע צרכני	מחקר על מידע משתמשים ללא חשיפה למידע עצמו	למידה מבוזרת, פרטיות דיפרנציאלית, חישוב רב-משתתפים	2023	קישור

כללי

חברות טוויטר (Twitter) ו-OpenMined שיתפו פעולה כדי לאפשר עבודה של חוקרים חיצוניים על מידע פנימי – גם כשלא ניתן לפרסם את הנתונים עצמם. הפרויקט התמקד בבניית סביבה מאובטחת המאפשרת לחוקרים לבצע ניתוחים על נכסים דיגיטליים שלא שוחררו, תוך הגנה על פרטיות המשתמשים. המטרה המרכזית של המחקר הייתה לאפשר מחקר מדעי איכותי, מבלי לחשוף מידע אישי רגיש.

תהליך

1. בניית סביבה בטוחה ומבודדת: טוויטר סיפקה תשתית הרצה מבוססת סביבה טכנולוגית שמבטיחה שאף גורם (כולל ספקי הענן) לא יוכל לצפות בנתונים שמנותחים.
2. שימוש בקוד פתוח: החוקרים כתבו את הקוד לניתוח הנתונים מחוץ לסביבה, ולאחר בדיקה של טוויטר, הקוד אושר והורץ בתוך הסביבה המאובטחת.
3. פלט של תוצאות העיבוד: רק תוצאות מעובדות וסטטיסטיות אושרו לצאת מהסביבה המאובטחת, לאחר תהליך סינון שנועד לוודא שאין חשיפה של מידע אישי.

תוצאות

1. שמירה על פרטיות המשתמשים: לא היה צורך לשחרר את המידע המקורי או לחשוף אותו בפני החוקרים. הגישה לנתונים הייתה תחת בקרה הדוקה ומוגנת באמצעות טכנולוגיות שמונעות זליגת מידע.
2. שחזור מחקרים באופן אתי: הפרויקט אפשר לחוקרים לשחזר מחקרים על בסיס אותם נתונים מקוריים, בהתאם לעקרונות של שקיפות מדעית, מבלי להפר את פרטיות המשתמשים.
3. הדגמת יכולת לפיקוח חיצוני: המנגנון מאפשר ביצוע ביקורת חיצונית על מערכות בינה מלאכותית, גם כשיש מגבלות רגולטוריות או חוזיות על פרסום פומבי של הנתונים.

ppa@justice.gov.il | 073-3928529 | *3103
קרית הממשלה, ת.ד. 33503, תל-אביב 6133401 | www.go.gov.il/ppa