INSS
המכון למחקרי ביטחון לאומי
THE INSTITUTE FOR NATIONAL SECURITY STUDIES
אוניברסיטת תל אביב
TEL AVIV UNIVERSITY
נייר מדיניות

להאיץ או לבלום? "דילמת האסיר" של הבינה המלאכותית והמלצות לישראל

הדס לורבר | 23 בספטמבר 2026

עיקרי הדברים

בימים האחרונים מתנהל בארצות הברית ויכוח חריג בעוצמתו סביב קצב התפתחות הבינה המלאכותית. הטריגר המרכזי הוא מאמרו של מנכ"ל אנתרופיק, דריו אמודאי, "We Must pace the Frontier", שבו הוא קורא לווסת את קצב השיפור ביכולותיהם של מודלי ה-AI המתקדמים ביותר כדי לאפשר למנגנוני הבטיחות, ההערכה והפיקוח להדביק את קצב ההתפתחות הטכנולוגית. קריאתו זכתה לתמיכה מצד מובילים נוספים בתעשייה, בהם סם אלטמן, אילון מאסק ודמיס הסביס. מנגד, הנשיא טראמפ דחה את האפשרות שארצות הברית תאט את המרוץ באופן שיסכן את יתרונה מול סין, הציב במרכז שיקול אסטרטגי אחר: "Whoever wins AI, wins", ואף הכריז על הקמת "AI Force" שיעסוק בנושא.

מאחורי המחלוקת מסתתר פרדוקס עמוק יותר. גם חברות החוששות מקצב ההתפתחות אינן יכולות בהכרח להרשות לעצמן להאט באופן חד־צדדי כאשר מתחרותיהן – ובוודאי סין – ממשיכות להתקדם. לכן, הוויכוח על Pacing (ויסות הקצב) עשוי להתפתח בהדרגה מדיון בבטיחות לדיון על שליטה בגישה ליכולות: מי רשאי לפתח את המערכות המתקדמות ביותר, מי יקבל גישה לשבבים, לכוח מחשוב ולמודלים, באילו תנאים, ומי יידרש להוכיח כי הוא שותף "מהימן". הגבול בין בטיחות, בקרות ייצוא, מדיניות תעשייתית וביטחון לאומי הולך אפוא ומיטשטש.

עבור ישראל, המשמעות רחבה בהרבה משאלת הרגולציה. כמדינה התלויה במידה רבה בטכנולוגיה, בשבבים, בתשתיות ענן ובמודלים שמקורם מחוץ לישראל, אשר בה בעת מבקשת לשמר עצמאות פעולה ויתרון טכנולוגי וביטחוני, עליה להיערך לעולם שבו גישה ל-AI מתקדם היא גם מקור לעוצמה ולחופש פעולה לאומי. האתגר אינו לבחור בין "האצה" ל"האטה", אלא לבנות כבר עתה את הנכסים – כוח מחשוב, אנרגיה, הון אנושי, יכולות מחקר, שרשראות אספקה ושותפויות אסטרטגיות – שיבטיחו לישראל מקום סביב השולחן שבו ייקבעו כללי המשחק. במובן זה, הדיון אינו עוד רק על "ריבונות דיגיטלית", אלא על הריבונות עצמה: יכולתה של ישראל לפתח, להפעיל ולהבטיח גישה רציפה ליכולות AI קריטיות, ובמקביל להשתלב במעגל המדינות המעצבות את כללי הגישה אליהן.

****

במשך רוב העשור האחרון התאפיינה תחרות ה-AI במרוץ כמעט ליניארי: חברות ביקשו להגדיל מודלים, להרחיב יכולות מחשוב (Compute), לגייס חוקרים ולשחרר יכולות חדשות במהירות. החשש העיקרי של כל חברה היה שמתחרה תגיע ראשונה לפריצת הדרך הבאה.

עם זאת, האירועים המתרחשים בספטמבר 2026 מערערים את ההיגיון הזה: במאמר שפרסם מנכ"ל אנתרופיק, דריו אמודאי, הוא אינו חוזר על הטיעון המסורתי שלפיו AI הוא טכנולוגיה מסוכנת ולכן יש לעצור את פיתוחה; להפך, אמודאי מדגיש את הפוטנציאל העצום של AI להאצת המחקר המדעי, לריפוי מחלות, לצמיחה כלכלית ולהרחבת היכולות האנושיות. אולם, לטענתו, בחודשים האחרונים השתנה קצב ההתפתחות באופן המחייב חשיבה מחודשת.

שני גורמים עומדים במרכז טיעונו:

הראשון הוא התקדמות ביכולת של מערכות AI לסייע בפיתוח הדור הבא של מערכות הבינה המלאכותית – תהליך שהוא מתאר כחלק מתחילתו של "שיפור עצמי רקורסיבי" (Recursive self-improvement). אם מערכות AI מסוגלות לכתוב קוד, לבצע ניסויים, לייעל תהליכי אימון ולסייע למחקר AI עצמו, המרוץ אינו מתקדם עוד רק בקצב של היכולת האנושית לפתח מודלים, משום שהטכנולוגיה מתחילה לשמש מכפיל כוח לפיתוח הטכנולוגיה הבאה. אמודאי חושש כי מנגנון זה עלול להביא לכך שקצב השיפור ביכולות יעלה על הקצב שבו ניתן להבין אותן, להעריכן ולשלוט בהן.

הגורם השני הוא התנהגות בלתי צפויה של מערכות סוכנים אוטונומיים. אמודאי מצביע במיוחד על אירוע ניסויי שבו נחיל של סוכני AI ביצע פעולות סייבר שלא נדרשו כחלק מהמשימה וניסה להשפיע על מנגנון ההערכה שלו. בעיניו, המשמעות אינה שהאירוע עצמו גרם נזק אסטרטגי, אלא שהרחבה משמעותית של יכולותיהן של מערכות דומות עלולה להפוך התנהגות בלתי מתואמת מסוג זה למסוכנת הרבה יותר. מכאן מגיעה מסקנתו החריגה: "עלינו להאט את הקצב שבו אנו משפרים את יכולותיהם של מודלי הבינה המלאכותית" ("We must slow the pace at which we improve the capabilities of AI models").

אמודאי מדגיש כי pacing (ויסות) אינו pause (עצירה). הוא אינו מציע להפסיק לאמן מודלים או לעצור את ההתקדמות המדעית, אלא ליצור מרווח גדול יותר בין התקדמות היכולות לבין פריסתן, כך שניתן יהיה לבצע הערכות תאימות ובקרות על התאמת התנהגות המודלים למטרות ולערכים אנושיים (Alignment), בדיקות תקיפה יזומות (Red teaming) וכן בדיקות בטיחות, לפני שמערכות בעלות יכולות חדשות מופצות בקנה מידה רחב.

מהאטה עצמית למשטר פיקוח חדש

חשיבות מאמרו של אמודאי אינה רק באזהרה, אלא במנגנון שהוא מתווה – מודל בן שלוש שכבות:

השכבה הראשונה – Embedded Evaluators: חברות Frontier AI (AI מתקדם) יעניקו לצוותי הערכה חיצוניים גישה רציפה למערכות ולתהליכי הפיתוח שלהן, ברמה הדומה לזו של עובדים פנימיים. אנתרופיק התחייבה באופן חד־צדדי ליישם עיקרון זה. הרעיון מזכיר במידה מסוימת פיקוח פיננסי: הרגולטור אינו מסתפק בדוח שמגישה החברה לאחר מעשה, אלא מקבל נראות לתהליכים עצמם.

השכבה השנייה – Democratic Coordination: חברות AI במדינות דמוקרטיות יתאמו סטנדרטים משותפים לבטיחות ואף מגבלות מסוימות על התקדמות בלתי מבוקרת של יכולות. כאן כבר נדרשת מעורבות ממשלתית, בין היתר משום שתיאום בין חברות מתחרות מעלה שאלות משפטיות ותחרותיות.

השכבה השלישית – Global Coordination: ארצות הברית ובעלות בריתה ינסו להגיע להבנות גם עם מדינות יריבות, ובראשן סין, לגבי מגבלות מסוימות על יכולות AI מסוכנות במיוחד, תוך הכרה בכך שאימות הציות להסדרים כאלה יהיה קשה.

הצעה זו זכתה בתוך ימים ספורים לתמיכה יוצאת דופן מצד מתחריו של אמודאי. סם אלטמן הודיע כי הוא מסכים שיש צורך "To pace the frontier" – לווסת את הקצב – ואף התחייב כי OpenAI תאמץ את רעיון הגישה למעריכים עצמאיים. אילון מאסק הסתפק באמירה קצרה: "Dario is right". דמיס הסביס מ- Google DeepMind אמר כי הכיוון נכון, תוך קישורו להצעת DeepMind לגוף סטנדרטים תעשייתי. גם מיקרוסופט הצטרפה לדיון, אם כי בגישה זהירה יותר, והדגישה כי כל משטר הערכה חייב לשמור גם על פרטיות הלקוחות ועל שליטתם במידע העסקי שלהם.

נוצר אפוא מצב חריג: חלק ניכר מהמנהלים המובילים את תעשיית הייצור של מערכות ה-AI המתקדמות בעולם מסכימים לפחות על העיקרון כי קצב ההתפתחות מחייב מנגנוני בקרה חזקים יותר. אולם דווקא בנקודה זו מתנגשת שאלת הבטיחות עם שאלת הביטחון הלאומי.

וושינגטון — "Whoever wins AI, wins"

תגובת הנשיא טראמפ ממחישה את המתח שנוצר סביב הסוגייה: טראמפ לא דחה לחלוטין את אפשרות חסמי הבטיחות (guardrails), אך דחה את ההנחה שלפיה ארצות הברית יכולה להרשות לעצמה להאט את המרוץ. לדבריו, ארצות הברית מובילה כיום על סין והוא מעוניין לשמר את היתרון הזה: "Whoever wins AI, wins" (מי שמנצח ב-AI, זוכה). מדובר בהמשך ישיר לדוקטרינת ה-AI הרחבה של ממשלו – "Build faster, regulate less, beat China" (בנייה מהירה יותר, פחות רגולציה, הבסת סין) – לא רק במודלים אלא גם בדאטה סנטרים, אנרגיה, שבבים ותשתיות המאפשרים את פיתוחם.

דייוויד סאקס, שכיהן כצאר ה-AI בממשל טראמפ, הציג את טיעון הנגד החריף ביותר לחברות. אם אנתרופיק ו-OpenAIסבורות שהמודלים שלהן מתקדמים מהר מדי, לשיטתו, הן יכולות פשוט להאט בעצמן; אין הצדקה להשתמש בוושינגטון כדי לחייב את המתחרות לעשות כך גם הן.

בבסיסו של טיעון זה נמצאת גם ביקורת עמוקה יותר: חלק מהתומכים בהאצה בעמק הסיליקון חוששים ש- AI safety (בטיחות AI) הופך לכלי של לכידה רגולטורית. החברות הגדולות מחזיקות כבר בתשתיות מחשוב, מודלים, הון וכוח אדם; רגולציה יקרה המחייבת מערכי בקרה, התאמה ודיווח עשויה להיות נסבלת מבחינתן, אך להעלות משמעותית את חסמי הכניסה עבור מתחרות חדשות.

בקונגרס, יו"ר בית הנבחרים מייק ג׳ונסון הציג עמדה מתונה מעט יותר. הוא דחה מורטוריום מיידי, אך גם לא ביטל את הסיכונים. לדבריו, הקונגרס צריך להיזהר מלקבוע כללים לפני שהוא מבין מהו הפתרון. הוא הציע לכנס את הנשיא, מנהיגי הקונגרס וראשי מעבדות ה- AI לניסיון לגבש גישה משותפת. מנגד, הדמוקרטים מתכנסים בתחושת דחיפות אך עדיין ללא דוקטרינה מוסכמת. חלק מהמחוקקים תומכים בצעדים מרחיקי לכת סביב בינת־על ואחרים מעדיפים בקרה, דיווח, ועדות ייעודיות ורגולציה ממוקדת של Frontier models. ההערכה היא שהפער הפוליטי הופך גיבוש רגולציית AI פדרלית כוללת לקשה מאוד בעתיד הקרוב.

הפרדוקס הגיאו-אסטרטגי: אפשר להאט אם היריב אינו מאט? "דילמת האסיר" הטכנולוגית

בלב המחלוקת נמצא פרדוקס שאינו ניתן ליישוב באמצעות רגולציה פנימית בלבד. אם AI הוא רק מוצר מסחרי, מדינה יכולה לקבוע את רמת הסיכון שהיא מוכנה לקבל ולהשית רגולציה על החברות בהתאם. אבל אם AI הוא גם טכנולוגיית יסוד של עוצמה לאומית, ההחלטה להאט הופכת אסטרטגית.

ארצות הברית וסין מתחרות כיום לא על המודל הטוב ביותר בלבד, אלא על כל שכבות ה-AI stack: שבבים, ציוד לייצור שבבים, hyperscale data centers , חשמל, cloud, מודלים, טאלנט, סטנדרטים ושרשראות אספקה. במקביל, הפער בין יכולות המודלים האמריקאיים והסיניים הולך ומצטמצם, למרות מגבלות הייצוא האמריקאיות על שבבים וטכנולוגיות מתקדמות. מכאן, שהבעיה שמציג אמודאי היא במידה רבה בעיית פעולה קולקטיבית: גם אם החברות האמריקאיות המובילות יאטו, חברות אחרות, או סין, תמשכנה להתקדם בקצב מלא. כל מי שמאט מסתכן באובדן היתרון.

אמודאי עצמו מכיר בדילמה הזו. לדבריו, הימנעות מפיתוח הטכנולוגיה עלולה למנוע מהאנושות את יתרונותיה – או להעביר את היכולת לידי משטרים אוטוריטריים. לכן הוא אינו מציע לעצור את המרוץ אלא להפוך אותו מ- Race to the bottom ל-Race to the top, שבמסגרתו בטיחות עצמה הופכת למדד של תחרות. גישה זו מסבירה מדוע השלב השלישי בתוכניתו הוא תיאום בינלאומי. ואכן, דיווחים בוושינגטון מצביעים על כך שבטיחות AI צפויה להיות על סדר היום במפגש המתוכנן בין הנשיאים טראמפ ושי ג'ינפינג ב-24 בספטמבר.

זהו אולי השינוי האסטרטגי המשמעותי ביותר בוויכוח: AI Safety הופך בהדרגה מסוגיית רגולציה טכנולוגית לסוגייה של בקרת נשק, אמון אסטרטגי ויחסים בין מעצמות.

הוויכוח מלמד בנוסף כי השאלה אינה יכולה להישאר מוגבלת ליצרניות המודלים. מערכת AI מתקדמת תלויה בשרשרת שלמה: שבבים מתקדמים, cloud infrastructure, data centers, תקשורת, אנרגיה, מודלים, נתונים ומשתמשים. ככל שמערכות AI הופכות agentic ואוטונומיות יותר, עולה גם חשיבותם של מי שיכול לגשת לכוח המחשוב, מאיזו מדינה, באיזה היקף ובאילו תנאים.

מכאן עשוי להתפתח מודל חדש של Trusted AI ecosystem: לא רק מודלים "בטוחים", אלא תשתיות, ספקי ענן, מפעילי data centers, שרשראות אספקה ולקוחות העומדים בסטנדרטים משותפים של Cyber security, Access control, Auditability Resilience ו-Customer screening. במובן זה, הוויכוח על אודות Frontier safety מצטרף למגמה רחבה יותר במדיניות האמריקאית: שבבים, כוח מחשוב ותשתיות AI נתפסים יותר ויותר כנכסים אסטרטגיים ולא כמשאבים מסחריים רגילים. הגבול בין בקרת ייצוא ומשילות AI (Export controls ו-AI governance) לבין ביטחון לאומי הולך ומיטשטש.

מאחורי הוויכוח על הקצב: מי יקבע את כללי הגישה ל-AI?

מכאן עולה גם אפשרות מורכבת יותר: אם לא ניתן להאט את המרוץ באמצעות ריסון עצמי של החברות, הקצב עצמו עלול להתפתח בהדרגה ממנגנון בטיחות למנגנון המסדיר גישה ליכולות. השאלה תהיה אז פחות האם ה-AI מתקדם מהר מדי,

ויותר מי רשאי לפתח את המערכות המתקדמות ביותר, מי יקבל גישה לשבבים, לכוח מחשוב ולמודלים, באילו תנאים, ואילו מדינות וחברות יידרשו להוכיח כי הן "אחראיות" או "מהימנות" כדי להישאר בתוך המעגל. כללים כאלה אינם חייבים להתגבש במסגרת אמנה בינלאומית פורמלית; הם עשויים להיווצר באמצעות שילוב של בקרות ייצוא אמריקאיות, תנאי גישה של ספקיות ענן ומודלים, סטנדרטים תעשייתיים והסדרים בין מדינות בעלות תפיסות דומות.

במובן זה קיימת אנלוגיה, גם אם בלתי מושלמת, למשטרי אי-הפצה דוגמת ה-NPT; בעלי היכולות המתקדמות ביותר ממשיכים להחזיק בהן ולפתחן, בעוד שמדינות אחרות נדרשות לעמוד בתנאים כדי לקבל גישה לטכנולוגיות הרגישות. אם יתפתח מבנה דומה בתחום ה-AI החלוקה המשמעותית במערכת הבינלאומית לא תהיה רק בין מדינות שמפתחות AI לבין מדינות שאינן מפתחות אותו, אלא בין מי שנמצאות בתוך המעגל שבו נקבעים כללי הגישה לבין מי שנדרשות לציית לכללים שנקבעו בלעדיהן.

הבעיה האסטרטגית והאתגר לישראל

עבור ישראל, השאלה המעשית אינה אם להצטרף למחנה של "מאיצים" או למחנה של "מאטים". האתגר הוא לבנות במהירות מספקת את הנכסים שיאפשרו לה להיות בעלת השפעה משמעותית כאשר ייקבעו כללי המשחק. בעולם שבו גם החברות וגם המדינות צפויות לעצב מנגנוני גישה וריסון, מי שאין בידיו יכולות, תשתיות, שוק, מומחיות ושותפויות לא בהכרח יזכה להכרה על איפוקו; הוא עלול פשוט למצוא עצמו כפוף לפשרות שנקבעו בין אחרים. מכאן נובעת דחיפותה של השקעה ישראלית בתשתיות מחשוב, אנרגיה, מודלים, הון אנושי ושרשראות אספקה, לצד העמקת השותפות עם ארצות הברית ובעלות בריתה והשתלבות בדיונים שבהם נקבעים סטנדרטים, מנגנוני הערכה וכללי גישה. המטרה אינה להבטיח שישראל תמיד תסכים עם הכללים שייקבעו, אלא להבטיח שהיא תהיה בין המדינות שיש להן את היכולת להשפיע עליהם.

בהקשר זה, ככל שה-AI הופך לתשתית של ביטחון, כלכלה, מחקר, אנרגיה ותעשייה, מדובר למעשה בשאלה רחבה יותר של ריבונות: היכולת של מדינה לפעול, לפתח, לקבל החלטות ולהפעיל מערכות קריטיות גם כאשר תנאי הגישה לטכנולוגיה משתנים. ריבונות כזו אינה מחייבת אוטרקיה או עצמאות מלאה מכל ספק זר; היא מחייבת יכולת, חלופות, שותפויות ואחיזה בשרשראות הערך הקריטיות. בהקשר של AI, משמעות הדבר היא לא רק להחזיק מודלים מקומיים, אלא להבטיח גישה רציפה לכוח מחשוב, לאנרגיה, לשבבים, לענן, לטאלנט ולידע, ובמקביל לפתח יכולות מקומיות בתחומים שבהם תלות חיצונית עלולה להפוך לנקודת כשל אסטרטגית.

מכאן עולה גם ממד גיאו-פוליטי עמוק יותר. במערכת שבה היכולות המתקדמות ביותר מרוכזות במספר מצומצם של מדינות וחברות, כללי בטיחות אינם מתקיימים בחלל ריק: הם משתלבים במערך רחב יותר של בקרות ייצוא, סינון לקוחות, הגבלת גישה לכוח מחשוב, סטנדרטים של Trusted infrastructure והחלטות בדבר זהותם של השחקנים שנתפסים כמהימנים. עבור ישראל, הסיכון אינו בהכרח בכך שתיקבע כלפיה מגבלה מפורשת, אלא בכך שתמצא את עצמה בעתיד בעמדה שבה כללי הגישה כבר נקבעו בלעדיה, והיא נדרשת להתאים עצמה אליהם.

המעבר האפשרי ממשטר המתמקד בבטיחות מודלים למשטר רחב יותר של שליטה בגישה ליכולות AI נושא משמעויות מיוחדות עבור ישראל. ישראל אינה נמנית כיום עם המדינות המפתחות Frontier foundation models בקנה המידה של ארצות הברית או סין, אך היא מדינה שבה ל-AI משמעות אסטרטגית חריגה בשל צפיפות האקוסיסטם הטכנולוגי, מערכת הביטחון, תעשיות ה-Dual-use, הסייבר והמחקר האקדמי. מכאן נגזרות שלוש משמעויות מרכזיות.

ראשית, ביטחון לאומי: מערכת הביטחון הישראלית צפויה להיות משתמשת משמעותית במערכות Agentic AI. הסיכון אינו רק שמודל "יטעה", אלא שמערכת אוטונומית המחוברת לרשתות, למודיעין או למערכות מבצעיות תפעל באופן שלא נחזה מראש. לכן ישראל זקוקה ליכולת עצמאית של בקרות על מערכות בעלות משמעות ביטחונית.

שנית, עוצמה גיאו-פוליטית: בעולם שבו גישה לשבבים ,כוח מחשוב ,מודלים ותשתיות הופכת בהדרגה לחלק ממערכת היחסים בין מדינות, היכולת הטכנולוגית של ישראל היא גם נכס מדיני. ככל שישראל תחזיק ביותר יכולות עצמאיות ותהיה משולבת עמוק יותר באקוסיסטם האמריקאי, כך יגדל משקלה בדיונים שבהם ייקבעו כללי הגישה העתידיים.

שלישית, הזדמנות כלכלית: המעבר למערכות AI מורכבות ואוטונומיות יוצר שוק חדש בתחומים שבהם לישראל יתרונות קיימים, לרבות: Cybersecurity for AI, Model evaluation, Observability, Secure infrastructure, Identity and management access ו-Agent security /Red teaming**.** לכן, AI Safety יכול להפוך מדרישה רגולטורית לתעשיית ייצוא ישראלית.

המלצות למדיניות ישראלית: האצה אחראית ומבוקרת

ראשית, לא לבחור בין האצה לבין בטיחות. ישראל זקוקה למדיניות של "האצה מבוקרת" – Responsible acceleration –האצת אימוץ ופיתוח AI תוך יצירת חסמי בטיחות במקומות שהסיכון בהם משמעותי. ניסיון להעתיק רגולציה אמריקאית או אירופית באופן גורף יהיה שגוי; באותה מידה, גישת laissez-faire מלאה אינה מתאימה למדינה שבה AI משתלב במהירות במערכות ביטחוניות וקריטיות.

שנית, להקים יכולת לאומית להערכת מודלים מתקדמים. ישראל צריכה לבנות גוף או קונסורציום שיכלול ממשלה, מערכת ביטחון, אקדמיה ותעשייה, ויהיה מסוגל לבצע הערכות עצמאיות למודלים מתקדמים. ניתן לבחון שיתוף פעולה עם גופי הערכה אמריקאיים ובריטיים.

שלישית, לכונן דיאלוג אסטרטגי עם ארצות הברית בנושא AI safety/trusted compute – נושא זה צריך להיות חלק מהיחסים הטכנולוגיים בין המדינות לצד Pax Silica, ומטרת ישראל צריכה להיות ברורה: להבטיח כי היא נתפסת כ-Trusted AI partner ולא כמדינת צד שלישי, שהגישה שלה ליכולות אמריקאיות נבחנת בכל פעם מחדש.

רביעית, למפות תלות בשרשרת אספקת ה-AI – יש למפות GPUs, cloud, data centers, networking, models, energyורכיבים קריטיים אחרים ולזהות single points of failure. בעולם שבו AI הוא נכס אסטרטגי, עמידות טכנולוגית היא חלק מהחוסן הלאומי.

חמישית, להפוך AI safety ליתרון תעשייתי ישראלי – מטה ה-AI, משרד הכלכלה, רשות החדשנות, ומערכת הביטחון צריכים לעודד חברות ישראליות לפתח טכנולוגיות Secure ,AI cyber defense ,Agent security ,Evaluation Observability וinference-. אלה עשויים להפוך לשכבה משמעותית בכלכלת ה-AI הבאה.

שישית, להשתלב בעיצוב הסטנדרטים ולא רק לאמץ אותם. אם ארצות הברית, בריטניה ומדינות דמוקרטיות נוספות יפתחו סטנדרטים משותפים ל-Frontier AI, על ישראל להיות בחדר שבו הם נקבעים. מעמדה כחברה קטנה יחסית אך בעלת אקוסיסטם טכנולוגי וביטחוני מתקדם מאפשר לה לתרום, במיוחד בנושאי Cyber, Dual-use ו- Testing בסביבות מבצעיות.

סיכום: המרוץ אינו נעצר — הוא משנה צורה

הוויכוח הער בארצות הברית אינו עוד פרק בעימות הישן בין "חדשנות" ל"רגולציה". העובדה שמנהלי החברות המתקדמות ביותר קוראים בעצמם לוויסות מצביעה על שינוי משמעותי בהערכת הסיכון בתעשייה. אך תגובת ממשל טראמפ מצביעה על מגבלה חשובה לא פחות: אי אפשר לנהל את סיכוני ה-AI במנותק מהתחרות הבין־מעצמתית. מבחינת וושינגטון, האטה אמריקאית בעוד סין ממשיכה להתקדם עלולה להיות בפני עצמה סיכון לביטחון הלאומי. מבחינת החברות, המשך מרוץ בלתי מוגבל עלול ליצור סיכונים שהמערכת הפוליטית והטכנולוגית לא תספיק לנהל.

הפתרון שיתגבש ככל הנראה לא יהיה בחירה בינארית בין עצירה להאצה. סביר יותר שהוא יתבסס על שילוב של בקרות ייצוא, הערכות וסטנדרטים תעשייתיים (לרבות Evaluation, Trusted infrastructure, Export controls, Industry standards) ותיאום בין מדינות בעלות תפיסות דומות. הסיכוי ל"משטר בינלאומי" בנושא ה-AI נמוך, אך יש להמתין לפגישת טראמפ-שי הקרובה, שסביר שהנושא יעלה בה.

עבור ישראל, המסקנה היא כי יכולת AI אינה עוד סוגיה טכנולוגית בלבד אלא נכס של עוצמה לאומית. ישראל צריכה להאיץ את בניית יכולותיה לא רק כדי להשתמש ב-AI אלא כדי להבטיח שיהיו בידיה הנכסים, השותפויות והמנופים שיאפשרו לה להשפיע על כללי הגישה לעידן ה-AI. במובן זה, השאלה אינה עוד רק של ריבונות דיגיטלית גרידא — אלא של הריבונות עצמה.