

פיתוי הבינה המלאכותית הסינית:

אל לה לישראל להסתכן בשיכון מעמקים בעקבות DeepSeek

אריאל סובלמן, מיכאל גנקין | 18 בפברואר, 2025 | גיליון 1951

תחרות המעצמות לעליונות בבינה המלאכותית שועטת קדימה והמהלומה האחרונה של סין - השקת מודל הבינה המלאכותית DeepSeek - הביאה לטלטלה עזה ברחבי העולם. תעשיית הטכנולוגיה, בורסות המניות, אקדמיה וממשלות בוחנות האם פריצת הדרך ההנדסית הסינית תגרום גם לשינוי המאזן הגיאופוליטי. בצל קשיי תכנית הבינה הלאומית הישראלית, נבחנות יעילות המודל הסיני החדש ועלותו האמיתית מול זו שהתפרסמה ומדגיש כי אסור להסתנוור מהפיתוחים הסיניים ולהתפתות לקיצורי דרך טכנולוגיים על חשבון הקמת תשתית טכנולוגיה לאומית וביצור ריבונותה הדיגיטלית של ישראל.

קצב האירועים בתחום הבינה המלאכותית ממשיך להדהים ולהכות גלים ברחבי העולם. ההתפתחות המשמעותית האחרונה נרשמה עם השקתו של מודל בינה מלאכותית בשם R1 שיצרה חברת DeepSeek הסינית בסוף ינואר 2025, בצמוד להשבעת טראמפ לנשיא ארצות הברית וחגיגת ראש השנה הסיני. השיח התקשורתי המערבי על אודות המודל התמקד בעלות האיומן הנמוכה עד כדי להדהים אשר פורסם כי עמדה, לכאורה, על פחות משישה מיליון דולר, ביעילותו החישובית, הגדולה משמעותית ממודלים מערביים מתחרים, ובעלות השימוש המסחרי בו, הנמוכה בעשרות אחוזים ממחירי השוק הנוכחיים של החברות המתחרות דוגמת ChatGPT הפופולרי.¹ מיד עם השקתו עלתה טענה כי לפריצת דרך מדעית והנדסית כזו יש משמעויות הרבה מעבר לטכנולוגיה עצמה. לפי אומדני הביצועים של החברה עולה שסין, הנתונה למשטר קשוח של הגבלות ייצוא אמריקאיות, או עקפה אותן, או - מדהים אף יותר - הצליחה בכוחות עצמה להגיע לפריצת הדרך באמצעות כלי מחשוב מדורות ישנים שכבר יש ברשותה ואינם כפופים למשטר הפיקוח. עד לחשיפת DeepSeek R1, הנחת היסוד בעולם הטכנולוגיה הייתה שעל מנת לבצע את כמות החישובים העצומה הנדרשת למערכות בינה מלאכותית, אין ולא יהיה תחליף למעבדים המתקדמים ביותר, המיוצרים רובם כולם על ידי חברות במערב, בראשן חברת nVidia האמריקאית. התגובות הראשונות בעולם היו קיצוניות

¹ Kyle Wiggers, "DeepSeek Claims Its 'reasoning' Model Beats OpenAI's O1 on Certain Benchmarks," *TechCrunch* (blog), January 27, 2025, <https://techcrunch.com/2025/01/27/deepseek-claims-its-reasoning-model-beats-openais-o1-on-certain-benchmarks/>; DeepSeek-AI et al., "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning" (arXiv, January 22, 2025), <https://doi.org/10.48550/arXiv.2501.12948>.

והיו אף אלה שהשוו את ההפתעה לשיגור לוויין הספוטניק הסובייטי ב-1958 וכינו את המעמד "רגע ה-DeepSeek" או "רגע הספוטניק של הבינה המלאכותית".² ואכן, תגובת השווקים להכרזה הייתה מיידית ומניית ענקיות הטכנולוגיה ברחבי העולם רשמו ירידות חדות שגרמו להפסדים עצומים.

מאז ההשקה ניצת דיון ציבורי חובק עולם על אודות המשמעויות הטכנולוגיות וההשלכות הגיאו-פוליטיות בתחרות בין סין לארצות הברית על ההובלה הטכנולוגית. הדיון לא פסח על ישראל, הנתונה במשבר משל עצמה בתחום הבינה המלאכותית, כשדירוגה במדדים העולמיים נמצא בירידה מתמדת והתוכניות הלאומיות בתחום מתקשות לממש את יעודן לאפשר לתעשייה הישראלית להשיג מקום מוביל, כפי שהצליחה בגלים קודמים של טכנולוגיות מפציעות, תחום הסייבר למשל.

חמור מכך, מעמדה של ישראל נפגע קשות בשבוע האחרון לכהונתו של הנשיא ביידן, שהוציא צו המדרג את מדינות העולם לפי נכסיותן הטכנולוגית, מידת האמון בהן והתייצבותן לצד ארצות הברית בתחרות מול סין. בדרג א' יש 18 מדינות, אשר ארצות הברית מחשיבה כבעלות בריתה האמינות ביותר מחד גיסא, ושהן נתפסות גם כנכסיות ביותר לשיתוף פעולה סביב בינה מלאכותית וטכנולוגיות מתקדמות, מאידך גיסא. דרג ג' כולל את סין, רוסיה, איראן וצפון קוריאה, שחל איסור מוחלט על ייצוא טכנולוגיה לבינה מלאכותית אליהן. על פי הצו, ארצות הברית תשית הגבלות על ייצוא מעבדים מתקדמים המאפשרים פיתוחי בינה מלאכותית למדינות בדרג ב', והגבלות אלו כוללות - לראשונה - גם את ישראל. מבחינות מסוימות, ישראל סופגת נזק אגבי ניכר מהתחרות בין ארצות הברית לסין. אך זהו הסבר חלקי ומקל מכיוון שהסיבות להכללתה ברשימת המדינות המוגדרות רשמית כדרג ב' הן עמוקות יותר ונוגעות בהיבטים רחבים בהרבה, ובעצם מהוות הבעת אי-אמון אמריקאית ביכולתה של ישראל לפתח טכנולוגיית בינה המלאכותית מתקדמת, או באופן שאינו פוגע ביעדים האסטרטגיים של ארצות הברית. מן הסתם, זהו גם מעגל קסמים שלילי, מכיוון שהכללתה של ישראל בדרג ב' והגבלות הייצוא שיחולו עליה יביאו למצב שללא גישה מספקת לשבבים מתקדמים יהיה קשה אף יותר לפתח את תכנית הבינה המלאכותית ולמנוע את המשך הירידה בדירוגים העולמיים. הצו של הנשיא ביידן הביא את ישראל, לראשונה, להכיר בחומרת המצב והניע פעילות אינטנסיבית בממשלה, כנסת, תעשייה ומערכת הביטחון, שתכליתה למצוא דרך להתמודד עם האתגר, לנסות להשפיע על הנשיא הנכנס טראמפ לבטל את הנחיית קודמו (דבר שראש הממשלה נתניהו אכן עשה בביקורו בוויניגטון בתחילת פברואר. נותר עוד לראות האם הנשיא ייעתר בחיוב

Jack Burnham and Craig Singleton, "AI's Sputnik Moment: Chinese AI Model DeepSeek-R1² Reportedly Surpasses Leading U.S. AI Models," *FDD Policy Brief*, January 30, 2025, https://www.fdd.org/analysis/policy_briefs/2025/01/30/ais-sputnik-moment-chinese-ai-model-deepseek-r1-reportedly-surpasses-leading-u-s-ai-models/.

לבקשתו להחריג של ישראל מההגבלות), או, לחילופין, לגבש תכנית שתעלה את ישראל לדרג א'.

לנוכח ההקשר הרחב והלך הרוח בישראל אחרי הצו של בידן, קל להבין איך ההכרזה הסינית הציתה באחת תקווה גם בישראל שאם סין הצליחה לפתח מודל בינה מלאכותית מוביל בעלות זולה ובלי גישה למעבדים מתקדמים – אזי ייתכן שגם ישראל תוכל להצליח באותה השיטה. המעבר מפסימיות קיצונית בעקבות המהלך של בידן לאופטימיות אופורית אחרי ההכרזה הסינית שיקף, למעשה, את מטוטלת הרגשות הישראלית. לדוגמא, כבר ביום הכרזת הצו, איגוד התעשיינים בישראל יצא בקריאה דחופה לראש הממשלה לטפל במשבר, שיפגע אנושות בתעשיית הטכנולוגיה הישראלית. כשבוע לאחר מכן, המגמה התהפכה ומשרד החדשנות והמדע הזדרז להתייחס להשקת המודל R1 של DeepSeek בציינו כי "עבור ישראל, המודל הסיני מסמן אופק חיובי באשר לסיכוייה להיות מהמובילות בבינה מלאכותית בעולם גם ללא השקעות עתק, שהיוו עד עתה חסם כניסה משמעותי למועדון זה באופן עצמאי".³

המציאות, כמובן, מורכבת יותר ודורשת העמקה טכנית, ניתוח גיאואסטרטגי והערכת המשמעויות עבור ישראל. מאמר זה מתמקד בשלושה נושאים עיקריים אלה. הראשון, בחינה טכנית מעמיקה של המודל הסיני החדש, עוצמת החישוב שלו וניתוח עלותו האמיתית מול זו שהתפרסמה. לאחר מכן, נבחן האם פריצת הדרך הסינית אכן עשויה לערער את יסודות התחרות הטכנולוגית בין סין לארצות הברית והאם ניתן להסיק גם על יעילות משטר הפיקוח וההגבלות האמריקאי, שייתכן שמיצה את עצמו. כנגזרת של שני ממדים אלה, יסקרו המשמעויות לישראל ובכללן האם אכן השקת DeepSeek טומנת בחובה הזדמנות עבור ישראל, והאם בעקבותיה יש תוחלת ממשית שעשויה לאפשר לישראל להגיע לעמדת הובלה בבינה מלאכותית בלי להיזקק למעבדים המתקדמים ביותר, אשר כיום אינם זמינים לה עקב מיקומה בדרג ב' וההגבלות האמריקאיות.

עד כמה חזק ויעיל DeepSeek ומה היתה עלות פיתוחו האמיתית?

הסוגייה הראשונה היא האם יכולות Deep Seek ועוצמתו הן כפי שנטען על ידי החברה הסינית עצמה בפרסומיה, או פרסומים וניתוחים טכניים מערביים, שבעצם התבססו על נתוני החברה שפורסמו (ולא על בדיקות עצמאיות). ככל שחולפים הימים ומתבררים נתוני אמת כתוצאה מבדיקות עומק, התמונה מתבהרת מעט. בדיקות עצמאיות ראשונות מעידות כי ייתכן

³ אגף קו האופק, "השקת מודל R1 DeepSeek מערערת את פרדיגמת פיתוח ה-AI במערב ומאיימת לשנות את המאזן בתחרות הגאו-טכנולוגית בין ארה"ב לסין" משרד החדשנות, המדע והטכנולוגיה. (January 28, 2025),

שביצועי המודל אינם טובים כפי שנטען על ידי החברה בעת פרסומו.⁴ בחינה מדוקדקת של הפרטים הטכניים שפורסמו על המודל מגלה כי אומנם הארכיטקטורה של DeepSeek מייצרת ייעול עלויות אימון משמעותי, אבל היא לא עושה זאת בזכות הצגת רעיון חדש, אלא באמצעות שילוב יצירתי של רעיונות קיימים – כולם כאלה שמומשו בעבר על ידי חברות מערביות.⁵ ובייחס לטענה כי הצלחת DeepSeek מעידה כי משטר הגבלות הייצוא האמריקאי נכשל בהשגת מטרותיו (מניעת הובלה סינית בתחום השבבים והבינה המלאכותית), יש לזכור כי חברת DeepSeek עצמה חשפה כי המודל אומן באמצעות מעבדים שנרכשו לפני כניסת הגבלות הייצוא לתוקפן באוקטובר 2023, ולכן השפעתן טרם מורגשת במוצרים הנוכחיים.⁶ יתרה מכך, בראיון שקיים לאחרונה, מנכ"ל DeepSeek, הוא ציין כי להערכתו, המגבלה המשמעותית ביותר שהחברה צפויה להתמודד איתה בעתיד הנראה לעין היא זמינות מעבדים לאימון והפעלת המודלים שלה.⁷

בסוגיית העלות, בשורה התחתונה, DeepSeek הצהירה ש"הוא" עלה פחות משישה מליון דולר. אך קריטי להדגיש כי הנתון הזה הוא רק של העלות הסופית של אימון המודל, ללא רכיב התשתיות או עלות הפעלתו. מעבר לכך, חברת DeepSeek השיקה שני מודלים של בינה מלאכותית במקביל, R1 - שהוא מודל הסקה (Reasoning) ו-V3 - שהוא מודל שפה (LLM). סכום השישה מליון דולר שהחברה נקבה בו מתייחס ספציפית למודל V3, וזאת בעוד שההפתעה המשמעותית יותר הייתה דווקא מביצועי המודל R1. את העלויות של מודל R1 החברה הסינית לא חשפה. העלות האמיתית של אימון מודל V3 מוערכת בחצי מיליארד דולר, הכוללים את עלות חוות השרתים, מעבדים, חשמל וכח אדם - רכיבים שהם בלתי ניתנים

Jenia Jitsev  UA IL [@JJitsev], "DeepSeek R1 Distilled Llama 70B & Qwen 32B Models Claim to Solve Olympiad Level Math & Coding Problems, Matching O1-Mini Which Claims Same. Can They Handle Versions of AIW Problems That Reveal Generalization & Basic Reasoning Deficits in SOTA LLMs?(<https://arxiv.org/abs/2406.02061>)  1/n <https://t.co/uYmdCK0y0j>," Tweet, Twitter, February 3, 2025, <https://x.com/JJitsev/status/1886210118594760744>.

Ben Thompson, "DeepSeek FAQ," *Stratechery* (blog), January 27, 2025,⁵ <https://stratechery.com/2025/deepseek-faq/>.

Lennart Heim, "The Rise of DeepSeek: What the Headlines Miss," January 25, 2025,⁶ <https://blog.heim.xyz/deepseek-what-the-headlines-miss/>; DeepSeek-AI et al., "DeepSeek-V3 Technical Report" (arXiv, December 27, 2024), <https://doi.org/10.48550/arXiv.2412.19437>.

Jordan Schneider et al., "Deepseek: The Quiet Giant Leading China's AI Race," January 28, 2025,⁷ <https://www.chinatalk.media/p/deepseek-ceo-interview-with-chinas>.

להפרדה מחישוב העלויות האמיתיות וכלולות בעלויות האימון, שאותן נוהגות לדווח חברות מערביות.⁸

כשניגשים לחישוב סדור יותר של עלויות האימון עצמן, רואים כי מחירי האימון המצוטטים (כ-5.6 מיליון דולר ו/או כ-2.8 מיליון שעות או 57 ימי מעבד) בוצעו באמצעות מחשב על המונה 2048 מעבדי H800 של חברת nVidia והם מתייחסים למודל שפה גדול (LLM) קודם של החברה אשר שוחרר בסוף דצמבר 2024 – DeepSeek V3. לטענת DeepSeek, מודל זה משיג ביצועים מקבילים או טובים יותר בהשוואה למודלי ה-LLM המובילים של חברות מערביות, ביניהן OpenAI GPT 4o, Anthropic Claude 3.5 ו-B5META Llama 3.1 40B. Sonnet⁹ לשם השוואה, בעת פרסום מודל ה-LLM Llama 3.1 405B ביולי 2024, חברת META הצהירה כי נדרשו כ-30.8 שעות אימון באמצעות מחשב על המונה כ-16,000 מעבדי H100 של חברת nVidia – בעלות אימון משוערת של 92.4 מיליון דולר.¹⁰ עלות האימון המשוערת של מודל GPT 4o, אשר שוחרר במאי 2024, היא כ-78 מיליון דולר.¹¹ על פניו, נתונים אלו אכן מראים כי עלותו של המודל V3 הינה נמוכה בסדר גודל, ועוד על בסיס עוצמת חישוב מופחתת שהייתה זמינה לסינים. וזה אכן הנרטיב הראשוני שהתקבע.

אך יש להאיר מספר סייגים. ראשית, בתעשיית בינה המלאכותית מקובלת מגמה של שיפור ביצועי אימון והסקה של פי 3 בשנה, אשר ניתן להגיע אליהם רק משיפורי תוכנה ואלגוריתמיקה (דהיינו, בלי שימוש בשבבים חדשים). השיפורים של DeepSeek תואמים את קו המגמה הכללית ולכן יש לשקול נתון זה כשבזדקים עד כמה החברה פרצה דרך. בנוסף, מעבדי H800, ש-DeepSeek הצהירה כי היא משתמשת בהם, אינם "גרסה מוחלשת" של מעבדי H100 הנחשבים "מעבדים מתקדמים" (ומוטלות עליהם מגבלות ייצוא), אלא גרסה מיוחדת המותאמת לייצוא לנוכח ההגבלות של הממשל האמריקאי לפני אוקטובר 2023. על מנת לעמוד בהגבלות שקבע הממשל, חברת nVidia פיתחה שבב המאפשר ביצועים דומים באימון מודלים קטנים, אבל לא באימון מודלים גדולים.

Dylan Patel et al., "DeepSeek Debates: Chinese Leadership On Cost, True Training Cost, Closed Model Margin Impacts," *SemiAnalysis* (blog), January 31, 2025, <https://semianalysis.com/2025/01/31/deepseek-debates/>.

DeepSeek-AI et al., "DeepSeek-V3 Technical Report."⁹

Llama Team, "The Llama 3 Herd of Models" (AI @ Meta, July 23, 2024), ראה את <https://ai.meta.com/research/publications/the-llama-3-herd-of-models/>. לפרטי מחשב העל והעלות החישובית. עלות דולרית ע"ס הערכת עלות שימוש במעבד H100 העומדת על \$3/שעה.

Dennis Normile, "Chinese Firm's Faster, Cheaper AI Language Model Makes a Splash," *Science* 387, ¹¹ no. 6731 (January 15, 2025): 238, <https://doi.org/10.1126/science.z18lyzw>.

מבחינת עוצמת החישוב העומדת לרשות חברת DeepSeek, הערכות שונות גורסות כי בידיה כ-50,000 מעבדי nVidia Hopper (כלומר H100, H800 או H20) בנוסף למחשב בעל כ-10,000 מעבדי nVidia A100.¹² מעבר לכל אלה, עלו השערות והאשמות שלפיהן ייתכן ש-DeepSeek משתמשת גם בשבבים המתקדמים ביותר, דוגמת H100, שייתכן שהצליחה להבריח לתוך סין תוך עקיפת הגבלות היצוא והפיקוח. אין בשלב זה אפשרות להפריך או לאשש את הטענות הללו ולכן, כלל החישובים במאמר זה מתבססים על הנחה שדיווחי החברה על אודות משאבי המחשוב העומדים לרשותה הם אמיתיים ואמינים.

אם לפשט את הנתונים, אפשר לסכם אותם בכך שנראה שחברת DeepSeek אכן ביצעה התקדמות הנדסית מרשימה מאוד, אבל הפרשנויות שפורסמו בתקשורת על אודות המשמעויות שלה, ובפרט ההשוואה לביצועי מודלים מערביים הינן מוגזמות. DeepSeek לא הצליחה להראות שיפור מהותי מעבר לקו המגמה המקובל בעולם ולא הייתה פריצת דרך מדעית יסודית שמרעידה את אמות הסיפים של עולם האלגוריתמיקה וההנדסה.

החולשה המשמעותית ביותר ב-DeepSeek היא שהועלמו לחלוטין רכיבי התשתית שנדרשה – בראשם עוצמת החישוב והעלות הכוללת האמיתית של ייצור המודל, וכן הניסיון לגרום לעולם לחשוב שהייתה פריצת דרך על גבול האלכימיה, שבאמצעותה ניתן לבצע כמות חישובים עצומה באמצעות טכנולוגיות ישנות יחסית. כקנה מידה לעוצמת המחשוב האמיתית שהראנו בפרק זה, ניתן להשוותה לעוצמת החישוב המצופה ממחשב העל הישראלי, הנמצא כיום בשלבי מכרז.

בשנת 2021, מדינת ישראל החליטה להקים מעבדה לאומית לפיתוח מחשב-על שיעניק שירותי חישוב מתקדמים לצורכי המדינה והציבור. עתה נערוך השוואה בין המשאבים שנדרשו לפיתוח המודלים המדוברים של חברת DeepSeek לבין יכולותיו של מחשב העל הלאומי, המתוכנן לקום בישראל באמצעות מכרז של רשות החדשנות.

לפי התוכנית, במהלך 2026 אמור לקום מחשב שיידע לספק כ-75% מעוצמת החישוב שנדרשה על מנת לאמן את המודלים של DeepSeek (החישוב מתבסס על הערכה של כ-1500 מעבדי H100 במחשב הישראלי, לעומת 2048 מעבדי H800 במחשב הסיני המספקים ביצועי חישוב דומים) ולכל היותר יספק 44% מעוצמת החישוב (המוצהרת, בהנחה שאין לה עוד יכולות שהוסתרו, למשל כדי לא לחשוף הפרת סנקציות) הזמינה לה כעת. יתרה מכך, אם נשקלל את קו המגמה הכללי של עליה בדרישות המחשוב עבור מודלים גדולים מחד גיסא, והגידול במשאבי המחשוב של חברות הבינה המלאכותית, שממשיך וימשיך בפרק הזמן עד להקמתו בפועל של מחשב העל הלאומי, הרי שהפער בין יכולות המחשב לבין צורכי השוק

Patel et al., "DeepSeek Debates."¹²

הולך ומתרחב. לכן, למרות שיפורים אלגוריתמיים או רעיונות מפציעים לשיפור יעילות, מחשב העל הלאומי לא ידע לתת מענה תחרותי לאימון מודלים דוגמת R1 של DeepSeek, גם כאשר יפעל במלוא תפוקתו. כשם ש"חוק מור" מתאר דפוס קבוע בתעשייה של מזעור רכיבים כל שנתיים בשעה שתפוקתם מוכפלת, כך "פרדודקס גיבונס" מראה כי יעילות שימוש במשאב אינה מביאה להקטנת השימוש בו, אלא להיפך – השימוש בו גדל יותר מאשר קצב ההתייעלות, וחוזר חלילה. במקרה של DeepSeek, ניתן לחזות, אם כך, כי ההתקדמות בייעול החישוב לא צפויה להקטין את הדרישה למשאבי מחשוב (במקרה זה שבבים), אלא להגדיל את הדרישה להם. כשחלחלה ההבנה הזו בעולם, החרדה לעתידן של חברות השבבים נעלמה בהתאמה, הירידות החדות במניות של חברות דוגמת nVidia אחרי השקת DeepSeek נבלמו ושווין חזר ועלה.

לסיכום נושא זה, להערכתנו, בחישוב האופטימי ביותר, DeepSeek נדרשה לעוצמת חישוב גדולה פי 2 מעוצמתו המקסימלית של מחשב העל המתוכנן. בחישוב מחמיר יותר, יתכן שעוצמת החישוב שנדרשה ועמדה לרשות DeepSeek גדולה כמעט פי 10. מכאן אנו מסיקים, שתהיה זו טעות לתלות תקוות ב-DeepSeek ולראות בה מודל שיוכל לשרת את ישראל ולאפשר לה להגיע להישגים בבינה מלאכותית בלי להשקיע בתשתיות הנדרשות.

משמעויות

כמצוין בראשית הדברים, יש שלושה ממדים עיקריים להתפתחויות, הממד הטכנולוגי עצמו, הממד הגיאואסטרטגי וזה הישראלי.

את הממד הטכני עצמו סקרנו והצגנו את הפערים בין התפיסות לגבי DeepSeek לבין הביצועים והעליות האמיתיים. בממד הגיאואסטרטגי עולות מספר מסקנות אפשריות לגבי יעילות הפיקוח האמריקני. מפתה לחשוב שהפיקוח כשל, אבל זו מסקנה פשטנית, בראש ובראשונה מכיוון שמרבית ההצלחה של DeepSeek מבוססת על טכנולוגיה ורכיבים שהחברה השיגה לפני שההגבלות הנוכחיות הושטו. מסקנה מדויקת יותר תהיה שמשטר הפיקוח בעצם הצליח טוב מדי וגרם לאפקט לא צפוי, מכיוון שהוא אילץ את סין לחפש ולמצוא דרכים להתקדם שלא דווקא תלויות בהגדלת עוצמת החישוב הזמינה. למרות הפיקוח (ואולי בגללו) ארצות הברית לא הצליחה למנוע מסין מלהראות שיש ביכולתה להגיע איתה לאיזון אסטרטגי בתחום הבינה המלאכותית באמצעות טכנולוגיה ושיטות שאינן כפופות למשטר הפיקוח. עם זאת, הנתונים שהראינו מלמדים שסין לא פרצה דרך מדעית חדשה שמשנה את יסודות המחשוב, אלא שהצליחה לסגור את הפער בשילוב בין שיטות יצירתיות שאפשרו לה להתקדם ביחד עם קו המגמה והביצועים הצפויים בתחום. עדות נוספת לכך שהבעיה אינה נובעת מכשל טכני של הגבלות היצוא היא כי חברת DeepSeek עצמה חשפה את מגבלותיה להמשיך להתקדם ללא חומרה מתאימה - כזו שאין באפשרותה להשיג כיום בגלל ההגבלות. במצב הנוכחי,

במערב יש סדר גודל כזה של מעבדים מתקדמים (שאינן בסין, ככל הנראה) שכבר בשבועיים שעברו מאז יציאת DeepSeek לשוק, הצליחה חברת Open AI (המפתחת את Chat GPT) לפרסם מודל הסקה חדש, o3-mini – אשר מנצח את DeepSeek R1 בכלל המודלים וזאת רק בזכות נגישות עדיפה למעבדים מתקדמים. אפשר לצפות שבעתיד המודלים המערביים יצליחו להציג ביצועים טובים אף יותר באמצעות הטמעת השיטות בהן השתמשה DeepSeek. לכן, כל עוד סין נאלצת ומצליחה לפתח שיטות חדשות כדי להתמודד עם היעדר משאבי מחשוב מחד גיסא, וארצות הברית מנצלת את השיטות הללו ביחד עם עוצמת המחשוב העדיפה שלה מאידך גיסא, נראה כי שתי המעצמות ימצאו בנקודת איזון אסטרטגית.

בצד האמריקאי, הדבר מחייב התבוננות מחודשת על משטר הפיקוח. אמנם סין לא תוכל, ככל הנראה, להגיע להובלה בלי גישה לחומרה האמריקאית, אבל היא הצליחה לנתק את הקשר בין סנקציות למצב בו היא מצליחה מבחינת מחקר ופיתוח לשמור על קו המגמה המקובל בעולם גם במגבלות החומרה החלות עליה. מצד אחד, זאת תוצאה לא רעה ותואמת את ההגבלות אשר נועדו, לפחות ברמה המוצהרת, למנוע מסין לעקוף את ארצות הברית. אך באותה מידה ניתן לומר כי משטר ההגבלות ההולך ומחריף בעצם אילץ את סין למצוא אלטרנטיבה ולבסס את עצמאותה הטכנולוגית, אפילו אם לא תהיה לה יכולת להשיג הובלה. מבחינה זו, ייתכן וניתן לראות ב-DeepSeek סנונית ראשונה לאוטרכיה סינית טכנולוגית, שאיפה אשר סין הכריזה עליה כבר לפני שנים. דוגמאות לאוטרכיה המתהווה היא שחברת DeepSeek מעסיקה רק סינים שלמדו באוניברסיטאות סיניות והיא גם ובוחנת שימוש במעבדי Huawei סינים להפעלת המודלים שלה. לסיכום, ברמה הטכנולוגית, כנראה שסין לא תוכל לעקוף את ארצות הברית, אבל ארצות הברית לא תוכל למנוע את שמירת השוויון על ידי סין באמצעות הגבלות יצוא.

אלו אפשרויות עומדות בפני ארצות הברית להתמודד עם המצב החדש? ניתן אולי לדמות את המצב היום לזה שארצות הברית היתה בו בשנים שקדמו לפיתוח פצצת האטום. כשהבינה את הפוטנציאל היישומי של הפיזיקה הגרעינית, המדע הפך מתחום אקדמי פתוח לתחום מסווג וכל המחקר האמריקאי נכנס תחת תוכנית מנהטן.¹³ זה אפשר לארצות הברית לשלוט בתפוצת הידע לפרק זמן מוגבל וקריטי. כיום, מדע הבינה המלאכותית נמצא בנקודה דומה. לכן, יתכן והכלי המשמעותי ביותר העומד לרשות ארצות הברית יהיה הגבלת פרסום מחקרים מדעיים בסיסיים על ידי הכנסת כל תחום הבינה המלאכותית תחת תכנית סטארגייט שהנשיא טראמפ הכריז עליה ביומו הראשון בתפקיד. וזו תהיה, בעצם, תכנית מנהטן של הבינה המלאכותית. כך או אחרת, שתי המעצמות – והעולם כולו, למעשה – יצטרכו ללמוד לחיות

¹³ Michael S. Sweeney, *Secrets of Victory: The Office of Censorship and the American Press and Radio in World War II* (Chapel Hill, NC: University of North Carolina Press, 2001), 196–200.

ולהתמודד עם מצב חדש של איזון אסטרטגי בין סין וארצות הברית ולשקלל זאת לתוך חישוביהן האסטרטגיים.

בממד המשמעותי עבור ישראל, עולות מספר מסקנות עקרוניות. החשובה שביניהן היא, שלמרות הפיתוי לראות ב-DeepSeek פתרון קסם, אסור שישראל תסתנוור ותחשוב שיש בנמצא קיצור דרך טכנולוגי שייתר את הצורך בעוצמת חישוב על מנת לקיים תכנית בינה מלאכותית מובילה ובת-קיימא. כפי שהראינו, DeepSeek עשתה שימוש, לכל הפחות, בעוצמת חישוב שהיא גדולה בעשרות אחוזים מאשר זו שישראל תוכל לספק באמצעות מחשב העל שיוקם במסגרת התכנית הלאומית הקיימת. וזאת, מבלי להתחשב בדרישות החישוב שמכילות עצמן בערך כל שישה חודשים וגורמות לכל מחשב להתיישן במהירות ולאבד את הרלוונטיות שלו בלי השקעה מתמדת בשדרוגו.

במרחב שיקוליה הגיאופוליטיים, ישראל מוצאת את עצמה בדילמה. כיום, ישראל נמצאת בנחיתות בתחום הבינה המלאכותית והיא חייבת לפלס דרך חדשה כדי להתקדם. משום שארצות הברית, בעלת בריתה החשובה ביותר, השיתה עליה הגבלות, עלולים להיות כאלה שיתפתו לחפש פתרונות המבוססים על טכנולוגיות סיניות, למשל שבבים של חברת Huawei, או שיתופי פעולה עם חברות סיניות בתחום על מנת לקדם את יכולותיה של ישראל. תהיה זו תוצאה עגומה ביותר וחיוני שישראל וארצות הברית תעשנה כל מאמץ למנוע ממנה להתממש.

דרך המלך עוברת בעלייתה של ישראל לדרג א' והסרת ההגבלות מעליה. ישנם שלושה מסלולים אפשריים לשם. הראשון הוא הסתמכות על מערכת היחסים בין ממשלת ישראל והנשיא טראמפ על מנת שזה יקדם את ישראל לדרג א', או לכל הפחות יחריג את ישראל מההגבלות. מסלול אפשרי נוסף הוא שיפור מעמדה של ישראל כחלק מהתמורה עבור הסדר מדיני רחב יותר במזרח התיכון. והדרך השלישית, ייתכן שהיא הסבירה ביותר, היא להאיץ את דיאלוג ה- Trusted Technologies בין המדינות, שהחל בימי ממשלת בנט לאחר החתימה על הצהרת ירושלים על ידי הנשיא ביידן בירושלים ומתקדם בעצלותיים (בלשון המעטה) – וליישמו בהקדם. כל אלה, במקביל להתאמת התוכניות הלאומיות לצרכים שנמנו במאמר זה, על מנת לבסס הובלה בבינה המלאכותית ולהגביר את נכסיותה של ישראל עבור ארצות הברית.

סיכום

האירועים וההתפתחויות בתחום הבינה המלאכותית בתקופה האחרונה, וביתר שאת בשבועות האחרונים, מביאים למסקנה כי אל לה לישראל להסתנוור ולהידבק בשיכרון המעמקים של DeepSeek. במקום זאת, מתגבשת ההכרה בכך שיש צורך לעדכן את התוכניות הלאומיות ולאמץ תוכנית טכנולוגיה לאומית, שתבטיח תשתית טכנולוגיה מינימלית אשר תבצר את ריבונותה הטכנולוגית, תיצור את התנאים ההכרחיים עבור התעשייה הישראלית למצות את

יכולותיה, תחזק את נכסיותה של ישראל בעיני ארצות הברית, תיצור הזדמנויות מלהיבות לשיתופי פעולה אזוריים עם מדינות המפרץ ובכללן ערב הסעודית, ותאפשר לישראל לתפוס את מקומה הראוי בשורה הראשונה של מדינות העולם ולנסוק לגבהים חדשים.

עורכי הסדרה: ענת קורץ ואלדד שביט