The Clean Paper

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

האם "מודלי יסוד" גדולים לרובוטים באמת עובדים טוב יותר? תשובה זהירה — במידה מתונה כן, ורוב המחקרים לא יכולים לדעת

אימן "מודלי התנהגות גדולים" — מדיניות רובוטית שעברה קדם-אימון על *Toyota Research Institute* כ-*1,700* שעות של נתוני מניפולציה מגוונים — והשווה אותם למדיניות למשימה יחידה שאומנה מאפס בפרוטוקול מחמיר במיוחד. ניסויים עיוורים, אקראיים ובעלי מדגם גדול (כ-*1,800* בעולם האמיתי ויותר מ-*47,000* בסימולציה), עם ניתוח סטטיסטי ממשי. לאחר כיוונון לכל משימה המודלים הגדולים הצליחו בממוצע טוב יותר, נזקקו בערך לפי *3–5* פחות נתונים ספציפיים למשימה והיו עמידים יותר כשהתנאים השתנו, הביצועים השתפרו באופן חלק ככל שנוספו נתוני קדם-אימון. אבל ללא כיוונון הם לא ניצחו בעקביות מודלים למשימה יחידה, כמה מהאפקטים היו קטנים מספיק שרק המדגמים הגדולים חשפו אותם, ובחירת נרמול נתונים שגרתית השפיעה יותר מהארכיטקטורה. זו תמיכה מדודה בכיוון של מודלי יסוד לרובוטים — לא רובוט לשימוש כללי, לא גנרליסט *zero-shot* ולא "זינוק מגיח" — ובמקביל אזהרה חדה שחלק ניכר ממחקרי הרובוטיקה אולי מודדים רעש.

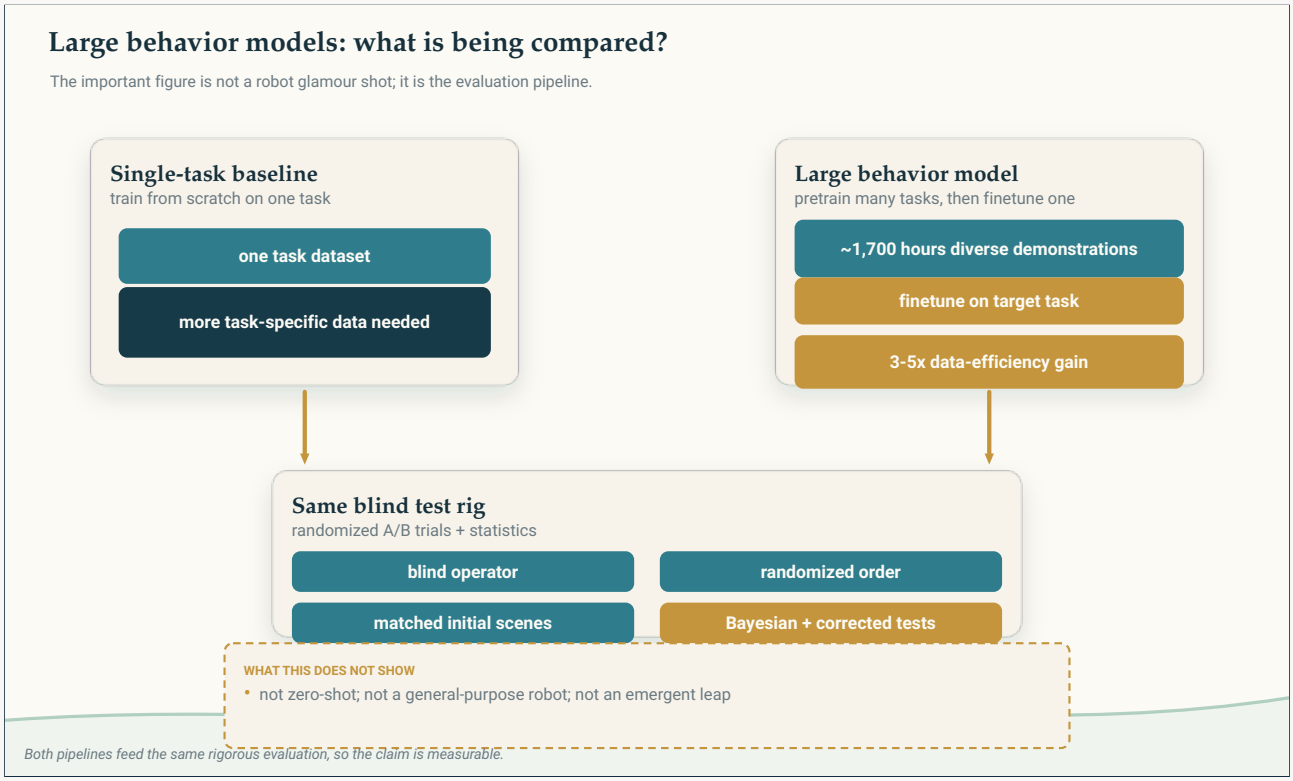
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *A Careful Examination of Large Behavior Models for Multitask Dexterous Manipulation*, Toyota Research Institute Large Behavior Model Team — J. Barreiros, A. Beaulieu, et al.; senior authors incl. R. Ambrus, B. Burchfiel, S. Feng, H. Kress-Gazit (Cornell), R. Tedrake, Science Robotics ;(2026) preprint arXiv:2507.05331 · DOI: 10.1126/scirobotics.aea6201

מאמר על רובוטים שהנושא האמיתי שלו הוא כנות

תחום הרובוטיקה נמצא בעיצומו של מרוץ אחר "מודלי יסוד" (foundation models). הרעיון, שהושאל מעולמות הבינה המלאכותית לשפה ולתמונות, מפתה מאוד: במקום לאמן רובוט לכל משימה בנפרד, מאמנים **"מודל התנהגות גדול"** (large behavior model, LBM) אחד על מאגר עצום ומגוון של הדגמות, ומקבלים מערכת שאמורה להיות בעלת יכולת רחבה ולהסתגל במהירות. ההתלהבות — וגם ההשקעה — עצומות. הכותרת כמעט כותבת את עצמה: המוחות הרובוטיים לשימוש כללי כבר כאן.

המאמר הזה, של Toyota Research Institute, מעניין דווקא מפני שהוא מסרב לכתוב את הכותרת הזאת. התרומה האמיתית שלו אינה רובוט נוצץ יותר, אלא בחינה קשוחה של שאלה שנראית פשוטה אך מטעה — האם גישת המודל הגדול באמת עובדת טוב יותר, ואיך בכלל נדע? — תוך שימוש ברמת הקפדה סטטיסטית שלפי המחברים עצמם אינה שכיחה בתחום. התוצאה היא "כן, אבל" שימושי באמת, וגם אזהרה שחלק ניכר ממחקרי הרובוטיקה אולי מודדים רעש.



שתי הגישות נכנסות לאותה הערכה עיוורת ואקראית. טענת המאמר אינה שמודלי יסוד לרובוטים הם גנרליסטים ללא התאמה מוקדמת, אלא שקדם-אימון יכול לשפר את יעילות הנתונים ואת העמידות כאשר מודדים זאת בקפידה.

מה עשו המחברים

הם בנו LBMs במובן מסוים וקונקרטי: **מדיניות ויזו-מוטורית מבוססת דיפוזיה** (Diffusion Transformer) שקורא תמונות מצלמה, הוראת שפה קצרה ואת מצבי המפרקים של הרובוט, ומפיק מקטעים קצרים של פקודות תנועה בקצב 10 הרץ). המודלים עברו **קדם-אימון על כ-1,700 שעות** של הדגמות רובוטיות — יותר מ-500 משימות שונות

שנאספו במעבדה, יחד עם מאגרי נתונים ציבוריים — ולאחר מכן עברו **כיוונון עדין** (finetuning) למשימות בודדות. לאורך כל המאמר ההשוואה היא מול **מדיניות למשימה יחידה שאומנה מאפס** רק על הנתונים של אותה משימה.

אבל לב המאמר הוא דווקא ההערכה, שהמחברים מתייחסים אליה כתוצאה המרכזית. כדי לא להטעות את עצמם הם השתמשו ב:

- **בדיקת A/B עיוורת ואקראית** בעולם האמיתי — האדם שהפעיל את הרובוט לא ידע איזו מדיניות נבדקת, וסדר הניסויים היה אקראי.
- **תנאי פתיחה מבוקרים וחוזרים על עצמם** — המפעילים התאימו את הסצנה לתמונה שקופה שהוצגה מעליה לפני כל ניסיון.
- **מספר גדול של ניסויים** — 50 הרצות בעולם האמיתי לכל משימה, לכל מדיניות ולכל תנאי; 200 לכל משימה בסימולציה. בסך הכול: כ-1,800 **הרצות עיוורות בעולם האמיתי ויותר מ-47,000 הרצות בסימולציה**.
- **סטטיסטיקה ראויה** — אומדנים בייסיאניים להסתברות ההצלחה, ובדיקות השערה בזוגות עם תיקון להשוואות מרובות, במקום להסתכל בעין על פסי שגיאה חופפים. הם אף ביצעו בקרת איכות לרבע מהניסויים שדורגו בידי בני אדם כדי למדוד את שגיאת הדירוג.

[video pending]

השוואה זה לצד זה של מודלים המסדרים שולחן לארוחת בוקר: משמאל, קו הבסיס למשימה יחידה; מימין, LBM. שני הסרטונים מוצגים במהירות רגילה. זו משימה אחת מתוך ההערכה, ולא הוכחה לאוטונומיה כללית.

זה לב העניין. המאמר כולו הוא טיעון שלפיו בלי המנגנון הזה אי-אפשר לדעת אם שיפור הוא אמיתי או פשוט מזל.

מה הם מצאו

מודלים גדולים שעברו כיוונון עדין ניצחו, בממוצע, מודלים למשימה יחידה שאומנו מאפס. כאשר מצרפים את התוצאות על פני משימות, LBM שעבר קדם-אימון ולאחר מכן כיוונון למשימה מסוימת עלה באופן עקבי על מדיניות שאומנה מאפס לאותה משימה, הן בסימולציה והן בעולם האמיתי, וההבדל היה מובהק סטטיסטית. ברמת המשימות הבודדות, ה-LBM המכוונן היה טוב לפחות כמו המודל שאומן מאפס כמעט בכל המקרים (3 מתוך 3 משימות בעולם האמיתי, 15 מתוך 16 בסימולציה).

היתרון הגדול והברור ביותר הוא יעילות נתונים. LBM מכוונן הגיע לביצועים שקולים לאימון מאפס תוך שימוש בכמות קטנה בערך פי **3–5 של נתונים ספציפיים למשימה**. במשימה אחת בעולם האמיתי — עריכת שולחן לארוחת בוקר — LBM שעבר כיוונון על רק 15% מההדגמות עקף מדיניות שאומנה מאפס על 100% מהן.

קדם-אימון עוזר במיוחד כאשר התנאים משתנים. כאשר סביבת הבדיקה הוזזה במכוון מתנאי האימון ("distribution shift"), היתרון של ה-LBM המכוונן דווקא גדל. באחת מקבוצות הסימולציה הוא גבר באופן מובהק על המודל שאומן מאפס ב-3 מתוך 16 משימות בתנאים רגילים, אך ב-10 מתוך 16 תחת שינוי התפלגות. מאחר שבפריסה אמיתית התנאים תמיד סוטים במידת מה מתנאי האימון, העמידות הזאת היא אולי הממצא החשוב ביותר מבחינה מעשית.

יותר נתוני קדם-אימון עזרו באופן חלק. הביצועים עלו בהדרגה ככל שנוספו נתוני קדם-אימון — **בלי קפיצה פתאומית או "זינוק מגיח"** בקני המידה שנבדקו. שימושי, צפוי ולא דרמטי.

אבל הסיפור על גנרליסט שלא זקוק לכיוונון לא החזיק מעמד. LBM שעבר קדם-אימון ושימש *zero-shot* —

ללא כיוונון ייעודי למשימה — לא ניצח באופן עקבי מודלים למשימה יחידה. רשת אחת יכלה לבצע משימות רבות, אבל החלום של "פשוט תגידו לו מה לעשות" לא קיבל כאן תמיכה; המחברים מייחסים חלק מכך לשבריריות של מקודד השפה הקטן שבו השתמשו.

והשיפורים היו קטנים מספיק כדי שיהיה קל לפספס אותם — או לדמיין אותם. רבים מהאפקטים נעשו גלויים רק בזכות המדגמים הגדולים מהמקובל והבדיקות הקפדניות. המחברים אומרים במפורש שלנוכח גודל האפקטים והרעש, **קיים סיכון משמעותי שמאמרים רבים ברובוטיקה מודדים רעש סטטיסטי.** הם גם מצאו שבחירה שגרתית לכאורה — האופן שבו הנתונים עוברים נרמול — השפיעה על התוצאות יותר משינויים בארכיטקטורה, ושבאג בנרמול בזמן קדם-האימון התגלה רק לאחר סיום ההערכות.

מה זה כנראה אומר

הפרשנות שאפשר להגן עליה היא שקדם-אימון בקנה מידה גדול על נתוני רובוטיקה מגוונים הוא מרכיב אמיתי וכדאי: הוא מפחית את כמות הנתונים הדרושה לכל משימה חדשה והופך את המדיניות לעמידה יותר כשהעולם אינו תואם בדיוק את תנאי האימון. זה אכן תומך בכיוון שעליו מהמר התחום. אבל השיפור הוא מתון ותלוי תנאים — הוא מופיע בעיקר אחרי כיוונון, ובולט במיוחד כאשר מצרפים משימות או מפעילים לחץ על המערכת — ולא מבשר על הופעתו של רובוט כללי שאפשר פשוט להציב בכל מקום.

המשמעות השקטה והחשובה יותר היא מתודולוגית. בפועל, המאמר מציע סרגל מדידה: הוא מראה כמה ראיות באמת צריך כדי לטעון טענה אמינה על מדיניות רובוטית, ומרמז שחלק גדול מההתלהבות שפורסמה נשען על מעט מדי נתונים. זה תיקון שהתחום זקוק לו יותר מאשר לעוד מודל.

מה זה לא מוכיח

- זה **לא רובוט לשימוש כללי.** היתרונות הודגמו עבור ארכיטקטורה מסוימת (מדיניות דיפוזיה) שעוברת כיוונון נפרד לכל משימה, בתנאים מבוקרים, על בסיס הדגמות בטל-אופרציה — לא רובוט אוטונומי שמבצע כל משימה חדשה לפי הוראה.
- זה **לא** מאשש שימוש zero-shot. ללא כיוונון, המודל הגדול לא ניצח באופן עקבי את קווי הבסיס למשימה יחידה.
- זו **לא** עדות ל**"זינוק מגיח"**. הגדלת הקנה שיפרה ביצועים בהדרגה; אין כאן אי-רציפות שתומכת בסיפור של "ואז פתאום הוא נעשה מסוגל".
- המספרים **יחסיים ומוגבלים למעבדה.** שיעורי ההצלחה האבסולוטיים כוונו בכוונה לסביבות 50% כדי להפוך את ההשוואה לרגישה יותר; הם אינם מדד לאמינות בעולם האמיתי, והמחקר עוסק בארכיטקטורה אחת ממעבדה אחת.
- המחקר **לא** קובע **למה** מדיניות מסוימת מצליחה או נכשלת, וכמה משימות שבהן המודל הגדול ביצע גרוע יותר מדווחות אך אינן מוסברות.
- הוא לא אומר **דבר על בטיחות, אוטונומיה או פריסה** מחוץ למערכת ההערכה.

עד כמה הראיות חזקות?

לגבי הטענות ההשוואתיות המרכזיות — ש-LBM מכוונים עולים בממוצע על קווי בסיס שאומנו מאפס, זקוקים לכמות נתונים קטנה פי כמה ועמידים יותר לשינוי התפלגות — הראיות חזקות וגם מבוקרות באופן חריג: עיוורות, אקראיות, מבוססות מדגם גדול, נבדקות סטטיסטית ועם בקרת איכות של הדירוג. זה מקרה נדיר שבו המתודולוגיה טובה מספיק כדי לקבל את המסקנות הראשיות כמעט כפשוטן.

ההסתייגויות הכנות הן בדיוק אלה שהמחברים מעלים בעצמם. פסי השגיאה שלהם מייצגים את האקראיות של ההערכה, אבל **לא** את האקראיות של האימון — אם מאמנים את אותו מודל פעמיים, ייתכן שמקבלים מדיניות שונה באופן משמעותי, והשונות הזאת אינה כלולה בסטטיסטיקה. במשימות בעולם האמיתי היו 50 ניסויים לכל משימה: מספיק כדי לגלות אפקטים בינוניים, אך לא בהכרח קטנים. מיזוג הוראות השפה נעשה בעזרת מקודד צנוע, ולכן מסקנות לגבי "פשוט תגידו לרובוט מה לעשות" עשויות להשתנות במערכות גדולות יותר. ויש גם הגילוי הישיר של באג בנרמול שנמצא לאחר מעשה. אף אחת מהנקודות הללו אינה ממוטטת את הממצאים המרכזיים, אבל הן בדיוק מסוג הדברים שהמאמר טוען שהתחום נוהג לעתים קרובות לדחוק הצדה.

הערת מקור, באותה רוח: ההסבר הזה מבוסס על ה-preprint של המחברים. לא הצלחנו להשיג את הגרסה שפורסמה בכתב העת, ולכן לא בדקנו אם חלו שינויים בין ה-preprint לטקסט שפורסם.

למה זה חשוב

"מודלי יסוד לרובוטים" הוא ביטוי שכמעט מזמין הגזמה, וקל לפרש מחקר כזה בצורה קיצונית לשני הכיוונים — כ-"זה עובד!" ניצחוני או כ-"זה רק הייפ" מזלזל. הקריאה המדויקת שימושית יותר משתיהן: קדם-אימון על נתונים מגוונים נותן יתרונות אמיתיים, מדידים אך מתונים — בעיקר פחות נתונים לכל משימה ויותר עמידות — והביצועים משתפרים באופן צפוי עם קנה המידה.

הסיבה העמוקה יותר לחשיבותו היא שהמאמר מפנה את הקפדנות שלו אל התחום עצמו. בכך שהוא מראה שהאפקטים האמיתיים קטנים מספיק כדי להיעלם תחת הערכה רשלנית, ושבחירה משעממת כמו נרמול נתונים יכולה להשפיע יותר מארכיטקטורה חדשה ומתוחכמת, הוא טוען שחלק גדול מההתקדמות בלמידת רובוטים זקוק למדידה מוצקה יותר לפני שאפשר להאמין לו. מאמר שמבזבז חלק מההילה שלו על שמירת הגבול בין תוצאה לבין משאלה עושה משהו נדיר יותר — ויקר ערך יותר — מאשר רק לטפס לראש טבלת דירוג.

בקצרה

חוקרים ב-Toyota Research Institute אימנו "מודלי התנהגות גדולים" — מדיניות רובוטית מבוססת דיפוזיה שעברה קדם-אימון על כ-1,700 שעות של נתוני מניפולציה מגוונים — והשוו אותם למדיניות למשימה יחידה שאומנה מאפס, בפרוטוקול קפדני במיוחד: ניסויים עיוורים, אקראיים ובעלי מדגם גדול (כ-1,800 בעולם האמיתי ויותר מ-47,000 בסימולציה), עם ניתוח סטטיסטי ממשי. לאחר כיוונון לכל משימה, המודלים הגדולים הצליחו בממוצע טוב יותר, נזקקו לכמות קטנה בערך פי 3–5 **של נתונים ספציפיים למשימה**, והיו **עמידים יותר כאשר התנאים השתנו**; הביצועים השתפרו בהדרגה ככל שנוספו נתוני קדם-אימון. אבל **ללא כיוונון** הם לא ניצחו באופן עקבי מודלים למשימה יחידה, כמה מהאפקטים היו קטנים מספיק שרק המדגמים הגדולים חשפו אותם, ובחירת נרמול נתונים שגרתית השפיעה יותר מהארכיטקטורה. זו תמיכה מוצקה ומדודה בכיוון של מודלי יסוד לרובוטים — **לא** רובוט לשימוש כללי, לא גנרליסט

zero-shot ולא "זינוק מגיח" — ובמקביל אזהרה חדה שחלק ניכר ממחקרי הרובוטיקה אולי מודדים רעש.

בדיקה מפוכחת

מה המאמר מראה: בהערכה קפדנית, עיוורת ובעלת כוח סטטיסטי (כ-1,800 הרצות בעולם האמיתי + יותר מ-47,000 בסימולציה), מדיניות דיפוזיה שעברה קדם-אימון רב-משימתי ואז כיוונון (LBMs) עולה בממוצע על מדיניות למשימה יחידה שאומנה מאפס, מגיעה לביצועים דומים עם בערך פי 3–5 פחות נתונים ספציפיים למשימה, ועמידה יותר לשינוי התפלגות; הביצועים משתפרים באופן חלק עם הגדלת נתוני הקדם-אימון.

מה סביר אך לא הוכח: שהיתרונות האלה יעברו גם למודלי vision-language-action גדולים בהרבה (מקודד השפה כאן היה קטן); שהשיפור החלק עם קנה המידה יימשך מעבר לטווח הנתונים שנבדק.

מה הוא לא מראה: רובוט לשימוש כללי או zero-shot (בלי כיוונון → אין יתרון עקבי); קפיצה "מגיחה" ביכולת; הסבר לכישלונות ספציפיים ברמת המשימה; אמינות אבסולוטית בעולם האמיתי (שיעורי ההצלחה כוונו לסביבות 50% כדי להגביר רגישות); או משהו על בטיחות ופריסה אוטונומית.

המגבלות העיקריות: הסטטיסטיקה כוללת אקראיות בהערכה אך לא שונות בין ריצות אימון; 50 ניסויים בעולם האמיתי לכל משימה עלולים לפספס אפקטים קטנים; ארכיטקטורה אחת ומעבדה אחת; מקודד שפה צנוע; באג בנרמול נתונים התגלה אחרי ההערכות; הניתוח מבוסס על ה-preprint (הגרסה שפורסמה לא נבדקה).

כמה ביטחון צריך להיות לקורא כללי? גבוה בכך שקדם-אימון רב-משימתי יחד עם כיוונון נותנים יתרונות אמיתיים אך מתונים — בעיקר יעילות נתונים ועמידות — ושהיתרונות נמדדו כאן בזהירות חריגה. גבוה גם בכך שזה **לא** רובוט כללי או zero-shot **ולא** זינוק מגיח. בינוני לגבי השאלה עד כמה השיפור יימשך במודלים גדולים יותר. וכדאי לקחת ברצינות את האזהרה של המחברים עצמם: האפקטים בתחום קטנים מספיק כך שמחקרים חסרי כוח סטטיסטי עלולים לדווח על רעש. הגישה המתאימה: אופטימיות מדודה לגבי הכיוון, וספקנות בריאה כלפי תוצאות ברובוטיקה שאין מאחוריהן בסיס סטטיסטי דומה.

מקורות

arXiv:2507.05331
https://arxiv.org/abs/2507.05331

10.1126/scirobotics.aea6201 Robotics, Science — retrieved) (not version Peer-reviewed
https://doi.org/10.1126/scirobotics.aea6201

page Project
https://toyotaresearchinstitute.github.io/lbm1/