The CLEAN PAPER

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

هل تعمل «النماذج التأسيسية» الكبيرة للروبوتات بصورة أفضل فعلًا؟ إجابة دقيقة: نعم، بشكل متواضع — ومعظم الدراسات لا تستطيع التمييز

درّب *Toyota Research Institute* «نماذج سلوك كبيرة» — سياسات روبوتية دُرّبت مسبقًا على نحو 1,700 ساعة من بيانات المناولة المتنوعة — واختبرها ضد سياسات لمهمة واحدة مدربة من الصفر بصرامة غير معتادة: تجارب عمياء عشوائية كبيرة العينة (نحو 1,800 حقيقية وأكثر من 47,000 بالمحاكاة) مع إحصاء فعلي. بعد الضبط الدقيق لكل مهمة، كان أداء النماذج الكبيرة أفضل في المتوسط، واحتاجت بيانات خاصة بالمهمة أقل بنحو 3–5 مرات، وكانت أكثر متانة عندما تغيرت الظروف؛ وارتفع الأداء بسلاسة مع مزيد من بيانات التدريب المسبق. لكن من دون ضبط دقيق لم تتفوق باستمرار على نماذج المهمة الواحدة، وكانت عدة تأثيرات صغيرة لدرجة تطلبت عينات كبيرة لرؤيتها، كما أن خيارًا عاديًا لتطبيع البيانات كان أهم من المعمارية. إنها أدلة موزونة لصالح اتجاه النماذج التأسيسية للروبوتات — لا روبوتًا عامًا، ولا نموذجًا *zero-shot*، ولا «قفزة ناشئة» — وتحذير من أن كثيرًا من المجال قد يقيس الضوضاء.

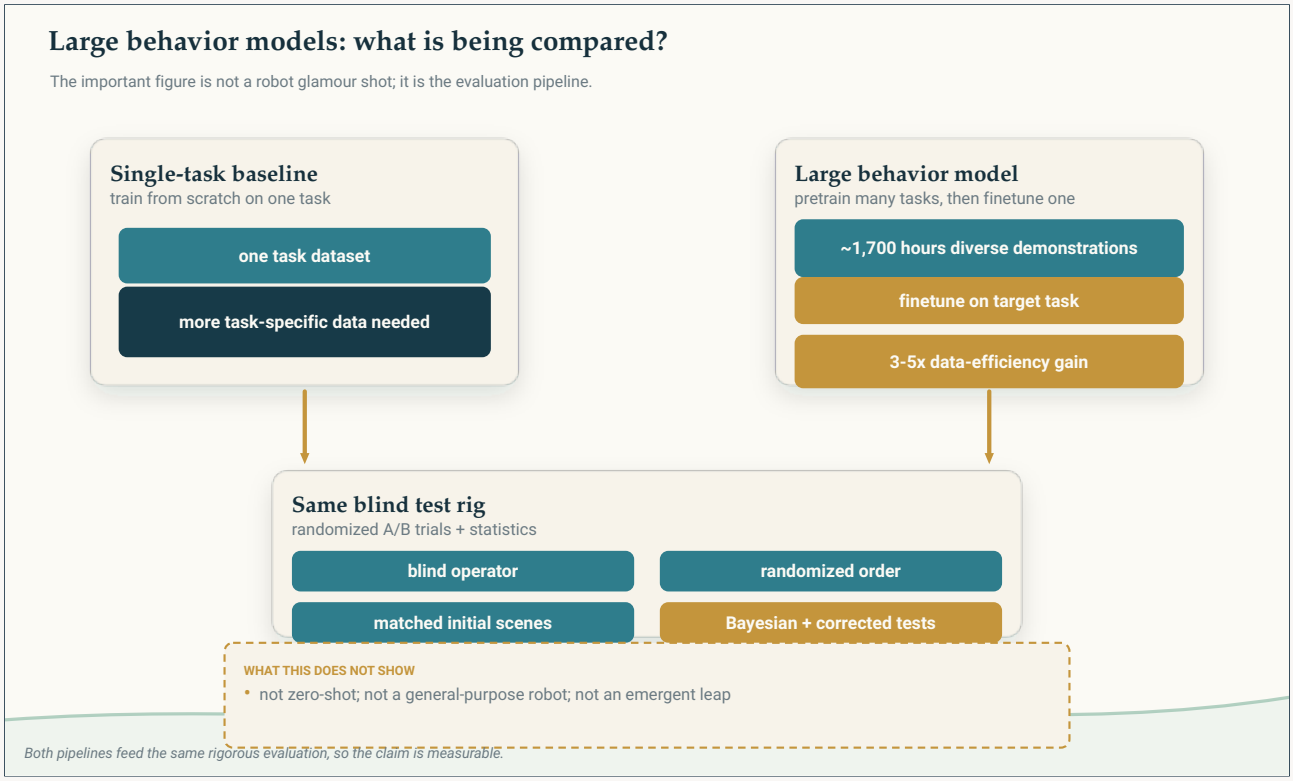
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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: scirobotics.aea6201/1126.10

ورقة عن الروبوتات موضوعها الحقيقي هو الصدق

يمر مجال الروبوتات حاليًا بحمى «النماذج التأسيسية». الفكرة، المستعارة من نماذج اللغة والصور، مغرية: بدل تدريب روبوت لمهمة واحدة في كل مرة، درّب **«نموذج سلوك كبير» (LBM)** واحدًا على كم هائل ومتنوع من العروض العملية، فتحصل على نظام واسع القدرات وسريع التكيف. الحماس — والاستثمار — هائلان. والعنوان يكتب نفسه: وصلت أدمغة الروبوتات ذات الأغراض العامة.

هذه الدراسة من Toyota Research Institute مثيرة تحديدًا لأنها ترفض كتابة ذلك العنوان. مساهمتها الحقيقية ليست روبوتًا أكثر استعراضًا. بل فحص صارم لسؤال يبدو بسيطًا على نحو خادع: هل يعمل نهج النموذج الكبير فعلًا بصورة أفضل، وكيف يمكننا أصلًا أن نعرف؟ والإجابة جاءت بدرجة من العناية الإحصائية يقول الباحثون أنفسهم إنها غير معتادة في المجال. النتيجة «نعم، ولكن» مفيدة حقًا، وتحذير من أن كثيرًا من أبحاث الروبوتات قد تكون تقيس الضوضاء.



يخضع النهجان للتقييم الأعمى العشوائي نفسه. لا تدّعي الدراسة أن النماذج التأسيسية للروبوتات عموميون يعملون بلا أمثلة سابقة، بل إن التدريب المسبق قد يحسن كفاءة البيانات والمتانة عندما يُقاس بعناية.

ما الذي فعله الباحثون؟

بنوا نماذج LBM بمعنى محدد وملموس: **سياسات بصرية-حركية قائمة على الانتشار** — Transformer Dffiusion يقرأ صور الكاميرا، وتعليمًا لغويًا قصيرًا، ومواضع مفاصل الروبوت نفسه، ثم يُخرج دفعات قصيرة من أوامر الحركة عند 10 Hz. دُرّبت هذه النماذج **مسبقًا على نحو 1,700 ساعة** من عروض الروبوت — أكثر من 500 مهمة مختلفة جُمعت داخليًا، إضافة إلى مجموعات بيانات عامة — ثم خضعت **للضبط**

الدقيق على مهام منفردة. والمقارنة في الدراسة كلها هي مع **سياسة لمهمة واحدة تُدرَّب من الصفر** على بيانات تلك المهمة وحدها.

لكن قلب الدراسة هو التقييم، الذي يعامله الباحثون بوصفه النتيجة الرئيسية. ولتجنب خداع أنفسهم استخدموا:

- **اختبار A/B أعمى وعشوائي** في العالم الحقيقي — الشخص الذي يشغّل الروبوت لا يعرف أي سياسة تُختبر، وترتيب الاختبارات عشوائي.
- **شروط بداية مضبوطة وقابلة للتكرار** — يطابق المشغلون المشهد مع طبقة صورة مرجعية قبل كل تجربة.
- **أعدادًا كبيرة من التجارب** — 50 تشغيلًا حقيقيًا لكل مهمة ولكل سياسة ولكل حالة؛ و200 لكل مهمة في المحاكاة. المجموع: نحو **1,800 تشغيل أعمى في العالم الحقيقي وأكثر من 47,000 تشغيل بالمحاكاة.**
- **إحصاءً فعليًا** — تقديرات Bayesian لاحتمال النجاح، واختبارات فرضيات زوجية مع تصحيحات للمقارنات المتعددة، بدل الاكتفاء بالنظر إلى أشرطة خطأ متداخلة. وأجروا حتى مرورًا لضمان الجودة على ربع التجارب التي قيّمها بشر لقياس خطأ التقييم.

[video pending]

مقارنة جنبًا إلى جنب لنموذجين يرتبان مائدة إفطار: إلى اليسار خط أساس لمهمة واحدة، وإلى اليمين LBM. كلا الفيديوين يعمل بسرعة 1×. هذه مهمة واحدة خضعت للتقييم، وليست دليلًا على استقلالية عامة الغرض.

هذه المنهجية هي النقطة. الدراسة كلها حجة تقول إنه من دونها لا يمكنك التمييز بين تحسن حقيقي وضربة حظ.

ماذا وجدوا؟

بعد الضبط الدقيق، تفوقت النماذج الكبيرة على نماذج المهمة الواحدة المدربة من الصفر — في المتوسط. عند تجميع النتائج عبر المهام، تفوق LBM دُرّب مسبقًا ثم ضُبط دقيقًا على مهمة بصورة موثوقة على سياسة دُرّبت من الصفر على المهمة نفسها، في المحاكاة والعالم الحقيقي، وكان الفارق ذا دلالة إحصائية. وعلى مستوى المهام الفردية كان LBM المضبوط دقيقًا مساويًا أو أفضل إحصائيًا في جميع الحالات تقريبًا: 3/3 من مهام العالم الحقيقي و15/16 من مهام المحاكاة.

أكبر مكسب وأوضحه هو كفاءة البيانات. وصل LBM المضبوط دقيقًا إلى أداء يعادل التدريب من الصفر باستخدام **بيانات خاصة بالمهمة أقل بنحو 3–5 مرات.** وفي مهمة حقيقية واحدة — ترتيب مائدة إفطار — تفوق LBM ضُبط على **15% فقط** من العروض على سياسة مدربة من الصفر باستخدام **100%** منها.

يفيد التدريب المسبق أكثر عندما تتغير الظروف. عندما غُيرت بيئة الاختبار عمدًا بعيدًا عن ظروف التدريب — «انزياح التوزيع» — ازداد تفوق LBM المضبوط دقيقًا. وفي مجموعة محاكاة واحدة، تفوق إحصائيًا على التدريب من الصفر في 3 من 16 مهمة ضمن الظروف العادية، لكن في 10 من 16 تحت انزياح التوزيع. ولأن ظروف النشر الحقيقي تنحرف دائمًا عن التدريب، فقد تكون هذه المتانة أهم نتيجة عمليًا.

ساعد المزيد من بيانات التدريب المسبق، بصورة سلسة. ارتفع الأداء تدريجيًا مع إضافة بيانات التدريب المسبق، من دون **قفزة مفاجئة أو «انبثاق» لقدرة جديدة** عند المقاييس المختبرة. مفيد، قابل للتنبؤ، وغير درامي.

لكن قصة «نموذج عام بلا ضبط دقيق» لم تصمد. لم يتفوق LBM مدرّب مسبقًا عند استخدامه *zero-shot* — من دون ضبط خاص بالمهمة — باستمرار على سياسات المهمة الواحدة. أمكن لشبكة واحدة تنفيذ مهام كثيرة، لكن حلم «أعطه أمرًا فحسب» لم يتحقق هنا، ويعزو الباحثون جزءًا من ذلك إلى هشاشة المشفر اللغوي الصغير لديهم.

وكانت المكاسب صغيرة بما يكفي لتُفقد بسهولة — أو لتبدو موجودة وهي ليست كذلك. لم تصبح كثير من التأثيرات مرئية إلا بفضل أحجام

العينات الأكبر من المعتاد والاختبارات الدقيقة. ويقول الباحثون بوضوح إنه، بالنظر إلى حجم التأثيرات والضوضاء، **هناك خطر كبير من أن عددًا من أوراق الروبوتات يقيس ضوضاء إحصائية.** كما وجدوا أن خيارًا عاديًا جدًا — كيفية تطبيع البيانات — أثّر في النتائج أكثر من تغييرات المعمارية، وأن **خطأ برمجيًا** في التطبيع أثناء التدريب المسبق لم يظهر إلا بعد انتهاء التقييمات.

ماذا يعني هذا على الأرجح؟

القراءة التي يمكن الدفاع عنها هي أن التدريب المسبق واسع النطاق على بيانات روبوتية متنوعة عنصر حقيقي ومفيد: يقلل البيانات التي تحتاج إليها لكل مهمة جديدة، ويجعل السياسات أكثر صلابة عندما لا يطابق العالم ظروف التدريب. وهذا يدعم فعلًا الاتجاه الذي يراهن عليه المجال. لكن المكاسب متوسطة ومشروطة — تظهر في الغالب بعد الضبط الدقيق، وتكون أوضح عند التجميع وتحت الضغط — وليست وصول روبوت عام جاهز للاستخدام.

أما المعنى الأهدأ والأهم فهو منهجي. تعمل الدراسة عمليًا كمسطرة قياس: تبيّن مقدار الدليل المطلوب فعلًا لتقديم ادعاء موثوق عن سياسة روبوتية، وتوحي بأن قدرًا كبيرًا من الحماس المنشور قائم على أدلة أقل من اللازم. وهذا تصحيح يحتاجه المجال أكثر من نموذج جديد آخر.

ما الذي لا تثبته الدراسة؟

- هذا **ليس روبوتًا عامًا.** المكاسب مثبتة لمعمارية محددة — سياسات انتشار — تُضبط لكل مهمة، في ظروف مضبوطة، ومن عروض يتحكم بها البشر عن بعد، لا لروبوت مستقل ينفذ أي مهمة جديدة عند الطلب.
- لا تثبت صلاحية الاستخدام **zero-shot.** من دون الضبط الدقيق، لم يتفوق النموذج الكبير باستمرار على خطوط أساس المهمة الواحدة.
- هذه **ليست** أدلة على **«قفزة ناشئة».** حسّن التوسع الأداء بصورة سلسة؛ لا يوجد انقطاع يدعم رواية «ثم أصبح فجأة قادرًا».
- الأرقام **نسبية ومقيدة بالمختبر.** ضُبطت معدلات النجاح المطلقة عمدًا قرب ~50% لجعل المقارنات حساسة؛ وهي ليست مقياسًا للموثوقية في العالم الحقيقي، والعمل معمارية واحدة من مختبر واحد.
- لا تحسم **لماذا** تنجح سياسة أو تفشل، وتعرض الدراسة عدة مهام محددة كان أداء النموذج الكبير فيها أسوأ من دون تفسير السبب.
- ولا تقول **شيئًا عن السلامة أو الاستقلالية أو النشر** خارج منصة التقييم.

ما مدى قوة الأدلة؟

بالنسبة إلى الادعاءات المقارنة المركزية — أن LBM المضبوط دقيقًا يتفوق على خطوط الأساس المدربة من الصفر عند التجميع، ويحتاج بيانات أقل عدة مرات، ويكون أكثر متانة تحت انزياح التوزيع — فالأدلة قوية **ومضبوطة بصورة غير معتادة:** أعمى، عشوائي، بعينات كبيرة، واختبارات إحصائية، مع مرور لضمان جودة التقييم. هذه من الحالات النادرة التي تكون فيها المنهجية سليمة بما يكفي لأخذ الاستنتاجات الرئيسية قريبًا من ظاهرها.

أما التحفظات الصادقة فهي التي يثيرها الباحثون أنفسهم. أشرطة الخطأ لديهم تمثل عشوائية التقييم لكنها **لا** تمثل عشوائية التدريب — إذا دربت النموذج نفسه مرتين فقد تحصل على سياستين مختلفتين بصورة مهمة، وهذا التباين غير موجود في الإحصاءات. كان لكل مهمة حقيقية 50 تجربة، وهو عدد يكفي لالتقاط تأثيرات متوسطة لكنه قد يفوت الصغيرة. واستخدم التكييف اللغوي مشفرًا متواضعًا، لذلك قد تختلف

ادعاءات «قل للروبوت ما يفعل» في أنظمة أكبر. وهناك الكشف الصريح عن خطأ التطبيع بعد انتهاء التقييم. لا شيء من ذلك يهدم النتائج الرئيسية، لكنه بالضبط من نوع التفاصيل التي تقول الدراسة إن المجال يميل إلى تجاهلها.

وملاحظة مصدرية بالروح نفسها: يعتمد هذا الشرح على النسخة الأولية للباحثين. لم نتمكن من الحصول على نسخة المجلة المنشورة، ولذلك لم نتحقق من وجود تغييرات بين الpreprint والنص المنشور.

لماذا يهم هذا؟

عبارة «النماذج التأسيسية للروبوتات» تكاد تكون مصممة للمبالغة، ومن السهل إساءة قراءة دراسة كهذه في الاتجاهين — إما بانتصار «إنه يعمل!» أو بازدراء «مبالغ فيه». القراءة الأدق أكثر فائدة من كليهما: التدريب المسبق على بيانات متنوعة يقدم فوائد حقيقية قابلة للقياس لكنها متوسطة — خصوصًا بيانات أقل لكل مهمة ومتانة أكبر — والمسار يتحسن بصورة قابلة للتنبؤ مع زيادة المقياس.

والسبب الأعمق لأهميتها أن الدراسة توجه صرامتها إلى مجالها نفسه. بإظهار أن التأثيرات الحقيقية صغيرة بما يكفي لتختفي في تقييم رديء، وأن خيارًا مملًا مثل تطبيع البيانات قد يفوق أثر معمارية جديدة ذكية، فهي تدافع عن فكرة أن كثيرًا من «تقدم» تعلم الروبوتات يحتاج إلى قياس أصلب قبل تصديقه. ورقة تنفق مصداقيتها في حراسة الفرق بين نتيجة وأمنية تفعل شيئًا أندر وأكثر قيمة من تصدر لوحة ترتيب.

الخلاصة الواضحة

درّب باحثون في Toyota Research Institute «نماذج سلوك كبيرة» — سياسات روبوتية قائمة على الانتشار دُرّبت مسبقًا على نحو 1,700 ساعة من بيانات المناولة المتنوعة — وقارنوها بسياسات لمهمة واحدة مدربة من الصفر باستخدام بروتوكول صارم على نحو غير معتاد: أعمى وعشوائي وبعينات كبيرة، نحو 1,800 تجربة حقيقية وأكثر من 47,000 تجربة محاكاة، مع إحصاءات فعلية. بعد الضبط الدقيق لكل مهمة، كان أداء النماذج الكبيرة أفضل بصورة موثوقة عند التجميع، واحتاجت **بيانات خاصة بالمهمة أقل بنحو 3–5 مرات**، وكانت **أكثر متانة عند تغير الظروف**، مع تحسن سلس للأداء كلما زادت بيانات التدريب المسبق. لكن عند استخدامها **من دون ضبط دقيق** لم تتفوق باستمرار على نماذج المهمة الواحدة، وكانت عدة تأثيرات صغيرة إلى حد لم تكشفه إلا أحجام العينات الكبيرة، كما أن خيارًا عاديًا في تطبيع البيانات كان أهم من المعمارية. إنها أدلة قوية وموزونة لصالح اتجاه النماذج التأسيسية للروبوتات — **وليست** روبوتًا عامًا، ولا نموذجًا عامًا zero-shot، ولا «قفزة ناشئة» — إضافة إلى تحذير مباشر من أن كثيرًا من أبحاث الروبوتات قد تكون تقيس الضوضاء.

مراجعة بلا تهويل

ما الذي تُظهره الدراسة: باستخدام تقييم صارم وأعمى وذي قوة إحصائية — نحو 1,800 تشغيل حقيقي + أكثر من 47,000 تشغيل محاكاة — تتفوق سياسات الانتشار المدربة مسبقًا على مهام متعددة ثم المضبوطة دقيقًا (LBMs) على سياسات المهمة الواحدة المدربة من الصفر عند التجميع، وتصل إلى أداء مماثل ببيانات خاصة بالمهمة أقل بنحو 3–5 مرات، وتكون أكثر متانة تحت انزياح التوزيع؛ كما يتحسن الأداء بسلاسة مع زيادة بيانات التدريب المسبق.

ما هو معقول لكنه غير مثبت: أن تنتقل هذه الفوائد إلى نماذج رؤية-لغة-فعل أكبر بكثير — كان المشفر اللغوي لديهم صغيرًا — وأن يستمر

التحسن السلس خارج نطاق البيانات المختبر.

ما الذي لا تُظهره: روبوتًا عامًا أو zero-shot — لا ضبط دقيق يعني لا تفوقًا ثابتًا — ولا قفزة قدرة «ناشئة»، ولا تفسيرًا لإخفاقات مهام بعينها، ولا موثوقية مطلقة في العالم الحقيقي — فقد ضُبطت معدلات النجاح قرب 50% لزيادة الحساسية — ولا أي شيء عن السلامة أو النشر المستقل.

القيود الرئيسية: تمثل الإحصاءات عشوائية التقييم لا التباين بين دورات التدريب، وقد تفوت 50 تجربة حقيقية لكل مهمة التأثيرات الصغيرة؛ معمارية واحدة ومختبر واحد، مشفر لغوي متواضع؛ اكتُشف خطأ في تطبيع البيانات بعد التقييم، والتحليل مبني على الـpreprint، إذ لم تُراجع النسخة المنشورة.

ما مقدار الثقة المناسب للقارئ العام؟ ثقة عالية بأن التدريب المسبق متعدد المهام مع الضبط الدقيق يقدم فوائد حقيقية ومتوسطة — وخاصة في كفاءة البيانات والمتانة — وأنها قِيست بعناية غير معتادة. وثقة عالية بأن هذا **ليس** روبوتًا عامًا أو zero-shot **وليس** قفزة ناشئة. وثقة متوسطة في مدى استمرار المكاسب مع نماذج أكبر. ويستحق تحذير الباحثين نفسه أن يؤخذ بجدية: تأثيرات المجال صغيرة بما يكفي لأن الدراسات ضعيفة القوة الإحصائية قد تبلغ عن ضوضاء. الموقف المناسب: تفاؤل موزون تجاه النهج، وشك صحي في نتائج روبوتات الذكاء الاصطناعي التي تفتقر إلى هذا النوع من الأساس الإحصائي.

المصادر

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

scirobotics.aea6201/1126.10 Robotics, Science — retrieved) (not version Peer-reviewed
https://doi.org/10.1126/scirobotics.aea6201

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