Опитни разработчици се чувствали по-бързи с AI, докато измеримо работели по-бавно — истинският резултат и присъдата за AI програмирането, която той не е

В рандомизирано контролирано изпитване на METR шестнадесет опитни разработчици на софтуер с отворен код изпълняват 246 реални задачи в кодови бази, които познават добре, като при случайно избрана половина от задачите могат да използват AI инструменти от началото на 2025 г. (Cursor Pro плюс Claude 3.5/3.7 Sonnet). Те очакват AI да намали времето за задача с около 24%; вместо това времето за изпълнение се увеличава с 19% — а след експеримента те продължават да вярват, че AI ги е ускорил с около 20%. Тази разлика между усещане и измерване е най-ясният и устойчив резултат от изследването. Но става дума за шестнадесет разработчици, тесен контекст и фиксирана снимка на началото на 2025 г.: авторите изрично казват, че това не показва, че AI не помага на повечето разработчици, а собственото им проследяване от 2026 г. вече клони към ускорение, макар и със свои уговорки.

Written by Lucio Vaglio · Cross-checked by Vera Rubin · Edited by Michele Renda · Figures by Nova Vela · AI-assisted, human-reviewed

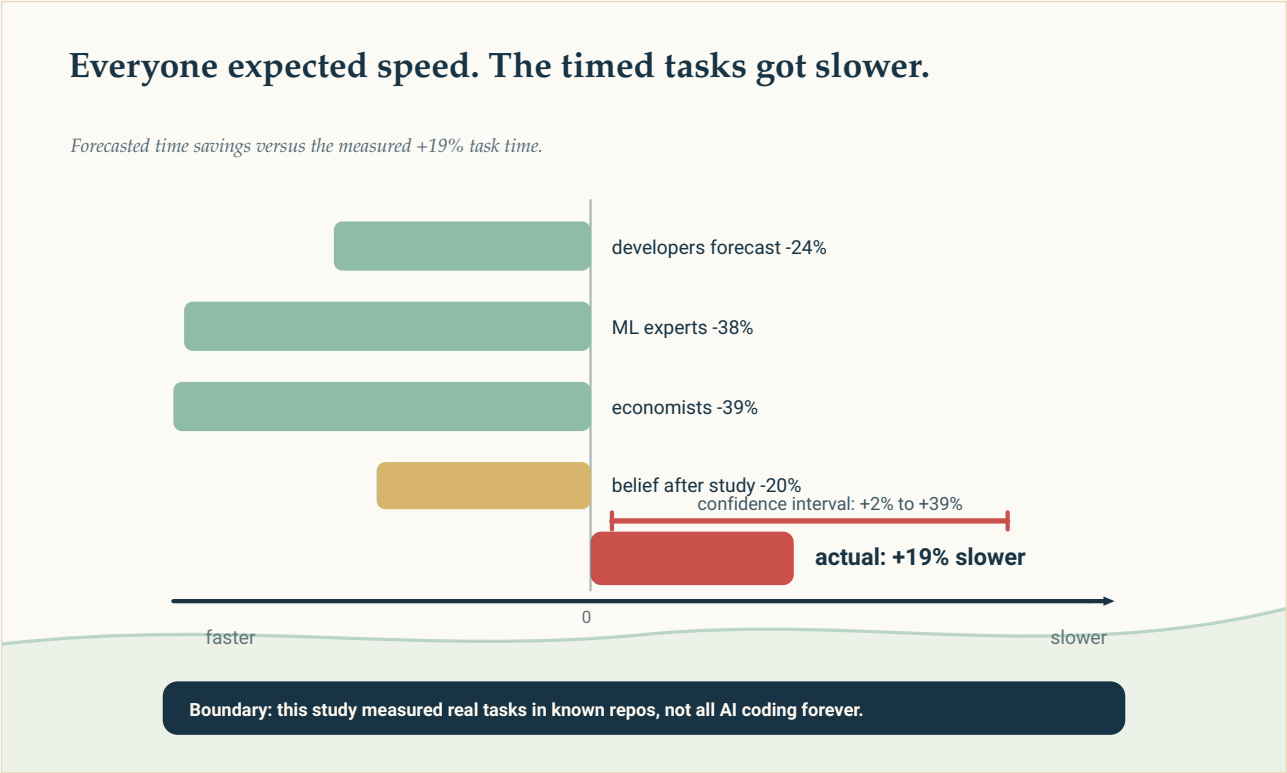
Paper: *Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity*, Joel Becker, Nate Rush, Beth Barnes, David Rein (METR), arXiv:2507.09089 [cs.AI] (preprint) · DOI: 10.48550/arXiv.2507.09089

Опитни разработчици се чувствали по-бързи с AI, докато измеримо работели по-бавно — истинският резултат и присъдата за AI програмирането, която той не е

Попитайте опитен програмист дали AI помощник за писане на код го ускорява и обикновено ще получите число: спестява ми двайсет, трийсет процента. Попитайте икономист или изследовател по машинно обучение и числото става още по-голямо. В началото на 2025 г. екип от METR прави бавната и скъпа проверка. Взема шестнадесет опитни разработчици на софтуер с отворен код, дава им 246 реални задачи от големи кодови бази, които познават добре, и — задача по задача, чрез случайно разпределение — или им позволява да използват AI инструменти, или не. След това измерва времето.

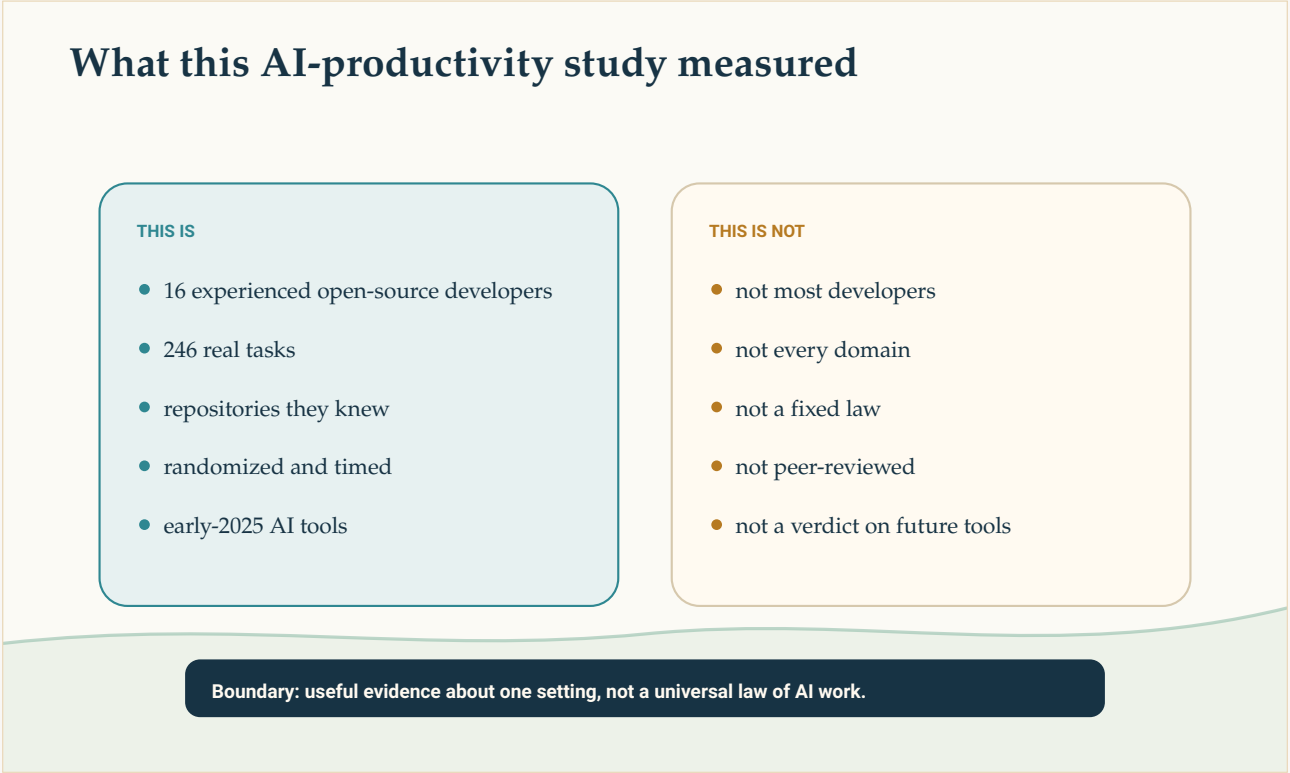
Разработчиците предварително прогнозират, че AI ще намали времето им за задача с около 24%. Случва се обратното: задачите, изпълнявани с AI, отнемат **19% повече време**. А ето частта, върху която си струва да се замислим — след приключването същите разработчици продължават да вярват, че AI ги е ускорил с около 20%. Работели са по-бавно, но са се чувствали по-бързи, и разстоянието между тези две числа е най-интересното в изследването.

Това е реален, внимателно измерен резултат. Но това са и само шестнадесет разработчици, работещи в хранилища, които познават в детайли, с инструментите от началото на 2025 г. — и резултатът не се свежда до плоското твърдение „AI прави разработчиците по-бавни“. По-късните данни на същия екип вече сочат в другата посока.



Всички прогнозират, че AI ще ускори работата — разработчиците с −24%, експертите по машинно обучение с −38%, икономистите с −39%, а собствената оценка на разработчиците след експеримента е −20%. Хронометърът показва +19% по-бавно, с интервал от +2% до +39%. Разликата между усещането и измерването е същинският резултат.

Original diagram — The Clean Paper CC BY 4.0



Какво измерва изследването — 16 експертни разработчици на софтуер с отворен код, 246 реални задачи, познати хранилища, инструменти от началото на 2025 г., рандомизация — и какво не измерва: не повечето разработчици, не други области, не неизменен закон и не рецензиран резултат.

Original diagram — The Clean Paper CC BY 4.0

Какво е измерено тук и какво ви дава едно рандомизирано изпитване

Рандомизираното контролирано изпитване (RCT) е инструментът, който медицината използва, за да различи реален ефект от такъв, на който просто се надяваме. Тук всяка от 246-те задачи е разпределена на случаен принцип за изпълнение с разрешен или забранен AI, така че средно единствената системна разлика между двете групи е самият AI. Това позволява да се каже, че AI *е причинил* промяната във времето, вместо просто да се забележи, че хората, които посягат към AI, по други причини са по-бързи или по-бавни. Това е важно, защото обичайните доказателства за печалба от AI при програмиране — самооценки и резултати от бенчмаркове — не могат да направят това: бенчмаркът не е реална работа, а самооценката, както показва това изследване, може уверено да греши. Разработчиците тук не са начинаещи, които се мъчат с нова играчка. Те са

> утвърдени сътрудници в големи и зрели проекти с отворен код, които познават добре, и вече имат известен предишен опит с инструментите.

Какво са направили авторите

- Провеждат **рандомизирано контролирано изпитване** (METR: Joel Becker, Nate Rush, Beth Barnes, David Rein). Шестнадесет опитни разработчици на софтуер с отворен код, всеки работещ в голямо хранилище, за което редовно допринася и което познава добре — средно по пет години опит с конкретните проекти.
- Използват **246 реални задачи** — поправки на бъгове, нови функции и рефакториране, взети от собствените системи за задачи на проектите. Всяка задача е разпределена на случаен принцип като „AI разрешен“ или „AI забранен“.
- „AI разрешен“ означава **инструменти от началото на 2025 г.: Cursor Pro с Claude 3.5/3.7 Sonnet.** Основната мярка е действителното време за изпълнение на всяка задача. Наред с него екипът събира предварителни прогнози от разработчиците, оценки след изпълнението, както и прогнози от експерти по икономика и машинно обучение.

Какво са установили

- **С AI задачите отнемат 19% повече време.** Не по-бързо — по-бавно. 95-процентният доверителен интервал е приблизително от +2% до +39%, така че посоката е стабилна, дори точният размер да остава несигурен.
- **Всички са предвидили обратното.** Разработчиците прогнозират 24% ускорение; експертите по машинно обучение — около 38%; икономистите — около 39%. И трите групи очакват AI да спести много време; хронометърът показва, че то всъщност струва време.
- **Разликата във възприятието.** След като свършват работата по-бавно, разработчиците все още оценяват, че AI ги е ускорил с около 20% — разлика от приблизително 40 процентни пункта между усещането и измереното време.
- **Възможни причини, претеглени, а не доказани.** Авторите посочват фактори, които биха могли да обяснят забавянето: тези разработчици познават собствените си кодови бази много добре, така че помощникът има по-малко какво да добави; зрелите проекти имат високи, често неявни стандарти за качество; хранилищата са големи и пълни с контекст, който моделът няма; а реално време се изразходва за подсказки, после за преглед и поправяне на AI изхода. Това са насоки, не окончателни присъди.

Какво не показва това

- То **не** показва, че AI не успява да ускори повечето разработчици. Авторите го казват пряко: шестнадесет експерти, работещи върху код, който познават почти наизуст, не са средният разработчик със средната задача.
- То **не** показва, че AI е безполезен или че забавя хората в други ситуации — новодошли в кодова база, нови проекти от нулата, непознати езици или изцяло други области.
- То **не** замразява инструментите във времето. Това са Cursor и Claude 3.5/3.7 Sonnet от началото на 2025 г.; авторите изрично отбелязват, че по-добри инструменти или по-добри начини за използването им могат да променят резултата дори в същата среда.
- Това е **препринт** (публикуван през юли 2025 г., все още нерецензиран) и авторите отбелязват, че не могат напълно да изключат експериментални артефакти — макар резултатът да се запазва при различните им анализи.
- То **не** оправдава успокояващото тълкуване, че разработчиците сигурно са спечелили по някакъв друг начин — научили са повече, чувствали са се по-добре, написали са по-качествен код. Единственото подобно нещо, измерено тук — усещането за ускорение — е точно онова, на което данните противоречат.

Колко силни са доказателствата

- **Дизайнът е необичайно честен.** Рандомизация, реални задачи, реални хранилища, действително измерване на времето — голяма крачка напред спрямо самооценките и класациите от бенчмаркове, върху които се опират повечето твърдения за AI програмиране. Забавянето от 19% остава при проверките за устойчивост на авторите.
- **Най-преносимият резултат е разликата във възприятието.** Експертната интуиция за собственото ускорение с AI греши с около 40 пункта, при това в оптимистичната посока. Това е предупреждение за *всяка* самоотчетена печалба в продуктивност от AI — включително числата в самото това изследване.
- **Това е моментна снимка, не тенденция.** Собственото проследяване на METR от февруари 2026 г. със същия тип разработчици, но с по-нови инструменти, сочи към ускорение — много приблизително −18% за завръщащи се разработчици и −4% за нови — но авторите го определят като слабо доказателство, изкривено от това кой е бил склонен да участва (разработчиците все по-често отказват да работят без AI, заплащането е намалено, а подборът на задачи се измества). Честният прочит е, че картината се движи и дори движението е докладвано предпазливо.

Защо е важно

Голяма част от спора за AI и програмирането се движи от демонстрации и усещания: лъскав запис на екрана, уверено твърдение, насрещно завъртане на очи. Рядкото — и причината това изследване да си струва четенето — е, че някой е провел скучния експеримент: рандомизирай, измери времето за реална работа, после попитай хората как е минало. Отговорът е неудобен и за двата лагера. Той пробива разказа, че AI еднакво турбоускорява експертните разработчици върху труден и познат код. Но пробива и подредената противоположност — „доказано е, че AI прави разработчиците по-бавни“ — защото по-новите данни на същия екип вече клонят в другата посока. Най-трайният урок е и най-малкият, и най-човешкият: хората, които вършат работата, се чувстват по-бързи, докато измеримо са по-бавни. „Усеща се по-бързо“ не е доказателство, че е така. Измерете го.

Кратко обобщение

В рандомизирано контролирано изпитване METR възлага на шестнадесет опитни разработчици на софтуер с отворен код да изпълнят 246 реални задачи в кодови бази, които познават добре, като за случайно избрана половина от задачите са разрешени AI инструменти от началото на 2025 г. (Cursor Pro плюс Claude 3.5/3.7 Sonnet). Разработчиците очакват AI да намали времето им за задача с около 24%; вместо това то се увеличава с 19% — а след това те все още вярват, че AI ги е ускорил с около 20%. Тази разлика между възприятие и измерване е ясното и устойчиво ядро на изследването. Но това са шестнадесет разработчици, една тясна среда и фиксирана снимка на началото на 2025 г.; авторите изрично казват, че резултатът не показва, че AI не помага на повечето разработчици, а собственото им проследяване от 2026 г. вече сочи към ускорение, със свои уговорки. Внимателно измерване, което заслужава да бъде взето насериозно — не присъда за AI програмирането.

Източници

J. Becker, N. Rush, B. Barnes, D. Rein (METR), Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity, arXiv:2507.09089 [cs.AI] (2025)
https://arxiv.org/abs/2507.09089

METR, ‘Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity’ (study write-up, July 2025)
https://metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study/

METR, developer-uplift follow-up (February 2026)
https://metr.org/blog/2026-02-24-uplift-update/