RESEARCH REPORT · NOVEMBER 2026

When AI Does All The Labor

A conditional timeline from the Computational Transition Model

Eli Mikel
Independent Researcher; Founder and CEO, Clockwise Capital, LLC, Miami, FL

MEDIAN ESTIMATE

2037

AI performs the majority of compensation-weighted labor in advanced economies. 80% interval: 2031 to 2051.

2034	2045	No date
Most cognitive work	Most physical work	All work, with no indispensable human anywhere in the loop

Conditional model output, not a forecast of record. See Sections 7 and 9 and the disclosures.

SUMMARY

The short version

The usual question is how fast AI gets smarter. The better question is how fast machine interpretation becomes sufficient relative to the organizations it has to run. Capability is racing ahead. Integration and physical capacity are not, and the slowest of them sets the date.

1. **"All the labor" never gets a date.** Under the model, a complete handoff is a limit that no finite observation can confirm. What can be dated are majority milestones, and they arrive in sequence.
2. **Capability is close to settled as a constraint.** At METR's measured doubling rates, frontier systems likely handle month-long software tasks at 50% reliability during 2027 and year-long tasks at 80% reliability around 2029.
3. **The real lag is integration.** In October 2026 about one in four US businesses reported using AI in the prior two weeks, and most adopters use it in a handful of functions. Authority, memory and liability have to be rebuilt firm by firm.
4. **Physical work binds the system.** Roughly 13,000 humanoid robots shipped worldwide in 2025. A physical-labor majority in high-income economies plausibly needs on the order of 100 million general-purpose machines. The central scenario gets there around 2045.
5. **The headline estimate is 2037.** That is the median year by which AI performs most compensation-weighted labor in advanced economies, with an 80% interval of 2031 to 2051. The probability is about 2% by 2030, 36% by 2035 and 66% by 2040.
6. **The most valuable unmeasured number is ρ.** It sets how well the economy routes around weak stages. High substitution pulls the date toward the mid-2030s. Near-serial dependence pushes it toward the mid-2040s.

Table 1. The milestone sequence

Ranges for M1 and M2 span the fast and slow capability scenarios. Ranges for M3 to M5 are 80% intervals, rounded outward to whole years.

Milestone	Central	Range	Gated by
M1 Frontier AI completes month-long software tasks at 50% reliability	Q2 2027	Q1 2027 to Q1 2028	Capability
M2 Frontier AI completes year-long tasks at 80% reliability	Q3 2029	Q4 2028 to Q3 2031	Capability
M3 AI performs most cognitive labor value in advanced economies	2034	2030 to 2042	Integration
M4 AI performs most of all labor value in advanced economies	**2037**	**2031 to 2051**	Integration and physical
M5 Machines perform most physical labor in advanced economies	2045	2037 to 2058	Physical
M6 No indispensable human anywhere in the loop	No date	Limit case	Every stage

SECTION 1

The question, restated

“When will AI do all the labor?” sounds like one question. It is at least three.

Capability is one: what a frontier system can do on a well-specified task. Deployment is another: whether that system holds authority, memory and permissions inside a real organization. Physical execution is a third. Each moves at its own rate, and the slowest one that cannot be routed around sets the pace for the whole.

This report answers with the Computational Transition Model (ACTM; Mikel 2026). ACTM does not forecast compute. It asks when effective machine interpreter capacity outruns the complexity of what it must reconstruct. That framing produces dates as conditional outputs of stated assumptions, which is the only honest form a date like this can take.

Four definitions hold throughout.

- **Labor value** is compensation-weighted work. A surgeon’s hour counts for more than a cashier’s.
- **Advanced economies** means high-income economies broadly. Others likely lag, for reasons outside this model.
- **AI performs** the work when software, machines or both do it without a human in an indispensable interpreting role for that work. People may still supervise, own, or be employed elsewhere.
- **Majority** means more than half. “All” is treated separately, because the model says it behaves differently.

SECTION 2

The model in one page

Four quantities carry the argument.

Quantity	Meaning
$C(t)$	Raw machine capability. What frontier systems can do in isolation.
$I_A(t) = \Omega[C, A_r]$	Effective interpreter capacity. Capability filtered through integration architecture: access to canonical state, authority to act, persistent memory, portability, auditability.
$K(t)$	Reconstructive complexity. The burden of rebuilding the organization’s working form, not its size.
$\chi = I_A / K^{\Upsilon}$	Relative interpreter sufficiency. Capacity measured against the burden it carries.

$$d \ln \chi / dt = g_I - \Upsilon g_K$$

When interpreter capacity grows faster than complexity-adjusted burden, sufficiency rises. When it grows slower, sufficiency falls, even while capability improves. Absolute progress and a stalled transition can coexist. That is the model’s core claim.

Version 2 of the model decomposes χ across seven reconstruction stages: access, interpretation, variation, evaluation, instantiation, maintenance and recursive incorporation. Stage-level closure C_s

aggregates through a substitution parameter ρ:

$$C(\tau) = \left(\Sigma_s \, w_s \, C_s(\tau)^{\rho} \right)^{1/\rho}, \; \rho \leq 1$$

As ρ falls toward negative infinity, the weakest stage binds everything. At ρ = 1, stages substitute freely. In the near-serial case the binding stage, the one with the lowest sufficiency, governs the system, and it moves over time because stages improve at different rates.

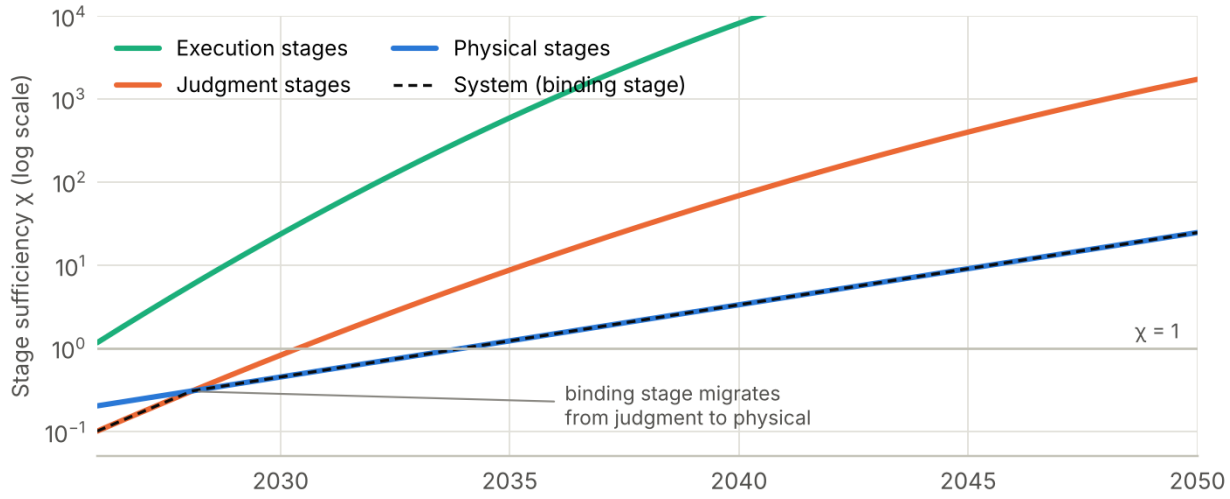


Figure 1. Stylized stage sufficiency. Execution stages (interpretation, variation, software build) improve fastest. Judgment stages (evaluation, goal-setting) start low but improve quickly. Physical stages (physical build, maintenance) improve slowly and become the binding constraint. The dashed line is the system under near-serial aggregation. Parameters are illustrative, not measured.

Three consequences follow for the labor question.

1. Capability benchmarks mostly measure execution stages. They say little about the binding stage.
2. The economy-wide date is set by the slowest stage the economy cannot route around.
3. A single "AI takes over" date is the wrong unit. The model yields a sequence.

SECTION 3

Input one: capability

METR's task-completion time horizon is the best public series for execution-stage capability. It measures the length of task, in human-expert time, that a model completes at a stated success rate (Kwa et al. 2025).

The anchor comes from METR's Frontier Risk Report for February to March 2026. The public frontier sat near 12 hours at 50% success. The internal frontier likely sat at 16 hours or more at 50%, and between 3 and 4 hours at 80% (METR 2026b). We anchor at 16 hours and 3.1 hours in March 2026.

Three doubling rates are on the table, each reported by METR: about 212 days across 2019 to 2025, 129 days from 2023 onward, and 105 days for frontier models released after January 2024 (Kwa et al. 2025; METR 2026b, 2026c). We use 129 days as central and the other two as bounds.

Table 2. Projected crossing dates

Threshold	105-day doubling	129-day (central)	212-day doubling
50% horizon reaches a work week (40 h)	Q3 2026	Q3 2026	Q4 2026
50% horizon reaches a work month (167 h)	Q1 2027	**Q2 2027**	Q1 2028
50% horizon reaches a work year (2,000 h)	Q1 2028	Q3 2028	Q1 2030
80% horizon reaches a work week	Q2 2027	Q3 2027	Q2 2028
80% horizon reaches a work month	Q4 2027	Q1 2028	Q3 2029
80% horizon reaches a work year	Q4 2028	**Q3 2029**	Q3 2031

Projected, not measured. The 40-hour crossings fall before this report's date but cannot be confirmed with current instruments.

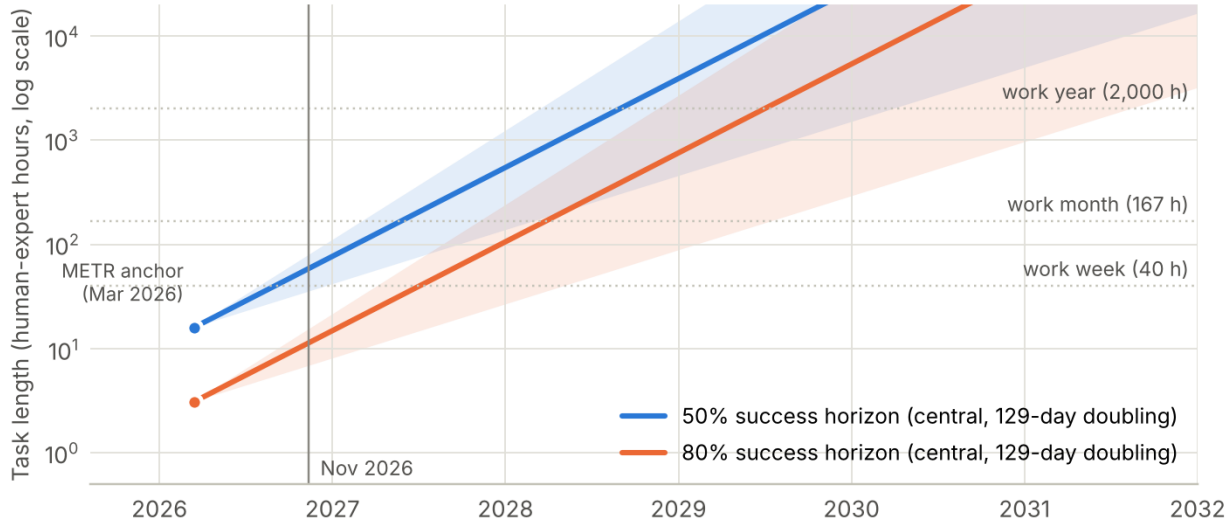


Figure 2. Projected frontier time horizons from the March 2026 METR anchor. Solid lines use the 129-day doubling. Shaded bands span the 105-day and 212-day fits. Everything to the right of the anchor is extrapolation.

Three cautions keep this from being the answer.

- **The instrument is saturating.** METR states that its Time Horizon 1.1 suite cannot reliably measure horizons above 16 hours, and its time-horizons page stopped active updates in September 2026 (METR 2026b, 2026c). On the central path the frontier would sit near 58 hours today, beyond what the instrument can see.
- **Benchmarks flatter.** Agents are strongest where progress is cheap to verify and much weaker on open-ended work that needs judgment. METR found that only about half of AI-generated SWE-bench Verified solutions graded as passing in late 2025 would have been accepted in real code review (METR 2026b).
- **Execution is not judgment.** In METR's assessment, no company relied on agents to set research agendas, make hiring or budget decisions, or merge critical code without human review (METR 2026b). Those are evaluation-stage functions, exactly where ACTM expects the binding stage to sit first.

The conclusion is narrow. On execution stages g_I is so large that the convergent regime is close to certain. Capability is unlikely to be what holds the transition back. The constraint lives elsewhere.

SECTION 4

Input two: integration

ACTM's central distinction is between C and I_A. A model that cannot hold authority, persistent state and permissions inside an organization contributes little to how that organization reproduces itself. Integration is where most of the calendar goes.

The adoption data say integration is early.

- In the Census survey wave of October 5 to 18, 2026, 24.6% of US businesses reported using AI in the prior two weeks, up from about 17% when the question was broadened in November 2025 (US Census Bureau 2026a).
- In the Census AI supplement covering November 2025 to January 2026, 18% of firms used AI in a business function, or 32% weighted by employment. Among users, 57% used it in three or fewer functions (US Census Bureau 2026b).
- Among firms that replaced tasks with AI, 71% did so for only a small number of tasks (US Census Bureau 2026b).

Even inside AI developers, integration is narrow at the top. Agents write code, run experiments and analyze data with substantial autonomy. They do not set direction or hold final authority (METR 2026b).

History supports a lag. Electricity took decades to show up in productivity because factories had to be redesigned around it (David 1990). General purpose technologies tend to trace a productivity J-curve: heavy intangible investment first, measured gains later (Brynjolfsson et al. 2021). AI should move faster than either, since it ships over existing software rails and needs no new grid. It should still move slower than benchmarks imply.

Our prior is 3 to 7 years from execution-stage capability to majority deployment in cognitive work. With M2 near 2029, that puts M3, AI performing most cognitive labor value, at a median of 2034 with an 80% interval of 2030 to 2042.

One structural point deserves emphasis. In ACTM, lowering K works as well as raising I_A. Firms that modularize, standardize and encode their rules in machine-readable form will likely cross earlier than firms running the same models on messier operations. The transition is probably distributed across organizations rather than synchronized across the economy.

SECTION 5

Input three: physical work

Physical work is where the binding stage most likely ends up.

The installed base is small. Omdia counted roughly 13,000 humanoid robots shipped worldwide in 2025 (Omdia 2026), and TrendForce projects more than 50,000 in 2026 (TrendForce 2026). Conventional industrial robots are a larger base, about 4.66 million operational units in 2024, but that stock has grown about 11% a year since 2019 and consists mostly of single-purpose machines (IFR 2025).

The arithmetic

High-income economies employ roughly 600 million people. If something like 40% of that work is predominantly physical, a majority means about 120 million worker-equivalents. If each general-purpose machine substitutes for 1.5 to 2 workers through multi-shift operation, the requirement is 60 to 80 million units deployed in those economies. With production and deployment spread globally, a global stock on the order of 100 million is the right marker. These are order-of-magnitude assumptions, stated so they can be attacked.

The scenarios

We start from about 50,000 installed units at the end of 2026 and let annual stock growth decay from an early-phase rate toward a long-run floor.

Scenario	Initial growth	Long-run floor	Decay (years)	Reaches 10M	Reaches 100M
Fast	150%	35%	6	2033	2038
Central	120%	25%	5	2036	**2045**
Slow	80%	15%	4	2048	2064

Table 3. Robot stock scenarios. Growth parameters are assumptions, not estimates.

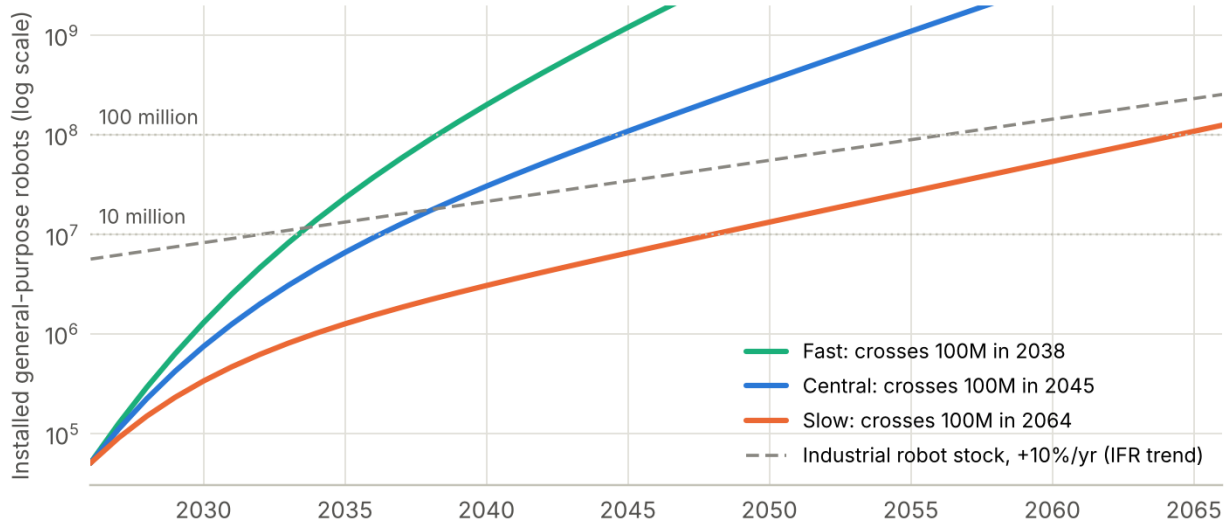


Figure 3. Installed general-purpose robot stock under three growth scenarios. The dashed line extends the industrial robot stock at 10% a year for comparison.

The central path is not exotic, but it is demanding. In the year it crosses 100 million it implies about 24 million units added, roughly a quarter of current global vehicle output. The fast path needs about 45 million a year by 2038. Both require actuator, battery and assembly supply chains that do not yet exist at that scale, plus site-by-site integration that today looks more like construction than software.

This puts M5, machines performing most physical labor in advanced economies, at a median of 2045 with an 80% interval of 2037 to 2058. The median is the central scenario’s 100 million crossing.

SECTION 6

Input four: substitution

The model's largest open parameter is ρ. It decides which of the dates above is the real answer to the headline question.

Take five informational stages at a closure of 0.9 and two physical stages at 0.2, weighted equally. Only ρ changes across the rows.

ρ	Reading	System closure
1	Full substitution	0.70
0	Partial substitution (geometric mean)	0.59
−1	Limited substitution	0.45
−5	Near serial	0.26
−∞	Strict weakest link	0.20

Table 4. Same stages, same capabilities. System closure runs from 0.20 to 0.70 on ρ alone.

Across a whole economy ρ is probably higher than inside any single production process. Value can shift toward cognitive output, services can be redesigned to need less physical presence, and consumption can reallocate. That argues the labor-value date sits closer to the cognitive date than to the physical one.

Inside specific supply chains ρ is probably low. A logistics network with a fully automated planning layer and human loading docks still moves at the speed of the docks.

The composition matters too. If cognitive work accounts for something like 60% of labor compensation in advanced economies, a labor-value majority with no physical progress at all would require automating more than 80% of cognitive value. The realistic path combines deep cognitive coverage with partial physical coverage, which is why M4 lands between M3 and M5.

Our M4 median of 2037 assumes a moderate ρ, well above the serial limit and well below full substitution. Nobody has measured it. Estimating ρ from architectural variation across sibling organizations may be the single most valuable empirical project this framework points to.

SECTION 7

The estimate

Combining the four inputs gives a sequence rather than a single date.

Milestone	Median	80% interval	By 2030	2035	2040	2045	2050
M3 Cognitive work	2034	2030 to 2042	6%	60%	87%	96%	99%
M4 All labor value	**2037**	**2031 to 2051**	2%	36%	66%	81%	89%
M5 Physical work	2045	2037 to 2058	<1%	2%	23%	52%	74%

Table 5. Probability that each milestone is reached by the given year.

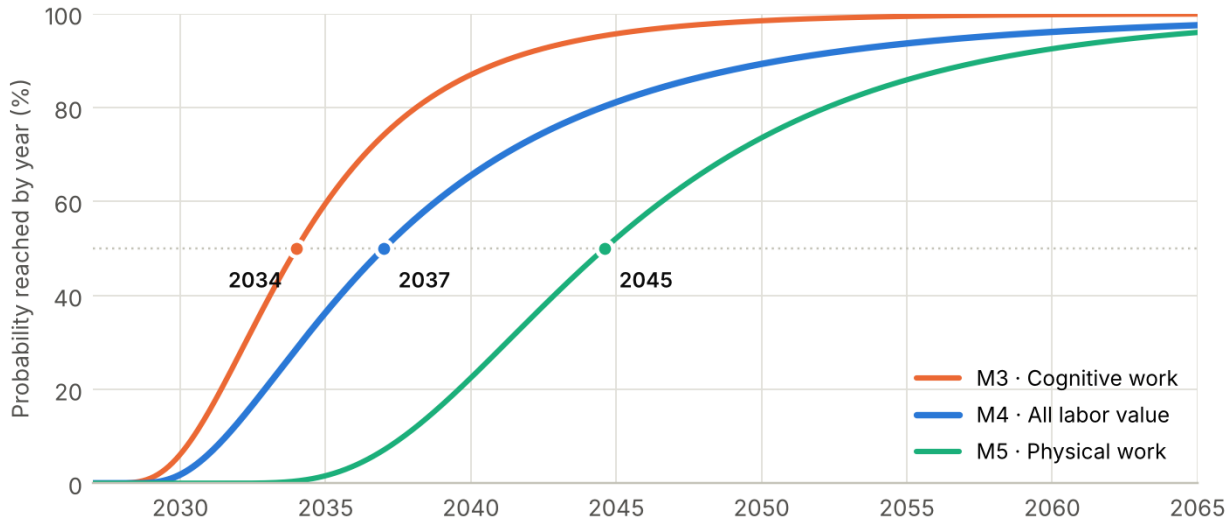


Figure 4. Cumulative probability of each milestone. Dots mark the medians. The distributions are shifted lognormals whose parameters are listed in Appendix A.

Reading the headline

The best single answer to "when does AI do the labor" is 2037: more likely than not, AI performs most compensation-weighted work in advanced economies by then. The probability passes one in three around 2035 and four in five around 2045.

The answer to "when does AI do *all* of it" is that the question has no finite date under the model. Residual human dependence at a stage over τ generations is $1 - p_s^{\tau}$. For any per-generation closure below one, that figure climbs toward one as the horizon lengthens. Full independence requires sufficiency to grow without bound at every indispensable stage. It is a limit, not an event.

What sits in the tails

The left tail, before 2031, is where fast-takeoff scenarios live: AI systems accelerating their own integration and robot production. ACTM does not rule them out. It says they need g_l to jump at the integration and physical stages, not only at execution stages, and that is a checkable claim.

The right tail, past 2051, holds capability plateaus, legal limits on machine authority and supply-chain ceilings on physical scale. Any one of them can keep the system in the divergent regime for years.

SECTION 8

Signposts

Each input has an observable. The table lists what would move the estimate and roughly how far. Magnitudes are judgments, not computations.

Signal	What it would mean	Effect on M4
Measured horizon doubling stretches past 7 months	Execution capability slowing	Later by 3 to 5 years
Doubling holds near 105 days through 2027	Fast capability path	Earlier by about 1 year

Signal	What it would mean	Effect on M4
Census AI use passes half of firms, with use spread across many functions	Integration accelerating	Earlier by 2 to 3 years
Firms give agents final authority over budgets, hiring or production code	Judgment stage closing	Earlier by 2 to 4 years
Humanoid shipments exceed 1 million units in a single year before 2030	Fast physical path	Earlier by 1 to 2 years (M5 by 3 to 5)
Liability or licensing rules bar machine authority in regulated work	Architecture constrained by law	Later by 3 to 10 years
Withdrawal tests show a stable floor in human dependence	A possible hard limit	Undated; challenges the model itself

Two signals deserve more weight than the rest. The capability series has lost its main instrument, so watch for METR's successor benchmarks and for real-world task evidence rather than leaderboards. And the Census survey asks only whether a firm uses AI, not how deeply, so the supplement's scope-of-use questions matter more than the headline rate.

SECTION 9

What this report does not claim

- **It is not an unemployment forecast.** AI performing most labor value does not mean most people stop working. Hours, wages and new categories of work are separate questions.
- **Labor is not control.** Who owns the productive system, and who holds claims on its output, is analytically separate from who or what performs the work.
- **It is not a measurement.** None of I_A, K, γ, ρ or the stage closures has been estimated for any economy. Every date here is a conditional output of stated priors.
- **It is not global.** The estimates cover advanced economies. Others likely lag.
- **It is not a benchmark extrapolation.** METR's suite measures software tasks and is saturating. It anchors one input of four.

Limitations

The integration lag is a prior, not an estimate, and it carries most of the uncertainty in M3. The robot arithmetic rests on assumptions about workforce composition and substitution ratios that could each be off by half. The milestone distributions treat the inputs as roughly independent when they are not: fast capability likely accelerates robot development, and legal limits on authority would slow both cognitive and physical deployment. Correlated shocks would fatten both tails.

The model's value is in what it makes measurable. If a preregistered estimate of χ predicts which organizations cross first, and withdrawal designs locate the binding stage where the model says it should be, the dates here can be replaced with estimates. If not, the model should be revised, and the dates with it.

APPENDIX A

Equations and parameters

Core model

$$I_A(t) = \Omega[C(t), A_r(t)]$$

$$\chi(t) = \chi_0 \cdot (I_A/I_{A,0}) / (K/K_0)^{\gamma}, \; d \ln \chi/dt = g_I - \gamma g_K$$

$$p_s = \chi_s^{\beta} / (1 + \chi_s^{\beta}), \; D_s(\tau) = 1 - p_s^{\tau}$$

$$C(\tau) = (\Sigma_s \, w_s \, C_s(\tau)^{\rho})^{1/\rho}, \; L_\varepsilon = \max\{\tau : 1 - C(\tau) < \varepsilon\}$$

Computations in this report

$$\text{Capability crossing: } t^* = t_0 + \log_2(H^* / H_0) \cdot d / 365.25$$

$$\text{Robot stock: } S_{y+1} = S_y \, [1 + g_f + (g_0 - g_f) \, e^{-(y - 2026)/\tau}]$$

$$\text{Milestones: } T = s + \exp(\ln m + \sigma Z), \; Z \sim N(0, 1)$$

Parameter	Value	Basis
50% horizon anchor H_0	16 h, March 2026	METR (2026b), internal frontier, likely at least 16 h
80% horizon anchor H_0	3.1 h, March 2026	METR (2026b), low end of the 3 to 4 h range
Doubling time d	105 / 129 / 212 days	METR fits, post-2024 / post-2023 / 2019 to 2025
Initial robot stock S_{2026}	50,000 units	Assumption from 2025 shipments and 2026 projections
Robot growth g_0, g_f, τ	See Table 3	Assumption
Physical majority marker	100 million units	Order-of-magnitude arithmetic, Section 5
M3: s, m, σ	2027, 7.0, 0.55	Prior: 3 to 7 year integration lag after M2
M4: s, m, σ	2028, 9.0, 0.72	Prior: moderate ρ, between M3 and M5
M5: s, m, σ	2030, 14.6, 0.50	Median at central 100M crossing
CES illustration	0.9 × 5, 0.2 × 2, equal weights	Illustrative

Ordering check: across 2027 to 2075 the cumulative probability of M3 is never below M4, and M4 is never below M5.

References

Brynjolfsson E, Rock D, Syverson C (2021) The productivity J-curve: how intangibles complement general purpose technologies. American Economic Journal: Macroeconomics 13(1):333–372. https://doi.org/10.1257/mac.20180386

David PA (1990) The dynamo and the computer: an historical perspective on the modern productivity paradox. American Economic Review 80(2):355–361

IFR (2025) World Robotics 2025: Industrial Robots, executive summary. International Federation of Robotics, Frankfurt. https://ifr.org/img/worldrobotics/Executive_Summary_WR_2025_Industrial_Robots.pdf

Kwa T, et al. (2025) Measuring AI ability to complete long tasks. arXiv:2503.14499

METR (2026a) Time Horizon 1.1. January 29, 2026. https://metr.org/blog/2026-01-29-time-horizon-1-1/

METR (2026b) Frontier Risk Report (February to March 2026). May 19, 2026. https://metr.org/blog/2026-05-19-frontier-risk-report/

METR (2026c) Time-horizon benchmark results and analysis code, Time Horizon 1.1 data release. Accessed November 2026. https://github.com/METR/eval-analysis-public

Mikel E (2026) The Computational Transition Model: when machine interpreter capacity outruns organizational complexity, version 2.0. Manuscript.

Omdia (2026) Omdia ranks AGIBOT No.1 worldwide in humanoid robot shipments in 2025. Press release, January 8, 2026. https://www.prnewswire.com/news-releases/omdia-ranks-agibot-no1-worldwide-in-humanoid-robot-shipments-in-2025-302656788.html

TrendForce (2026) Humanoid robots move toward commercialization; solid-state batteries set to break through power bottlenecks. Press release, January 28, 2026. https://www.trendforce.com/presscenter/news/20260128-12902.html

US Census Bureau (2026a) Business Trends and Outlook Survey, national share of businesses using AI in the prior two weeks: 17.3% for the two weeks ending November 30, 2025, the first period after the November 17, 2025 question change, and 24.6% for October 5–18, 2026. https://www.census.gov/hfp/btos/data

US Census Bureau (2026b) Bonney K, Breaux C, Dinlersoz E, Foster L, Haltiwanger J, Pande A. The microstructure of AI diffusion: evidence from firms, business functions, and worker tasks. Center for Economic Studies Working Paper CES-WP-26-25, April 2026. https://www.census.gov/library/working-papers/2026/adrm/CES-WP-26-25.html

Disclosures

This report reflects the author’s independent research and personal views. It is not investment advice, a recommendation, or an offer or solicitation for any security or advisory service. Estimates are conditional model outputs based on the stated assumptions, carry substantial uncertainty, and may differ materially from actual outcomes. The author is Founder and CEO of Clockwise Capital, LLC, an SEC-registered investment adviser; registration does not imply a certain level of skill or training. Clockwise Capital had no role in this research. Research and drafting were assisted by Anthropic Claude; the author is responsible for the final content.