

More code, larger changes, growing review pressure

What several industry datasets tell us about code review in the AI era

Dr. Michaela Greiler · Awesome Code Reviews

Participant handout · Evidence checked October 27, 2026

Writing code faster changes the work that reaches review. Across several vendors' datasets, we see a related set of signals: more changes, larger pull requests (PRs), longer waits for review, and more changes merging without recorded review.

The percentages differ because the reports measure different organizations, periods and outcomes. A median PR size is not an average; waiting for the first review is not time spent reviewing. The useful finding is the **convergence of signals**, rather than a single universal estimate of AI's impact.

Different measurements, a shared direction

Signal	Finding	What is being compared?
More changes	+23% monthly merged PRs · GitHub [1]	Octoverse 2025 reports 43.2 million merged PRs per month on average, up year over year. Platform-wide activity, not output per developer.
Larger changes	+122% median PR size · DX [2][3]	Year-over-year growth reported in DX's Q3 2026 research summary.
Larger changes	+79% median PR size · Greptile [4]	68 to 122 lines changed per PR, June 2025 to June 2026.
Larger changes	+71.8% average PR size · Faros [5]	Reported in the Q3 2026 analysis of deepening AI adoption.
Longer review queues	+105% wait for first review · DX [2]	Median 82 versus 40 minutes in organizations with 80%+ versus 20-40% AI code, July-September 2026. A group comparison, not growth over time.
Longer review workflow	+199.6% average time in review · Faros [6]	Organizations' periods of highest versus lowest AI adoption in its earlier 2026 report.
More unreviewed merges	+76.3% PRs merged without review · Faros [5]	A relative increase reported in Q3 2026. It does not mean that 76.3% of all PRs lack review.

How to connect the findings: a growing stream of changes can place more demand on verification. Larger changes may require more context, and longer queues can create pressure to ship. This is a workshop interpretation of the combined evidence, not a causal chain demonstrated by one study.

Why the numbers differ

These are first-party industry datasets, not representative samples of all software teams. Their customers may overlap. Agreement across reports is useful corroboration, but it does not make the studies independent replications.

Source and data vintage	Scope and method	What to keep in mind
GitHub Octoverse 2025 [1]	Platform activity reported for 2025 against the preceding year.	Growth can reflect more developers and projects. It does not isolate AI's contribution. The +23% label belongs to the monthly-average series; do not substitute the article's separate annual-total growth figure.
DX Q3 2026 [2] [3]	500+ R&D companies. Default report window: July-September 2026; PR-size growth is year over year.	The review-wait comparison groups organizations by AI-code percentage. Work mix, PR size and team practices may also differ. It does not establish that AI caused the extra waiting.
Greptile Q2 2026 [4]	Fixed March 2025 repository cohort; PR size counts nonzero additions plus deletions on merged and closed PRs, grouped by creation date.	June measurements were provisional. Connected repositories are the observed population; the public methodology does not state a cohort count for this metric.
Faros: Acceleration Whiplash, 2026 [6]	Two years of telemetry from 22,000 developers and 4,000+ teams; compares each organization's lowest and highest AI-adoption periods.	The public summary does not give exact observation boundaries. Elapsed review time includes workflow delay, not just active attention.
Faros: Speed Trap, Q3 2026 [5]	Most recent 12 months of telemetry from 22,000 developers across 4,000 teams, examining deeper use within high AI adoption.	Exact observation dates and the absolute unreviewed-merge baseline are not supplied in the public summary. This is a different comparison design from the earlier report.

Four distinctions that matter

- Waiting versus working.** Time to first review captures a queue. Time in review captures a broader workflow interval. Neither directly measures hours of human reading or the quality of the review.
- Relative increase versus share.** If 10 of 100 PRs lacked review, a 76.3% relative increase would correspond to about 17.6 of 100, assuming the same denominator. This is an illustration, not Faros's baseline. It is not an increase of 76.3 percentage points.
- Recorded review versus all checking.** A PR merged without recorded review may still have been inspected locally or discussed elsewhere. Conversely, a recorded approval does not tell us how carefully the code was understood.
- More discussion and more skipped reviews can coexist.** Some changes can attract more scrutiny while others merge without recorded review. Growth in unreviewed merges does not, by itself, prove that the total number of reviews fell.
- These reports describe workflows, not controlled comparisons of individual models. Publication in 2026 does not mean every observation involved the latest 2026 coding agent.

Convergence does not mean every study finds the same result

Chen and Stratton's August 2026 working paper uses Jellyfish data covering January 2021-March 2026 across 718 firms. Following coding-agent adoption, it estimates longer submission-to-merge times, but **no statistically significant change in PR size measured by lines added**. [7] That is a different population, metric and design. It prevents us from claiming that larger PRs are an inevitable result of AI use everywhere.

What this means for a team

We need risk-based, multi-stage reviews as implementation becomes more agentic. More and larger changes call for deliberate decisions about where review happens, who contributes, and what needs careful verification. Not every change needs the same oversight.

This is the idea behind **SCOPE: Staged Change Oversight with Proportionate Escalation**, a review model I am developing for agentic software development. Review starts before implementation: the team considers the change's risk and alignment needs, agrees the oversight, and revisits those decisions as more becomes known. Agent Review, Developer Review and Independent Review then serve as building blocks, with their depth and involvement adapted to the change.

Higher risk calls for stronger verification and independent challenge. Higher alignment needs call for more collaboration and shared understanding. AI can contribute to implementation, investigation, tests and review; the developer remains responsible for understanding and verifying the work. Peer review should add an independent perspective where it is needed, informed by the evidence already established.

The aim is to direct limited human attention where it brings the most value as code volume grows. SCOPE is under active development; this is a proposed response to the findings, not an intervention validated by these reports.

A few practical checks support this approach:

Signal to inspect locally	A practical response
Growing PR volume or size	Split unrelated concerns; use smaller reviewable steps where risk or alignment needs call for closer oversight.
Longer first-review waits	Agree review ownership early, surface aging work, and separate queue time from active review rounds.
More unreviewed merges	Examine a sample, including local or agent checks; make review and escalation requirements explicit.
Fast approvals but weak shared understanding	Use walkthroughs and targeted questions on consequential changes; record what was checked, the decision and remaining concerns.

Sources

All links were checked October 27, 2026. Figures above use publicly accessible summaries; no gated report was assumed to have been read.

1. **GitHub**. [Octoverse 2025: A new developer joins GitHub every second as AI leads TypeScript to #1](#). See opening activity figures.
2. **DX**. [State of AI Impact in Engineering: Q3 Report](#). See "Review queues are saturated" and "About this data."
3. **Grace Fu and Scott Carey, DX**. [The gap between code velocity and trust is growing](#), October 14, 2026. Source of the +122% median PR-size figure.
4. **Greptile Research**. [The State of AI Coding: Q2 2026](#), July 2, 2026. See chart 1.1 and "Methodology and Sources."
5. **Faros Research**. [The Speed Trap: 8 takeaways from our latest AI engineering research](#), Q3 2026 report summary. See sections 4-5 and its stated sample scope.
6. **Faros Research**. [Ten takeaways from the AI Engineering Report 2026: The Acceleration Whiplash](#). See methodology introduction and section 7.
7. **Fiona Chen and James Stratton**. [Artificial Intelligence in the Firm: Bottlenecks in Software Production](#), working paper, August 4, 2026. See section 4.1 and Appendix Figure 5.