

A check for AI code review suggestions

AI code assistants suggest changes faster than people can review them. We measured what the redundant ones cost and built a check that catches them.

~\$49K

modeled technical debt from redundant AI suggestions over 4 months

87.5%

agreement with senior engineers, best case

+47%

more consistent than ad hoc prompting

THE CHALLENGE

Suggestions pile up unchecked

AI review tools produce a steady stream of code suggestions, and only 2.6% of them ever get human feedback. The redundant ones slip through and turn into maintenance work later.

WHAT WE DID

Measured the cost, then built a check

- ✓ Analyzed AI code review suggestions from a public dataset
- ✓ Estimated the maintenance cost of redundant suggestions
- ✓ Built an evaluation check on our ICI framework that scores each suggestion against team standards
- ✓ Ran it on an open model (Qwen2.5-Coder 7B) that can run in-house

RESULTS

- ✓ Up to 87.5% agreement with senior engineers
- ✓ 47% more consistent than ad hoc prompting
- ✓ Flags redundant suggestions before they ship

Watch the talk: youtu.be/stdgG-7ZqBY

WHAT IT SUPPORTS

Defect escape rate

Technical debt ratio

Reviewer hours

The same approach works anywhere AI output needs checking before it is used: code, support replies, reports, or content.

Using AI to write or review code?

Start with a free 30-minute call. You leave with a clear next step.

[Book a free call →](#)