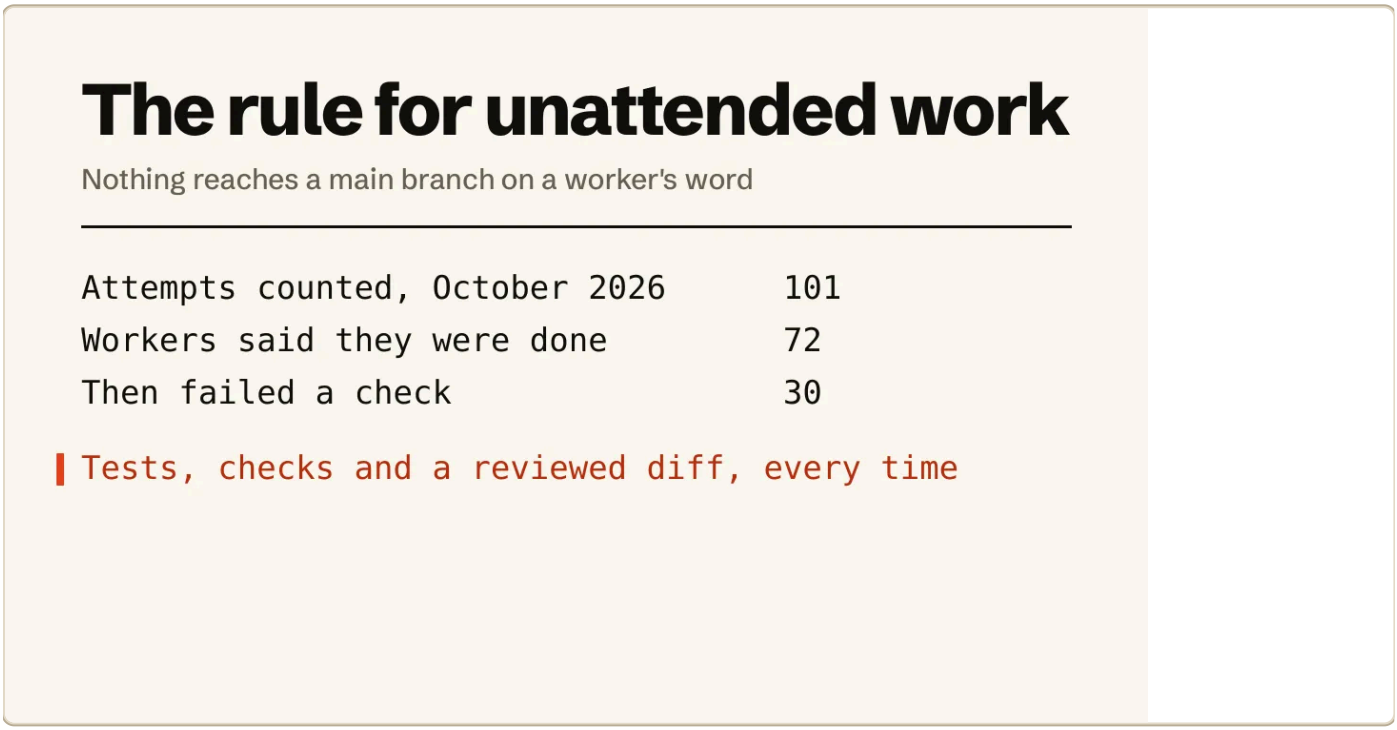
The rule that lets me trust work I'm not watching

Unattended AI work is fine when nothing reaches a main branch without passing tests, passing checks and a reviewed diff. Why I hold to that rule, with two numbers and one catch that explain it.



Published
4 November 2026

Reading time
1 minute

About
Mullet

Tags
AI, Testing, Security

More on: Mullet

I'm happy for AI to work while I'm not watching. I'm not happy for it to change a main branch on its own say-so. The rule that reconciles the two is short: nothing reaches a main branch without passing its tests, passing its checks and having its diff reviewed.

A worker's word counts for nothing

When I counted 101 attempts in my job runner's record in October 2026, the workers said they were done 72 times. Thirty of those then failed a check. That is the number I think about whenever someone says a model has finished.

Tests alone aren't enough

On 3 and 4 November 2026 a cheaper model's change passed its tests, and the review still caught a real security hole that the tests had missed. That attempt was fixed by hand, and it went no further.

Whoever writes a change doesn't get to judge it. Something that took no part in writing it reads the diff, and I read it again before it merges.

What the rule doesn't do

It doesn't prove the code is right. It makes a wrong change hard to get through, and it makes sure I find out before it lands, not after.

That is the bar I set for unattended work, and it's why I'm comfortable leaving it running overnight. It's also the reason cheap models are safe to use at all, which I wrote about in software that gets built and tested while I sleep.