Tom Hert - INET @ HAW Hamburg

Tipping Scales: Analyzing the impact of LLM on FOSS Maintainers

Analyzing 254 pull requests in RIOT from 1 Jun – 31 Aug 2026

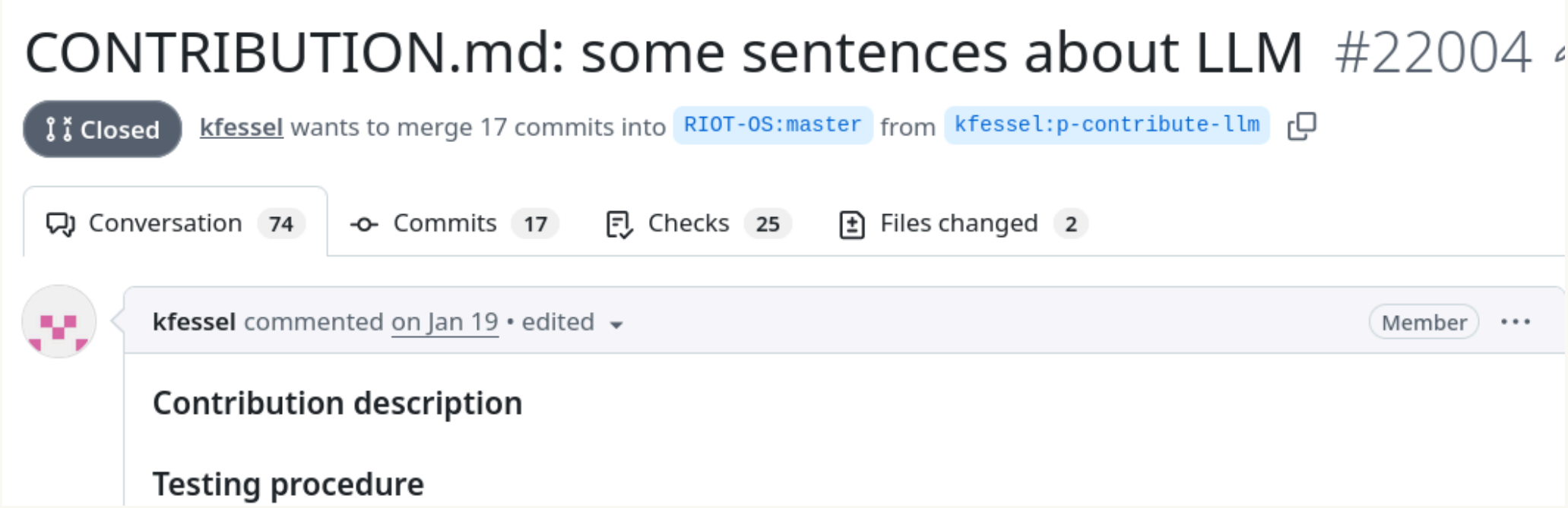
00/20 Agenda

- Measurement Goal
- Recap of AI Policy in RIOT
- Cool Graphs ✨
- Takeaways

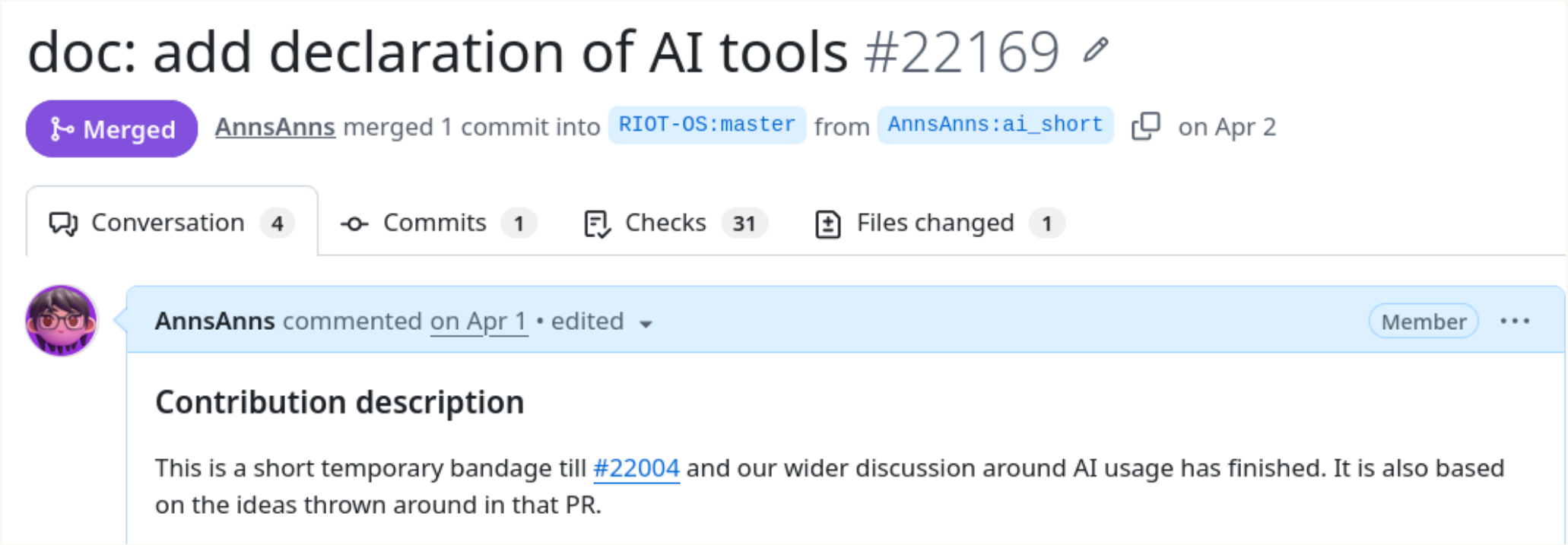
01/20 What we hoped to find out

- What is the level of AI usage in PRs?
- Is AI-assistance a factor in PR acceptance or rejection?
- Is AI-assistance a burden on Maintainers?
- What are the most common tools and use cases?

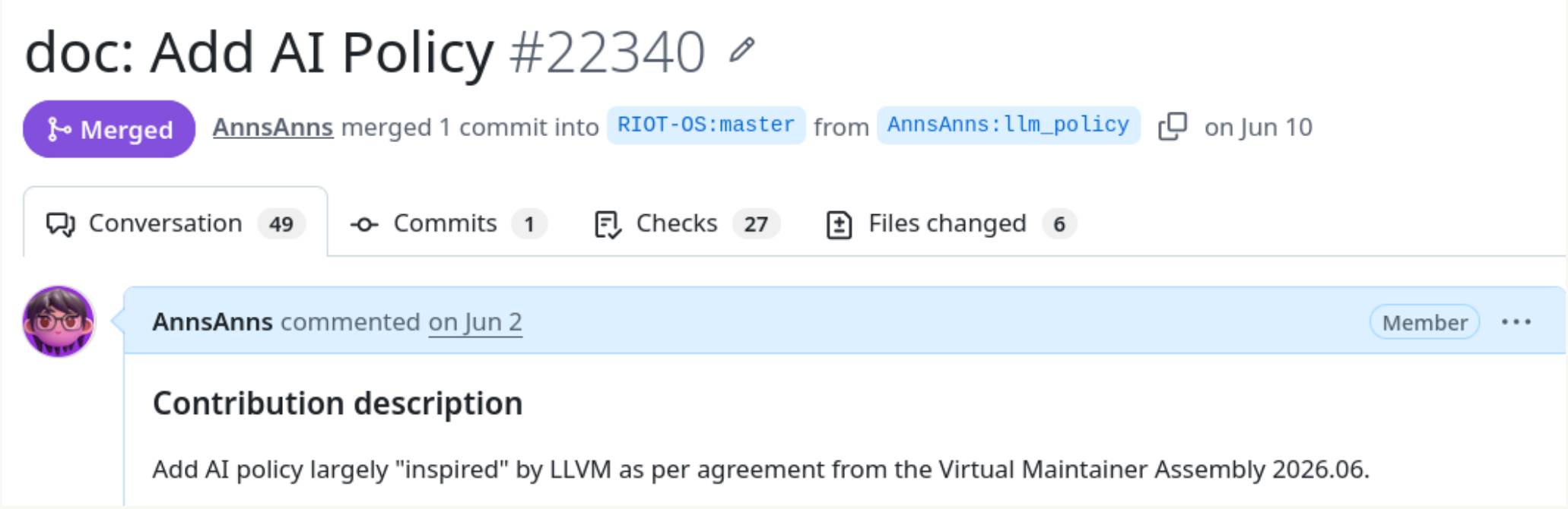
02/20 AI Declaration First Attempt



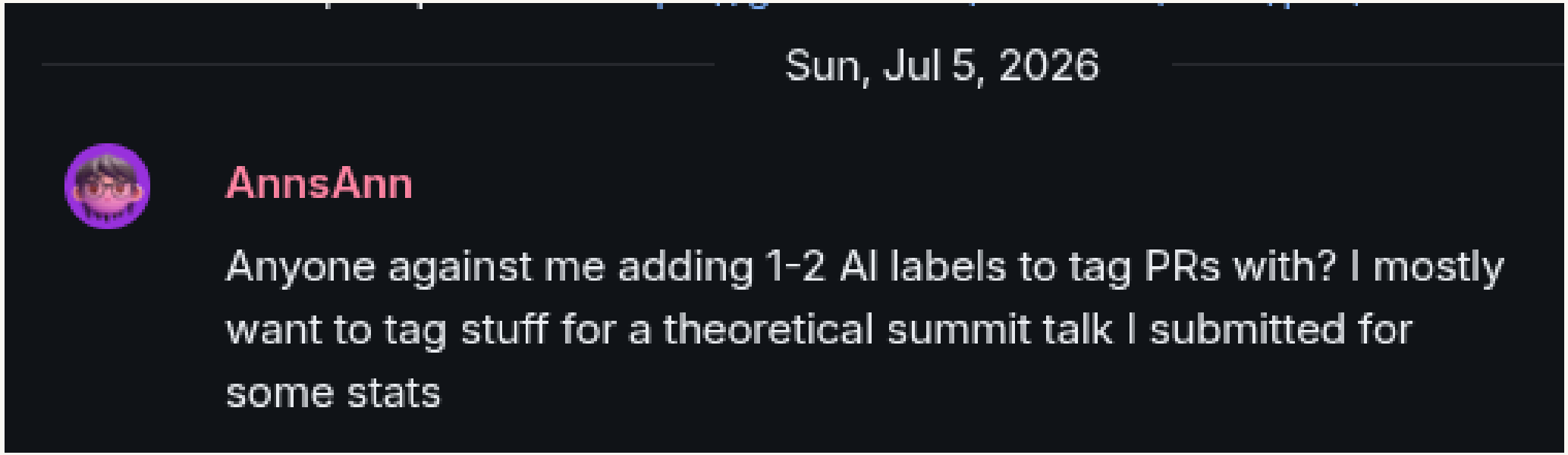
03/20 AI Declaration Second Attempt



04/20 AI Policy ala LLVM



05/20 Labelling PRs



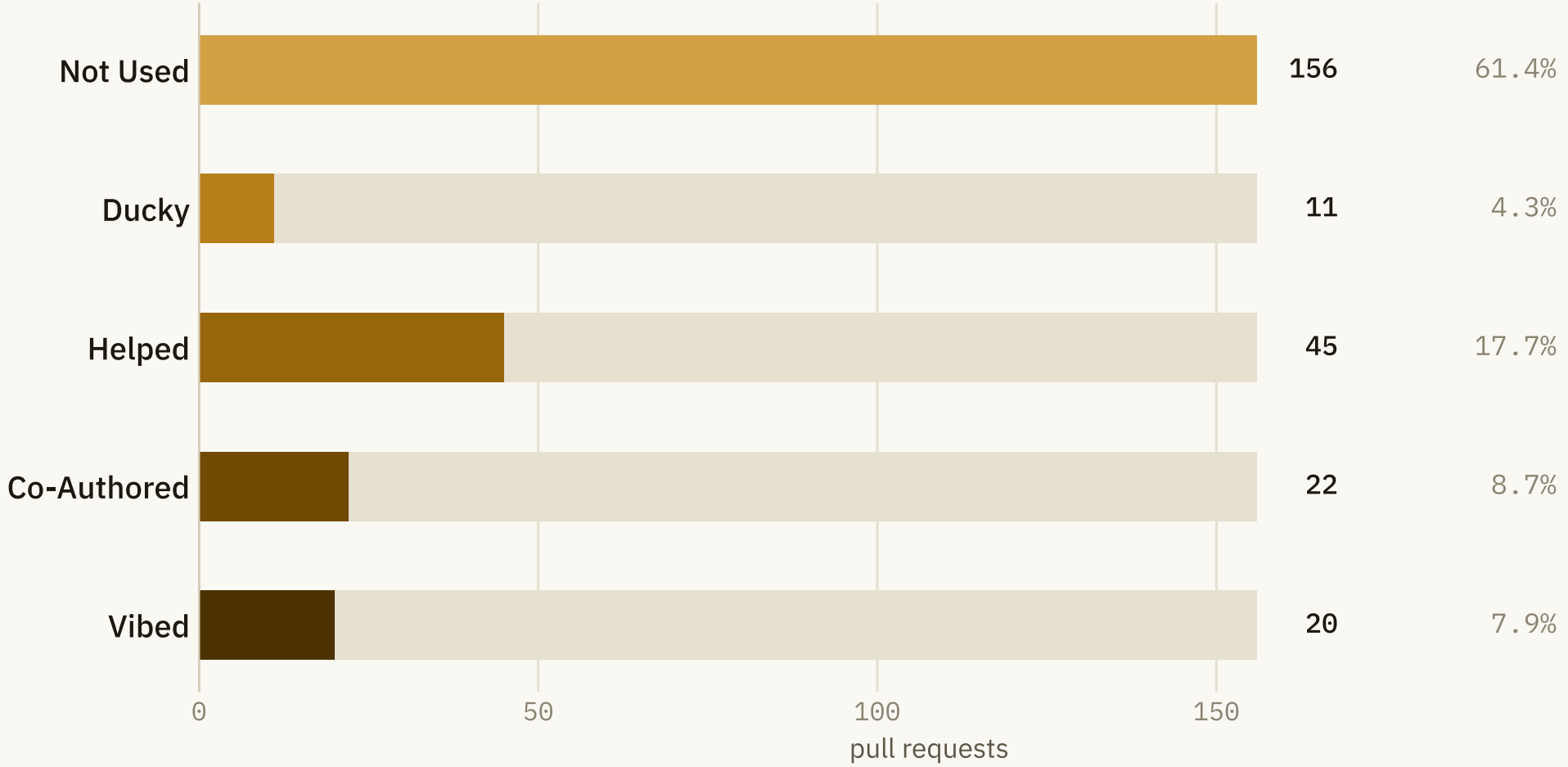
06/20

Limitations

- **Self-reported, then graded by hand** mostly by 2 Maintainers
- **Under-declaration of AI usage can be assumed**
- **Only 254 PRs** means AI tags thin out when filtered

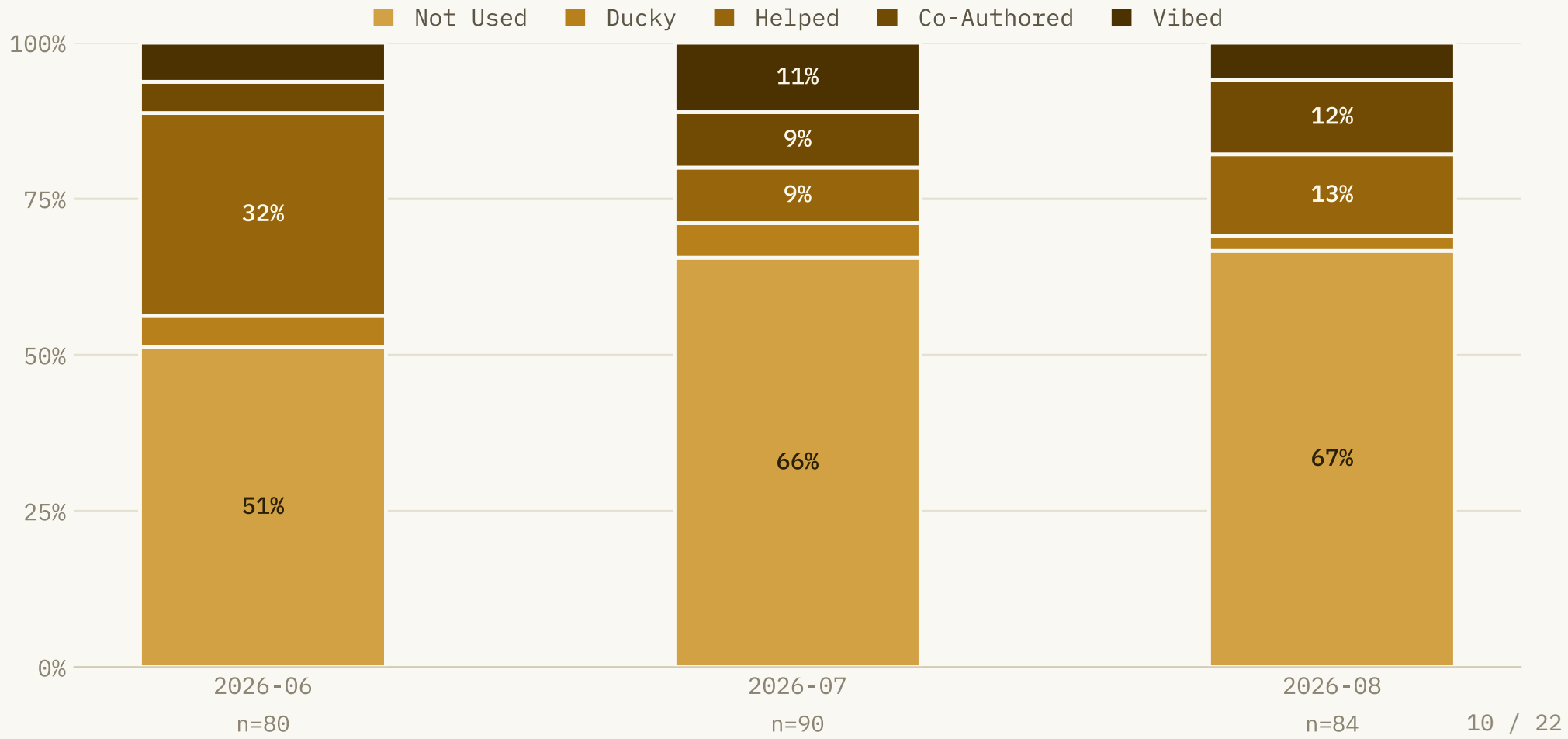
07/20

How much AI is declared



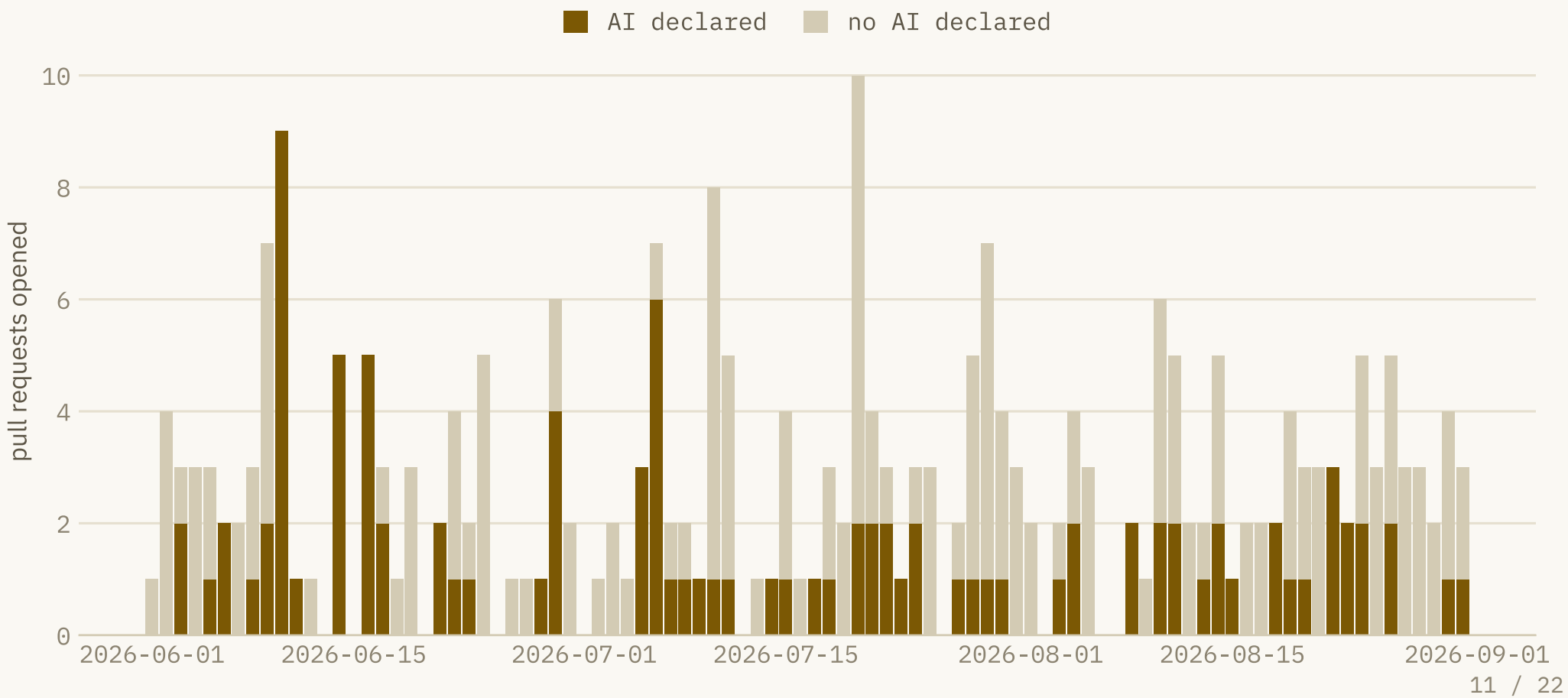
08/20

Level mix by month



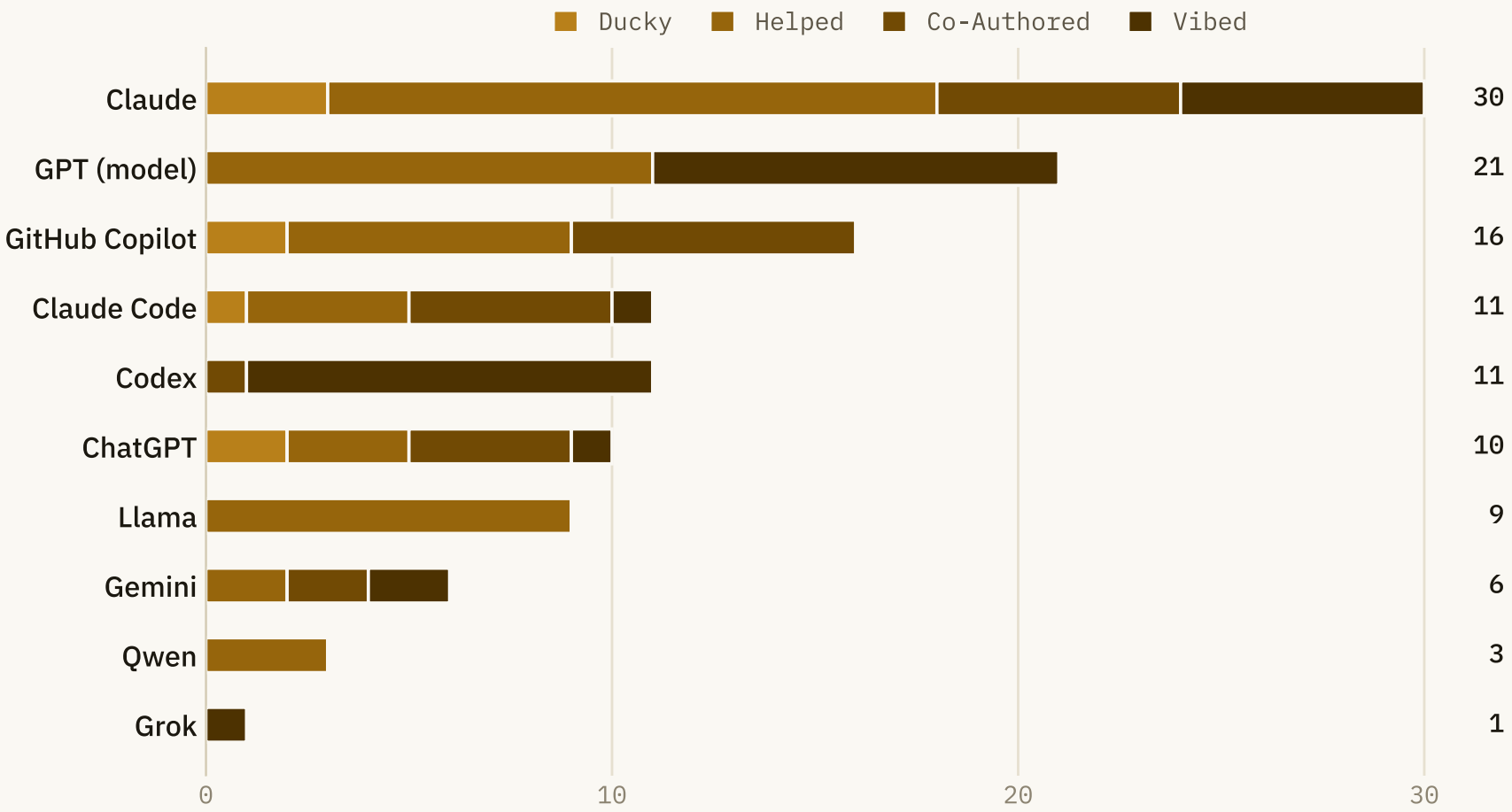
09/20

Pull requests opened per day

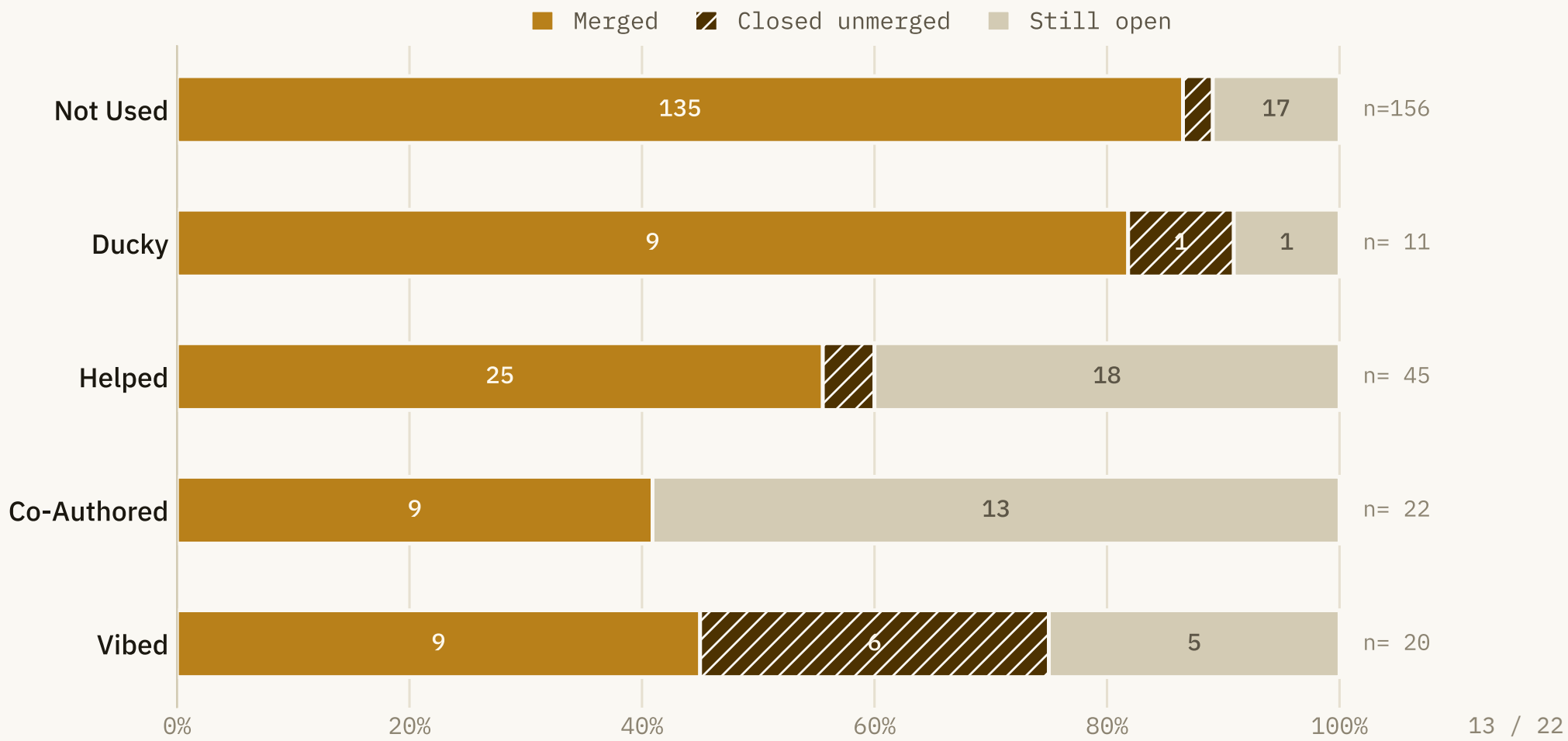


10/20

Tools named

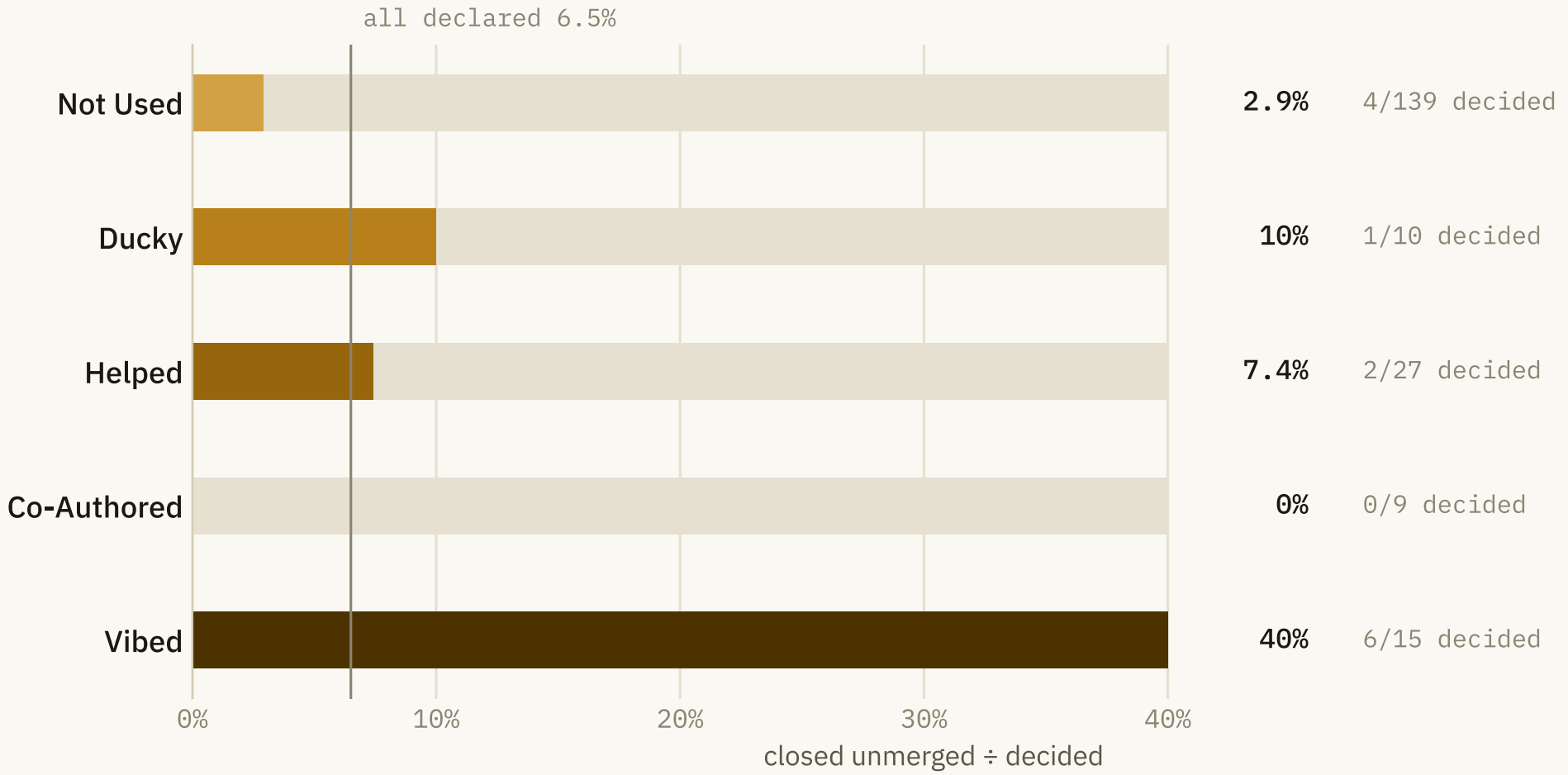


Outcome by level



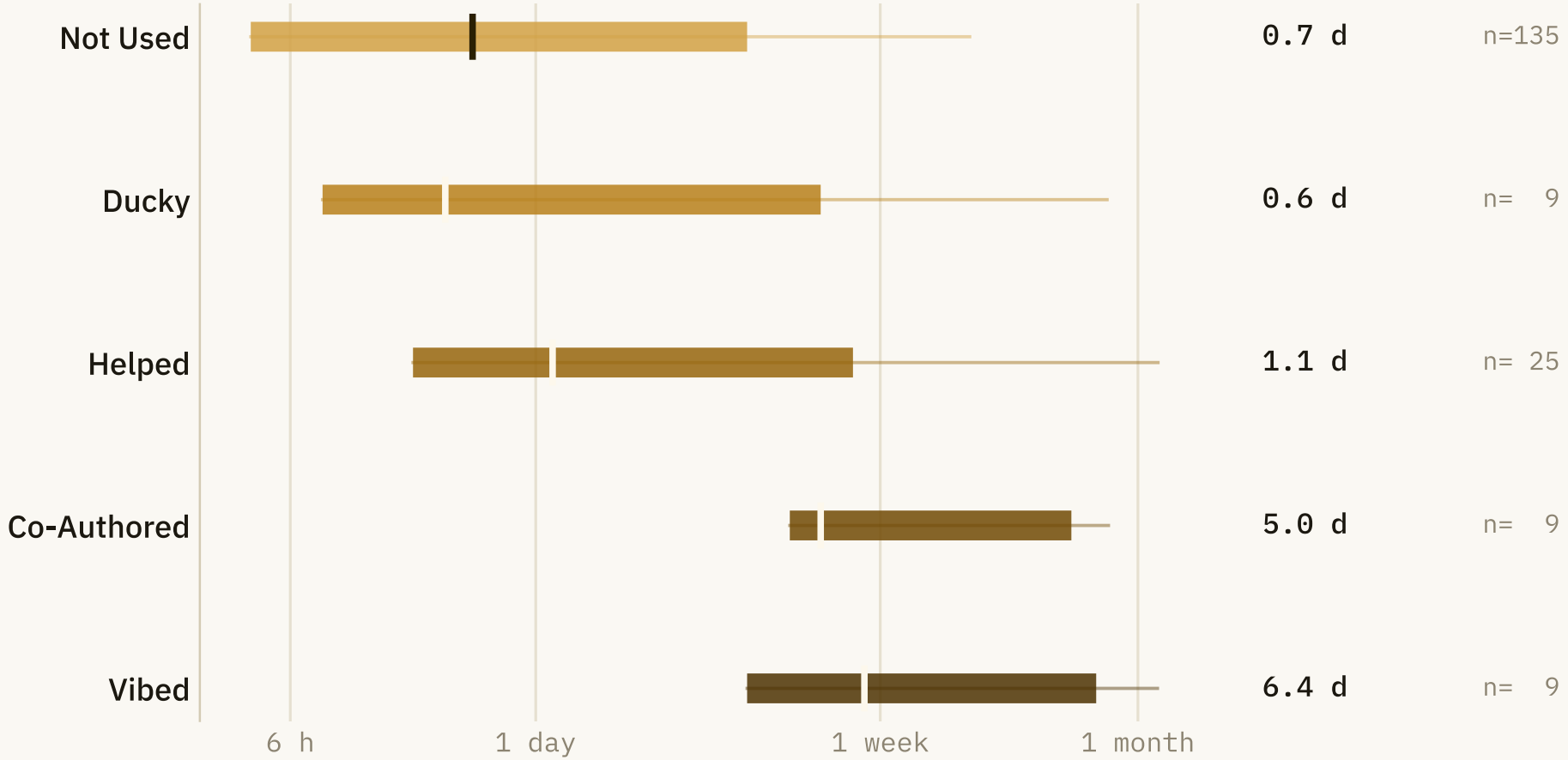
12/20

Rejection rate by level



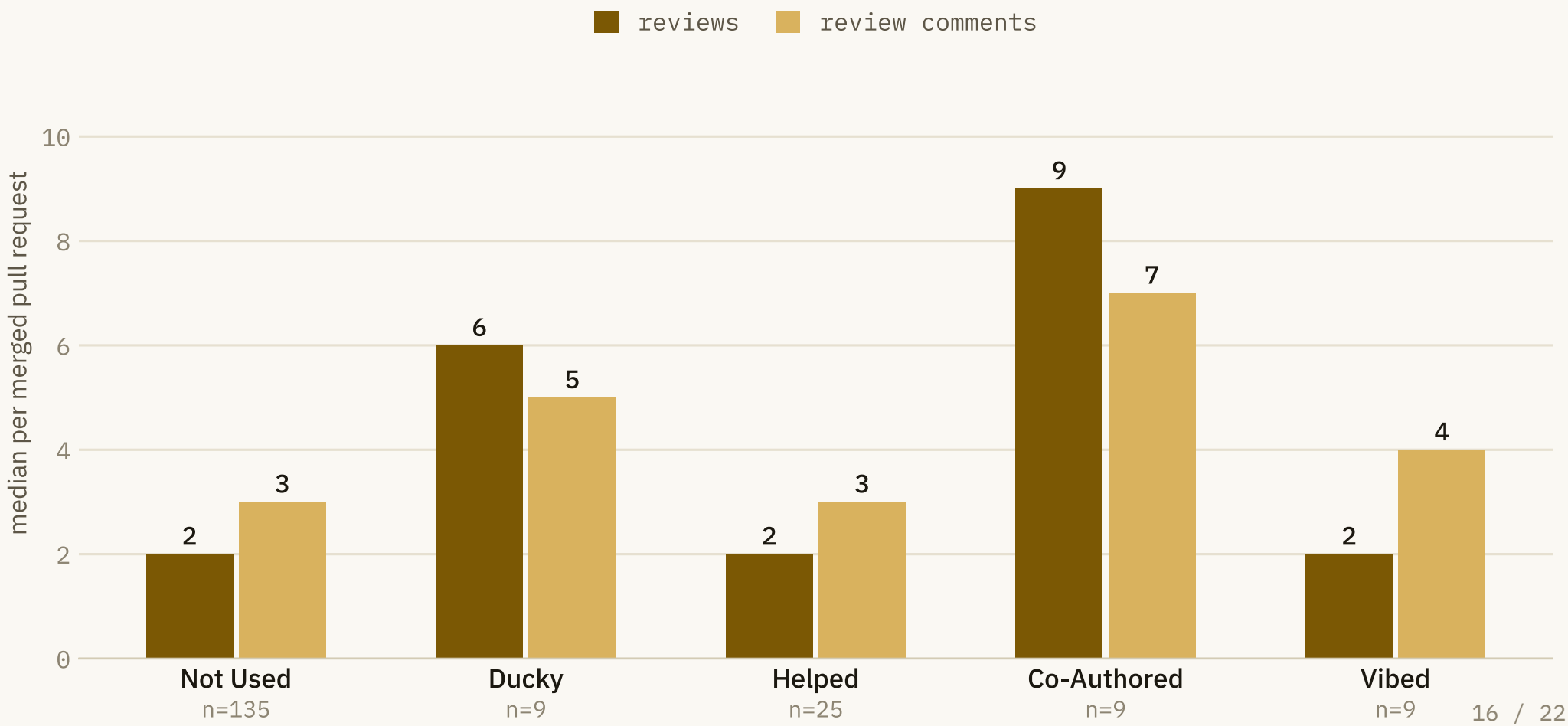
13/20

Time to merge by level

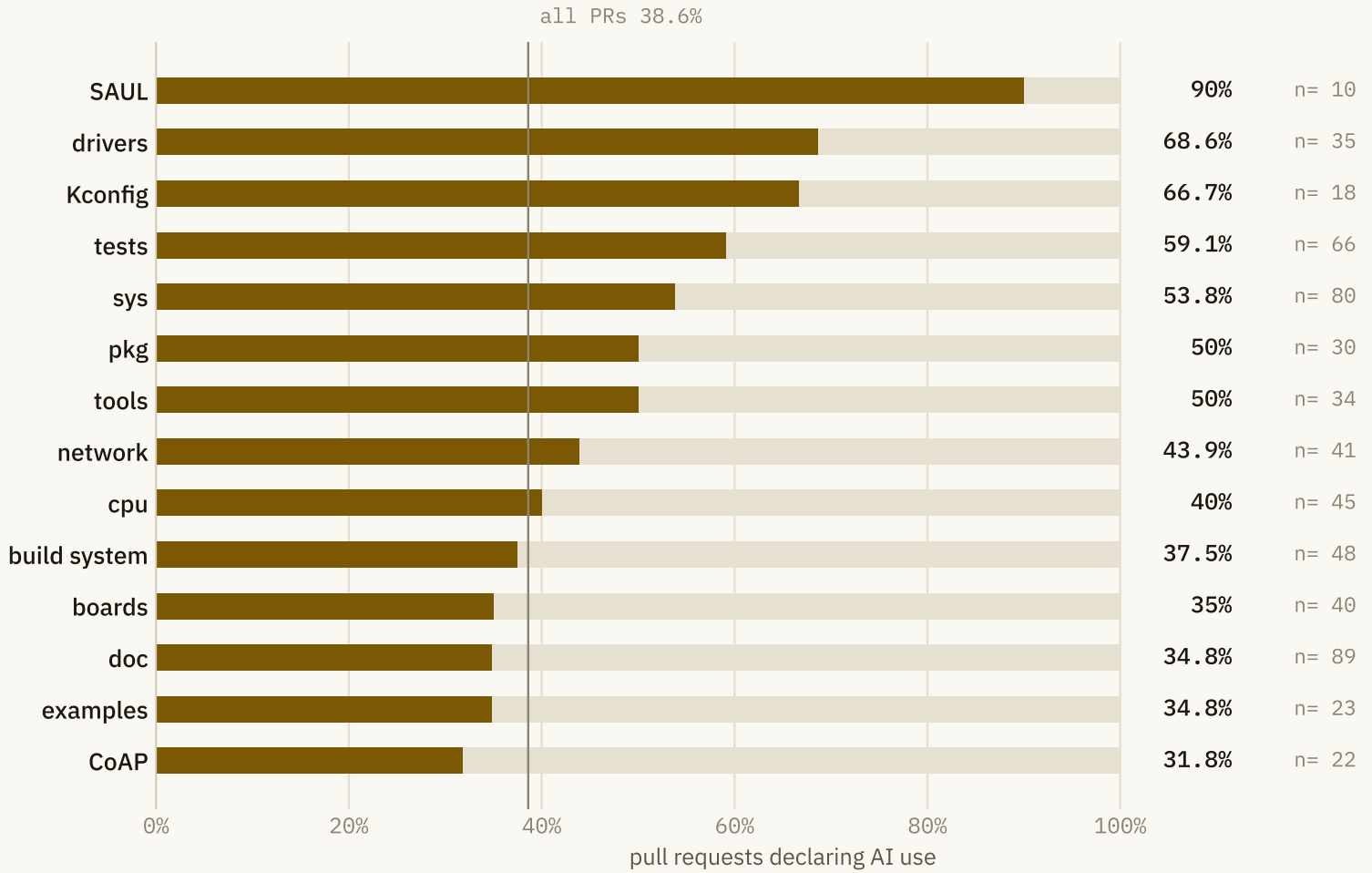


time from opening to merge — bar p25–p75, line to p90, tick median

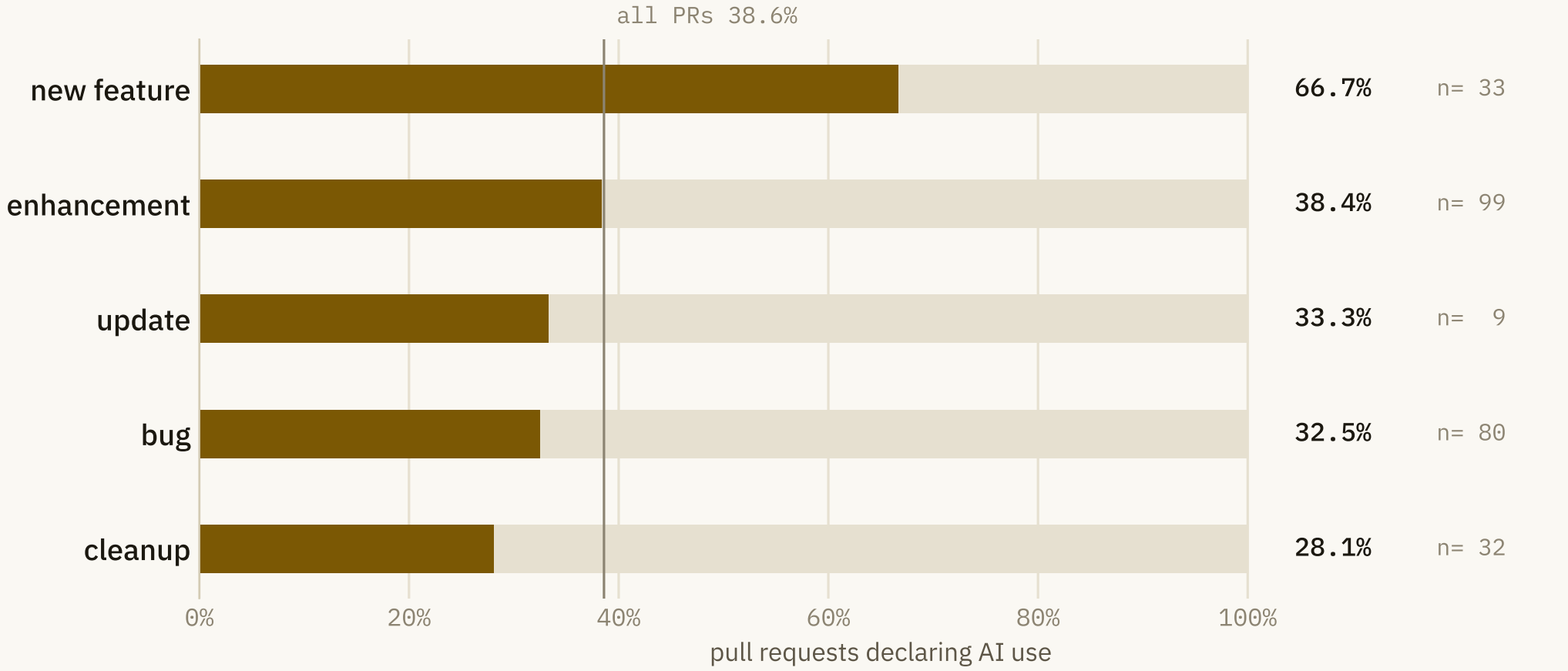
Review effort by level



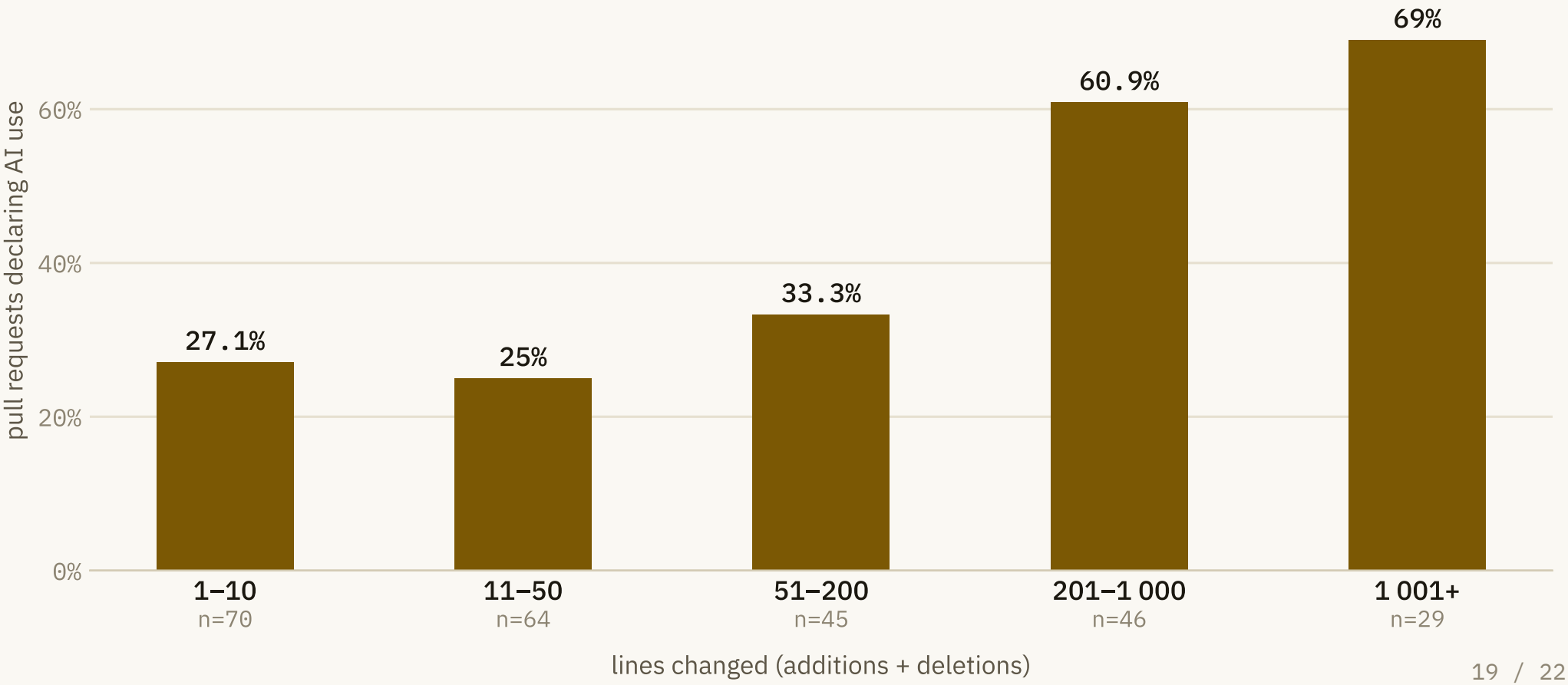
AI share by area



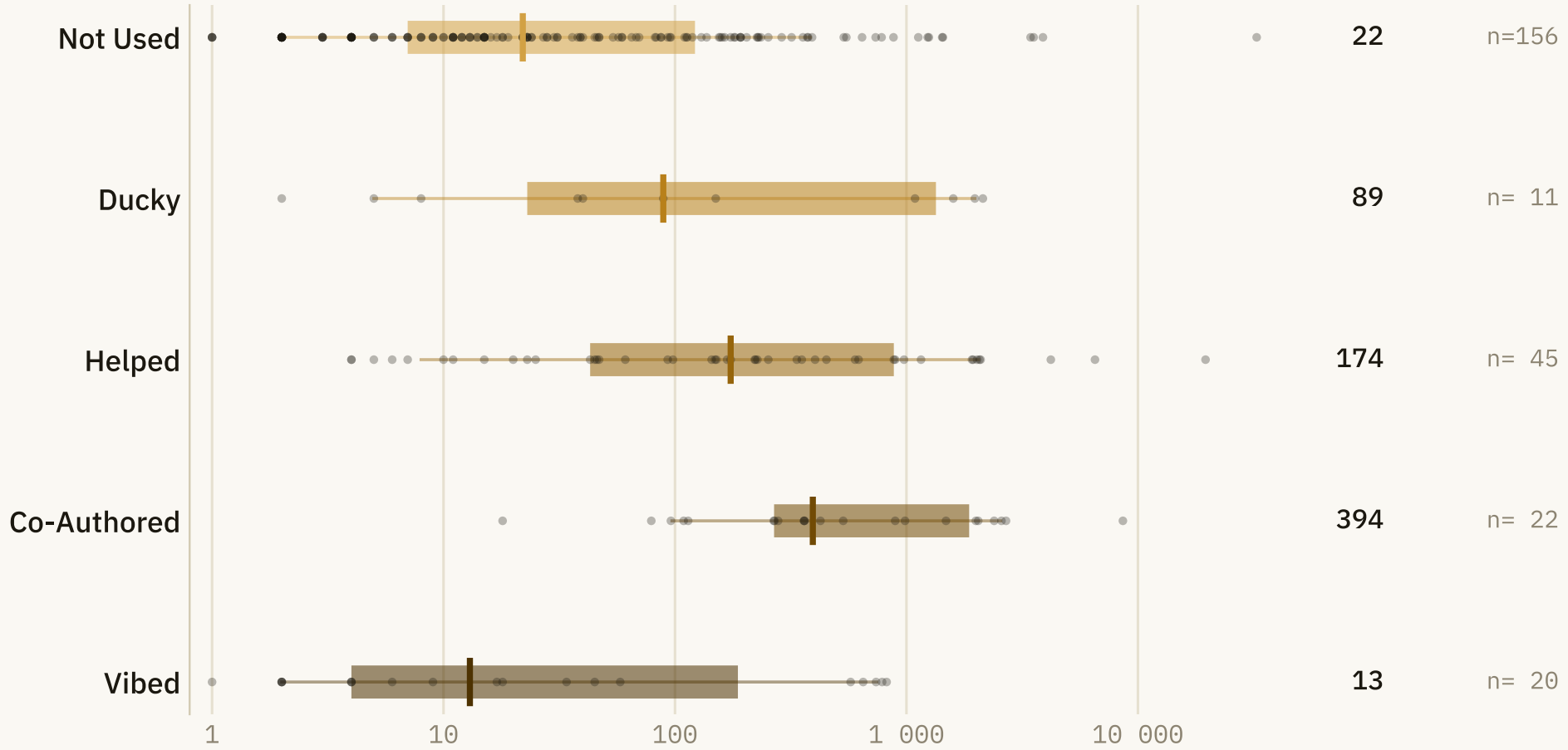
AI share by kind of change



AI share by pull request size

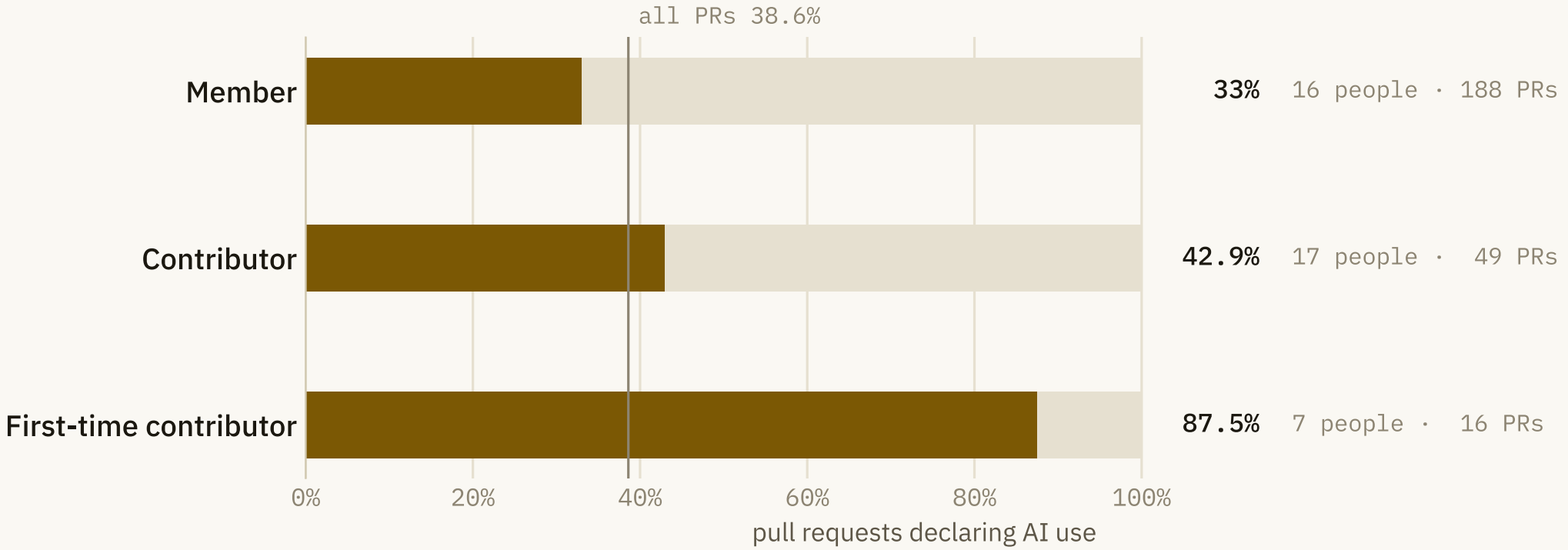


Size distribution by level



19/20

AI share by standing with the project



20/20 Takeaways

- **Over a third of PRs declare AI** (38.6%)
- **The more Code the more AI** (27% AI under ten lines, 69% over a thousand)
- **The less familiar the more AI** (33% under Members, 87.5% under First Timers)
- **New Features & Drivers usually with AI, Maintenance tends to be human**
- **AI Review Impact still Inconclusive** (Needs longer term data)