RQs and Methodology:Understanding Automation and Human Intervention in Agentic Pull Requests

PHAN TRUONG PHUOC NGUYEN & SAGE YANG

We study how automated agents and human developers collaborate on pull requests (PRs) in open-source repositories. We quantify how many PRs are fully automated versus human-involved, how work type and complexity relate to automation levels, and which bots tend to require the most human feedback and corrections.

1 Research Questions (RQs)

We study the following research questions about automation and human intervention in agentic pull requests (PRs).

RQ1 Automation

To what extent are pull requests (PRs) fully automated (all commits by bots, no human comments or reviews) versus requiring human participation?

RQ2 Work type and complexity

What types of work and complexity levels correspond to fully automated PRs compared to PRs that involve humans?

RQ3 Bot-level differences

Which bots generate PRs that require the most human feedback and corrections?

2 Methodology: How We Answer Each RQ

RQ1 Automation extent

We join PRs with comments, reviews, and commits to set human-participation flags (using the bot/human rules and excluding merge commits). Based on these flags, we assign each PR to Level 0/1/2 and report counts and proportions of each level, both overall and per repository.

RQ2 Work type and complexity

We join PRs with issues, labels, and task-type annotations (e.g., `pr_task_type`) plus `commit_details` to obtain work categories and size metrics (numbers of commits, lines, and files). We then compare task-type distributions and size distributions across automation levels to characterize which kinds of work and PR sizes tend to be fully automated versus human-involved.

RQ3 Bot-level differences

We assign each PR with bot commits to a primary bot, compute human-interaction measures per PR (human comments, reviews, and whether it reaches Level 2), and aggregate these measures per bot. We compare bots, stratified by task type and coarse PR size buckets, to identify which bots' PRs systematically require more human feedback or corrections.

Author's Contact Information: PHAN TRUONG PHUOC NGUYEN & Sage Yang.