Mining Software Repositories: An Exploratory Study of AI-Generated Pull Requests Using the AIDev Dataset

Nihed Jamaoui
National Computer Science School (ENSI)
University of Manouba, Tunisia
Email: nihed.jamaoui@ensi-uma.tn

***Abstract*—AI-based programming agents are increasingly involved in software development by autonomously generating pull requests in open-source repositories. This paper presents an exploratory Mining Software Repositories (MSR) study of the AIDev dataset, which captures large-scale AI-generated pull requests on GitHub. Using a Goal–Question–Metric (GQM) approach, we analyze the contribution, outcomes, and processing time of agentic pull requests. The study provides quantitative insights into the behavior of AI-generated contributions and discusses limitations related to dataset scope and validity.**

***Index Terms*—Mining Software Repositories, AI Programming Agents, Pull Requests, GitHub, MSR Challenge**

I. Introduction

The emergence of AI-based programming agents such as OpenAI Codex, GitHub Copilot, and Devin has significantly transformed modern software development practices. These agents are capable of autonomously generating code contributions and submitting them as pull requests (PRs) to open-source repositories.

Mining Software Repositories (MSR) provides a systematic methodology for analyzing software artifacts and development activities at scale. While prior MSR studies have extensively investigated human-generated pull requests, the characteristics and outcomes of fully AI-generated pull requests remain largely unexplored.

Goal and Contribution. The goal of this study is to analyze the behavior and outcomes of AI-generated pull requests in open-source GitHub projects. The main contribution of this work is an exploratory quantitative analysis of agentic pull requests using the AIDev dataset, focusing on their contribution, merge outcomes, and processing time.

II. Related Work

Several empirical studies have investigated the dynamics of pull requests in GitHub. Rigby et al. examined the influence of social and technical factors on the evaluation of contributions, identifying key patterns in acceptance and review processes [1]. Their work established foundational insights in MSR analysis of collaborative development.

More recently, Chen et al. studied the capabilities and limitations of AI code generation agents on GitHub, analyzing how such agents perform in real development workflows [2]. While these studies focus on human-assisted and machine-assisted contributions, our work extends this line of research by exclusively analyzing pull requests autonomously generated by programming agents using the AIDev dataset.

III. Study Design

A. Goal–Question–Metric (GQM)

Goal: Analyze the contribution and outcomes of AI-generated pull requests in open-source GitHub repositories to better understand their integration into software development workflows.

B. Research Questions

- **RQ1:** What is the outcome of AI-generated pull requests in terms of merge and closure without merge?
- **RQ2:** What is the processing time of AI-generated pull requests?
- **RQ3:** How is the contribution distributed among different AI programming agents?

C. Dataset and Subset

The AIDev dataset is a public dataset released as part of the MSR Mining Challenge and hosted on Hugging Face. It contains metadata describing pull requests generated by AI-based programming agents in GitHub repositories.

To answer our research questions, we define a subset including pull requests that:

- have a valid agent identifier,
- contain complete lifecycle timestamps,
- are closed (merged or closed without merge).

The final subset contains **859,927 pull requests** described by 14 attributes.

D. Metrics

The following metrics are used:

- **Merge rate**
- **Closed without merge rate**
- **Processing time** (days)
- **Contribution proportion per agent**

IV. Analysis and Results

A. Contribution of AI Agents (RQ3)

Figure 1 shows the distribution of pull requests among AI programming agents. The results indicate a strong dominance of OpenAI Codex, followed by GitHub Copilot, Devin, and other agents.

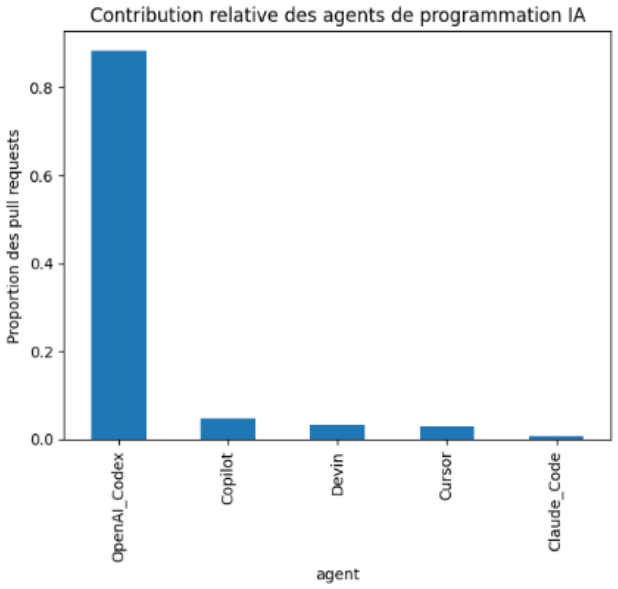

Fig. 1. Contribution distribution of AI programming agents

B. Processing Time (RQ2)

The average processing time of AI-generated pull requests is approximately **0.33 days**. Figure 2 illustrates the distribution of processing times, indicating short review durations with limited variability.

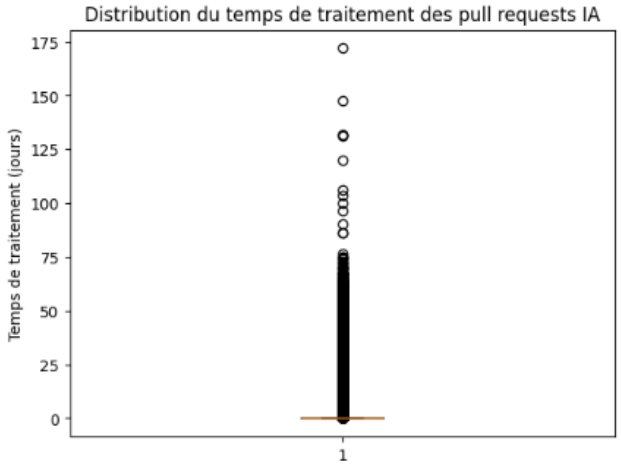

Fig. 2. Distribution of processing time for AI-generated pull requests

C. Outcomes of AI-Generated Pull Requests (RQ1)

The analysis reveals that approximately **91.9%** of AI-generated pull requests are merged, while **8.1%** are closed without merge. This high merge rate suggests that agent-generated contributions are generally accepted in the studied repositories.

V. Discussion

The results indicate that AI-generated pull requests are rapidly processed and frequently merged, suggesting that programming agents are primarily used for focused and manageable development tasks. The dominance of certain agents highlights ecosystem effects, where a limited number of tools account for most automated contributions.

For developers, these findings suggest that AI agents can effectively accelerate development workflows. For tool designers, the results emphasize the importance of seamless integration and usability. For researchers, this study demonstrates the value of dedicated datasets such as AIDev for advancing empirical research on agentic software development.

VI. Threats to Validity

Construct validity is limited by the use of quantitative metrics that do not capture code quality or semantic correctness.

Internal validity is affected by the absence of human-generated pull requests in the dataset, preventing direct comparisons between AI and human contributions.

External validity is limited since the dataset covers a specific set of agents and repositories, which may not generalize to all open-source projects.

VII. Conclusion

This paper presented an exploratory MSR study of AI-generated pull requests using the AIDev dataset. The results highlight a high merge rate, short processing times, and unequal contribution among agents. Despite dataset limitations, the study provides valuable insights into agentic software development and demonstrates the potential of the AIDev dataset for future MSR research. Future work may integrate human-generated pull requests or perform qualitative analyses of code changes.

References

[1] P. C. Rigby, D. M. German, and M.-A. D. Storey, "Understanding the influence of social and technical factors for evaluating contribution in github," in *Proceedings of the 36th International Conference on Software Engineering*. ACM, 2014, pp. 356–366.

[2] Y. Chen, W. Li, X. Qiu, D. Zhang, and H. Li, "Codexbots: Understanding the capabilities and limitations of ai code generation agents on github," in *Proceedings of the 20th International Conference on Mining Software Repositories (MSR)*. ACM, 2023, pp. 553–565.