

RELATÓRIO

# Caracterizando a Atividade de Code Review no GitHub

## Characterizing the Code Review Activity on GitHub

Lorrayne Oliveira [Pontifícia Universidade Católica de Minas Gerais | [lorrayne.marayze@gmail.com](mailto:lorrayne.marayze@gmail.com)]  
Pedro Pires [Pontifícia Universidade Católica de Minas Gerais | [pedro.pires@gmail.com](mailto:pedro.pires@gmail.com)]

**Resumo.** Este trabalho apresenta um estudo empírico sobre a prática de code review em repositórios populares do GitHub. A atividade de revisão de código, fundamental para assegurar a qualidade e manutenibilidade do software, é analisada sob diferentes dimensões, considerando métricas associadas ao tamanho, tempo de análise, descrição e interações de pull requests (PRs). O objetivo central é compreender como essas variáveis influenciam o resultado das revisões (merge ou rejeição) e o número de revisões realizadas. Para isso, foi construído um dataset contendo PRs dos 200 repositórios mais populares do GitHub, avaliando somente aqueles que passaram por processos de revisão humana e tiveram tempo mínimo de análise superior a uma hora. As correlações entre as variáveis foram examinadas por meio de análises estatísticas de Spearman, de modo a identificar padrões significativos no comportamento das revisões de código. Os resultados obtidos fornecem evidências quantitativas sobre fatores que impactam o sucesso das contribuições, oferecendo insights sobre as dinâmicas colaborativas de revisão em projetos open-source.

**Abstract.** This work presents an empirical study on the practice of code review in popular GitHub repositories. The code review activity, essential for ensuring software quality and maintainability, is analyzed through multiple dimensions, considering metrics related to the size, review time, description, and interactions of pull requests (PRs). The main goal is to understand how these variables influence the review outcomes (merge or rejection) and the number of reviews performed. A dataset was built containing PRs from the 200 most popular GitHub repositories, considering only those with human reviews and analysis time greater than one hour. Correlations among variables were examined using the Spearman statistical test to identify significant patterns in review behavior. The results provide quantitative evidence about factors influencing contribution success, offering insights into collaborative review dynamics in open-source projects.

**Palavras-chave:** Code Review; Pull Requests; GitHub; Revisão de Código; Engenharia de Software Colaborativa

**Keywords:** Code Review; Pull Requests; GitHub; Code Quality; Collaborative Software Engineering

## 1 Introdução

A revisão de código (code review) é uma das práticas mais consolidadas em processos ágeis de desenvolvimento de software. Ela consiste na inspeção do código por revisores antes de sua integração ao repositório principal, com o intuito de detectar defeitos, melhorar a legibilidade e assegurar a aderência aos padrões de qualidade. No contexto de plataformas open-source como o GitHub, essa atividade é operacionalizada por meio de pull requests (PRs), que representam contribuições submetidas por desenvolvedores e avaliadas por colaboradores do projeto.

Neste trabalho, busca-se caracterizar empiricamente o comportamento das atividades de code review em repositórios populares do GitHub. Especificamente, este estudo visa compreender a forma com que os fatores, como o tamanho do PR, o tempo de análise, a extensão da descrição e o volume de interações influenciam o resultado final da revisão (merge ou rejeição) e o número de revisões realizadas.

A partir da análise, pretende-se identificar padrões recorrentes e variáveis críticas que impactam o sucesso de revisões, contribuindo para a melhoria de práticas de colaboração e qualidade de código em ambientes distribuídos de desenvolvimento.

### 1.1 Hipóteses Informais

### 1.2 Objetivos

## 2 Metodologia

### 2.1 Coleta de Dados

A

### 2.2 Processamento dos Dados

### 2.3 Exportação e Análise

### 2.4 Limitações

## 3 Resultados

## 4 Conclusão