

# **Análise de Características de Repositórios Populares**

**Bruno Gomes Ferreira, João Pedro Mairinque de Azevedo,  
Matheus Vieira dos Santos, Marcio Lucas Machado Pereira**

<sup>1</sup> Instituto de Ciências Exatas e Informática (ICEI)  
Pontifícia Universidade Católica de Minas Gerais  
Engenharia de Software  
Belo Horizonte – MG – Brazil

## **1. Introdução**

Os sistemas open-source tem como premissa trazer colaboração e transparência entre os usuários, tendo isso em vista, este projeto tem como objetivo analisar repositórios de código aberto focando em sistemas que possuem uma maior atenção e admiração pela comunidade, ou seja, os repositórios com a maior quantidade de estrelas no GitHub.

Ao longo deste estudo, busca-se percorrer algumas características destes repositórios com o intuito de examinar o processo de desenvolvimento, avaliar a regularidade das contribuições externas, analisar a frequência de lançamentos de versões e investigar outras características associadas a esses elementos.

## **2. Metodologia**

Inicialmente, foram traçadas características em comum de repositórios populares e formulou-se as seguintes perguntas de pesquisas (Research Questions):

- RQ 01. Sistemas populares são maduros/antigos?
- RQ 02. Sistemas populares recebem muita contribuição externa?
- RQ 03. Sistemas populares lançam releases com frequência?
- RQ 04. Sistemas populares são atualizados com frequência?
- RQ 05. Sistemas populares são escritos nas linguagens mais populares?
- RQ 06. Sistemas populares possuem um alto percentual de issues fechadas?

Diante as perguntas formuladas, existem algumas hipóteses informais que podem ser levantadas, sistemas populares têm uma tendência a serem maduros já que passaram por diversas revisões até se tornarem um sistema confiável para a comunidade utilizar no dia-à-dia, por se tratar de um sistema open-source a partir do momento que esse repositório é relevante para a comunidade ele irá receber contribuições externas.

A taxa de releases deve ser frequente, pois isso traria um maior engajamento da comunidade. Esses repositórios também irão ter uma taxa de atualização maior pois a tecnologia está em constante evolução, seja com novas features, correções de bugs ou atualização de dependências.

Repositórios tendem a ser escritos em linguagens populares, pois o engajamento da comunidade muitas vezes está correlacionado a uma tecnologia que se tornou popular e este repositório pode agregar alguma coisa para essa tecnologia. E finalmente, sistemas populares tendem a ter um percentual de issues fechadas, já que ele deverá se manter constantemente atualizado.

Com o objetivo de caracterizar os repositórios populares do GitHub verificar a veracidade das hipóteses informais, utilizou-se a linguagem de programação Python, onde foi feito um script que comunica-se com a API do GitHub e faz uma chamada no GraphQL através de uma query para obter as respostas das perguntas de pesquisa. Inicialmente foi utilizado o script para pesquisar os 100 primeiros repositórios com maior quantidade de estrelas no GitHub, em seguida aumentou-se o valor para 1000 repositórios.

### **3. Resultados**

Após executar o programa em Python, foi possível obter o resultado da query contendo os 1000 repositórios,

### **4. Discussão**

### **5. Conclusão**

### **References**