

Características de Repositórios Populares

Ana Júlia Teixeira Cândido
anajuliateixeiracandido@gmail.com
Pontifícia Universidade Católica de
Minas Gerais - PUC Minas
Belo Horizonte, MG, Brasil

Marcella Ferreira Chaves Costa
marcellafccosta@gmail.com
Pontifícia Universidade Católica de
Minas Gerais - PUC Minas
Belo Horizonte, MG, Brasil

Davi José Ferreira
daviferreiradev@gmail.com
Pontifícia Universidade Católica de
Minas Gerais - PUC Minas
Belo Horizonte, MG, Brasil

1 INTRODUÇÃO

O GitHub é a principal plataforma de hospedagem de projetos *open-source* e concentra milhares de repositórios populares. Este laboratório tem como objetivo analisar os 1.000 repositórios mais estrelados, a fim de identificar padrões de popularidade, manutenção e evolução do software.

As questões de pesquisa abordam se sistemas populares tendem a ser mais antigos (RQ01), recebem muitas contribuições externas (RQ02), lançam *releases* com frequência (RQ03), são atualizados regularmente (RQ04), utilizam linguagens amplamente usadas (RQ05) e apresentam alto percentual de *issues* fechadas (RQ06). Opcionalmente, investiga-se também a influência da linguagem (RQ07).

- **RQ01 – Sistemas populares são maduros/antigos?**
 - *Métrica:* Data de criação do repositório.
 - *Hipótese:* Quanto maior a idade/maturidade de um repositório, maior será sua popularidade no mercado.
- **RQ02 – Sistemas populares recebem muita contribuição externa?**
 - *Métrica:* Quantidade de *pull requests* mergiados.
 - *Hipótese:* Repositórios com maior popularidade recebem significativamente mais *pull requests* de contribuidores externos.
- **RQ03 – Sistemas populares lançam *releases* com frequência?**
 - *Métrica:* Total de *releases* por repositório.
 - *Hipótese:* Repositórios populares lançam *releases* em intervalos curtos, mantendo frequência regular.
- **RQ04 – Sistemas populares são atualizados com frequência?**
 - *Métrica:* Tempo até a última atualização (dias desde o último *commit*/atividade).
 - *Hipótese:* Repositórios populares apresentam curto tempo até a última atualização.
- **RQ05 – Sistemas populares são escritos nas linguagens mais populares?**
 - *Métrica:* Linguagem de programação principal do repositório.
 - *Hipótese:* A popularidade de um repositório está diretamente relacionada à popularidade de sua linguagem de programação principal, com linguagens como JavaScript, Python e Java sendo predominantes.
- **RQ06 – Sistemas populares possuem um alto percentual de *issues* fechadas?**
 - *Métrica:* Percentual de *issues* fechadas, calculado como $(\text{issues fechadas} / (\text{issues abertas} + \text{issues fechadas})) * 100$
 - *Hipótese:* Repositórios populares tendem a ter um alto percentual de *issues* resolvidas, indicando boa manutenção e capacidade de gerenciar o feedback da comunidade.

2 METOLOGIA

3 RESULTADOS

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

4 DISCUSSÃO

5 CONCLUSÃO