

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Unidade Praça da Liberdade

Curso: Engenharia de Software

Prof.: Laerte Xavier

Disciplina: Laboratório de Experimentação de Software

Grupo: Anderson Heráclito e Cássia Ramos

- **Goal:**

Existe algum impacto das issues dos repositórios mais populares do github no stackoverflow?

- **Questions:**

- **RQ 01:** Com que frequência issues do GitHub são discutidas no StackOverflow?
- **RQ 02:** Qual o impacto das discussões de issues do GitHub no StackOverflow?
- **RQ 03:** Existe alguma relação entre a popularidade dos repositórios e o buzz gerado?
- **RQ 04:** Repositórios populares no github são mencionados em issues do stackoverflow?
- **RQ 05:** Linguagem mais utilizada nos 100 repositórios mais populares do github também é muito citada em perguntas do stackoverflow?
- **RQ 06:** Linguagem menos utilizada nos 100 repositórios mais populares do github também é pouco citada em perguntas do stackoverflow?

- **Metrics:**

- **RQ 01:** total de perguntas relacionadas (*PostType: Question*)
- **RQ 02:** total de respostas / total de perguntas relacionadas (*PostTypes: Answer / Question*)
- **RQ 03:** número de estrelas vs total de perguntas relacionadas

- **RQ 04:** Número de menções aos repositórios mais populares do github no stackoverflow;
- **RQ 05:** linguagem primária dos repositórios mais populares do github
- **RQ 06:** Quantidade de issues abertas no stackoverflow por linguagem
- **Dataset:**
 - As 10 issues mais antigas dos top-100 repositórios mais populares do github.
- **Análise**
 - **RQ 01:** Foram analisadas através do ID, as 10 issues mais antigas do top-100 repositórios mais populares do github, e dentro deste dataset não foram encontradas referências as issues do github no stackoverflow.
 - **RQ 02:** Como não foram encontradas referências de issues do github no stackoverflow não há impacto de issues nas discussões do stackoverflow.
 - **RQ 03:** Como não foram encontradas referências de issues do github no stackoverflow não há relação entre popularidade do repositório com o buzz gerado.
 - **RQ 04:** Foram coletados os nomes dos top-100 repositórios mais populares do github e o resultado gerado foi de 685.153 menções dos repositório nas discussões do stackoverflow.
 - **RQ05 e RQ06:** A linguagem mais utilizada nos top-100 repositórios encontrada foi a linguagem javascript sendo utilizada por 35 dos repositórios analisados e teve um total de 78998 menções nas discussões do StackOverflow, enquanto a linguagem PHP teve um total de 79396, ou seja mais citações do que javascript, e foi utilizada por apenas um repositório analisado. O que demonstra para as questões 5 e 6 não haver relação entre as linguagens mais usadas e menos usadas dos repositórios com as mais discutidas ou menos discutidas no stackoverflow.
- **Ameaças a validade**

Para que os resultados deste experimento sejam considerados válidos é importante analisar uma questão com relação às ameaças a validade, que são categorizados das seguintes maneiras:

 - Validade de construção: é o grau com o qual as variáveis independentes e dependentes medem o que elas se propõem a medir. As variáveis independentes são as issues coletadas por meio de query de consulta no graphql, de acordo com o dataset definido acima, enquanto que as variáveis dependentes são as menções a essas issues nas discussões do

StackOverflow, coletadas através de api de pesquisa própria do StackOverflow. Como todo o processo de coleta dos dados se deu de forma automatizada, pode-se considerar mínimas as ameaças a validade de construção.

- Validade Interna: esta ameaça está ligada ao quão confiável é a conclusão obtida pelo efeito das variáveis independentes sobre as variáveis dependentes.
- Validade Externa: diz respeito a replicabilidade dos resultados obtidos por este estudo. O uso de apenas um par de estudantes em todas as fases desta pesquisa pode restringir a forma como os resultados obtidos podem ser generalizados para a população em estudo e/ou outros contextos.

- **Conclusão**

Com base nos resultados obtidos por este experimento podemos concluir que não existe relação ou impacto das issues e linguagens utilizadas nos top-100 repositórios do GitHub nas discussões que são abertas no StackOverflow.