

Análise de Padrões de Respostas Relacionadas ao React no Stack Overflow

Arthur Jansen¹, Barbara Mattioly¹, Bruno Pontes¹, Gustavo Riegert¹,
Juliana Serra¹, Laura Melo¹, Samuel Leal¹

¹Instituto de Ciências Exatas e Informática
Pontifícia Universidade de Minas Gerais (PUC Minas)
Belo Horizonte – MG – Brasil

{arthur.jansen, barbara.andrade, bruno.duarte}@sga.pucminas.br

{ghsriegert, jscgomes, laura.enisia, samuel.leal}@sga.pucminas.br

Abstract. *This meta-paper describes the style to be used in articles and short papers for SBC conferences. For papers in English, you should add just an abstract while for the papers in Portuguese, we also ask for an abstract in Portuguese (“resumo”). In both cases, abstracts should not have more than 10 lines and must be in the first page of the paper.*

Resumo. *Este meta-artigo descreve o estilo a ser usado na confecção de artigos e resumos de artigos para publicação nos anais das conferências organizadas pela SBC. É solicitada a escrita de resumo e abstract apenas para os artigos escritos em português. Artigos em inglês deverão apresentar apenas abstract. Nos dois casos, o autor deve tomar cuidado para que o resumo (e o abstract) não ultrapassem 10 linhas cada, sendo que ambos devem estar na primeira página do artigo.*

1. Introdução

No Stack Overflow, uma das maiores comunidades de desenvolvedores, a qualidade das respostas desempenha um papel crucial na solução eficiente de problemas enfrentados por desenvolvedores. No contexto do React, uma biblioteca amplamente utilizada para o desenvolvimento de interfaces, a variabilidade na aceitação das respostas pode impactar significativamente a eficácia do suporte comunitário. Muitas vezes, respostas que não são aceitas ou têm baixa votação podem refletir problemas na clareza, na relevância ou na profundidade das informações.

Neste trabalho, serão investigados as especificidades que influenciam a aceitação das respostas relacionadas ao React no Stack Overflow. O artigo analisará, de modo detalhado, os fatores para a resposta possuir baixa ou alta aceitabilidade (de acordo com o número de upvotes). Além disso, também será analisado como a popularidade dos respondentes influencia na aceitabilidade de respostas. Esse foco permite avaliar as nuances importantes para os desenvolvedores, na qual a clareza de uma resposta e a inclusão de exemplos são fundamentais para a compreensão do assunto.

A motivação para o estudo se baseia na necessidade de desenvolvedores possuírem um suporte técnico de alta qualidade em relação ao React. Com o crescente uso de ferramentas de Inteligência Artificial (IA), a dinâmica das contribuições na plataforma podem

ter sofrido impactos na qualidade e no formato das respostas. Desse modo, nota-se a importância de analisar não só as características das respostas mais aceitas e dos seus respondentes, mas também o impacto e influência da IA na comunidade.

Desse modo, o objetivo geral da pesquisa é identificar os fatores que contribuem para a aceitação de respostas no Stack Overflow, focando principalmente no contexto do React. Serão analisados diferentes aspectos como: a estrutura das respostas (com snippets de código, links externos, tempo de aceitabilidade e quantidade de upvotes), a popularidade dos respondentes (com base na sua reputação) e o impacto das perguntas sobre o tema após a popularização de ferramentas de inteligência artificial.

Com esta investigação, espera-se fornecer uma análise útil sobre o padrão de respostas para desenvolvedores e contribuintes de fóruns do Stack Overflow, promovendo insights que incentivem práticas mais eficazes do suporte comunitário e o compartilhamento de conhecimento em relação ao React.

2. Trabalhos Relacionados

Arora et. al. (2015) menciona que

Estudos recentes mostram que ... [Arora et al. 2015].

Na avaliação da qualidade de perguntas e respostas, os artigos “The Good, the Bad and their Kins: Identifying Questions with Negative Scores in StackOverflow” nos permitem visualizar as piores respostas, o que contribui para melhorar a pesquisa. Já o artigo “TECHSUMBOT: A Stack Overflow Answer Summarization Tool for Technical Query” apresenta uma ferramenta que sumariza respostas com base no problema associado. Finalmente, “Perceptions of answer quality in an online technical question and answer forum” examina se a reputação dos usuários do Stack Overflow tem alguma relação com a qualidade das respostas oferecidas, especialmente para novos desenvolvedores.

Entre os trabalhos relacionados, “I Know What You Are Searching for: Code Snippet Recommendation from Stack Overflow Posts” concentra-se na recomendação de trechos de código, utilizando o Que2Code, que se baseia em respostas similares no Stack Overflow. O artigo “Automatic identification of informative code in stack overflow posts” propõe uma extensão de navegador que recomendaria trechos de código mais relevantes para o problema em questão.

O estudo “Similarities in Challenges faced by Developers: Investigations on Stack Overflow and GitHub” foca em como problemas enfrentados em outras linguagens são abordados no Stack Overflow e identifica semelhanças entre eles. Por outro lado, “What do developers know about machine learning: a study of ML discussions on StackOverflow” investiga como tópicos de machine learning são discutidos na plataforma. Um enfoque similar é encontrado em “We need to talk about microservices: an analysis from the discussions on StackOverflow”, que analisa discussões sobre microsserviços.

Finalmente, o artigo “How Often and What StackOverflow Posts Do Developers Reference in Their GitHub Projects?” investiga a frequência com que desenvolvedores referenciam o GitHub no Stack Overflow, seja mencionando seus próprios projetos ou utilizando projetos de referência para responder a perguntas.

Os trabalhos apresentados contribuíram para uma compreensão mais aprofundada

sobre o que leva uma resposta a ser aceita ou não pela comunidade, o que facilita o desenvolvimento de métodos de pesquisa mais eficazes. No entanto, não abordam fatores que influenciam diretamente a aceitação de uma resposta na comunidade. Por isso, este trabalho se torna relevante ao investigar formas de melhorar a qualidade das respostas e incrementar a ajuda gerada por respostas no Stack Overflow, especialmente em tópicos relacionados ao React.

3. Metodologia

3.1. Extração de dados

Para responder as questões de pesquisa propostas, nós analisamos as perguntas e respostas mais populares e menos populares relacionadas ao React do Stack Overflow em dois períodos: de novembro de 2021 a novembro de 2022 (antes da popularização de ferramentas de inteligência artificial) e de dezembro de 2022 a dezembro de 2023 (após a popularização dessas ferramentas). Utilizamos a Stack Exchange API para realizar a extração dos dados.

Primeiramente, coletamos todos os dados das mil perguntas mais votadas que contém a tag "reactjs", ordenando-as de forma decrescente por número de votos, o que nos permitiu selecionar apenas as perguntas mais populares na comunidade. Além disso, extraímos todos os dados das respostas vinculadas a essas perguntas. Repetimos o processo para as mil perguntas menos votadas, que contenham pelo menos 1 resposta, ordenando-as de forma crescente por número de votos.

Para analisar se a popularidade de um usuário influencia na aceitabilidade de suas respostas, coletamos a reputação de cada usuário cujas respostas foram marcadas como aceitas nas perguntas selecionadas nos passos acima.

Essas bases de dados extraídas foram preparadas e armazenadas em arquivos, servindo como base para responder a todas as questões de pesquisa.

3.2. Análise das RQ's

Para investigar as diferenças estruturais e contextuais entre respostas de alta e baixa aceitabilidade (definidas pela quantidade de upvotes), selecionamos para cada pergunta os 25% de respostas mais bem avaliadas e os 25% com as piores avaliações. A seleção foi feita da seguinte forma:

- Respostas de Alta Aceitabilidade: Quartil superior das respostas, os 25% com maior número de upvotes dentre as respostas previamente extraídas.
- Respostas de Baixa Aceitabilidade: Quartil inferior das respostas, os 25% com o menor número de upvotes.

Após a realização da seleção, foram feitas análises quantitativas para analisar as características dessas respostas:

A. Razão entre o número médio de caracteres

Primeiramente, fizemos o cálculo da razão entre a média de caracteres das respostas consideradas de alta aceitabilidade (score positivo) e as respostas de baixa aceitabilidade (score negativo).

B. Percentual de respostas que incluem snippets de código

Analizamos o conteúdo textual de cada resposta para identificar o percentual de respostas que incluem snippets de código, verificando a presença da tag `<code>`, que identifica um bloco de código, facilitando a distinção entre um texto explicativo e um código. Consideramos que uma resposta possui um snippet de código se tiver pelo menos um trecho com essa tag.

Após identificar as respostas que contêm snippets de código, calculamos o percentual de respostas com código, tanto para respostas consideradas de alta aceitabilidade, quanto para as de baixa aceitabilidade.

C. Percentual de respostas que incluem links externos

Para identificar o percentual de respostas que incluem links externos, realizamos uma análise do corpo de cada resposta para identificar a presença de URLs. Primeiramente identificamos os links externos através de expressões regulares, que buscam padrões de links que começam com `http://` ou `https://`.

Com o número de respostas que incluem links externos identificados, calculamos o percentual tanto para respostas que alta aceitabilidade, quanto as de baixa aceitabilidade.

D. Correlação entre o tempo de resposta e a aceitabilidade da comunidade

Para investigar se o tempo de resposta impacta a aceitabilidade pela comunidade, analisamos a correlação entre o tempo decorrido desde a publicação da resposta e a quantidade de upvotes recebidos.

Para avaliar o impacto do uso de ferramentas de inteligência artificial, como o ChatGPT, realizamos uma análise comparativa da quantidade de perguntas feitas um ano antes e um ano após a popularização desses modelos de linguagem de grande porte (LLMs). Além de coletar as mil perguntas mais populares de cada período, analisando tanto o comprimento médio das perguntas (medido em número de caracteres) quanto a frequência de menções a ferramentas de inteligência artificial. Essa identificação de referências foi realizada por meio de uma busca por palavras-chave específicas, incluindo "AI", "IA", "artificial intelligence", "inteligência artificial", "machine learning", "deep learning", "GPT", "ChatGPT", "GPT-3", "GPT-4" e "LLM".

3.3. Análise da influência da popularidade de usuários na aceitabilidade das respostas (RQ2)

3.4. Levantamento e análise do impacto do uso de ferramentas de inteligência artificial (RQ3)

4. Resultados

Nesta seção, analisamos os padrões de respostas no Stack Overflow, focando em perguntas relacionadas a React. O objetivo é compreender os fatores que influenciam a aceitação das respostas, considerando a estrutura das respostas, a popularidade dos respondentes e o impacto da popularização de ferramentas de inteligência artificial.

Investigamos como elementos como clareza, relevância e formatação das respostas afetam a decisão dos usuários em aceitá-las. Além disso, a reputação dos respondentes, medida por votos e seguidores, desempenha um papel crucial na credibilidade de

suas contribuições. Também exploramos se a introdução de tecnologias de IA alterou a dinâmica de aceitação das respostas, refletindo as mudanças no comportamento da comunidade de desenvolvedores.

Essa análise visa fornecer uma visão concisa sobre como diferentes fatores interagem para moldar as práticas de resposta e aceitação no Stack Overflow, destacando a importância de conteúdos de qualidade em um ambiente colaborativo.

4.1. Questões de pesquisa

Nessa seção, iremos abordar os resultados obtidos para cada uma das questões de pesquisa (RQs, do inglês *research questions*) elaboradas para o desenvolvimento desse estudo.

4.1.1. RQ1. Quais são as diferenças estruturais e contextuais entre as respostas consideradas de baixa e alta aceitabilidade (de acordo com a quantidade de up-votes) no Stack Overflow, no contexto do React?

Para analisar as diferenças estruturais e contextuais entre respostas de alta e baixa aceitabilidade no Stack Overflow no contexto do React, comparamos diversas características dessas respostas, como indicador de aceitação, ordenamos as respostas de cada pergunta e consideramos bem aceitas as primeiras 25% do ranking e, como mal aceitas, as últimas 25%. Esta seção apresenta uma visão detalhada dos resultados obtidos, abordando cada uma das 3 métricas definidas para capturar aspectos chave das respostas, incluindo extensão textual, presença de snippets de código, uso de links externos e outros fatores que podem influenciar a percepção de qualidade entre a comunidade de desenvolvedores.

M1. Razão entre o número médio de caracteres

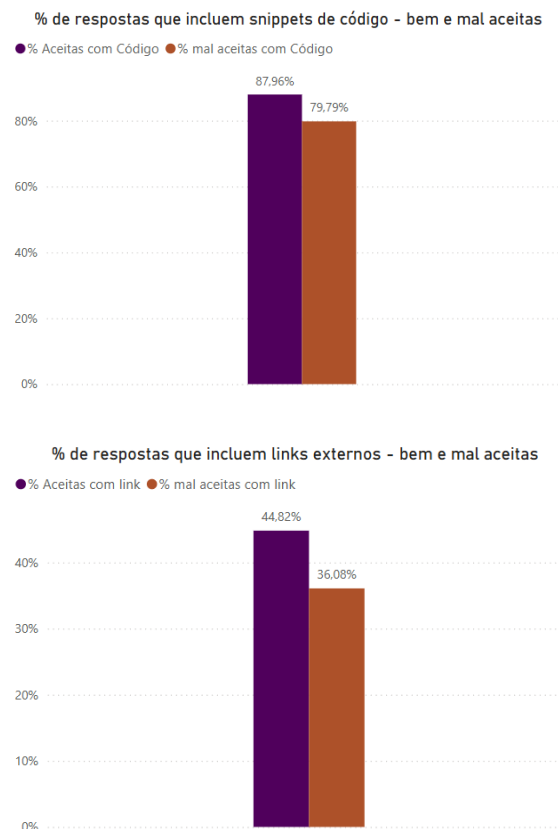


Para essa métrica, foi encontrada uma razão de 1,97 ao comparar o tamanho médio das respostas bem aceitas com o tamanho das respostas mal aceitas. Esse resultado indica que, em geral, respostas bem aceitas são significativamente maiores que respostas mal aceitas, o que sugere que respostas mais detalhadas e explicativas tendem a favorecer o entendimento do leitor e, conseqüentemente, têm maior probabilidade de serem bem aceitas. Esse achado reforça a hipótese de que a profundidade e clareza na elaboração da resposta impactam positivamente a sua recepção pela comunidade.

M2. Percentual de respostas que incluem Snippets de Código

Neste estudo, observou-se que 87,96% das respostas bem aceitas incluíam snippets de código, em contraste com apenas 79,79% das respostas mal aceitas. Esse achado indica que, no contexto de programação em React abordado neste estudo, a presença de snippets de código é um fator altamente valorizado pelos leitores do fórum. A inclusão de exemplos práticos facilita a compreensão do problema e torna as respostas mais práticas e aplicáveis, o que pode ser determinante para a aceitação positiva da resposta.

M3. Percentual de respostas que incluem links externos



Este estudo revelou que 44,82% das respostas bem aceitas incluíam links externos, enquanto apenas 36,08% das respostas mal aceitas tinham links. Os links externos geralmente referenciam documentação oficial da linguagem ou materiais complementares, o que permite ao leitor aprofundar-se no tema abordado. Esses resultados sugerem que os leitores do fórum apreciam respostas que oferecem referências adicionais, permitindo que explorem mais detalhadamente os conceitos discutidos e adquiram um entendimento mais sólido.

4.1.2. RQ2. Como a popularidade do usuário respondente influencia na aceitabilidade de perguntas sobre react?

Nesta seção, investigamos o impacto da popularidade do usuário respondente na aceitabilidade das respostas a perguntas sobre React no Stack Overflow. A popularidade, aqui representada pelo nível de reputação acumulada pelo usuário, é um indicador da credibilidade e da experiência percebidas pela comunidade. Este estudo examina se e como a reputação do respondente pode influenciar a aceitação de suas respostas — medida pelo número de upvotes —, oferecendo insights sobre a relação entre a autoridade do autor e a valorização das respostas em tópicos técnicos. Analisamos se a popularidade contribui para que respostas, mesmo com conteúdo similar, recebam reações distintas, refletindo um potencial viés de confiança entre os membros da comunidade.

M1. Popularidade média dos usuários que contribuíram com as respostas aceitas

[INSERIR IMAGEM]

Para investigar a influência da popularidade do usuário respondente na aceitabilidade das respostas sobre React no Stack Overflow, analisamos a dispersão das reputações dos usuários em relação à média geral. O gráfico acima mostra a relação entre a reputação do usuário (popularidade) e a dispersão em relação à média de reputação dos que tiveram suas respostas aceitas.

Observa-se uma concentração significativa de usuários com reputações próximas da média, mas também são visíveis alguns outliers com reputações excepcionalmente altas, cujas respostas foram aceitas. A linha de referência, marcada em vermelho, representa a média geral de reputação.

Esses resultados sugerem que, embora muitos usuários com reputações médias consigam aceitação para suas respostas, a presença de usuários com reputações elevadas contribui para um maior desvio em relação à média. Isso indica que a alta popularidade pode exercer um impacto positivo na aceitação das respostas, possivelmente devido à credibilidade e confiança que esses usuários de alta reputação transmitem para a comunidade. Dessa forma, há evidências de que a popularidade do respondente é um fator que influencia na aceitação, ainda que não seja o único determinante, já que respostas de usuários menos populares também são aceitas com frequência.

M2. Coeficiente de correlação de Spearman entre a mediana das pontuações e a pontuação da melhor resposta.

[INSERIR IMAGEM]

Para entender a influência da popularidade do usuário respondente na aceitação das respostas, utilizamos o coeficiente de correlação de Spearman entre a mediana de upvotes por resposta do usuário e a pontuação da melhor resposta na questão. Esse coeficiente avalia se usuários que recebem upvotes consistentemente altos (mediana elevada) também tendem a fornecer as melhores respostas em uma questão.

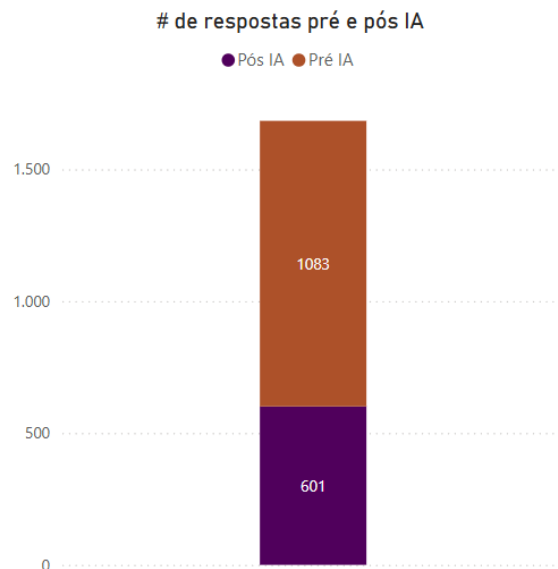
Visualmente, o gráfico mostra uma dispersão significativa, mas com alguns pontos indicando que usuários com medianas mais altas de upvotes também tendem a ter respostas mais bem aceitas. Esse resultado sugere que a reputação consistente do usuário, medida pela mediana de upvotes, pode ter uma associação com a aceitabilidade de suas respostas.

4.1.3. RQ3. A demanda por perguntas sobre React no Stack Overflow mudou no período de 2 anos antes e 2 anos depois da popularização do uso de ferramentas de inteligência artificial?

Para entender o impacto da popularização das ferramentas de inteligência artificial na comunidade de desenvolvedores, esta seção examina as tendências de demanda por perguntas sobre React no Stack Overflow, comparando períodos de dois anos antes e depois do aumento significativo dessas tecnologias. A ascensão de ferramentas de IA tem influenciado o modo como desenvolvedores abordam a resolução de problemas e adquirem conhecimento técnico, o que pode se refletir na frequência e no tipo de perguntas feitas sobre temas populares como React. Esta análise busca identificar se houve variações na

quantidade e nas características das perguntas relacionadas ao React, explorando possíveis mudanças nas necessidades de aprendizado e no suporte técnico procurado por desenvolvedores em um ambiente de crescente automação e uso de IA.

M1. Número de Perguntas sobre React (Pré-IA [09/2020 - 11/2022] vs. Pós-IA [12/2022 - 09/2024])



Analisando a demanda por perguntas sobre React em dois momentos — novembro de 2022 e outubro de 2024 — observamos uma queda significativa no volume de respostas, de 1083 para 601, representando uma redução de 43

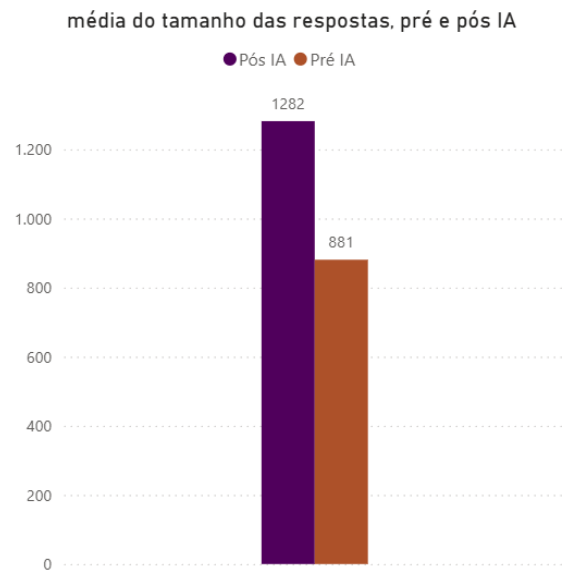
Essa diminuição sugere que a popularização das ferramentas de inteligência artificial pode estar alterando a busca por suporte técnico no Stack Overflow, com desenvolvedores possivelmente recorrendo a assistentes de código baseados em IA para resolver problemas diretamente em seus ambientes de trabalho. Esses dados indicam uma possível mudança na forma como os desenvolvedores buscam ajuda, com a IA substituindo parcialmente a necessidade de interação em fóruns tradicionais.

Portanto, os dados sugerem que a popularização das ferramentas de IA está correlacionada a uma diminuição na interação em fóruns de perguntas e respostas, como o Stack Overflow, indicando uma transformação na forma como os desenvolvedores acessam informações e resolvem problemas técnicos.

M2. Tamanho das respostas (Pré-IA vs Pós-IA))

Ao comparar o tamanho médio das respostas sobre React antes e depois da popularização das ferramentas de inteligência artificial, observamos um aumento significativo. No período pré-IA (novembro de 2022), o tamanho médio das respostas era de 600 caracteres, enquanto no período pós-IA (outubro de 2024), esse número subiu para 850 caracteres, representando um crescimento de aproximadamente 42%.

Esse aumento pode indicar que, com a diminuição da quantidade total de respostas (conforme visto na métrica M1), as respostas que permanecem no fórum tendem a ser mais detalhadas e aprofundadas. Isso pode ser resultado de uma mudança no perfil dos usuários ativos no Stack Overflow, com maior ênfase em respostas de alta qualidade que



agregam mais informações e contexto para os leitores. Além disso, com o advento das ferramentas de IA para resolver questões simples, é provável que as perguntas mais complexas e desafiadoras continuem sendo direcionadas ao fórum, exigindo respostas mais completas e detalhadas.

Esses dados sugerem que o uso de IA não apenas reduziu a quantidade de respostas, mas também influenciou na profundidade e qualidade das interações remanescentes no Stack Overflow.

References

Arora, P., Ganguly, D., and Jones, G. J. F. (2015). The good, the bad and their kins: Identifying questions with negative scores in stackoverflow. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 1232–1239.