Computer Science and Information Technologies
Vol. 4, No. 1, March 2023, pp. 14~23
ISSN: 2722-3221, DOI: 10.11591/csit.v4i1.pp14-23

A preliminary empirical study of react library related questions shared on stack overflow

Ganno Tribuana Kurniaji[1], Yusuf Sulistyo Nugroho[1], Syful Islam[2]
[1]Department of Informatics Engineering, Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia
[2]Department of Computer Science and Telecommunication Engineering, Faculty of Engineering and Technology, Noakhali Science and Technology University, Noakhali, Bangladesh

Article Info

Article history:

Received May 17, 2022
Revised Dec 16, 2022
Accepted Dec 30, 2022

Keywords:

Empirical study
Question
React
Stack overflow

ABSTRACT

React is a JavaScript library to develop user interfaces for single-page applications. Developers utilize react to build large web apps that allow users to update data without refreshing the page. Despite its benefits, many developers face react-related issues in the implementation. To find a solution, developers commonly shared and discussed their issues on stack overflow (SO). Although recent studies have demonstrated the benefits of utilizing react in web development, the trends of the users' attentions remain unknown. In this study, we conducted a preliminary empirical study of react library-related questions shared on SO. We applied an exploratory data analysis technique to investigate the distribution of problems shared by the developers. The findings reveal that although the quantity of react-related topics on SO has risen over time, community interest is beginning to decrease. This is shown by the increase of the unsolved questions and the decrease of the number of views per year. Regarding the react users' activity, most of them are more active in providing answers rather than commenting and providing scores. The findings of this study might point to future research that recommends approaches to assist the react community in overcoming issues while using react in the early phases.

Corresponding Author:

Ganno Tribuana Kurniaji
Department of Informatics Engineering, Faculty of Communication and Informatics
Universitas Muhammadiyah Surakarta
Surakarta, Indonesia
Email: l200184092@student.ums.ac.id

1. INTRODUCTION

React is a JavaScript library that allows users to develop custom user interfaces for single-page apps. Since react offers the benefits regarding its response, scale, and simplicity, developers have made use react to develop web apps so that the users can change the data without reloading the webpage [1]–[3]. React's main goal is to be quick, scalable, and simple. Despite the benefits, many developers face react-related issues in their implementation. To find solutions, developers share and discuss their issues on question-and-answer platforms like stack overflow (SO). Developers discuss coding issues and share code samples on SO [4], a famous Q&A website [5]. This website has grown into a vast knowledge base for developers to discuss technical difficulties and solve programming issues [6]. In addition, users may share their knowledge and experience on SO.

Several studies on the quality and content of online question-answering platforms, such as SO, have shown their significance. A research that investigated at discussions on SO about network simulators was

able to learn about the issues and difficulties that users have faced [7]. Prior research [8] describes a method for analyzing the textual content of SO discussions using latent dirichlet allocation (LDA), a topic modeling technique using statistics, to define the topics of developer communication automatically. In another research that conducted a qualitative study to categorize unanswered questions [9] reveals the characteristics of questions that would be difficult to find answers and also conducted experiments to determine the duration of questions remain unanswered on SO. A previous study used SO to combine the ideas of questions, types, and code to correlate programming concepts and identifiers (such as the String class) with particular question categories, such as "how to code" [10]. Other research [11] also analysed stack overflow (SO) to understand challenges and confusions developers face while dealing with privacy-related topics. Several researches also analyzed the topics of SO discussion. For example, a study that aims to automatically classify the SO question posts into seven question categories [12], a work that investigates the mobile-related questions topics asked by mobile developers on SO using LDA [13], research that presents an empirical study on the prevalence and severity of application programming interface (API) misuse on stack overflow [14], investigation about how the badge system steers the revision behavior of users on stack overflow [15], a survey that examines how secure coding practice is supported on stack overflow [16]. Despite recent research demonstrating the popularity and frequency of adopting react in web development, the trends of react-related discussions are still unknown.

In this paper, we performed a preliminary empirical study of react-related question posts shared on SO to analyze the trends of react-related discussions. We applied an exploratory data analysis (EDA), a technique in which the researcher examines the data without any preconceived ideas to discover what the data can say about the phenomenon being studied [17]. EDA is done by interactively applying various analysis operations (such as filtering, aggregation, and visualization) [18]. By applying this technique, we investigated the distribution of react-related questions shared by users and the popularity of the react-related questions. The dataset is extracted from Stack Exchange data archive, a network of question-and-answer (Q&A) web-based system covering a wide range of topics, with each site addressing a specific topic and a reputation awarding mechanism for questions, answers, and users. From the collected dataset, we filtered the dataset based on the specified tags related to react. Out of 22,087,195 question posts on SO, we obtained 447,542 react-related questions with 14 attributes.

The findings of this study reveal that although the quantity of react-related questions on SO has risen over time, the community's attentiveness to these topics has remained low. This is shown by the increase of the unsolved questions per year. We also found that most react-related topics covered on SO suggest the developers to consider user experience improvement. The attraction to react-related problems is starting to get no attention from the community. This means react topics that are discussed on stack overflow are starting to experience a decline in popularity. Regarding the responses of react users on SO, most of them prefer to interact with the community by providing answers rather than commenting and providing scores.

The rest of this paper is structured. Section 2 describes the research methodology in which we might perform the study. In detail, we explain the research questions, and data collection. The results of the study and their interpretation are described in section 3. In section 3 also present our implications and threats to validity. Lastly, section 4 presents the conclusion of this paper.

2. METHOD

In this section, we present the study setting to analyze the frequency and popularity of react-related questions shared on SO. In detail, we define research questions and data collection procedures. To guide the study, we formulate two main research questions with their motivations. In the data collection section, we present the results of data extraction of each step.

2.1. Research questions

2.1.1. RQ1: What is the frequency of react-related questions posted on SO?

RQ1.1: What is the trend of react-related questions on SO between 2012 and 2021?

Motivation: The objective of this RQ is to extract insights about the trend of react-related questions from 2012 to 2021. To analyze the trends, we utilize the total number of react-related questions shared by developers for the last 10 years. In addition, we also analyze the distribution of the questions that relate to react per year.

RQ1.2: How many react-related questions have accepted answers?

Motivation: Our research question is primarily motivated by a desire to determine the frequency of react-related question posts that have an accepted answer. There is a reason for this since the answers posted by users in each post are not necessarily the most appropriate solution for the problem being addressed. The answer of this question will show the SO users on the availability of react-related solutions.

2.1.2. RQ2: How popular are react-related questions on SO?

RQ2.1: How many answers, comments, and scores do react-related questions get on SO per year?

Motivation: Stack overflow questions contain important variables such as answers, comments, and scores. The scores of these variables will increase depending on the users' reactions. Thus, by analyzing the sum of each of these variables per year, we aim to determine the trend of the popularity of react-related questions on SO.

RQ2.2: What react-related questions are popular on SO?

Motivation: This research question aims to identify the most popular questions related to react on stack overflow. In order to conduct the analysis, two parameters are taken into consideration, namely the number of views and the number of scores. The number of views indicates the number of users who faced the same problems, while the number of scores describes the number of users who are interested in the question.

RQ2.3: How do the users respond to react-related questions?

Motivation: This research question aims to identify the most common strategies used by SO users to handle react-related issues. As a measure of the quality of the questions, we consider the number of questions that have answers. In addition, we also count the number of questions that have comments, and the number of questions that have a score.

2.2. Data collection

Our main goal is to empirically analyze react-related question posts shared on SO. We applied the EDA to analyze the frequency and popularity of react-related question posts shared on SO [19], [20]. As shown in Figure 1, we initially extracted the latest data from Stack Exchange data archive that is publicly available.

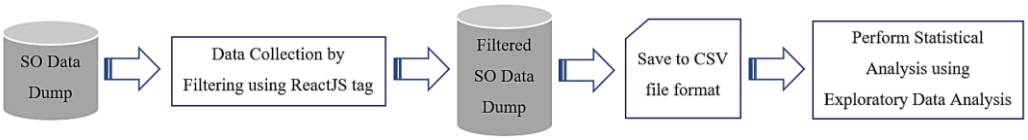

Figure 1. Overview of the study steps

As described in Figure 1, the steps of data collection are distributed into 3 main steps. First, we initially extracted the raw data from SO data dump. As shown in Table 1, the size of the initial collected data is 22,087,195 questions shared on SO that were recorded from January 2012 to December 2021.

To obtain the specific react-related questions on SO, in the second step, we defined 26 tags as presented in Table 2. By applying these 26 specified tags in the filtration, we obtained 568,419 question posts. The dataset of questions produced in this step has 14 attributes, as described in Table 3.

Table 1. Result of each step of data collection

Step	# SO Question Posts
Step 1: Initial raw data collection from SO	22,087,195
Step 2: Extraction of react-related questions using 26 specified tags	568,419
Step 3: Duplications removal	447,542

Table 2. Frequency of posts for each tag used in the filtration

No	Tag	# Posts	No	Tag	# Posts
1	reactjs	348,288	14	react-testing-library	2,085
2	react-native	110,093	15	react- javascript xml (jsx)	1,668
3	react-redux	20,642	16	react-select	1,618
4	react-router	20,274	17	react-dom	898
5	react-hooks	18,379	18	reactjs-flux	820
6	azure-active-directory	13,874	19	create-react-app	486
7	react-navigation	7,818	20	reactjs.net	169
8	react-native-android	7,208	21	reactjs-testutils	89
9	react-bootstrap	3,658	22	konvajs-reactjs	79
10	react-native-ios	3,395	23	reactjs-native	53
11	react-props	2,394	24	video-reactjs	20
12	react-native-flatlist	2,219	25	reactjs-popup	4
13	react-apollo	2,186	26	applicationinsights-react-js	2

Table 3. The attribute names and descriptions of the dataset

No	Attribute Name	Description
1	ID	The ID number of the post
2	PostTypeId	ID number of post type
3	AcceptedAnswerId	The ID number of the accepted answer in the post
4	OwnerUserId	The ID number of the post owner
5	AnswerCount	The number of answers in the post
6	CommentCount	The number of comments on the post
7	FavoriteCount	Number of favorites given by the user to the post
8	CommunityOwnedDate	The date and time the post was converted to community wiki
9	CreationDate	The date and time the post was created
10	Score	Number of favorites given by the user to the post
11	ViewCount	Number of users who viewed the post
12	Title	Post title
13	Body	Post explanation
14	Tags	topic hashtags used in discussion posts

Finally, since there are duplicated question posts in the filtered raw data, we removed the duplications. This procedure aims to reduce bias in our analyses. As described in Table 1, the duplication removal produced 447,542 SO posts.

3. RESULTS AND DISCUSSION

3.1. Results

The frequency and popularity of react-related SO posts were analyzed in this section. As described in section 2.1, we propose 2 research questions that are distributed into sub-questions. Each research question is presented in detail, along with the methodology and results.

3.1.1. RQ1: What is the frequency of react-related questions posted on SO?

RQ1.1: What is the trend of react-related questions on SO between 2012 and 2021?

Approach: In this research question, we used the final data of 447,542 react-related SO questions after removing the duplications, as described in Table 1. We classified this dataset based on the year that the questions were initially posted. From the results, we then analyzed the trends of the questions from 2012 to 2021.

Results: As shown in Figure 2, there was an increase in the number of react-related questions posted by users on SO in each year between 2012 and 2021. In comparison with the increase in the previous two years, a sharp growth in the number of react-related questions occurred in 2014 and 2020. This describes that the number of react-related questions asked on stack overflow forums is increasing steadily.

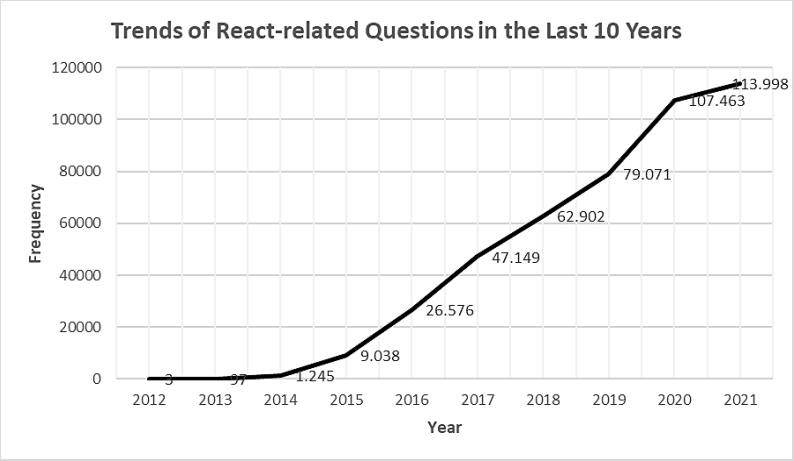

Figure 2. Trends of react-related questions from 2012 to 2021

RQ1.2: How many react-related questions with accepted answers?

Approach: Any questions asked on SO forum is open for users to respond to. However, the answers provided by the users are not always able to overcome the shared problems. To investigate the frequency

with which the askers solved questions, we compared the number of questions with accepted answers to the number of questions without accepted answers. When the user who asked the question marks an answer as accepted, it is considered an accepted answer [21]. To support the investigation, we used the same 447,542 react-related questions, as described in Table 1.

Results: Figure 3 shows that the frequency of questions with accepted answers is lower than the frequency of questions without accepted answers, with 41.76% to 58.24%, respectively. This means that most of the respondents are unable to provide the appropriate solutions to the react-related problems faced by the askers on SO. Since the availability of appropriate solutions is necessary, thus, based on this result, SO users are suggested to not fully dependent on SO discussion when facing react-related problems.

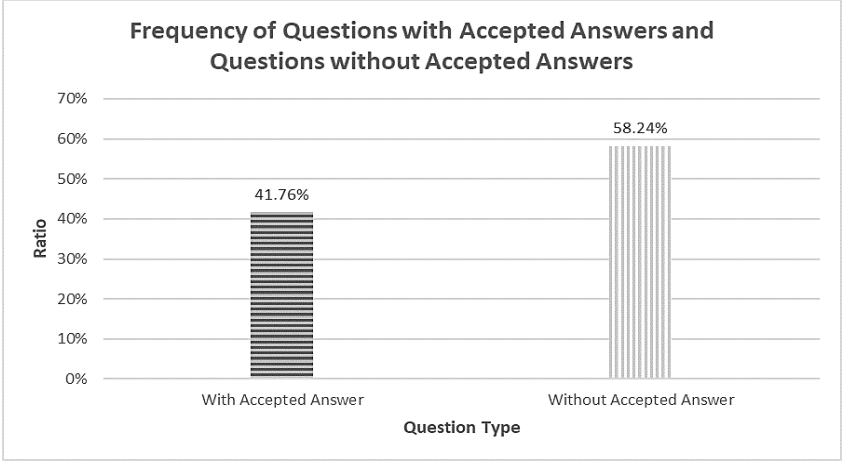

Figure 3. Comparison of the number between questions with accepted answers and questions without accepted answers in the last 10 years

To further investigate the users' ability in solving the problems, we also present the distribution of the questions with accepted answers per year. The result as shown in Figure 4, describes that the most react-related questions that have appropriate solutions to solve the problems occurred in 2013 and 2014, as many as 68% and 65.8% respectively. Despite the fact that the percentage of questions with accepted answers increased significantly in 2013, it began to decline the following year. This is inversely proportional to the frequency of questions that do not have an accepted answer. This can be interpreted that although the number of react-related questions on SO has increased over time, not many appropriate solutions are provided from the community. This leaves many of the issues discussed still unsolved.

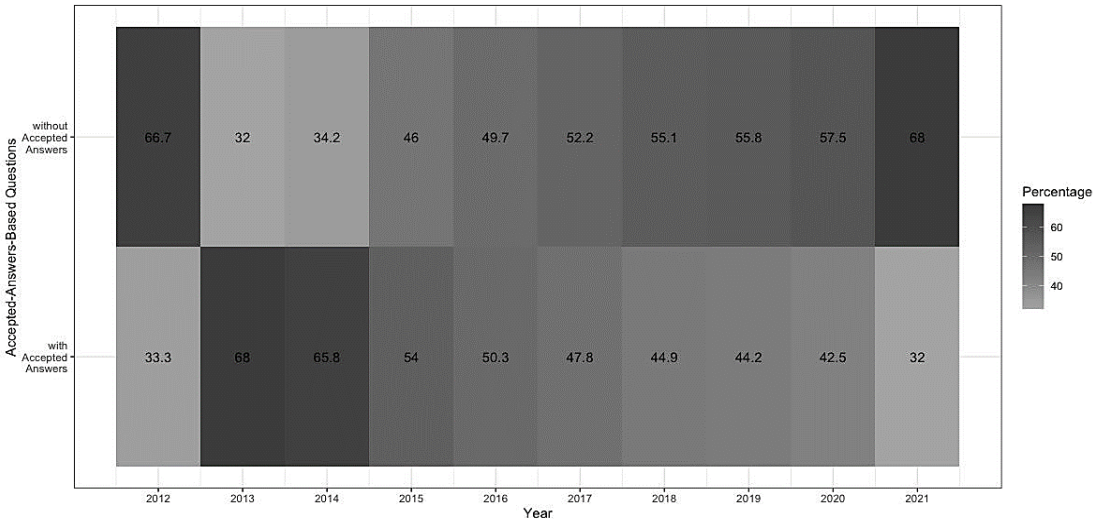

Figure 4. Ratio of questions with accepted answers and without accepted answers each year

3.1.2. RQ2: How popular are react-related questions on SO?

Beside the topics, the number of visitors on stack overflow might be also affected by several other factors, such as the score of the questions [22]. In addition, each question posted on stack overflow has included attributes, such as answer count, comment count, favorite count, scores, and view count. For users, these features can be utilized to consider the quality of both questions and answers to refer to. Thus, to understand the popularity of the react-related questions shared on SO, we further empirically investigate the question scores, the number of answers, comments, favorites, and views [23] for each post related to react. To facilitate understanding the popularity of react-related questions, we break down the analysis into 3 sub-questions.

RQ2.1: How many answers, comments, and scores of react-related questions on SO per year?

Approach: To find out how most react users respond to questions, we collect the number of answers, comments, and scores that are contained in the dataset. Furthermore, we also classify these parameters by year to see the statistics generated in each year. From these three variables, we do a comparison to find out how most react users respond to the questions.

Results: From the graph shown in Figure 5, it can be seen that the number of questions that have the most answers occurs in 2020, accounting for 86,883 questions. The questions with the most comments also occurred in 2020, as many as 58,006 questions. For questions that have the highest score also occur in 2020, as many as 47,418 questions. From 2012 to 2017, the number of questions that have a score is more than the number of questions that have comments. However, from 2018 to 2021, the number of questions that have a score is less than the number of questions that have comments.

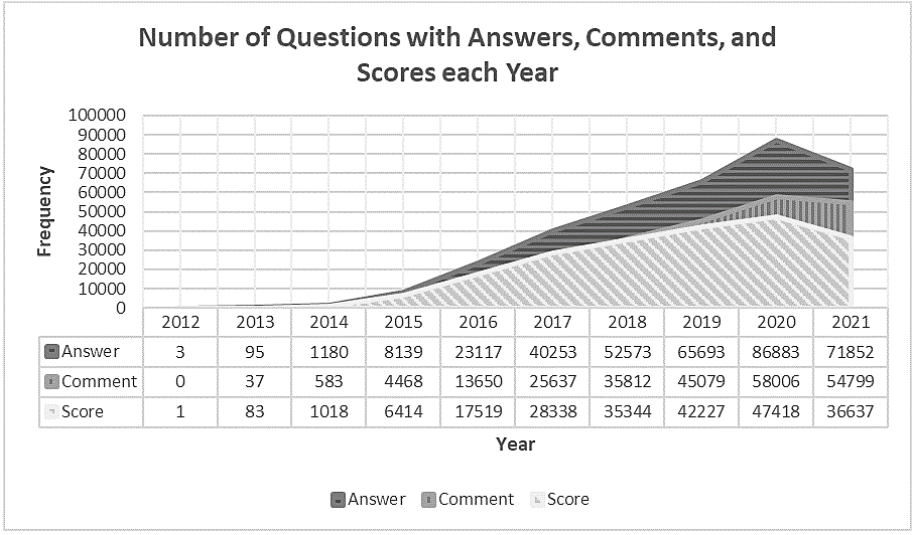

Figure 5. Number of questions with answers, comments, and scores each year

Figure 5 also shows the number of questions that have answers is always more than the number of questions that have comments and the number of questions that have a score. This indicates that react users tend to prefer to interact by providing answers rather than commenting and giving scores. Based on the data obtained, it can be concluded that 2020 is the culmination of all aspects, namely the number of questions that have answers, the number of questions that have comments, and the number of questions that have scores. In that year, the interaction between the askers and responders was interactive. This can be seen from the high activity in answering questions, also being active in providing comments on the question, and giving scores. In previous years, the number of all parameters always increased every year until 2020. However, in the following year, the trends started to decline.

RQ2.2: What react-related questions are popular on SO?

Approach: To understand the most popular react-related questions posted on SO discussion forums, we analyzed the data based on the most views and the highest scores. From the data, we extracted both attributes, score and view count, from all questions. We then extracted the top 10 react-related questions based on the highest number of views and top scores.

Results: As described in Table 4, it can be seen that the question that has the most views is "Loop inside react JSX" with 1,215,263 views. Judging from the number of views, this question has caught the

attention of many react users. From the description in Table 5, it can be seen that the question that has the highest score is "Programmatically navigation using react router" with a score of 1,917. With such a large number of scores, it can indirectly conclude that the question is a very high-quality question. Whereas the question "Loop within react JSX" gets the most views, the amount of scores collected is only the second highest. On the other hand, the question "Programmatically navigate using react router" has the highest number of scores but the number of views obtained is only the fifth highest. This shows that questions that have a large number of views do not necessarily get a large number of scores. and vice versa, questions that get a large number of scores do not necessarily also have a large number of views. So, the popularity of the question does not guarantee that it will be useful for solving problems faced by react users.

Table 4. Top 10 questions with most views

No	Title	View Counts
1	Loop inside react JSX	1,215,263
2	How to get parameter value from query string?	1,210,237
3	Can you force a react component to rerender without calling setState?	1,157,385
4	"SyntaxError: Unexpected token < in JSON at position 0"	1,155,750
5	Programmatically navigate using react router	1,109,796
6	Show or hide element in react	1,001,782
7	Understanding unique keys for array children in react.js	911,107
8	How to push to History in react Router v4?	836,258
9	React.js inline style best practices	761,918
10	React js onClick can't pass value to method	735,004

Table 5. Top 10 questions with most scores

No	Title	Score
1	Programmatically navigate using react router	1,917
2	Loop inside react JSX	1,638
3	What do these three dots in react do?	1,253
4	How to pass props to {this props children}	1,198
5	Why use Redux over Facebook Flux?	1,124
6	How to dispatch a Redux action with a timeout?	1,004
7	React-router urls don't work when refreshing or writing manually	971
8	What is the difference between react Native and react?	928
9	Can you force a react component to rerender without calling setState?	888
10	Understanding unique keys for array children in react.js	865

RQ2.3: How do the users respond to react-related questions?

Approach: To answer this question, we summarized the quantity of answers and comments of each question, number of questions marked as favorite, number of questions that have a score, and the total number of views. To further investigate the popularity of the react-related questions, we empirically studied these factors per year.

Results: As can be seen in Figure 6, generally, the attention from the react-related developers on SO increased in the last decade. This is shown by the growth of the number of answers and comments to respond to the react-related problems. However, although the attention to the react-related questions rose, the shared problems were not attractive to the community. This can be proven by the decrease of the number of questions marked as favorite and get scores since 2016. Furthermore, the react-related questions that have the most answers and comments happened in 2020, accounting at 23.37% and 24.44%, respectively. This shows that in 2020, react-related developers paid more attention to the problems shared on SO than that in previous years. On the other hand, the highest frequency of react-related questions that are marked as favorite occurred in 2016, as many as 19.89%. The number of questions that received more scores is found in 2017, accounting at 18.51%. This describes that the react-related problems shared by the developers in both 2016 and 2017 are more attractive for the community compared to the other periods.

The popularity of the react-related questions can be seen from their number of views on SO. Figure 7 depicts that the number of views of react-related questions increased between 2012 and 2017, in which it reached the most views in 2017 at 21.53%. The number of views of the questions amongst developers subsequently declined from 2018 to 2021. This can be interpreted that although the attraction of developers to react-related problems increased until 2017, however, it was starting to get no attention from the community since then. This can cause the askers to find it increasingly difficult in getting responses from the community to solve the problems.

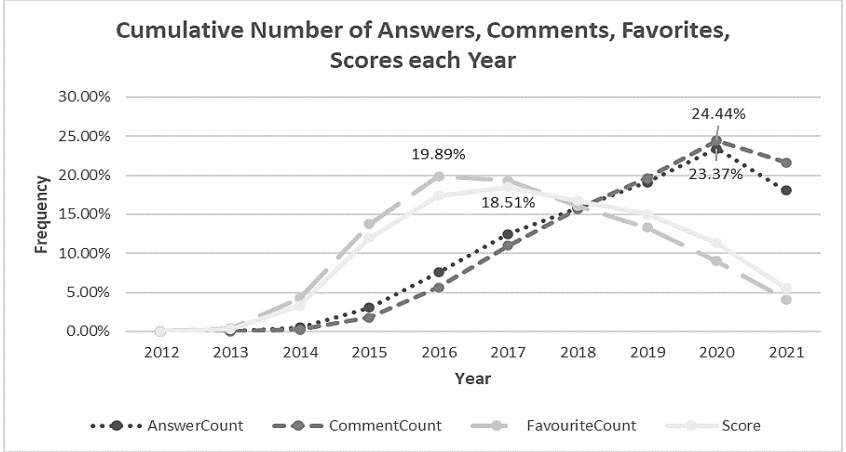

Figure 6. Cumulative number of answers, comments, favorite questions, and scores each year

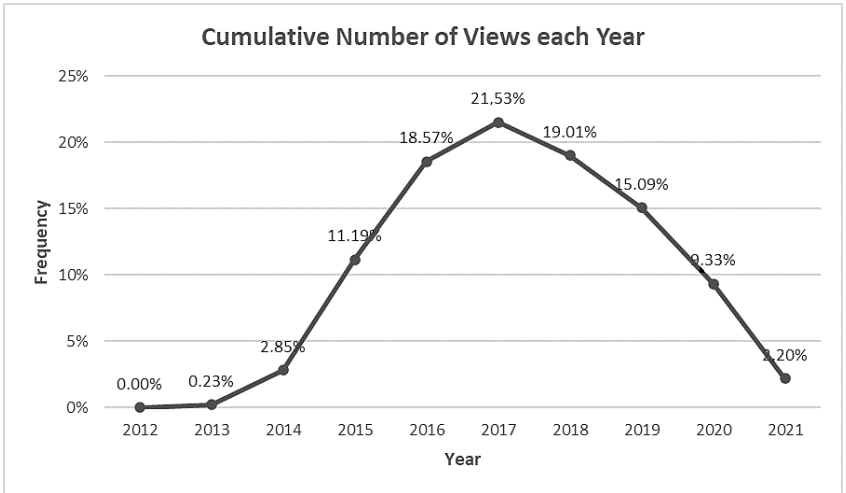

Figure 7. Cumulative number of react-related question views each year

3.2. Implications

The findings of this study will assist react users in better understanding and concentrating on the most important react library issues. This section explains how our findings can benefit practitioners to make use of specific discussion forums more intense. Furthermore, the results of this study can also be useful to researchers as a starting point for their future works on react-related discussion shared on SO.

3.2.1. Practitioners

Based on Figure 3, there are still a lot of react-related questions that do not have accepted answers. As shown in Figure 4, the frequency of questions that do not have accepted answers increases every year. This of course will make it difficult for react users to find solutions to address their problems.

The findings in this paper suggest that software developers need to consider about and prepare a specific react-related discussion forum. As observed in research, despite the lack of community interest, the number of react-related queries on SO has risen over time, as seen in Figure 2. Sharing react-specific issues in a community forum increases the probability of finding solutions. Communication channels may be used by react-related projects to collect new knowledge and update current knowledge. Due to its importance for the community to discuss difficulties, a dedicated react discussion forum is required [24], [25].

3.2.2. Researchers

Our preliminary empirical analysis shows the top views on react-related issues addressed on the SO platform. Based on Table 4 and Table 5 we find some of the top react-related topics covered on SO, namely

loops inside react JSX, getting parameter values from query strings and programmatically navigating. Therefore, we can suggest researchers to investigate the nature of challenges, such as classification of the questions and the reputation of the users who asked those questions.

3.3. Threats to validity

There are various threats that might compromise the validity of this study. In this section, we present the threats to validity in depth. We specified the threats into 2 types, that are, construct validity and external validity.

3.3.1. Construct validity

In our research, we discovered several possible challenges to construct validity. Regarding the tags that we used for topic searches related to react, it is possible that we might have missed some tags that could affect the number of the extracted topics. Since the increase in prevalence by this problem is low, the impact of missing tags is negligible.

3.3.2. External validity

External validity issues arise during data preparation. Although we investigated large data of react-related topics on stack overflow, the results cannot be generalized to other question-and-answer forums, and other discussion topics. More specific, in this paper, we did not perform an in-depth analysis to study the factors to identify the rationale why the shared react-related problems were not attractive to the community. This may open the opportunity for other researchers to perform further study on react-related discussions on SO.

4. CONCLUSION

In this paper, we conducted a preliminary empirical analysis of react library-related question posts on SO. To accomplish the objective, we applied exploratory data analysis technique to investigate the frequency and popularity of react-related problems shared by developers on SO. The findings of this preliminary study show that the quantity of react-related questions on SO has risen over time, but the community's interest has remained low. As a result, many of the problems raised remain unresolved. Our study also provides the top views of react-related discussions on SO. In addition, react users prefer to provide answers rather than give comments and provide scores. Our findings may open future research to investigate the most error types of questions asked by users and the reputation of the users who asked the questions.

REFERENCES

[1] P. Rawat and A. N. Mahajan, "ReactJS: A modern web development framework," *International Journal of Innovative Science and Research Technology*, vol. 5, no. 11, pp. 698–702, 2020, [Online]. Available: www.ijisrt.com.

[2] S. Aggarwal, "Modern web-development using reactJS," *International Journal of Recent Research Aspects*, vol. 5, no. 1, pp. 133–137, 2018.

[3] S. Domes, "Progressive web apps with react and firebase," *Progressive Web Apps with Angular*, 2017.

[4] Y. Wu, S. Wang, C. P. Bezemer, and K. Inoue, "How do developers utilize source code from Stack Overflow?," *Empirical Software Engineering*, vol. 24, no. 2, pp. 637–673, 2019, doi: 10.1007/s10664-018-9634-5.

[5] J. Zhou, S. Wang, C. P. Bezemer, and A. E. Hassan, "Bounties on technical Q&A sites: a case study of Stack Overflow bounties," *Empirical Software Engineering*, vol. 25, no. 1, pp. 139–177, 2020, doi: 10.1007/s10664-019-09744-3.

[6] L. An, O. Mlouki, F. Khomh, and G. Antoniol, "Stack Overflow: A code laundering platform?," *SANER 2017 - 24th IEEE International Conference on Software Analysis, Evolution, and Reengineering*, pp. 283–293, 2017, doi: 10.1109/SANER.2017.7884629.

[7] S. Islam, Y. S. Nugroho, and M. Javed Hossain, "What network simulator questions do users ask? A large-scale study of Stack Overflow posts," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, pp. 1622–1633, 2021, doi: 10.11591/ijeecs.v21.i3.pp1622-1633.

[8] A. Barua, S. W. Thomas, and A. E. Hassan, "What are developers talking about? An analysis of topics and trends in Stack Overflow," *Empirical Software Engineering*, vol. 19, no. 3, pp. 619–654, 2014, doi: 10.1007/s10664-012-9231-y.

[9] M. Asaduzzaman, A. S. Mashiyat, C. K. Roy, and K. A. Schneider, "Answering questions about unanswered questions of Stack Overflow," *IEEE International Working Conference on Mining Software Repositories*, pp. 97–100, 2013, doi: 10.1109/MSR.2013.6624015.

[10] M. Allamanis and C. Sutton, "Why, when, and what: Analyzing Stack Overflow questions by topic, type, and code," *IEEE International Working Conference on Mining Software Repositories*, pp. 53–56, 2013, doi: 10.1109/MSR.2013.6624004.

[11] M. Tahaei, K. Vaniea, and N. Saphra, "Understanding privacy-related questions on Stack Overflow," *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1–14, 2020, doi: 10.1145/3313831.3376768.

[12] S. Beyer, C. Macho, M. Di Penta, and M. Pinzger, "What kind of questions do developers ask on Stack Overflow? A comparison of automated approaches to classify posts into question categories," *Empirical Software Engineering*, vol. 25, no. 3, pp. 2258–2301, 2020, doi: 10.1007/s10664-019-09758-x.

[13] C. Rosen and E. Shihab, "What are mobile developers asking about? A large scale study using Stack Overflow," *Empirical Software Engineering*, vol. 21, no. 3, pp. 1192–1223, 2016, doi: 10.1007/s10664-015-9379-3.

[14] T. Zhang, G. Upadhyaya, A. Reinhardt, H. Rajan, and M. Kim, "Are code examples on an online Q&A forum reliable?: A study

of API misuse on Stack Overflow," *Proceedings - International Conference on Software Engineering*, pp. 886–896, 2018, doi: 10.1145/3180155.3180260.
[15] S. Wang, T. H. Chen, and A. E. Hassan, "How do users revise answers on technical QA websites? A case study on Stack Overflow," *IEEE Transactions on Software Engineering*, vol. 46, no. 9, pp. 1024–1038, 2020, doi: 10.1109/TSE.2018.2874470.
[16] T. Lopez, T. Tun, A. Bandara, L. Mark, B. Nuseibeh, and H. Sharp, "An anatomy of security conversations in Stack Overflow," *Proceedings - 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society, ICSE-SEIS 2019*, pp. 31–40, 2019, doi: 10.1109/ICSE-SEIS.2019.00012.
[17] A. T. Jebb, S. Parrigon, and S. E. Woo, "Exploratory data analysis as a foundation of inductive research," *Human Resource Management Review*, vol. 27, no. 2, pp. 265–276, 2017, doi: 10.1016/j.hrmr.2016.08.003.
[18] T. Milo and A. Somech, "Automating exploratory data analysis via machine learning: An overview," *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 2617–2622, 2020, doi: 10.1145/3318464.3383126.
[19] W. L. Martinez and A. R. Martinez, "Exploratory data analysis with MATLAB®," *Exploratory Data Analysis with MATLAB*, pp. 1–405, 2004, doi: 10.1198/tech.2006.s392.
[20] K. Sahoo, A. K. Samal, J. Pramanik, and S. K. Pani, "Exploratory data analysis using python," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, pp. 4727–4735, 2019, doi: 10.35940/ijitee.L3591.1081219.
[21] M. Mofijul Islam *et al.*, "RAiTA: Recommending accepted answer using textual metadata," *Advances in Intelligent Systems and Computing*, vol. 813, pp. 119–131, 2019, doi: 10.1007/978-981-13-1498-8_11.
[22] A. Merchant, D. Shah, G. S. Bhatia, A. Ghosh, and P. Kumaraguru, "Signals matter: Understanding popularity and impact of users on Stack Overflow," *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, pp. 3086–3092, 2019, doi: 10.1145/3308558.3313583.
[23] S. Meldrum, S. A. Licorish, C. A. Owen, and B. T. R. Savarimuthu, "Understanding Stack Overflow code quality: A recommendation of caution," *Science of Computer Programming*, vol. 199, 2020, doi: 10.1016/j.scico.2020.102516.
[24] J. Tantisuwankul *et al.*, "A topological analysis of communication channels for knowledge sharing in contemporary GitHub projects," *Journal of Systems and Software*, vol. 158, 2019, doi: 10.1016/j.jss.2019.110416.
[25] Y. S. Nugroho *et al.*, "How are project-specific forums utilized? A study of participation, content, and sentiment in the eclipse ecosystem," *Empirical Software Engineering*, vol. 26, no. 6, 2021, doi: 10.1007/s10664-021-10032-2.

BIOGRAPHIES OF AUTHORS

Ganno Tribuana Kurniaji is an undergraduate student at Universitas Muhammadiyah Surakarta, Indonesia. His research interests include empirical software engineering and mining Stack Overflow. He can be contacted at email: l200184092@student.ums.ac.id.

Yusuf Sulistyo Nugroho is a lecturer at the Department of Informatics, Universitas Muhammadiyah Surakarta, Indonesia. He received his Ph.D. degree from Nara Institute of Science and Technology in 2020. His research interests include empirical software engineering, software documentation, and mining software repositories. He can be contacted at email: yusuf.nugroho@ums.ac.id.

Syful Islam received the M.E., and Ph.D. degree in Engineering from Nara Institute of Science and Technology, Japan. At present, he is working as an assistant professor at Noakhali Science and Technology University, Bangladesh. His research interests include software ecosystem and mining Stack Overflow. He can be contacted at email: syfulcste@nstu.edu.bd.