

Enriching Release Notes with Reddit Feedback: Understanding How User Feedback Can Augment Maintaining Software Updates

B. Singh^{a*}, S. M. Daftary^a, M. Rudra^a, S. P. Veed^a, S. Berhe^a, M. Maynard^b, F. Khomh^c

^aUniversity of the Pacific, 3601 Pacific Ave., Stockton CA 95211, USA

^bData Independence LLC, 3 Settlers Way. Ellington, CT 06029, USA

^cPolytechnique Montreal, 2500 Chem de Polytechnique, Montreal, QC H3T 1J4, Canada

ABSTRACT

The past decade has shown that the number of software components has increased, as well as their inter-connectivity and update cycles. Maintaining daily software updates is a time-consuming, arduous, and risky task. One common question is which update to review and perform next. Our previous work leveraging software release notes showed limitations in classifying and prioritizing breaking and high-impact updates. In this research, we enrich release notes with Reddit-based user feedback to understand how it potentially improves maintaining software updates. Towards the objective, we review five research questions useful to a clinical ecosystem use case scenario, involving nine software components, each enriched with at least 100 Reddit posts. The results show that enriched Reddit user feedback complements release notes and improves the classification of breaking and high-impact updates. For example, identifying an average of 5% critical breaking updates posted by Reddit users improves the early evaluation of potentially dysfunctional software components. Additionally, linking an update to an average of 10 other component helps in identifying potential dependencies among related components. Given the results, our prototyping efforts illustrate how, for example, a clinical ecosystem can accelerate maintaining software updates by proposing next update recommendations. By integrating user feedback from Reddit, we aim to better understand where user feedback is useful in maintaining software updates.

Keywords: Software Maintenance, Reddit, Impact, Updates

I. INTRODUCTION

Today's software ecosystems comprise numerous independently developed, operated, and maintained software components and stakeholders. The variability of third-party user devices, firmware, and operating systems complicates achieving comprehensive pre-release test coverage. Consequently, software ecosystem teams require continuous monitoring, classification, and impact evaluation of user and market feedback to maintain software quality and comply with legal requirements in various domains. However, given the frequency of daily device and software updates, manually performing these tasks is not feasible. In recent years, beyond app store reviews, active social media platforms such as Twitter and Reddit have become popular for sharing user feedback on software updates (see Tab. 1).

Our previous work mined software release notes and Common Vulnerabilities and Exposures (CVE) references to identify and prioritize high-impact updates (Berhe, Kan, Khan, Pader, Farooqui, Maynard, & Khomh, in press). However, for example, critical updates due to major bugs are often not documented in release notes or CVEs, especially when the root cause is not well

understood or when the issue does not meet the specific criteria for CVE classification (see Fig. 1). This lack of documentation can hinder the timely identification and resolution of critical issues, leading to potential legal compliance failures, degraded system functionality, and a negative user experience. Effective monitoring and documentation of market user feedback can be essential to mitigate these risks and maintain the integrity and reliability of software ecosystems. To illustrate, a recent operating system update released in May 2024 led to reports in June 2024 from Reddit users that deleted photos were resurfacing on their phones (u/ios-user, 2024). While there is no technical review process to verify these user posts, they can still be critical, especially in sensitive clinical ecosystems, where such bugs could lead to significant legal compliance failures. Monitoring these potential critical bugs is challenging as they can both occur at and impact any phase of the software engineering lifecycle.

- **Requirements:** Identifying posts on critical updates affecting requirements (e.g., legal requirements or configuring minimum supported OS versions).
- **Architecture:** Identifying posts on critical updates affecting architecture (e.g., choosing frameworks, libraries, or databases).
- **Implementation:** Identifying posts on critical updates affecting development (e.g., updating the IDE).
- **Testing:** Identifying posts on critical updates affecting testing priorities (e.g., frequent bugs on specific devices).
- **Maintenance:** Identifying posts on critical updates affecting maintenance (e.g., legal, or operational impacts).

This paper evaluates an automated process for enriching software release notes with Reddit-based user feedback to accelerate the impact evaluation and prioritization of software updates by reviewing a set of research questions. In particular, to demonstrate the practical application of our approach, we apply our methodology to a clinical ecosystem software use case application. In the clinical domain, rapid and accurate response to software updates is critical due to stringent privacy regulations and sensitive healthcare data. Our aim is to lead to faster identification and rectification of software issues, preventing potential breaches of patient confidentiality and ensuring compliance with regulations such as HIPAA (Health Insurance Portability and Accountability Act).

* Corresponding author E-mail: b_singh13@u.pacific.edu

Table 1: Comparison of Related Work

Author	Data Source	Classification	Objective	SE Phase
Horawalavithana, 2022	Reddit	Growth Prediction	Predict discussion	Post-Release
Li, 2021	Reddit	Privacy Discussion	Understanding discussions	All Phases
Shrestha, 2020	Twitter, Reddit	Vulnerability Spread	Analyzing spread	Post-Release
Achimescu, 2021	Reddit	Disinformation	Monitoring disinformation	All Phases
Aggarwal, 2021	Reddit	Sentiment Analysis	Analyzing user interactions	All Phases
Iqbal, 2021	Reddit	Feedback Type	Investigating usefulness for RE	Requirements
Srewuttanapitikul, 2016	User Feedback	Feedback Type	Prioritizing maintenance	All Phases
Our Work	Reddit, Release Notes	Feedback Impact	Impact evaluates updates	All Phases

The objective of this work is to develop and evaluate an automated process for enriching software release notes with Reddit-based user feedback to improve the classification and impact evaluation of software updates. To achieve this objective, we aim to address a set of research questions and evaluate them against the clinical application use case.

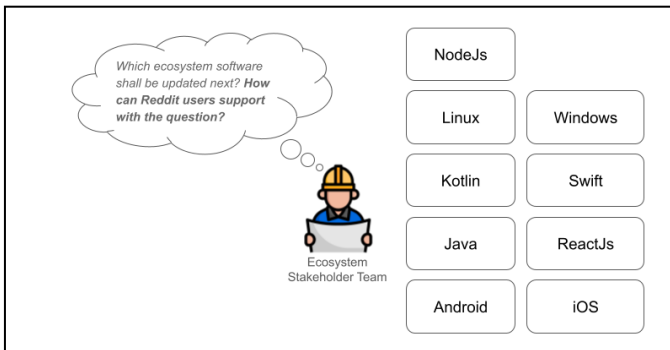


Fig. 1. Ecosystem Stakeholder Team

The remainder of this paper is organized as follows: Section II reviews related work. Section III describes the methodology, including data collection and analysis techniques. Section IV presents the results of our study. Section V discusses the implications of our findings, particularly in the context of clinical ecosystem software. Finally, Section VI concludes the paper and suggests directions for future research, emphasizing the potential for broader applications of our methodology in other high-stakes domains.

II. RELATED WORK

The related work in the field primarily focuses on various aspects of user feedback and data analysis from social media platforms like Reddit and Twitter. Srewuttanapitikul and Muengchaisri (2016) proposed a method that uses natural language processing (NLP) to analyze user feedback in order to extract defect related keywords, highlighting the importance of user feedback in software maintenance planning. Horawalavithana, Shrestha, and Iqbal (2019) examined feedback types on Reddit and Twitter to compare posts before

and after software updates, emphasizing post-release phases. Shrestha, Shrestha, and Aggarwal (2020) analyzed the spread of vulnerability discussions across GitHub, Twitter, and Reddit, focusing on post-release phases. Achimescu, Shrestha, and Li (2021) monitored perceived disinformation on Reddit, employing disinformation classification across all phases of the software lifecycle. Aggarwal, Horawalavithana, and Li (2021) conducted sentiment analysis to understand user interactions and sentiments on Reddit, relevant to all phases. Iqbal, Achimescu, and Horawalavithana (2021) explored the usefulness of Reddit data for software engineering, specifically targeting the requirements phase. Li, Aggarwal, and Muengchaisri (2021) investigated developer discussions about personal data on Reddit to understand implications for user privacy. Horawalavithana (2022) predicted the growth of discussion threads on Reddit, highlighting post-release monitoring. To contrast, our work integrates and enriches Reddit user feedback with release notes data through an automated classification of feedback impact and prioritization of related software update review. Our research focuses on the impact evaluation of software updates, encompassing all phases of the software lifecycle. The aim is to not only improve the granularity of feedback classification but also accelerate the evaluation and prioritization process, addressing the critical need for timely updates and compliance in high-stakes domains like healthcare.

III. METHOD

To understand how Reddit user posts can improve the maintenance of software updates, we will first design a clinical application ecosystem use case scenario encompassing high impact software components. Next, we will formulate a set of research questions to evaluate the effectiveness of our classification and prioritization model. Finally, we will introduce an Extract, Load, and Transform (ELT) process tailored to both the release notes dataset and the Reddit dataset, with a focus on meeting the specific requirements of the clinical use case. This process aims to seamlessly integrate and analyze user feedback and release notes, aiming comprehensive support for maintenance activities in a clinical software ecosystem (see Fig. 2).

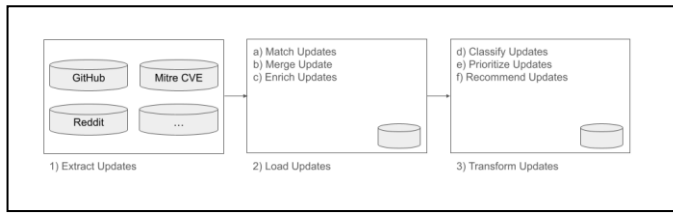


Fig. 2. Process Flow for Extracting, Loading, and Transforming Reddit and Release Notes Updates Data

A. Evaluation

1) Use Case Design

The clinical application ecosystem is a complex network of software and hardware components used for communicating clinical digital content, including photos, voice, videos, and text, between the phone and the server. Ensuring the security, privacy, and stability of this ecosystem is critical. The stakeholder team is responsible for overseeing these aspects, continuously monitoring for risks, and prioritizing updates. A common significant challenge in this ecosystem is that phone users are able to independently apply available updates and communicate their findings on the Reddit platform. Towards this objective, the following use case study scenario is designed.

1. Clinical application ecosystem stakeholder team overseeing the ecosystem security, privacy, and stability.
2. Clinical application includes features communicating digital content, such as photos, voice, videos, and text between the phone and the server.
3. Clinical application ecosystem team is accountable for ensuring that content adheres to clinical security, privacy, and stability.
4. Clinical application ecosystem components include Android, React, Java, Kotlin, iOS, Swift, Linux, Windows, NodeJS.
5. Clinical application ecosystem components must be continuously monitored for security, privacy, and stability risks.
6. Clinical application ecosystem stakeholders want to evaluate how well the monitoring and update prioritization can be automated.
7. Clinical application ecosystem stakeholders want to get an automated scheduling recommendation of updates specifically on updates that potentially impact legal regulations.

2) Research Question Design

To evaluate the effectiveness of our model and prototype, we will assess the results against a defined set of research questions.

RQ1: How many clinical application component release notes were enriched with Reddit data?

RQ2: How does the categorization of updates change before and after integrating Reddit data?

RQ3: How does user feedback vary across different operating systems and types of updates?

RQ4: How do different types of updates (e.g., major, minor, security) receive feedback on Reddit?

RQ5: What are the most common concerns raised by users on Reddit about new updates?

RQ6: How much did the inclusion of Reddit data improve the detection of critical issues?

IV. RESULT

A. Research Questions

RQ1: How many release notes were enriched with Reddit data? Each software component was enriched with at least 100 Reddit posts. The number of enriched release notes for each component is as follows: Android (1000), Linux (200), iOS (149), Kotlin (51), Windows (22), ReactJs (16), Swift (8), Java (8), NodeJs (3). The majority of the enrichment was performed by component, ignoring the actual version. For most Reddit posts, more than 50% had no matching semantic versioning information. Android and Java posts had more than 50% with versioning information. Due to the fragmented Android ecosystem, this is very valuable information for maintaining software updates. All operating systems (Android, iOS, Linux, Windows) had each 100 related posts between June and July of 2024, indicating both the need to continuously monitor high-impact OS components and the active Reddit user engagement in posting related content. In contrast, NodeJs posts were not as recent, indicating that developers use other platforms to communicate issues (see Tab. 2).

RQ2: How does the categorization of updates change before and after integrating Reddit data? After integrating Reddit data, a significant observation is the increased count in the categorization of components. For instance, the number of categories for Android increased from 6 to 12, for iOS from 5 to 6, and for Linux from 1 to 12. This indicates that Reddit posts now encompass a broader scope for each component and highlight potential technical dependencies on other non-core components. With respect to predicting security-related software updates, classifying posts led to the identification of security-related posts for each component. Unlike Common Vulnerabilities and Exposures (CVEs), these posts have not been formally reviewed yet, but they provide clinical stakeholders an opportunity for a potential early evaluation. The most notable change pertains to the identification of breaking updates. Breaking updates refer to issues where core functionalities such as Bluetooth, Internet, or account management fail after an update. Although these issues do not

Table 2: Component Data Overview

Component	Release Note Count	Reddit Post Count	Reddit with Version and Version Match	Reddit Post Upvote	Predicted Component Type (Release Note)	Predicted Component Type (Reddit Post)
nodejs	3 earliest: Aug 2022 latest: Feb 2023	101 earliest: Jun 2014 latest: Mar 2019	versions: 10 major: 4 unmatched: 6	75 (of 101)	WEB: 100%	WEB: 67% SERVER: 6% Security (True): 0 Breaking (True): 0
java	8 earliest: Jul 2011 latest: Jun 2024	101 earliest: Oct 2020 latest: Jul 2024	versions: 46 unmatched: 46	84 (of 101)	LANGUAGE: 100%	LANGUAGE: 83% LIBRARY: 12% DEVOPS: 4% API: 2% Security (True): 0 Breaking (True): 0
swift	8 earliest: Feb 2017 latest: Mar 2024	102 earliest: Jan 2021 latest: Jul 2024	versions: 12 unmatched: 12	78 (of 102)	MOBILE: 75% OS: 25%	MOBILE: 55% APP: 13% OS: 12% LANGUAGE: 5% API: 4% DEVOPS: 3% Security (True): 1 Breaking (True): 0
kotlin	51 earliest: Jul 2020 latest: Jul 2024	100 earliest: Jun 2024 latest: Jul 2024	versions: 7 major: 2 unmatched: 5	74 (of 100)	MOBILE: 100%	MOBILE: 69% DEVOPS: 5% DATABASE: 4% OS: 3% LANGUAGE: 3% WEB: 3% Security (True): 0 Breaking (True): 0
react	16 earliest: Mar 2020 latest: Apr 2024	101 earliest: Jan 2021 latest: Jul 2024	versions: 7 unmatched: 7	75 (of 101)	MOBILE: 6% LIBRARY: 88% WEB: 6%	LIBRARY: 68% WEB: 9% APP: 6% LANGUAGE: 3% DEVOPS: 2% Security (True): 0 Breaking (True): 0
linux	200 earliest: Aug 2020 latest: Jul 2024	102 earliest: Jun 2024 latest: Jul 2024	versions: 34 major: 2 unmatched: 32	78 (of 102)	OS: 99.4% REPOSITORY: 0.6%	OS: 80% BROWSER: 4% APP: 3% IDE: 3% Security (True): 4 Breaking (True): 2
windows	22 earliest: Mar 2024 latest: Jul 2024	101 earliest: Jul 2024 latest: Jul 2024	versions: 30 major: 1 unmatched: 29	71 (of 101)	CONTAINER: 27% IDE: 9% OS: 50% SERVER: 5% SECURITY: 5%	OS: 82% APP: 5% DEVOPS: 2% Security (True): 1 Breaking (True): 0
android	1000 earliest: Sep 2008 latest: Jun 2024	101 earliest: Jun 2024 latest: Jul 2024	versions: 58 unmatched: 58	78 (of 101)	REPOSITORY: 75% OS: 4% CLOUD: 0.4% OTHER: 20.4% MOBILE: 0.1% DATA: 0.1%	OS: 80% DEVOPS: 4% APP: 4% Security (True): 2 Breaking (True): 0
ios	149 earliest: Feb 2010 latest: Jul 2024	102 earliest: Jul 2024 latest: Jul 2024	versions: 25 major: 4 unmatched: 21	75 (of 102)	OS: 99.3% DATA: 0.7%	OS: 83% APP: 9% BROWSER: 3% DEVOPS: 2% Security (True): 1 Breaking (True): 0

affect all devices and operating system versions uniformly, they are critical because they can block the operation of clinical applications. For example, iOS users reported problems such as "iOS 17.5.1 Freezing and Apps not working," "Bluetooth Issues since iOS update," and "Help! My phone doesn't turn all the way on after the iOS 17.5.1 software update."

Integrating Reddit user feedback is crucial for the clinical application ecosystem. It improves the identification and prioritization of security and breaking updates, which is essential for maintaining system stability and ensuring compliance with security and privacy standards.

RQ3: How does user feedback vary across different operating systems and types of updates? In general, there was no major difference between software component types. However, recent posts indicate that phone operating systems are most impacted by breaking updates. For example, 13 (component average was 5) out of 105 iOS posts were related to breaking updates. Consequently, clinical applications that have a mobile component should prioritize these types of updates. Additionally, iOS-related posts had the highest number of matching versioning-related information, allowing for a better understanding of which component versions are affected.

All Reddit posts are typically concise, with an average of ten words in the title summarizing the main issue. This brevity facilitates prompt scanning and comprehension of the issues at hand. Integrating this Reddit feedback into clinical ecosystems is crucial for promptly identifying and addressing critical updates that may affect system stability and compliance with healthcare regulations.

For instance, if a clinical application relies on mobile devices for data entry or patient monitoring, a breaking update that affects core functionalities like Bluetooth or Internet connectivity can have severe implications. In this context, the high occurrence of breaking updates in iOS posts underscores the need for vigilant monitoring and swift action to ensure uninterrupted service and patient safety. By prioritizing these updates, clinical stakeholders can proactively mitigate risks and maintain the integrity of their systems.

RQ4: How do different types of updates (e.g., major, minor, security) receive feedback on Reddit? Android and Java Reddit users post most versioning details, with iOS users frequently posting detailed patch versions. Over 50% of Android-related Reddit posts include some form of version data. In contrast, more technical components such as Kotlin or React were less likely to include version details. This suggests that discussions about these technical components are either more generalized or that detailed, version-specific queries are directed to more specialized technical forums.

RQ5: What are the most common concerns raised by users on Reddit about new updates? Frequent technical issues posts raised are about being unsure why some phone-related functionality does not work correctly (WiFi, Bluetooth, Configurations). Generally, Reddit users report issues when a phone functionality that affects all applications fails to work. Typically, these questions are not posted on app store reviews, so Reddit is a good source for clinical stakeholders to understand such critical bugs. Moreover, usability, performance, and compatibility-related concerns are frequently posted. Security-related concerns are mostly posted if they impact the user.

RQ6: How much did the inclusion of Reddit data improve the detection of critical issues? The integration of Reddit data led to a significant increase in the detection of critical issues, with a total of 58 identified in the clinical use case scenario. Some of these posts highlight potential future legal impacts. Early identification of these risks substantially improves the ability to mitigate them and allocate resources effectively for maintenance efforts. These types of updates are typically absent from regular release notes and may be delayed in CVEs if included at all. Consequently, critical updates identified through Reddit posts can potentially improve the user experience and usability of clinical applications.

B. Prototype

To illustrate the effectiveness of the proposed model, we implemented a prototype web application available at <https://releasetrain.io/reddit?q=ios,android,linux,windows,kotlin,swift,java,nodejs,react>. The application prioritizes recommendations by sorting security and breaking updates first, as these are typically high-priority reviews in clinical settings due to legal compliance (see Fig. 3). For each component type, it lists the number of Reddit posts, comments, and upvotes, and compares impact classifiers between Reddit posts and release notes. Since the web application follows an extract and load process initially, the transform step remains performant.

V. DISCUSSION

Reddit provides a broader and more diverse range of user feedback compared to traditional app store reviews, which is particularly valuable for clinical applications. While app stores are limited to feedback on specific applications, Reddit allows for discussions on various software components and systems, capturing user experiences and issues across different platforms. This broader scope is crucial in identifying and addressing non-core dependent components that could affect clinical ecosystem software.

Breaking updates at the phone level can disrupt multiple applications, including critical clinical software. Issues such as Bluetooth malfunctions, internet connectivity problems, and account-related errors are common on Reddit. Monitoring these updates helps identify these issues early, allowing for swift mitigation and ensuring the stability of clinical applications.

Recent Most Critical Update:

Component: LINUX | Date: Jul 3, 2024 (17 days ago) | Title: [Linux's DRM Panic "Screen of Death" Sees Patches For QR Code Error Messages](#) | Type: Security

Fig. 3. Automatic Reddit Enrichment and Update Prioritization

Reddit users often provide more specific and explicit content when describing issues. Instead of vague feedback, users detail the exact nature of problems, making it easier to understand and address the root cause. This detailed feedback is crucial for identifying and prioritizing updates that may not be thoroughly documented in release notes or CVEs, which is essential for maintaining the reliability of clinical software.

The language, intention, and content of Reddit users offer a comprehensive view of software issues. The platform supports both technical discussions and casual user experiences, providing a good insight into user experiences. This diversity helps highlight problems that might not be evident through formal channels, ensuring a more thorough understanding of the issues affecting clinical applications.

The prototype application presented enables rapid review and prioritization of critical updates. By prioritizing real time updates based on security and breaking changes, the prototype facilitates quick identification of high-impact issues. This capability is particularly valuable in clinical settings, where timely updates are crucial for maintaining compliance with security and privacy standards. The prototype's ability to prioritize critical updates helps in efficiently managing resources and mitigating risks.

The clinical use case shows that while not all posts on Reddit are critical, the 5% highlighting security and breaking updates are essential for mitigating risks. These posts provide early warnings about potential issues, allowing for proactive measures to address them. By enriching the feedback with these critical updates, the overall user experience and reliability of clinical applications can be significantly improved.

In summary, leveraging Reddit for user feedback in the clinical use case offers a unique and valuable perspective on software issues. The ability to quickly review and prioritize critical updates ensures that significant problems are addressed promptly, ultimately enriching software quality and user satisfaction.

VI. CONCLUSION

In this study, we demonstrated the significant benefits of integrating Reddit user feedback into software release notes across all phases of the software engineering lifecycle. Our automated process for enriching release notes with user feedback identified 5% more critical and breaking updates out of 900 Reddit posts. This improvement is crucial for

maintaining clinical applications, as early detection of these updates helps mitigate risks and improve system stability. We recommend adopting automated user feedback integration in the software maintenance process to improve software quality and user satisfaction. Future work will explore applying this methodology in real application domains and further refining the automation process to handle a broader range of feedback sources.

VII. ACKNOWLEDGEMENT

We would like to thank Reddit for user feedback data, as well as GitHub, MITRE, and Wikipedia for release note references. Their support and resources were crucial to the successful completion of this study.

REFERENCES

- Berhe, S., Kan, V., Khan, O., Pader, N., Farooqui, A. Z., Maynard, M., & Khomh, F. (in press). Attribute-driven maintenance cost estimation in graph-based software ecosystem models. *International Journal of Ubiquitous Systems and Pervasive Networks (JUSPN 2024)*.
- u/ios-user. (2023, July 15). Regarding the iOS 17.5 photo glitch. Reddit. https://www.reddit.com/r/ios/comments/1cwgljj/regarding_the_ios_175_photo_glitch/
- Horawalavithana, S., Bhattacharjee, A., Liu, R., Choudhury, N., Hall, L. O., & Iamnitchi, A. (2019). Mentions of security vulnerabilities on Reddit, Twitter and GitHub. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI '19)* (pp. 200–207). Association for Computing Machinery. <https://doi.org/10.1145/3350546.3352519>
- Shrestha, P., Sathanur, A., Maharjan, S., Saldanha, E., Arendt, D., & Volkova, S. (2020). Multiple social platforms reveal actionable signals for software vulnerability awareness: A study of GitHub, Twitter and Reddit. *PLOS ONE*, 15(3). <https://doi.org/10.1371/journal.pone.0230250>
- Aggarwal, A., Gola, B., & Sankla, T. (2021). Data mining and analysis of Reddit user data. In V. K. Gunjan, P. N. Suganthan, J. Haase, & A. Kumar (Eds.), *Cybernetics, Cognition and Machine Learning Applications* (pp. 1–14). Springer. https://doi.org/10.1007/978-981-33-6691-6_24
- Achimescu, V., & Chachev, P. D. (2021). Raising the flag: Monitoring user perceived disinformation on Reddit. *Information*, 12(1), 4. <https://doi.org/10.3390/info12010004>
- Iqbal, T., Khan, M., Taveter, K., & Seyff, N. (2021). Mining Reddit as a new source for software requirements. In *Proceedings of the 2021 IEEE 29th International Requirements Engineering Conference (RE)* (pp. 128–138). IEEE. <https://doi.org/10.1109/RE51729.2021.00019>
- Li, T., Louie, E., Dabbish, L., & Hong, J. (2021). How developers talk about personal data and what it means for user privacy: A case study of a developer forum on Reddit. *Proceedings of the ACM on Human-Computer Interaction*, 4, 1–28. <https://doi.org/10.1145/3432919>
- Horawalavithana, S., Choudhury, N., Skvoretz, J., & Iamnitchi, A. (2022). Online discussion threads as conversation pools: Predicting the growth of discussion threads on Reddit. *Computational and Mathematical Organization Theory*, 28(2), 112–140. <https://doi.org/10.1007/s10588-021-09340-1>
- Srewuttanapitikul, K., & Muengchaisri, P. (2016). Prioritizing software maintenance plan by analyzing user feedback. In *2016 International Conference on Information Science and Security (ICISS)* (pp. 1–5). IEEE. <https://doi.org/10.1109/ICISSEC.2016.78858>