

An Empirical Study of Obsolete Answers on Stack Overflow

H. Zhang, S. Wang, T. -H. Chen,
Y. Zou, A. E. Hassan

in IEEE Transactions on Software Engineering,
vol. 47, no. 4, pp. 850-862, 1 April 2021

発表者: 井町 瑠貴

背景

現在、ソフトウェアシステムは急速なペースで進化しており、Stack Overflow(SO)の回答にある知識を時間とともに陳腐化させる可能性がある

回答が陳腐化していることに気づかず参考にし、ソフトウェアの品質に問題が生じる可能性がある

目的

SOのアンサーズレッド52,177件を調査し、アンサーズレッドに埋め込まれた知識がどのように陳腐化していくのか、陳腐化した回答の特徴を理解し、実行可能な提案を行う

また、ユーザーが陳腐化したと判断した根拠や、その後ユーザーが行う行動を理解するために、定性的な調査を行った

リサーチクエスト

- RQ1: 回答が陳腐化したと観測された場合、
何が起こるか?
- RQ2: 特定のタグが付いた質問に対する回答は、
他の回答よりも陳腐化しやすいか?
- RQ3: 回答が陳腐化する潜在的な理由は何か?
- RQ4: 誰が陳腐化した回答を観察し、
どのような証拠を提供するのか?

データ収集

Stack Exchangeコミュニティが2017年8月31日に
公開したデータを利用

バッジ, コメント, 投稿履歴, 投稿リンク, 投稿, タグ, ユーザー, 投票に関する情報が含まれている

データ収集

以下の2つの基準の両方を用いて、陳腐化したアンサーズレッドを特定

1. アンサーズレッドのコメントに, "deprecated", "outdated", "obsolete", "out of date" のいずれかのキーワードが含まれている
2. 基準1と同じキーワードが質問や, その回答のいずれにも現れていない

データ収集

調査の結果, 52,177のアンサーズレッドを発見し, その中には陳腐化について言及している58,201のコメントが含まれており, また, 12,629個のタグが含まれていた

RQ1:

回答が陳腐化したと観測された場合、何が起こるか？

観察の結果、回答の陳腐化が観測された後、行われる可能性のある行動は2種類である

1. 陳腐化した回答を更新する(update)
2. 新しい回答を作成する(new)

RQ1:

回答が陳腐化したと観測された場合、何が
起こるか?

定量的分析(陳腐化が観測されるタイミング)

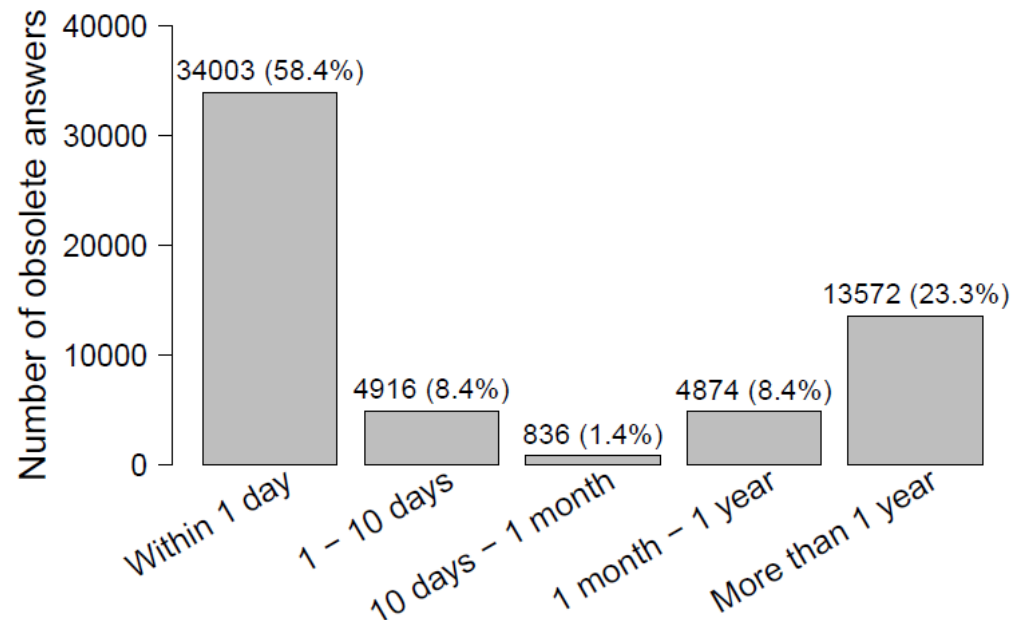


Fig. 4: Number of obsolete answers vs number of days to point out the obsolescence of the answers.

RQ1:

回答が陳腐化したと観測された場合、何が起こるか？

定量的分析(陳腐化が観測されるタイミング)

- ・回答の58.4%が、回答作成後24時間以内に陳腐化の指摘を受けている
→投稿された時点で陳腐化していた可能性がある
- ・承認された回答のうち、陳腐化したと指摘された後に更新されたのは30.7%、未承認の回答で更新されたのは24.8%であった
→承認された回答の方が更新されやすい

RQ1:

回答が陳腐化したと観測された場合、何が起こるか？

定量的分析(陳腐化が観測されるタイミング)

- ・今後の研究では、承認されていない回答にも注目する必要がある
- ・また、コメントで陳腐化が確認された後、陳腐化した回答を更新するまでに平均227日、新しい回答を追加するまでに平均198日かかることがわかった

RQ1:

回答が陳腐化したと観測された場合、何が起こるか？

定性的分析(updateとnewの出現率)

- ・陳腐化した回答を更新したケースが20.5%, 新しい回答を追加したケースが6.3%であった
- ・平均して, 回答の陳腐化が観測されてからユーザーが反応するまでには118日かかることが分かった

RQ2:

特定のタグが付いた質問に対する回答は、他の回答よりも陳腐化しやすいか？

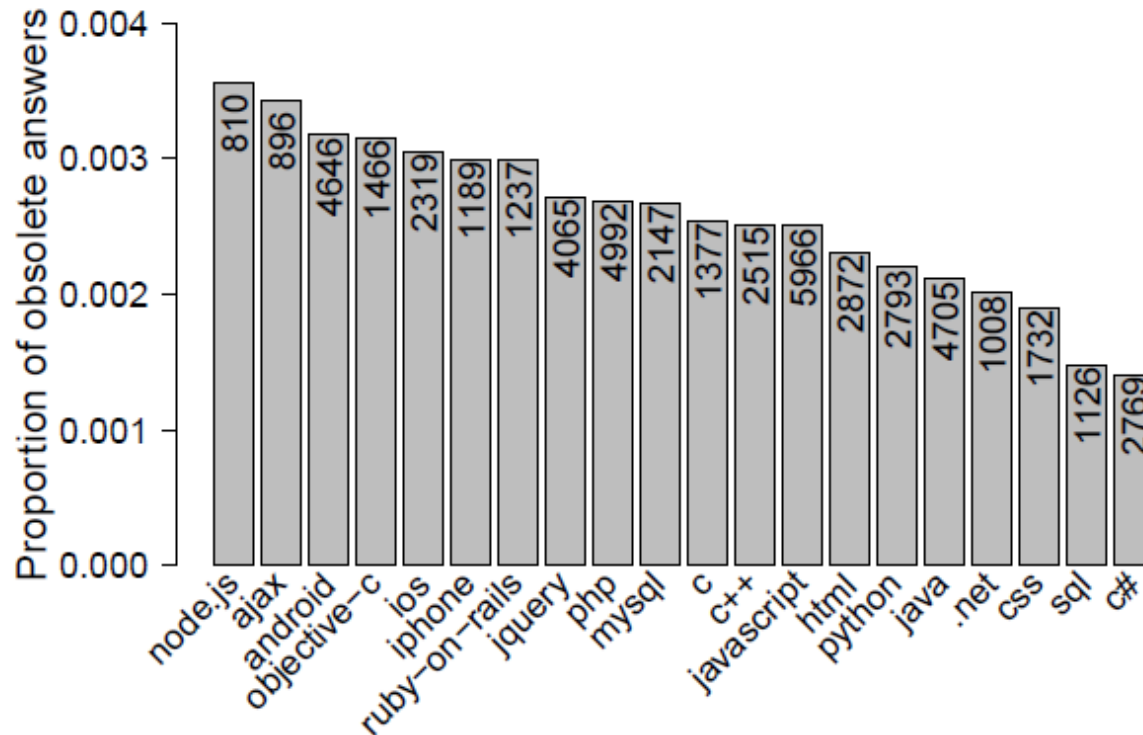


Fig. 7: The top 20 tags ranked by the ratio of obsolete answers to the total number of answers in each tag.

RQ2:

特定のタグが付いた質問に対する回答は、他の回答よりも陳腐化しやすいか？

- ・最も陳腐化しやすいタグは node.js で、回答の 0.36% が陳腐化していると指摘されている
- ・モバイルアプリの人気により、SOではモバイルアプリ関連の質問と回答が増加している
- ・モバイルアプリ技術に関連する回答は、技術の進歩が速いため、陳腐化する可能性が高い

RQ3: 回答が陳腐化する潜在的な理由は何か?

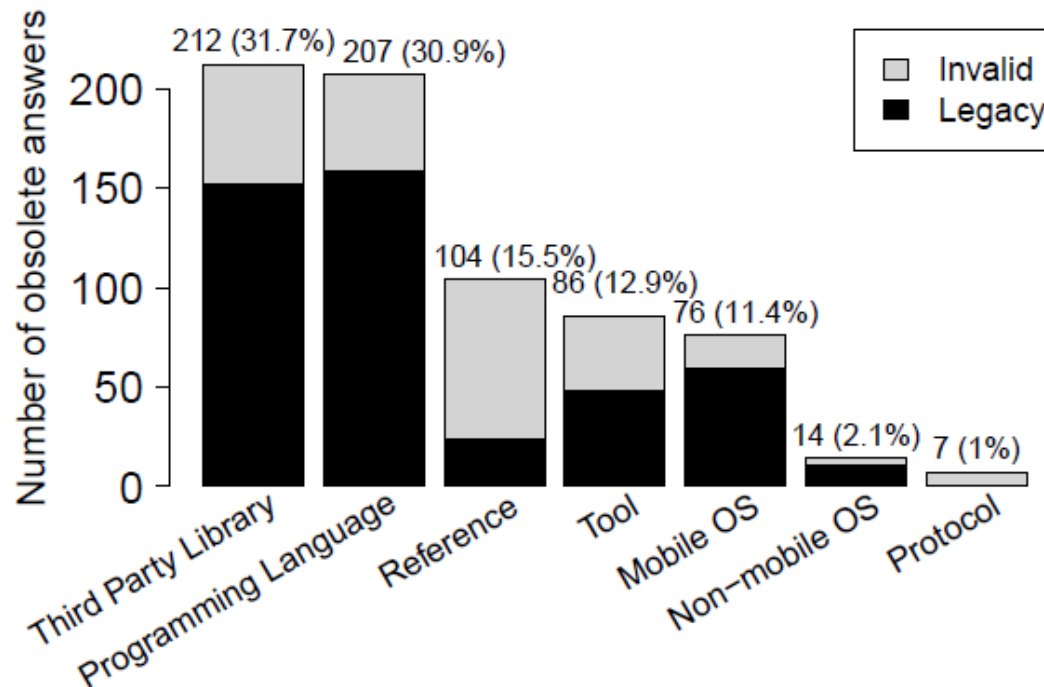


Fig. 8: Number and percentage of each obsolescence reason based on our manual analysis. The figure also shows the proportion of legacy (black) and invalid (gray) obsolescence.

RQ3:

回答が陳腐化する潜在的な理由は何か？

- ・調査対象となった回答の31.7%が、関連するサードパーティ製ライブラリの進化によって陳腐化していた
- ・30.9%の回答がプログラミング言語の進化によって陳腐化していることがわかった
- ・SOでは、投稿時に有効なバージョンを記録することができるので、回答にバージョン番号を提供することを推奨する

RQ4:

誰が陳腐化した回答を観察し、どのような証拠を提供するのか？

観察者を以下の5つに分類する

1. 質問者: 質問を投稿したユーザー
2. 回答者: 陳腐化した回答を投稿したユーザー
3. 他の回答者: 未承認の回答を投稿したユーザー
4. コメンター: 質問スレッドにコメントを投稿したユーザー
5. アウトサイダー: スレッドで一度も活動をしたことがないユーザー

RQ4:

誰が陳腐化した回答を観察し、どのような証拠を提供するのか？

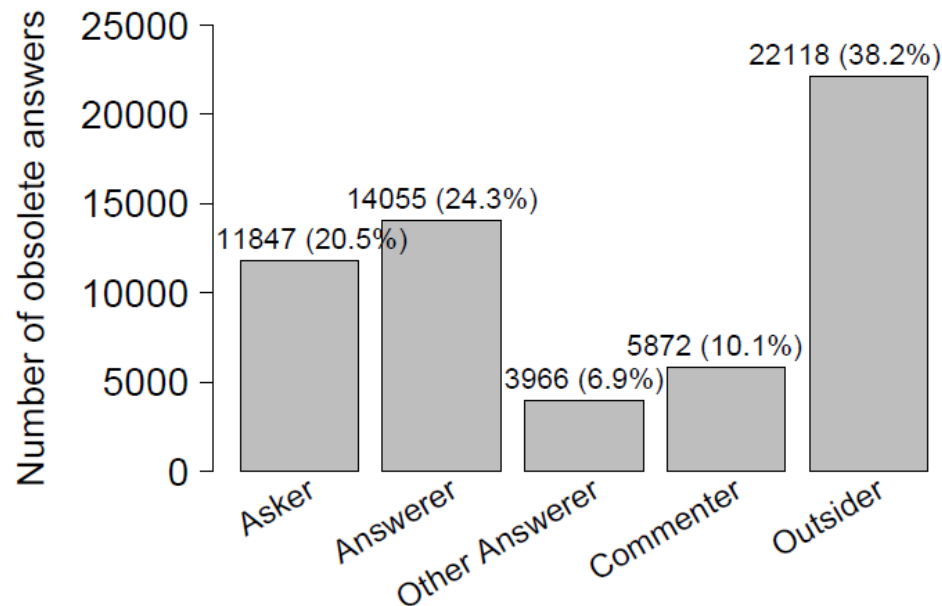


Fig. 9: The number (as well as the percentage) of the obsolete answers that are observed by each type of user. A role of user is assigned using the following priority: asker > answerer > other answerer > commenter > other user. For example, if a user has multiple roles, such as an answerer and a commenter, we consider the user as an answerer.

RQ4:

誰が陳腐化した回答を観察し、どのような証拠を提供するのか？

- ・回答が陳腐化していることを最も観察していたのはアウトサイダーであった

- ・回答者の方がすくない理由

- 1.SOで活動しなくなった

- 2.長期間にわたって自分の回答を最新の状態で維持することをあまり望んでいない

RQ4:

誰が陳腐化した回答を観察し、どのような証拠を提供するのか？

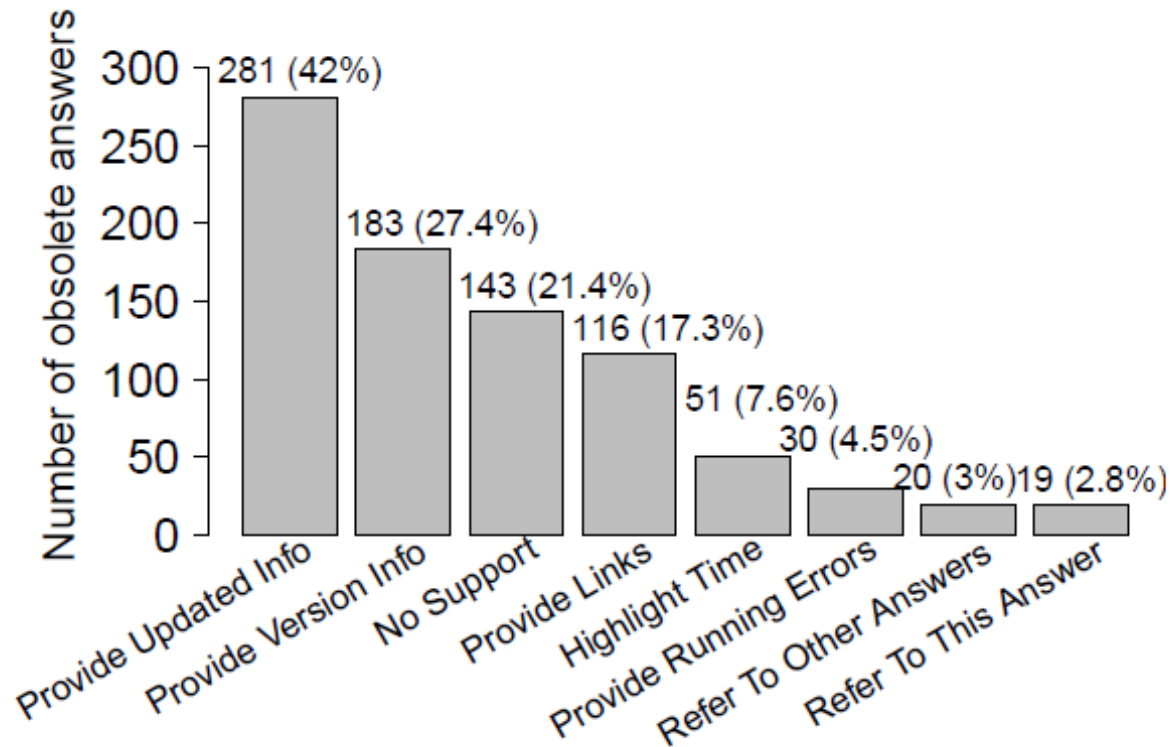


Fig. 10: The proportion of each type of evidence that users provide when pointing out obsolescence.

RQ4:

誰が陳腐化した回答を観察し、どのような証拠を提供するのか？

- ・78.6%の陳腐化の観察は、証拠によって裏付けられていた
- ・1つの陳腐化の観測に対して、複数の証拠を提示している場合が観察された

SOに対する実用的な提案

- ・ユーザーは自分の回答の陳腐化をあまり気にしておらず、頻繁に回答をメンテナンスしていない

ユーザー(特に質問スレッドの内部の人)が回答の陳腐化にもっと注意を払い、陳腐化した回答を維持する努力を促すメカニズムを開発すべき

ユーザーへの実行可能な提案

- ・陳腐化が確認された回答の44.8%では、コメントが回答を更新するための解決策を提供している
- ・回答の陳腐化を示すコメントの73.5%は、陳腐化した回答の上位1位のコメントであることが確認されている

そのため、回答が陳腐化した場合に備えて、回答に関連付けられたコメントに注意深く目を通すことが推奨される

結論

今回の研究から以下のようなことが分かった

1. 特定のタグ(node.js, ajax, android, objective-c)の回答は, サードパーティのライブラリやプログラミング言語の進化によって陳腐化する可能性が高い
2. 陳腐化した回答のほとんどは, 回答者以外の人によって指摘され, 証拠によって裏付けられている。
3. 陳腐化した回答が特定された場合, その後更新されるのはごく一部である

結論

今回の研究から以下のようなことを提案する

1. SOは、コミュニティ全体が陳腐化した回答を特定したり維持したりすることを奨励する仕組み(インセンティブシステム)を開発すべきである
2. 回答者は回答を作成する際に、有効なバージョンや時間の情報を含めることが推奨される
3. ユーザーは、回答の陳腐化が指摘された場合に備えて、回答スレッドの情報に注意深く目を通すことが推奨される

所感

- ・半数以上の回答が回答した時点で陳腐化していたのは興味深かった
- ・回答者などのメンテナンスがなくても、回答の陳腐化についての情報がユーザーに分かりやすくなると便利になるのでは、と考えた