

Understanding the Role of Images on Stack Overflow

Dong Wang[†], Tao Xiao[‡], Christoph Treude[§], Raula Gaikovina Kula[‡], Hideaki Hata[¶], Yasutaka Kamei[†]
Kyushu University, Japan[†]
Nara Institute of Science and Technology, Japan[‡]
University of Melbourne, Australia[§]
Shinshu University, Japan[¶]

Proceedings of 20th International Conference on Mining Software Repositories (MSR '23)
pp.??-??

発表者：井町 瑠貴

背景 | SOの質問の質

開発者は、ウェブ上の情報を検索するためにStack Overflow (SO) を利用することが多い

- 効果的に回答を得ることは開発者にとって不可欠

しかし、いくつかの研究で、質の高い質問を作成し、効果的に回答を得ることは容易ではないことを指摘

- SO上のかなりの質問が質が極めて悪い^[5]
- 47%の質問にaccept answerがない^[6]

[5] D. Correa and A. Sureka, “Chaff from the wheat: Characterization and modeling of deleted questions on stack overflow,” in Proceedings of the 23rd international conference on World wide web, 2014, pp. 631–642.

[6] M. Yazdaninia, D. Lo, and A. Sami, “Characterization and prediction of questions without accepted answers on stack overflow,” in Proceedings of the 29th International Conference on Program Comprehension, 2021, pp. 59–70.

既存研究 | SOにおける質の高い質問

読みやすくわかりやすい情報を提供することが質の最も強い指標とみなされる^[7]

- 質の高い質問には質の高いコードスニペットが必要

通常のテキスト（タイトル，説明文など）とは別に，コードスニペットの役割を調査するための研究が多く行われている^[8,9]

- コードスニペットを含む質問は回答比率が高く，1つ以上の良い回答がある可能性が観察されている^[10]

[7] M. Duijn, A. Kucera, and A. Bacchelli, “Quality questions need quality code: Classifying code fragments on stack overflow,” in Proceedings of the 12th Working Conference on Mining Software Repositories, 2015, pp. 410–413.

[8] F. Calefato, F. Lanubile, M. C. Marasciulo, and N. Novielli, “Mining successful answers in stack overflow,” in Proceedings of the 12th Working Conference on Mining Software Repositories, 2015, pp. 430–433.

[9] Z. Gao, X. Xia, J. Grundy, D. Lo, and Y.-F. Li, “Generating question titles for stack overflow from mined code snippets,” ACM Transactions on Software Engineering and Methodology, vol. 29, no. 4, pp. 1–37, 2020.

[10] M. Asaduzzaman, A. S. Mashiyat, C. K. Roy, and K. A. Schneider, “Answering questions about unanswered questions of stack overflow,” in Proceedings of the 10th Working Conference on Mining Software Repositories, 2013, pp. 97–100.

先行研究 | SO上の画像に関する調査[1]

開発者がソーシャルコーディング環境（SOなど）で画像を共有することが増えている

- SOでは過去5年間で、画像を含む投稿の割合が直線的に増加している傾向にある

しかし、画像が持っている特徴や、画像の役割に関する調査は行われていない

- 本研究は、これらを明らかにすることを目的としている
 - 画像の特徴、どのように使用されているか、画像の役割

[1] M. Duijn, A. Kucera, and A. Bacchelli, “Quality questions need quality code: Classifying code fragments on stack overflow,” in Proceedings of the 12th Working Conference on Mining Software Repositories, 2015, pp. 410–413.

RQ

RQ1：使われる画像の特徴

- － 動機：画像がどのような特徴（ソース、コンテンツ、目的など）を持っているのか明らかにする

RQ2：画像はどのように使われているのか

- － 動機：質問のタイプによって画像がどの程度使われているのか明らかにする

RQ3：画像を使った質問はどのように回答されるか

- － 動機：画像の存在が回答に与える影響を明らかにする

データセット | データ収集

SOTorrent (2019/12 時点) を利用

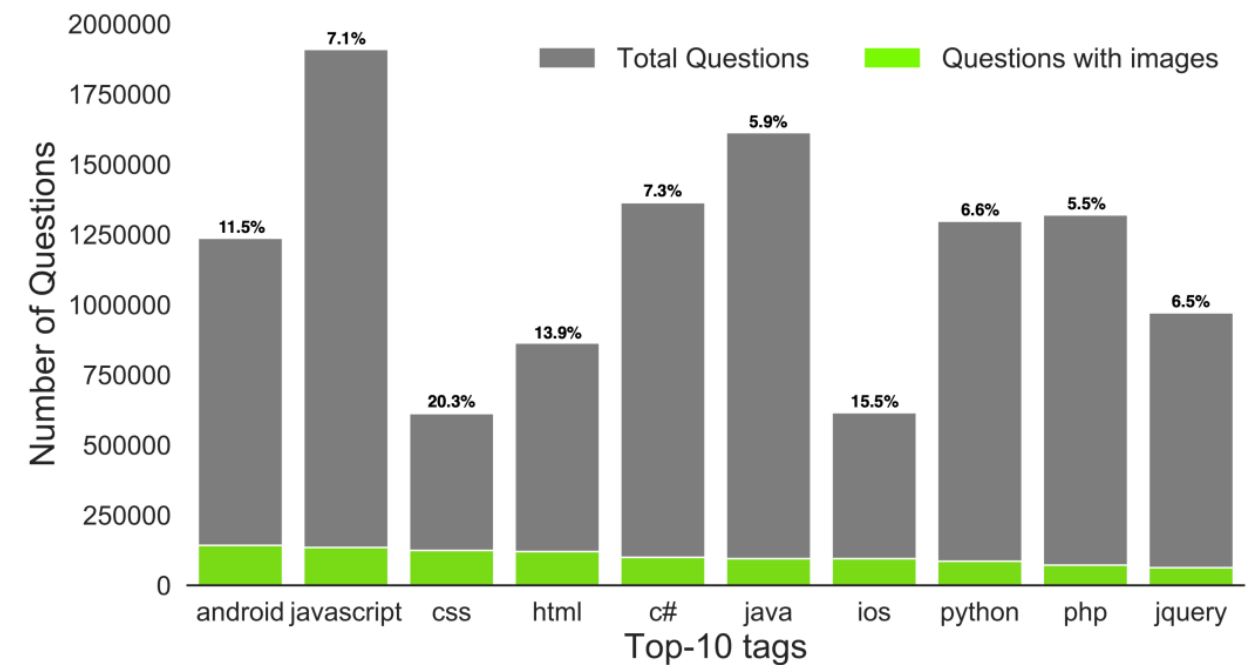
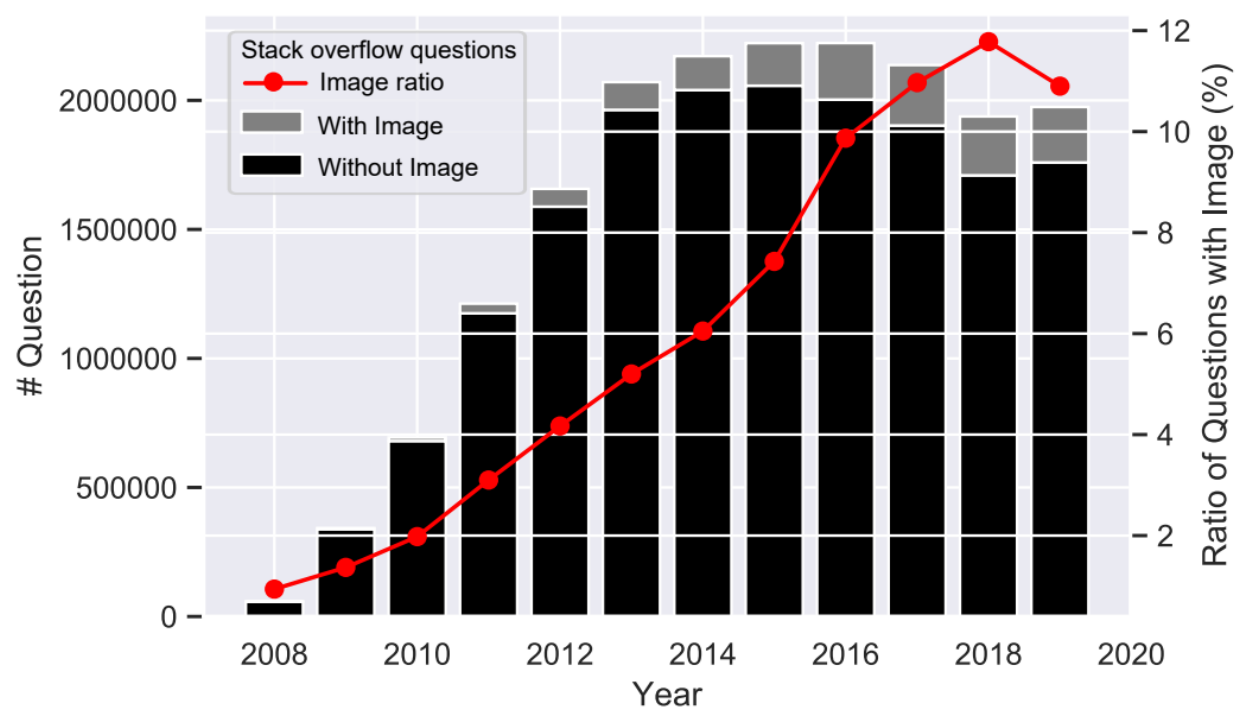
- 画像を含む質問は正規表現を用いて抽出

Studied Period	2008.06 – 2019.12
# Questions	18,699,426
# Questions Contain Images	1,426,890 (7%)
# Total Images	2,026,090
# Images per Posts (1st Qu./Median/3rd Qu.)	1/1/2

データセット | 概要

以下の2分析を行なった

- 画像を含む質問の傾向（左図）
- 画像を含む質問の人気のあるタグ（右図）



RQ1 | アプローチ

定性的な分析を実施

- 以下の3つの次元を採用
 - 画像ソース（出典），画像コンテンツ（内容），目的
- さらに画像の重要性を手動で分類

代表サンプルの作成

- 信頼水準95%、信頼区間5%に基づき，無作為にサンプリング
- 384の回答をサンプリング

RQ1 | アプローチ

手動コーディング（3人の著者で実施）

– 第1ラウンド

- 30件を独立してコーディング
- その後ガイドについて議論し，10件でガイドをテスト・改良
- コードの統合・見逃したコードの追加

– 第2ラウンド

- 別の30件を独立してコーディングし，カッパ係数を計算

– 第3ラウンド

- 別の20件を用いて，コーディング基準を再度検証

RQ1 | 結果

以下の表のような結果が得られた

	Category	Description/Example	Frequency
<i>Source</i>	Desktop application	shorturl.at/jnoHS	180 (47%)
	Web browser	shorturl.at/iqrGM	99 (26%)
	User-created	shorturl.at/doBGR	42 (11%)
	Mobile application	shorturl.at/iqrGM	40 (10%)
<i>Content</i>	User interface	such as buttons or an input screen, e.g., shorturl.at/rsDPS	224 (58%)
	Diagram	such as designs or charts, e.g., shorturl.at/ksvIQ	43 (11%)
	Result	such as output from a program, e.g., shorturl.at/fiwMN	38 (10%)
	Error code	such as a warning or error log, e.g., shorturl.at/auzCU	33 (9%)
	Source code	such as code snippets, e.g., shorturl.at/gjzJT	19 (5%)
	Configuration	such as build or dependency files, e.g., shorturl.at/kFGSU	6 (2%)
<i>Purpose</i>	Undesired output	an image presents an output that a developer would like to avoid, e.g., shorturl.at/arFX1	174 (45%)
	Additional context	an image is used to further support understanding of the developer question, e.g., shorturl.at/akyU	141 (37%)
	Desired output	an image show an output that a developer would like to achieve, e.g., shorturl.at/cgmsF	48 (13%)
<i>Essentialness</i>	Essential	the question is not understandable without images.	263 (68%)
	Not essential	the question is understandable without images	101 (26%)

RQ2 | アプローチ

どのようなタイプの質問が画像を含む可能性が高いかを理解するために質的分析を実施

- 既存研究^[11]に則り, 質問タイプを9つに分類

手動コーディング (2人の著者で実施)

- 2ラウンド行い, カッパ係数は0.90であった

対照群

- 比較のために, 画像なしの質問384件を無作為に選択し, 同様にコーディングを実施

[11] C. Treude, O. Barzilay, and M.-A. Storey, “How do programmers ask and answer questions on the web? (nier track),” in Proceedings of the 33rd International Conference on Software Engineering, 2011, p. 804–807.

RQ2 | アプローチ

さらに以下の2つの質問の特徴について分析

- 説明の長さ（単語数）
- 質問者の経験（過去に投稿した画像付きの質問数）

以下の仮説を立てて検証

- 画像がある質問には説明が少ない
 - 既存研究^[19]：Issueは画像があると説明が少なくなる
- 画像がある質問は、経験豊富なユーザによって投稿される可能性が高い
 - 画像から情報をよりよく要約できるため頻繁に用いられるのではないか

[19] H. Kuramoto, M. Kondo, Y. Kashiwa, Y. Ishimoto, K. Shindo, Y. Kamei, and N. Ubayashi, “Do visual issue reports help developers fix bugs?” in Proceedings of the 30th International Conference on Program Comprehension, 2022, pp. 511–515.

RQ2 | 結果

質問のタイプ

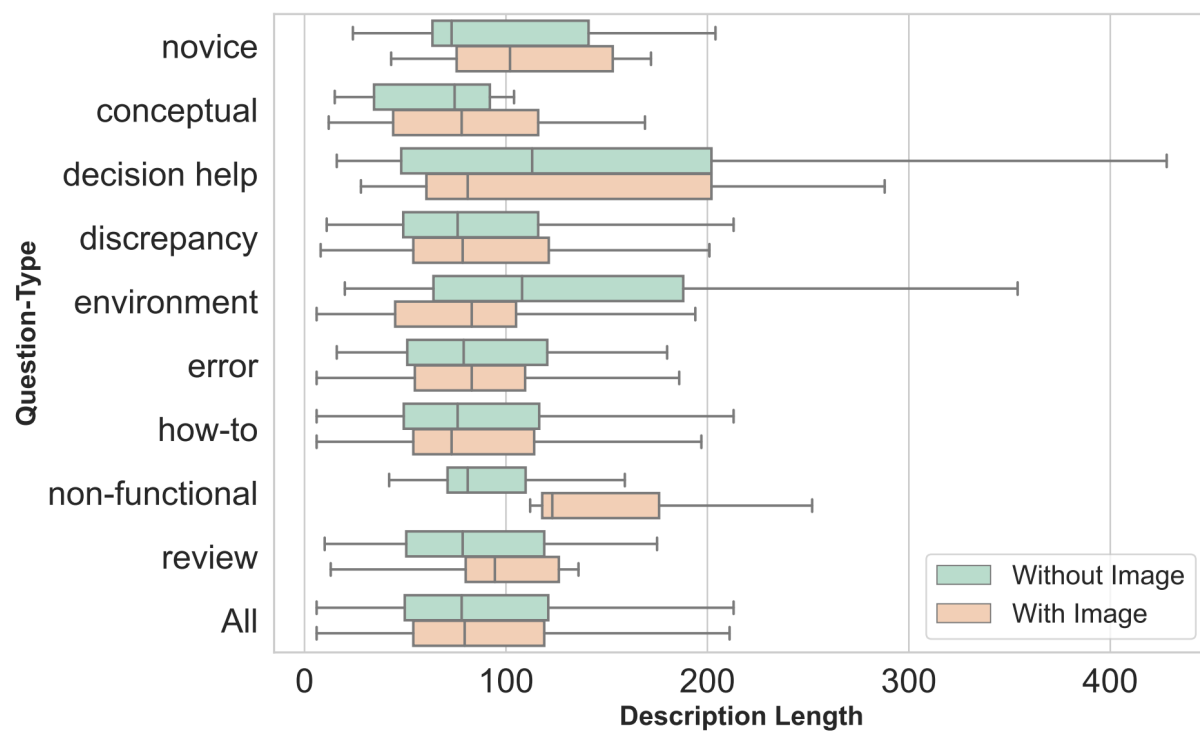
- how-to, discrepancyの頻度が高い
- ユーザが何らかの予期せぬ状態に対して説明を求めて画像をアップロードする可能性が高いことを示唆

		Category	With image	Without image
<i>Question Types</i>		how-to	185 (46%)	162 (42%)
		discrepancy	93 (24%)	69 (18%)
		error	41 (10%)	55 (14%)
		environment	13 (3%)	16 (4%)
		conceptual	13 (3%)	10 (3%)
		novice	13 (3%)	11 (3%)
		decision help	7 (2%)	25 (7%)
		non-functional	8 (2%)	10 (3%)
		review	7 (1%)	14 (4%)

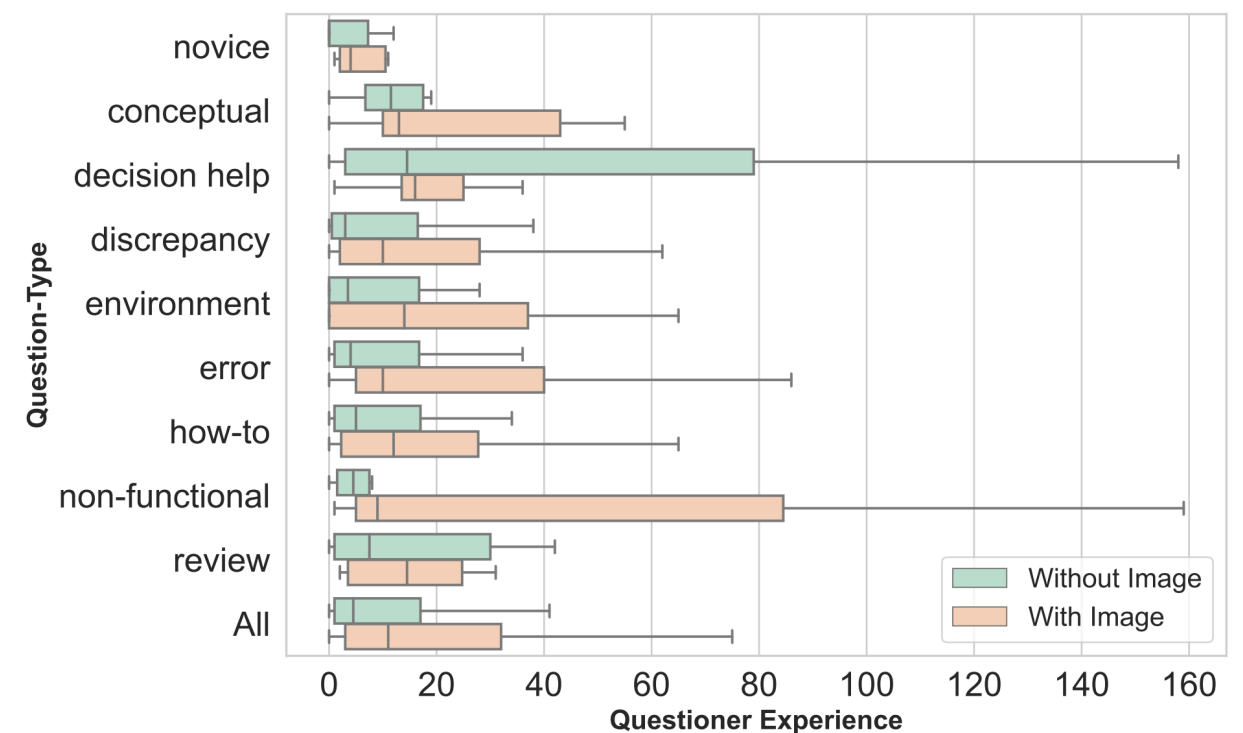
RQ2 | 結果

質問の特徴

- 説明の長さには大きな違いは見られない
- 画像は経験豊富なユーザによって使用される傾向がある



説明の長さ



質問者の経験

RQ3 | アプローチ

以下の3つのメトリクスから，画像のある質問と画像のない質問の違いを分析する

- 回答の有無
 - 画像が回答率を向上させるのに役立つと考えられる
- accept answerの有無
 - 画像がaccept answerを得る助けになると考えられる
- 回答時間
 - 画像は補足的な情報として機能するため，より早く回答を得るのに役立つと考えられる

RQ3 | 結果

accept answerの有無と回答の有無

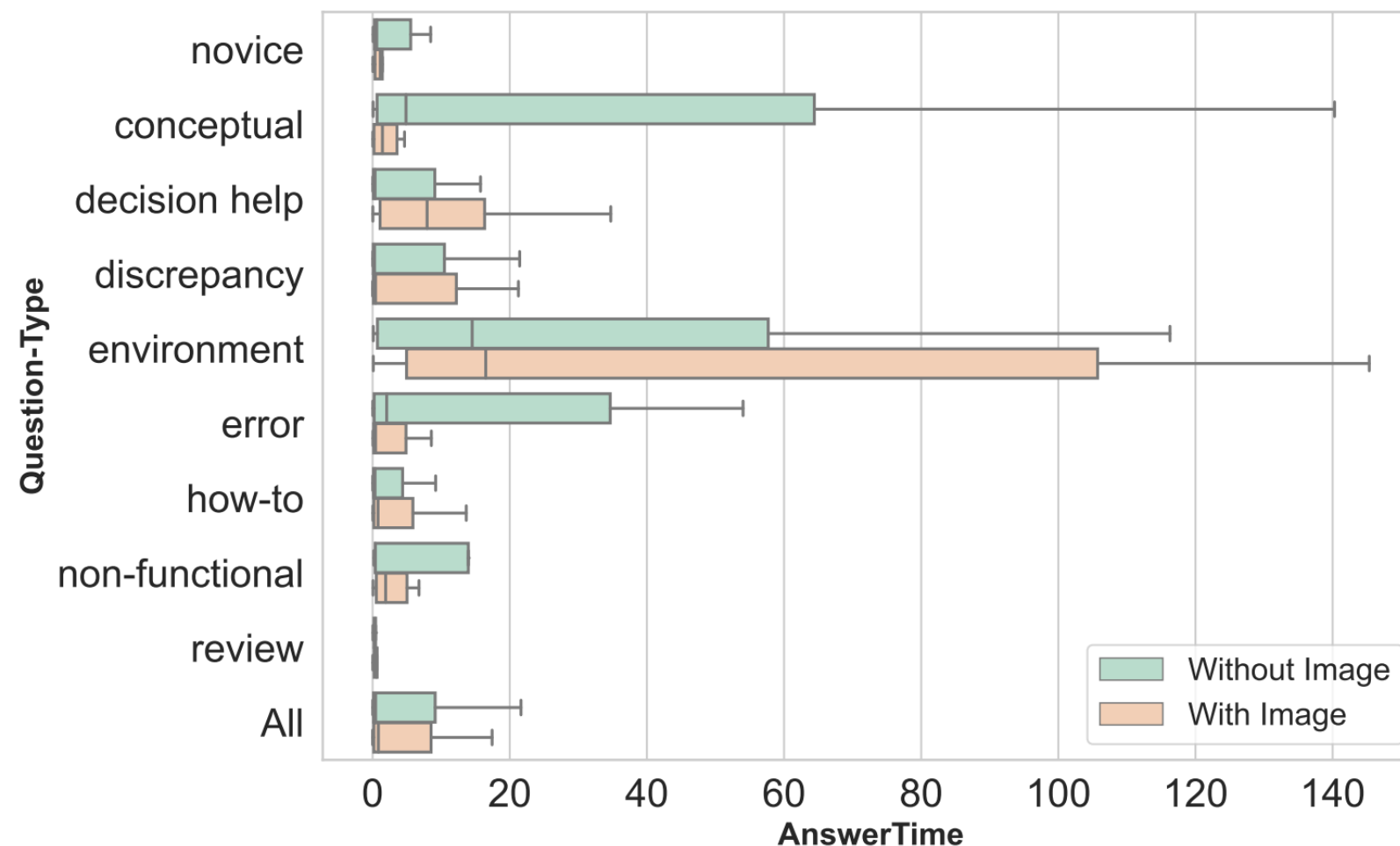
- 画像は回答を得る確率を高めることには役に立っていないと考えられる
- 画像を含む質問はaccept answerを得る可能性が高い

	Accepted Answer		Has Answers	
	With image	Without image	With images	Without image
how-to	110 (60%)	78 (48%)	158 (88%)	147 (91%)
discrepancy	45 (48%)	32 (46%)	75 (82%)	57 (82%)
error	21 (51%)	21 (38%)	30 (77%)	45 (82%)
decision help	6 (86%)	14 (56%)	6 (86%)	23 (92%)
novice	5 (45%)	5 (45%)	10 (91%)	10 (91%)
conceptual	9 (69%)	3 (30%)	10 (83%)	8 (80%)
environment	6 (46%)	10 (63%)	11 (85%)	16 (100%)
non-functional	3 (50%)	4 (40%)	5 (83%)	9 (90%)
review	4 (67%)	8 (57%)	6 (100%)	13 (93%)
Total	205 (56%)	175 (47%)	311 (85%)	328 (88%)

RQ3 | 結果

回答時間

- 画像の有無で明らかな差は見られず，画像は回答を得るまでの時間をはやめない可能性を示している



まとめ

画像の役割を理解するためにSOで実証的な調査を行なった

調査の結果、以下のことが明らかになった

- SOでは様々な画像が共有されており、多くの質問を理解するために画像が不可欠である
- 画像を含む質問タイプは、how-to, discrepancyであったが、説明の長さに大きな違いはない
- 画像を含む質問は、画像を含まない質問よりもaccept answerを得られる可能性が高いが、時間が短縮されることはない

所感

画像はあくまでも補助的な役割を担っている

- 画像は直感的であるが、テキストによる説明が基本的に必要
- テキストで必要な情報は？
 - 環境情報, 実現したいこと, など？

動画 (gif) だとどうなる？

- 情報量は画像よりも多いが、テキストが必要であることに変わりはない
- 不具合の場合、動きがない（動作しない）場合が多く、画像と変わらない？