

Stack Overflow における検索支援を目的とした 参照関係可視化システムの提案

白木 大貴¹ 楨原 絵里奈^{2,a)} 小野 景子³

概要： Stack Overflow (以下 SO) はプログラミングに関する QA プラットフォームであり、ユーザは SO における質疑や質疑回答の閲覧を通して、ソフトウェア開発に対する知見の深化が可能である。ユーザは SO 内の投稿を検索できるが、SO には大量のデータが蓄積されており、ユーザは求める情報にたどり着くための適切な単語や道筋が分からない場合がある。そこで本研究では、SO 内における質疑の投稿に含まれるリンクに着目し、投稿同士の関係性を可視化するツール SO-RVG を提案する。先行研究より、リンクが集中する投稿やリンクを多く含む投稿は、ユーザの検索内容に対する前提や基礎的な知識を多く含むため、SO-RVG によって投稿同士の関係性が可視化されると、SO における情報検索を通じたソフトウェア開発教育をユーザに提供できると考える。SO-RVG の評価実験では、SO-RVG を使用しない場合に比べて、20 名中 13 名が検索性の向上を述べ、また、可視化された投稿同士の関係性を基に、ユーザは期待した回答に加えて、さらに新規の知識取得に成功した。以上より、SO-RVG によるリンクを基にした投稿の参照は、ユーザへ検索性の向上だけでなく、ソフトウェア開発に関わる新規知識の獲得にも寄与できると考える。

キーワード： Stack Overflow, 参照関係, 可視化, ネットワークグラフ

RO-RVG: Supporting tool for Visualizing Relationship of Posts on Stack Overflow

DAIKI SHIRAKI¹ ERINA MAKIHARA^{2,a)} KEIKO ONO³

1. はじめに

近年、様々な技術の登場に伴いソフトウェア開発が活発化している。ソフトウェア開発における知識獲得の手段として、Stack Overflow^{*1} (以下、SO) というプログラミングに関する QA サイトが存在する。SO には 2022 年 10 月時点で約 2300 万の質問と約 3400 万の回答が投稿されており、開発者にとっての事実上のフォーラムとして機能している [1]。SO に集まった投稿は SO の検索エンジンを通

して検索が可能である。SO の検索機能はテキスト検索やフィルターを使用した高度な検索をサポートしており、他者の投稿の閲覧により知識獲得が可能である。ソフトウェア開発者はプログラミングに関する疑問に遭遇した際に、一般的に SO を利用して回答を見つけようとするわかっている [2]。さらに Storay らは、多くのソフトウェア開発者にとって SO を利用した情報習得は重要な日常的行動であると結論付けている [3]。以上より、SO における検索による知識獲得はソフトウェア開発において重要な行動であると考えられる。

しかし、検索エンジンを活用しても求めている情報の取得は容易でない [4][5]。例えば、数語のテキスト検索で意図した内容の表現は容易でなく、人間の検索意図と実際の検索テキストの間には一定の乖離が発生する [6][7]。また、特に開発初学者は自身の疑問の正確な言語化が困難である

¹ 同志社大学大学院
Doshisha University, Kyotanabe, Kyoto 610-0321, Japan

² 立命館大学
Ritsumeikan University, Ibaraki, Osaka 567-8570, Japan

³ 同志社大学
Doshisha University, Kyotanabe, Kyoto 610-0321, Japan

^{a)} makihara@fc.ritsumeai.ac.jp

^{*1} <https://stackoverflow.com>

他、エラーが発生した際に問題がどこにあるかを判断できないケースが存在する [8]。SO の知識情報は特定の質疑応答から成り立つという性質上、一つの投稿で獲得できる知識は限定されている [9]。その一方で、開発者はドキュメントに対して多様性を求めていると報告されており [10]、SO を用いたソフトウェア開発における知識習得では、投稿を基に多くの情報を提供することが重要であると考えられる。

そこで本研究では、開発初学者が検索を行う際に、検索結果の質を向上させる支援を目指す。具体的には、ユーザにとって興味のある記事を提示することで、主体的な学びを促進しつつ、プログラミングやソフトウェア開発に関する知識の拡張や底上げを目指す。本研究では、提示情報として SO で重要性が示唆されている参照の関心に着目した。本研究で言う参照とは、SO の投稿内に存在するハイパーリンクを用いた他投稿への遷移行動を指すものとする。参照は情報の結び付けを行っているため、情報の多様性を苦手とする SO において重要な役割を担っており [11]、参照によって結び付けられる情報の提示は検索結果の質向上につながると考える。

SO ではユーザはプログラミングに関する質問や、質問に対する回答を投稿できるほか、全てのユーザはこれらの投稿に対して投票による評価が可能である。この評価は投稿スコアとして質問や回答の良し悪しを示す指標となるが、我々が行った先行研究より、被参照投稿において投稿スコアと参照された回数（以下、被参照回数）の間には相関関係がほとんど存在しないと明らかになった [12]。例えば、PHP における計算時の小数点以下の扱いに関する返信*2は、投稿スコアは 4 だが 314 回の参照が行われている [12]。投稿スコアが低い理由は、類似した質問の存在、回答先の投稿が人が興味を持つタイトルであるかなど複数の要因が存在する。参照が多く行われている原因も様々考えられ、原因追及にはより詳細な調査が必要だが、他に頻繁に参照された投稿には、プログラミングに関する質問を行うときの良い例示について*3や、HTML と正規表現に関する回答*4などが存在する。これらを参照する投稿をランダムに確認したところ、前者は SO だけでなくオープンソース開発など多くのソフトウェア開発の場において汎用的に活用可能な知識である点、後者は回答の内容に加え記述のユニークさが参照されやすい要因になっていると考えられる。以上より、被参照投稿は既存の評価では別の尺度で評価されていると考えられ、被参照投稿をユーザに提示することで、ユーザは単純に類似した質問だけではなく、ソフトウェア開発やプログラミングに対する広い知識を養えると期待される。

そこで本論文では、SO 上に参照関係の可視化や被参照

回数のような参照関係を用いての検索を可能とするシステムである SO-RVG を提案する。具体的には、開発者に SO 上の投稿同士の参照関係や参照に関する情報を提示しつつ、それらを利用した検索が可能であるブラウザの拡張機能を提案する。そして、SO-RVG の有効性の検証として被験者実験を用いたシステムの評価を行い、SO における検索行動に対する支援の実現を目指す。SO-RVG によってユーザは情報探索が容易になるだけでなく、リンクの活用により、検索内容に関わる基礎知識や新しい知見など、セレンディピティ [13] に満ちた投稿を発見でき、SO を利用しソフトウェア開発について学ぶことが可能になると考える。

本稿の構成は以下の通りである。まず 2 章で Stack Overflow の説明や SO に関する既存研究を紹介した上で、SO における検索支援の課題を整理する。次に 3 章で提案ツール SO-RVG の各機能を紹介する。そして 4 章では SO-RVG を評価するため Research Question（以下 RQ）を設定し、5 章でそれぞれの RQ に回答する形で SO-RVG を評価する。最後に 6 章でまとめを述べる。

2. 準備

2.1 Stack Overflow

SO とはプログラマやソフトウェア開発者などがプログラミングに関する質問を投稿し回答を得るための場を提供する Web サイトである。SO には 2022 年 10 月時点で約 2300 万の質問と約 3400 万の回答が投稿されている。SO に登録しているユーザは質問や回答の投稿を通してプログラミングやその関連技術について議論を行う。

議論内容は全てのユーザが閲覧可能であり、投稿を閲覧したユーザは投稿に対して投票が可能である。投票は投稿にとっての評価となり、投票数を元にユーザは投稿の良し悪しを判断できる。本研究では以後投稿の投票数を投票スコアとして扱う。質問には複数の回答が投稿されることもあり、質問者は一つの投稿を accepted answer として選択することで解決状態の提示が可能である。さらに投稿にはコメントを付与可能である。コメントは投稿そのものに対して改善の意見や質問に対する説明要求を可能とし、投稿の質向上および適切な質疑の実現を促す効果がある [14]。

SO では既存の投稿を検索するための検索機能*5が提供されている。検索機能を用いてテキスト検索やフィルターによる条件を追加した高度な投稿検索を実行可能である。例えば、投稿にはタグが付与されているため、タグを指定し検索を行うと検索結果の情報の絞り込みが可能である。

2.2 SO における参照行動

SO では多くの情報を結びつける手段として、ハイパーリンクの共有による外部情報の参照が行われる。図 1 は

*2 <https://stackoverflow.com/a/48056129> (2024/07/24 参照)

*3 <https://stackoverflow.com/q/5963269> (2024/07/24 参照)

*4 <https://stackoverflow.com/a/1732454> (2024/07/24 参照)

*5 <https://stackoverflow.com/search>

SO の投稿のうち、投稿に対するコメントにリンクを含む例である*6。図 1 の赤枠で囲った箇所は “Here’s an example that I found” や “here’s a recent example.” などの文言とともにリンクが提示されている。また、SO の各質問のページには Linked や Related という項目があり、参照された記事は Linked に記載されている。

Vincent らは SO を含む大規模コミュニティと Wikipedia の間に存在する関係性について調査を行った [15]。具体的には、Wikipedia への参照を含む投稿について調査を行った。Vincent らは調査結果として、参照を含む投稿は価値のある投稿として観察されており、参照を含まない投稿と比較して投稿スコアが高くなる傾向があると示している。Gomez らはソフトウェア開発者の知識探索行動の調査を目的とし、外部情報の付与を目的とした SO におけるリンク共有行動について分析を行った [11]。調査結果より、参照は投稿同士を結び付けるエコシステムを形成するうえで重要な役割を果たしていると結論付けた。実際に、SO は質疑の形式で情報が保存されているため解決手段の提示に長けている一方で、情報の内容は限定的なものであると確認されている [2][9]。Baltes らは SO の Android と正規表現に対する投稿における、外部リンクの参照やその目的を調査した [16]。調査の結果、リンク参照の目的は多岐にわたり、単純な過去の投稿の参照から、質問の背景にあたる知識の提供、情報の要約など様々であった。

また、我々は先行研究において SO 内の投稿の参照について調査を行った [12]。先述の通り、SO の各質問投稿には Linked という項目で、該当の質問に対し、回答やコメント内でリンク、すなわち参照された投稿が表示されている。一方、特定の質問や回答に対し、それら自体を参照する投稿について、ユーザは知る手段が無い。そのため、1 章で述べた PHP における計算時の小数点以下の扱いに関する返信のような、投稿内にリンクが存在しないが、多くの質問や回答で引用される投稿の存在は不明瞭であった。多く引用されるということは、内容が有益であるに加え、SO ユーザが質問や回答を行う際の良い例として用いることができると考える。そこで被参照の特徴を調査した結果、被参照回数と投稿スコア間に相関関係は見られなかった。また、被参照回数が多い投稿には、正規表現のように特定の言語に依存しない汎用的な知識に加えて、投稿スコ

アが低いものも存在した。

以上より、参照は SO において重要な役割を果たしており、参照を行う投稿は開発者の情報取得に貢献するに値する投稿であると考えられる。

2.3 SO の開発者支援に関する研究

近年、SO におけるソフトウェア開発者の支援を目的とした研究は多く存在する。mondal らは SO における未回答質問に着目し、回答が付与されない可能性がある投稿の検出を目的とした質問投稿分析を行った [17]。調査結果として、投稿内容の可読性やトピック、質問投稿者の経験値が回答の付与されやすさに関係していると発見した。Kumar らは SO における未回答質問の回答を提供するツールを提案している [18]。具体的には質問と回答のデータをグラフベースのクラスタリング手法を用いて分類し、未回答質問に対して同じクラスタ内の他の投稿済みの質問を利用して回答を提供するツールを提案した。Cao らは SO のテキスト検索効率上昇を目的とし、深層学習を用いた検索テキスト候補の再構築機構を提案した [19]。提案モデルは 5 つの評価指標で基準値を上回っており、投稿検索におよび検索投稿の推薦の実現が見込めると結論付けた。

以上より、SO において情報取得を目的とした開発者を支援する試みは多数行われており、質問投稿のみならず投稿検索においても開発者が適切な情報を獲得できるような支援は重要である。

2.4 提案システム構築に向けた要件定義

先行研究 [12] や先述した既存研究より、ユーザへ被参照投稿を提示することにより、検索支援につながると結論づける。そこで本論文では参照関係を利用した検索を支援するシステムの提案を行う。一方で、現状の SO における検索システムは参照関係の情報や被参照回数について考慮していない。また、テキスト検索ではユーザの意図とテキストの間、およびテキストと検索結果の間に乖離が発生する可能性がある。そこで、参照関係を通じて関連する投稿を視覚的に捕捉可能とすることで、より効率的な投稿検索が可能となると考える。視覚的な情報提示は理解力の向上や意思決定の改善に寄与するものであり、SO における参照関係の視覚的な補足は有効であると考えられる [20]。以上より、本稿では以下の要件を満たすことで、参照に関する情報を提供し SO における検索を支援するシステムを提案する。

- R1 投稿ごとの被参照回数が確認可能である
- R2 参照関係を視覚的に捕捉可能である
- R3 参照関係を元に投稿を検索可能である

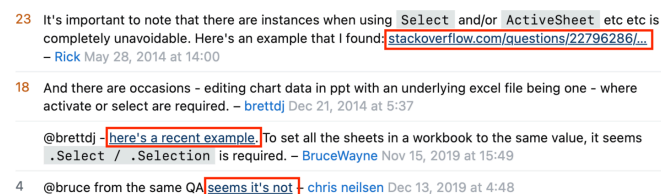


図 1: Stack Overflow におけるリンク (赤枠部分) の使用例

*6 [https://stackoverflow.com/q/10714251\(2024/05/30参照\)](https://stackoverflow.com/q/10714251(2024/05/30参照))

3. 提案ツール

3.1 SO-RVG: Stack Overflow Reference Visualizer with Graph

本稿では SO における投稿検索の支援を目的とし、以下の機能を有したシステムである Stack Overflow Reference Visualizer with Graph(以下、SO-RVG)を構築する。F1 から F3 の各機能は 2.4 節において述べた要件 R1 から R3 へそれぞれ対応している。

F1 閲覧ページを参照する投稿一覧表示機能

F2 参照関係ネットワークの可視化機能

F3 参照関係を用いた投稿検索機能

図 2 に SO-RVG の概要図を示す。SO-RVG はユーザが操作を行うフロントエンド部とデータの処理を行うバックエンド部に分割されている Web アプリケーションである。フロントエンド部はインターネットブラウザである Chrome の拡張機能によって構成される。Chrome 拡張機能は Google Chrome ブラウザ^{*7}の拡張により、新たな機能の提供や既存のページに情報を付与するツールおよびアプリケーションを指す。バックエンド部は Web サーバおよびデータベースによって構成されており、Web サーバは Python の Web フレームワークである Flask^{*8}、データベースは SQLite を使用した^{*9}。投稿参照のデータ取得に関して、表示速度の向上と、投稿間の参照関係の取得に時間が要する点から、リアルタイムな更新は行っていない。SO-RVG は SO の投稿の中でも、他の SO 内の投稿の参照を扱うため、後述するネットワーク図は短期間で大きく変化しにくいと考える一方、トレンドなツールや技術についても扱う場合、データベースの定期的な更新が必要である。参照関係の取得が容易になった際にはリアルタイムな更新も可能になると考えるが、表示速度を含めたユーザビリティの向上は今後の課題としたい。

SO-RVG は、投稿閲覧時に投稿 ID を基に投稿同士の参

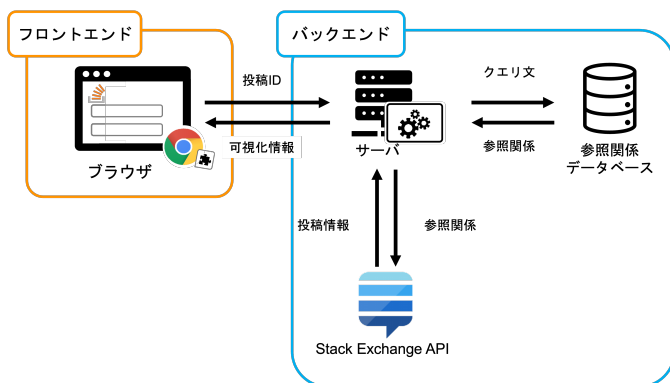


図 2: SO-RVG の概要図

*7 <https://www.google.com/intl/ja-jp/chrome/>

*8 <https://flask.palletsprojects.com/en/3.0.x/>

*9 <https://www.sqlite.org/index.html>

照関係をデータベースから取得する。次に、参照関係を基に Stack Exchange API(以下 SE-API)を通して投稿情報の取得を行い可視化情報を完成させる。そしてフロントエンド部が可視化情報を受け取り、Chrome 拡張機能を通してユーザへ視覚的な参照関係の提示を行う。ユーザは SO-RVG を導入したブラウザを通じて F1 から F3 の各機能を利用可能である。

また、SE-API から取得した情報を以下に示す。

- 投稿 ID
- 投稿タイトル
- 投稿内容
- 投稿区分（質問、回答、コメント）
- 投稿に付与されたタグ名
- 投稿スコア

なお、投稿内容は被参照情報のデータベースを作成するのみに用いており、後述する SO-RVG の各機能では、投稿内容を除いた 5 つの情報を利用している。

被参照情報については下記の手順で取得を行った。

手順 1: SE-API からリンク情報を元に参照を使用している投稿を全て取得する。

手順 2: 取得した投稿の参照部分の URL から参照先の投稿を取得する。参照には Wikipedia や公式ドキュメントの参照など復習種類が存在するが、本研究では“<https://stackoverflow.com>”を含むもののみを取得した。

手順 3: 手順 2 と同時に参照回数を取得し、投稿毎に保存する。

3.2 システム機能

図 3 に SO-RVG が提供する機能の全体像を示す。なお、図 3 は実際の SO における特定の投稿ページ（図 3 におけるノード 1）^{*10}へ SO-RVG を埋め込んだものである。このように SO-RVG は既存の SO の質問および回答に追加する形式で機能提供を行うものとなっている。SO-RVG は該当ページに対し参照を行っている投稿内容の一覧表示領域（F1 に対応）、グラフビュー領域（F2 に対応）、検索機能領域（F3 に対応）に分割されている。以下、各機能の詳細を述べる。

F1: 閲覧ページを参照する投稿一覧表示機能

本機能では被参照回数を含んだ投稿内容を提示する。被参照回数は先行研究 [12] と同様に、ハイパーリンクの共有によって他投稿から参照を行われた回数とする。そして、投稿ごとにタイトル、投稿スコア、被参照回数が提示され、ユーザはグラフビューで提示された投稿の内容を確認できる。具体的には、投稿スコアはオレンジ色の枠組みで、被参照回数は水色の枠組みで認識可能である。本機能によ

*10 <https://stackoverflow.com/q/2707322> (2024/07/24 参照)

り、ユーザは被参照回数を一つの指標として認識しながらの検索を行える。

F2: グラフビュー

参照関係の可視化において、本研究ではネットワークグラフを採用する。ネットワークグラフを採用した理由は、主に大規模なコミュニティの可視化に適している点と、直感的に投稿同士の引用・繋がりが分かる点があげられる。SO には Linked といった参照先の情報が一覧表示されているが、該当投稿を参照している投稿については公開されていない。また、参照されている投稿がさらに他の投稿を参照する場合も多く、SO には単純なリスト表示では表せない複雑な参照関係が構築されている。以上より、1つの投稿に対し必要な知識や、参考となる知識が複数の投稿に分散する機会が多いため、これら一連の投稿を可視化できるネットワークグラフが適していると考えた。また、構築したネットワークグラフに対しグラフサンプリングなどの既存のグラフ分析手法を適用すると、局所部分の構造や特徴が分析でき、参照される理由の分析や、個人の質問傾向に特化した関連トピックの抽出なども将来的に可能であると考える。

以上より、本機能では参照関係ネットワーク可視化を目的とし、投稿および投稿同士の参照関係に関するデータ構造を属性付き有向グラフ $G = (V, E, A)$ として構成する。 V はノードの集合、 $E \subseteq V \times V$ は有向辺の集合、 A は属性の集合を表す。 V の要素 $v_{start}, v_{end} \in V$ はノードを表し、 E の要素 $e = (v_{start}, v_{end}) \in E$ は有向辺を表す。ここで v_{start} は参照を行っている投稿、 v_{end} は被参照投稿とする。 v は写像 $attribute: V \rightarrow 2^A$ によって属性と結び付けら

れる。また、属性は $A = \{ \text{投稿タイトル, 投稿スコア, タグ, 被参照回数} \}$ とする。そして、本機能では、Web サーバから受け取った可視化情報を基にグラフビューを D3.js^{*11} を用いて描画し可視化を行う。

グラフの描画は情報過多にならず、一覧で情報を取得できる UI が望ましい [21]。そのため、本機能では被参照回数に基づいて表示するノードの数およびエッジ数の制限を行う。なお、ノードの数およびエッジの数については後述の検索機能にて変更可能である。また同様の理由により、投稿ごとのタイトルは、ノードに対しマウスホバー操作によって表示非表示を調節できる。

加えて、ノードの色の濃淡は中心ノード、すなわち閲覧しているページからの距離を示している。図 3 は Java における NULL の役割を問う投稿ノード 1 (以下投稿を示すノードを “N[番号]” と示す) の表示画面であるが、例えば N[2] や N[5] は N[1] の投稿内で直接リンクが張られており、N[1] の質問に直接関連する投稿である確率が高い。実際、N[1] の回答の一部は直接 N[5] に用いられている。一方、N[3] や N[7] は N[1] を直接参照しておらず、N[3] が参照する N[2] が N[1] を参照、あるいは N[1] を参照する N[6] の異なる参照先が N[7] である。したがって、これらノードは N[1] の関連知識として有用である場合が考えられる。例えば N[7] は Java における NULL チェックの避け方に関する質問であり、N[1] に比べると具体的な問題の解決法である。

本機能により、ユーザは参照関係の可視化を通して確認できる情報の質や量が増加し、検索を行いやすくなると期待できる。さらに、グラフを用いた可視化により、グラフ探索によるセレンディピティに富んだ知識獲得が期待できる。セレンディピティとは「偶然によって思いがけず価値のあるものを発見する能力」と言われ [13]、本提案システムを通し、ユーザは特定の記事の中で参照された新規の知見や意外性のある関連投稿などを知った結果、ソフトウェア開発に関する知識の深化が期待される。

F3: 検索機能

本機能ではグラフビューで示された投稿に対し、検索可能な機能を提供する。ユーザは検索したい文字列をテキストボックスに入力し、その文字列を含む投稿タイトルを持つ投稿を取得できる。検索結果は、グラフビューにおいてノードが強調表示されるため、ユーザは複数あるノードから知りたい情報を取得できる。

次に、タグ一覧から投稿の絞り込みが可能である。タグ一覧はグラフビューに含まれる投稿のタグを表示しており、チェックボックスに対してチェックを入れると、該当のタグが付与されている投稿の絞り込みが可能である。絞り込み結果は投稿タイトルの検索機能と同様に、ノードの強調

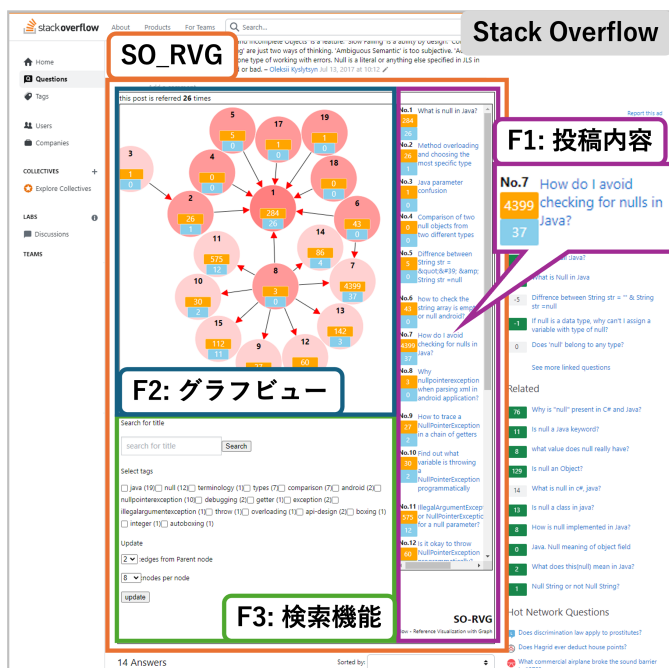


図 3: SO-RVG における機能全体像

*11 <https://d3js.org/>

表示によって認識可能である。さらに、投稿ごとのエッジ数やノード数の最大数を設定できる。そのためユーザは適切な描画量のグラフビューのもとで検索を実現する可能である。本機能により、ユーザは参照関係を利用して適切な検索を実現できると期待される。

3.3 使用シナリオ

本研究の目的は、SO における検索性の向上、およびユーザが自ら興味のある記事を探し出し、プログラミングやソフトウェア開発に関する主体的・意欲的な学びを育むツール・仕組みの提供である。加えて、単純に特定のエラーやバグの解決方法について知るだけでは無く、該当エラーが生じた原因やそれらに関連する知識など、知識の底上げを目指す。

特定のエラーに関する単純な解決策を求める場合、書籍やプログラミング解説の Web サイト、近年では ChatGPT をはじめとする大規模言語モデルの活用が考えられる。一方、これらツールを用いてユーザがエラーを解決した場合、そのエラーが出た根本的な原因、同様の原因によって生じる他のエラー、必要と思われる背景知識など、該当のエラーに関する幅広い情報の調査、取得はユーザにとって煩雑な作業である。ChatGPT はエラーの原因を段階的に解説する場合もあるが、エラーの根本的な原因や、該当エラーに関係する知識の有無はユーザによって様々である。そこで、特定のエラーや知識に関し、ユーザが興味を持つ類似事例を提示したり、ユーザにとって既知の情報と現在生じたエラーの関係を可視化すると、ユーザは自身の知識や興味を基に、該当のエラーについて自学自習が可能だと考える。

SO の場合、例えば Java のオーバーロードに関する質問^{*12}に対し、類似した質問への解説^{*13}、さらにオーバーロードの際の扱いが複雑である NULL の説明^{*14}がそれぞれリンクで繋がっている。また、これらの質問は Android 開発に関する NULL のエラー事例や、String における NULL と空の違いなど、Java に関する異なる投稿ともリンクで繋がっている。SO-RVG でリンクを可視化することで、ユーザは自身の疑問に対する関連知識や基礎知識を深掘りでき、さらにリンクで繋がる投稿がユーザにとって意外であるほど、ユーザの興味を引き出せると考える [13]。

以上より、宿題などで生じたプログラムのエラー解決時の使用は、SO-RVG の主目的とは異なる。SO-RVG は自学自習時などユーザが主体的にプログラミングやソフトウェア開発について学ぶ際、特定のエラーや事象について、意外性のある投稿の提供、自身の持つ知識と問題との関係などの可視化により、ユーザの SO における検索を指示し、

特定のエラーに関する知識の拡張、底上げを支援する。

4. SO-RVG の有効性検証に向けた評価実験

4.1 実験概要

本実験の目的は SO-RVG が SO において検索支援を実現できるかどうかの検証である。本実験では SO の検索を必要とするタスクを準備し、SO-RVG の有無で検索内容やタスクの結果がどのように変化するかについて調査を行う。具体的にはタスク実行において SO-RVG を用いる場合と SO-RVG を用いない場合を全ての被験者において調査する。また、全タスクの実行後にアンケート調査を行い、SO-RVG の主観的評価を行う。なお、本検証ではタスクの妥当性を調査するために事前実験（以下、実験 A）を行い、実験 A の結果を元にタスクの調整をした後に本実験（以下、実験 B）を行う。実験 A および実験 B では、タスクの内容以外における実験条件は同等のものとした。

具体的な実験手順を以下に示す。

- 手順 1 実験目的の説明および SO-RVG の導入 (10 分)
- 手順 2 SO-RVG の使用方法の確認 (5 分)
- 手順 3 タスクの実行 (70 分)
- 手順 4 本実験に関するアンケートを実施 (無制限)

4.2 実験条件

被験者は、C および Java におけるプログラミング経験がそれぞれ 1 年以上ある 20 代大学生であり、実験 A で 10 人、実験 B で 10 人、計 20 人に協力いただいた。実験 B では実験 A の内容を踏まえ、タスク内容の一部変更およびタスク実行時の参考投稿の記述要求を行う調整を行った。本調査では SO-RVG の有無が生み出す差異の調査を行うため、同種のタスクを 2 つ準備し、SO-RVG の有無による差異を調査する。さらに、本実験では異なる状況下における検索行動を考慮して調査を行うため、本実験では 2 種類のタスクを 2 個ずつ、計 4 個のタスクを準備する。タスクは SO-RVG 無の場合は SO の検索システムのみを用いて行い、SO-RVG 有の場合は SO の検索システムと SO-RVG を併用して実行する。

表 1 に実験 B で扱うタスクの詳細を示す。T1-1 および T2-1 は複数プログラミングに関する語句の意味を調べるタスクであり、T1-2 および T2-2 はプログラミング言語に関する想定質問をできる限り挙げるタスクとなっている。本調査では順序効果を考慮し、被験者の半数は SO-RVG 有りであり T1-1 および T1-2 を、SO-RVG 無しでタスク T2-1 および T2-2 を行い、残りの半数は SO-RVG の有無の順序を逆にして同様の実験を行う。

タスク選定にあたり、下記 2 点を考慮した。まず 1 つ目に SO-RVG を用いて投稿を検索するためには、参照関係を多く含む題材を選ぶ必要がある。先行研究 [12] より、SO には Git, C, string など、特定のツールやプログラミング

^{*12} <https://stackoverflow.com/q/13237114> (図 3 における 3)

^{*13} <https://stackoverflow.com/q/9361639/> (図 3 における 2)

^{*14} <https://stackoverflow.com/q/2707322> (図 3 における 1)

言語、データ構造等に関する投稿が多いと分かった。また、被験者全員が C 言語、Java の使用経験があったため、基本文法や用語などで差が生じる可能性が低いと考え、本タスクでは C 言語と Java 言語を対象とした。2 つ目に、エラー解決をタスクとした場合、簡単なエラーの解決は、被験者が自身の持つ知識でのみ回答し SO-RVG を使用しない恐れがある一方、難しいエラーは被験者が回答の妥当性を把握できない可能性がある。先述した本研究の目的より、本来被験者は SO-RVG で可視化された問題間の参照関係より、自身の興味がある質問や投稿の深掘りが期待される。しかしながら、各被験者にとって興味のある質問をそれぞれ設定する場合、記事の多さや参照数が異なるため、被験者によって検索性に差が出てしまい、評価実験として公平性、妥当性にも欠けると考えた。

以上より、T1-1、T2-1 では参照投稿が十分にある単語を選定しつつ、被験者の事前知識が少ない単語を与え、SO-RVG を用いて各単語に関し検索してもらい SO-RVG の検索性を評価した。T1-1、T2-1 で被験者が特定の単語に関し多くの学びを得た場合、SO-RVG は通常の検索エンジンなどと同様に、SO における投稿の検索が満足に行えると考える。また、T1-2、T2-2 では被験者全員へ「あなたはプログラミング初学者へプログラミングを教えてください。T1-2 では C 言語、T2-2 では Java 言語を初学者が学習する上で、想定される質問と質問に対する回答を考えてください」と伝えた。本指示を行った理由は、特定のデータ構造や機能について指定しなかった場合、被験者は本来の SO-RVG の目的に則り、SO-RVG を通して自身が興味のある C 言語・Java 言語の要素について、自由な検索が促されると考えたためである。これらの回答については著者およびプログラミング教育経験が豊富な教員で確認し、記述の多さに差はあるものの、内容として誤った回答、不適切な回答は無いと判断した。

一方、上記タスクは SO-RVG が本来想定する、ユーザの意欲による自学自習とは異なる。SO-RVG で該当要素への知識・興味の向上や知識の底上げがユーザへ及ぼす影響については、長期的な使用を必要とするため、今後の課題としたい。

4.3 Research Question

本実験では以下の 3 つの RQ の観点から有効性の検証を行う。本調査で用いたアンケート内容は「システムを用いて希望通りに検索を行えたと思うか」などの質問形式と、自由記述で分かれている。質問形式の場合は、5 段階評価とし、ネガティブ評価を 1、ポジティブ評価を 5 とした。

RQ1：SO-RVG は容易に検索を行えるか

SO-RVG はユーザに参照関係を可視化したグラフを提示した結果、検索においてユーザが求める適切な投稿を選択可能となると期待する。そのためには、SO-RVG が検索

表 1: 実験タスク詳細

タスク番号	内容	時間
T1-1	語句「floating point」「casting」「reference」「environment variables」「callback」に関する説明を作成する	15 分
T1-2	「C 言語」の学習における想定質問と答えの組みをできる限り作成する	20 分
T2-1	語句「scope」「null」「regex」「error handling」「asynchronous」に関する説明を作成する	15 分
T2-2	「Java」の学習における想定質問と答えの組みをできる限り作成する	20 分

支援システムとして成立することが重要である。そこで、SO-RVG の有無による検索体験を比較し、SO-RVG の使用を通して検索を容易に行えるかどうかを調査する。本 RQ ではアンケート調査の結果を基に考察を行う。なお、本調査は実験 A および実験 B の双方を用いて考察を行う。

RQ2：SO-RVG の有無は獲得情報に差異を生じるか

SO-RVG は開発者に参照関係を可視化したグラフを提示するため情報量の増大が期待できる。そのため、SO-RVG の有無によってユーザが獲得する情報内容に差異が生じるかを調査する。本 RQ では、アンケート調査の結果による定性調査、およびタスク実行時の投稿閲覧ログの取得による定量調査を踏まえて、獲得情報量の差異について調査する。なお、本調査は実験 A および実験 B の双方を用いて考察を行う。

RQ3：SO-RVG は作業にどのような影響を及ぼすか

SO-RVG はユーザに追加で情報を提示するため、既存の SO と比較し検索によって生じる作業内容に差異を発生させると考える。そこで、SO-RVG の有無によってタスクによる成果物の内容にどのような差異が発生するかを調査する。本 RQ では、SO-RVG の有無によってタスクにおける回答作成数がどのように変化するかについて着目し考察を行う。なお、本調査はタスク内容に焦点を当てた考察を行うため、実験 B の結果のみを用いて考察を行う。

5. 実験結果

5.1 RQ ごとの結果と考察

以下、RQ に従って結果と考察を述べる。

RQ1：SO-RVG は容易に検索を行えるか

図 4 および図 5 に実験 A および実験 B における SO-RVG の検索に関するアンケートの結果を示す。本調査は SO-RVG の検索性について調査を行った。そのため、被験者が通常の検索エンジン同様に、SO-RVG を通して SO 内の投稿を検索ができたと感じるかについて、被験者の主観に基づきアンケートに回答してもらった。SO における投稿の内、被験者の理解度に最も適した投稿が選択されたかに

については、被験者のプログラミング能力と併せたより詳細な調査が必要だと考える。

図 4 より、平均して 3.7 の評価を受けている他、全体の半数以上が 4 および 5 の評価をしているとわかる。この結果より、SO-RVG は検索支援システムとしての役割を果たしたと考える。さらに、図 5 より実験 A および実験 B の結果を合算して 20 人中 13 人が SO-RVG 有の方が検索が行いやすいと回答した。そのため、SO-RVG により検索がより行いやすくなると考える。

変わらない、またはシステム有と回答した被験者に対して評価理由の調査を行ったところ、SO-RVG の処理速度の遅さやそれによるユーザ体験の低下について言及していると分かった。そのため、上記の意見は SO-RVG の主目的である参照関係の可視化とは直接関係するものではないと考える。以上より SO-RVG は検索支援システムとして成立しており、かつ既存の SO と比較して検索が容易に行えるシステムであると考えられる。

RQ2：SO-RVG の有無は獲得情報に差異を生じるか

図 6 に実験 A および実験 B における SO-RVG で獲得した情報に差異を感じたかに関するアンケート結果について示す。結果より、およそ半数の被験者において SO-RVG を用いた検索において差異を感じたとわかった。差異を感じた被験者に対して理由の調査を行ったところ、被参照回数的大小によって表示内容が異なるという意見を 2 人の被験者から得た。具体的には、被参照回数が多い投稿はその

内容が汎用的な知識である一方で、被参照回数が少ない投稿はその内容が限定的であると感じたというものであった。これは、先行研究 [12] の被参照投稿における調査結果と同等の内容であり、調査内容の裏付けを行う内容となっていると考える。また、被参照回数が多い投稿は含んでいる情報量が多いと感じたという意見も得た。そのため、SO-RVG は被参照回数的大小を通してユーザが汎用的な内容を含む投稿や、情報量に富む投稿を能動的に選択可能となると考える。

次に、図 7a に実験 A および実験 B において SO-RVG によって閲覧した投稿と SO の関連質問機能との関係性を示した結果を示す。本結果は、SO-RVG を用いて閲覧した投稿が SO の関連質問機能によっても提供されている投稿でも存在したかどうかを、被験者の閲覧投稿ログを対象に実施した結果である。なお、SO の関連質問機能とは、既存の SO に存在する、投稿閲覧時に表示される関連質問をユーザに提示する機能である。結果として、20 人中 19 人において SO-RVG を通して閲覧した投稿が関連質問機能では得られない投稿であるとも判明した。さらに、図 7b に実験 B において SO-RVG によって閲覧した投稿をタスクの回答作成における参考内容として使用したかの結果結果を示す。結果として 10 人中 8 人において、SO-RVG でのみ提供された投稿を回答作成の参考に使用している旨を確認した。そのため、SO-RVG ではユーザが求める情報であるが、既存の SO では提示されない情報を獲得可能であると考えられる。

RQ3：SO-RVG は作業にどのような影響を及ぼすか

表 2 に実験 B における SO-RVG の使用がタスクの回答作成数に与えた影響について示す。結果として、T1-1 および T2-1、T1-2 および T2-2 の双方において、10 人中 9 人の回答作成数が同等もしくは増加したと確認した。これは SO-RVG の使用により、一定時間あたりの作業量が向上したために回答作成数が変化したと考える。なお、回答作成数が減少した被験者についてその原因を調査したところ、

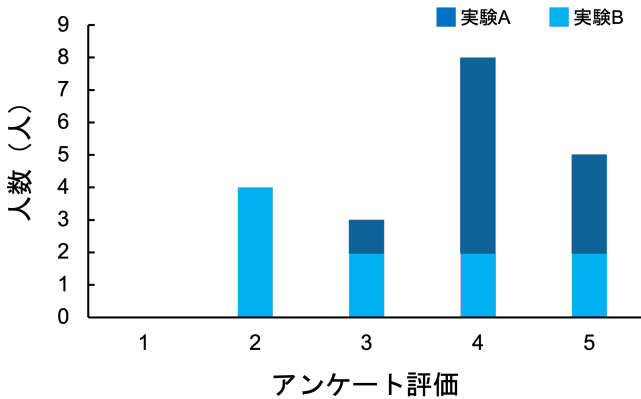


図 4: SO-RVG を用いて検索が実現できたか (1: できなかったと感じる, 5: できたと感じる)

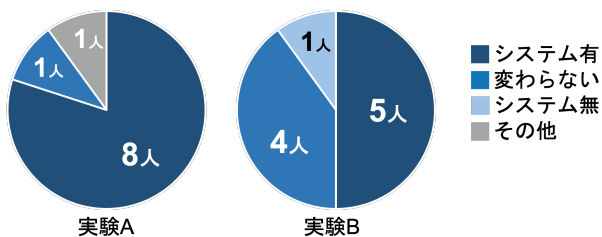


図 5: SO-RVG の有無による検索のしやすさの優劣

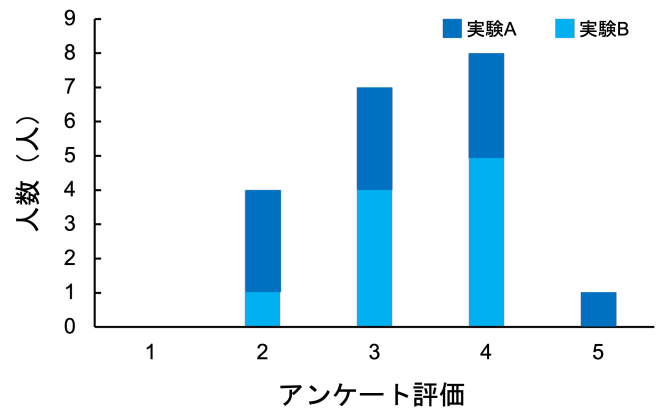


図 6: SO-RVG の有無によって獲得情報に差異を感じたか (1: 特にならと感じた, 5: 差異があると感じた)

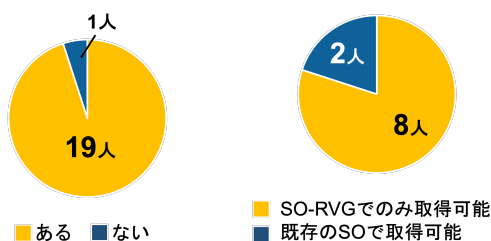
表 2: SO-RVG の使用によるタスクごとの回答作成数の変化

被験者	T1-1 および T2-1	T1-2 および T2-2
被験者 A	増加	増加
被験者 B	同等	増加
被験者 C	同等	同等
被験者 D	同等	減少
被験者 E	同等	同等
被験者 F	減少	同等
被験者 G	増加	同等
被験者 H	同等	増加
被験者 I	増加	同等
被験者 J	増加	同等

T1-1 および T2-1 においては、システムを用いた際のタスク実行において「regex」の回答作成時に多くの時間を消費しているとわかった。具体的には、SO の検索システム使用の時点で検索において多くの手戻りが発生していた。そのため、本被験者における回答作成数の差異はタスク内容が影響していると考えられる。また、T1-2 および T2-2 においては被験者から SO-RVG の使用に至る前に作成した回答が多く存在したという意見を得た。そのため、SO-RVG の有無が回答作成にあまり影響を及ぼさなかったと考える。したがって、回答作成数が減少した理由は SO-RVG の使用によるものではないと考える。以上より、SO-RVG の使用により、一定時間あたりの作業量が向上すると考える。

5.2 SO-RVG に対する考察

前節より、SO-RVG の使用により多様な情報を獲得できる他、検索速度の向上が見込めるとわかった。そのため、SO-RVG の使用は検索の質を向上させ、結果として検索支援につながると結論づける。一方で SO-RVG の処理速度の遅さやそれによるユーザ体験の低下が存在すると確認した。そのため、SO-RVG の処理速度が向上した場合、ユーザからより高い評価が獲得できる可能性がある。さらに、一部の被験者からは調べたい内容とは関係がない内容の



(a) SO-RVG を通してのみ閲覧できた投稿の有無
 (b) SO-RVG を通してのみ閲覧できた投稿を回答作成に参考にしてたか

図 7: SO-RVG によって閲覧した投稿と SO の関連質問機能との関係性

表示が存在するという意見を得た。そのため、ユーザの意図に沿った内容を優先的に提示する機能を実装した場合、より検索の質が向上するシステムが実現すると考える。また、本実験は短期的な使用による評価を行っているため、システム操作の慣れは考慮していない。そのため、今後は長期的な使用を踏まえた実験を通し、多様な観点からシステム評価を行う必要がある。

RQ3 において、SO-RVG の使用を通し検索の過程で新たな知見や意外性のある記事が発見されると期待したが、SO-RVG を用いる前に自身の知識を持って回答する被験者も見受けられた。一方で、RQ1 と RQ2 の結果より、SO が推薦する関連投稿以外の投稿も、SO-RVG は推薦可能だと考える。したがって、SO-RVG におけるセレンディピティについては今後長期的な使用を通して調査を行う必要がある。

本稿では定性的な調査が主となっているため、定量的な側面からも SO-RVG の有効性を調査する必要がある。例えば、SO 内の関連する質問とリンクで繋がった質問の重複率や、ユーザが閲覧した投稿数、検索結果をまとめたユーザの回答の質の調査などが考えられる。

5.3 妥当性の脅威

外部妥当性として、タスクの調査対象が考えられる。今回対象とした単語以外を調査する場合、リンクがより効果的に働く、あるいは働かない可能性がある。一方、先述の通り、本タスクで調査する単語は先行研究の結果を基に十分なリンクが含まれる投稿を利用した。そのため、検索性に関し本調査内容は一定の妥当性を有すると考える。

また、内部妥当性として実験設計があげられる。本調査ではユーザが SO-RVG を通して投稿を閲覧し、それによって学びを深めていき、タスクに回答する形式を取った。一方、ユーザによっては最初に目に付いた記事のみ閲覧し、それ以上関連投稿の検索を行わない場合があった。RQ2 より、20 名中 19 名が SO-RVG でのみ閲覧できた投稿があり、うち 10 名中 8 名は SO-RVG を通してのみ閲覧できた投稿を参考にしているため、SO-RVG でリンクによる参照を元に提示された投稿はユーザの知識獲得に寄与したと考えられるが、学びの深化に関しては継続的な実験を通して調査する必要がある。

6. おわりに

本研究では、SO における投稿の検索性の向上、およびユーザのソフトウェア開発に関する知識獲得の支援のため、SO で重要な役割を担う参照行動に着目した。そして参照される投稿に着目した調査によりその重要性を再確認し、参照関係を活用した検索支援システムである SO-RVG を提案した。そして、評価実験を通して SO-RVG が実際に SO における検索行動を支援可能かを検証した。

結果として SO-RVG は検索支援システムとして有効であり、既存の SO と比較して容易な検索が実現可能であり、さらに多様な情報の獲得や検索速度の向上にも貢献できた。SO-RVG における課題として、処理速度の向上やユーザの求める情報を考慮した投稿の提示により、ユーザ体験を向上させる必要がある。加えて、本稿では短期的な使用における評価を行ったため、今後は長期的な SO-RVG の使用を踏まえた実験により、更に多角的な視点からシステムを評価可能であると考えられる。

謝辞 本研究は JSPS 科研費 24K20905 の助成を受けた。

参考文献

- [1] Lena Mamykina, Bella Manoim, Manas Mittal, George Hripcsak, and Björn Hartmann. Design lessons from the fastest q&a site in the west. In *In Proc. of the SIGCHI conference on Human factors in computing systems*, pp. 2857–2866, 2011.
- [2] Christoph Treude, Ohad Barzilay, and Margaret-Anne Storey. How do programmers ask and answer questions on the web?(nier track). In *In Proc. of the 33rd international conference on software engineering(ICSE)*, pp. 804–807, 2011.
- [3] Margaret-Anne Storey, Alexey Zagalsky, Fernando Figueira Filho, Leif Singer, and Daniel M German. How social and communication channels shape and challenge a participatory culture in software development. *IEEE Transactions on Software Engineering(TSE)*, Vol. 43, No. 2, pp. 185–204, 2016.
- [4] Xin Xia, Lingfeng Bao, David Lo, Pavneet Singh Kochhar, Ahmed E Hassan, and Zhenchang Xing. What do developers search for on the web? *Empirical Software Engineering(ESE)*, Vol. 22, pp. 3149–3185, 2017.
- [5] Chunyang Chen and Zhenchang Xing. Towards correlating search on google and asking on stack overflow. In *In Proc. of the IEEE 40th Annual Computer Software and Applications Conference(COMPSAC)*, Vol. 1, pp. 83–92. IEEE, 2016.
- [6] Ritendra D, Dhiraj J, Jia L, and James Z W. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, Vol. 40, No. 2, pp. 1–60, 2008.
- [7] Zheng-Jun Zha, Linjun Yang, Tao Mei, Meng Wang, Zengfu Wang, Tat-Seng Chua, and Xian-Sheng Hua. Visual query suggestion: Towards capturing user intent in internet image search. *ACM Transactions on Multimedia Computing, Communications, and Applications(TOMM)*, Vol. 6, No. 3, pp. 1–19, 2010.
- [8] Preetha Chatterjee, Minji Kong, and Lori Pollock. Finding help with programming errors: An exploratory study of novice software engineers’ focus in stack overflow posts. *Journal of Systems and Software(JSS)*, Vol. 159, p. 110454, 2020.
- [9] Chris Parnin, Christoph Treude, Lars Grammel, and Margaret-Anne Storey. Crowd documentation: Exploring the coverage and the dynamics of api discussions on stack overflow. Technical report, Georgia Institute of Technology(Technical Report), 2012.
- [10] Michael Meng, Stephanie Steinhardt, and Andreas Schubert. How developers use api documentation: An observation study. *Communication Design Quarterly(CDQ)*, Vol. 7, No. 2, pp. 40–49, 2019.
- [11] Carlos Gómez, Brendan Cleary, and Leif Singer. A study of innovation diffusion through link sharing on stack overflow. In *In Proc. of the 10th working conference on mining software repositories(MSR)*, pp. 81–84, 2013.
- [12] 白木大貴, 横原絵里奈, 小野景子, 谷祐貴, 花本凜. Stack Overflow における参照利用に関する実証調査～プログラミングにおける QA コミュニケーション支援に向けて～. 研究報告ソフトウェア工学 (SE), Vol. 2023-SE-214, No. 10, pp. 1–7, 2023.
- [13] A. Foster and N. Ford. Serendipity and information seeking: an empirical study. *Journal of Documentation*, Vol. 59, No. 3, pp. 321–340, 2003.
- [14] Subhasree Sengupta and Caroline Haythornthwaite. Learning with comments: An analysis of comments and community on stack overflow. In *In Proc. of the 53rd Hawaii International Conference on System Sciences(HICSS)*, 2020.
- [15] Nicholas Vincent, Isaac Johnson, and Brent Hecht. Examining wikipedia with a broader lens: Quantifying the value of wikipedia’s relationships with other large-scale online communities. In *In Proc. of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018.
- [16] Sebastian Baltes, Christoph Treude, and Martin P Robillard. Contextual documentation referencing on stack overflow. *IEEE Transactions on Software Engineering(TSE)*, Vol. 48, No. 1, pp. 135–149, 2020.
- [17] Saikat Mondal, CM Khaled Saifullah, Avijit Bhattacharjee, Mohammad Masudur Rahman, and Chanchal K Roy. Early detection and guidelines to improve unanswered questions on stack overflow. In *In Proc. of the 14th Innovations in software engineering conference(ISEC)*, pp. 1–11, 2021.
- [18] Abhishek Kumar, Deep Ghadiyali, Sridhar Chimalakonda, and Akhila Sri Manasa Venigalla. Socluster-towards answering unanswered questions on stack overflow via answered questions. In *In Proc. of the 16th Innovations in Software Engineering Conference(ISEC)*, pp. 1–5, 2023.
- [19] Kaibo Cao, Chunyang Chen, Sebastian Baltes, Christoph Treude, and Xiang Chen. Automated query reformulation for efficient search based on query logs from stack overflow. In *In Proc. of the IEEE/ACM 43rd International Conference on Software Engineering(ICSE)*, pp. 1273–1285. IEEE, 2021.
- [20] Karin Eberhard. The effects of visualization on judgment and decision-making: a systematic literature review. *Management Review Quarterly(MRQ)*, Vol. 73, No. 1, pp. 167–214, 2023.
- [21] Miriam Arnold, Mascha Goldschmitt, and Thomas Rigotti. Dealing with information overload: a comprehensive review. *Frontiers in Psychology*, Vol. 14, p. 1122200, 2023.