

Stack Overflow における編集活動：知識の公開という観点から

田島逸郎 (慶應義塾大学大学院)

niryuu@keio.jp

I 背景

プログラミングに関する質問回答サイト Stack Overflow(SO) は、Web 上で幅広く参加者が集まり質問回答を行うことを通じて、その記録が広く Web 上でアクセス可能にすることを旨としたサイトである¹⁾。「世界中の中の良いプログラミングの知識の総和を集合的に増やす」目的をもって設立され、実際に 2019 年 10 月現在で 1800 万件の質問が公開されている。それを実現するために、様々な機能やガイドラインなどが設定されている。

特に、SO では質問者と回答者に限らない参加者による共同作業を重視しており、編集やコメントなどを通じて、質問者や回答者以外も質問回答に参加できるようになっている。それは多様な人々の参加を可能にすると同時に、互いに異なる視点を持った人々が断片的に質問から始まる場に関与するという点でもある。

SO での編集は Wikipedia の影響を受けたものであり、「Stack Overflow の核」であるともされている。SO の質問や回答は、一定の貢献を行った利用者が編集することができる。質問や回答が編集された場合、編集履歴へのリンクが表示され、編集履歴画面では第 1 図のようにどこが編集されたかと、編集者によるコメントなどが表示される。SO の「Help Center」画面における「Why can people edit my posts? How does editing work?」の項目では、質問をすべき場合について「文法やスペルの間違いを直す」「投稿の意味をより明確にする」「コメントで追加された情報を含める」「小さい間違いを直し、古い投稿を更新する」「関連する情報源やハイパーリンクを加える」の 5 種類が挙げられている。

編集は、質問回答をより改善する手段として、質問回答の一部として行われていると考えられる。このため、上記に挙げたガイドラインは存在するが、実際の編集は質問回答というやりとりの文脈の中で適切な形で行われる必要がある。そして、それは質問回答の趨勢や、さらには知識の公開にもかかわっていると考えられる。では、それはどのように行われているのだろうか。

II 関連研究

SO における編集に焦点を当てた研究は 2 つある。Guo Li らは、SO における編集モデルが質問回答の品質を上げる一方で、新参加者の参加を妨げるのではないかという問いを立て、各種の編集行動を定量的に分析した²⁾。その結果、質問の表題、内容、タグの編集が質問に対する回答や投票を増やすこと、回答の編集は投票を増やすものの、新参加者の回答の編集は将来的な回答行動を減らすことなどが明らかになった。Jie Yang らは、SO において良い質問を作ることが協働で行われることに焦点を当て、質問行動の分析を行った³⁾。まず、編集をソースコードの改善、質問の文脈の追加、詳細の追加、入力と出力の例の追加、問題の明確化、試したことの追加、の 6 種類に類型化した。その上で編集行動が品質と関連していること、コミュニティに長くいるにしたがって編集行動が減ることなどを明らかにした。しかし、両研究はともに、具体的な質問回答の中で編集行動がどのように行われているかについては注目していない。

III 目的と方法

本研究は、Stack Overflow における編集が、質問回答という共同作業の中でどのように行われているかという実践の側面を理解することを目的とする。

編集を活動として捉える場合、それは明らかに質問回答の一連の流れの中にある。このため、編集という現象を捉える際には、その参加者が編集をするまさにそのとき、質問から始まった場の何を理解し、どのように編集を行ったかということ、つまり実践が焦点となる。

例えば質問者が編集を行った場合、質問から始まるコメントや回答などの一連のやりとりを理解し、その上で質問の編集を行うという時間的な幅を持ったものになるかもしれない。一方で、既に解決された質問回答を編集する参加者は、経緯を知らないまでもその記録から場面を自身なりに理解し、編集を行うかもしれない。いずれにせよ、編集者がその場その状況に即した形で編集を行うという事実があるからこそ、



第 1 図 Stack Overflow の編集履歴ページ (<https://stackoverflow.com/questions/5597060/> より)

第 1 表 本稿で提示する SO の投稿

質問番号	質問 ID	タイトル
1	19669786	Check if element is visible in DOM
2	10865025	Merge/flatten an array of arrays
3	5597060	Detecting arrow key presses in JavaScript
4	12789231	Class type check

編集という活動が成立するのである。このため、編集が何を指してどのように行われているかを研究することが、SO を理解する手掛かりになる。

この視点は、エスノメソドロジエ的ワークの研究、すなわちある特定の場面がどのように成立しているかを、その活動の成り立ちの中から見出していく研究の視点である⁴⁾。ワークの研究においては、ある場面に参加している人がその場面を理解し、自身も適切に物事を行っていく方法が、研究者にとってのその場面を理解する方法にもなる。本研究では、実際の編集を含む事例を集め、SO 特有の編集の特徴を成立させている方法を記述し、例示していく。その際には、個々の編集行為を理解する必要に応じて「テキストを読み書きすること」「プログラムを書くこと」などへの言及も行いながら関連研究についても言及する。

本研究で用いるデータセットは、質問数が多いカテゴリで、なおかつアクセス数の多い質問を選ぶことにした。具体的には、初心者から

熟練者まで幅広い利用者の存在するプログラミング言語”javascript”のタグを含む、アクセス数の多い質問を抽出した。Google Bigquery 上にホストされている 2019 年 5 月までのデータセットを利用し、質問 ID の抽出を行った。その上で、具体的な質問の編集活動を観察し、その特徴を例示した。本稿で扱うデータを第 1 表に示す。質問 ID は SO において与えられた値であり、”<https://stackoverflow.com/questions/質問ID/>”の形でアクセスできる。なお、次節以降では質問番号によって質問を参照する。

IV Stack Overflow における編集の実践

Rod Watson は、人々の活動の中で文献やサインなどの様々なテキストがどのように使われているかについて述べた⁵⁾。その中で、テキストは複数の読まれうる人々に向けてデザインされること、テキストはそれを読むという活動によって活性化され、それゆえにテキストと読むことは切り離せないということについて指摘した。その観点で SO における編集という活動を

考えると、編集前に既に書かれた質問などの文章(これまでの編集履歴を含む)を読み、そして他の人に読めるように変更を加えるものと捉えることができる。

SOにおける編集を観察した結果、「誰に向けて編集をしているか」に関して、大きく2つに分かれることがわかった。想定される回答者に向けた回答を促す編集と、公開された情報源として読まれることを考慮し、既に成立した質問回答を管理するための編集である。

A 回答を促すための質問の編集

質問行為は、潜在的な回答者が答えられるような質問として成立するように行われる⁶⁾。編集もまた、回答を得るため、もしくはより望ましい回答を促すために行われることがある。

質問1においては、質問が行われた9分後に質問者自身によって編集が行われている。最初の質問では実現したいことと試したことについて述べたが、コメントで提示された解決策に対して「より一般的な方法でやりたい」と返信した。その後編集において「何をチェックしたらよいかわからない」と述べ、思いついた案を例示することで、より期待している解決策を明確にした。

質問1では、まだ回答が存在しない状態で、質問者がコメントでのやりとりにおいて欠けている情報を認識し、追加している。これは質問者が潜在的な回答者から回答を得るために編集をしていると見ることができる。

B 情報源としての質問回答の質の管理

一方で、質問者が既に満足できる回答を獲得した後も、編集は行われる。それは、何を指してどのように行われるのだろうか。

質問2においては、2012年に行われた既に回答の存在する質問において”How would I go about making this just:”となっている部分が、2018年に”How would I go about merging the separate inner arrays into one like:”とより明確な形に編集されている。

質問1と異なり、質問2では既に多くの回答がなされ、質問者も回答を受け入れた状態で質問を明確化している。つまり、質問2では質問者や、潜在的な回答者、コメントをしている者など、質問回答の場に参加している者に向けて編集をしているとは限らない。むしろ、ただ質問回答の内容を改善することに主眼が置かれている。それは、特定の閲覧者を想定していないが、編集によって質問回答を情報源として利用

するものにとってより「明確に」なりうる。

それが明確に述べられた事例も存在する。質問3においては、2012年になされた質問に対し2012年に編集が加えられている。その編集コメントで編集者は「この質問はgoogleで高いランクに位置している。だからコードを修正し、いくつかの文法を修正し、新しいタグを加えた」と、当該質問がGoogleで検索されることを見越して編集をしたことを述べている。

また、質問3においては、「コードの修正」が編集内容に加わっている。SOの質問、回答においては、問題や解決策となっているプログラムのソースコードを掲載することが質問を明確にするという点で推奨されている。実際にソースコードが掲載されていることは一般的であり、ソースコードが編集されることもある。質問3の場合は、プログラムが入れ子構造になっているが、その構造が見やすいようにスペースを付与し整形する編集がなされている。

Graham ButtonとWes Sharrockは、コンピュータプログラムをプログラマーの実際のワーク”lived work”を示すものであり、また作業を行う際に用いられるものであるとした⁷⁾。その際、「ハッカー」は単にプログラムを書くのに対し、「プロフェッショナル」はそれに加えてプログラムやその構造を、他の人にとって読みやすく書くと指摘した。その上で、様々な実践を例示したが、その中で構造を見やすくする整形についても指摘されている。

その点で、質問3で行われている編集は「プロフェッショナル」による実践である。「googleで高いランクに位置している」から整形を行うということは、質問回答が他の「プロフェッショナル」にアクセスされることを見越し、より読みやすいように改善するということであるといえる。

質問4においては、TypeScriptに関する質問がなされ、それに対し情報源と説明を付加した回答がなされた。Javascriptには、他の言語からJavascriptに変換できる言語(AltJS)が存在し、その中にTypeScriptも含まれている。この質問は2012年になされたが、質問者は2017年に他のAltJSを使うようになったことと、その理由を編集によって追加した。しかし、これは本来のTypeScriptに関する質問とは関係ないため、他の参加者によって編集前のものに「巻き戻し」された。それに対して質問者は何度も追加部分を編集することを繰り返したが、その

たびに巻き戻しされた。

このことは、回答がなされた後であっても、質問内容と回答内容が整合性の取れたものになっているかが参加者によって考慮され続けることを示している。質問者が回答を受け入れたあとにアプローチを根本的に変えたとしても、それは以前に行われた質問と回答のペアとは関係がない。一旦なされた質問と回答は、その後も根本的な部分については維持される。それは、回答がなされた後にも公開された情報源として質問回答にアクセスされうることを考慮しているとも考えられる。

V 議論と結論

本研究では、SOにおける編集活動が、質問回答の中で何を指してどのように適切に行われているかを記述し、例示した。そして、その中に質問者や回答者などのその場に居合わせた参加者に向けたものと、公開された情報源として訪れた者に向けた編集があることを明らかにした。その上で、他の人にもわかるようにプログラムを書く「プロフェッショナル」としての編集、問題が解決した後も質問回答の形を守り続ける編集などが存在することを明らかにした。

これらのことから、SOにおける編集は、質問回答という形を取りながらも、それに直接参加していない、「プロフェッショナル」なプログラマーの共同体に向けて広く公開されることを志向して行われることもあることがわかった。これは、SOにおいて「知識の総和を集合的に増やす」ことに実際に参加者が従事していることを示唆する。そして、それは質問と回答の整合性を維持することによって可能になっている。

以上のことは、「SOで何が行われているか」に関して、参加者が何を指してどのように物事を行っているかについて記述するワークの研究によって明らかにできたことである。特にSOのようなサイトにおいては、システムの複雑さや参加者の多様性、参加の断片性などにより、行動をその文脈から切り離して分析したり、もしくは参加者が実際に参照することのないような行動過程によって分析されることが多い。その中でワークの研究のアプローチは、編集などの小さく断片的な活動であっても、十分に合理的で様々なことを考慮に入れた活動であることを明らかにすることができる可能性がある。

参考文献

- 1) 田島逸郎. "Stack Overflow における「知識の総和の増加」への志向と達成”. 三田図書館・情報学会研究大会発表論文集. 2017, p.25-29.
- 2) Li, Guo; Zhu, Haiyi; Lu, Tun; Ding, Xi-anhua; Gu, Ning. "Is It Good to Be Like Wikipedia?: Exploring the Trade-offs of Introducing Collaborative Editing Model to Q&A Sites". CSCW '15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. ACM, 2015, p.1080-1091.
- 3) Yang, Jie; Hauff, Claudia; Bozzon, Alessandro; Houben, Geert-Jan. "Asking the Right Question in Collaborative Q&A systems". Proceedings of the HT'14. ACM, 2014, p.179-189.
- 4) Rouncefield, Mark; Peter Tolmie eds. Ethnomethodology at work. Ashgate, 2011.
- 5) Watson, Rod. Analysing practical and professional texts: A naturalistic approach. Routledge, 2016.
- 6) 田島逸郎. "専門的 Q&A サイトにおける質問行為の分析”. 三田図書館・情報学会研究大会発表論文集 2015 年度. 三田図書館・情報学会. 2015, p.1-4.
- 7) Button, Graham; Sharrock, Wes. "The Mundane Work of Writing and Reading Computer Programs". Have, Paul ten; Psathas George eds. Situated Order: Studies in the Social Organization of Talk and Embodied Activities. University Press of America, 1995, p.231-258.