

質問で不足している情報を回答で指摘されたユーザが再投稿した質問の調査

小島 正裕 水野 友之 渡辺 靖彦 岡田 至弘

龍谷大学大学院 理工学研究科 情報メディア学専攻

{t10m101,t060627}@mail.ryukoku.ac.jp, {watanabe,okada}@rins.ryukoku.ac.jp

1 はじめに

質問を投稿しておく他ユーザが答えてくれるコミュニティベース質問応答サービス(以下 Q&A サイト)がさかんに利用されている。例えば、Yahoo!知恵袋は2004年から1年あまりの間に300万件以上の質問が投稿されている。このように、大量の質問が投稿されているQ&Aサイトでよい回答を得るためには、回答者にとって回答する負担が少ない質問、特に回答するのに十分な情報が書かれている質問を投稿することが重要である。例えば、以下の(質問1)は、(回答1)で指摘されているように、回答するのに十分な情報が書かれていない。

(質問1) PCが青い画面になって起動しません。どうしたらいいでしょうか？

(回答1) OSは何を使っていますか？OSが分からないと答えられません。

回答するのに重要な情報((質問1)では「OSは何か」)を説明していないことが質問を投稿する前に指摘されれば、この質問を投稿したユーザはその情報を質問で説明したであろう。そこでわれわれは、質問での説明が不十分であるかもしれない情報を指摘する質問作成支援システムを作成した[3]。このシステムでは、質問での説明が不十分であるかもしれない情報を指摘するのに、Yahoo!知恵袋に投稿された質問と回答から取り出した知識を用いる。しかし、Yahoo!知恵袋から取り出した知識は膨大で、どの情報をどのように提供すればいいのかわからず、さらにそれらの情報をユーザがどのように利用するのかはまだ明らかになっていない。これらを明らかにして効率的な質問作成支援を実現するため、Yahoo!知恵袋^{*1}の「パソコン・周辺機器」カテゴリに投稿された質問と回答を対象に、質問での説明が不十分なことを回答で指摘されたユーザがその後質問を再投稿したケース(図1)について

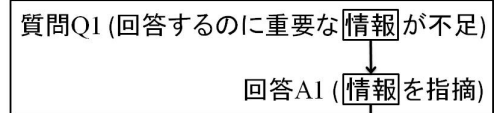
質問での説明が不十分なことを回答者はどのように指摘したのか

回答での指摘を質問者はどのように利用して質問を再投稿したのか

について調べる。Yahoo!知恵袋の公開データに収録されている質問には1件以上の回答が必ずあり、そのうち1件がベストアンサーに選ばれている。また、質問および回答が投稿された時刻、そして回答の受付が締め切られてベストアンサーが決定された時刻も秒単位で記録されている。

^{*1} <http://research.nii.ac.jp/tdc/chiebukuro.html>

スレッド1



スレッド2

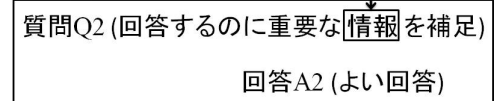


図1 質問での説明が不十分なことを回答で指摘されたユーザが、その指摘を利用して質問を再投稿し、よい回答を得るまでの流れ

2 回答で指摘された情報を利用して再投稿した質問

本研究では、Yahoo!知恵袋の公開データの「パソコン・周辺機器」カテゴリに収録されている質問と回答を対象に予備調査を行った。

2.1 質問での説明が不十分な情報についての回答での指摘の仕方

予備調査の結果、表1に示すように、質問(図1では質問Q1)での説明が不十分なことを回答(図1では回答A1)で指摘する場合、その指摘の仕方には以下の5つのタイプがあった。

TYPE (A1-1) 質問者が問題解決のために行ったことを確認する。

(質問2) シマンテックからセキュリティの警告が来て、延長サービスの更新をせねばならないのですが、延長キーを忘れてしまいました。どうしたらいいですか？

(回答2) 延長キーを購入しましたか？

TYPE (A1-2) 問題の状況について確認する。

(質問3) ネットをしていて、画面を最大化しても、画面いっぱいおきくならず、上が二センチほど空きます。ワードなど、ネット以外のときは画面いっぱいになるんですが、どうすれば元にもどりますか？

(回答3) その2センチの空白部分には何が表示されていますか？もしかして真っ黒ですか？

TYPE (A1-3) 対象の属性について確認する。1章の(回答1)での指摘の仕方は、このタイプに分類される。

表1 質問での説明が不十分な情報についての回答での指摘の仕方のタイプ

| TYPE | 説明 |
|------|---------------------------|
| A1-1 | 質問者が問題解決のために行ったことを確認 |
| A1-2 | 問題の状況について確認 |
| A1-3 | 対象の属性について確認 |
| A1-4 | A1-1~3 以外について確認 |
| A1-5 | 解決方法を提案(ただし、その方法では解決できない) |

TYPE (A1-4) A1-1、A1-2、A1-3 以外のことで説明が不十分な情報について確認する。

(質問4) パソコン初心者なので、接続タイプが分かりません。どこをどうやればいいですか？

(回答4) 何の接続タイプのことでしょうか。

TYPE (A1-5) 問題を解決する方法を提案する。ただし、提案された方法では問題は解決できない。このタイプの回答では、質問で説明が不十分な情報は直接的には指摘されていない。しかし本研究ではこれを、説明が不十分なことを指摘する仕方の1つとして取り扱う。これは、回答で提案された方法では問題を解決できないという情報が質問では十分に説明されていないことを間接的に指摘していると考えたからである。

(質問5) 円の記号は¥(今の) 昔の記号はy にヨコ2本線がついてましたね。この記号をパソコンの中から探したいのですが、どこにあるでしょうか。急ぎますので、わかる方は教えてください。よろしくお願いします。

(回答5) 文字コードとしてはいいです。Windows や Mac でもその字形を持つ書体はインストールされていません。お急ぎでしたらソフトによりますが、y と=をサイズ調整して重ねるのが手っ取り早いと思います。

2.2 回答で指摘された情報についての再投稿した質問での利用の仕方

予備調査の結果、表2に示すように、再投稿した質問(図1では質問Q2)における回答(図1では回答A1)での指摘の利用の仕方には以下の5つのタイプがあった。

TYPE (Q2-1) 指摘された情報を質問に追加する。例えば、(質問6)の質問者は(回答6)で指摘された情報を追加して(質問7)を投稿している。(質問7)の下線部は(回答6)で指摘された情報を追加した部分である。このタイプの場合、質問者は知っていたのに質問では十分に説明しなかったことを回答で指摘されたと考えられる。

(質問6) iMac 初期化について教えて下さい m(-_-)m 買った時の状態もどしたいのですが IE と Outlook の情報はそのままの状態でもストックしておきたいのです。(メールアドレスの設定情報など) Macintosh HD のどのフォルダのファイルを保存しておけばよいのでしょうか？

(回答6) OS のバージョン、IE のバージョン、Outlook Express のバージョン、これがわからない事には正確な回答が出来ません。

表2 回答で指摘された情報についての再投稿した質問での利用の仕方のタイプ

| TYPE | 説明 |
|------|-------------------------|
| Q2-1 | 指摘された情報を質問に追加 |
| Q2-2 | 指摘された情報以外の情報を質問に追加 |
| Q2-3 | 指摘された情報では問題を解決できないことを説明 |
| Q2-4 | 指摘された情報そのものについて質問 |
| Q2-5 | 指摘された情報には触れず、同じ質問を再投稿 |

(質問7) iMac の初期化について教えて下さい m(-_-)m OS は9。IE と Outlook のバージョンは5です。買った時の状態にもどしたいのですが IE と Outlook の情報はそのままの状態でもストックしておきたいのです。(メールアドレスの設定情報など) Macintosh HD のどのフォルダのファイルを保存しておけばよいのでしょうか？

TYPE (Q2-2) 指摘された情報以外の情報を質問に追加する。例えば、(質問8)の質問者は(回答8)で指摘された情報が問題の解決に重要ではないと判断し、その情報を(質問9)には追加しなかった。しかし、(回答8)の指摘がなされた原因を解消するため、指摘された情報以外の情報((質問9)の下線部)を質問に追加している。このタイプの場合も、質問者は知っていたのに質問では十分に説明しなかったことを回答で指摘されたと考えられる。

(質問8) XP のパソコンに 98 の HDD つなげて起動して、ちゃんと立ち上がりませんか？異常がでたりしませんかね？ CPU は PEN4 です。スペックは XP を軽く動かせます。HDD は、今は SATA で 98 は ATA でつなぎます

(回答8) XP が先に入っているのなら、98 が起動できないでしょう。先に 98 を入れてから、もう一度 XP をインストールしないと、起動画面で OS を選べないです。

(質問9) 先ほど、XP のパソコンに 98 の HDD つなげて起動して~ の質問をしたものです。回答ありがとうございます。http://knowledge.yahoo.co.jp/service/question_detail.php?queId=1134820 デュアルブートしようとかじゃないんです。XP の HDD を取り払って、98 をつないで起動できますか？ってことです。

TYPE (Q2-3) 指摘された情報では問題を解決できないことを説明する。例えば、(質問10)の質問者は(回答10)で指摘された情報では問題を解決できなかった。そこでその情報を追加して(質問11)を投稿している。このタイプの場合、質問者が知らなかったことを回答で指摘されたと考えられる。

(質問10) 人が作った物をコピーして買ったエクセルのデータなのですが、開こうとするとマクロを有効にしますか？と出ます。マクロは有効にしたいしセキュリティーは高のまま使用したいです。そのままに電子証明して下さいと出ます。どのようにしたら良いですか？

(回答10) マクロを無効にするか、セキュリティーを低くするしかないのではないのでしょうか？

(質問11) 先ほど「人から買ったデータ、マクロを有効に、セキュリティーを高で、電子証明が」で質問したものです。

エクセルを開いてマクロを有効を選択して、その後エクセルのデータ更新リンクが出たので更新するを選択しました。セキュリティを低いで設定しましたが、開きません。やはり電子証明がないとダメですか？当方で電子証明を取ると開く事は出来ますか？無知ですいません。

TYPE (Q2-4) 指摘された情報そのものについて質問する。例えば、(質問 12) の質問者は (回答 12) で指摘された情報を活用する方法がわからない。そこで、(質問 13) では指摘された情報の活用方法について質問している。

(質問 12) エクセルでマクロを作成した後、何日か経過してマクロ操作が不要になったのでこのブック内のマクロをすべて削除しましたが、次回からブックを開くたびにマクロを有効にするかどうかのメッセージが毎回出てきます。メッセージが出てこないようにしたいのですがどうすればよいでしょうか。

(回答 12) モジュールを開放してないのでは？ VISUAL BASIC EDITER を開いてモジュールの開放をしてください。

(質問 13) エクセルでマクロを作成した後、何日か経過してマクロ操作が不要になったのでこのブック内のマクロをすべて削除しましたが、次回からブックを開くたびにマクロを有効にするかどうかのメッセージが毎回出てきます。メッセージが出てこないようにしたいのですがどうすればよいでしょうか。モジュールの解放が有効らしいですが、操作手順が分かりません。おかげでこのエクセルでの保存時間が毎回とても長く、重くなっています。どなたか詳しい方、具体的にお教えてください。

TYPE (Q2-5) 指摘された情報には触れず、同じ質問を再投稿する。例えば、(質問 14) の質問者は (回答 14) で指摘された情報には触れず、(質問 15) を投稿している。

(質問 14) 光学式のマウスを使っていますが、最近、動きがにぶくなってきました。ボール式のマウスで言うと、うまくボールが転がらない状態。イライラして何度もコロコロ(?) しているうちにマウスの矢印が固まってしまって、動かなくなります。これは寿命でしょうか？それとも PC 側に原因があるのでしょうか？教えて下さい。

(回答 14) マウス裏の赤く光っている部分に糸ゴミとかがついていませんか？光学式マウスは光を出しテーブル等に反射した光を読み取ることでポインター(矢印)を動かしています。よって、光を出す部分と光を読み取る部分(光っている部分に向かい合うようになっています)に糸ゴミ等がついていると出した光をうまく読み取ることができずあなたの言うような症状がでます。光っている部分に息を吹きかける等してゴミを払えば改善されると思います。

(質問 15) 光学式のマウスを使用していますが、最近、動きがにぶいです。ボール式マウスでいうと、ボールがうまく転がらない状態。そのうち固まって反応しなくなります。これはマウスの寿命なんですか？それとも PC 側に問題があるのでしょうか？教えて下さい。お願いします。

3 質問での説明が不十分な情報について指摘した回答とその指摘を利用して再投稿された質問の検出

Yahoo!知恵袋の「パソコン・周辺機器」カテゴリに投稿された質問と回答を対象に、質問での説明が不十分なことを回答で指摘されたユーザがその後質問を再投稿したケースについて調査するため、

質問とその質問での説明が不十分な情報について指摘した回答(図 1 の質問 Q1 と回答 A1)

その指摘を利用して再投稿された質問とその回答(図 1 の質問 Q2 と回答 A2)

を以下の手順で取り出す。

(手順 1) 質問での説明が不十分であることを指摘する文を含む回答(図 1 の回答 A1)を取り出す。質問での説明が不十分であることを指摘する文は、磯貝らの方法で検出する [1] [2]。

(手順 2) (手順 1) で取り出した回答が投稿された質問(図 1 の質問 Q1)を取り出す。

(手順 3) (手順 2) で取り出した質問の投稿者によって投稿された質問のうち、以下の 2 つの条件を満たすもので投稿時刻が最も早いものを取り出す。

(手順 2) で取り出した質問が投稿された時刻以降に投稿された質問

(手順 2) で取り出した質問に対する回答の受付が締め切れ、ベストアンサーが決定されてから 72 時間経過した時刻よりも以前に投稿された質問

(手順 4) (手順 3) で取り出した質問を(手順 2) で取り出した質問と比較し、一致する自立語の数が以下のいずれかの条件を満たすものを回答での指摘を利用して再投稿された質問(図 1 の質問 Q2)として取り出す。

(手順 2) と(手順 3) で取り出した質問がいずれも自立語を 20 個以上含む場合、一致する自立語が 10 個以上ある

それ以外の場合、一致する自立語が 5 個以上ある

(手順 5) (手順 4) で取り出した質問に対して投稿された回答(図 1 の A2)を取り出す。

4 調査結果と検討

Yahoo!知恵袋の「パソコン・周辺機器」カテゴリに 2004 年 4 月から 2005 年 10 月までに投稿された約 17 万件的質問と約 47 万件的回答を対象に 3 章で述べた方法で、

質問とその質問での説明が不十分なことを指摘した回答(図 1 の質問 Q1 と回答 A1)

その指摘を利用して再投稿された質問とその回答
(図1の質問 Q2 と回答 A2)

を 4271 組取り出した。そのうち 200 組を無作為に取り出し、質問での説明が不十分なことを回答で指摘されたユーザがその後どのように質問を再投稿したのかを調査した。無作為に取り出した 200 組のうち、後で投稿した質問が先に投稿した質問の再投稿であったのは 133 組であった。その 133 組を対象に以下の調査を行った。その結果を表 3 に示す。

1. 質問を再投稿して、よい回答は得られたのか
2. 質問での説明が不十分なことを回答者はどのように指摘したのか
3. 質問での説明が不十分なことを指摘した回答に質問者は高い評価 (ベストアンサー) を与えたのか
4. 回答での指摘を質問者はどのように利用して質問を再投稿したのか

最初に、質問を再投稿してよい回答は得られたのか、特に回答での指摘は有効であったのかについて検討する。表 3 に示すように、質問を再投稿した結果、よい回答が得られたのは 133 組中 122 組であった。このうち、回答での指摘に触れずに質問を再投稿した Q2-5 の場合、26 組中 16 組でよい回答が得られた。一方、回答での指摘を利用して質問を再投稿した Q2-1~4 の場合、107 組中 106 組でよい回答が得られた。したがって、回答での指摘を利用して再投稿した質問ではよい結果が得られやすいことがわかる。

次に、質問での説明が不十分なことを回答者はどのように指摘したのかについて検討する。質問を再投稿するのに最もよく利用された回答での指摘のタイプは A1-5(解決方法の提案)である。ただし A1-5 では、回答者は質問での説明が不十分なことを直接的には指摘していない。したがって、質問者が提案された解決方法 (ただし、その方法では問題は解決できない) を見て、質問での説明が不十分なことに間接的に気づくことが多いことがわかる。一方、質問での説明が不十分なことを回答者が直接的に指摘するタイプでは、A1-2(状況の確認)と A1-3(属性の確認)が多い。

次に、質問での説明が不十分なことを指摘した回答に質問者は高い評価 (ベストアンサー) を与えたのかについて検討する。表 3 に示すように、質問での説明が不十分なことを直接的 (A1-1~4) あるいは間接的 (A1-5) に指摘した回答に質問者がベストアンサーを与えたのは 133 組中 112 組であった。「パソコン・周辺機器」カテゴリに投稿された回答約 47 万件のうちベストアンサーに選ばれたのは約 17 万件であるので、質問での説明が不十分なことを指摘した回答に質問者は高い評価 (ベストアンサー) を与えることが多いことがわかる。

最後に、回答での指摘を質問者はどのように利用して質問を再投稿したのかについて検討する。質問者は知っていたのに質問では十分に説明しなかったことを回答で指摘された Q2-1(指摘された情報を質問に追加)と Q2-2(指摘された情報以外の情報を質問に追加)のケースは 71 組ある。一方、回答で指摘されるまで質問者は知らなかった情報を質問を再投稿するのに利用している Q2-3(指摘された情報

表 3 質問での説明が不十分なことを回答者はどのように指摘したのか、そしてその指摘を質問者はどのように利用して質問を再投稿したのかについての調査結果。() はベストアンサーに選ばれた回答の数を示す。

(a) 質問を再投稿して、よい回答が得られたケース

| | A1-1 | A1-2 | A1-3 | A1-4 | A1-5 | 合計 |
|------|-------|---------|---------|---------|---------|-----------|
| Q2-1 | 5 (5) | 15 (14) | 21 (16) | 4 (4) | 10 (10) | 55 (49) |
| Q2-2 | 0 (0) | 1 (1) | 1 (1) | 7 (6) | 6 (4) | 15 (12) |
| Q2-3 | 1 (1) | 0 (0) | 0 (0) | 0 (0) | 18 (14) | 19 (15) |
| Q2-4 | 1 (1) | 4 (3) | 0 (0) | 0 (0) | 12 (10) | 17 (14) |
| Q2-5 | 0 (0) | 4 (4) | 1 (1) | 1 (1) | 10 (8) | 16 (14) |
| 合計 | 7 (7) | 24 (22) | 23 (18) | 12 (11) | 56 (46) | 122 (104) |

(b) 質問を再投稿してもよい回答が得られなかったケース

| | A1-1 | A1-2 | A1-3 | A1-4 | A1-5 | 合計 |
|------|-------|-------|-------|-------|-------|--------|
| Q2-1 | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Q2-2 | 0 (0) | 0 (0) | 0 (0) | 1 (1) | 0 (0) | 1 (1) |
| Q2-3 | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Q2-4 | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Q2-5 | 0 (0) | 2 (1) | 2 (1) | 1 (1) | 5 (4) | 10 (7) |
| 合計 | 0 (0) | 2 (1) | 2 (1) | 2 (2) | 5 (4) | 11 (8) |

では問題を解決できないことを説明)と Q2-4(指摘された情報そのものについて質問)のケースは 36 組ある。したがって、質問者は知っていたのに質問では不十分にしか説明しなかった場合が多いことがわかる。また、説明が不十分であるとの指摘だけで、質問者がその説明を自力で直して質問を再投稿した Q2-1~3 のケースは 90 組もある。このことから、詳細な説明を提供しなくても、質問での説明が不十分なことを指摘するだけでも有効な質問作成支援になることがわかる。

謝辞 本研究の一部は、日本学術振興会科学研究費補助金 (課題番号 20500106 と 22650021) の助成を受けて行われたものです。

参考文献

- [1] 磯貝, 西村, 渡辺, 岡田: Q&A サイトへの質問の作成を支援するための情報の抽出, 情報社会学会第 1 回知識共有コミュニティワークショップ, pp.19-28, (2008).
- [2] 磯貝, 西村, 渡辺, 岡田: Q&A サイトへの質問の作成を支援するための情報の複数のカテゴリからの抽出, 言語処理学会第 15 回年次大会, P1-8, pp152-155, (2009).
- [3] 小島, 西村, 磯貝, 渡辺, 岡田: 不足している情報を指摘する質問作成支援システム, 人工知能学会情報編纂研究会第 3 回研究会, (2010).