DEIM 2020 D6-1 (day2 p35)

QA サイトにおける質問応答に着目した気づきを促す問いかけの分析

齊藤　史明†　　山本　祐輔†

† 静岡大学総合科学技術研究科　〒 432–8011 静岡県浜松市中区城北 3–5–1
E-mail: †saito@design.inf.shizuoka.ac.jp, †yamamoto@inf.shizuoka.ac.jp

あらまし　本研究では，気づきを促す良い質問の生成・検索方法を検討するために，質問応答サイト（以下，Q&A サイト）において質問者に大きな気づきを与えた，回答者からの問いかけに関する分析を行う．一般に Q&A サイトでは，質問者が回答に対する感想や意見をまとめた返礼コメントを行うことができる．本研究では，分析用コーパスとして Yahoo!知恵袋を利用し，質問・回答文の言語的分析を行う．具体的には，返礼コメントに「考えてもいなかった」「ハッとした」等の文言があった場合，質問者は予想だにしていなかった気づきを得たと仮定し，質問と対となる回答から気づきを促すに至った問いかけ文を抽出する．その上で，それら問いかけ文の言語的な特徴を分析し，気づきを促す良い質問の同定・生成のための方法を検討する．本研究で得られた知見をもとに気づきを促す良い質問を生成・検索することができれば，対話システムや情報検索システムにおいて，ポジティブな態度・行動変容を促すインタラクションの設計が可能となる．

キーワード　QA データ処理，ソーシャルビッグデータ，対話分析

1　は じ め に

コミュニケーションで日常的に使われる質問の中には，質問を受けた人に気づきを促す良い質問が存在する．気づきとは自身の知らなかったことを新しく知ることだけではない．自身とは違う姿勢を理解することや、考えてもいなかった観点を認知することが挙げられる．また，質問を受けた人がそのトピックについて知識が不足していた・深く考えられていなかったことを認識するようになることも挙げられる．本稿では、このような効果を気づきを定義する．

トピックについて自身とは違う姿勢を知ることで，相手を慮る態度の形成や建設的な意見の提案が可能になる．また，多角的な視点で物事を考えることで様々なアイデアを生み出すことが可能になる．

ニュース記事や SNS 上の投稿といった情報は，書き手の立場や情報元などの違いにより，1 つの情報のみでそのトピックを網羅的に理解することは困難である．気づきによってトピックについて自身の見識が狭いことを自覚すれば，追加情報を得るための情報探索行動を行うようになると考える．

良い質問について粟津は，問われた人が思わず答えたくなる，新しい気づきを与えてくれる質問であると定義しており，本稿ではこの定義を引用する [1]．

気づきを促す良い質問を生成・検索することが可能になれば，対話システムや情報検索システムにおいて，ユーザに良い質問を提示することでポジティブな態度・行動変容を促すインタラクションの設計が可能となる．

本研究では，気づきを促す良い質問の生成・検索方法の検討のために分析用コーパスとして Q&A サイトの Yahoo!知恵袋 [1] から質問・回答文の言語的分析を行う．Yahoo!知恵袋では，投稿された回答の中で一番参考になったとされるものをユーザが選択し，ベストアンサーとして強調表示する機能が存在する．質問者がベストアンサーを選択した場合には，回答者や回答に対する感想や意見をまとめた返礼コメントを書くことができる．

Q&A サイトの流れと，本稿の調査対象の概略を図 1 に示す．

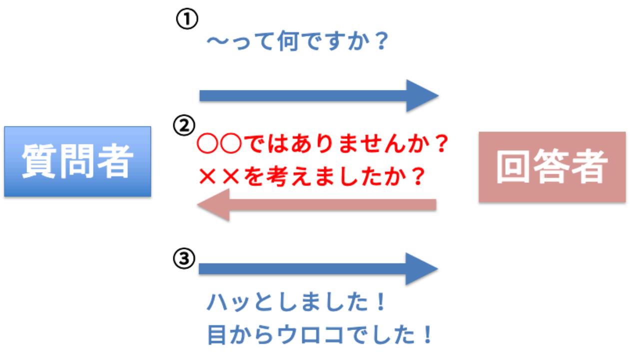

図 1　Q&A サイトの流れと調査対象の概略

返礼コメント (図 1 の 3）には回答者への感謝を示す目的のコメントが多いが，中には「考えたこともありませんでした」「〇〇だと思っていましたが〜」「勉強になりました」のように質問者が気づきを得たことを表現しているコメントが存在する．本稿では，その返礼コメントの内容から質問者が回答から新たな気づきを得たか判断する．

本稿は，返礼コメント内に出現した表現に注目し,「考えてもいなかった」「ハッとした」等の表現があった場合，質問者は予想だにしていなかった気づきを得たと仮定する．そして，その気づきを得ることになった回答から,「〜について調べたことはありますか？」「〜は試しましたか？」「〜はいかがでしょうか？」といった質問者に問いかけを行なっている文 (図 1 の 2）

1：https://chiebukuro.yahoo.co.jp/

を抽出する．この文が、本稿で考える人に気づきを与える問いかけ文であると仮定する．最終的に，それら問いかけ文に使われている品詞の割合や，使われている表現など言語的な特徴を分析し，気づきを促す良い質問の同定・生成のための方法を検討する．

本論文の構成は以下の通りである．次章で関連研究について述べる．3 章では分析データの概要および分析方法について述べる．4 章では分析結果について述べる．5 章では分析結果および気づきを促す良い質問の同定・生成のための方法の考察を述べる．最終章では本稿のまとめを行う．

2 関 連 研 究

2.1 Q&A サイトの活用・利便性向上

Q&A サイトの活用・利便性向上のための研究が数多く行われている．

Adamic らは Yahoo Answers（アメリカの Q&A サイト）[2] から，質問特性やユーザ間のコミュニケーションから，ユーザのクラスタリングを行なった [2]．

石川らは，Yahoo!知恵袋に投稿された回答の中からベストアンサーに選ばれるものを推定する手法を提案した [3]．クラウドソーシングによる実験を行い，どのような要素から質問者がベストアンサーを選択するのか分析した．実験の結果，特に回答の詳しさ，根拠の有無や種類，文章の丁寧さに基づいて評価を行なっていることが分かった．そして，その要素に基づいた特徴量設計で機械学習を行い，ベストアンサーを推定する学習器を作成した．

Li らは Q&A サイトのサービス向上のために，投稿された質問の質を半教師あり学習の Label Propagation を用いて測定する手法を提案した [4]．Ravi らも質問の質を測定する手法を提案したが，ラベルデータとして使う質の定義を，プログラミングに関する質問サイト「Stack Overflow」[3] の質問評価機能を元にすることで教師データ数を多く確保している [5]．

日暮らは，内容が分かりやすく伝わる質問タイトルを生成するために，ランク学習を用いた手法を提案した [6]．現在の Yahoo!知恵袋では，質問タイトルを本文と別に入力することができないため，本文の先頭部分が見出しとして表示される．クラウドソーシングによる評価を教師データとしてランク学習を行い，質問の文章の中で見出しにふさわしい 1 文を予測した．

本研究では，気づきを促す効果を持っていると考えられる質問を取得・分析するために Q&A サイトに投稿された回答と質問者の反応を利用する．

2.2 質問文の生成

Kunichika らは英語の長文読解学習の支援を目的として，中学生の英語教材の文章からその内容に関する質問を生成する手法を提案した [7]．英語知識の定着と英語による対話訓練のために 5 つのモジュールを作成し質問を生成した．その結果，約 93%の質問文が意味の通ったものになっており，中学生レベルの文法の約 80%を網羅していることが分かった．

Mostafazadeh らは，入力された画像に関する会話が成立するような質問文を生成する手法を提案した [8]．

Du らは，Sequence-to-Sequence [9] と呼ばれる系列データの処理に向いた学習モデルを使った質問文作成手法を提案している [10]．

Nishida らは，質問文を知りたい情報を正確に表現できるように修正する手法を提案した [11]．入力された質問に複数の回答候補を取得し，それぞれの回答候補のみを回答とする質問文をニューラルネットワークを使い出力する．

本研究では，問いかけることで気づきを促す質問に限定し，その同定・生成のための方法を検討する．

2.3 ウェブ上の行動変容への説得・インタラクションに関する研究

Bateman らは検索エンジンの有効的な利用を促進するために，ユーザに検索行動のフィードバックを行う Search Dashboard を提案した [12]．検索回数や閲覧したウェブページ数などの行動を単に数字で見せるのではなく，速度計を模した図で見せることで行動の変容を狙っている．

Saito らは，批判的検索行動を促進するために，ウェブページ上の文章で曖昧な表現が使われている文を強調表示して情報の信憑性について注意を促す手法を提案した [13]．

Yamamoto らは， 暗黙的に批判的検索行動を促進するために，検索クエリの推薦機能にデータの確認や比較などを促すクエリを挿入する“ クエリプライミング ” システムを提案した [14]．

本研究は，実際にインタラクションや説得についての効果を測るものではないが，気づきを促す質問の同定・生成の方法を検討することで，ポジティブな態度・行動変容を促すインタラクションの設計に貢献ができると考える．

3 分 析 方 法

3.1 分析データの概要

3.1.1 データセット

分析データには，ヤフー株式会社が提供した国立情報学研究所の情報学研究データリポジトリ「Yahoo!知恵袋データ（第 2 版）」[4] を利用する．Yahoo!知恵袋では，回答の中からベストアンサーの選択がされた質問を解決済みの質問と定義している．その解決済みとなった質問を抽出し，質問・回答文や質問カテゴリ，投稿日時などをデータセットとして提供されている．抽出の対象期間は 2004 年 4 月から 2009 年 4 月の 5 年間である．データ数は，質問データが 16,098,580 件，回答データが 49,673,309 件である．

3.1.2 質問のカテゴリとベストアンサー

質問を分類するためのカテゴリは大カテゴリ，中カテゴリ，小カテゴリの 3 段階に分けられている．質問にはカテゴリを必

2：https://answers.yahoo.com/
3：https://ja.stackoverflow.com/
4：https://www.nii.ac.jp/dsc/idr/yahoo/chiebkr2/Y_chiebukuro.html

表 1　質問数の多い中小カテゴリ名 (上位 10 カテゴリ) のベストアンサーと返礼コメントに関するデータ

詳細カテゴリ名	質問件数	平均回答件数	質問者がベストアンサーを決定した質問件数 (割合)	返礼コメントの入力件数 (割合)
生き方と恋愛，人間関係の悩み>恋愛相談，人間関係の悩み>恋愛相談	556,418	3.88	292,736(52.6%)	225,103(76.9%)
生き方と恋愛，人間関係の悩み>恋愛相談，人間関係の悩み	444,521	5.12	373,104(83.9%)	89,262(23.9%)
健康，美容とファッション>健康，病気，病院>病気，症状，ヘルスケア	442,925	2.26	289,514(65.4%)	148,814(51.4%)
Yahoo! JAPAN>Yahoo!知恵袋	380,200	4.39	342,368(90.1%)	98,773(28.9%)
ニュース，政治，国際情勢>政治，社会問題	367,043	4.05	238,487(65.0%)	94,771(39.7%)
スポーツ，アウトドア，車>スポーツ>プロ野球	294,603	3.56	213,134(72.4%)	93,537(43.9%)
インターネット，PC と家電>パソコン	292,657	2.45	256,441(87.6%)	49,530(19.3%)
その他>アダルト	263,043	3.16	179,431(68.2%)	78,773(43.9%)
Yahoo! JAPAN>Yahoo!オークション	255,615	5.76	235,458(92.1%)	30,183(12.8%)
インターネット，PC と家電>パソコン>Windows 全般	248,811	2.01	166,588(67.0%)	104,259(62.6%)

ず中カテゴリまたは小カテゴリまで設定をする必要がある．

質問者がベストアンサーを選択した場合にのみ，回答へのお礼や感想をまとめた返礼コメントを入力することができる．

全質問のうち，11,151,247 件 (約 69.3%) が質問者がベストアンサーを選択している．質問者がベストアンサーを選択し，かつ返礼コメントを入力した質問は 5,505,876 件であった．これは全質問の約 34.2%であり，質問者がベストアンサーを選択した質問の約 49.4%である．

表 1 に，質問数の多い中・小カテゴリごとの，そのカテゴリに属する質問のうち質問者がベストアンサーを選択した件数，その質問の中で実際に返礼コメントを入力した質問の件数，平均回答件数を示す．カテゴリによってベストアンサーを質問者が決定する割合に差が見られる．これは，明確な答えを求めるものや意見や共感を求めるものといった質問のタイプの違いによるものと考える．返礼コメントの有無についてもカテゴリによって差が見られる．

3.1.3　返礼コメント

質問者が気づきを得た回答を抽出するために，返礼コメントの内容を利用する．本稿では，返礼コメント中に特定の表現が 1 つ以上出現した場合に，質問者が気づきを得たと仮定する．質問者が気づきを得たと仮定する特定の表現の一覧は，筆者が決定した．表 2 に，気づきを得たと仮定する表現の一覧と，その表現が出現した回数を示す．抽出の結果，58,493 件の返礼コメントが該当した．気づきを得たと仮定する表現は 1 つの返礼コメントに複数出現する場合もあり，各表現の出現回数の合計とは異なる．抽出した返礼コメントの例を表 3 に示す．

3.2　問いかけ文の抽出

質問者が気づきを得たと仮定した質問への回答から，気づきを得ることになったと考えられる問いかけ文を取得する．ベストアンサーに選択された回答を正規表現を使い 1 文づつに分け，文末に疑問符が付いた文を抽出する．その結果，12,839 文が抽出された．抽出した問いかけ文の例を 3 に示す．

表 2　気づきを得たと仮定する返礼コメントの表現と出現件数

表現	件数
知らなかったです	3918
知りませんでした	31129
考えてなかったです	194
考えたことがありませんでした	63
考えたことがなかった	86
ハッとしました	455
目からウロコ	1152
目からうろこ	1017
驚きました	10529
びっくりしました	6378
ビックリしました	3441

表 3　抽出した返礼コメントの例

・意外な内容で驚きました（笑）．ありがとうございました．
・そんなのあるんですね！知りませんでした．ありがとうございます★
・お 2 人の意見にハッとしました．前向きに日々感謝の気持ちを
忘れずにがんばります．
・そんなもんなんですか！ちょっとびっくりしました …
・ありがとうございます．祝儀なんて頭にも浮かびませんでした ‥
目からうろこでした．本当お恥ずかしいです..
もう少し勉強したほうがいいですね．とても参考になりました !!

表 4　抽出した問いかけ文の例

・ホームページ見てみてはいかがでしょうか？
・いっその事電話番号を変えたら如何ですか？
・このサイトに分かりやすく書かれてますので
読んでみられてはどうでしょか？
・人を見た目で判断しているのはどっち？
・今 7ヶ月という事ですが，体重は平均と比較してどうですか？
・商品説明には発送方法や送料はどのように
記載されていたのでしょうか？

3.3　問いかけ文の分析

気づきを与える問いかけ文の問いかけ文について，形態素解

表 5　質問の大カテゴリ名，問いかけ分の数，出現した単語の種類数

大カテゴリ名	問いかけ文の数	単語の種類
Yahoo! JAPAN	849	1699
インターネット，PC と家電	1086	2450
エンターテイメントと趣味	2281	5940
コンピュータテクノロジー	59	290
スポーツ，アウトドア，車	880	3082
ニュース，政治，国際情勢	329	1512
ビジネス，経済とお金	354	1195
マナー，冠婚葬祭	308	1058
健康，美容とファッション	1214	2948
地域，旅行，お出かけ	698	2408
子育てと学校	894	2273
教養と学問，サイエンス	836	3117
暮らしと生活ガイド	1784	4321
生き方と恋愛，人間関係の悩み	886	1936
職業とキャリア	247	907
その他	213	855

析を行い分析を行なった．形態素解析には MeCab[5]を用いた．単語的な特徴を調べるために，元々の質問が属しているカテゴリの大カテゴリ（全 16 カテゴリ）ごとに問いかけ分を分類し，出現した単語とその回数を記録したリストを作成した．その後，各 16 カテゴリの問いかけ文に 1 回以上出現した単語を，問いかけ文に典型的な特徴語とした．なお，出現回数を記録する単語は名詞と動詞に限定し，品詞細分類で「数」「非自立」に該当するものは除外した．

次に，系列パターンマイニング手法を用いて，問いかけ文によく使われる言語パターンを抽出した．問いかけ文の 1 文を形態素を要素とする系列データと見なし，全問いかけ文から頻出する系列パターンを取得した．系列パターンマイニングには Modified PrefixSpan 法を用いた [15]．Modified PrefixSpan 法は，PrefixSpan 法 [16] の改良手法であり，パターン中のワイルドカードの有無が区別されたパターンを取得することが可能である．例えば，単語 “A” と単語 “B” が「AB」の順番で出現するパターンについて，PrefixSpan 法では単語 “A” と “B” の間に不特定の単語が挟まれる「A*B」というパターンと同じ扱いをしてしまうが，Modified PrefixSpan 法では区別される．本研究では出現頻度が 5 回以上の系列パターンを頻出系列パターンとし，以後の分析対象とした．

その後，抽出した頻出系列パターンから，単語頻度の分析で作成した，特徴単語を含むパターンを絞り込んだ．

4　分 析 結 果

問いかけ文を形態素解析した結果，取得できた問いかけ文の中から 17,726 種類の単語が抽出された．表 5 に，大カテゴリ名とその質問に対して投げかけられた問いかけ文の数，出現した単語数を示す．

問いかけ文を質問の属する大カテゴリごとに分け，各 16 カテゴリの問いかけ文に 1 回以上出現した単語を抽出した結果，30 種類の単語が抽出された．この 30 種類の単語を頻出単語とする．表 6 に抽出された頻出単語を示す．Modified PrefixSpan 法を用いて問いかけ文の単語パターンを分析した結果，19,795,334 件のパターンが抽出された．このパターンのうち，頻出単語が 2 種類以上存在した単語パターンを抽出した結果，9,385 件のパターンが抽出された．

表 6　抽出した頻出単語

する, 時, できる, 出る, ある, どこ, 使う, れる, 問題, ここ, なる, 者, 持つ, それ, やる, 見る, 何, 後, いかが, 知る, いう, 場合, せる, 的, 思う, 自分, 質問, 前, 関係, 自身

抽出した 9,385 件のパターンから，パターンの自由度に制約を付けるためにワイルドカードが 2 回以上出現するパターンを省いた結果，358 件のパターンを取得した．表 7 に，358 件のパターンの中から，いくつかのパターン例と実際の問いかけ文の一部を示す．

表 7　抽出したパターンと問いかけ文の例

パターン	問いかけ文の例
質問者*でしょうか	・質問者さんは最近の作品を聴かれたのでしょうか？ ・質問者様は（中略）目に負担のかかる生活をしてませんでしょうか？
質問者*ですか	・質問者さんは「自分がこうしたい」という理想からはずれるとテンパってしまうタイプですか？
質問者さんは*か	・質問者さんは神経質な方ではありませんか？
自分*いかがでしょうか	・ご自分で交換が難しければ，メーカーの交換サービスを使用してはいかがでしょうか？
ある*いかがでしょうか	・無料でカウンセリングなどあるので行ってみてはいかがでしょうか？

5　考　　察

5.1　頻出単語について

本稿では，質問の分野によって回答に使われる単語の傾向は変わるため，問いかけ文に使われる単語を分析するためには単語が全ての大カテゴリで出現しているかを判定する必要があると考えた．そのため，問いかけ文に出現した単語を質問カテゴリ毎に分け，全てのカテゴリに横断して出現した単語を抽出した．

しかし，大カテゴリによって質問数が大きく変わるため，問いかけ文が少ない大カテゴリが存在する．そのため，本稿で採ったアプローチで抽出した頻出単語以外にも，有用な問いかけ文で使われている単語が多く存在していることが考えられる．今後は，カテゴリのサイズを考慮して，問いかけ文で汎用的に用いられる特徴語を抽出する方法を検討する必要がある．

5 : https://taku910.github.io/mecab/

5.2 Modified PrefixSpan 法による単語パターンの取得

Modified PrefixSpan 法を用いて問いかけ文の単語パターンを取得した．その結果「質問者*でしょうか」「質問者さんは*か」といった，質問者の状態や質問の詳細を尋ねるパターンが取得された．他にも,「自分*いかがでしょうか」というパターンでは，問いかけを行なっている人が質問者の視点に立った問いかけを行なっているパターンが取得された．気づきを与える問いかけについて，単なる指摘ではなく，問題点や考えるべき点について，問いかけられた人自身が考えるような問いかけが気づきに重要な要素を持っていることが考えられる．しかし，単語パターンの取得について,「質問者*でしょうか」に対して「質問*でしょうか」といった似たパターンを複数取得しており，取得数に対して問いかけ文の傾向を多く取得することができなかった．今後は Modified PrefixSpan 法のパターン検出方法を変えることや，ストップワードを用いてパターン検出を行う必要があると考える．

5.3 気づきについての状態の判定

本稿では，問いかけられた人が気づきを得たかどうかを返礼コメントに特定の言語表現が含まれているかを手がかりに判断した．しかし，それら手がかり表現がなくとも，回答から質問者が気づきを得ることも十分に考えられる．返礼コメントの入力は質問者の任意であり，回答へのお礼を表す目的のみで書かれる場合も多い．質問者が気づきを得たか否かを確認するためには，本稿で採った方法とは別のアプローチも必要と考えられる．本稿では，気づきの有無による問いかけ文の比較を行うことができなかった．今後は，特定の表現に限らず返礼コメントの内容から気づきの有無を判断する手法が必要であると考える．

気づきの有無を判定できるようになれば，気づきを得ることになった問いかけ文と，気づきをもたらすことができなかった問いかけ文を用意し，出現した単語を特徴量として分類モデルを構築することで，どのような単語が気づきに有効であるか分析を行うことができると考える．また，単語以外も特徴量とすることで，気づきに必要な要素を分析することができると考える．

6 ま と め

本研究では，気づきを促す良い質問の生成・検索方法を検討するために，Q&A サイトの Yahoo!知恵袋をデータセットとし，質問者に気づきを与えた問いかけに関する分析と考察を行なった．質問者がベストアンサーを選択した際に返礼コメントを送ることができる機能から，返礼コメントの内容で質問者が気づきを得たと仮定し，そのベストアンサーから気づきを促すに至った問いかけ分を抽出した．系列パターンマイニングを行なった結果，気づきを与える問いかけ文の単語パターンを得ることができたが，単語の品詞による制約やパターン生成の条件などをさらに考慮する必要がある．

今後は，気づきを促す効果がなかった問いかけ文との比較を行うことで，気づきに有効な問いかけの特徴をより詳細に分析することが可能になると考える．気づきを促す効果が無かったことを測定するためには，返礼コメントの内容分析を特定表現の有無だけでなく，感情分析のような文章単位の分析を行う必要がある．

謝 辞

本研究は JSPS 科研費 JP16H02906，JP18KT0097，JP18H03243，JP18H03244，JP18H03494，および課題設定による先導的人文学・社会科学研究推進事業の助成を受けたものです．また，本研究では，国立情報学研究所の IDR データセット提供サービスによりヤフー株式会社から提供を受けた「Yahoo! 知恵袋データ（第 3 版）」を利用しました．ここに記して謝意を表します．

文 献

[1] 粟津恭一郎. **「良い質問」をする技術**. ダイヤモンド社, 2016.

[2] Lada A. Adamic, Jun Zhang, Eytan Bakshy, and Mark S. Ackerman. Knowledge sharing and yahoo answers: Everyone knows something. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, pages 665–674, 2008.

[3] 石川 大介, 栗山 和子, 酒井 哲也, 関 洋平, and 神門 典子. Q&A サイトにおけるベストアンサー推定の分析とその機械学習への応用. **情報知識学会誌**, 20(2):73–85, 2010.

[4] Baichuan Li, Tan Jin, Michael R. Lyu, Irwin King, and Barley Mak. Analyzing and predicting question quality in community question answering services. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12 Companion, pages 775–782, 2012.

[5] Sujith Ravi, Bo Pang, Vibhor Rastogi, and Ravi Kumar. Great question! question quality in community q&a. In *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[6] 日暮立, 小林隼人, 村尾一真, and 増山毅司. ランク学習による yahoo! 知恵袋の見出し生成. In **人工知能学会全国大会論文集 第 32 回全国大会** *(2018)*, pages 2E203–2E203. 一般社団法人人工知能学会, 2018.

[7] Hidenobu Kunichika, Tomoki Katayama, Tsukasa Hirashima, and Akira Takeuchi. Automated question generation methods for intelligent english learning systems and its evaluation. In *Proc. of ICCE*, 2004.

[8] Nasrin Mostafazadeh, Chris Brockett, Bill Dolan, Michel Galley, Jianfeng Gao, Georgios Spithourakis, and Lucy Vanderwende. Image-grounded conversations: Multimodal context for natural question and response generation. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Taipei, Taiwan, November 2017. Asian Federation of Natural Language Processing.

[9] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.

[10] Xinya Du, Junru Shao, and Claire Cardie. Learning to ask: Neural question generation for reading comprehension. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1342–1352, Vancouver, Canada, July 2017. Association for Computational Linguistics.

[11] Kyosuke Nishida, Itsumi Saito, Atsushi Otsuka, Hisako Asano, and Junji Tomita. Retrieve-and-read: Multi-task

learning of information retrieval and reading comprehension. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, pages 647–656, 2018.

[12] Scott Bateman, Jaime Teevan, and Ryen W. White. The search dashboard: How reflection and comparison impact search behavior. In *Proc. of CHI 2012*.

[13] Fumiaki Saito, Yoshiyuki Shoji, and Yusuke Yamamoto. Highlighting weasel sentences for promoting critical information seeking on the web. In *International Conference on Web Information Systems Engineering*, pages 424–440. Springer, 2019.

[14] Takehiro Yamamoto, Yusuke Yamamoto, and Sumio Fujita. Exploring People's Attitudes and Behaviors Toward Careful Information Seeking in Web Search. In *Proc. of CIKM 2018*.

[15] Hajime Kitakami, Tomoki Kanbara, Yasuma Mori, Susumu Kuroki, and Yukiko Yamazaki. Modified prefixspan method for motif discovery in sequence databases. In *Pacific Rim International Conference on Artificial Intelligence*, pages 482–491. Springer, 2002.

[16] Jiawei Han, Jian Pei, Behzad Mortazavi-Asl, Helen Pinto, Qiming Chen, Umeshwar Dayal, and Meichun Hsu. Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth. In *Proc of the 17th international conference on data engineering*, pages 215–224, 2001.