

ウェブ上のコラムページからのノウハウ質問応答事例の収集

Collecting Examples of Know-How Question Answering from Column Pages on the Web

前田 竜治^{*1}

Tatsuya Maeda

陳 騰揚^{*1}

Tengyang Chen

大川 遥平^{*1}

Yohei Ohkawa

宇津呂 武仁^{*1}

Takehito Utsuro

河田 容英^{*2}

Yasuhide Kawada

^{*1}筑波大学大学院システム情報工学研究科

Grad. Sch. Sys. & Inf. Eng. Univ. of Tsukuba

^{*2}(株) ログワークス

Logworks Co., Ltd.

In the area of factoid question answering, it is known that the state-of-the-art technology has achieved the accuracy comparable to human. In the area of non-factoid question answering, however, there exists a small number of data set for training question answering (i.e., machine comprehension) models. Considering such a situation, within the area of the non-factoid question answering, this paper studies how to develop a data set for training Japanese know-how question answering models.

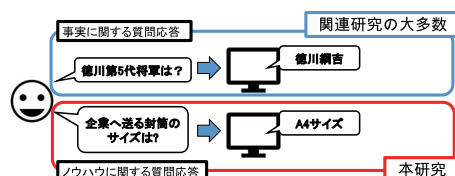
表 1: ノウハウサイト候補のドメインに対する評価基準

ノウハウ サイト	ドメインそのものがクエリ・フォーカスに関するノウハウを提示する個別ページへのリンクを一覧するページである	A 群
	ドメインそのものはクエリ・フォーカスに関するノウハウを提示する個別ページへのリンクを一覧するページではないが、ノウハウを提示する個別ページへのリンクを一覧するページが存在する	B 群
	ドメイントップページから容易に辿り着ける	C 群
	ドメイントップページから容易には辿り着けない	D 群
非ノウハウ サイト	クエリ・フォーカスを含む話題に関するノウハウを提示するページへのリンクを一覧するページ(ドメイントップページ, ドメイントップページから容易に辿り着ける, ドメイントップページから容易には辿り着けない)が存在する	E 群
	ノウハウを提示するページが存在しない	

1. はじめに

質問応答とは、ユーザからの質問に対して、文書中から回答部分を探し出すタスクである。特に、自然言語で記述された質問文とコンテキストを与え、質問に対する回答を探すタスクを読解タスクと呼ぶ。読解タスクにおいては、近年、ニューラルネットワークを用いた読解モデルが提案され、代表的な読解タスク用データセットの一つである SQuAD [Pranav16] に対しては、人間の回答精度を上回る結果が得られることが報告されている [Devlin18]。ここで、質問応答のタスクは、固有名詞や数量等といった事実に関する内容を回答対象とする質問応答、および、目的や理由、方法等の非事実を回答対象とする質問応答の二種類に大別される(図 1)。このうち、現状の研究動向としては、例えば Wikipedia に明示的に書かれるような事実を回答対象とする研究 [Yi15, Pranav16, Pranav18] が大半を占める。そのため、非事実の典型として、ものごとのやり方を答えるノウハウや、意見を問う質問に対して的確に回答することは容易ではない。

その一方で、インターネット上には、物事のやり方に関するノウハウが多数掲載するウェブサイト(本論文では、ノウハウサイト [李 18] と呼ぶ)が多数存在し、有用なノウハウを効率よく集めることを目的として、それらのノウハウサイトを効率よく収集する手法 [李 18] も提案されている。そこで本論文で



	Wikipedia記事に書かれている 歴史的事実や出来事に関する質問応答	左以外の質問応答 (非事実に関する質問応答)
英語	- SQuAD [Pranav et al. 16, 18]: 約10万事例規模 - 回答を必ず含む場合の精度: 87.4% (人間の精度 82.3% を超える) - 回答を含まない場合の精度: 84.2% (人間の精度 86.8% に迫る)	MS MARCO [Bajaj et al. 18]: 検索エンジンのクエリログ (事実+非事実[(例)料理の作り方])
日本語	- 早押しクイズ大会の質問 - 約2万5,000事例 - 精度: 52.4% (鈴木, 2018)	本研究 ノウハウに関する質問応答

図 1: 事実に関する質問応答・非事実に関する質問応答

は、ニューラル読解モデル(図 2)の中でも、特にノウハウに関する質問応答を対象とするニューラル読解モデル訓練のための質問応答事例データセットを作成することを目的とする。具体的には、上述のノウハウサイトにおいてノウハウが掲載されたウェブページを情報源として、ニューラル読解モデル訓練のための質問応答事例を作成する方式を提案する。特に、本論文では、ノウハウサイト中のウェブページの中でも、コラムページからのノウハウ質問応答事例作成が容易であることを示すとともに、トピックモデルを用いることにより、多様な話題のノウハウ質問応答事例を効率よく作成する方式を提案する。

2. ノウハウサイトの収集

本論文では、クエリ・フォーカスとして、「就活」、「結婚」、「花粉症」を対象として、[李 18] に従い、図 3 の手順に沿って、クエリ・フォーカス(図 3 の例では「就活」)に関するノウハウを多く掲載するノウハウサイトを収集する。

はじめに、Google 検索エンジンを用いて、クエリ・フォーカス「就活」についての検索エンジン・サジェストを網羅的に収集する。一文字列につき最大 10 個のサジェストが収集できること、および、一クエリ・フォーカスにつき約 100 通りの文字列(五十音、濁点、半濁点、開やく音を含む)を指定することから、一クエリ・フォーカスあたり最大約 1,000 語のサジェストを収集できる。ウェブページ収集においては、収集したサジェストを用いて、「クエリ・フォーカス AND サジェスト」を検索

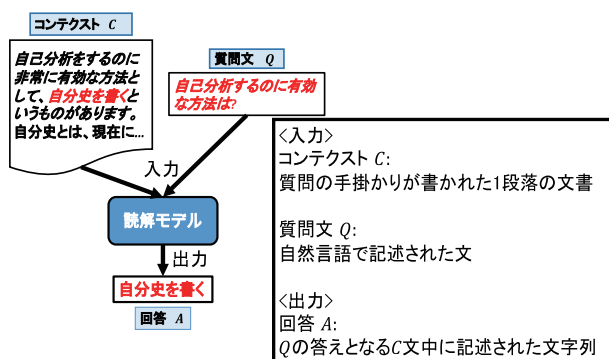


図 2: ニューラル読解モデルの概要

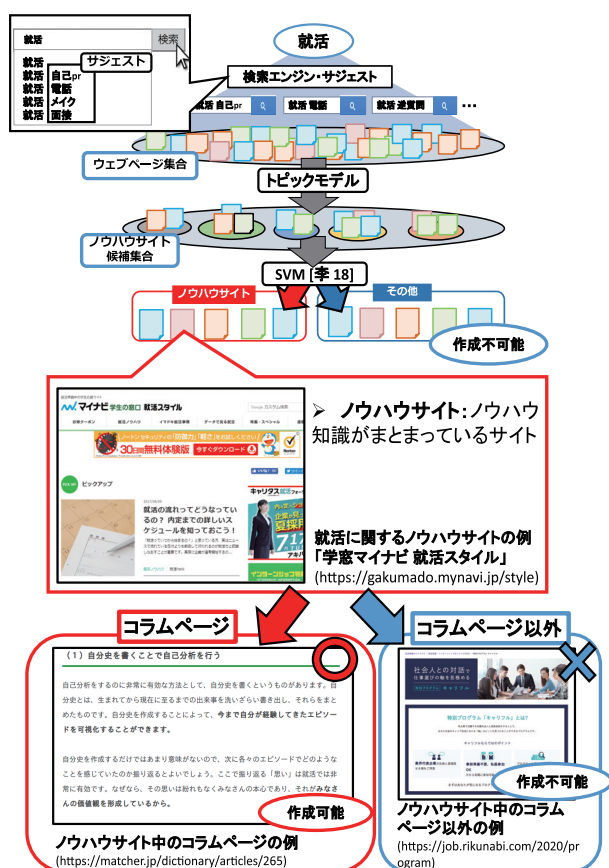


図 3: ノウハウ質問応答事例作成のためのウェブページ選定手順

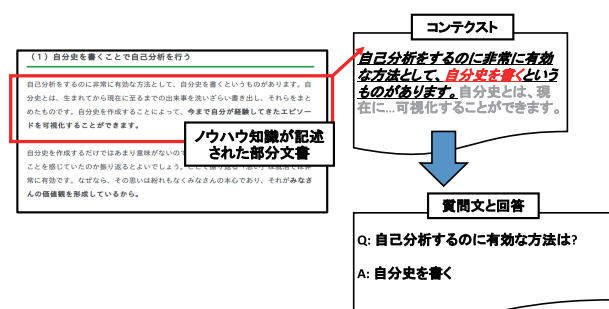


図 4: ノウハウ質問応答事例の作成例

表 2: 一ページから作成された質問応答事例数

質問応答事例数	1	2	3	4	5	計
ページ数 (%)	116 (35)	95 (28.7)	56 (16.9)	22 (6.6)	42 (12.7)	331 (100)

語とする AND 検索によって, Google Custom Search API^{*1} を用いて各検索結果の上位 10 件を収集し, ウェブページ集合 D^{est} を作成する. このとき, 異なるサジェストに対して検索されるウェブページが重複する際には重複を削除する^{*2}.

[李 18] および本論文では, 収集された各ウェブページのドメインに対して, 「どの程度の数のノウハウを掲載しているか」という観点での分類を行うために, 各ドメインを表 1 に示す A 群 ~ E 群に分類する. 表 1 の分類は, 当該ドメインにおいてノウハウを一覧するページが存在するか否か, および, そのノウハウ一覧ページがドメインのトップページから容易にアクセス可能か否か, の二つの観点に基づいている. 具体的には, A 群, B 群, C 群においては, ドメインのトップページからのアクセス容易性に差はあるものの, ノウハウを一覧するページが存在する点は共通している. 本論文の評価においては, A 群, B 群, C 群のドメインは, ノウハウを多く掲載する「ノウハウサイト」である. 一方, D 群においても, クエリ・フォーカス (図 3 の例では「就活」) そのものについてのノウハウには限定されないが, より広範囲の話題についてのノウハウを一覧するページが存在する. それらのページには, クエリ・フォーカスに関するノウハウの記述も一部含まれるため, 本論文の評価においては, D 群のドメインも, ノウハウを多く掲載する「ノウハウサイト」とみなす^{*3}. それに対して, 「ノウハウ知識を提示するページが存在しない」E 群のサイトにおいては, 本論文の評価においても, ノウハウ質問応答事例が作成できないため, 「ノウハウサイト」とはみなさない^{*4}.

なお, [李 18] においては, 上述の A 群, B 群, C 群, (および, D 群) のノウハウサイトを効率よく収集する手法として, トピックモデルを用いてウェブページ集合の話題集約を行った結果を利用する. 具体的には, まず, 収集されたウェブページ集合に対して, (本論文では, トピック数 $K = 50$ として) 潜在的ディリクレ分配法 (LDA: Latent Dirichlet Allocation) [Blei03] を適用することにより, ウェブページ集合の話題集約を行う. そして, 複数のトピックを横断してウェブページを持つウェブサイトに限定して, 表 1 の A 群, B 群, C 群, D 群のサイトを人手で収集する. この手順をふまえて, 本論文において, クエリ・フォーカス「就活」を対象とした場合には, ノウハウサイトの候補が 86 サイト得られ, そのうち, ノウハウサイトである A 群, B 群, C 群, D 群のサイト数は合計 57 サイトとなった. さらに, 本論文では, 以降の過程を効率的に行うため, A 群, B 群, C 群, D 群のサイトの中でも特に良質なサイトと思われるものを優先的に選定し, ノウハウ質問応答事例作成のための情報源とする. 具体的には, 上述の 57 サイトのうち, 当該サイトのウェブページが 10 トピック以上にまたがって存在するサイトに限定した. さらにその後, 以降の作業対象とするサイトを無作為に選定した結果, A 群, B 群, C 群のサイト合計 11 サイト, および, D 群のサイト 3 サイトの合計 14 サイト

*1 <https://developers.google.com/custom-search/>

*2 「就活」に対しては, 916 種類のサジェストに対して, 6,762 件のウェブページが収集された.

*3 [李 18] においては, 「ノウハウを掲載する個別ページは存在するものの, それらのページへのリンクを一覧するページが存在しない」ドメインを D 群と定義している. しかし, 精査の結果, 本論文で述べる D 群の定義がより実情に即していることが判明した.

*4 以上の議論の前提として, [李 18] では, 特定のジャンルのノウハウの一覧が参照できる可能性が低いサイトとして, i) 質問回答サイト, ii) ニュースサイト, iii) 通販サイト, iv) ブログホスト中のブログサイトのウェブページ, はあらかじめ分析対象から除外しており, 本論文でもこれに従う.

表 3: ノウハウ質問応答事例 (コンテキスト C ・質問 Q ・回答 A の組) の例

コンテキスト C	質問 Q	回答 A
履歴書に短所を書く時は前向きにまとめるようにします。「工夫して克服した」「直すように努力している」などと書けば悪いイメージの短所で好印象を与えることも可能です。自分の短所の中で努力すれば改善しそうなものを選ぶと書きやすいでしょう。	履歴書に短所を書く時のポイントは？	前向きにまとめる
結婚式の節約術の1つとして、ウエディングアイテムを持ち込むという方法があります。これは、結婚式に必要なものをネットやお店で自分で探して購入したり、自分自身で手作りしたりして当日使用するということです。	結婚式の費用を節約する方法は？	ウエディングアイテムを持ち込む
そのため花粉の季節は、室内の湿度を 50～55%ほどに保てるように加湿器を使用しましょう。	花粉の季節に保つべき室内の湿度の目安は？	50～55%

イトを以降の作業対象のサイト集合 T とした。

3. トピックモデル訓練時に未知のウェブページのトピック分類

次に、ノウハウ質問応答事例作成の情報源として用いるサイト集合 T の各サイトに対して、そのサイト中の全ウェブページを収集し、ウェブページ集合 $D^{\text{inf}}(T)$ を作成する。次に、ウェブページ集合 $D^{\text{inf}}(T)$ に対して、「検索エンジン・サジェスト」を用いて収集したウェブページ群によって訓練したトピックモデル^{*}のトピック群へ分類する方式 [Blei03] を適用する。そして、各ウェブページ $d(\in D^{\text{inf}}(T))$ に対して、トピックの確率分布 $P(z_n|d)$ ($n = 1, \dots, K$) において確率最大となるトピック z_n を割り当てることにより、トピック z_n が割り当てられたウェブページの集合を $D^{\text{inf}}(z_n, T)$ とする^{*5}。

$$D^{\text{inf}}(z_n, T) = \left\{ d \in D^{\text{inf}}(T) \mid z_n = \underset{z_u (u=1, \dots, K)}{\operatorname{argmax}} P(z_u|d) \right\}$$

ここで、実際には、サイト $t (\in T)$ ごとに、確率値 $P(z_n|d)$ の上位 5 ページのみを集めた部分集合 $D^{\text{inf}}(z_n, T, \leq 5)$ を以降の作業対象とする。

4. ノウハウ質問応答事例作成の情報源としての「コラムページ」

前節までの手順においてノウハウサイト中のウェブページ集合をトピックに分類した結果を選定した。次に、ノウハウ質問応答事例作成の情報源として有効なページ種類について分析を行った。その結果、ノウハウサイトにおけるウェブページの種類は、ノウハウが記述されたコラム、企業サイトにおける商用ページ、口コミや体験談のページに大別されるが、ノウハウ質問応答事例作成において有用なページ種類は「コラムページ」のみであることが判明した。そこで、前節において作成した $D^{\text{inf}}(z_n, T, \leq 5)$ について、要素となるウェブページを「コラムページ」のみに限定した部分集合 $D^{\text{inf}}(z_n, T, \leq 5)$ を作成し^{*6}、この集合中の全ウェブページを以降のノウハウ質問応答事例作成の情報源として用いる。

5. ノウハウ質問応答事例の作成手順

本節では、ノウハウ質問応答事例の作成手順 (図 4) について述べる。前節で作成したウェブページ集合 $D^{\text{inf}}(z_n, T, \leq 5)$ 中の全ページに対して、各ページ中の段落を対象として、以下で定義されるコンテキスト C 、質問 Q 、および、回答 A の組を手で作成する。

コンテキスト C : 一段落で構成される短い部分文書であり、画像などの非テキスト情報を含まない。

^{*5} 「就活」においては、ウェブページが一つ以上分類され、 $|D^{\text{inf}}(z_n, T)| > 0$ となるトピックは、全 50 トピック中 42 トピックであった。
^{*6} 「就活」において前節で $|D^{\text{inf}}(z_n, T)| > 0$ となった 42 トピック中で、コラムページを含み、 $|D^{\text{inf}}(z_n, T, \leq 5)| > 0$ となるトピックは 36 トピックであった。

質問 Q : 自然言語で記述された文。

回答 A : コンテキスト中に記述された任意の文字列。

具体的には、一ページあたり最大 5 組を上限として^{*7}、ページ中の各段落に対して、コンテキスト C 、質問 Q 、および、回答 A の組の作成可否を判定し、作成可能な場合には、当該段落をコンテキスト C とし、質問 Q および回答 A の組とともに記録する。一例として、ノウハウサイトの「コラムページ」^{*8} からノウハウ質問応答事例を作成する場合の模式図を図 4 に示す。この例では、コンテキスト C として

「自己分析をするのに ... 可視化することができます。」

という段落を抽出し、質問 Q として「自己分析に有効な方法は?」、 Q に対する回答 A として「自分史を書く」を抽出した。クエリ・フォーカスとして「就活」、「結婚」、「花粉症」を用いた場合のノウハウ質問応答事例各一例ずつを表 3 に示す^{*9}。表 4 に示すように、クエリ・フォーカス「就活」の場合には、ノウハウ質問応答事例作成対象となる「コラムページ」1,559 ページ中 331 ページからノウハウ質問応答事例 772 例が作成された。一方、クエリ・フォーカス「結婚」および「花粉症」を対象としては、各 50 例のノウハウ質問応答事例を作成した。

6. トピックごとのノウハウ質問応答事例作成効率の分析

次に、クエリ・フォーカス「就活」において、3. 節の「トピックモデル訓練時に未知のウェブページのトピック分類」によってウェブページが一つ以上分類され、 $|D^{\text{inf}}(z_n, T)| > 0$ となった 42 トピックに対して、 $D^{\text{inf}}(z_n, T, \leq 5)$ 中の全ページに対してノウハウ質問応答事例の作成を行い、ノウハウ質問応答事例が一つ以上作成できたページがどの程度含まれたか、を表す「ノウハウ質問応答事例作成効率」を評価する。そして、トピックごとに、この「事例作成効率」と各トピックの特徴との間の相関を分析する。

6.1 トピックごとのノウハウ質問応答事例作成効率

ウェブページ d に対して、5. 節の手順によりノウハウ質問応答事例が作成できたか否かを $qa(d) = 1$ (作成できた) または $qa(d) = 0$ (作成できなかった) によって表す。そして、 $D^{\text{inf}}(z_n, T, \leq 5)$ 中のウェブページのうち $qa(d) = 1$ (ノウハウ質問応答事例が作成できた) となるページ d のみから構成される部分集合

$$D^{\text{inf}}_{qa}(z_n, T, \leq 5) = \{ d \in D^{\text{inf}}(z_n, T, \leq 5) \mid qa(d) = 1 \}$$

^{*7} 一ページから作成された質問応答事例数の分布を表 2 に示す。
^{*8} <https://matcher.jp/dictionary/articles/12>
^{*9} 「就活」については、サイト集合 T に対して、3. 節の「トピックモデル訓練時に未知のウェブページのトピック分類」を経て収集したウェブページ集合 $D^{\text{inf}}(T)$ を情報源として作成したノウハウ質問応答事例を示す。一方、クエリ・フォーカス「結婚」および「花粉症」においては、本論文執筆段階では、3. 節の「トピックモデル訓練時に未知のウェブページのトピック分類」の手順が未適用であるので、本論文では、2. 節において収集したウェブページ集合 D^{est} に対してトピックモデルを適用した段階で、各トピックにおける「コラムページ」を情報源としてノウハウ質問応答事例を作成した結果を示す。

表 4: 「トピックごとのノウハウ質問応答事例作成効率」 r_{qa}^{inf} の分析 (クエリ・フォーカス「就活」)

r_{qa}^{inf}	トピック数	トピックの説明	質問応答事例が作成できたページ数	質問応答事例数
≥ 0.3	15	「最終面接」,「服装・髪型」,「履歴書」,「自己PR」,「OBOG 訪問」,「面接・長所短所」,「電話マナー」,「証明写真」,「面接・逆質問・質問」,「グループディスカッション」,「ビジネスメール」,「SPI」,「封筒」,「志望動機」,「就活メイク」	818 ページ中 260 ページ	634
$> 0, < 0.3$	11	「選考時期・身だしなみ」,「業種紹介」,「就活情勢」,「不安・性格」,「企業選び」,「交通費」,「休暇制度」,「(特定企業の)ES 対策」,「フリーター」,「就活用語」,「人事目線の面接・体験談」	529 ページ中 54 ページ	101
	8	複数の話題が混在	212 ページ中 17 ページ	37
計	34	上記全トピック	1,559 ページ中 331 ページ	772

表 5: ノウハウ質問応答事例が作成できたトピック数 (クエリ・フォーカスごと)

クエリ・フォーカス	ノウハウ質問応答事例が作成できたトピック数		トピックの説明
	合計	個別	
就活	34	26	—
		8	複数の話題が混在
結婚	32	18	「婚姻届」,「ご祝儀」,「フォトウェディング」,「身内・顔合わせ」,「披露宴・ゲスト」,「結婚式・費用」,「ゲストスピーチ」,「式場選び」,「服装・身だしなみ」,「招待状・返信マナー」,「予算・時期」,「内祝い・お返し」,「海外ウェディング」,「お車代」,「結納」,「招待状」,「結婚式挨拶」,「結婚式 BGM」
		14	複数の話題が混在
花粉症	36	21	「乳酸菌・花粉症」,「抗アレルギー薬」,「お茶」,「原因と対策」,「薬の選び方」,「ビタミン」,「かゆみ対策・生活習慣」,「子供の花粉症治療」,「薬の副作用」,「花粉症・漢方」,「目のケア」,「咳・くしゃみ」,「仕事場でのかゆみ対策」,「飲み物」,「耳のケア」,「鼻のケア」,「病院・治療」,「注射治療」,「食べ物」,「薬の種類」,「家での対策・掃除」
		15	複数の話題が混在

を作成する。そして、次式によって「トピックごとのノウハウ質問応答事例作成効率」を定義する。

$$r_{qa}^{inf}(z_n, T, “\leq 5”) = \frac{|D_{qa}^{inf}(z_n, T, “\leq 5”)|}{|D_c^{inf}(z_n, T, “\leq 5”)|}$$

6.2 分析結果

経験的には、 $r_{qa}^{inf} \geq 0.3$ となるか、 $0 < r_{qa}^{inf} < 0.3$ となるかによって、トピックごとの「事例作成効率」の良否の判断が分かれる。そこで、クエリ・フォーカス「就活」における分析対象の 36 トピック (実際に対象となるのは 34 トピック)*10 を、 $r_{qa}^{inf} \geq 0.3$ となる 15 トピック、および、 $0 < r_{qa}^{inf} < 0.3$ となる 19 トピックに二分し、各トピックの特徴を列挙した結果を表 4 に示す。「ノウハウ質問応答事例作成効率」が相対的に高い 15 トピックは、いずれも、「就活」のノウハウにおいて必要不可欠であり、かつ、中心的话题として位置付けられる内容で占められている。一方、「ノウハウ質問応答事例作成効率」が相対的に低い 19 トピックは、単一的话题で占められる 11 トピック、および、「複数の話題が混在」する 8 トピックに分けられる。このうち、単一的话题で占められる 11 トピックにおいても、 $r_{qa}^{inf} \geq 0.3$ となる 15 トピックと比較すると、ノウハウとしての重要度が相対的に低いトピックが観測される。その他、「結婚」、「花粉症」に対して「ノウハウ質問応答事例が作成できたトピック数」を表 5 に示す。

7. 関連研究

本論文に関連して、事実に関する読解タスク用データセットとして、英語を対象とした SQuAD [Pranav16, Pranav18]、および、日本語を対象とした解答可能性付き読解データセット [鈴木 18] が挙げられる。SQuAD は、英語版 Wikipedia 記事中の段落をコンテキストとして、クラウドソーシングにより質問・回答を付与した約 10 万件の質問応答事例から構成される。同様に、解答可能性付き読解データセット [鈴木 18] は、日本

語の早押しクイズ大会の約 12,000 件の質問文・回答に対して、回答の文字列が含まれる日本語版 Wikipedia 記事中の段落から選定したコンテキストを付与した質問応答事例から構成される。一方、[山本 19] においては、本論文において作成したノウハウ質問応答事例を訓練・評価事例として、ニューラル読解モデル*11 への適用を行い、ノウハウ質問応答において一定レベルの性能を達成している。

8. おわりに

本論文では、ノウハウサイト中のウェブページの中でも、コラムページからのノウハウ質問応答事例作成が容易であることを示すとともに、トピックモデルを用いることにより、多様な話題のノウハウ質問応答事例を効率よく作成する方式を提案し、クエリ・フォーカス「就活」を対象として、提案方式の有効性を検証した。今後の課題としては、クエリ・フォーカスとして、「結婚」、「花粉症」を含む多様な話題を対象として大規模なノウハウ質問応答事例を作成するとともに、[山本 19] におけるノウハウ質問応答の評価を大規模化して、本論文のアプローチの有効性を検証することが挙げられる。

参考文献

[Blei03] Blei, D. M., Ng, A. Y. and Jordan, M. I.: Latent Dirichlet Allocation, *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022 (2003).
[Devlin18] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *CoRR*, Vol. abs/1810.04805 (2018).
[李 18] 李佳奇, 趙辰, 林友超, 丁易, 川畑修人, 宇津呂武仁, 河田容英: トピックモデルおよび分類器学習を用いたノウハウサイトの同定, 第 10 回 DEIM フォーラム論文集 (2018).
[Pranav16] Pranav, R., Jian, Z., Konstantin, L. and Percy, L.: SQuAD: 100,000+ Questions for Machine Comprehension of Text, *Proc. EMNLP*, pp. 2383–2392 (2016).
[Pranav18] Pranav, R., Robin, J. and Percy, L.: Know What You Don't Know: Unanswerable Questions for SQuAD, *Proc. 56th ACL*, pp. 784–789 (2018).
[Seo17] Seo, M. J., Kembhavi, A., Farhadi, A. and Hajishirzi, H.: Bidirectional Attention Flow for Machine Comprehension, *Proc. 5th ICLR* (2017).
[鈴木 18] 鈴木正敏, 松田耕史, 岡崎直樹, 乾健太郎: 読解による解答可能性を付与した質問応答データセットの構築, 言語処理学会第 24 回年次大会論文集, pp. 702–705 (2018).
[山本 19] 山本航平, 前田竜治, 陳騰揚, 川畑修人, 大川遼平, 宇津呂武仁, 河田容英: ノウハウ質問応答におけるニューラル読解モデルの評価, 言語処理学会第 25 回年次大会論文集 (2019).
[Yil15] Yi, Y., Wen-tau, Y. and Christopher, M.: WikiQA: A Challenge Dataset for Open-Domain Question Answering, *Proc. EMNLP*, pp. 2013–2018 (2015).

*10 4. 節で $|D_c^{inf}(z_n, T, “\leq 5”)| > 0$ となった 36 トピックのうち、「複数の話題が混在」する 2 トピックにおいて、結果的には「ノウハウ質問応答事例」が作成できず、 $r_{qa}^{inf} = 0$ となった。

*11 事実に関する大規模読解用訓練事例データセットである SQuAD に対する適用実績のある BiDAF [Seo17] を採用した。