

## 就職活動支援のための対話型システム構築の一手法

### A Method of Making the Interactive System for Recruiting Support

小松 隆行\*

Takayuki Komatsu

#### 概要

本稿では、北海道科学大学の学生の就職活動において、その支援をするための対話型システム（チャットボット）を、クラウド AI を利用して効率よく構築するために考案した手法について述べる。システム構築にあたり、本学の Web サイト内のキャリアアシストサイトに掲載されている文書を利用して、クラウド AI への入力データ（FAQ 形式）を作成しプロトタイプシステムを作成しテストし問題点を確認した後、これを自然言語のテキストを分散表現として学習する word2vec を用いてキーワード抽出し、それらと典型的な FAQ での疑問文形式のテンプレートと組み合わせながら、表計算ソフト Excel の関数とその VBA プログラムによってクラウド AI への入力学習データである FAQ 文を生成し、先の問題を軽減するシステムを制作した。

#### 1. はじめに

近年、50 年来の AI のブレークスルーと称されるディープラーニング<sup>(1)(2)</sup>とその実用化に象徴されるように、人工知能の進歩は目覚ましく、機械学習<sup>(3)</sup>と組み合わされて多くのアプリケーションが開発されて利用されている。ここ数年はクラウド型 AI として大手 IT 企業が、自然言語処理などの多くの種類の人工知能サービスが提供されてきており、一部無料のものがあるなど、一般のユーザが手軽に利用でき、かつ時間をかけずに AI アプリと AI システムが構築できるようになってきた<sup>(4)(5)(6)</sup>。

一方、知的対話システムが、チャット（会話）形式での自動技術「チャットボット」として実用化されてきた<sup>(7)</sup>。日常会話の感覚でできる宅配業の再配達依頼の手続きや金融取引での顧客対応などで利用されてきている。また、チャットボット「りんな」<sup>(8)</sup>や、ヴァーチャル区民<sup>(9)</sup>などの仮想キャラクターへの応用もある。

また、大学生の就職活動においても、IT の利用や自動化、および人工知能利用が進んでいる。求人に対する応募書類の選考業務に AI が導入されてきている<sup>(10)(11)</sup>。また、AI による面接が実用化され、普及し始めている<sup>(12)(13)(14)</sup>。就職活動支援サイトでの AI 利用も始まっている<sup>(15)</sup>。学生の資質の変化、新しい就職活動形態や企業の新しい採用方式、学生の

適性の多様化、学科の特性などから、従来からの就職指導では対応できないケースも増加していると考えられる。したがって、大学においても AI などを利用した新しい就職活動支援システムなどを構築利用することで対応する必要があると考えられる。AI を使った学生毎の修学アドバイスサービス「学生の“自己成長支援システム”」を開始した大学もある<sup>(16)(17)</sup>。

本報告では、本学の学生の就職活動を支援するための対話型システム（チャットボット）を、クラウド AI を利用して効率的に構築するために考案した構築手法について述べる。

#### 2. システムの内容

本学では、本学学生のための就職活動支援サイト「キャリアアシストサイト」<sup>(18)</sup>を開設している。これは、就職指導のポータルサイトとして機能しており、就職活動の流れや各種手続き、企業説明会などのイベントや過去の就職実績などのデータを公開して、学内外から閲覧できるようになっている。学内専用のシステムとしては、キャリアデザインサポートシステム（CDSS）<sup>(19)</sup>があり、企業情報や年度ごとの求人情報が公開され閲覧できるようになっている。

今回制作するシステムは、このキャリアアシストサイトの情報を基にしたシステムである。自然言語

\* 北海道科学大学未来デザイン学部メディアデザイン学科

で入力可能なチャットボットシステムであり、キャリアアシストサイト利用学生ためのフロントエンド的役割を担い、チャット形式で簡潔な答えを返し、関連するキャリアアシストサイトのページへ誘導するリンクなども持つものである。今回構築したシステムにおいては、学外へも公開されている部分をシステム化することとした。

### 3. システム構築方法

基本となるシステムは、Microsoft 社製のクラウド AI である Microsoft Azure（以下で Azure と略記）<sup>(4)</sup>を用いる。質問に対する答えを返す部分の機能は Azure の QnA Maker<sup>(20)</sup>を使用する。これは、FAQ（よくある質問とその回答）のデータを基にして、入力した質問を理解して回答してくれるものである。QnA Maker に入力するデータは質問と回答がペアになっており、質問文の行頭は英字の「Q」、回答の行頭は「A」を識別子として入力しておく。以下では、学習用の入力データの中の質問文を「Q 文」と呼び、回答文を「A 文」と呼ぶことにする。これらの FAQ データを入力として学習させて、チャットボットシステムを生成できる。このシステムを動作すると、質問文を自然言語で入力でき、入力された質問文と形式的に同じ Q 文が学習データ内に無くても、質問の主旨が同じ Q 文に対応する適切な回答文 A 文を返してくれる。ただし、必ずしも適切な回答とはならないため、学習データをより適切なものにすることで、回答の適切さを向上させることができる。

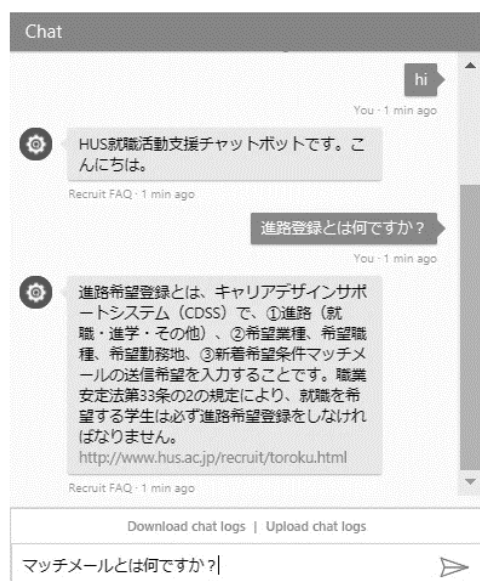


図 1 就活支援用の対話型システムの動作例

以下では、実際のシステム構築手法の手順に従って、3.1 節でプロトタイプ制作、3.2 節でプロトタイプのテスト、3.3 節で本システム生成のための FAQ データの作成、3.4 節でシステム生成とテストおよび再学習、という流れで述べる。

#### 3.1 プロトタイプの制作

まず、プロトタイプのシステムを作成する。学習データとして本学の Web サイト内の就職支援ページである「キャリアアシスト」に掲載されている文章情報を用いる。そのままのテキストデータでは、QnA Maker が学習できないので、フォーマットを変換する。いくつかのフォーマット形式が可能であるが、ここでは「Q」「A」の接頭文字による識別形式を用いる。キャリアアシストサイトのテキストデータを自動的に Q&A 形式に変換することはできないので、手作業で変換する。これは Microsoft Word (\*.docx 形式)で作成する。まずは、キャリアアシストサイトの内容を単語と文章をほぼそのまま以下の図 2 のような Q&A 形式データに変換した。Q&A のペアは、全部で 188 ペアになった。Q を接頭文字とする質問の Q 文は、特に意識せず自然言語で「〇〇は〇〇ですか?」「〇〇とは?」という形式で作成し、A を接頭文字とする回答の A 文は、キャリアアシストサイトの全ページの文章から抜き出して、ほぼそのまま使用した。このデータを学習データとして QnA Maker に入力し学習させ、プロトタイプの対話型システムを自動生成した。

- Q. 進路希望登録とは何ですか?。A. 進路希望登録とは、キャリアデザインサポートシステム (CDSS) で、①進路 (就職・進学・その他)、②希望業種、希望職種、希望勤務地、③新着希望条件マッチメールの送信希望を入力することです。職業安定法第 33 条の 2 の規定により、就職を希望する学生は必ず進路希望登録をしなければなりません。http://www.hus.ac.jp/recruit/toroku.html
- Q. 進路希望登録の方法 (やり方) は?。A. 進路希望登録は、キャリアデザインサポートシステム (CDSS) の「個人設定」画面から指示項目を入力して行います。http://www.hus.ac.jp/recruit/toroku.html
- Q. 就職相談とは何ですか?。A. 就職支援センターと各研究室担任は、皆さんの就職に関する種々の相談に応じてくれます。http://www.hus.ac.jp/recruit/toroku.html
- Q. 就職相談の内容はどんなものがありますか?。A. ①就職に不安がある、②どのような会社を選んでよいかわからない、③会社研究の方法が判らない、④進路決定に悩みがある、⑤就職試験の対応に困っている、などです。このような場合には自分自身で悩まず気軽に相談に来てください。就職課も利用してください。http://www.hus.ac.jp/recruit/toroku.html
- Q. キャリアデザインサポートシステム (CDSS) とは何ですか?。A. 就職活動支援のためのシステムです。キャリア支援機能として、PF 面談およびキャリア講座受講記録、業界研究・自己 PR・自己紹介等の各種シート、企業情報、求人情報の閲覧と支援機能があります。ログインは https://cdss.hus.ac.jp/users/sign\_in を、使用方法は https://cdss.hus.ac.jp/manual/login.pdf を参照ください。
- Q. 進路の進学とは何ですか?。A. 学部卒業後、別の学校組織に入学することです。例えば、大学院、専門学校などに進学希望の場合です。

図 2 学習データ用の FAQ 文の一部 (docx 形式)

### 3.2 プロトタイプのテスト

3.1 節で作成したシステムを動作させて、多数の質問文を自由に入力して、その回答が的確かどうかを確認した。その結果、以下のような問題点が確認できたので、以下にそれぞれの対応策を併せて示す。

- ① 学習データの QA の文章が丁寧な書き言葉であると、入力文章が話し言葉のような場合、不適切な回答になる場合がある。対応策としては、好ましい質問の形式例を事前に提示するなどが必要であると考えられる。
- ② 学習データ中のキーワードと同義語を使用すると不適切な回答になる場合がある。対応策として、学習データの中に同義語を盛り込む必要があると考えられる。例えば、「方法」に対応する「やり方」、「流れ」に対応する「プロセス」「手順」、「合同企業説明会」に対応する「合説」などである。
- ③ スマートフォン利用を意識したチャットボットシステムで表示するため、すなわちウィンドウサイズが小さいデバイスでの利用も前提とするため、長い回答文は好ましくない。対応策としては、A 文の数を必要最小限にし、できるだけ短く簡潔に作る。どうしても長くなる場合や、補足説明が必要である場合、外部参照を奨める場合は、外部 Web ページの URL のリンクなどを組み込む必要がある。
- ④ 学生が知りたい情報が網羅されていない。例えば、「キャリアアシストサイトとは?」、「〇〇〇〇を教えてください。」、「進路を決めたら」、「複数内定したら」、「内定したらどうする」、「会社説明会への参加方法は?」、「学内で説明会はありますか?」、「合説とは?」、「履歴書の値段は?」、「就職支援サイトとは?」、「就活対策の授業はあるのか?」などの質問に対して、適切な回答が返らない。また、最新の就職事情に対応していない。例えば、「オワハラ」「お祈りメール」「雑談面接」などの新しい用語に対応した FAQ も必要である。対応策として、不足している情報を A 文として補い、それを回答とするような Q 文を作成し学習データに組み込むことが考えられる。

### 3.3 本システム生成のための FAQ データの作成

3.1 と 3.2 で作成・テストしたプロトタイプのシステムにおいて不十分であったことを補強するこ

とと、より一般的な事項や最新の就職事情にも対応するように追加の学習データを用意することを考慮して、学習データ (FAQ 形式の文章) を作成することにする。以下では、その手順に従って説明する。

手順 1) 疑問文形式テンプレート作成:

インターネット上の Web サイトなどの FAQ の事例を参照して、疑問文の形式のテンプレートを作成する。具体的には、以下の図 3 のような形式である。

	A	B
1	とは?	
2	はありますか?	
3	について教えてください	
4	は可能ですか?	
5	条件を知りたいのですが	
6	してもらうことは可能ですか?	
7	はどのようになっていますか?	
8	が知りたいのですが?	
9	はしてもらえますか?	
10	をお願いできますか?	
11	までお願いできますか?	
12	作ってもらえますか?	
13	してくれませんか?	
14	をお願いしたいのですが	
15	:	
16	続く	

図 3 日本語の疑問文形式テンプレートの一部

このように Excel のワークシート形式にしておき、後述するように就職活動関連のキーワードや事項を問い合わせる疑問文を Excel 関数で自動生成するために用いる。

手順 2) 就職活動に関わる重要語句リストの作成

本学キャリアアシストサイトの記事の文章ファイルを入力として、Python のライブラリである gensim<sup>(22)</sup> の言語学習ニューラルネットワークである word2vec<sup>(23) (24)</sup> で文章のモデルを学習させる。word2vec はテキスト処理を行う 2 層のニューラルネットワークであり、テキストコーパスを入力として、コーパスにある単語の特徴ベクトルが出力される。この特徴ベクトルを用いて、単語間の類似度を数値で算出することができ、それを用いてある単語をキーワードとして、それに類似した単語を類似度の高い順に出力することが容易にできるものである。

この word2vec で文章のモデルを学習生成する。この際、文章中に登場する回数が  $n$  回以下の単語を除外するというパラメータを、 $n = 1, 3, 5$  と変えて 3 つのモデルを生成しておく。これらの学習後のモデルをそれぞれ使って、word2vec の most\_similar 命令により就職活動関連のキーワードを抽出する。具体的にはまず、キーワード「就職」





これらの中から、適切で重要と考えられるものをFAQのQ文の原型として選び出す。日本語の表現として不適當であったり、あるいは文法的に誤りがあったりする場合は、適宜手作業で整形し質問文として適切な正しい日本語の文章に修正する。

手順3) 英文での質問文の典型的な形式である5W3Hを考慮した質問を生成する。具体的には、Microsoft Excelにおいて、縦方向に質問のキーワードを配置し、横方向には質問の基本要素である5W3H(What/When/Where/Who/Why/How/How long/How much)を項目として配置する。表の各要素として組み合わせから生成した質問文をExcelのVBAマクロプログラムで自動生成し、対応する回答文を手入力した(図7、図8)。

	A	B	C	D	E	F	G
1							
2							
3		質問の					
4		キーワード	What	When	Where	Who	Why
5		休憩	休憩は、何をやるのですか?	休憩時間は、いつですか?	休憩の場所は、どこですか?	休憩に関わる人は、どんな人ですか?	休憩をやる理由は、何ですか?
6		応募	応募は、何をやるのですか?	応募時期は、いつですか?	応募の場所は、どこですか?	応募に関わる人は、どんな人ですか?	応募をやる理由は、何ですか?
7		内定	内定は、何をやるのですか?	内定時期は、いつですか?	内定の場所は、どこですか?	内定に関わる人は、どんな人ですか?	内定をやる理由は、何ですか?
8		会社	会社は、何をやるのですか?	会社時期は、いつですか?	会社の場所は、どこですか?	会社に関わる人は、どんな人ですか?	会社をやる理由は、何ですか?
9		卒業	卒業は、何をやるのですか?	卒業時期は、いつですか?	卒業の場所は、どこですか?	卒業に関わる人は、どんな人ですか?	卒業をやる理由は、何ですか?
10		進路	進路は、何をやるのですか?	進路時期は、いつですか?	進路の場所は、どこですか?	進路に関わる人は、どんな人ですか?	進路をやる理由は、何ですか?
11		採用	採用は、何をやるのですか?	採用時期は、いつですか?	採用の場所は、どこですか?	採用に関わる人は、どんな人ですか?	採用をやる理由は、何ですか?

図7 Excelによる5W3H疑問文原型の生成例

	A	B	C	D	E	F	G	H	I	J	K	L
1					1	2	3	4	5	6	7	8
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
32												
33												
34												
35												
36												
37												
38												
39												
40												
41												
42												
43												
44												
45												
46												
47												
48												
49												
50												
51												
52												
53												
54												
55												
56												
57												
58												
59												
60												
61												
62												
63												
64												
65												
66												
67												
68												
69												
70												
71												
72												
73												
74												
75												
76												
77												
78												
79												
80												
81												
82												
83												
84												
85												
86												
87												
88												
89												
90												
91												
92												
93												
94												
95												
96												
97												
98												
99												
100												

図8 5W3H疑問文原型によるQ文とA文の自動生成の例

回答文は、質問文に対応する答えであることが明確に伝わるように、質問文のキーワードに関わる部分を繰り返すようにしている。Q文に対してA文を作成する。A文は、既存の文章を流用したり、手入力で作成したりする。

最後に、就職活動における新しい単語や同義語を盛り込む。例えば、「合同企業説明会」という単語は、「合説(ごうせつ)」という短縮形で使用されることもある。本学の就職支援サイトの「キャリアアシストサイト」に掲載されている文章には、この「合説」という単語が使用されていないことから、入力データとして「合説」を盛り込む必要がある。具体的には、同義語を「合同企業説明会(合説)」のような形で同じQ文に組み込む。

手順4) 手順2から手順3と同様の工程を、インターネット百科事典 Wikipedia の就職カテゴリ<sup>(21)</sup>に掲載されている用語についても実行した。例えば、「お祈りメール」「オワハラ」「社会人基礎力」など、本学のキャリアアシストサイトには掲載されていない用語も多数含まれている。これらの用語のリンク先の各ページの記事をそれぞれダウンロードし、一つのファイルとして結合し、手順2から手順3と同様の工程で処理する。

### 3.4 システム生成とテストおよび再学習

3.3 節までで生成した FAQ 文のデータを学習データとして QnA Maker に学習させる。十分な学習データを用意できても、的確でない回答を返す場合がある。そのような場合をテストによって明らかにし、その都度適切な回答をするようにシステムを再学習させる。図 9 で示すように、左側の吹き出しの候補の中から、的確な回答を選択し選んで指示する。これにより適切な回答をシステムに教えたことになる。これである特定の質問に対して再学習させたことになる。図 9 では、的確な回答である内定辞退の手順についての回答を指示している。

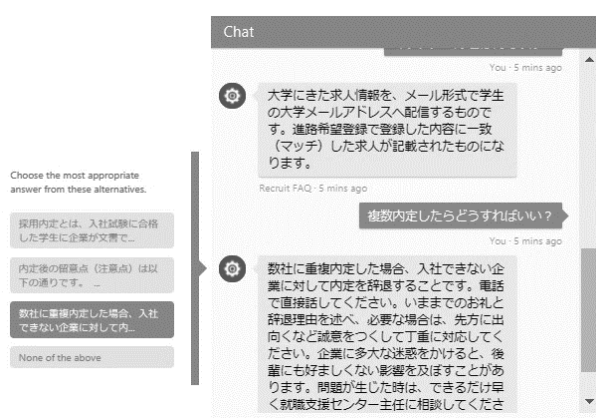


図 9 再学習が必要な例

## 4. まとめ

本報告では、大学 4 年生の就職活動を支援するための対話型システムを、クラウド AI である Microsoft Azure の QnA Maker を用いて構築する際に、word2vec, Excel 関数、および VBA プログラムを組み合わせた効率的な手法について説明した。しかしながら、実際に就職活動中の学生に使用させて検証を行っていないため、できるだけ多くの学生に使用してもらい評価をし、適切な回答をしない質問文を収集して、再学習や学習データの FAQ データの補強をする必要があると考えられる。

今後の課題としては、キャリア・カウンセラーの知識やノウハウを知識として追加し、再学習させること、あるいは内定をもらいにくい学生タイプ毎に異なる模擬面接を実施できる対話型の面接練習(個人と集団)システムの構築、大学としてのキャリア支援プログラムとの連携、なじみやすいインターフェースやアバターなどの導入などが考えられる。また、今回は Microsoft Azure の QnA Maker を使用したが、Azure の中の他の機能との連携によるシス

テム拡張、あるいは他社のクラウド AI (例えば IBM Watson や Amazon AWS など) を用いたシステム構築と性能比較が考えられる。

就職活動支援をはじめとして、大学での様々な学生サービスをチャットボット化し、如何にうまく活用していくかを検討することが重要となってきたと考えられる。学内向けだけではなく、大学の入試広報などの学外向けのサービスや、大学の業務や授業支援の AI 化は、チャットボット以外の他の AI 技術の進展に伴って、実現していかなければならない課題であるとも考えられる。

今後は、AI 面接官のようなシステムを教育目的で制作することや、面接指導のための AI 模擬面接システムなどの構築が意義あるものと考えられる。さらに、感情面でのケアができるようなメンタルサポートができる相談機能を加えていくことも有益であると考えられる。

## 参考文献

- (1) Yann LeCun, Yoshua Bengio, Geoffrey Hinton : Deep Learning, Nature 521, pp. 436-444, 2015.
- (2) 松尾豊 : 人工知能は人間を超えるか ディープラーニングの先にあるもの, KADOKAWA/中経出版, 2015.
- (3) K. P. Murphy : Machine Learning: A Probabilistic Perspective: The MIT Press, 2012.
- (4) Microsoft Azure : 2018 年 5 月 4 日, <https://azure>.
- (5) IBM Watson : 2018 年 5 月 4 日, <https://www.ibm.com/watson/>.
- (6) Google Cloud Platform : 2018 年 5 月 4 日, <https://cloud.google.com/>.
- (7) 日経コンピュータ「何でも答える「自動対話 AI」ー人手不足、解消の切り札ー: 日経コンピュータ, 5 月 11 日号, pp. 34-4, 12017.
- (8) りんな : 2018 年 5 月 4 日, <https://www.rinna.jp/>.
- (9) マイナビニュース: AI アカウント「渋谷みらい」くん誕生, 小学 1 年生のおませさん, 2018 年 5 月 4 日, <https://news.mynavi.jp/article/20171106-a230/>.
- (10) 日経コンピュータ「リクルート, IT 人材採用に AI」: 日経情報ストラテジー, 6 月号, pp. 26-27, 2017.

- (11) ソフトバンク : 「新卒採用選考における IBM Watson の活用について」, 2018 年 5 月 4 日, [https://www.softbank.jp/corp/group/sbm/news/press/2017/20170529\\_01/](https://www.softbank.jp/corp/group/sbm/news/press/2017/20170529_01/).
- (12) 日経ビジネスオンライン採用活動で活躍し始めた「AI 面接官」: <http://business.nikkeibp.co.jp/atcl/report/15/226265/060700137/>.
- (13) マイナビ 2018 模擬面接シミュレーター : 2018 年 5 月 4 日, [https://job.mynavi.jp/conts/2018/mensetsu/index\\_v.html](https://job.mynavi.jp/conts/2018/mensetsu/index_v.html).
- (14) SHaiN : 2018 年 5 月 4 日, <https://www.taleasse.co.jp/shain/>.
- (15) マイナビニュース : 「マイナビ, 人材領域に特化した AI を開発-学生向けサービスに初導入-」 2018 年 5 月 4 日, <https://news.mynavi.jp/article/20170718-a175/>.
- (16) IBM THINK WATSON 「金沢工業大学が取り組む コグニティブ・キャンパスとは」: 2018 年 5 月 4 日, <https://www.ibm.com/think/jp-ja/watson/cognitive-campus/>.
- (17) 金沢工業大学「国内初. 人工知能 IBM Watson 活用した学生一人ひとりに対応した修学支援サービス開始!」: 2018 年 5 月 4 日, [http://www.kanazawa-it.ac.jp/kitnews/2017/0728\\_watson.html](http://www.kanazawa-it.ac.jp/kitnews/2017/0728_watson.html).
- (18) 北海道科学大学 : キャリアアシストサイト (Career Assist Site), 2018 年 5 月 4 日, <http://www.hus.ac.jp/recruit/>.
- (19) 北海道科学大学 Web サイト : キャリアデザインサポートシステム (CDSS), 2018 年 5 月 4 日, [https://cdss.hus.ac.jp/users/sign\\_in](https://cdss.hus.ac.jp/users/sign_in).
- (20) Microsoft Azure : QnA Maker, 2018 年 5 月 4 日, <https://azure.microsoft.com/ja-jp/services/cognitive-services/qna-maker/>.
- (21) Wikipedia: Category:就職活動, 2018 年 5 月 4 日, <https://ja.wikipedia.org/wiki/Category:%E5%B0%B1%E8%81%B7%E6%B4%BB%E5%8B%95>.
- (22) genism topic modelling for humans : 2018 年 5 月 4 日, <https://radimrehurek.com/gensim/>.
- (23) Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean: Efficient Estimation of Word Representations in Vector Space, , 2018 年 5 月 4 日, <https://arxiv.org/abs/1301.3781>.
- (24) Xin Rong : word2vec Parameter Learning Explained, , 2018 年 5 月 4 日, <https://arxiv.org/abs/1411.2738>.