

令和4年度ベンチャー・ビジネス研究活動報告書

研究開発課題代表者 (所属・職名・氏名)	先端科学技術研究科・博士後期課程 東京社会人コース (知識科学)・奥田 聡
研究開発課題名 (研究開発期間)	外食企業向け需要予測システムの研究開発と機械学習応用システムの開発手法研究 (2020年度～2022年度)

1. 研究開発課題の概要

はじめに、VBL入居時は外食企業向け需要予測システムを対象としていたが、昨今の新型コロナウイルス感染症の影響により、当該産業の需要予測に対するニーズが低下した。

そのため本進捗報告では、別のプロジェクトである知能QAシステムを対象として報告を行う。

機械学習応用システムの開発手法は、近年「機械学習工学」として世界的に注目されている。本研究は、提案者が参加する「JST未来社会創造事業QAMLプロジェクト」の研究成果を活用したロードマップ活用型開発手法を、実際の機械学習応用システム開発に適用し、機械学習応用システム自体の開発と手法の開発を同時並行で行う。

新たに研究開発し、事業化する機械学習応用システムは、知能QAシステムである。これは会社内にある様々な業務マニュアル等のルールをコンピュータに読み込ませ、従業員がPC画面で質問をすると、コンピュータがAIを活用して、関連する回答を表示するシステムとなる。

知能QAシステムとは

- ・ ユーザーが入力した質問に対して、業務マニュアルやFAQリストの中から回答を提示するシステム。
- ・ 知能QAシステムを利用すると、ユーザーは膨大なマニュアルの中を探すことなく知りたい情報を手に入れられるようになる。
- ・ 旧来のChatbotや検索エンジンに対して「脳」の機能にあたる知能を提供する。

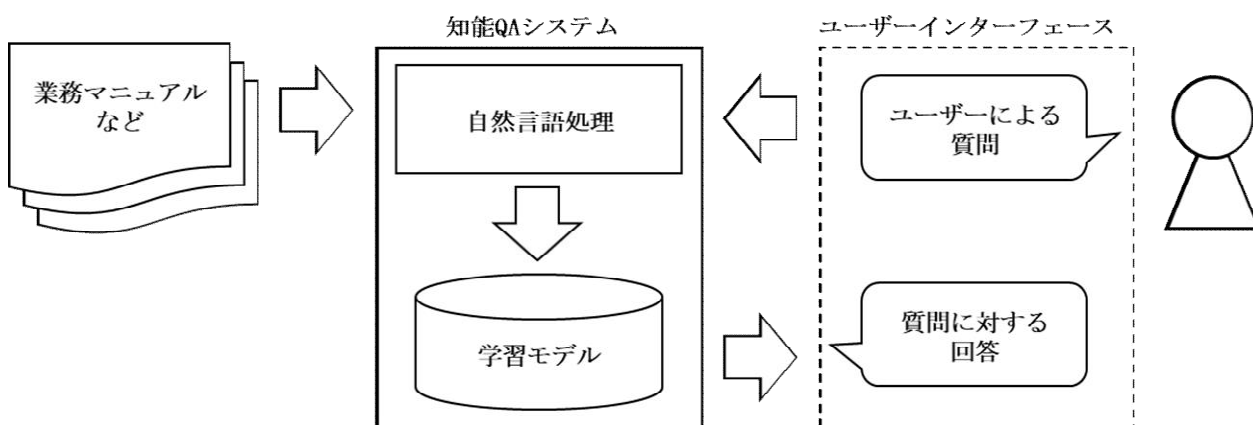


図1. 知能QAシステム概要

2. 研究成果（途中年度の場合は進捗状況）

■2020年度（1年目）の進捗

- ・R社とのディスカッションを元にアーキテクチャ選定・MVP開発 精度7割強をクリア（直近のテストでは85%をマーク）
- ・コアMVPについては開発終了（社会人学生とともにMVPをSさんの副テーマとして共著）
- ・QAシステムをもとにしたSaaS化を検証

■2021年度（2年目）の進捗

2020年度（1年目）終了時に立てたリサーチクエスチョン

1. PoC開発で実施した内容(85%)が他業界、業務でも同程度の精度が出せるか
2. 対象顧客と顧客価値はなにか？

1番の評価環境開発は6種類の環境で概ね8割の精度を出しており目標をクリアできた。
2番の対象顧客と顧客価値検証はデモ環境や構造化情報の取り扱いを求められることも多かったためその対策を実施した。

実施内容：

1. 複数のマニュアルを基にした精度評価

R社（健康関連サービス）、自治体、C社（メーカー）、F社（メーカー）、D社（通信）、z社（SaaS）でそれぞれの精度を確認

Accuracy (Z 76p, accuracy 89.47%, F241p, Accuracy84.21%, D 102Pp, Accuracy 95%)

概ね8割を超える結果を出しており、2020年度の目標としていた7割をクリアすることができている。

2. 対象顧客と顧客価値の検証

11件のインタビューを実施。一部ニーズヒアリングができたものの、まだアタック件数が不足しておりニーズの深堀が必要。ただし求められるコメント数が多いものは開発を実施した。

利用シーンと一言コメント

A. 業務システムのヘルプデスク	2 (M社、Z社)	更なる調査必要
B. 営業マンや店舗での利用	3 (K社、A社、N社)	可能性を感じる
C. コンタクトセンター	2 (D社、A社)	予算厳しく性能要件も高い
D. 社内問い合わせ	2 (R社、N社)	予算は総じて渋め

考察

ニーズが高くある程度の予算化が可能ではないかと思われる順に

・2番から営業マンや店舗で商品について問い合わせを受けて、その商品情報の検索や類するTI PS、業務マニュアルを検索対象にすることで商品知識の低い店員でも店舗対応ができるということにニーズがあるのではないかという仮説。ここでは社外の情報も含めて探し出せるのであれば精度を犠牲にすることも考えるか。

・1番から業務システムの開発が終了時にベンダーがマニュアルを作るがそのマニュアルを読まない社員が問い合わせをしてくるのでその問い合わせを本QAシステムでやるというユースケースも想定できる。これも第2の仮説といえる。

■2022年度（3年目）の進捗

2021年度の結果より高い精度を求められる利用シーンには合わないことから、対象とする利用シーンを新たに上場企業のIR資料や助成金情報の適切な検索・マッチングと定義した（表1）。これにより入手情報は社外に向くが、ここで外部情報を自動的に取得し構造化情報として保持する必要性があった。2022年度に新たに助成金13,879,000円を獲得。本原資を基にしてさらなる開発を進める。

ここでは上場企業4000社のデータをスクレイピングし情報構造化を行い、自然言語で問い合わせをされたテキストに対して、適切な情報を返すシステムを構築した。主要な機能を表2に示す。

表1：新たに設定したユーザー課題と提案

No.	利用シーン	顧客	課題	本提案
1	インサイドセールスにおけるBDR (Business Development Representative) 業務	SaaS企業等	セグメントを選定して顧客リストに対して営業がしたいが仮説ごとに営業リストを作るのが大変	脱炭素に積極的な企業は？など自然言語の検索で企業リストを出力
2	公共工事入札情報の社内適合可否マッチングシステム	建設会社	公共工事において営業マンが自治体の公告を都度確認しているが、手間がかかるし自社でできるかの判断も俗人的になりがちである。	自社が過去に行ってきた建設事例と近いスペックの公共工事を提案

表2：知能QAシステム主要機能

大項目	機能	要件
文書情報構造化	Webスクレイピング	収集したい分野のURL、スクレイピング深度、キーワードが設定可能であり、ターゲットとするファイルが取得可能であること
	ルールベースによる情報構造化	構造化データについて、あらかじめ設定した各項目と、該当するキーワードの抽出が可能であること
	自然言語処理による情報構造化	元データの自然言語と比較して、形態素解析、構文解析の構造化データの結果が網羅できていること
情報類似度判定	情報のアップロード機能	CSVなどの一覧データやPDFなどの非構造化データを持つ資料をアップロードできること
	類似度判定機能	アップロードした資料と構造化データの内容を比較し、類似度の高いレコードから任意の件数を抽出できること
	類似度一覧出力	ブラウザに出力できる形式で類似度のレコードを出力できること

研究期間全体の成果としてとして以下があげられる。

1. 3年間のプロジェクト期間中6名のJAIST学生が副テーマ、インターンシップ、アルバイトなど様々な形でかかわってくれ、就職活動にも一定の貢献ができたと考える
2. N氏 (Huynh研博士) が奥田の経営するP社に就職
3. 本プロジェクトはR社の支援により立ち上がったプロジェクトであったが、コロナにおける影響銘柄であったため1年で予算が凍結されてしまった。その後自然言語処理による検索システムとしての活路を探ったが、機能特性上8 - 9割の精度が限界であり、人の手では処理できない情報量をこなすという前提で回さないとビジネス化は難しいと判断 (2021年度)。その後クロール先の情報をネット上のIR資料及び助成金情報に切り替え、助成金を得て、開発を2022年7月から再開し、2023年3月末に無事開発が完了した。2023年4月から市場化に取り組む所存である。