

ソフトバンクが提供する対話型 FAQ システム : APTWARE

石川 誠一¹ 尾曲 博隆¹

¹ソフトバンク株式会社 ボイスプラットフォーム開発部

Abstract: 対話型 FAQ システム「APTWARE」は、ソフトバンクが開発した一問一答型の FAQ システムである。ユーザーは自然な会話調の文章でシステムに質問することによって必要な情報を引き出せることが可能となる。ここでは APTWARE の特徴、検索の仕組みについて説明する。

はじめに

ユーザーが製品などに疑問を持った場合、Web サイトやマニュアルを確認し、FAQ 検索を経て、それでも解決しない場合はコンタクトセンターに問い合わせる流れが多い。そのため FAQ システムでユーザーの自己解決率を高めることで、顧客満足度を向上させることができる。

しかしコールセンター白書のレポート[1]によると、約 80%の人が、コールセンターに電話する前にサポート用のサイトを参照・検索したが「見たが解決しなかった」と回答している(図 1)。ここから、FAQ システムに必要な情報が登録・更新されていない、もしくは登録されていても、どこに情報が載っているか分からず、探し出せないことがわかる。

またコンタクトセンターに電話した場合の「問題や疑問が解決するまでのプロセス」として約 73%の人が「1 回の電話で解決した」と回答している。ここからコンタクトセンターは解決のための「回答」を持っており、対話からユーザーの聞きたい内容を特定していることがわかる。

以上から APTWARE の設計にあたって、1) FAQ 作成担当者が、オペレータの持つ知識をいかに容易に登録・更新できるか、2) 使い方を気にすることなく、ユーザーを自然な対話で回答へ誘導できるかの 2 点を重視し、開発を進めた。

APTWARE とは

APTWARE は、ソフトバンクが提供する一問一答の FAQ システムである[2]。ユーザーは Web ブラウザなどから自然な文章を入力するだけで、回答を得る。APTWARE はソフトバンク社内の ERP システムにおける「コンシェルジュ」機能として、2014 年 10 月から本格的に利用されている(図 2)。

APTWARE は、目的があいまいな質問に対して聞き返す回答を返し、一問一答をしながら、まるで対話をしているかのように回答候補を絞込む仕組みを持つ。これによりユーザーの目的を明確にし、満足度の高い回答を表示することが可能となる。

また技術者でない担当者でも QA データのメンテナンスが可能とする観点から、機械学習などの「AI 的な手法」を利用せず「キーワード・マッチング」で回答を引き当てるアプローチをとる(図 3)。検索がヒットしない理由が一目でわかることや、事前の学習データを準備しなくても FAQ システムを立ち上げられるようにするためである。そこでキーワード自動抽出や自動テストツールなど、回答を引き当てるためのチューニング作業を支援する機能を充実させている。

ここでは APTWARE の主な特徴、検索の仕組みについて説明する。

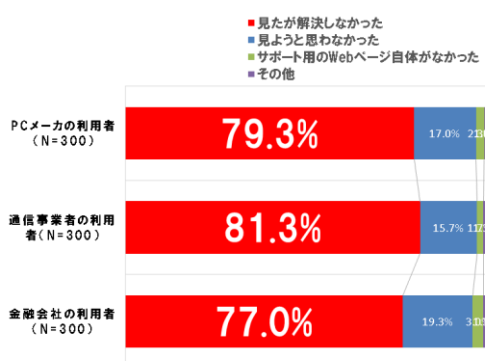


図 1 FAQ を事前に確認したか [1]

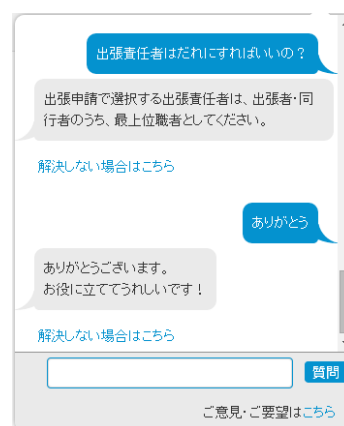


図 2 社内利用例

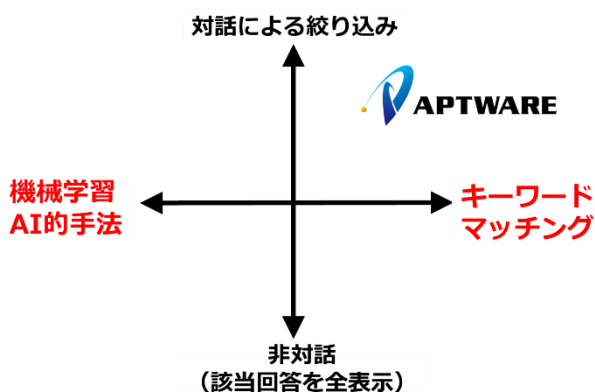


図 3 APTWARE のポジショニング

APTWARE の特徴

聞き返し処理と状態保持

例えば図 4 のように、「スマホについて教えて」という質問が来た場合、「スマホの何が知りたいですか?」と回答を返し、人間同士の自然な会話のように「聞き返し」を行う。この機能を使い、お客様の質問の真の意図を正確に判断することが可能とする。

また、APTWARE は直前の話題を記憶しておき「状態保持」を行うことができる。このように APTWARE は「スマホ」という話題を記憶できるため、「メールの操作について教えて」と問い合わせた場合には、「スマホ」の「メールの操作について」を回答することができる。

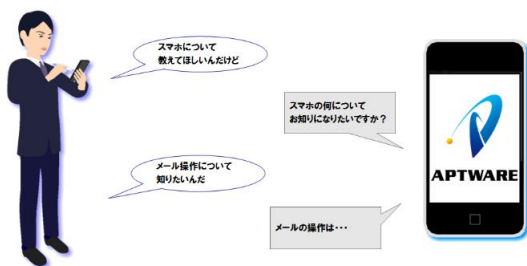


図 4 聞き返し機能のイメージ

大量 QA データに対する高い引当実績

APTWARE は、ソフトバンクの FAQ データ 4,045 件に対してテストを行った結果、他社を大きく上回る結果を記録した(図 5)。

116 件の質問を 3 つのシステムに対して同様にテストを行ったところ、APTWARE 以外のシステムは、

半分以上の質問が不正解だったことに対し、APTWARE はおよそ 85% の質問に対し正しい回答を得た*。このように APTWARE が持つ仕組みを利用することで、4,000 件もの大量 QA データに対しても、高い正答率を得ることができる。

*質問に対して模範回答を用意し、それらが検索上位 1 位になった割合を表す。また、このテストは 2013 年 3 月に実施した結果である。

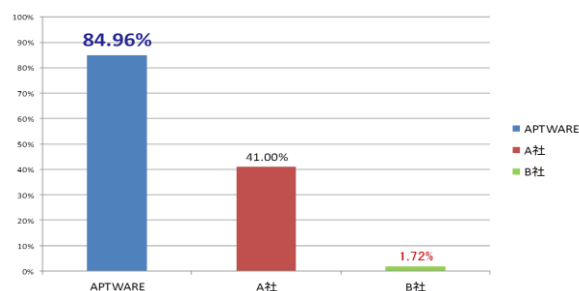


図 5 高い引当実績

APTWARE の機能

APTWARE の概要を図 6 に示す。FAQ 作成担当者は QA データおよび検索用キーワードを組み合わせて格納する。検索ユーザーが自然な文章で質問すると、APTWARE は質問に含まれる単語(もしくはフレーズ)を用いて回答を引き当てる。マルチテナント対応をしているため、社外向け FAQ や、社内の営業システム向け、経理システム向けなど複数の Web サイト用の FAQ を 1 つのシステムで対応できる。

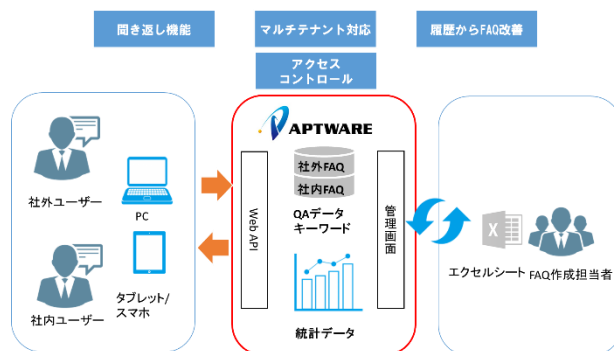


図 6 APTWARE 概要



図 7 機能概要

一般的に FAQ システムには、QA データの作成や管理を支援する機能、正答率の向上を支援する機能、および QA データの改善を支援する機能が必要となる(図 7)。

以下では、他システムと比較して特徴的な機能について説明する。

QA データ作成支援

1. データ作成 Excel シート

管理画面だけでなく、技術者ではない担当者でも普段から使い慣れた Excel を使って QA データやキーワードを登録することが可能である。既存の QA データをそのまま貼り付けてインポートすればそのまま使用できる。そのため QA データの追加/修正に IT 技術者が不要になり、より迅速な運用が可能になる。

2. キーワード自動抽出

用意した QA データから未登録のキーワード候補を抽出・提示する。QA データ作成担当者は提示された候補を選択するだけでキーワードの登録ができる。キーワードは、引き当てたい回答が一意に決まるような、回答を特徴づける単語(もしくはフレーズ)であることが望ましい。社内業務や商品の QA データを特徴づける言葉はその企業、業種特有の単語であり、名詞を合わせた複合名詞であることが多い[3]。

例えば、「光回線開通工事の工事代金について」という回答を登録する場合、キーワードとして「光」、「回線」、「工事」のような一般的な単体の単語ではなく「光回線開通工事」が望ましい。より回答を特徴づけるキーワードとして「光回線開通工事の工事代金」のようなフレーズを使うと、さらに特徴づけができ、一意に特定しやすくなる。

APTWARE では「回線」のような単体名詞の抽出だけでなく、「光回線開通工事」のような複合名詞や、「光回線開通工事の工事代金」、「SB 光を利用」のようなフレーズをキーワード候補として提示できる。

図 8 自動テストツール

3. 自動テストツール

QA データ作成担当者が事前に用意した複数の想定質問と模範回答の組み合わせを使い、APTWARE は正答率を一括で算出する。また正答率だけでなく、模範回答の順位、回答の引き当てに使われたキーワードなどを表示して、不正解となった理由を提示する(図 8)。

4. 類義語・表現のゆらぎ対応

QA データには社内用語やサービス・商品名が頻繁に記述されているが、それらを検索ユーザーが正しく質問してくれるとは限らない。そのため、質問された単語そのものでは検索がうまくいかないことが多い。

例) 質問での表現 : 請求書の精算

QA データでの表現 : 請求書払

APTWARE では、質問に含まれる単語を「請求書の精算=>請求書払」のように片方向に展開する「正規化辞書」を作成することで表現のゆらぎに対応する。

このような表記ゆれや類義語に関する課題に対して、類義語辞書を用いて対応を自動化している FAQ システムも存在する。これらは、検索がヒットしない「0件ヒット」対策のために行われることが多い。類義語のいずれかが含まれる文章を検索して検索漏れをなくとするアプローチである。

しかし類義語辞典に登録している単語は「一般的な単語」であることが多く、社内用語やサービス名などの「専門用語」をそもそも含まない。類義語辞典を利用することで、専門用語に対してではなく、それ以外の「一般的な用語」を展開してしまい、検索結果が大量に抽出される可能性がある。そのため社内利用などの「限定された領域」での FAQ システムにおいて、検索結果を上位に上げるチューニングには適さないと筆者らは考えている。

5. シミュレーション機能

正答率を高めるために「チューニング」が必要だが、これまでの FAQ システムはリアルタイムに成果率の変動を確認することが難しかった。APTWARE の「シミュレーション機能」は、チューニングをし

ながら、正答率を確認できる機能を提供する。精度を確認しながら作業ができるため、FAQ 作成担当者は手戻りなく、「自信」を持って作業を進めることが可能である(図 9)。



図 9 シミュレーション機能

正答率向上支援

6. 聞き返しと状態保持

回答が特定できない曖昧な質問に対しては「何が知りたいですか？ 次の中から選んでください」のように聞き返す回答を返すことで APTWARE では対話を実現する。QA データをカテゴリズ(分類分け)して、質問に含まれるキーワードから条件分岐しながら、検索対象となる回答文を絞り込んでいく。この作業をすべて Excel ツールだけで実現できる。

7. 質問のサジェスト

ユーザーは、自身が欲しい回答を得るために、どのような質問をすればいいのかわからないことがある。そこで最近の多くの検索システムでは、質問をリアルタイムに補完してくれる機能を持つ(図 10)。

APTWARE ではサジェスト用のデータを用意することなく、登録する QA データから質問の候補を提示する。APTWARE のサジェスト機能は次の特徴を持つ。

- ・ ひらがな、カタカナ、漢字、ローマ字など入力される文字種に関わらず、候補を提示
- ・ 入力した文字を、質問文の中でハイライト
- ・ 質問の文字列のみだけでなく、その回答に紐づくキーワードも考慮して、候補を提示
- ・ 先頭一致だけでなく、部分一致で候補を提示

なお、サジェスト機能の実装はアティリカ株式会社が開発したサジェストエンジン Akahai を用いた [4]。

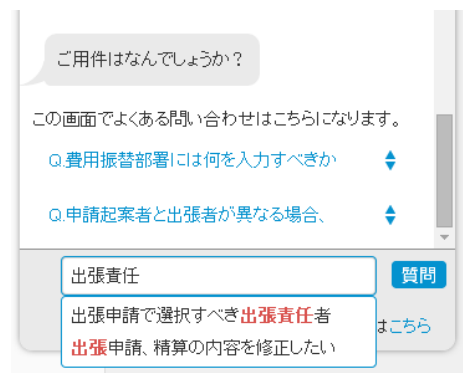


図 10 サジェスト機能

QA データ改善支援

8. レポート・統計情報

ユーザーのよく質問されるキーワードのランキングや、質問の履歴など利用状況をグラフで把握することができる。またそのデータを CSV ファイルで出力も可能である。公開中の回答が足りているか、よく質問されるキーワードなどから QA データの改善に反映させることができる(図 11)。

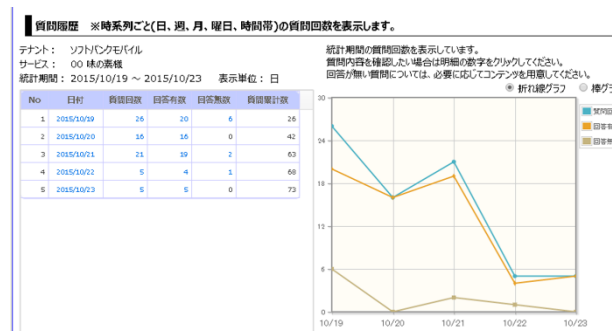


図 11 統計情報の一例

APTWARE の検索技術

APTWARE は検索の精度を向上させるために 2 種類の検索技術を使う。これらの実装には、検索のコアエンジンにオープンソースソフトウェアの Apache Solr (以下, Solr) を用いている。Solr は、実績のある全文検索ライブラリ Apache Lucene を用いた転置索引 (インデックス) 方式の全文検索エンジンである[5]。また形態素解析器としてアティリカ株式会社が開発した Kuromoji[6]を用いる。

図 12 に APTWARE の検索の流れを示す。大きく 3 つの部分で構成されており、それぞれ次の機能を持つ。

1. 受付・前処理部

入力された質問を形態素解析し、同義語や表記の揺れに対して、キーワードの正規化を行う。また不要語(ストップワード)の削除など、検索前処理を行う。

2. 検索部

登録済みの QA データに対して全文検索を行う「全文検索」と、QA データに紐づくキーワードを検索する「キーワード検索」を実行する。それぞれの検索手法を用いて、適合度(スコア)を算出して検索結果を出力する。なお、全文検索では Solr が算出したスコアをそのまま用いている。一方、キーワード検索では Solr のロジックの一部を改良して、APTWARE 独自のスコア計算を実装している。

3. 評価・応答部

APTWARE 独自のスコア正規化手法を用いて、それぞれの検索結果のスコアを正規化し、順序付けする。

今後の展開

APTWARE は、社内 FAQ や商品サービスサイトでの活用のみならず、対話のシナリオを工夫することによって、デジタルサイネージやロボットなどの様々なクライアント・インタフェースでも活用できると考えている。近年の訪日外国人向けのインバウンド・ソリューションの注目の高さを受け、多言語対応にも取り組んでいくことを予定している。2015 年 11 月に米語対応 ベータ版をリリース予定であり、今後は北京語、韓国語の対応も検討していく。

参考文献

- [1] コールセンター白書 2013 コールセンター利用者調査 コンピューターテレフォニー編集部・編
- [2] <https://rizbell.jp/>
- [3] 中川裕志, 森辰則, 湯本紘彰: "出現頻度と連接頻度に基づく専門用語抽出", 自然言語処理, Vol.10 No.1, pp. 27-45, (2003)
- [4] <http://www.atilika.com/en/products/akahai.html>
- [5] <http://lucene.apache.org/solr/>
- [6] <http://www.atilika.com/en/products/kuromoji.html>

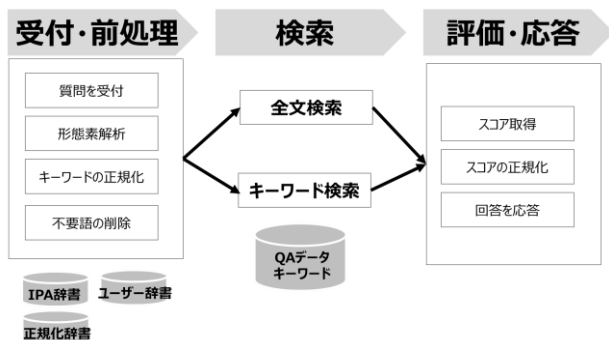


図 12 APTWARE 検索の概要