

# 大量のコールセンター対応記録に対する WEBベース検索・分析支援システムの開発

中野 治郎<sup>†</sup> 大園 忠親<sup>‡</sup> 清水 泰<sup>‡</sup> 服部 和彦<sup>†</sup> 新谷 虎松<sup>‡</sup>  
 (株) INAX 基礎研究所<sup>†</sup>  
 名古屋工業大学 知能情報システム学科<sup>‡</sup>

## 1. はじめに

コールセンターでは、主に商品・サービスに関するユーザーからの問い合わせに対応する役割を果たしている。コールセンターにおける迅速で誠実な対応や、そこに寄せられたさまざまな意見や要望を反映したより良い商品・サービスを提供していくことは、CS (Customer Satisfaction: 顧客満足) 向上の観点から企業にとって大変重要である。一方コールセンターに日々蓄積される質疑応答データはテキスト情報が中心で、且つ膨大であるため、そのデータを検索あるいは分析等に再利用するのは大変困難であった。

今回コールセンターにおける大量の質疑応答データを共有化し各部署で効果的に利用するため、WEB ベースでの検索・分析支援システムの構築を試みたので、その効果と問題点について報告する。

## 2. これまでの問題点と本システムの開発方針

当社のこれまでのシステムでは、検索する際に対応記録 DB に直接 SQL を発行してアクセスしていたため、レコード数の増加と共に検索速度が著しく低下していた。また分析用のツールが無かったため、分析するためには一度キーワードで検索し、得られた結果を手作業で集計しなければならなかった。そのため時間と手間が掛かるのと同時に、検索ノウハウを持った特定の人以外は扱づらいという欠点があった。

上記問題点を解決するため、以下の点を考慮したシステムを開発した。

- 1) WEB ベースで社内どこからでも利用できる
- 2) 検索速度が速く、且つ漏れを少なくする
- 3) 分析ツールは高速で使いやすいものにする

Development of a Web-based Retriever and Analyzer for a Large-scale Text Database for Question and Answer Records in a Call Center

Jiro Nakano<sup>†</sup>, Tadachika Ozono<sup>‡</sup>, Yasushi Shimizu<sup>†</sup>, Kazuhiko Hattori<sup>†</sup> and Toramatsu Shintani<sup>‡</sup>

<sup>†</sup> Basic Research Center, INAX Corp.

<sup>‡</sup> Dept. of Intelligence and Computer Science, Nagoya Institute of Technology

## 3. システム構成と検索・分析フロー

本システムは対応記録などのデータベース群、検索用 INDEX およびアプリケーションサーバから構成され、クライアントは WEB ブラウザを介して本システムを利用できる。全文検索と分析を高速に行うために、予め対応記録 DB から検索用の Suffix Array 型インデックス[1]と分析用の係り受け解析結果 DB を作成しておく。またそれとは別に類語検索用の辞書 DB やログ DB も用意しておく(図1)。

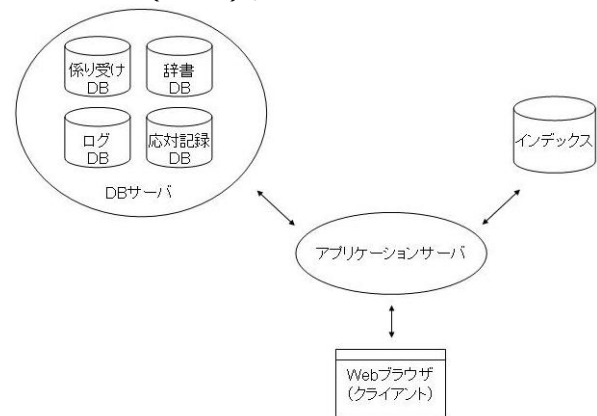


図1 システム構成

検索フローについては、クライアントから検索要求があるとまずインデックスを検索し、キーワードが含まれるレコード番号を抽出した後、対応記録 DB から目的のレコードを取得することで高速な検索を可能にした。また類語検索を行う場合は、辞書 DB を用いて同意異表記や書き間違い等を吸収できる仕組みにし、検索精度の向上を計った(図2)。

分析フローについては、まずキーワードや各種条件で調査したいレコードを絞り込んだ後、係り受け DB から解析結果を抽出し、集計する。表示部分については、必要に応じて結果を CSV ファイル形式で出力し、市販のグラフ描画ソフト等の機能を用いた(図3)。

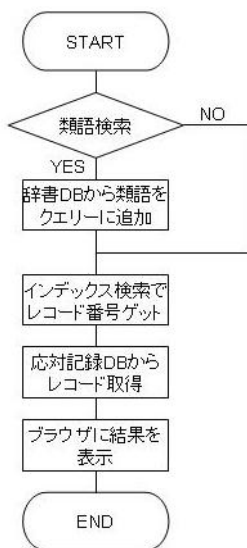


図2 検索フロー



図3 分析フロー

また分析結果の表示については、品種や品番など比較したい任意の項目をX軸に、係り受けの種類をY軸に、件数をZ軸に取ることで、データの傾向を視覚的に把握することができるようになった。分析結果の一例を図4に示す。

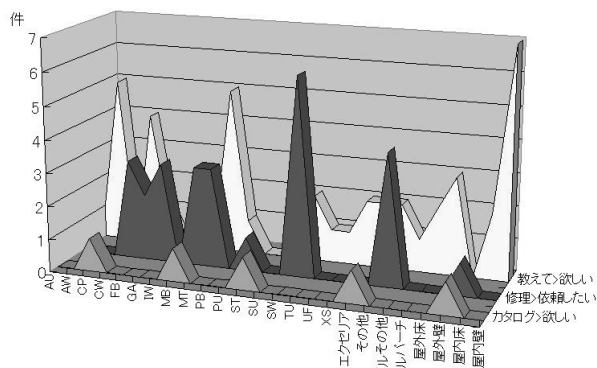


図4 分析結果例

#### 4. システムの評価と考察

##### [検索について]

検索速度については、Suffix Array 型インデックスを用いたことで検索漏れの無い高速な全文検索が可能となり、大幅な改善が見られた(表1)。また運用上、データ更新に伴うインデックスの更新は不可欠である。本システムではインデックス更新をデータの増加に対して漸増的に行うことで、インデックスの更新にかかる時間を短縮した。現在では1日1回夜間にインデックスの更新を行い、元DBとの整合性を取っている。

表1 検索速度(母数80万レコード)

| 検索キーワード | ヒット数 | 旧システム(秒) | 新システム(秒) |
|---------|------|----------|----------|
| サティス    | 8600 | 150      | 2        |
| サティス 寸法 | 236  | 165      | 2        |

検索性能については、全角/半角対応や類語検索機能を用いたことで向上した。企業であるため検索キーワードとして品番や製品名が使用される場合が多く、今後は品番 製品名に対応する辞書を整備することによって、より一層の検索精度の向上が期待される。

##### [分析について]

データの分析には、フリーソフトであるCaboChaの係り受け解析機能を用いて行った[2][3]。係り受け解析には通常かなりの時間がかかるが、解析結果DBを予め用意しておくことで、ユーザーは条件設定後すぐに分析結果が得られるようになり、分析にかかる時間を大幅に短縮することができた。

ただし現状では、無意味なレコードやデータ入力者毎の表現の違いなどの影響を取り除けない。これらの点に関しては自然言語処理によってある程度は解決可能であるが、実行時間とのトレードオフが問題となってくるため、適用に関してはさらなる研究が必要である。

#### 5. まとめ

コールセンターに蓄積された大量の質疑応答データに対して、「高速な検索機能」と「分析機能」をシームレスに統合したシステムを構築することにより、誰にでも使いやすい知識発見ツールが実現できた。

#### 参考文献

- [1] 山下 達雄：用語解説「Suffix Array」, 人工知能学会誌, Vol.15, No.6, pp.1142(2000)
- [2] 藤巻 伸洋, 大園 忠親, 新谷 虎松：Q & Aの係り受け解析に基づく顧客知識発見について, SICE SI2002 講演論文集(1), pp.135-136(2002)
- [3] T. Kudoh and Y. Matsumoto: Japanese Dependency Structure Analysis Based on Support Vector Machines, In Empirical Methods in Natural Language Processing and Very Large Corpora, V25, pp.18(2000)