

テキストマイニングを成功させるポイント

テキストとして蓄積された顧客の生の声を商品開発・サービス改善に活かそうという意識の高まりとともに、テキストマイニングツールの導入が急速に進んでいる。しかし、一方で新たな課題も浮き彫りになってきた。本稿では、野村総合研究所（以後、NRI）の経験に基づき、テキストマイニングを有効活用するポイントについて考察する。

増えるテキストマイニングの導入

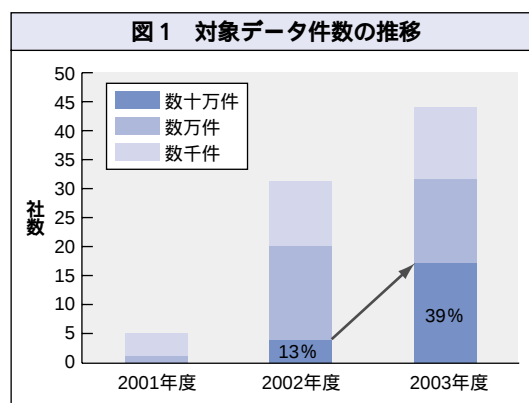
商品開発やサービスの改善に活かすことを目的に、顧客の生の声（日本語文章のテキストデータ）を分析するテキストマイニングは、ここ数年で、実験的な導入の段階からより具体的な成果が求められる段階に入っている。NRIが開発したテキストマイニングツール「TRUE TELLER（トゥルーテラー）」も、2003年度の1年間だけで約50社（累積約90社）の企業に導入されている。なかでも顧客からの問い合わせやクレームなどが蓄積されるコールセンターでの利用が6割を占めている。

図1は、TRUE TELLERが導入された企業での分析対象データ件数の推移を示したものである。2001年度に年間数千件から多くても数万件であったものが、2003年度には数十万件に達している。また、対象データが100万件を超える企業も増えてきている。これは、短期間に大量のアンケート回答が得られるインターネットリサーチの普及や、コールセンターでの膨大な問い合わせログを分析対象とする企業が増えてきていることなどが理由である。たとえば、ある大手通信企業のコールセンターでは、分析対象となる問い合わせな

どのデータ件数は月に100万件にも達する。そこから、顧客の年齢や地域、利用サービス別に問い合わせの特徴や時系列変化をとらえることになる。この場合には、分析は夜間バッチ処理で行い、翌日には結果が参照できるといった、大量データに耐えられるシステム構成が必要となる。

進化するテキストマイニングツールの機能

テキストマイニングでは、構文解析などの自然言語処理技術をベースにして、日本語の文章を単語ごと、あるいは主語と述語のフレーズに切り出し、そこから重要なキーワードを抽出する。また、疑問、要望、困難といった文脈もとらえることができるようになってきている。





初期のテキストマイニングツールは、文書ファイルの検索を中心とした検索系と、マーケティングや商品開発を目的とした分析系があり、用途に合わせて使い分けられていた。近年は、分析した結果や顧客の声の生データを全社員で検索、共有する仕組みが求められ、徐々に分析と検索の統合的な環境を提供するテキストマイニングツールに進化している。

こうしてツールの機能が高度になっていくにつれて、テキストマイニングの成功の障壁となる新たな課題が浮き彫りになってきた。

課題となる入力データの質の向上

「ゴミのデータからはゴミしか発見できない」。これは、データマイニングにおける多くの企業での失敗経験から広く認知されている言葉である。テキストマイニングについても同じことが言える。コールセンターでは、長い間、1コールの対応時間をいかに短くするかということが最優先とされ、電話を受けたオペレーターが入力する内容は、たとえば「8月3日、スイッチが故障、折り返しTEL」というように、日付や連絡事項のみといった管理的な情報になっている場合が多かった。これでは、どのような状況で不具合が起きたのか、といった詳しいことはわからない。また、句読点がなかったり、主語と述語の関係などがはっきりしない文章は、テキストマイニングの構文解析処理では正しい解析の障害となる。このようなことから、顧客の声を有

効に活用するために、オペレーターの教育によって入力データの質を向上させる試みが始まっている。

TRUE TELLERでも、「入力データ診断機能」を搭載し、2つの視点からデータの質を定量的に測定できるようになっている。その1つは文章の書き方に関して、句読点の欠如、並列表記などを検出・測定する。もう1つは文章の内容について、主語と述語がきちんと記述されているか、状況説明の文節が入っているかなど、情報量の豊富さを測定する。この機能を活用することで、オペレーターが入力するデータの質のばらつきを分析し、入力方法の改善に活かすことができる。こうした機能は、これまでのテキストマイニングツールにはない、より実践的な目的のなかから生まれた新しい機能と言える。

情報共有と全社横断的な組織が必要

テキストマイニングツールは、データを入れれば自動的に課題の解決策が得られるというようなものではない。ツールが出力する結果は、あくまでも課題発見のための「気づきのもと」であり、課題を発見してアクションに結び付けるのは、業務に携わる「人」である。また、顧客の声から発見される課題には部署をまたがって解決すべきものが多い。そのため、分析結果とその課題を全社的に共有する仕組みと、全社横断的な権限をもつ組織が必要となってきた。 ■