

CoreExplorer を活用した 故障情報分析ソリューション事例

Case Study of Failure Data Analysis Solution Using CoreExplorer

製造業である A 社では、日々寄せられる製品の故障情報を電子化し社内データベースに蓄積している。蓄積したデータはデータマイニングの手法を用いて分析を行っているが、故障情報に含まれる自由記述文章は定型でないため、集計できず分析を行えていないのが現状である。そこで、自由記述文章を自動分類するため、テキストマイニングツールである CoreExplorer・VOCAnalyzer の導入に至った。

本報告では A 社で開発している故障情報分析システムに CoreExplorer と VOCAnalyzer を導入した事例について紹介する。

宮内 秀彰 Miyauchi Hideaki
塚原 朋哉 Tsukahara Tomoya
佐藤 俊也 Sato Syunya

1. はじめに

近年企業では、様々なデータが電子化されており、これら进行分析することによって営業活動・業務改善を行っている。製造業のお客様においては、故障報告書などの電子化が進み、製造過程で発生した故障品の情報や、市場で発生した故障品の情報が社内に蓄積されている。市場に出てから発生した故障は直接クレームにつながり、社会的信頼を大きく損なう恐れがあることから、高品質な製品の提供が求められている。このことから社内に蓄積された故障情報を解析し、根本原因の特定や部品寿命を予測するといった試みがされている。¹⁾²⁾³⁾⁴⁾

しかし、故障報告書の”故障状況”や”故障原因”，”対策”といった項目は自由記述で書かれていることがほとんどである。大量の故障情報を人が一つ一つ読んで分析することは非常に時間がかかる。年間数万件のデータが蓄積される場合、すべてを読んで分析することは現実的に不可能である。

そこで、故障報告書の自由記述文章にテキストマイニングの技術を用いることで、定量的に分析を行うソリューションを行ってきた。

本報告では、テキストマイニングツールである、CoreExplorer, VOCAnalyzer を用いた、製造業の A 社における故障情報分析システムについて紹介する。

2. CoreExplorer・VOCAnalyzer 導入の背景

テキストマイニングを用いた故障情報分析システムの代表的な例として、キヤノン (株) の FITS がある。⁵⁾ FITS は、自由記述文章を含んだ製品の故障情報を収集し、これらのあるルールに従って自動的に分類し分析するシステムである。これにより定量データと自由記述のデータを組み合わせた分析を行うことができる。

今回紹介する A 社においても、自社の製品に対する故障情報を分析するシステムの刷新を行うこととなり、テキストマイニングを用いたシステムを構築することが検討された。

今回構築された品質情報管理システムは、全国で稼働している自社製品の不具合・故障情報を収集し分析するシステムである。部品寿命の予測や、不具合の原因となる製造のボトルネックを発見し、製造プロセスの改善を本システムで実現する。

テキストマイニングを利用したシステム開発にあたっては、分析ルールの作成や管理、分析の自動化の可否、製品の導入実績等の観点で検討を行った。検討の結果、以下の観点から CoreExplorer を利用しシステム化を行うこととなった。

- ・VOCAnalyzer による分類規則作成のし易さ
- ・CoreExplorer と VOCAnalyzer の連携のし易さ
- ・製造業への過去の導入実績の豊富さ

次章以降において、本システムの概要と効果について述べる。

3. システム概要

今回 A 社が構築した品質情報システムは、4つのサブシステムと1つの DBMS から構成されている。システム全体の構成図を図1に示す。

各サーバ・PCの構成と詳細は以下のとおりである。

- ・サーバ：仮想サーバ4台 (VMWare 上に構築)
 - ・故障情報登録システム
 - ・故障情報分析システム
- ・CoreExplorer
- ・Oracle Database 11g
- ・VOCAnalyzer 用 PC(管理者用)

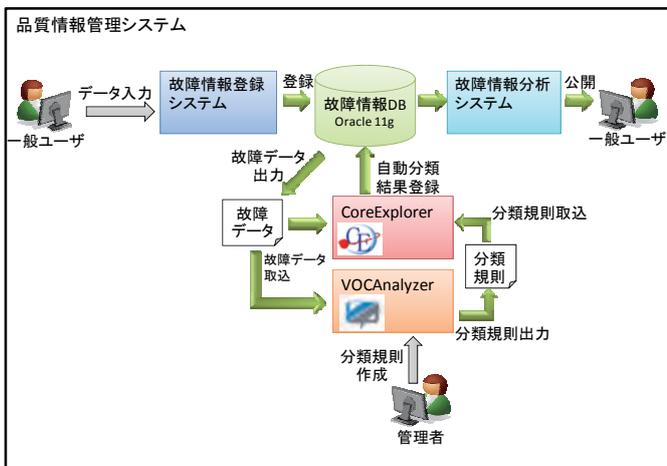


図1 全体システム構成図

3.1 故障情報登録システム

製品の故障が発覚した場合、故障情報を登録する仕組みとして本システムが導入されている。

登録される故障情報は下記のとおりである。

- 入力者：製品メンテナンス実施者
 入力時期：製品故障，異常発生後数日以内
 入力内容：故障情報入力フォームにて，以下内容を登録する。

“発生場所”，“導入年月日”，“故障発生年月日”，
 “故障発生時刻”，“故障完了年月日”，
 “故障完了時刻”，“故障番号”，“担当事務所”，
 “製品型番”，“エラーコード”，“部品型番”，
 “処置内容(自由記述)”，
 “顧客対応(自由記述)”，“備考(自由記述)”，
 “故障状況(自由記述)”，“故障原因(自由記述)”
 等...

3.2 故障情報分析システム

故障情報登録システムによって蓄積された故障情報をデータマイニング等の手法で分析することで，製品部品の故障原因の特定や部品寿命を予測する。また故障情報レポートを社内に公開し共有化を図る。

3.3 VOCAnalyzer

VOCAnalyzer は，CoreExplorer で自動分類を行うために必要な，“分類規則”を作成するテキストマイニングツールである。⁶⁾ CoreExplorer・VOCAnalyzer 共通の分類規則は大分類・中分類の2階層からなり，各分類に振り分けるためのキーワードを設定する。

VOCAnalyzer は類似文章をグループ化するクラスタリング分析機能を備えており，設定するキーワードの選出を支援する。また豊富な GUI で分類規則の作成を直感的に行うことができる。

3.4 CoreExplorer

CoreExplorer は登録したテキストデータや属性データを組み合わせた多角的な分析や，グラフ作成，レポート作成を行うことができるテキストマイニングツールである。今回の品質情報管理システムにおけるCoreExplorerの役割としては，故障データの中で自由記述された内容をテキストマイニングし，VOCAnalyzerで作成した分類規則に従い，文章の自動分類を行うことを目的としている。

4. CoreExplorer・VOCAnalyzerによる分析プロセス

品質情報管理システムにおいてCoreExplorerとVOCAnalyzerの役割はテキストデータの自動分類である。本章では，VOCAnalyzerによる分類規則作成からCoreExplorerでの自動分類までの分析プロセスについて説明する。

図2はCoreExplorerとVOCAnalyzerによる自動分類の分析プロセスを示している。

4.1 VOCAnalyzerによる分類規則作成

① 図3はVOCAnalyzerのデータ登録画面を示している。CSV形式の故障データを管理画面から登録する。その際にテキスト分析を行う対象列を指定する。

類によって文章に属性情報を付加したことにより、集計が可能となり従来のデータ項目+自由記述文章のデータを分析対象に含めることが可能となった。

以下に自由記述項目を分析に含めた場合の分析例について示す。

分析例 1：“導入期間”×“故障原因”のクロス分析 (図 6)

横軸：導入期間区分別 故障原因

縦軸：度数/1 カ月

データ件数：1675 件

導入してから故障原因がどのように変化するか推移を分析（初期不良，経年劣化等の分析）

図 6 の導入後から 3 カ月未満を見ると，全体的に故障が多く発生していることがわかり，原因 1 の頻度が最も高いことがわかる。また，原因 28 は 2 年以上～5 年未満に多く故障していることがわかる。同様に，原因 13 は 10 年以上～15 年未満が高く，原因 23 は 15 年以上～20 年未満が最も高い。

このように，初期不良，経年劣化の故障原因推移を「見える化」することができる。これにより，様々な改善策を立てることができる。

例えば，原因 28 が 2 年以上～5 年未満に多く故障しているが，部品の劣化によるものであれば，交換のサイクルを早める。また，点検時の要注意部品としてマニュアルに追加するといった対策を講じることができる。

また初期に原因 1 で多く故障が発生しているが，製品出荷時にある部品の品質にばらつきがあるのではないかと，といった仮説を立てることができる。原因 1 に関連する部品を調べることで，初期不良の根本原因を探ることができる。

分析例 2：“故障原因”×”部品型番”のクロス分析(図 7)

【棒グラフ】

横軸：部品型番別 故障原因

縦軸：故障頻度

データ件数：1675 件

【円グラフ】

各部品型番の故障原因内訳
部品ごとに発生している故障原因を知ることができる。

図 7 を見ると部品型番 12 においては，原因 1,原因 7,原因 12,原因 13 の割合が高いことがわかる。

このように部品ごとに発生している故障原因を追及することが可能である。

例えば，部品型番 12，19 において，原因 1 はほぼ同じ割合で発生していることがわかるが，部品型番 15 は割合が少ないことがわかる。例として，部品の生産工場が部品型番 12，19 は同じであるが，15 は違うといった場合，部品番号 12，19 における製造ラインに問題はないか，といった仮説を立てて調査をすることができる。

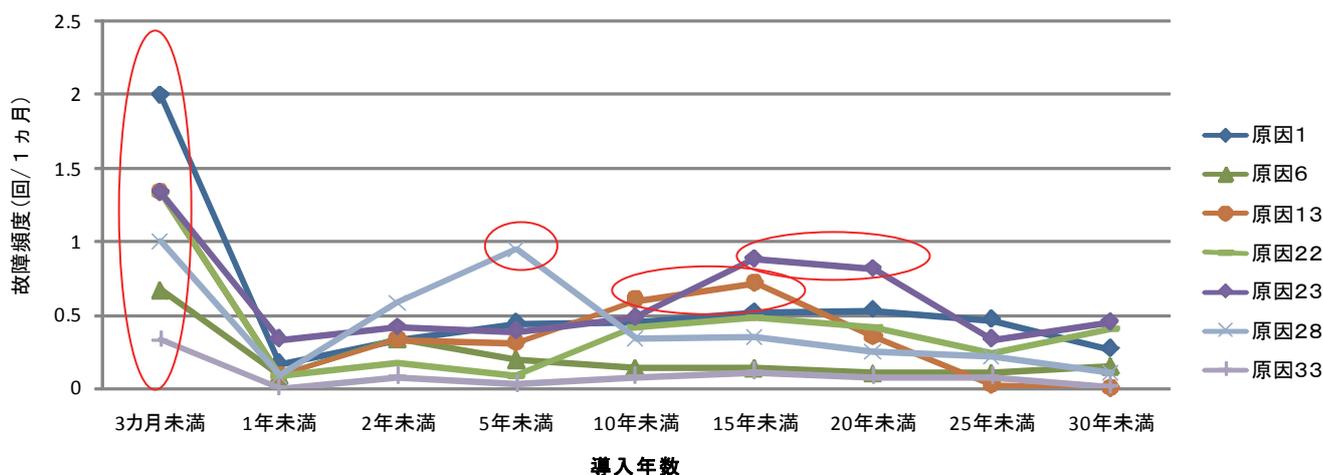


図 6 導入期間と故障原因のクロス分析結果

5.2 自動分類による分析の効率化

本システム導入前は、自由記述文章は一部の活用しか
なされていなかった。今回のシステム導入により、自由
記述文章を定量的に扱えるようになったため、すべての
項目を考慮した分析が可能になった。分析できる量が増
えたため、単純にシステム導入前と導入後の作業効率を
測ることは困難であるが、システム導入後の作業を手作
業で行った場合を参考値としてシミュレーションした。

・システム導入前

仮に 3000 件/月の文章の分類を手作業で行った場合、
1つの文章を 200 個の分類から最適な分類にあてはめる
ために必要な時間を平均時間 1分とすると、
 $3000(\text{件/月}) \times 1(\text{分}) \times 12(\text{カ月}) = 36000 \text{ 分/年} = 600 \text{ 時間/年}$
年間 600 時間の作業時間を要する。1 日を 7.75 時間と
し、ひと月 20 日 かけて分析したとすると、約 4 カ月に
相当する時間である。そのため、手作業で分析を行うこ
とは現実的に考えにくい。実際に行う場合は重要と思わ
れる内容のみを抽出してその部分のみの分析を行うと推
測される。

・自動分類システム導入後の場合

CoreExplorer での分類は自動のため、人手での作業は
VOCAnalyzer による分類規則の作成とメンテナンスの
みである。分類規則作成に約 5 人日かかるが、一度分類
規則を作成すれば、後はメンテナンスを行うだけで良い。
メンテナンスに 1 人日/月程度かかると仮定すると

$$7.75(\text{時間}) \times 5(\text{日}) + 7.75(\text{時間}) \times 12(\text{カ月}) = 131.2 \text{ 時間}$$

年間 131 時間の作業時間を要する。

単純比較をすると、手作業による分析の場合に比べ、
年間 469 時間の作業時間短縮が可能との見込みである。
さらに、自動分類の場合は文章がいくら増えても、分
類規則さえ作成されていれば、人手での時間は増加しな
い。

6. 今後の課題

自動分類を行う上で分類の精度や、分類規則の作成作
業の負荷は重要な課題である。ここでいう精度とは人が
分類した際に、分類されるべき分類項目に分類されるか
どうかを指している。

例えば、”良くない” といった言葉は否定する言葉で
あるが、”良い” というキーワードをもとに分類して

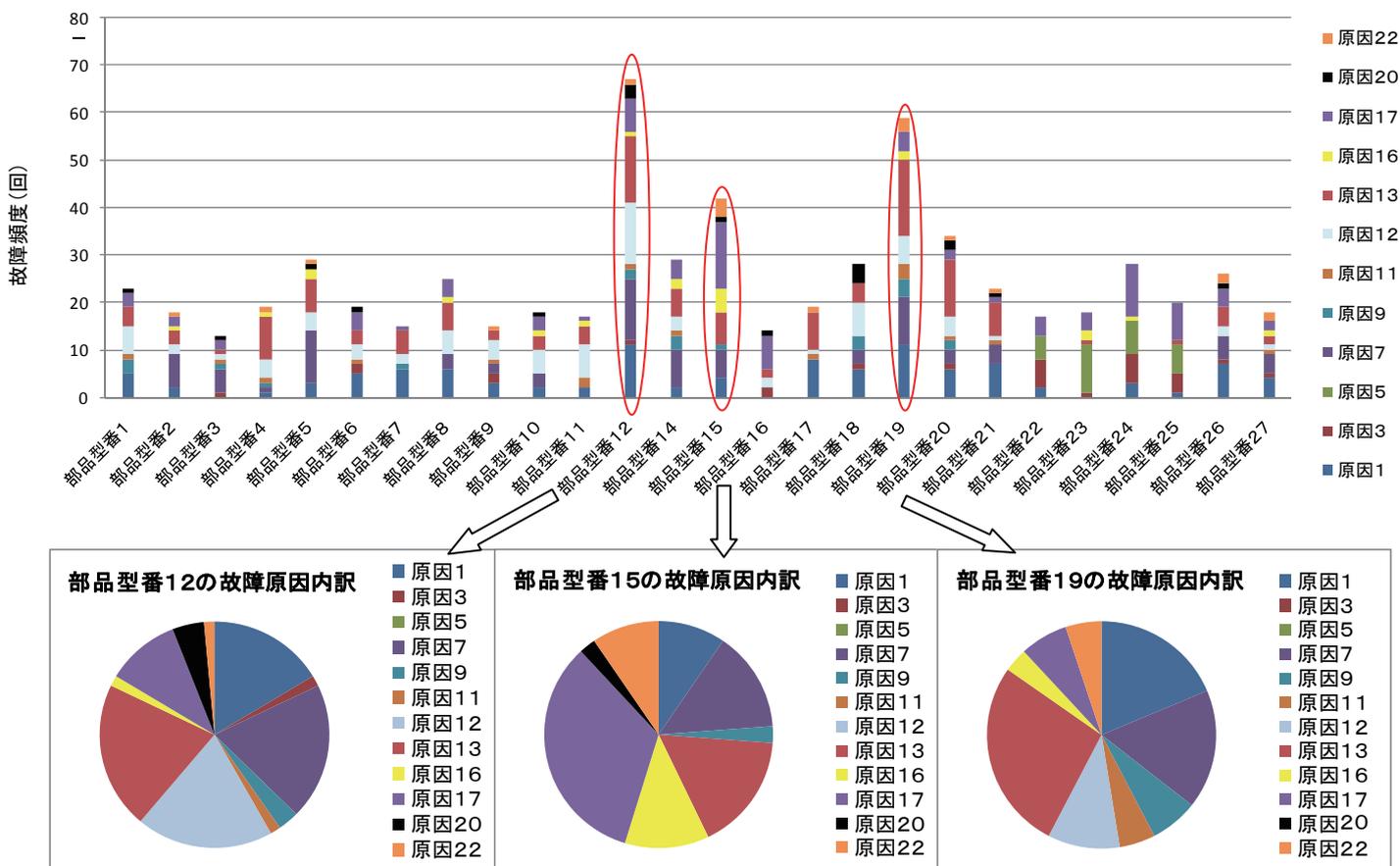


図 7 故障原因と部品型番のクロス分析結果

しまうと、肯定的な文章として分類されてしまう。

分類規則に”良いーない”というルール(「良くない」の構文解析を考慮したルール)で登録することで、否定的な文章を分類することは可能であるが、分類規則作成が専門的になりすぎたり、メンテナンスの負荷が増大すると運用に支障をきたす恐れがある。今後は分類精度を維持した上で、分類規則作成の負荷が増大しないシステムを検討する必要がある。

7. 今後の展開

A社では、故障情報だけでなく営業日報の活用や設計資料の分析も行いたいと考えているため、CoreExplorer活用の機会はさらに増える見込みである。

一方、A社の担当者からはCoreExplorerの機能としては多変量解析等による分析機能を強化してほしいといった要望がある。テキストマイニングからより有意な情報を素早く見つけるためには、多変量解析・データマイニングの要素が重要であり、現在機能の実装を進めている。

今後CoreExplorerでは、次期バージョンに向けてこのような顧客ニーズを踏まえ、分析機能の開発に注力する予定である。

8. おわりに

本報告では、製造業における故障情報分析の事例を紹介した。

一般的に、テキストマイニングはコールセンタに寄せられる声やアンケート分析が一般的であるが、製造業と密接な関係のある日立TOでは、このような故障情報に対する事例を多く行ってきたことで、同分析に対するノウハウを多数蓄積してきた。

今後は情報分析の分野でもお客様のお役にたてるソリューションを提供できるよう、研究・開発を進めていく所存である。

参考文献

- 1) 安藤 英幸他:テキストマイニングを用いた故障報告書分析手法の研究, 日本造船学会論文集 (192), pp.475-483, 2002-12
- 2) 堀 聡 他: データマイニングを用いた市場品質監視システム, 電気学会論文誌. C, 電子・情報・システム部門誌 121(8), pp.1289-1295, 2001-08-01

- 3) キヤノン IT ソリューションズ株式会社: 製造不良要因分析, <http://www.canon-its.co.jp/math/consultation/man/6011P.pdf>
- 4) A. Bergmann: Data Mining for Manufacturing: Preventive Maintenance, Failure Prediction, Quality Control, NDT.net, 2010
- 5) 石井 哲: テキストマイニング活用法, (株)リックテレコム(2002), pp.174-182
- 6) 宮内 他: お客様の声分析システム”VOCAnalyzer”の開発, 日立 TO 技報第 15 号, pp.17-21, 2009
- 7) Berry.Michael J.A.: データマイニング手法, 海文堂出版, 2005



宮内 秀彰 2008年入社
ナレッジソリューショングループ
CoreExplorer, テキストマイニングツールの開発
hideaki.miyauchi.01@hitachi-to.co.jp



塚原 朋哉 1997年入社
ナレッジソリューショングループ
CoreExplorer, テキストマイニングツールの研究開発
tomo@hitachi-to.co.jp



佐藤 俊也 1993年入社
ナレッジソリューショングループ
CoreExplorer, テキストマイニングツール, 知識管理システムの拡販, コンサルティング
shu_sato@hitachi-to.co.jp