



Title	イノベーションを支援するリスク分析ソリューションの研究
Author(s)	村山, 秀次郎
Citation	大阪大学, 2010, 博士論文
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/23476">https://hdl.handle.net/11094/23476</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

情 14330

イノベーションを支援するリスク分析  
ソリューションの研究

2010年1月

村山 秀次郎

**イノベーションを支援するリスク分析  
ソリューションの研究**

**提出先 大阪大学大学院情報科学研究科**  
**提出年月 2010年1月**

**村山 秀次郎**

# 発表論文リスト

## I. 発表論文（学術雑誌等掲載論文）

- (1) **Shujiro Murayama**, Koji Okuhara and Hiroaki Ishii, Framework of Risk Elimination from Reports on Unsafe Acts, Mechanical and Physical Hazards, International Journal of Japan Association for Management Systems(IJAMS), 掲載決定
- (2) **Shujiro Murayama**, Junko Shibata, Koji Okuhara, and Hiroaki Ishii, Data Mining for Hazard Elimination Through Text Information in Accident Report, Asia Pacific Management Review (APMR), 掲載決定

## II. 国際会議等講演論文（査読付き）

- (1) Koji Okuhara, **Shujiro Murayama** and Hiroaki Ishii, Expansive Rule Construction for Text Mining with Large Scale Data, Proceedings of the 11th Asia Pacific Management Conference, in CD-R, pp. 3D4-1-3D4-4, Tainan, Taiwan, (November 18-20, 2005).
- (2) **Shujiro Murayama**, Koji Okuhara and Hiroaki Ishii, Problem Discovery Framework by Text Mining from Accident Reports at Manufacturing Site, Proceedings of the 11th Annual International Conference on Industrial Engineering Theory, Applications, and Practice IJIE2006, in CD-R(ISBN: 0-9654506-2-7), pp. 835-840, Nagoya, Japan (October 24-27, 2006).
- (3) **Shujiro Murayama**, Koji Okuhara and Hiroaki Ishii, Risk Management by Problem Discovery Framework from Accident Report, Proceedings of the

12th Asia Pacific Management Conference, in CD-R(ISBN: 974-8257-30-4), pp. 259-262, Bangkok, Thailand (November 17-19, 2006).

- (4) **Shujiro Murayama**, Koji Okuhara and Hiroaki Ishii, Innovation in Manufacturing Premise by New Finding Obtained From Accident Relapse Prevention Report, Proceedings of the 13th Asia Pacific Management Conference, in CD-R, pp. 1124-1129, Melbourne, Australia (November 18-20, 2007).
- (5) **Shujiro Murayama**, Koji Okuhara, Hiroaki Ishii, THE MINISTRY OF LAND, INFRASTRUCTURE AND TRANSPORT OF JAPAN AND FINDING FROM SERIOUS, INCIDENT REPORT CONSIDERATION OF ACCIDENT RELAPSE PREVENTION PLAN DECISION MODEL-, The International Conference on Business & Information 2008, pp. 153-160, Honolulu, Hawaii (2008.9.2-3)
- (6) **Shujiro Murayama**, Junko Shibata, Koji Okuhara, Hiroaki Ishii, The Innovation of Manufacturing Premise based on Knowledge obtained by Quality Accident Report, in CD-R, ISDA2008, pp. 593-598, Kaohsiung, Taiwan (November 26-28, 2008)
- (7) Eri Domoto, **Shujiro Murayama**, Koji Okuhara, Hiroaki Ishii and Haruhiro Fujita, Modeling from Daily Menu to Physical Condition and Prediction of Condition Change, Proceedings of 9th APIEMS Conference and 11th Asia Pacific Regional Meeting of International Foundation for Production Research, pp. 135-139, (Dec., 2008).
- (8) Eri Domoto, Koji Okuhara and H. Ishii and **Shujiro Murayama**, echnology of Energy Saving Point Include Standardization and Awareness, Proceedings of EcoDesign 2008 Japan Symposium, D23-4, Tokyo, Japan (December 11-12, 2008).

- (9) **Shujiro Murayama**, Koji Okuhara and Hiroaki Ishii, Knowledge Discovery for Secure Traffic Management Based on Risk Analysis in Transportation Industry, Proceedings of 5th International Conference on Planning and Design, pp. 135-139, Tainan, Taiwan (May 25-29, 2009)

### Ⅲ. 学会全国大会および支部大会講演論文

- (1) 村山 秀次郎, 奥原 浩之, 石井博昭, 製造現場におけるイノベーション — 事故再発防止報告書から得た新たな知見による変化と対応 —, 第 38 回日本経営システム学会 全国研究発表大会、日本大学生産工学部 生産工学部 津田沼キャンパス (2007 年 6 月 26-27), 第 38 回日本経営システム学会 全国研究発表大会講演論文集, p. 30-33, 東京, (June 26-27, 2007)
- (2) 村山 秀次郎, 奥原 浩之, 石井博昭, 国交省 重大インシデント報告書からの知見 — 事故再発防止策策定モデルの一考察 —, 第 39 回日本経営システム学会 全国研究発表大会、中村学園大学 (2007 年 12 月 8-9), 第 39 回日本経営システム学会 全国研究発表大会講演論文集, p. 130-133, 九州, (December 8-9, 2007)

### Ⅳ. 学会技術研究会等講演論文

- (1) 村山 秀次郎, リスク・マネジメントを支援するテキストからの課題発見プロセスの研究 ～製造現場でのイノベーションをめざして～, 日本 OR 学会「OR と実践」研究部会・「ソフトコンピューティングと OR」部会 共催 「OR と実践」研究部会第 8 回研究集会・「ソフトコンピューティングと OR」部会第 3 回研究集会 (共催)
- 日 時：平成 19 年 9 月 19 日(水) 14:30～17:30
- 場 所：福井工業大学 FUT タワー多目的会議室 (〒910-8505 福井市学園 3 丁目 6 番 1 号)

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 研究の背景および目的	1
1.2 本論文の構成	3
<b>第2章 リスク・マネジメントとテキストマイニング</b>	<b>5</b>
2.1 企業経営とリスク・マネジメント	5
2.2 イノベーションの定義	13
2.3 テキストマイニング技術について	15
2.3.1 テキストマイニングの研究の経緯	15
2.3.2 テキストからのキーワード抽出法	21
2.4 キーグラフの概要	34
2.4.1 キーグラフの基本的な考え方	34
2.4.2 キーグラフ上での表示の仕方	39
<b>第3章 課題発見のプロセス構築</b>	<b>41</b>
3.1 チャンス発見	41
3.2 分析手法の概要	44
3.2.1 回帰分析	44
3.2.2 属性の序数性を考慮したラフ集合	46
3.2.3 自己組織化マップ (SOM)	52
3.3 提案する報告書の形式	54
3.4 提案する課題発見プロセス	60

<b>第4章</b>	<b>イノベーションを支援するリスク分析</b>	<b>63</b>
4.1	A食品会社における品質事故への適用	63
4.1.1	回帰分析の適用	63
4.1.2	回帰分析の結果と考察	68
4.1.3	属性の序数性を考慮したラフ集合（以下提案ラフ集合分析と記す）の適用	74
4.1.4	SOM の適用	80
4.1.5	分析手法と結果の比較評価	81
4.2	小型飛行機事故分析	83
4.2.1	回帰分析の適用	85
4.2.2	回帰分析の結果と考察	90
4.2.3	提案ラフ集合分析の適用	93
4.3	医療事故…ヒヤリハット情報の分析	95
<b>第5章</b>	<b>実システムの開発</b>	<b>99</b>
5.1	事故報告システム	99
5.2	事故報告システム操作マニュアル	103
<b>第6章</b>	<b>結論</b>	<b>109</b>
6.1	まとめ	109
6.2	今後の課題	110
<b>謝辞</b>		<b>113</b>
<b>参考文献</b>		<b>115</b>

# 目次

図 1.1	研究の全体像	3
図 2.1	リスク・マネジメントを取り巻く経営環境とリスク	6
図 2.2	変動の時代の新しい戦略的経営アプローチ	7
図 2.3	テキストマイニング処理プロセス	25
図 2.4	自然言語の階層的構造	27
図 2.5	日本語の品詞分類	29
図 2.6	Windows 版茶釜「WinCha2000」の処理画面	31
図 2.7	品詞情報を基にキーワードを取り出す	34
図 2.8	土台・屋根・柱の関係	36
図 2.9	キーグラフによる解析結果表示例	40
図 3.1	チャンス発見のプロセス	42
図 3.2	チャンス発見の二重らせんプロセスモデル	43
図 3.3	SOMの構造	53
図 3.4	作業の流れ	55
図 3.5	キーグラフからの出力	57
図 3.6	提案概要	58
図 3.7	課題発見プロセス	61
図 3.8	リスク対応プロセス	62
図 4.1	影響度グラフ	66
図 4.2	最適な項目での影響度グラフ	69
図 4.3	経験年数別影響度グラフ	70
図 4.4	発生曜日別影響度グラフ	71
図 4.5	発生時間別影響度グラフ	71
図 4.6	自己組織化マップ (SOM) での損害あり時の出力	80

図 4.7	分析手法の評価	81
図 4.8	ハインリッヒの法則	84
図 4.9	小型飛行機の影響度グラフ	88
図 4.10	Ru で選ばれた項目での影響度グラフ	90
図 4.11	風の影響度グラフ	91
図 4.12	天候の影響度グラフ	92
図 4.13	事故内容の影響度グラフ	92
図 5.1	デシジョンテーブル作成支援システムの概要	100
図 5.2	リスク情報登録および検索システムの概要	101

# 表目次

表 2.1	テキストマイニング関連研究	17
表 2.2	テキストマイニング関連研究の概要 (1/2)	17
表 2.3	テキストマイニング関連研究の概要 (2/2)	18
表 2.4	TF. ID 法とキーグラフ法の性能比較	23
表 2.5	茶釜による分析結果	32
表 2.6	品詞情報からキーワードとして適切かの判断例	33
表 3.1	チャンス発見事例	44
表 3.2	ミカンの品質に関する決定表	49
表 3.3	事故再発防止報告書(サンプル)	55
表 3.4	顧客固有語登録例	56
表 3.5	ストップワード(削除語)例	57
表 3.6	デシジョンテーブル第1版	59
表 3.7	デシジョンテーブル第2版	59
表 3.8	デシジョンテーブルにプロットされたサンプルデータ(一部)	60
表 4.1	多変量解析データ(部分)	64
表 4.2	SPSSでの回帰分析結果	65
表 4.3	影響度分析	66
表 4.4	説明変数選択基準(Ru)評価表	67
表 4.5	最適な項目でのSPSS回帰分析結果	68
表 4.6	最適な項目での影響度分析	69
表 4.7	経験年数別影響度	70
表 4.8	発生曜日別影響度	71
表 4.9	発生時間別影響度	71
表 4.10	最適な回帰式	73

表 4.11 提案ラフ集合分析用データ（一部）	74
表 4.12 分析解読用デシジョンテーブル	75
表 4.13 決定ルール結論 1（事故なし①）の時の出力リスト	75
表 4.14 決定ルール結論 1（事故なし②）の時の出力リスト	76
表 4.15 決定ルール結論 2（事故あり①）の時の出力リスト	76
表 4.16 決定ルール結論 2（事故あり②）の時の出力リスト	77
表 4.17 提案ラフ集合分析の出力例	78
表 4.18 事故なしのルールパターン	78
表 4.19 事故ありのルールパターン	79
表 4.20 提案ラフ集合分析により選択された項目	79
表 4.21 SOM のためのデータ（一部サンプル）	80
表 4.22 発生年別事故件数（2009/7/31 現在）	85
表 4.23 小型飛行機のデシジョンテーブル項目	86
表 4.24 小型飛行機の回帰分析結果	87
表 4.25 小型飛行機の影響度分析結果	88
表 4.26 小型飛行機の説明変数基準（Ru）評価表	88
表 4.27 説明変数選択基準（Ru）で選ばれた項目での回帰分析結果	89
表 4.28 説明変数選択基準（Ru）で選ばれた項目での影響度分析結果	90
表 4.29 風の影響度	91
表 4.30 天候の影響度	92
表 4.31 事故内容の影響度	92
表 4.32 提案するラフ集合分析結果	94
表 4.33 Ru で選ばれた項目群と提案するラフ集合での項目群対比表	94
表 4.34 事故報告システムで作成した医療向け提案するデシジョンテーブル	98

# 第 1 章 序論

## 1.1 研究の背景および目的

近年，多変量解析やデータマイニングなど，データをより積極的に分析・探索し，ビジネスに有益な何らかの知識や知見を引き出そうとする分析手法への注目が高まっている．これらの分析アプローチは，情報の電子化に伴い大量に蓄積されたデータに潜む「知見」をまさに宝を掘り出す（マイニングする）ように分析するというものである．これらの探索的な分析は，事前に想定しない知見や発見を導き出しうるため，増え続けるデータから知見を見つけるのに適したアプローチであり，ますます使用される場面に広がりを見せている．

筆者は，すでに職を辞しているが在職中，国産コンピュータ・メーカ系システム・エンジニアリング会社において 34 年間 SE（システム・エンジニア）として数百社の顧客システムの開発およびサポートを担当してきた．顧客サポートの中で，将来データは‘宝を生む’との確信からデータの長期保存を顧客に提案し，推進してきた経緯がある．企業を退職した現在，顧客の指導を通じてデータが‘宝を生む’ということに対して答えを出したのだろうかという疑念にいささか慄いていた．その問いに対する解法の一手法として，データマイニング・テキストマイニングを知るにおよび若干の安堵を得たのである．

データマイニングの技術は年々向上し，ベンダーもその重要性を認識し商品化を推進してきた．しかし，まだ普及には時間が必要であろう．その背景には，企業にとって，確実に富を生むツールとは位置位置づけされていない．すなわち投資対効果が明確に計測できない現実が存在するからである．また，実務で期待されているほどの成果はなかなか得られていないのである．

その原因の一つは分析者のスキル不足にある。ソフトウェアが高機能化し、また、操作が簡単になったとしても、「課題に適切な手法は何であり、何を引き出せばよいか」といった点は分析者が解決しなければならぬ。その上、分析結果を読み取り解釈するのは、あくまで人間であるため分析者のスキル向上が不可欠となる。もう一つはデータ自体にある。分析者が期待する知見の質は高まりつつあるが、単なるデータ上の結果ではなく、より実践的な解釈や応用が可能な知見が求められているのであり、その意味で分析データの質の向上抜きでは達成できないのである。

以上のような課題を抱えつつも多くの実務家がデータ分析に興味を持ち、分析を実際に経験したことで、分析環境は整えられたが、より知見を含んだデータを分析対象にしなければならないことが再認識された。「より多くの知見を含んだデータ」として、電子化された言葉の情報つまりテキストデータへの期待がある。

IT化の進展によって、日常生活の中で発生する、膨大な量の情報（自由記入された各種アンケートやネット上のブログでの自己主張、ネット上での仲間内のコミュニティ会話、コールセンターでの顧客とオペレータとの交渉記録、ネット上での社員同士の議論、営業活動での日次報告、製造活動における事故報告など）が溢れている。これらの言葉の情報から価値ある知見を生み出す技術として今日注目されているのが、テキストマイニングの技術である[1, 2]。

商品の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た。適用例ではコールセンターでの利用が多く、クレーム処理の短時間化に利用されている。ただ、テキストマイニングで出された結果はあくまでも課題発見のための「気付きのもと」であり、課題をどう見つけ行動につなげるかが、新たな研究課題となっている。今後の産業界への普及を鑑み、今こそ先行研究の必要性があり、新たなチャンス発見への一歩と位置づけ研究に取り組むことにした。

そんな折、広島県竹原市に本社を置くA食品会社は、品質管理と危機管理の必要性から事故再発防止報告書制度を採り、データの蓄積を行ってきたが有効利用ができていないことが問題点であった。そんな現状からの脱却をはかるべく、蓄積されたデータの有効利用のための解決策としてアカデミック的な分析手法を問

う相談を受けた。筆者にとって今回取り組む研究課題との調和もあり，また研究推進の検証の場として受け入れてくれることを前提に，再度にわたる提案を重ね，産学連携の形で取り組むことにした。

本研究の目的は，前述の問題点に対して，工場の現場で発生した品質事故（今回は人為的事故）の報告書である日報や事故再発防止報告書のテキストデータをもとに，テキストマイニングを適用して，事故報告者の主張（原因，防止策など）を的確に抽出し，その結果をもとに得られた知見，課題と現実にたどった結果との比較実証を行い，課題発見のプロセスを見出すことが目的である。本論文では，社員の主張を的確に抽出するキーワード抽出技術，および，抽出されたキーワードをいかに活用するかという適用技術について論ずる。図 1.1 に研究の全体像を示す。

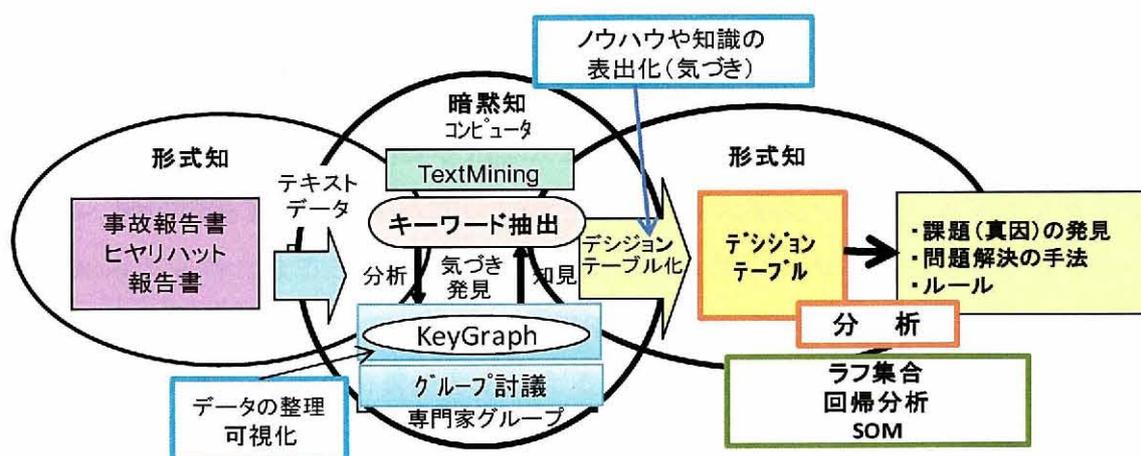


図 1.1 研究の全体像 [3]

## 1.2 本論文の構成

本論文の構成は次の通りである。

第 2 章では，リスク・マネジメントについて企業における基本的な考え方，および対処方法について論ずる。また，本研究のテーマであるイノベーションについて諸説を踏まえて定義を試みた。

さらに、テキストマイニングの技術の現状と、どのような分野でこの技術が適用されているかを解説する。不定形な文書からキーワードの抽出技術として、「共起グラフによるキーワード抽出法 KeyGraph」の有効性、および、適用する応用分野に依存した KeyGraph の適用方法の相違点について論ずる。

第 3 章では、本研究の主題であるリスク分析による課題発見のプロセスについて述べ、その実証として A 食品会社での課題発見プロセスの適用、および事故再発防止に向けた課題発見のための、分析方法として従来の分析手段である回帰分析を、また、量的データに対して属性間の序数性を考慮したラフ集合（提案ラフ集合）および自己組織化マップ（SOM）を提案する。

第 4 章では、連携先である A 食品会社の事故再発防止報告書をデシジョンテーブルにプロットし、第 3 章で提案した分析手法に沿って分析し解析を行い、結果と考察を述べ評価を行う。また、課題発見のプロセスを他の実業界への適用として航空機業界および医療事故に試みる。

第 5 章では、提案した課題発見のプロセスを広く実業界に提案するために、簡便に導入できるソリューションをめざした実システムとして開発を試みた。そのシステムの概要と利用マニュアル等を示す。

第 6 章では、各章で得られた研究結果を総括するとともに、今後の課題について述べる。

# 第2章 リスク・マネジメントと テキストマイニング

## 2.1 企業経営とリスク・マネジメント

経営環境が激しく変化する現代の企業にとって、過去の成功に学ぶよりも、失敗を避けて経営リスクをどのようにとらえてどう対応するかというリスク・マネジメントが重要な時代が到来してきた。そして、リスク・マネジメントには適切な仕組みが必要となる。なぜなら、その仕組みを動かす人がカギを握ることになるからである。企業のリスクは社会全体にも大きなダメージを与えかねず、リスク・マネジメントは社会的責任ともいえる。以下、文献[4,5]に基づいて説明する。

図2.1は企業に求められるリスク・マネジメントを取り巻く経営環境とさまざまな経営リスクを図示したものである。以下その変化の構造を分析し、まったく新しい市場の出現に企業がどのように対応していくべきかをリスク・マネジメントを通して解決策を考える。

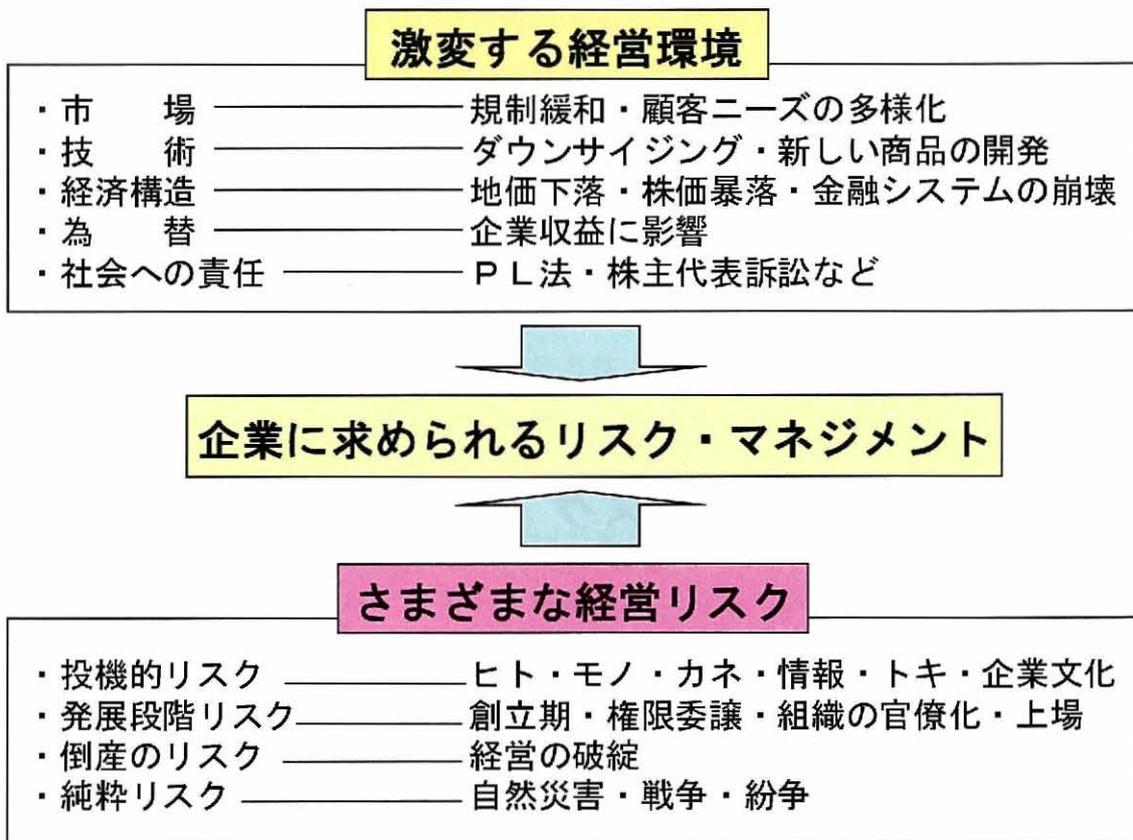


図 2.1 リスク・マネジメントを取り巻く経営環境とリスク

### (1) 変化の時代

戦後 50 年間の経営環境の変化と、これから 21 世紀前半までの 50 年間に予想される経営環境の変化を明確な形でとらえることが、経営者にとって重要である。

また、経営者は「変化を前提とした経営」と「変化によって生じるリスクを直視する経営」という 2 つの観点を踏まえ、特に経営戦略にリスク・マネジメントを組み込むことが必要となる。

物不足の時代から、物余りの時代へ、作れば売れた時代から売れない時代へ、そして安ければ売れた時代から、高くっても顧客ニーズに合えば売れる時代へと消費者の求めるものが大きく変化している。

このような時代には、このように多様化している顧客ニーズをいかに経営システムに取り組みかが、最大の経営リスクとなっている。そして、新しいパラダイムに合わせた戦略、すなわち、変化する環境の中、即断即決で戦略を実行する新

しい型の戦略的アプローチが最適となる（図 2. 2 参照）。そのためにも市場の変化に即した経営情報分析の見直しも必要となる。

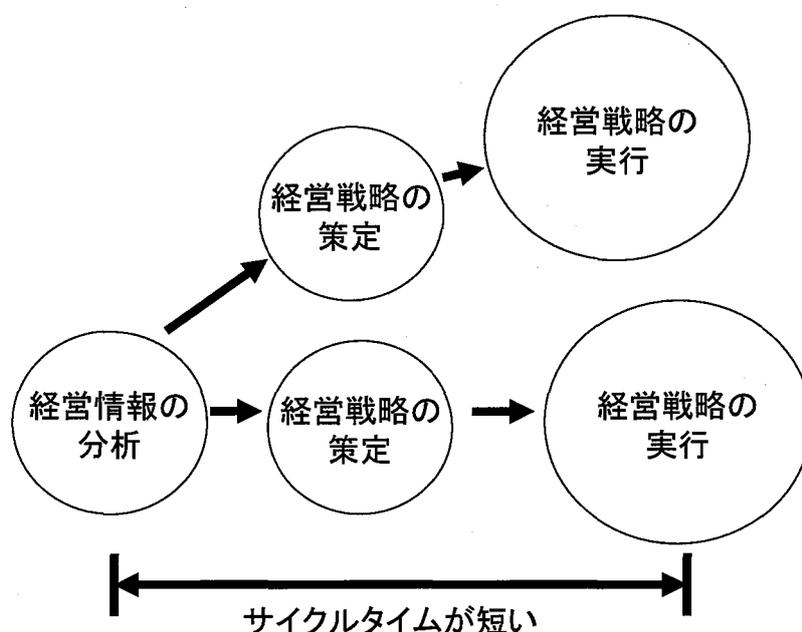


図 2. 2 変動の時代の新しい戦略的経営アプローチ

## (2) パラダイムシフトを招く環境条件

経営環境の劇的な変化によって、経営者は利益を出すために、従来とはまったく異なった経営手法の実行を余儀なくさせられる。パラダイムシフトとはこのような時期のことをいう。また、未来学者ジョエル・A・バーカー[6]は「パラダイムとは、境界を決めるルールや規則を意味し、その境界の中で何をすれば成功することができるかを教えてくれるもの」と定義し、パラダイムはそれに従わないと損を被ることになるビジネスのルールや枠組みである。そこで、パラダイムシフトが起きた場合はリスク・マネジメントが重要となる。

以下にパラダイムシフトを招く主な環境条件の変化をしてみる（図 2.1 参照）。

### ①市場の変化

大店法の改正や金融・保険・証券の事業規制の撤廃などに見られる規制緩和や、団塊の世代とその2世たちが大きな購買層になる人口構造の変化、通信販

売やインターネットなどの新しい販売チャネルの導入，グローバル化の進展による新規参入，などの影響で市場の環境が急激に変化してきている。

このような変化の激しい新しい市場に適切に対応するためには，コストの削減の必要性，新しい商品の研究・開発，消費者行動に合った販売方法の構築，競争激化に対応するための，新しい経営の仕組み作りなどが重要となる。

## ②技術の変化

技術の変化は，パソコンの普及に見られるようなダウンサイジングへの進展，半導体の発展と利用範囲の拡大，インターネットの出現などに象徴されるように著しいものがある。高度技術の導入は，製造工程や製品自体の小型化やスピード化に効果があるだけでなく経営管理や経営手法そのものに大きな影響を与えている。

## ③顧客の変化

生活のレベルがあがった消費者が，個性を主張し本物を求め始めた結果，消費行動が大きく変わり，顧客の質も急激に変化している。顧客の変化は，企業の商品供給体制を経営資源に基づく生産体制に合わせた固定的な少品種大量生産型から顧客の要求に合わせた柔軟な生産，つまり多品種少量生産型に変更せざるを得なくなっている。

## ④競争の変化

空白の 10 年に象徴されるように，先行き不透明な経済状態，各種規制緩和などによる競争条件の変化や，前述した市場，技術，顧客などの変化から，競争そのものの形が変わってきている。また，パラダイムシフトが起きているため，従来と同じ土俵での競争激化ではなく，まったく新しい場での新しい競争状況が出現している。

特に，顧客の変化が新しい市場での新しい型の競争優位を確立することを求めている。

## ⑤ 経済構造の変化

バブル経済の崩壊後、土地価格の下落や株価の暴落により担保価値が激減し、借入れができず資金繰りが悪化し結果として、倒産企業が増加し不良債権問題が表面化した。その結果、銀行などの経営破たんが起き、経済構造の変化がはじまった。特に土地や有価証券の担保価値がなくなり中小企業のファイナンスに問題が生じ、従来型の金融システムが崩壊した。その救済に国も中小企業支援政策としてベンチャー育成を掲げた。現在では民間による多くのベンチャーキャピタルが創設され、ベンチャーに投資し育成している。

## ⑥ 為替の変動

為替の変動は、輸入業者、輸出業者の収益に直接影響を及ぼす。その意味では為替の変動に対応できる体制が重要になる。

## ⑦ 責任の変化

「企業は誰のものか?」「企業の社会的責任とは?」など論議がなされる時代になり、また、株主代表訴訟やPL法などの法制度整備により、その責任範囲が急速に広がっている。つまり、企業内の業務活動を改善するなどの企業システムに対する内向きの責任から、地域社会や顧客などに対する外向きの責任へと拡大されている。

## (1) 投機的リスク

損失と利益の両方を発生させる可能性のあるリスクであり、一般に保険を掛けることによって損失をカバーできるタイプのリスクではなく、経営活動の結果によって実現するリスクである。したがって経営資源の獲得とその戦略的配分にかかわるヒト・カネ・モノ・情報・トキ・企業文化などは投機的リスクである。

この投機的リスクは、経営活動に直接かかわってくるリスクであるため、経営をいかに成功させるかという側面が重視され、失敗を避け、その結果成功するというリスク・マネジメントの側面が軽視されがちである。

## (2) 発展段階リスク

企業の設立からビジネスが軌道にのり発展していく各段階で、次のような特有のリスクが発生する。

### ①創成期の複合リスク

企業の設立時で、通常少ない資本と人材で新しい技術やノウハウ・アイデアを打ち出して船出する段階である。当然ながら、いくつかのビジネスとしての基本的な事柄が無視されており、あらゆる種類の経営リスクが存在しているという意味で、「複合リスク」であり、これらをいかに制御・管理するかが成長のカギとなる。

### ②成長期のリーダーシップ

ノウハウ・アイデアの商品化に成功し、ビジネスが順調に進み出す段階である。ヒト・カネを総動員して商売することになり、経営者が総務、営業、財務そのほか、ほとんどすべての業務に責任を持つことになり、リーダーシップ・リスクが起こる可能性がある。

この時点でのリーダーシップ・リスクの対策は、経営者が総合的な視野を持つことと将来の成長のための組織化を図ることである。

### ③経営管理による成長期の管理リスク

ビジネスの成功をおさめ利益が出る時期であるため、需要を満たすことに目が奪われ、組織や管理上の弱点を見過ごしがちである。この時期は原価管理や経営資源管理、予算編成、部門別の業績評価、管理会計などに重点を置く必要がある。

### ④権限委譲による成長期の統制リスク

企業規模が大きくなり、管理職のモチベーションをあげるために権限委譲による成長の持続が行われる。この段階で重要なのは、企業のビジョンが権限委譲された経営陣・管理職に明確に理解され実行されているかであり、事業部制

を編成するなど分権化による分散化した管理体制をいかに統括できるかである。

#### ⑤ 協調による成長期の官僚化リスク

地域的にも企業規模が拡大し、地方分権化が進むと組織の統一管理が困難になり官僚化が進行する。効率経営を妨げる官僚化を防ぐために、企業の原点に戻り、組織全体が一丸となり活性化するために、管理職が協調・協力して動けるような、管理職教育を含めた仕組みの構築が重要になる。

#### ⑥ 組織活性化による成長期の不活性化リスク

企業の巨体化・組織化が進むと、柔軟で創造的な経営の維持が難しくなると共に、組織の官僚化・保守化を含め構造上の無駄が著しく増え、組織に不活性化リスクが生じる。

組織の一員として、個々人が自ら進んで効率的な業務活動ができるような職場環境や仕組み作りが、古い体制を変えるための課題となる。

#### ⑦ 上場による成長期の株式公開リスク

少数株主による私的企業から多数株主による公的企業への成長は、どの企業も目指す目標の一つである。飛躍的な成長が望まれる一方、利害関係者が増え自由奔放な企業経営に歯止めが掛かることになる。この段階は、利害関係者へのディスクロージャー（情報公開）やその社会的責任が問題となる。

### （３）企業倒産のリスク

企業が経営破たん陥ると「倒産」という状態になる。この倒産の形態には、破産、特別清算、商法の会社整理、会社更生法、民事再生法の適用などがある。倒産の理由には以下の事項がある。

- ① 経営の失敗……商品開発の失敗，売り上げ低迷
- ② 相手企業の経営の失敗……債権先の倒産または経営破たんにより債権回収ができない場合
- ③ 不慮の災害による損失……地震や水害，事故

#### **(4) 株主代表訴訟のリスク**

社長に権限が集中するなど、取締役会の形骸化は企業文化・経営風土に帰する問題であるが、それに一石を投じているのが株主代表訴訟である。

代表訴訟では、取締役が損害賠償しなければならない場合以外にも、社会に対する企業イメージのダウンにつながる場合や、経営陣が退陣に追い込まれリーダーシップに問題が起きることで、社員の動揺などを引き起こし経営活動にマイナス影響を及ぼす懸念がある。

#### **(5) 純粹リスク**

自然災害のような、それが現実には起こった場合には損害のみを発生させるリスクを「純粹リスク」と定義する。その例として

- ①地震や水害などの自然災害
- ②戦争・紛争など
- ③海外駐在員などの不慮の事故（誘拐・交通事故）

これらのリスクに対しては、保険を掛けてある限り、損害の全部または一部がカバーできる。そこで、インシュランス・カバレッジ（付保率・保険担保範囲）が、重要なリスク・マネジメントのポイントとなる。

経営環境が変化する時代には、特に企業を取り囲むリスクをいかにコントロールするかが重要となる。そのためには適切な仕組みが必要であり、その仕組みを動かすヒトがカギを握ることになる。

リスク・マネジメントの適切な実行は、企業の存続にかかわるだけでなく、経営者や従業員の安全、さらには社会全体にも大きなダメージを回避できる可能性がある。このような意味でも、企業の社会的責任といっても過言ではない。

#### **(5) -1 リスク・マネジメントの定義**

一般にリスク・マネジメントとは、経営リスクをいかに管理するかの仕組みを

意味する。そこでリスクの表現の仕方によって危機管理・危険管理・リスク管理などの呼称が使われる。

ここでは、リスク・マネジメントの定義を「保険や安全対策、さらには経営戦略などを活用して事業の偶発的あるいは人為的な損失（リスク）を発生しないようにし、もしリスクが発生した場合には、それを最小化し、さらに実現したリスクに適切に対処する経営管理の方法」とする。

## **(5) -2 経営リスクを避けるには**

経営のリスクを避けるには、

- ①関係者がリスクに対してそれが起きるかもしれないことを常に認識すること。
- ②起きるかもしれないリスクに対して前もって理解し、いざという時の心構えができていること。
- ③リスクが起こるかもしれないとき、起きてしまったときの、さらに起きないように予防するための実際の行動ができる仕組みとその行動に精通していること。

つまり、経営リスクとその対応の仕方を認識して理解し、自分のものにしていくことが必須の心掛けである。

## **2.2 イノベーションの定義**

近年、イノベーションという言葉が氾濫している。いろんな場面で使われているが、イノベーションの定義はまちまちである。イノベーションという言葉を最初に定義したのは、オーストリアの経済学者シュンペータ(1883～1950)である。彼は、その著書「経済発展の理論」(1912) [7]で、経済発展は、人口増加や気候変動などの外的な要因よりも、イノベーションのような内的な要因が主要な役割を果たすと述べており、イノベーションの例として、(1)創造的活動による新製品

開発，(2)新生産方法の導入，(3)新マーケットの開拓，(4)新たな資源(の供給源)の獲得，(5)組織の改革などをあげている。また，いわゆる企業家(アントレプレナー)が，既存の価値を破壊して新しい価値を創造していくこと(創造的破壊)が経済成長の源泉であると述べている。また，P.F.ドラッカー(1909～2005)はその著書「イノベーションと企業家精神」(1985) [8]で，「イノベーションは，焦点を絞り，シンプルにしなければならない」また，一つのことに集中しなければならない。さもなければ焦点がぼける。シンプルでなければうまくいかない。

イノベーションに対する最高の賛辞は「なぜ自分には思いつかなかったか」とした上で，ドラッカーのイノベーションとは「既存の知識，製品，顧客のニーズ，市場など，すでに存在するものを，はるかに生産的な一つの全体にまとめる上で必要な小さな欠落を発見し，その提供に成功することである」。したがって，イノベーションが機会をもたらす領域を発見するには，すでに可能になっているにもかかわらず，欠落したままになっている致命的に重要なものは何か，経済的な効果を一変させるものは何かを問わなければならないと述べている。

P & G の A・G・ラリー(会長兼CEO)は著書「ゲームの変革者(The Game-Changer) (2009) [9]」の中で，イノベーションとは“どの会社にも，よって立つ原則というものがある。その原則に基づいて決断を下し，問題に立ち向かい，機会を生み出す”と述べ，また，イノベーションは，経営の根幹をなし，目標設定，戦略作成，組織構成，経営資源の配分，予算策定，リーダーの育成など，重要な意思決定を下す基盤となるとしている。

一方，日本国内では，イノベーション研究では先駆的な一橋大学イノベーション研究センターではイノベーションを「経済効果をもたらす革新」[10]と定義している。また，流通業の立場から木村達也は著書「流通業イノベーションの発生要因 [11]」の中で「生産のフロンティアを拡張させる経済的意味を持つ事象」と定義している。製造業の立場からヒロボー社長松阪敬太郎は著書「広島ものづくり先進企業 [12]」の中でイノベーションとは「内から変化をさせる」ものであり「既存のものに新しい価値，富や価値を創造する」ものであると定義し実践している。クレイトン・クリステンセンは著書イノベーションのジレンマ[13]では「破壊的イノベーション」によって既存の優良企業はそれまでの成功体験が足かせとなって追い詰められていくとしている。また，続編イノベーションの解[14]では

「破壊される側ではなく破壊者となる」ことによって成長路線がとれるとしている。

以上のように定義のさまざまを述べてきたが，当論文での定義は「新しいものに挑戦し，既存のものに新しい価値や創造を生み出す仕組みや方法を形成する行動」とする。

## 2.3 テキストマイニング技術について

### 2.3.1 テキストマイニングの研究の経緯

近年，電子化された文書（テキストデータ）は増加の一途をたどっており，また，WWWの普及にあいまってWEB上に公開されているレポートなど手軽に入手できる環境となった。一方，企業ではイベント開催時のアンケート，商品・サービスに対するクレーム，営業日報，製造日報，社内文書など日々発生するデータは電子化され関係部署に配布されている。しかしながら，配布先ではその膨大なテキストデータから迅速に欲しい情報を探したい，あるいは分析し傾向をつかみたいといった，利用者のさまざまな要求に対応できる情報アクセス手段はまだ乏しいのが現状である。

ところが，一部の先進企業ではテキストデータからクレーム対策や新規製品開発，次期企画などに利用し成果に結び付けている事例が出始めた。これらは，テキストデータが知識情報として，企業の関心の高さを反映し，さらに，蓄積されたテキストデータを分析し内容を瞬時に把握したいというニーズが一層高まりつつあることを物語っている。これらの解決手段として，自然言語処理技術を応用したテキストマイニングと呼ばれる分野が注目され始めた。

テキストマイニングあるいは，テキストデータマイニングという言葉が，論文に頻繁に登場するようになったのは，1997年ごろである。日本においては1999年ごろから論文数が増えはじめ研究が本格化し始めたことが窺える。

テキストマイニングは，データマイニングから派生した研究分野である。デー

タマイニングが、通常、構造化されたデータベースに格納された情報を抽出し、分析しているのに対し、テキストマイニングは構造化されていない、つまり、特定のデータ構造を持たない通常の文書を対象にしている。

このことから、狭い意味ではデータマイニングに、構造化されていないテキストデータのデータ化（構造化）というプロセスを加えたものだと見ることが出来る。分析の目的、適用する分析手法、分析結果の見せ方、対象とするデータによって、データ化の仕方が違ってくる。一方、広い意味では自然言語処理技術により、文書集合から分析者が気づかなかった知識や情報を得ることだといえる[15, 16]。

先にも述べたように、テキストマイニングはデータマイニングから派生した研究分野である。初期の研究では、データマイニングの技術あるいは手法をテキストデータに適用するというアプローチが取られた。近年では、さまざまな分野の研究者が、さまざまなアプローチや要素技術を用いて、テキストマイニングという新しい領域に参入してきており、研究領域も拡大している[15,17,18,19,20,21, 22,23,24]。

自然言語処理の研究者は、文書分類や情報検索を発展させて、文書集合から新たな知識発見に結び付くような研究に取り組んでいる。一方、データマイニングの研究者は、大規模データから思いがけないパターンを発見するという従来のデータマイニングを、数値データだけでなくテキストデータにも適用するという方向で発展させている。両分野の融合は、今後ますます進んでいくと思われる。

このような状況の中、テキストマイニングに関する研究がいくつも提案されている。タスクとしては、文書間や単語間の関係表示、日報や問い合わせ分析、経済市場の予測、WWW検索、単語間の相関ルールや単語の時系列パターンの抽出などが取り扱われており、技術としては、情報抽出、文書クラスタリング、文書分類、可視化、相関ルールなどが用いられている。

次に、個々の研究について、タスクごとに概要を紹介する（表 2.1～表 2.3 参照）。

表 2.1 テキストマイニング関連研究 [15]

タスク	対象文書	適用技術	研究事例
文書間の関係表示	一般文書	情報抽出	研究事例 1 研究事例 2
	新聞	文書クラスタリング、 可視化	研究事例 3
	論文	文書分類、可視化	研究事例 4
	Newsgroup	自己組織化、可視化	研究事例 5
単語間の関連表示	一般文書	可視化	研究事例 6
日報分析	営業日報	情報抽出、文書分類	研究事例 7
問合せ分析	コールセンター問合せ	情報抽出、関連ルール	研究事例 8
経済市場の予測	WWW、新聞	関連ルール	研究事例 9
WWW検索	WWW	WWWページ関連付け	研究事例 10
		検索結果の分類	研究事例 11
単語間の 関連ルールの抽出	一般文書	関連ルール	研究事例 12
			研究事例 13
単語の時系列 パターンの抽出	一般文書	関連ルール	研究事例 14

表 2.2 テキストマイニング関連研究の概要 (1/2) [15]

研究事例	研究概要
研究事例 1	5W1Hの情報に着目して文書から概念を抽出し、この情報を検索や文書分類に用いる事により、検索ノイズを削減し理解しやすい分類軸（イベントの時間軸に沿った遷移や比較）に分類する。
研究事例 2	パターンマッチングで広範囲の文書を高速にフィルタリングし、情報抽出でより精度の高い分類にリファインする。異なるユーザの要求に高い精度で適合する文書分類手法である。
研究事例 3	明示的な文書間の関係がない新聞記事どうしを、単語ベクトルにより計算される類似性に基づいて関係付け、グラフ構造として表現する。記事のスレッドやトピックスを表現する単語が表示される。記事の入力を時間順に並べるという制約を用いるために、少ない計算コストで実現できる。
研究事例 4	特定のドメインの論文間の関係を視覚的に表示し、サーベイの記述を支援することを目的とする。論文の引用部分を対象とし、他の論文を参照する部分の認識や参照関係の特定を、単語n-gramの出現頻度を用いて半自動化を実現する。
研究事例 5	自己組織化マップアルゴリズムに基づき文書集合を組織化して、2次元マップを作成する。類似する文書はマップ上の近い位置に配置される。

表 2.3 テキストマイニング関連研究の概要 (2/2) [15]

研究事例	研究概要
研究事例 6	単語間の連想関係を2次元マップ（ネットワーク図）として可視化することにより、文書群全体が持つ特徴・傾向を分析する手法。単語マップに出力される単語集合は、単語に付与された属性情報や単語の接尾文字列パターンを用いて絞り込まれ、また、連想関係計算に使用する文書集合は、文書に付与された属性情報や文書検索の結果を用いて絞り込まれる。
研究事例 7	営業管理職向け意思決定支援を目的として、営業日報から成功事例・機会損失を分析する。句と句や文と文の構成文に解釈するのは難しいので、製品名やキーとなる概念を個別に抽出し、次に、抽出されたどうしを組み合わせ、因果関係を有する構造を抽出するという2段階プロセスにて情報抽出を行なう。
研究事例 8	従来のクラスタリングの手法では、文書内の名詞句を中心とするキーワードを文書の特徴として扱うのが多いのに対して、当事例では、意図を含んだ情報を抽出するために、モダリティ情報と構文情報を利用している。係り受け関係を想定した2語組に関して、重要語ラベルと述語・評価ラベルとの評価パターンを用意して情報抽出を行なう。
研究事例 9	WWW上の新聞記事の情報を使って経済市場の予測を行なう。記事には、株価の数値データの動向だけでなく、発生理由も書かれているので、これらを使用することにより精度の高い予測が行なえる。テキストにはリード記事を用い、特定のフレーズの出現頻度とそのフレーズが現れるカテゴリ（株価の上昇、下落、横ばい）の区別を考慮した重みと、株価の値動きとの相関を学習する。
研究事例 10	複数の情報源を統合するために単純な表現を利用し、その表現にWWWページを関連付ける方法である。ドメインにおける知識表現や検索のためのインデックスページを作成するための機械学習、ユーザの検索要求をプリミティブに分解・実行するプランニングが用いられている。
研究事例 11	検索エンジンが返す膨大な結果を、ユーザの質問に対応した意味のあるグループに動的に分類する方法である。ユーザの質問を分類するために、質問のタイプやカテゴリのタイプについての知識からなる質問モデルを備えることと、文書にカテゴリを与えるために、個々の用語を一般的な概念にリンクする知識からなる専門用語モデルを備える。
研究事例 12	テキストデータの属性値間の相関ルールを求めるアルゴリズムの研究である。条件部・結論部はそれぞれ各文書に特定のキーワードが存在するかどうかを表す。キーワードの属性、あるいはキーワード間の関係情報を外部から与えることが可能であり、通常の高確信度・サポートによるフィルタリングに加え、条件部・結論部に出力されるキーワードをこのカテゴリによって絞り込むことを可能にした。
研究事例 13	研究事例12を発展させ、フレーズ間の相関ルールの抽出を可能とした。品詞の組み合わせにより、フレーズ候補となる単語列を選択し、接続する単語間の相関係数がしきい値より大きいペアを組み合わせることでフレーズ抽出を行なう。抽出されたフレーズには、統計量によるフィルタリングが適用され、マイニング結果として意味を持たないものが除去される。従来の単語ベースの方法では、意味のない相関ルールが大量に出力されることがあったが、フレーズの抽出およびフレーズ分類体系の定義によって、役に立つ情報を効率的に得られることが可能となる。
研究事例 14	順序データから、エピソードおよびエピソードルールを発見するアルゴリズムを適用し、テキスト情報を単語の系列データとしてとらえ、テキストマイニングを行なう手法。一定回数以上出現するエピソードおよび一定確率以上のエピソードルールを抽出し、さらに品詞情報の組み合わせによるフィルタリングを行なうことで、フレーズの抽出を可能とする。

テキストマイニングの研究が盛んになるにともない、関連するソフトウェアやソリューション・サービスも提供され始めている。企業の WWW ページや商品紹介パンフレットなどを通じて得た情報を基に、実用化されているツールの機能を紹介する。

テキストマイニングツールは、データマイニングツールからの拡張と情報検索ツールからの拡張の2つに分類できるが、どちらも装備されている機能は以下の9つにまとめられる[15,25]。

#### (1) 単語や概念の抽出

文書中の単語や概念を抽出して、リスト表示する。品種や種類を指定することができる。この機能を用いると、文書中で注目したい固有名詞などを抽出することができる。テキスト分析の基本機能であるが、単独で有効な機能というより、前処理的な機能である。

#### (2) 単語や概念マップの生成

抽出した単語や概念を単にリスト表示するのではなく、単語や概念間の意味的關係を考慮して、マップの形でグラフィカルに表示する。単語を手がかりにして、文書の全体を把握することができる。例えば、アンケートの自由記述文から全体の印象や大雑把な傾向を把握し、また新聞の話題変化などを把握するには有効と思われる。マップの解釈は人間に任される。

#### (3) 単語や概念の集計

抽出した単語や概念を集計する。例えば、アンケートの自由記述文から単語を切り出して、集計する。関連語をユーザ登録できるものもあるが、基本的には、意味的に同じ表現でも表記が異なれば別の表現として集計してしまう。

#### (4) 類似文書検索

与えられた文書に類似する文書を探す。この機能を用いて、FAQ 文書やマニュアルから、電子メールによるユーザの問い合わせに対する応答を自動生

成する機能を持つものもある。通常の検索では、適切な検索語を指定することが難しいという問題があるが、類似文書検索では、欲しい文書に近い文書を指定するだけで、似たような文書を探ることができる。

#### (5) 文書要約

与えられた文書を要約する。文書の長さ（文字数や圧縮率）の指定や、必要・不必要な単語リストの指定により、ユーザの好みに応じた要約を生成できる。本文を読まずに要約だけで内容を理解できるレベルには達していないが、読むべき本文を探すための手段として有効である。

#### (6) 文書分類（カテゴリゼーション）

あらかじめ用意したカテゴリに、文書を分類する。例えば、新聞記事を、スポーツ、政治、経済など、設定した分野に自動分類できる。分類先が分かっている場合に有効である。

#### (7) 文書クラスタリング

文書分類と異なり、カテゴリを用意せずに、文書をクラスタリングする。予期しない分類を発見する可能性があるが、生成されたクラスタの意味付けは人間に任されるので、必ずしも期待するクラスタが得られるとは限らない。

#### (8) シソーラスによる検索後展開

シソーラスを使って、検索語に異表記（ex.コンピュータとコンピューター）、同義語（ex.パソコンとパーソナルコンピュータ）、関連語（ex.パソコンとメモリ）などを追加する。これにより、検索もれを抑えられるが、誤った単語を追加すると、逆に検索精度の低下を招いてしまう。

#### (9) 検索結果のグラフィカル表示

検索結果をグラフィカルに表示して、注目する文書へのアクセスを容易にする。

## 2.3.2 テキストからのキーワード抽出法

読者は、文書の作成者がその文書を通して表現したかった内容を、文書中に現れるキーワードを介して的確に把握する。そのために、文書の中から作成者の主張の内容を表す重要なキーワードの抽出が必要となる。

従来のキーワード抽出の代表的な手法を次に示す[26,27,28].

- (1) 文書の見出し情報を用いるキーワード抽出法
- (2) 文書の中、段落中での位置情報を用いるキーワード抽出法
- (3) 自然言語解析を用いるキーワード抽出法
- (4) 手がかり表現を用いるキーワード抽出法
- (5) 統計量に基づくキーワード抽出法

次にその概要と本研究での適用評価を示す。

### (1) 文書の見出し情報を用いるキーワード抽出法

タイトルや見出しは、文章のポイントを簡潔に要約したものとなっていることが多い。そこで、タイトルや見出しの中の前置詞や冠詞以外の語をキーワードとする方法が考えられる。しかし、本研究で扱う文書は製造現場の担当者が瞬時に発生した事故、すなわち非常事態を緊急に伝えるものであるため決まった型をもたないので、タイトルとか見出しの情報だけでは十分なキーワード抽出は不可能であり、この考え方は本研究には適用できない。

### (2) 文書の中、段落中での位置情報からのキーワード抽出法

文書は、ジャンルに依存して、ある程度構造に規則性を持っている。新聞記事では記事の冒頭部に重要な語が多い。新聞記者は文章をそのように作るように訓練を受けているのでこの傾向は一般人の文書にあてはめることはできない。

実際、文書の種類によって主要部分の位置はさまざまで、重要な意味分類が少数の大段落に集中するとしてもそれがどの段落であるかは筆者の章立ての癖などに依存する。本研究で扱う文書も決まった型をもたないので、この考え方は本研

究には適用できない。

### (3) 自然言語解析を用いるキーワード抽出法

自然言語解析によって文書中のどの語が重要であるか判断できるならば、正確なキーワード抽出が実現可能となるが、文法ルールに正しく従うとは限らない文章から自然言語解析によって要点を的確に取り出すのは現時点で困難である。また、重要な語は太字で書かれているとか、重要な語の前に‘これは重要な…’などの前置きがあるという期待も、一般にはできない。本研究で扱う文書もその例外ではないので、この考え方は本研究には適用できない。

### (4) 手がかり表現を用いるキーワード抽出法

文書中の重要箇所を指示すると考えられる手がかり表現の語、接続詞などの手がかり語情報を自然言語解析に基づいて解析し、重要なキーワードを抽出する研究はいくつかある。その一つの研究では、理由、詳細化など文間の種々の結束関係を示す手がかり表現、同一・同義の語・句の出現、二文間の類似性、の三つの表層表現中の情報に着目して文章構造を分析しキーワードを抽出している。しかし、本研究で扱う文書は、簡潔を旨とした報告文書であり、接続詞の情報、重要箇所を明示するような表現に乏しい。したがって、この考え方は本研究には適用できない。

### (5) 統計量に基づくキーワード抽出法

情報検索分野では、文書中に頻繁に出現する内用語は文書の主題を示す傾向があるという仮定に基づいた研究が多数行われてきた。しかし、頻出語が文書の独自の主張を表現する語となることは実際には少ない。このような出現頻度だけではなく、文書中のあるキーワードの出現頻度と、その分野での平均出現頻度との相対比率をキーワードの重要度とするアプローチもある。一方、さまざまな単語の重み付け技法として、文書中で出現頻度の高い名詞をキーワードと考える TF 法、また、これに合わせて、出現する文書数も考慮することで、その文書固有の出現の度合いを計算する TF.IDF 法などが提案されている。TF.IDF 法では、単にキーワード出現回数の多さだけで重要度を定める手法よりは精度が高くなると報告さ

れている[29].

本研究で扱う文書は、前述のごとく製造現場の担当者が瞬時に発生した事故、すなわち非常事態を緊急に伝えるものであるため、新聞記事や学術論文といったいわば「定型」ではなく、まったく決まった型を持たない「不定形」である。したがって、定型の文書に現れる見出し、手がかり語、接続表現、位置情報などを対象とした手法は適用できないため、統計量に基づくキーワード抽出法を適用することにした。

最近、統計量に基づくキーワード抽出技術の中で、文書は著者独自の考えを主張するために書かれるという仮説をもとに、文書全体はその主張を目指して一つの流れを形成するという考え方から、文書を建物にたとえて説明できるとするキーワード抽出法「KeyGraph（以下キーグラフと記す）」[26,30,31]が注目されている。

この手法は、本研究の対象となる文書のような主張が明確な文書においては、TF.IDF法より高い精度で、重要なキーワードが抽出されたことが、次に示す表2.4のように報告されている。また、本研究でマイニングするテキストデータは、工程別、工場別、組織全体、専門家によるグループ討議議事録などの複数の文書を必要に応じて、一つの文書として扱うなどの理由から、複数の文書からキーワードの特徴を求めるというTF.IDF法は本研究には適さない。キーワードの抽出として「キーグラフ」のアルゴリズムを適用することにした。「キーグラフ」の詳細は2.4節「キーグラフの概要」で示す。

下記の表の結果[26]のように、従来の方法TF.IDF法より良好な結果が得られている。

表 2.4 TF.IDF法とキーグラフ法の性能比較

	TF・IDF法	キーグラフ法
suff	73.9%	86.4%
necc	66.5%	88.4%

①suff : 抽出されたキーワードの十分さ

$$\text{suff} = (A \cap K) / A$$

②necc : 抽出されたキーワードの必要度

$$\text{necc} = (A \cap K) / K$$

A : 著者の主張を表すキーワード集合

K : キーグラフによって得られたキーワード集合

実験対象文献は、コンピュータサイエンス、医学に関する論文または、その概要 23 件で長さは 200 単語台から 10000 単語台までほぼ一様に選んだ。

文書から重要なキーワードを抽出するまでのプロセスを図 2.3 に示す[15,18,19].

①第 1 フェーズ (データクリーニング)

複数の作成者が記述した文書には、作成者の好みに合った語の使い方があり、キーワードの抽出精度を悪化させるような語[33] (同義語、表記の揺れなど) が存在する。各適用分野別辞書を利用して語の表現を統一する必要がある。

②第 2 フェーズ (文書変換)

データクリーニングによって整形された文書を、キーワード抽出処理が可能なように、形態素解析ツール「茶筌」により、形態素に変換する。「茶筌」については 30 ページで紹介する。

③第 3 フェーズ (キーワード抽出)

形態素に変換された文書から、キーワード抽出ツール「キーグラフ」によりキーワードを抽出する。なお、「キーグラフ」については次項 2.4 節で紹介する。

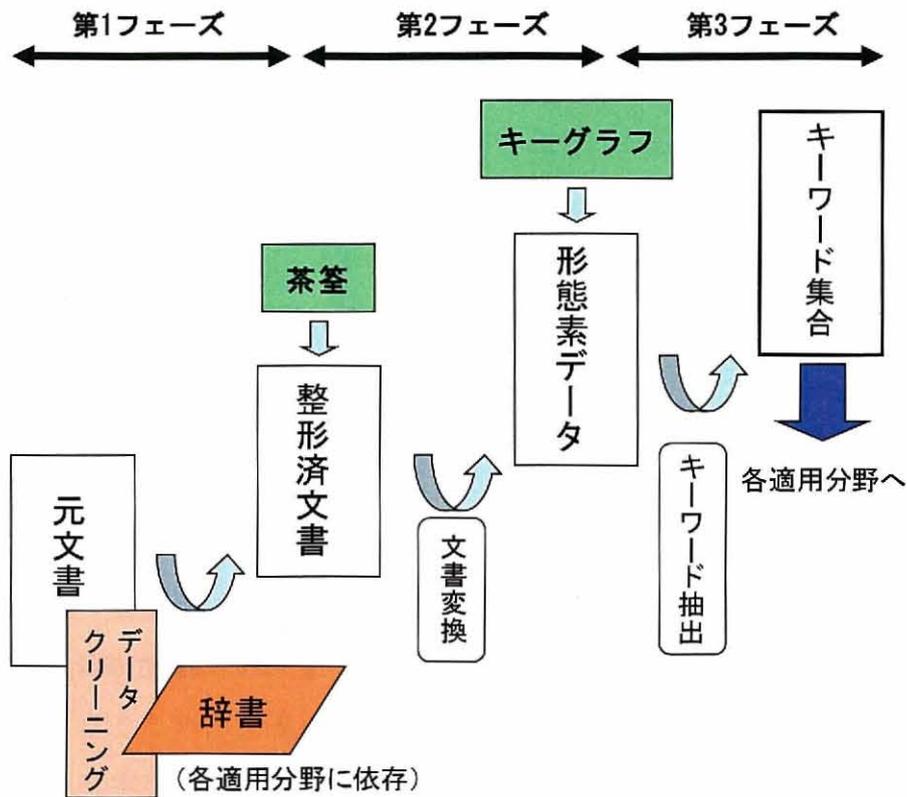


図 2.3 テキストマイニング処理プロセス

テキストマイニングでは、キーワードの出現頻度に関する統計的な分析が基盤となっている。このため、マイニングの精度をあげるためには、表層的に異なっている場合でも、意味が同じであれば、同一キーワードとして扱う必要がある。

表現の統一方法について次に述べる[34,35]。

(1) 誤字／脱字

修正する。

(2) 全角／半角，大文字／小文字

どちらか一方を統一表現とする。

(3) 同義語

統一表現とする。

① 外来語

打ち合わせ ↔ 会議 ↔ ミーティング

電子計算機 ↔ コンピュータ

写真機 ↔ カメラ

②省略語

特急 ↔ 特別急行

弱冷房車 (JR 東日本) ↔ 弱冷車 (JR 西日本)

③通称

首相 ↔ 内閣総理大臣

④年号

2006年 ↔ 平成18年 ↔ 平成十八年 ↔ H18

⑤用語の陳腐化

「女中」 → 「お手伝い」「ハウスキーパー」などに置き換えられており  
現在では差別語にまでなった。

⑥立場による言葉の違い

「公的資金」 ↔ 「血税」

(4) 表記の揺れ (同義語のうち発音も同じ)

統一表現とする。

①漢字と仮名による表記の揺れ

犬, イヌ, いぬ

②漢字表記の揺れ

沈殿, 沈澱

超電導 (JIS)

超伝導 (学術用語)

③外来語の表記の揺れ

インターフェース (新聞)

インタフェース (JIS)

インターフェイス (学術用語)

④古い記事を扱うときは異字体が問題

國語, 国語

⑤送り仮名の違いによる表記の揺れ

行う，行なう

打ち合わせ，打ち合せ，打合わせ，打合せ，打合

⑥句読点の書き方にも表記の揺れ

句点 (.) または (.)

読点 (,) または (,)

(5) 指示代名詞

該当する名詞に置換する.

本研究で扱う文書の選定が最初の課題となった。事故の状況は、工場の操業日報の中で第一報が報告され、正式な事故は現場で作成されたメモから起票された日報が対象となるため、文書の標準化はいたって困難であった。初期において過去のデータから学び、その欠点を補うことにした。

自然言語は図 2.4 に示すように階層的構造を持っている。

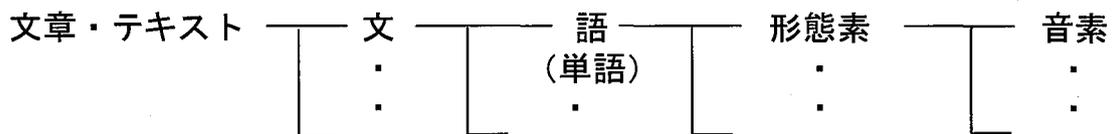


図 2.4 自然言語の階層的構造

文章・テキストはあるまとまった内容を表現するための文の順序付けられた集まりであり、隣接する文相互にはある種の関係性が存在する。文はあるまとまった内容を持ち、形の上で完結した（表記において句点が与えられる）言語単位であり、一つ以上の語（単語）から成り立っている。語（単語）は一つの意味のまとまりをなし、文法上一つの機能を持つ最小の言語単位であり、一つ以上の形態素から成り立っている。形態素は意味を持つ最小の言語単位であり、一つ以上の音素から成り立っている。音素は人間の意味（意志）伝達において音声をどのよ

うに使っているかを基に考えた音の単位である。

このうち、文字によって表記された自然言語においては形態素が最小の単位となるので、形態素解析[36]は自然言語処理の第一段階となる。

### (1) 形態素とは

言語の機械処理において、文字によって表記された文を対象とする場合には、表記に現れた語構成の最小単位を形態素といい、それ以上分割できない最小の言語単位のことである。

日本語の場合には語を区切る空白というものが存在しないため、語の厳密な定義を与えることは困難である。また、形態素という概念は欧米語の言語学から来たものであるため、日本語における形態素の定義もそれほど明確ではない。

形態素の集合である語は、文を構成する際の働きを基準とした場合、自立語と付属語に大別され、さらに細分化されて次頁の図 2.5 に示すように 10 の品詞に分類される。自立語は単独で文節を構成できる語（広辞苑第 6 版）で、動詞、形容詞、名詞、副詞などをいう。一方の付属語は、自立語に付属して文節を構成する語（広辞苑第 6 版）、助動詞、助詞などをいい、それ自体では意味をなさず、自立語と組み合わせさって初めて意味を持つ語をいう。なお、文節とは、文を意味上と発音上から不自然でない程度に区切った最小の言語単位をいう。

一方、語構成の立場から見ると、一つの要素からなる語と複数の要素からなる語がある。複数の要素からなる語については、次の 3 つの結合形態がある。

#### ① 活用語

文中での働きの違いによって語形が変化する語。変化しない部分を活用語幹、変化する部分を活用語尾という。例えば「食べる」では「食べ（活用語幹）」＋「る（活用語尾）」となる。

#### ② 派生語

ある語に付加的要素が付いてできる語。基の語を派生語幹、付加的要素を接辞（接頭辞または接尾辞）という。例えば、「寒さ」では、「寒（派生語幹）」＋「さ（接尾辞）」、「ま冬」では、「ま（接頭辞）」＋「冬（派生語幹）」とな

る。

### ③ 複合語

複数の語が結合して1語となったもの。「本+棚」、「うれし+涙」、「自然+言語+処理」など。

日本語においては前述の結合形態の各構成要素がほぼ英語の形態素に相当するものである。しかし、接尾辞と助動詞の区別は明確でなく、また、付属語（助詞、助動詞）を語と認めない考え方もある。さらにどのような基準で複合語を1語とみなすかも明確ではない。しかし、工学的に見たときの形態素解析の目的は、入力文をとにかく辞書中の項目の組み合わせに分解することである。

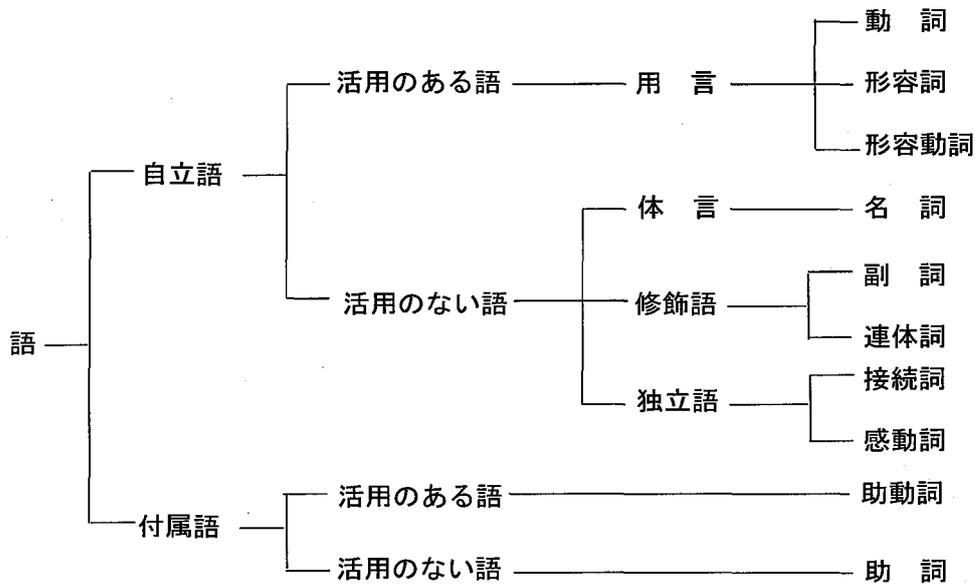


図 2.5 日本語の品詞分類

### (2) 形態素解析とは[37,38,39]

形態素解析とは、文を適切な形態素（それ以上分割できない最小の言語単位のこと）に分割する処理のことである。例えば、「私は大学院で学ぶ」という文では、計算機にとって見れば、ただの記号列にすぎない、そこで、この文を「私」「は」

「大学院」「で」「学ぶ」という形態素に分割することで、それぞれの形態素に意味を与え、構文分析や意味解析などへと処理をつなげる。英語などの、欧米言語では単語が空白によって区切られているため、単語の認識は大きな問題にならないが、日本語のように単語の間に区切り記号を持たない言語は、単語の認識が根本的な問題になる。単語の識別、活用語処理、品詞の同定は個別の処理とみなされることもあるが、日本語ではこれらの処理が密接に関連するため、同時に行うのが普通で、全体的な処理を形態素解析と呼んでいる。

形態素解析の問題点としては、例えば、その解析時間の問題や適切な分割による単語の認識があげられる。莫大な辞書を用意してどのような文でも解析が可能ということにしても、時間が掛かれば非効率的であり、時間が掛からないからといって適切でない分割をしてしまうのは片手落ちである。

したがって、その解析をより効率的に行なうために、より曖昧さを減らすために、さまざまな研究が行われている。さらに、形態素解析には辞書が必要不可欠であるが、すべての語を網羅することは不可能である。その、辞書に記載されていない単語が文に含まれていない場合の処理なども問題点の一つである。

形態素解行うためのツールとしては JUMAN が有名である。JUMAN は京都大学長尾研究室および奈良先端科学技術大学院松本研究室によって開発された日本語形態素解析システムのことである。また「茶筌 (Chasen)」は JUMAN (version2.0) をもとに、奈良先端科学技術大学院松本研究室で開発されたもので、現在フリーソフトとして流通しており、最新版は Ver.2.1 で UNIX 版、Windows 版が提供されている。本研究では開発環境が Windows であることと、性能が JUMAN より優れていることから日本語形態素解析システム「茶筌 (Chasen)」を用いることにした。

### (3) 「茶筌 (Chasen)」とは

奈良先端科学技術大学院松本研究室で開発されリリースされた、フリーソフトウェアの日本語形態素解析器である。辞書項目は約 23 万語、解析速度は Linux で約 40,000 文字/秒 (茶筌 2.0) であり 100M バイトの新聞記事 1 年分の全文を約 20 分で解析可能である。茶筌は他システムからの利用も可能であり、適用範囲の広がり構成している。また、Windows 版 (WinCha2000) ではインター

ラクティブ機能の提供と，Windows Excel（以下 Excel と表示）とのデータ互換性により，日本語形態素解析が容易にできるようになった。

次に図 2.6 に Windows 版茶釜「WinCha2000」の処理画面を示す。

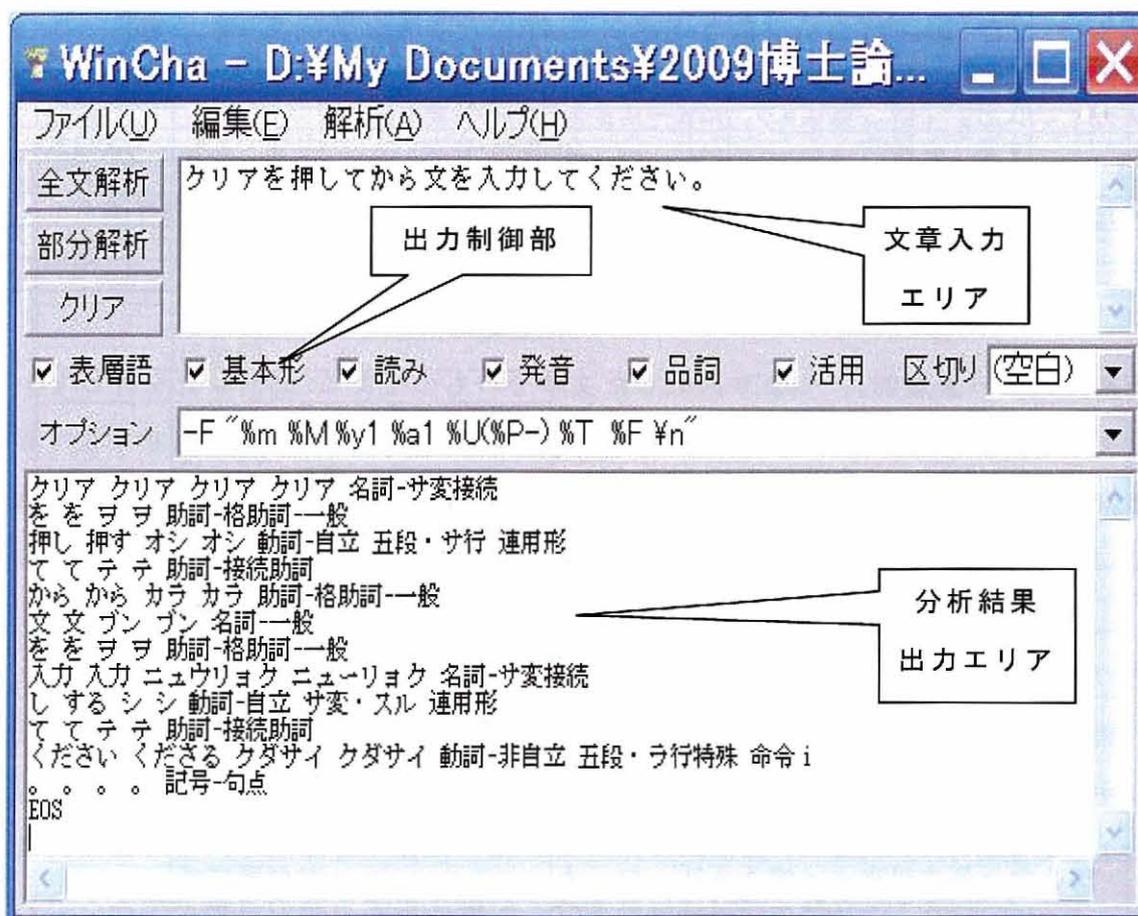


図 2.6 Windows 版茶釜「WinCha2000」の処理画面

画面に「クリアを押してから文を入力してください。」とある窓が，分析したい文章を入れる文章入力エリアである。分析結果は下の大きな窓，分析結果出力エリアに出てくる。出力制御部として「表層語」「基本形」「読み」「発音」「品詞」「活用」と6つのチェックボックスがある。図 2.6 は6つのチェックボックスを選んで「全文解析」のボタンを押したときの例である。分析結果は可変長で表示される。表 2.5 は Excel で加工した，分析結果である。

表 2.5 茶釜による分析結果

表層語	基本形	読み	発音	品詞	活用	
クリア	クリア	クリア	クリア	名詞-サ変接続		
を	を	ヲ	ヲ	助詞-格助詞-一般		
押し	押す	オシ	オシ	動詞-自立	五段・サ行	連用形
て	て	テ	テ	助詞-接続助詞		
から	から	カラ	カラ	助詞-格助詞-一般		
文	文	ブン	ブン	名詞-一般		
を	を	ヲ	ヲ	助詞-格助詞-一般		
入力	入力	ニューリョク	ニューリョク	名詞-サ変接続		
し	する	シ	シ	動詞-自立	サ変・スル	連用形
て	て	テ	テ	助詞-接続助詞		
ください	くださる	クダサイ	クダサイ	動詞-非自立	五段・ラ行特殊	命令 i
。	。	。	。	記号-句点		
EOS						

最初の列が「表層部」、つまり入力した文章そのものである。「クリアを押してから文を入力してください。」という例文を分析したので、これを語の最小単位（形態素）に分けた結果が一行ずつ表示されている。最後の行に「EOS」とあるのは End Of Sentence、つまり一文の終わりを意味する。次に「基本形」の列であるが、これは活用する前の形である。「押し」という語は「押す」の活用形なので、そのもとの形—基本形が表示されている。次の「読み」「発音」の列は、それぞれの語をどのように読むか、ということで、「入力」は「ニューリョク」と読み、発音する際は「ウ」を延ばして「ニューリョク」となることがみてとれる。「品詞」および「活用」の列は、字義どおり、その語はどのような品詞で、どのように活用するかを表示している。「押し」は動詞で、5 段活用する語の連用形であることが示されている。

これらの分析結果は、メニューバーの「ファイル」から、名前を付けて保存することができる、また、「編集」からクリックボードにコピーすることもできる。各列の区切り方は、デフォルトで「タブ」になっているが、カンマや空白も選ぶことができる。タブやカンマで区切りを入れておけば Excel で読み取ることができる。また、「部分解析」ボタンを使うことにより、文章入力エリアで数行選択し解析することもできる[40]。

茶釜で解析され出力された形態素を使って、テキストマイニングを行う場合、解析された形態素のどの「品詞」が求めるキーワードとして適用できるかを見極めることが重要である。茶釜では日本語の品詞を最大 5 階層、534 種類に分類し

ており適用には知識が必要である。次表 2.6 は品詞情報からキーワードとして適切かの判断例であり、図 2.7 は品詞情報を基にキーワード抽出を行う手順であり、筆者のキーグラフ適用前の手法でもある。

表 2.6 品詞情報からキーワードとして適切かの判断例 [41]

品詞区分	詳細	キーワードとして適切かどうか
記号		キーワードとして意味を持ちそうにない
形容詞	自立	キーワードとして絶対に必要
	接尾	
	非自立	
助詞	格助詞—一般	キーワードとして意味を持ちそうにない
	格助詞—引用	
	格助詞—連語	
	係助詞	
	接続助詞	
	副詞化	
	副助詞	
	副助詞—並立助詞—終助詞	
	並立助詞	
	連体化	
	助動詞	
接頭詞	名詞接続	キーワードとして意味を持ちそうにない
動詞	自立	キーワードとも取れる表現を多く含んでいる (判断が分かれるところ)
	接尾	キーワードとして意味を持ちそうにない
	非自立	キーワードとして採用したい表現もあるが、解釈が難しい語を多く含んでいる
副詞	一般	キーワードとして採用したい表現もあるが、解釈が難しい語を多く含んでいる
	助詞類接続	キーワードとも取れる表現を多く含んでいる (判断が分かれるところ)
名詞	サ変接続	キーワードとして絶対に必要
	一般	
	形容動詞語幹	
	固有名詞—組織	
	数	キーワードとして採用したい表現もあるが、解釈が難しい語を多く含んでいる
	接尾—一般	キーワードとして採用したい表現もあるが、解釈が難しい語を多く含んでいる
	接尾—助数詞	
	接尾—助動詞語幹	
	接尾—特殊	
	代名詞—一般	
	非自立—一般	
	非自立—副詞可能	
	副詞可能	
	連体詞	

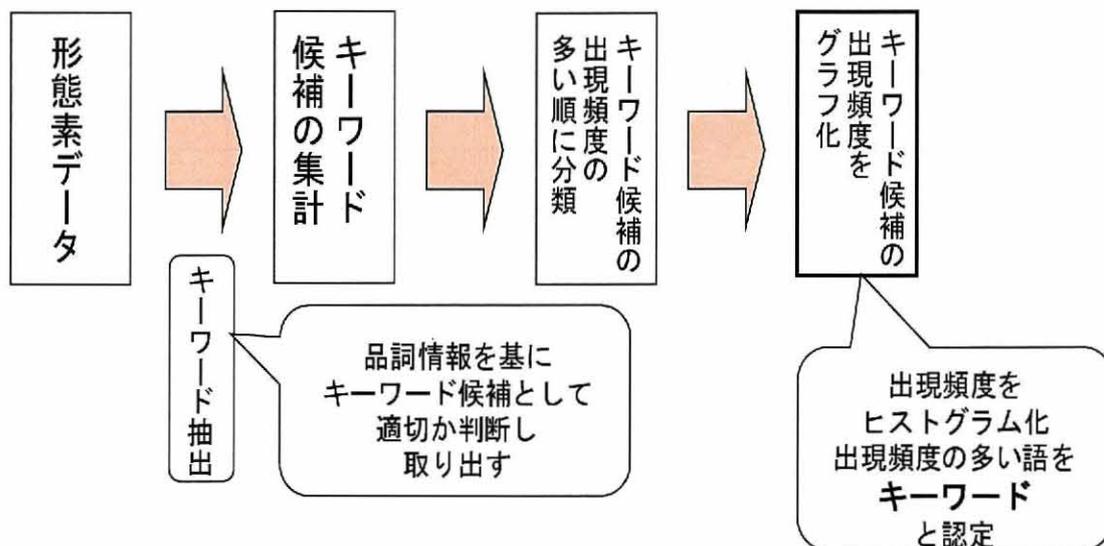


図 2.7 品詞情報を基にキーワードを取り出す

この手法は、分析する文章によってキーワード候補が違う点が難点であり、誰も行える手法ではない。この手法は本研究の標準プロセスの探求には適用不可と判断した。

## 2.4 キーグラフの概要

### 2.4.1 キーグラフの基本的な考え方

キーグラフ[26,42]は、文書が著者独自の考えを主張するために書かれるという仮説を基にして、文章全体はその主張を目指して一つの流れを形成しているとして、文章を建物にたとえて説明できる。

建物が建つには土台（文章を基にしている基本概念）が必要である。壁（文章の構成に必要な説明部分）、ドアや窓（詳細な記述）、さまざまな装飾（比喩や例など、付加的な記述）もある。しかし、建物の本質は日射や雨から住人を守る屋根（主張点）があって、屋根を支えるために柱（主張の主な展開）がある。

文章の中で繰り返される頻出語には文章の主張として著者が用いたキーワード

も含まれるが、それ以外にも文章の主張を支えるための重要な概念として文章の「土台」を形成するキーワードが数多く含まれている。これらの土台の上に立つ「柱」に支えられて文章全体の論点となっているのが主張（「屋根」）である。ここで多くの「土台」に支えられている「主張」を表すキーワードを抽出するために、土台と主張の関係である「柱」を基にしてキーワードを抽出するのがキエグラフである。キエグラフのアルゴリズムは次の3フェーズからなる。

#### (1) 土台の形成

文書形成の準備あるいは前提となる基本概念（具体的には、共起グラフにおいて強く連結しあう語の集まり）を土台とする。

#### (2) 屋根の形成

(1) で取り出した土台たちに強い力で支えられて文章を統合する語を屋根とする。

#### (3) キーワードの抽出

土台と屋根を結ぶ強い柱が多く集まった語をキーワードとする。

3つのフェーズを実行する前に、キーワードの候補としてふさわしくない（文章の土台や屋根となり得ない）単語を対象の文書から削除する準備フェーズを実行する。以下にその詳細を説明する。

まず、土台の形成は文書  $D$ （以下  $D$  と記す）を表す共起グラフ  $G$ （以下  $G$  と記す）をノードと枝(リンク)から生成する。

#### ・ノード

$D$  中の単語は、その出現回数によってソート（分類）される。このソートでの上位  $M$ （経験的に 30 程度が妥当）語からなる集合を **HighFreq** と呼び、はじめに  $G$  中のノード群として与えておく。（図 2.8 中の黒いノード）

**HighFreq** 中の語は土台を形成する要素として用いられる。なぜなら、土台す

なわち基礎概念を表す語は D 中で何度も用いられることが多いからである。

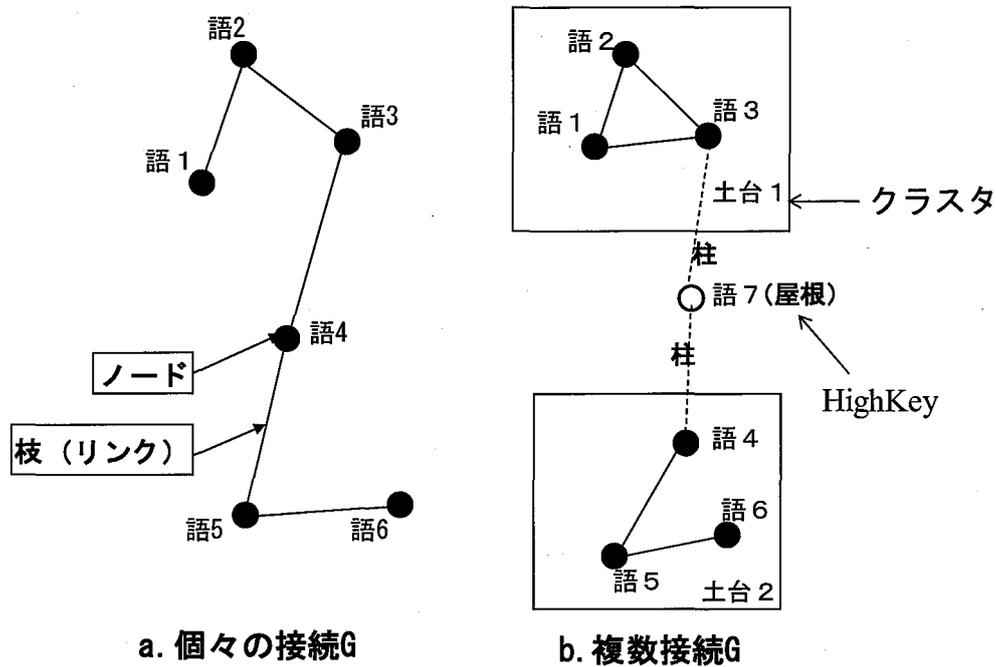


図 2.8 土台・屋根・柱の関係

・ 枝 (リンク)

HighFreq 中で共起度の高い語の対をそれぞれ枝で結ぶ。ここで、語の対  $(w_i, w_j)$  の D における共起度  $co(w_i, w_j)$  を次のように定義する。

$$co(w_i, w_j) = \sum_{s \in D} |w_i|_s |w_j|_s \quad (2.1)$$

ここで、 $|x|_s$  は、文  $S$  における要素  $x$  の出現回数で、 $x$  が語の場合に  $|x|_s$  は文  $S$  中の語  $x$  の出現回数である。式 (2.1) は、ある文  $S$  に出現した  $w_i$  は  $S$  中のすべての  $w_j$  と共起しているとみなした共起度を表す。

2 語間の共起度の定義として、文単位で共起度を測定した理由は次のとおりである。文単位で共起度を測定した理由を述べる。今、共起度を、文章中で連続する長さ  $w$  の範囲中で 2 語両方が出現する頻度とする。 $w$  が 1 文より短い場合には、

倒置や疑問文において語順が変わると、著者にとって関連の深い語同士が遠く離れ  $w$  内で共起しなくなることがある。逆に  $w$  を 1 文より長くすると、2 語のうち片方の内容が指示語の形で複数文にまたがって現れても、その内容ともう片方の語との共起をとらえることができ精度があがることもある。しかし、キーグラフのグラフ構築では、その原因で精度が上がるよりも多くのケースにおいて、実際には意味のつながりの弱い語の対にまでリンクが張られてしまい、この後で得られるキーワードの精度を劣化させる原因となると考えたためである。

HighFreq 中の語の各対は  $co$  の大きさによってソートされ、上位  $M-1$  (本研究では  $M$  はキーワードの抽出結果から判断して 30 とした) 番目までの語対に  $G$  中で枝を張る。なぜなら、それが  $G$  中のノードを冗長な ( $D$  の内容の展開を表す上で不用な) リンクなしに結び合わせるため必要最小限の枝数だからである。すなわち、 $G$  が連結グラフであれば、 $M-1$  本の枝だけ  $G$  中に存在することは、 $G$  が単結合であることと同値であり、それは  $D$  が冗長さなしに内容が展開する文章としてグラフ化される (例えば図 2.8-a のように語 3 が語 1 からただ一つのパスでたどれるという図になる) ことにあたる。

もし、図 2.8-b のように複数のパスを通して語 1 から語 3 がたどられるならば、語 3 と語 4 の関係の強さが土台 1 の内部の語同士の関係よりも弱いとみなせる。そこで、上記の  $co$  の上位  $M-1$  本の枝からなるグラフ  $G$  のうち、自分以外の連結部分グラフ (自分が含むどの 2 ノード間にもパスを持つ  $G$  の部分グラフ) に含まれない連結部分グラフ、すなわち極大連結部分グラフを一つの土台とする。

この土台の抽出法は、一つの文は一貫した概念を表現し、その概念をめぐって文中の語が関係しあうという仮説 (語彙的連鎖) に立っている。語彙的連鎖は文書を意味の一貫した部分 (土台のように) に分割する目的でも用いられているが、どの部分が文書中で重要であるかは残された課題であった。

キーグラフではこの語彙的連鎖を用いて文書の土台をつかみ、その後文書全体の流れにとって重要な主張を表すキーワードを抽出するのである。

土台、すなわち文書の基礎となる概念は、文書全体から見ると著者が主張しようとすることを導き出すために関連しあっている。抽出したいキーワードは文書の主張を表す語であるから、屋根として土台たちに強い力で支えられて文章を統合する語 (図 2.8 の語 7) でなくてはならない。

キーグラフでは語  $w$  が土台たちに支えられる力を  $\text{key}(w)$  で表す.  $\text{key}(w)$  は 0G から 1 までの実数で, 「G 中のすべての土台を著者が考慮しているときに  $w$  が用いられる」という条件付き確率を定式化したものである.

まず,  $\text{key}(w)$  を構成する次の 2 つの関数を定義する.

$$f(w, g) = \sum_{s \in D} |w|_s |g - w|_s \quad (2.2)$$

$$F(g) = \sum_{s \in D} \sum_{w \in s} |g - w|_s \quad (2.3)$$

$S$  : 語の示す添え字

$W$  : 語を示す添え字

$|g|_s$  : 土台  $g$  に含まれる語の  $s$  中の出現回数

$$|g - w|_s = \begin{cases} |g|_s - |w|_s & \text{if } w \in g \\ |g|_s & \text{if } w \notin g \end{cases} \quad (2.4)$$

すなわち,  $f(w, g)$  は語  $w$  と土台  $g$  中の語と共起度である. 式(4)で  $w$  が  $g$  に含まれるときに,  $w$  の出現回数を  $g$  の出現回数から差し引くのは,  $w$  と  $g$  の中の  $w$  以外の語との共起度を調べるためである.

$\text{Key}(w)$  は次のように定義する.

$$\text{key}(w) = \left[ 1 - \prod_g^{\text{bases}} \left( 1 - \frac{f(w, g)}{F(g)} \right) \right] \quad (2.5)$$

ここで  $\text{bases}$  は土台の個数である.  $f(w, g)/F(g)$  は  $g$  中の語と共起する語が  $w$  である確率であり, これを作成者が土台  $g$  の表す概念を考慮しているときに語  $w$  を書く確率として用いている.

文書  $D$  中の  $\text{key}$  の値の上位の語の集合(「屋根」にあたる)を  $\text{HighKey}$  と呼ぶ.

HighKey 中の語  $G$  にまだ含まれていなければ新しいノードとして  $G$  に加える。

キーグラフで取り出すキーワード群は、HighKey(屋根)そのものではない。というのは、土台中の語でも、屋根に強く結び付いている語は屋根を表現し文書  $D$  を要約する上で重要だからである。これらの語は  $key$  の値は小さいかもしれないが、HighKey 中の語と同様に接する「柱」(土台と屋根を結ぶリンク) の強さの和では上位にランクされるものと考えられる。そこで、HighKey 中の語との間の柱の強さ(下記) の和の大きな語を最終的にキーワードとする。

HighKey 中の語  $w_i$  と、いずれかの土台に含まれる語  $w_j$  を結ぶ柱の強さ  $c(w_i, w_j)$  は次の式で表す。

$$c(w_i, w_j) = \sum_{s \in D} |w_i|_s |w_j|_s \quad (2.6)$$

そして、 $G$  中のノードで、接する柱すべての  $c$  の和が上位の語をキーワードとする。

## 2.4.2 キーグラフ上での表示の仕方

前項で述べたようにキーグラフは土台の形成、屋根の形成、キーワードの抽出という 3 つのプロセスから成り立っている。これらのプロセスを次のような規約の上に可視化している。キーグラフは可視化のツールとしても優れた機能を提供している。

次にキーグラフ上での表示の仕方を示す[43]。

- ① 黒丸(小) ----- ノード (HighFreq) 土台を形成する
- ② 赤丸(小) ----- ノードに含まれていない HighKey で屋根を形成
- ③ 緑丸(大) ----- 抽出されたキーワード
- ④ 実線 ----- 枝(リンク)
- ⑤ 赤点線 ----- 柱(土台と屋根を結ぶリンク)

図 2.9 はキーグラフによる解析結果表示の例である。キーグラフではノードの位置やノード間の距離は、適当に見やすいところに表示されるだけである。しかし、その結果として、強く繋がったノードは近い位置に現れることが多い。ノード数が多くなるとノード同士の重なりなどが生じるので、ノードの移動が簡易に行えるよう配慮されている（マウスでドラッグし移動できる機能）。

JaJa[18-35-38-9]

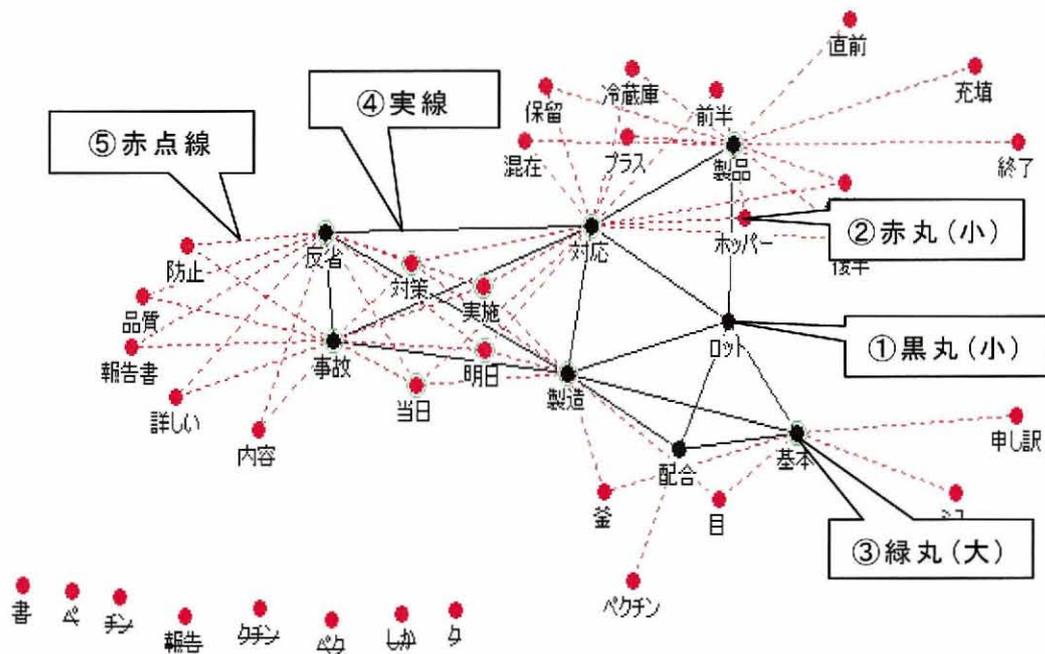


図 2.9 キーグラフによる解析結果表示例

# 第3章 課題発見のプロセス構築

## 3.1 チャンス発見

近年ビジネス界において様々な成功を導いているチャンス発見プロセスが注目されている[30,31,32,44]. チャンス発見プロセスでは, 意思決定を左右する可能性を有する事象または, 状況をチャンスと定義している.

ある文脈の中で生起する一連の事象系列をシナリオと呼ぶと, 新たなシナリオの部分として欠かせない事象あるいは状況を知ることが, 意思決定, すなわち未来のシナリオを一つ選択する作業に強く左右する. したがって, そのような事象や状況は上記定義によればチャンスである. 企業におけるチャンスとは, 例えば顧客や社員が一言で表した言葉であり, その言葉で新しいビジネスアクションのシナリオを描くことができるようなものである. その言葉は顧客や社員からニーズの発生という事象を表しており, そのニーズを満たせば描いたシナリオは実現できるかも知れない.

このような背景からチャンス発見では, 発見主体である自分のおかれた環境での様々なシナリオを発案し, 新しいシナリオを探索し選択してゆくプロセスを辿ってきた. ここで選択されるシナリオに実現可能性を持たせるためには, シナリオマップが有用となる. シナリオマップとは, 環境における様々な事象をどのように関連し合うかを可視化したものである. この可視化にキーグラフが有用なツールと位置づけられる.

### (1) チャンス発見のプロセス

図 3.1 のチャンス発見のプロセスについて述べる.

チャンス発見のサイクルは

- ・予兆は稀なチャンスであるかも知れない, その稀さゆえに, 最初, 人々はその価値について理解できていない状態にあるかもしれないこと.
- ・しかし, チャンスを発見したいという関心さえないならば, 人はチャンスに気づくことさえできなということ.

の二点を反映して, 「関心」からチャンスへの気づきを深めて「理解」し, そしてそのチャンスを用いた行動シナリオの「提案」に至り, その段階で実環境と相互作用することで, 初めて「評価」が可能となる様子を表している. この評価において,

場合によっては新しいチャンスに関心を寄せるようになって次のサイクルに向かい、ここで同時に環境に新しい目を向けてデータを収集するのが「観測」である。

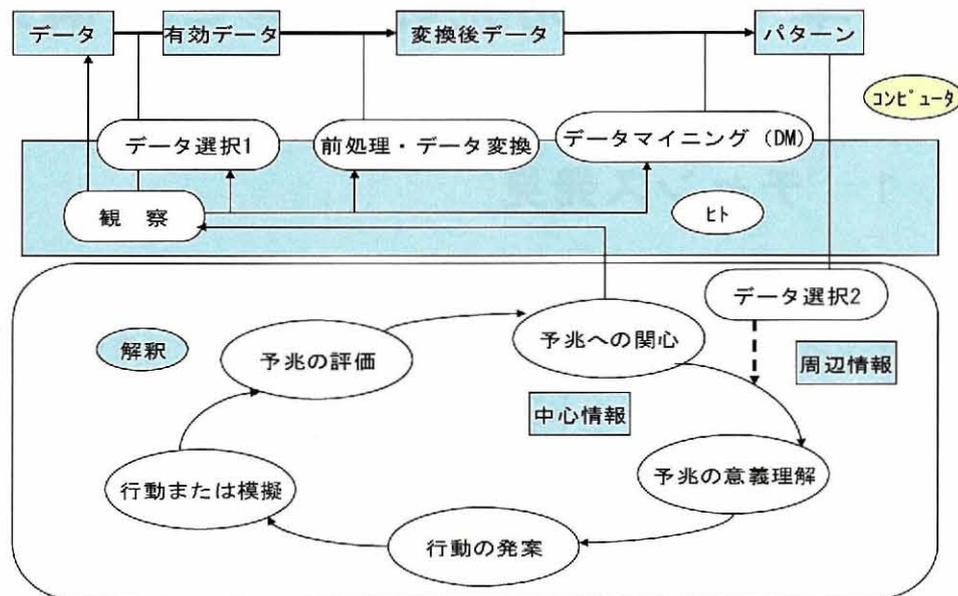


図 3.1 チャンス発見のプロセス

すなわち、チャンス発見の特徴は、次の各状態からなる発見者の心理コンテキストの遷移が主役を果たしている点にあるといえる。

- ・ **関心**：これから遭遇するかもしれない予兆について、自分の意思決定にとっての意義を新たに理解したいと感ずる状態、すでに実行あるいは模擬を行った行動をさらに改善するために、環境についてこれまでよりも深い理解を得ようとする。
- ・ **理解**：ある予兆が自分の意思決定に対して持つ意義を理解した状態。
- ・ **発案**：予兆をチャンスとして利用した行動のシナリオを発案し、場合によりグループの他者に提案する。
- ・ **行動または模擬**：上記のシナリオに基づき、実際または仮想的に（シミュレーション世界で）行動をとる。
- ・ **評価と選択**：実行または模擬の結果に基づき、行動を評価し選択する。この評価において、もっとも望ましいシナリオへの新たな関心が生まれる場合は、次のサイクルの「関心」に遷移する。

ここで、「行動」というのは、自分に接することのできる外部環境と相互作用を行うことをさす。

## (2) チャンス発見の二重らせんプロセス

図 3.1 の不足を補うために、新たに二重らせんモデルを提案している。意思決定者でありチャンスの発見者となるべき人が、自分の頭を通り過ぎる理解過程を言葉に残しておくだけで、新しいチャンスを活かすための戦略が発見できることは意外に多い。自分が扱っている環境から得られる「環境データ」に加えて、意思決定主体の思考過程をとらえた「主体データ」をあわせて使ってチャンス発見を進めるのが図 3.2 の「二重らせん」というプロセスである。

二重らせんモデルでは

- ①チャンスへの関心を抱く。
- ②目前の事象への意味を理解し、これをチャンスとして未来に活かすシナリオを発案する。
- ③チャンスを選択し、行動またはシミュレーションを行う。
- ④新しいチャンスに関心を抱く。

というサイクルを繰り返しながら螺旋状に進む「人のチャンス発見プロセス」と並行して、その人が自分の関心に基づいて選んだ環境データと、その人自身の思考内容を記録した「主体データの解析」をコンピュータが繰り返す。そのコンピュータの仕事は「データマイニング」と呼ぶ。主体データのようなテキストデータにたいするデータマイニングを「テキストマイニング」という。人が環境データのマイニング結果を理解している最中も、その理解過程までコンピュータに主体データとして取り込み、人とコンピュータが常に並列に働き続けることから「二重」らせんと呼ぶ。

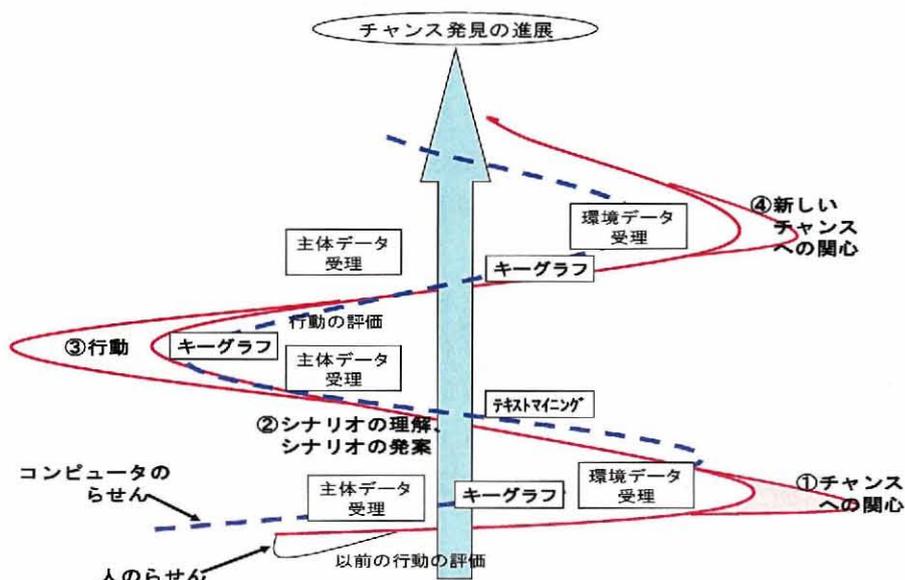


図 3.2 チャンス発見の二重らせんプロセスモデル

### (3) 応用事例

近年ビジネス界において様々な成功を導いているチャンス発見プロセスが注目されている。この項では実際に応用した事例(表 3.1)を示す[31,32,45,46]。

表 3.1 チャンス発見事例

応用事例	事例概要
360度評価における自由回答・選択式回答の混合データからの人事評価尺度発見	テキスト情報を含む360度評価の人事データから、管理職層の構成員に求められるニーズを抽出するプロセスを提案。実証ではこれまで設問項目に含まれていなかったが社内の重要なニーズに関する人事評価尺度が発見された。
二重螺旋のモデルを用いたスーパーの顧客行動変の予兆発見	顧客購買(POS)データからスーパーマーケットの購入金額増加のトリガとなる商品や、そのむせの経営状態の変化を示す予兆を発見することができた。

以上の考え方、応用事例から判断して、本研究でのプロセスに適用可能であり、課題・解決策のキーワード検出、デシジョンテーブル作成時、共有化に適用する。

## 3.2 分析手法の概要

### 3.2.1 回帰分析

回帰分析は、1つの計量的な従属変数と1つ以上の独立変数の間にどのような関係があるかを分析する、強力かつ柔軟性のある手法である。たとえば、以下の場合に用いられる[48]。

1. 従属変数における優位な変動を独立変数によって説明できるかどうか、つまり両者の間に何らかの関係が存在するかどうかを見る。
2. 従属変数の変動のうち独立変数によって説明できる変数がどれくらいあるか、つまり両者の関係の強さがどれくらいかを見る。
3. 独立変数と従属変数の間にどのような関係があるか、つまり両者の関係を示す数式を明らかにする。
4. 従属変数の値を予測する。
5. 他の独立変数による効果を制御しながら、特定または一組の独立変数による

寄与率をみる。

当研究では、回帰分析においてデータ間の関連性から「ルール」を見出すことができることに注目している。

今回扱うデータは質的変数であるため、ダミー変数にコード化することで独立変数として使用可能となる。この方法は数量化理論 I 類の適用である。

(数量化理論 I 類 (Dummy variables regression : 目的変数が定量データ, 説明変数が定性データ (カテゴリ・データ) の解析技法のひとつである)

解析に出るおもな用語の意味について [49]

① アイテムとカテゴリ

アイテムは性, 年齢, 地域など説明変数自体の項目を表し, カテゴリは, アイテムの内訳, たとえば, 性では男性と女性を指す。

② カテゴリスコア

どのカテゴリが予測値に影響しているかを表わすスコア。この値の絶対値の大きいほど目的変数への影響度が高い。

③ アイテムレンジ

アイテムの各カテゴリスコアのなかで最大値から最小値を引いた値。アイテムレンジの高いアイテム (説明変数) ほど目的変数に与える影響度が高い。

この分析における最適なアイテム (説明変数) グループを見つけるために, 説明変数選択基準 (Ru) を適用した [50,51,52,53,54]。

説明変数選択基準には赤池の AIC [55] や和光の選択基準 [56] など幾多の方法が提案されているが, 少量のデータでの適応を可能としている上田案を適用した。

また, 重回帰分析を適用して, 被説明変数に影響を与える因子群を求めることができる。

この因子を求めることで予測や要因分析が行える。

その要因での, 回帰式は以下の式で求めることができる。

$$y = \alpha_1 \chi_1 + \alpha_2 \chi_2 + \dots + \alpha_k \chi_k + c + \epsilon \quad k = (1, \dots, k) \quad (3.1)$$

(y : 被説明変数,  $\chi$  : 説明変数,  $\alpha$  : 回帰係数, c : 定数,  $\epsilon$  : 誤差)

説明変数選択基準 (Ru) は説明変数の最適化を図るための方式でありその式は以下である。

$$Ru = 1 - \frac{(1 - R^2) * (n - k - 1)}{n - k - 1} \quad (3.2)$$

(R : 関係数 n : データ数 k : 計算で用いた説明変数の総数)

最適な項目の組み合わせは Ru が正の数で最大となるときの組み合わせである。

### 3.2.2 属性の序数性を考慮したラフ集合

量的データに対して属性間の序数性を考慮したラフ集合によるルール抽出について説明する[36]. いま,  $S$  個のサンプルについて  $N$  個の条件属性と  $M$  個の決定属性からなる決定表を考える. 条件属性のうち序数性が仮定できるものを基準と呼ぶ. このとき,  $O_q$  を  $q$  番目の基準に基づいたアウトランキング関係とする. すなわち,  $x O_q y$  は  $q$  番目の基準に基づけば,  $x$  は少なくとも  $y$  と同程度によいことを表す.  $O_q$  は比較可能で推移的な関係すなわち弱順序関係であるとする. さらに  $m$  番目の決定属性において,  $S > t$  に対して  $C_m^s$  の任意の要素  $C_m^t$  がすべての要素よりも好ましいという性質をもつ  $U$  の分割  $r = \{C_m^1, C_m^2, \dots, C_m^R\}$  を考える.

ここでは, 関係  $(x \in C_m^s, y \in C_m^t, s > t \Rightarrow x O_m y \text{かつ} \sim y O_m x)$  が成立している.

いま,  $m$  番目の決定属性を  $R$  個のクラス

$$C_m^s \cap C_m^t = \phi, \quad (s \neq t), \quad C_m^R \succ \dots \succ C_m^r \succ C_m^1 \quad (3.3)$$

に分類する. このとき,  $x \in U$  が与えられると, 少なくともクラス  $C_m^r$  に属している  $U$  の要素の集合である上側累積集合  $C_m^{\geq r}$  と, たかだかクラス  $C_m^r$  に属している  $U$  の要素の集合である下側累積集合  $C_m^{\leq r}$  が,

$$C_m^{\geq r} = \bigcup_{s \geq r} C_m^s, \quad (3.4)$$

$$C_m^{\leq r} = \bigcup_{s \leq r} C_m^s$$

で定義できる.

すべての基準の集合を  $W$  とするとき,  $V \subseteq W$  を考える. 任意の  $v \in V$  について,  $x O_m^v y$  が成立するとき,  $x$  は  $V$  において  $y$  を支配するといい,  $x O_m^V y$  で表し

$$xD_m^V y \leftrightarrow g(x, n) \succ g(y, n), \quad (\forall v \in V), \quad (3.5)$$

と定義する。ここで、 $g(x, n)$  でサンプル  $x$  の属性  $n$  に関する属性値を表す。  
 $D_m^V = \bigcap_{v \in V} o_m^v$  となり、 $o_m^v$  は弱順序関係であったので、 $D_m^V$  は半予順序関係となる。  
 このとき、 $m$  番目の決定属性において、 $x \in U$  が与えられると、 $V$  において  $x$  を  
 支配する  $U$  の要素の集合  $D_m^{+V}(x)$  と、 $V$  において  $x$  を支配される  $U$  の要素の集合

$$D_m^{-V}(x) \text{ が} \quad (3.6)$$

$$D_m^{+V}(x) = \{y \in U \mid y D_m^V x\},$$

$$D_m^{-V}(x) = \{y \in U \mid x D_m^V y\}.$$

で定義できる。

支配集合  $D_m^{+V}(x)$  による累積集合  $C_m^{\geq r}$  の下近似集合  $V_*(C_m^{\geq r})$  と上近似集合

$$V^*(C_m^{\geq r}) \text{ は} \quad (3.7)$$

$$V_*(C_m^{\geq r}) = \{x \in U \mid D_m^{+V}(x) \subseteq C_m^{\geq r}\},$$

$$V^*(C_m^{\geq r}) = \bigcup_{x \in C_m^{\geq r}} D_m^{+V}(x)$$

で定義できる。この下近似集合から、 $V_*(C_m^{\geq r})$  に属している  $x$  を支配しているデータ  $x^*$  は必ずクラス  $r$  以上に属しているというルールが導かれる。すなわち、ある  $x^* \in C_m^{\geq r}$  について、次のような if-then ルールが得られる。

$$\text{IF } g(x^*, n_1) \geq g(x, n_1) \text{ and } g(x^*, n_2) \geq g(x, n_2) \quad (3.8)$$

$$\dots \text{and } g(x^*, n_N) \geq g(x, n_N), \text{ THEN } x^* \in C_m^{\geq r}$$

$V^*(C_m^{\leq r})$ は

$$\begin{aligned} V_*(C_m^{\leq r}) &= \{\chi \in U \mid D_m^{-v}(\chi) \subseteq C_m^{\leq r}\} \\ V^*(C_m^{\leq r}) &= \bigcup_{x \in C_m^{\leq r}} D_m^{-v}(x) \end{aligned} \quad (3.9)$$

で定義できる。やはり、この下近似集合から、 $V_*(C_m^{\leq r})$ に属している  $x$  を支配しているデータ  $\chi^*$  は必ずクラス  $r$  以下に属しているというルールが導かれる。すなわち、ある  $\chi^* \in C_m^{\leq r}$  について、次のような if-then ルールが得られる。

$$\begin{aligned} \text{IF } g(\chi^*, n_1) \leq g(\chi, n_1) \text{ and } g(\chi^*, n_2) \leq g(\chi, n_2) \\ \dots \text{and } g(\chi^*, n_N) \leq g(\chi, n_N), \text{ THEN } \chi^* \in C_m^{\leq r} \end{aligned} \quad (3.10)$$

いま、 $C_m^{\geq r}$  と  $C_m^{\leq r}$  について境界が

$$\begin{aligned} B_V(C_m^{\geq r}) &= V^*(C_m^{\geq r}) - V_*(C_m^{\geq r}) \\ B_V(C_m^{\leq r}) &= V^*(C_m^{\leq r}) - V_*(C_m^{\leq r}) \end{aligned} \quad (3.11)$$

で定義できるため、 $C_m^{\geq r}$  と  $C_m^{\leq r}$  について近似の精度が

$$\begin{aligned} \alpha v(C_m^{\geq r}) &= \frac{|V^*(C_m^{\geq r})|}{|V_*(C_m^{\geq r})|} \\ \alpha v(C_m^{\leq r}) &= \frac{|V^*(C_m^{\leq r})|}{|V_*(C_m^{\leq r})|} \end{aligned} \quad (3.12)$$

と定義され、分割  $r$  が部分基準集合  $V$  によって正しく分類できた対象の割合である近似の質は次のように定められる。

$$\beta v(T) = \frac{|U - (U_{r=1}^n B_V(C_m^{\geq r}) \cup U_{r=1}^n B_V(C_m^{\leq r}))|}{|U|} \quad (3.13)$$

$\beta_v(r) = \beta_v(r)$  が成立する極小集合  $V \subseteq W$  を縮約と呼ぶ。縮約は複数が存在し、それらの共通集合を核と呼ぶ。縮約に帰属する属性を用いることにより、

近似の質を低下させることなく，決定表を最も簡略化することができる。

ここで，表 3.2 のミカンの品質に関して得られたデータに対して，感性工学で多用されている従来のラフ集合における簡略化手法を適用して簡潔な if-then ルールに集約することを考える。

表 3.2 ミカンの品質に関する決定表

No	Weight	Size	Brix	Acid	Quality
1	very heavy	medium	high	medium	good
2	heavy	small	very high	low	good
3	medium	small	very high	medium	bad
4	light	large	low	very high	bad
5	light	medium	high	high	bad
6	medium	small	high	medium	good

従来法では，上近似を考慮せずに下近似のみから導出される近似の質に依存して縮約が求められるため，近似の質が 1 の場合に縮約が Weight, Brix と Brix, Acid となり，Brix が核となる。最初の縮約にもとづく

IF Weight is very heavy, THEN Quality is good.

IF Weight is heavy, THEN Quality is good.

IF Weight is light, THEN Quality is bad.

IF Weight is medium and Brix is very high, THEN Quality is bad.

IF Weight is medium and Brix is high, THEN Quality is good.

というルールが抽出される。しかしながら，一般には Brix が高くなるほど Quality は良くなるのが妥当であると考えられるため，このルールは受け入れがたい。このことは，従来の研究では各属性値は名義的な尺度として取り扱っているため，通常の簡略化手法では決定表が矛盾している場合には理解しがたい if-then ルールが得られる可能性があることを示している。

そこで，属性値間の序数性を考慮したラフ集合における簡略化手法を適用すると，近似の質が 2/3 の場合に Weight が縮約かつ核となる。この縮約から

IF Weight is better than heavy, THEN Quality is good.

IF Weight is worse than light, THEN Quality is bad.

というルールが抽出可能となる。この結果は従来法で得られた結果に比較して妥当なものと考えられることができる。しかしながら，条件属性が質的データである場

合には、通常の属性値間の序数性を考慮したラフ集合における簡略化手法は困難である。

条件属性が質的データである場合には、通常の属性値間の序数性を考慮したラフ集合における簡略化手法は適用できないため、ここでは、質的データに対して区間回帰分析を適用することにより、質的データに順序関係を与える方法を説明する。

まず、次のような区間回帰モデルを考える。

$$Z = \sum_{n=1}^N A_n X_n \quad (3.14)$$

ただし、 $A_n$  は属性  $n$  に関する区間効用値ベクトル  $(c_n, w_n)(n=1,2,\dots,N)$  を表し、

$$c_n = (c_{n1}, c_{n2}, \dots, c_{ni_n})^T \quad (3.15)$$

は中心ベクトル、

$$w_n = (w_{n1}, w_{n2}, \dots, w_{ni_n})^T \quad (3.16)$$

は幅ベクトルである。ここで、 $i_n$  は属性  $n$  の要因数であり、 $N$  は属性の数である。

$x_n$  は属性  $n$  の要因ベクトルである。区間演算により区間出力  $Z$  は

$$Z = \sum_{n=1}^N c_n^T X_n, \sum_{n=1}^N w_n^T X_n \quad (3.17)$$

で表すことができる。

$J$  個の入力データ  $x_{nj}$  と出力データ (評価値)  $z_j$  が与えられたとき、 $z_j$  を包含し、

区間線形モデルの幅の合計が最小となる区間係数  $A_n$  を求めることになる。

この問題は次の線形計画問題に帰着される。

$$\min_{C_n, W_n} \sum_{n=1}^N C_n^T X_n, \sum_{n=1}^N W_n^T X_n \quad (3.18)$$

$$\text{Subject to} \quad \sum_{n=1}^N C_n^T X_{nj} - W_n^T X_{nj} \leq Z_j$$

$$Z_j \leq \sum_{n=1}^N C_n^T X_{nj} + W_n^T X_{nj}$$

$$C_n, W_n \geq 0 \quad (n=1, 2, \dots, N)$$

区間効用値  $A_n = (C_n, W_n)$  が得られると、区間値  $S[s_1, s_2] T[t_1, t_2]$  が与えられたとき、属性のカテゴリ間の順序付けは

$$S[s_1, s_2] \succeq T[t_1, t_2] \leftrightarrow s_1 \geq t_1, s_2 \geq t_2 \quad (3.19)$$

で行うことができる。

本研究では、 $g(x, n)$  が区間値であるので、 $x$  が  $V$  において  $y$  を支配することを

$x D_m^V y$  で表し、

$$x D_m^V y \leftrightarrow g(x, n), \quad (\forall_v \in V) \quad (3.20)$$

$$\leftrightarrow \min_{s_1 \in g(x, n)} s_1 \geq \min_{t_1 \in g(y, n)} t_1$$

$$\text{and} \quad \max_{s_2 \in g(x, n)} s_2 \geq \max_{t_2 \in g(y, n)} t_2$$

で定義する。

さらに、 $V_*(C_m^{sr})$  に関する下近似集合から得られたルールと  $V_*(C_m^{sr})$  に関する下近似集合から得られたルールの条件部において、すべての属性においてカテゴリが同じとなる集合から、確実にある  $x^* \in U$  がクラス  $\gamma$  に属するという次の if-then ルールが得られる。

$$\begin{aligned} & \text{IF } g(\chi^*, n_1) = g(\chi, n_1) \text{ and } g(\chi^*, n_2) = g(\chi, n_2) \\ & \dots \text{and } g(\chi^*, n_N) = g(\chi, n_N), \text{ THEN } \chi^* \in C_m^r \end{aligned} \quad (3.21)$$

この抽出されたルールを満たすデータは確実にクラス  $r$  に属していることとなる。

表 3.2 のミカンの品質に関する決定表に、質的データにおける属性値間の序数性を考慮したラフ集合における簡略化手法を適用すると、量的データにおける場合と同じ近似の質、ルールが導かれる。ただし、Weight におけるカテゴリの順序関係が

$$\text{Light} \leq \text{medium} \leq \text{heavy} = \text{very heavy} \quad (3.22)$$

となり、Weight に関して heavy と very heavy は Quality にとっては同じであることが明らかにできる。

### 3.2.3 自己組織化マップ (SOM: Self-Organizing Map)

自己組織化マップ (SOM; Self-Organizing Map 以下 SOM と略記) は、1982 年に T.Kohonen によって提案された教師なしの競合近傍学習によって学習するニューラルネットワークの一つであり、多次元データの分類、解析に効果的な技術である。これは、高次元空間から低次元空間への写像を自己組織的に形成し、高次元データを特徴マップと呼ばれる 2 次元平面上へ視覚化することができる [61,62]。

SOM は、入力層と出力層の 2 層によって構成されるニューラルネットワークである。出力層は、競合層とも呼ばれる。図 3.3 のように、入力層のベクトルが

$$\chi_j = (\chi_{j1}, \chi_{j2}, \chi_{j3}, \dots, \chi_{jn}), \text{ 出力層のノード } i (i = 1, 2, \dots, l) \text{ が}$$

$m_i = (m_{i1}, m_{i2}, m_{i3}, m_{in})$  で表わされるとする。ベクトル  $m_i$  は参照ベクトル  $m_i$  と呼ばれ、学習が進むにつれて修正される。

#### SOM のアルゴリズム [61]

STEP 1: すべての参照ベクトルを初期化する。

STEP 2: 入力データ集合の先頭から順番に入力ベクトルを選択する。ステップ数が入力データの数より大きい場合は、先頭から選択する。

STEP 3: 選択した入力ベクトルをすべての参照ベクトルと比較し、ユークリッド距離を最小にする参照ベクトルをもつノードを最整合ノードとする。最

整合ノードは以下の式によって定義される。

$$C = \arg \min_i \{\|x_i - m_i\|\} \quad i = 1, 2, \dots, I \quad (3.23)$$

STEP 4: 最整合ノードとその近傍ノードの参照ベクトルを以下の式に従い更新する。

$$m_i(t+1) = \begin{cases} m_i(t) + h_{ci}(t) [X_j(t) - m_i(t)] & i \in N_c \\ m_i(t) & i \notin N_c \end{cases} \quad (3.24)$$

ただし,  $h_{ci}(t)$ は  $t(t = 1, 2, \dots, T)$ 回目の学習における近傍関数

$$h_{ci}(t) = \alpha(t) \exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)}\right) \quad (3.25)$$

である。  $\alpha(t)$ は学習率係数,  $\sigma^2(t)$ はノード  $C$ の近傍領域  $N_c$ を調整する関数であり, 学習回数の増加とともに単調に減少する。  $r_c$ と  $r_i$ は, それぞれノード  $C$ と  $i$ の 2次元平面上の座標ベクトルである。

STEP 5: 学習回数  $t$ が  $t = T$ であれば終了。 そうでなければ, STEP 2に戻る。

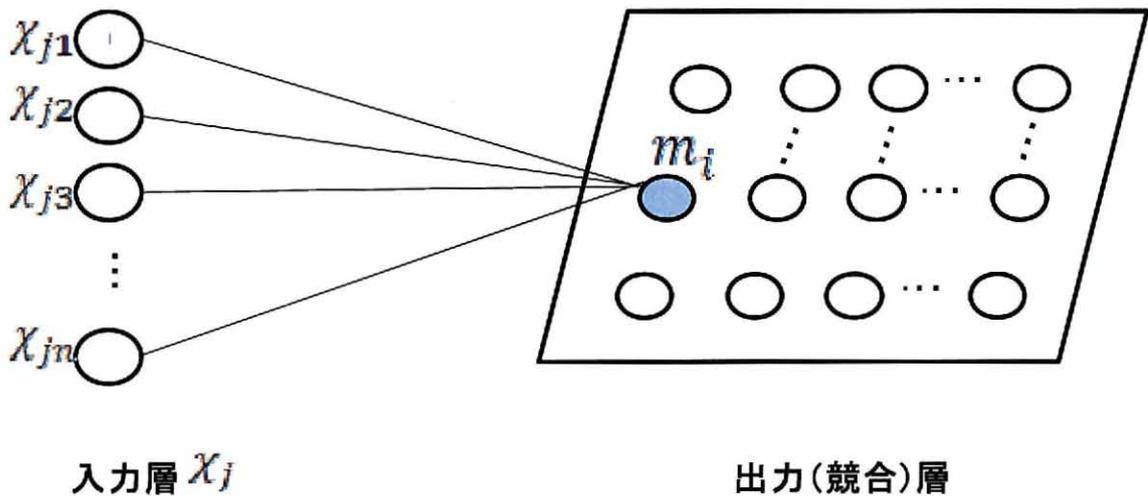


図 3.3 SOMの構造

### 3.3 提案する報告書の形式

本論文では事故再発防止報告書(表 3.3)の事故原因部分をテキストマイニングの対象とした[47].

テキストマイニングを行なうには

- ・テキストの話題の中心をどのようにして見つけるか
- ・前処理でどのようにクリーニングできるか
- ・どのようなマイニング技術を使用するか

が重要な問題である.

A食品会社への提案を実行するにあたり, 事故再発防止報告書のテキスト化されたデータを提供してもらい分析作業に取り掛かった.

事故再発防止報告書は, 問答式になっており, 決められた項目に沿って書けば出来上がる仕組みになっている, しかし, 表現に統一性がなかったり, ひたすら謝罪文であったり, A食品会社社内独自語が多かったりで, 次の改善要件を提案した.

#### 事故再発防止報告書の改善要件

##### (1)書き方の統一

##### (2)用語の統一

##### (3)フォーマットの再考

(1)の書き方の統一はすぐには実行できないので, 気づいた点を書き出して置いて, 最終的には書き方指南書を作成することにした. ただし, 美文化されることで本質が失われる可能性もあるので, 注意が必要である. 基本事項は前項 2.6.1 データクリーニングの表現の統一を参考にした.

(2)の用語の統一については, プロジェクトのメンバーで作成可能分から始めることにした. いずれは固有語辞書とする.

例) シェフイール, デニーズ, 内箱, 端数あわせ, 荷抜け.....など

(3)のフォーマットの再考については, すぐにはできないので今後の課題とした.

なお, A食品会社プロジェクトのメンバーには図 3.4 のような作業の流れを示し, 各作業の内容の説明, 作業分担を依頼した.

表 3.3 事故再発防止報告書(サンプル)

検印									
<b>事故再発防止報告書</b>									
[提出先] XX XX 工場長 殿					作成年月日 年 4 月 9 日				
					報告者 佐藤 隆 氏 名 XXX X				
(注)記入内容は簡易書に 製品名置その他 全て製法化									
ライン停止時間	0分								
損失金額									
事故内容	事故原因	再発防止対策							
(いつ どこで 誰が、なにを、どうした) いつ 200X/3/24 製造終了後 どこで F3・BMMラインで 誰が F3スタッフ藤崎が 何を 製造終了後の、当日使用を入力作業中にロイヤル アルカマ向けイチゴジャムのベタチン使用量が どうした 違っていることに気付いた。 内容 101AS 正規 ミス グラフが不正 770g 770g 330g 770g ベタチン発注時に、依頼書に記入ミスしていた。 22.8g上がりベタチン溶液と伝票に記載していたので、 製造当日使用時に、上記ミスした重量を高精度し、上がり 228gに合わせた。正確な重量であれば、5%溶液である が、グラフが不正が多いため、Brixは7%となった。作 業者も7%と認識したが、思い込みで、7%で良いと思 い、そのまま使用した。その為ベタチン含有量が下記のよ うになった。 正規 配合ミス 101AS 309mg% 309mg% グラフが不正 133mg% 309mg%	(直接および間接原因) ベタチン発注時の記入ミス ・発注時にF3で作成した配合表を元にベ タチン在庫依頼書を記入しているが、単純に記入 ミスした。 ・また、Wチェックシステムを行っていなかつ た為、記入ミスしたものがそのまま依頼され いた。 入庫時、使用時に間違いに気付かなかった。 ・F3に入庫時、持って来たスタッフが、伝票 と袋に書いてある品名 ・数量(連番)をチェックしているため、伝票 が間違っていると、そのままだま間違いに気付かなか った。 ・管理表に、ベタチン名、重量、数量を記入す る際も、伝票を元に記載していた。 ・昨年の1月、2月、3月ベタチン取り違えミス の対策として行っていた、ベタチン溶解比率表 が当時準備できていなかった。(機口が修復、 準備忘れ) ベタチン溶解時に5%のところを、7%であり ながら気付かなかった。 ・当時は品質管理で忙しく、切替後の1番 目の準備で忙しており、Brix7%と見ていな がら、思い込みでそのまま行ってしまった。 ・ベタチン溶解比率表がなかった。 ・配合表でちゃんとチェックしていなかった。	(緊急および恒久対策を具体的に) いつまでに 誰が ・QC発行の配合表を元に、依頼書を記入するよ うに変更。 200x/3/3 XXXXXX 0 ・APMがQC発行配合表と照らし合わせ、W チェックする体制に変更 ・品名、数量、重量を配合表と照らし合わせ チェックする。 ・ベタチン溶解比率表と、伝票を見て管理表に記 入する。 ・再度、奥産さんと相談し、見やすい形式に変 更し、ベタチン発注時に新様式で全アイテム比 率表を作成する。(溶解重量ごとに1枚ずつ) ・重量合わせ、Brix確認時は、ベタチン溶解比率 表で必ず確認することとする。							
事故反省会実施日 年 3 月 28 日					作業手順書の有無 有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/>				
(コメント) この度は、前回の事故の対策が徹底されず、同様な事故を繰り返してしまい申し訳ございません。 再度原因を調べ、根本的に手順を改正し、私1人がやる対応でなく、皆が意識して抜けてく 出来る対応で進めてまいります。					a.「有」の場合、見直しが必要 <input type="checkbox"/> 否 <input type="checkbox"/> 「要」の場合、いつまでに( )、誰が( ) b.「無」の場合、新編作成の要否 <input checked="" type="checkbox"/> 否 <input type="checkbox"/> 「要」の場合、いつまでに( )、誰が( )				
部署責任者: 担当者一人が間違ると、気付けない、気付かぬシステムになっていました。 他部署の事例を参考に体制を整理します。(職名該当)					対応表編の最終確認者( ) (署名) 対応表編の確認日 200X年 7 月 14 日 対応表編の確認日 年 月 日 承認者 ( )				

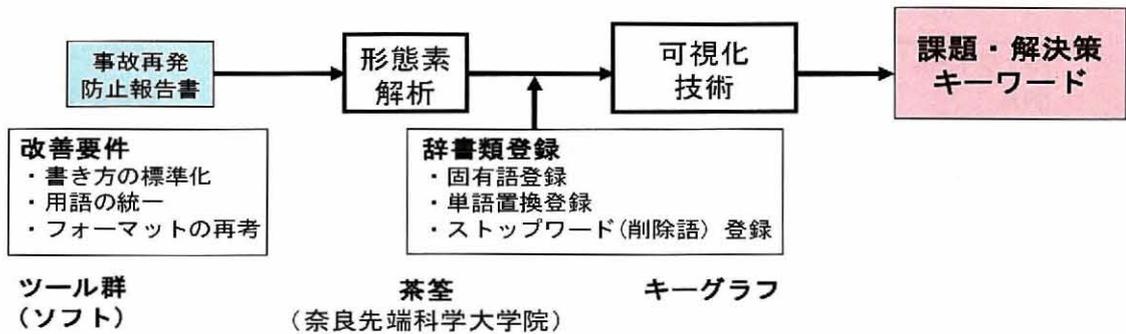


図 3.4 作業の流れ

辞書類の登録

(1)固有語登録

(2)単語置換登録

(3)ストップワード (削除語) 登録

(1)の固有語は事故再発防止報告書の改善要件の用語の統一の件で書き出した語を登録した。今後は発生都度追加とする (表 3.4 参照)。

(2)の単語置換登録は略語の標準語化のために使用することにした。

例) 秤→デジタル秤,

(3)ストップワード (削除語) 登録はキーグラフ作成時に登録が可能なことか

ら、その機能を利用することにした（表 3.5 参照）。

次の作業は茶釜，キーグラフのフリーソフトのインストール段階であるが，A社では，各自にノートパソコンを支給し，その管理基準が明確になっており，原則指定以外のソフトのインストールは禁止となっていた．今回は，特別の許可を得てインストールを行ない，ツールの使い方，キーグラフの見方などの教育を行なう．プロジェクトのメンバーはキーグラフから，シナリオ創生ができるまでには，そう時間は掛からなかった．

事故再発防止報告書のテキストデータ（実際は Excel データ）を茶釜で読めるようにテキスト変換を行ない，用語等の整形をし，キーグラフを実行することでキーワードがキーグラフ上（図 3.5 参照）に抽出される．また，キーワードや単語の関連性が線で連結されるので，そこから事故の原因がシナリオとして読み解かれる．キーワードからは事故原因の源泉がうかがえる．ここまでは順調に推移したが，そこからどのように知見に結び付けられるかが課題となった．

そこで，事故が起こった時の，条件や用件が明確にできるなら，その条件または要因を取り除くか，回避することで事故を減らせるのではないかと仮説を立ててみた．そのためには，事故の状況を明確に評価でき，分析可能なものの必要性を感じた．その方法について提案概要を述べる．

表 3.4 顧客固有語登録例

シェフイール	デニーズ	ロイヤルホテル	よつ葉乳業	GFBB
内箱	スタッフ	使用量	配合室	端数合わせ
荷抜け	半端	依頼書	凍結シビルジュース	外箱
収箱	ペール缶	含有量	配管内	荷受
製品数	釜数	発注専用	ストレージ	ケーサーオペレーター
積みつけ	保証書	配合表	ホッパー	コーダージェット
デジタル秤	早出セットアップ担当者	記入ミス	冷凍原料	内部時計
廃棄枚数	OPQC: パレターキューブ	Wチェックシステム	初歩的	
明確化	チェック機能	出庫依頼書	類似事故	
未設置	プリメルト煮熟管理表	入庫時	認知	
最終包装作業	半端ロット	管理表		
計量作業	明確化	溶解比率表	再確認	
カートンシーラー	品名	溶解	配合ミス	
うっかりミス	昼連	切替製造	アラハタ興産	
ケース重量	缶蓋	根本的	殺菌温度	
シートカウンター	配列状態	連動	充填温度	
使用枚数		Brix	1号釜	
包装最終記録表			2号釜	
作業内容				
仕組み				

表 3.5 ストップワード（削除語）例

(	い	から	する
)	いう	が	た
,	いただき	こと	だ
.	いる	これ	て
「	う	ごさい	で
」	おmoi	さ	です
?	お願い	さらに	ところ
→	か	さん	どの

JaJa[15-21-15-15]

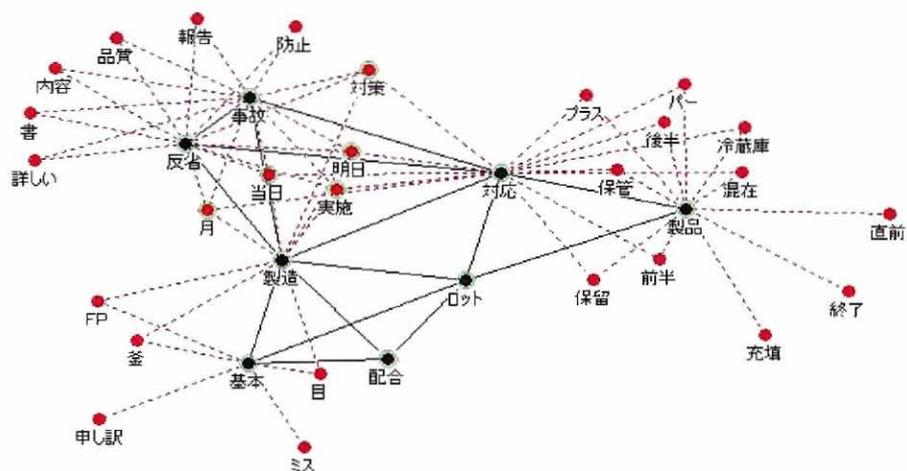


図 3.5 キーグラフからの出力

図 3.6 の提案概要はその解決策として、A 食品会社に提案したものである。提案時にそのサンプルとして筆者が事前に事故再発防止報告書からキーグラフを用いて作ったデシジョンテーブルを成果物のイメージとして提示した（表 3.6 デシジョンテーブル第 1 版）。

提案の要点は

- ①課題・解決策キーワードから事故を明確に見える形にするための表を作る。判断・判定の意味も持つデシジョンテーブルと命名した。
- ②デシジョンテーブルに事故再発防止報告書からデータを抽出し、プロットされたデータからルールまたは要因を取り出す手段として、ラフ集合法、多変量解析法を用いる。
- ③取り出された、ルール、または要因が事故再発防止策となり、現場が実行（守る）ことで事故防止となるはずである。
- ④デシジョンテーブルは事故発生時の日報報告に追加資料として記入することでチェックリストにもなり事故の環境が明確になるとともに、分析に利用できる。
- ⑤デシジョンテーブルは時間とともに見直され、更改され、そこから新たな知見

がなされ、ナレッジ化へと発展できる。  
と、説明し了承を得る。

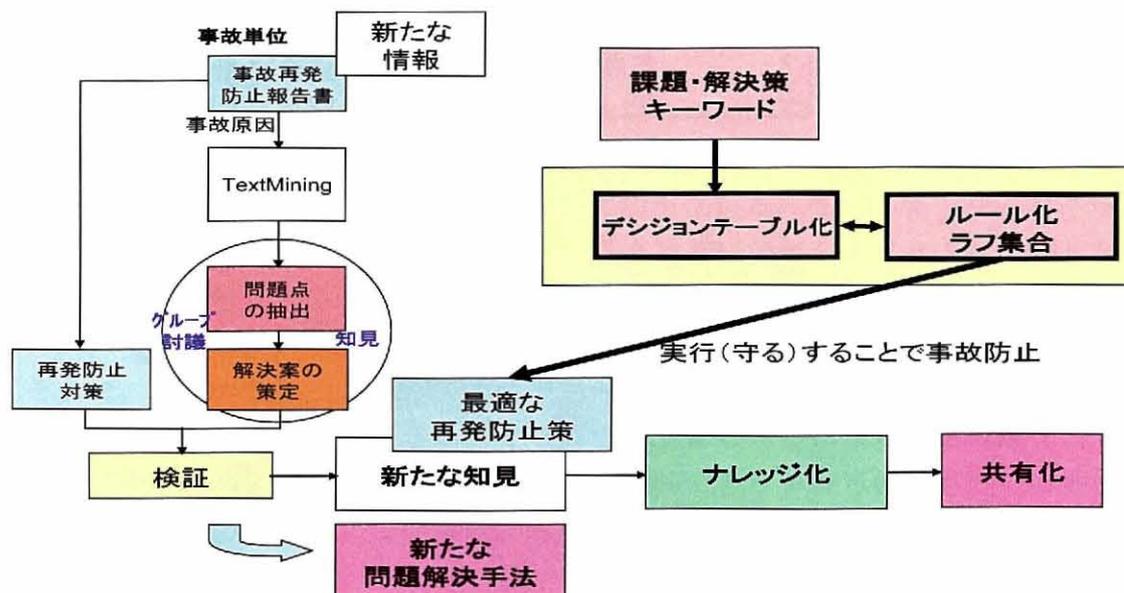


図 3.6 提案概要

### 1. 事故再発防止報告書の共通認識

改善要件

- ・書き方の標準化
- ・用語の統一
- ・フォーマットの再考
- ・固有語登録
- ・単語置換登録
- ・ストップワード(削除語)登録

### 2. キーグラフの見方, シナリオ創生

3. キーワードの抽出・・・デシジョンテーブルの条件属性決め (以下デシジョン項目と記す)

4. 条件属性の指標決め (以下カテゴリと記す)

5. サンプルデータをテーブル化

6. デシジョンテーブルの最適化

上記, 1, 2 については作業続行とし, 3 についてプロジェクトのメンバーとキーグラフを見ながら, 第1版のデシジョンテーブルの整合性を検討した. メンバーにて討議したことをキーグラフに入力し, チャンス発見で云う二重螺旋プロ

セスの実践を行い、デシジョン項目の決定、カテゴリの決定作業に入る。その結果デシジョンテーブルの第2版（表3.7参照）が作成された。

表 3.6 デシジョンテーブル第1版

デシジョンテーブル		2007.9 第一版				
事故No						
工程名						
作業者	自社社員:1	興産社員:2	派遣社員:3			
作業グループ 人数						
熟練度	半年以内:1	1年以内:2	3年以内:3	5年以内:4	5年以上:5	
工程の 変化	多い:1	少ない:1				
製造 度合い	毎日:1	週1~3回:2	半月:3	月:4	半年:5	
作業 指図書	あり:1	なし:2				
作業 手順書	詳細あり:1	概要あり:2	なし:3			
作業 チェック表	あり:1	なし:2				
作業 チェック者	あり:1	なし:2				
作業者 判断	あり:1	なし:2				
作業確認手段	計量:1	分析:2	その他:3	なし:4		
製造ノウハウの熟練度	高い:1	標準:2	低い:3			
事故発生	あり:1	なし:2				
発生時間	午前:1	午後:2	夜間:3			
ライン停止時間(分)						
損失金額(円)						
発生日時						
前回発生日時						
コメント						

表 3.7 デシジョンテーブル第2版

事故報告時のデシジョンテーブル		2008.12 第二版						
2008年度								
項目	カテゴリ-1	カテゴリ-2	カテゴリ-3	カテゴリ-4	カテゴリ-5	カテゴリ-6	カテゴリ-7	
発生年月日	年 月 日		報告者:					
工程名	1:配列 8:殺菌・冷却	2:配合 9:包装	3:濃縮 10:出荷	4:検査 11:その他	5:殺菌 12:スタッフ	6:充填	7:密封	
デシジョン項目	作業者	1:自社	2:興産	3:協力	4:アルバイト			
	作業人数	1:1人	2:2人	3:3人	4:4人	5:5人	6:5人以上	
	作業経験年数	1:半年以内	2:1年以内	3:3年以内	4:5年以内	5:5年以上		
	製造品目	1:17アイテム	2:27アイテム	3:37アイテム以上				
	製造頻度	1:週1	2:週2以上	3:半月1	4:月1	5:半年1		
	作業交替	1:なし	2:あり					
	手順書	1:なし	2:あり					
	チェック表	1:なし	2:あり					
	チェック者	1:なし	2:あり					
	作業者判断	1:なし	2:あり					
	作業確認手段	1:なし	2:人	3:機械・分析	4:不明			
	熟練度	1:低い	2:標準	3:高い	4:不明			
発生時間	1:スタート	2:午前	3:昼連	4:午後	5:二部	6:終了	7:不明	
発生曜日	1:日曜日	2:月曜日	3:火曜日	4:水曜日	5:木曜日	6:金曜日	7:土曜日	
評価項目	事故発生	1:なし	2:あり					
	ライン停止時間		分					
	損失金額		円					
	前回発生日時		年月日					
コメント								

5のサンプルデータをテーブル化は第2版のデシジョンテーブルに過去に発生した、事故再発防止報告書からプロットを依頼した。結果は、2007年、2008年の事故再発防止報告書から68件のサンプルデータが作成されたが、埋まらない項目(欠損値)があることが判明したので、54件のデータを分析の対象とした。

表 3.8 デシジョンテーブルにプロットされたサンプルデータ (一部)

デシジョンテーブル 2007年10月~2008年9月

No	発生日	工程名	作業者	作業人数	作業経験年数	製造品目	製造頻度	作業交替	手順書	チェック表	チェック者	作業者判断	作業確認手段	熟練度	発生時間	発生曜日	事故発生	ライン停止	損失金額
		1:配列	1:自社	1:1人	1:半年以内	1:17行ム	1:週1	1:なし	1:なし	1:なし	1:なし	1:なし	1:なし	1:低い	1:スタート	1:日曜日	1:なし		
		2:配合	2:異産	2:2人	2:1年以内	2:27行ム	2:週2以上	2:あり	2:あり	2:あり	2:あり	2:あり	2:人	2:標準	2:午前	2:月曜日	2:あり		
		3:濃縮	3:協力	3:3人	3:3年以内	3:37行ム以上	3:半月1						3:機械・分析	3:高い	3:昼連	3:火曜日			
		4:検査	4:7Mバ	4:4人	4:5年以内		4:月1						4:不明	4:不明	4:午後	4:水曜日			
		5:殺菌		5:5人	5:5年以上		5:半年1								5:二部	5:木曜日			
		6:充填		6:5人以上											6:終了	6:金曜日			
		7:密封													7:不明	7:土曜日			
		8:殺菌・冷却																	

No	発生日	工程名	作業者	作業人数	作業経験年数	製造品目	製造頻度	作業交替	手順書	チェック表	チェック者	作業者判断	作業確認手段	熟練度	発生時間	発生曜日	事故発生	ライン停止	損失金額
1	2003.12.1	釜	1	1	5	1	4	1	2	2	2	2	3	2	7	2	2		
2	2004.10.27	釜	2	5	3	1	1	1	2	1	1	1	2	2	2	4	2	40	
3	2004.2.27	ベラボ量室	12	6	5	1	4	1	1	1	1	2	2	2	3	6	2		4,800,000
4	2004.2.3	洗浄	1	6	5	2	4	1	1	1	1	1	2	1	7	3	1		2,200,000
5	2004.4.1	包装	1	1	5	1	1	1	1	1	1	2	2	2	7	5	2		12,000
6	2004.4.19	ピン供給	1	1	5	1	1	1	1	1	1	2	3	2	7	2	2	22	

表 3.8 は実際にデシジョンテーブルにプロットされたサンプルデータの一部である。

### 3.4 提案する課題発見プロセス

前項まで述べたことを、この項では課題発見プロセス(図 3.7 参照)として述べる。本研究の目的は産業界に適用可能なテキストからの課題発見プロセスをいかに構築するかにある。しかし、研究途中で新たな課題に突き当たった、それは課題発見という行為をどんな形で表現できるのか、また、産業界で受け入れてもらえるには

- ① 結果が形で見えるものでなければならない
- ② 何か役に立つアウトプットが必要条件
- ③ 新たな問題解決への足がかりの提供

などではないかと考えられた。

図 3.7 は課題発見プロセスとしてその流れと機能を表したものである。今回は、A 食品会社との連携ということでより具体的なプロセスで表現した。

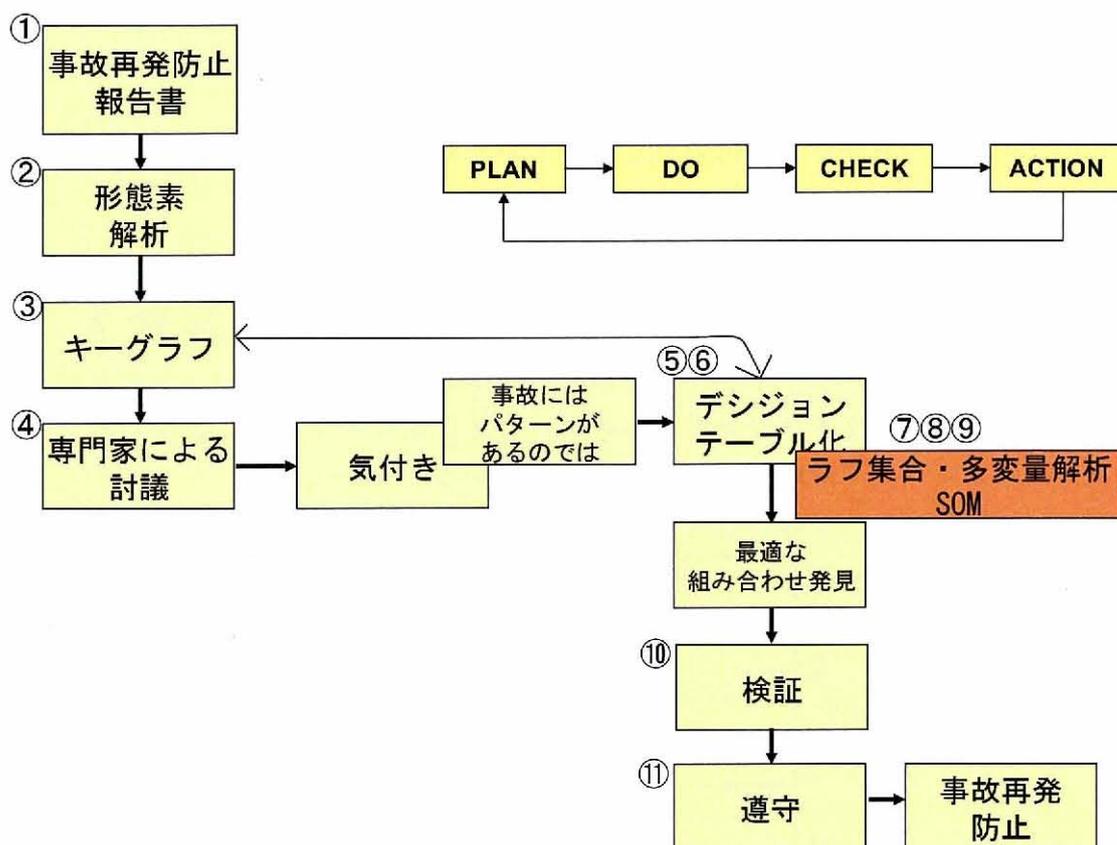


図 3.7 課題発見プロセス

課題発見プロセスの流れを以下に述べる。

- ① 対象テキストデータのどの部分を使用するかが重要である。  
本研究では事故再発防止報告書の事故報告のテキストが対象である。考慮すべきことは報告書の書き方の統一、用語の統一などが必要である。
- ② 形態素分析を行なう。  
今回は、オープンソフトの茶筌を使用した。
- ③ 茶筌で出力された、形態素データを固定語辞書、単語置換辞書、ストップワード(削除語)辞書を使用して単語の整形を行ない、キーグラフにてキーワードの抽出、可視化機能でキーグラフを表示する。
- ④ 表示されたキーグラフを見ながら専門家によるグループ討議がなされ、討議された会話(気づきという)を、チャンス発見プロセスの二重螺旋プロセスを経て、共通認識されたキーワードから課題の抽出が行なわれる。

今回は事故にはパターンが存在するのではないかと素人考えが功を奏した。

- ⑤ 共通認識されたキーワードからデシジョンテーブル項目が導かれる。
- ⑥ デシジョンテーブルに実データをプロットする。
- ⑦ 分析作業のためのデータ形成をおこなう。
- ⑧ 形成されたデータを分析ツールにて分析する。
- ⑨ 分析した結果の読み取り。

今回は、事故を未然に防ぐための手段を読み取る。

- ⑩ この手段の有効性の検証をおこなう。
- ⑪ テスト的に適用し、現場ではその方式を遵守する。

このプロセスはPDCAプロセスに則り繰り返されることにより、精度向上が図れる。

また、リスク・マネジメントからの図 3.8 のリスク対応プロセスからみると、④以降のリスクの発生からのプロセスに重なりがあり、⑨から⑪までが課題発見プロセスで対応可能であることがわかる。

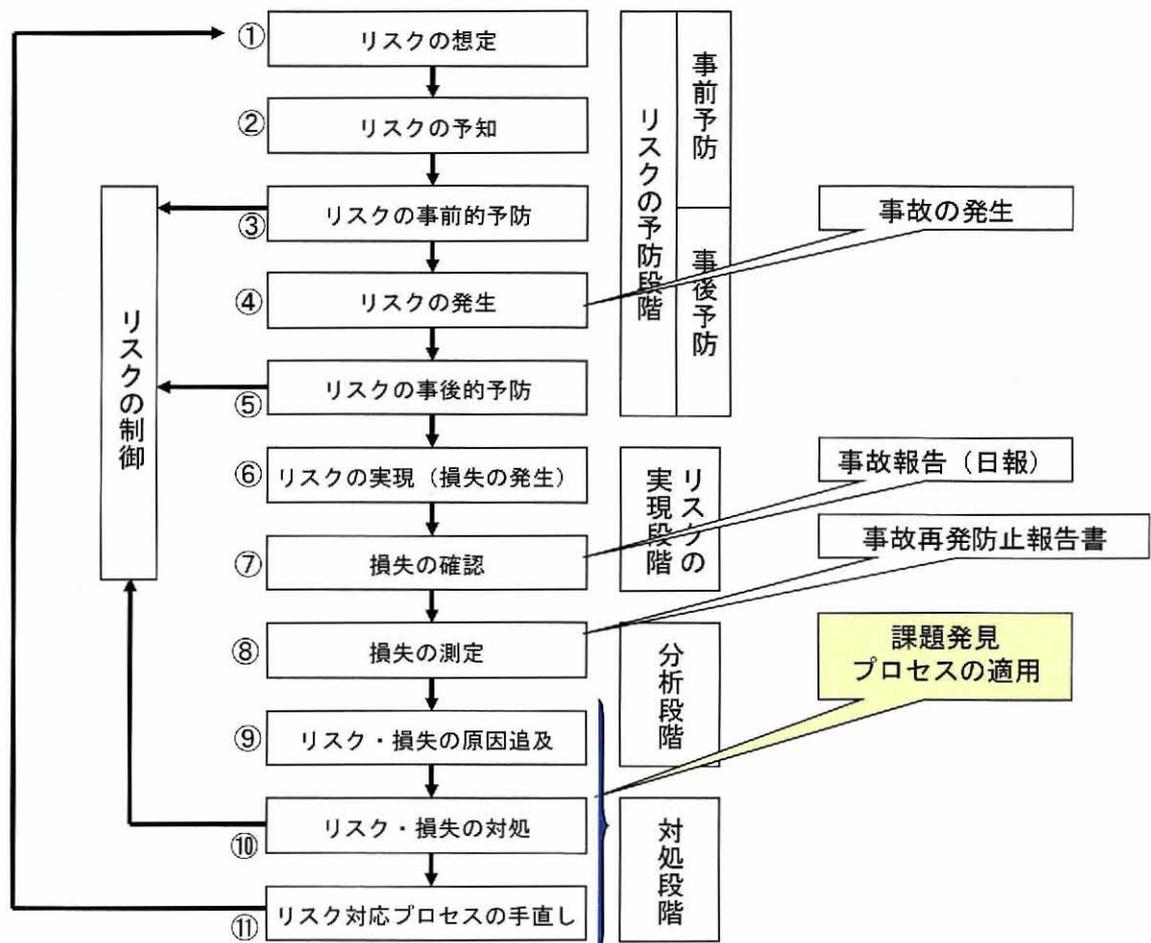


図 3.8 リスク対応プロセス

# 第4章 イノベーションを支援する リスク分析

## 4.1 A食品会社への適用

A食品会社では、過去の事故再発防止報告書をもとに、新たに作成された事故報告時のデシジョンテーブル(表 3.7 参照)にサンプルとしてプロットしてみた。このデータを使用して、さらに目論んでいた課題発見からの新たな解決策を探るべく、分析を行った。分析には欠損値のない 54 件のデータを選び、第 3 章で提案した 3 つの方法で分析し、分析結果の解析から、興味ある答えを得た。その報告と評価をこの章で述べる。 [63,64,65,66,67,90]

### 4.1.1 回帰分析の適用

#### 4.1.1.1 回帰分析のためのデータ構造

##### 1. 入力データ

デシジョンテーブルにプロットされたサンプルデータを回帰分析の多変量解析法を適用するためにデータ形成を行う。要因分析が目的であるが要因をより詳細に分析するために、ダミー変数(表 4. 1 多変量解析データ参照)を用いた。この方式は数量化理論 I 類に沿ったものである。その結果、説明変数の数が 39 個となり Excel の制限である 16 個を超過したので、SPSS Statistics 17 (以下 SPSS と記す) を Windows の環境で線形回帰分析機能を適用した。

データ件数は 54 件である。内訳は事故データが 29 件、事故なし（ヒヤリハット）が 25 件である。A 食品会社の事故再発防止報告書での報告は全社で年間 100 件程度である。今回、適用した本社工場では年間約 50 件である。

回帰分析での注意点は、カテゴリのデータが微小の場合、大きな回帰係数が計算され、間違った判断になりかねない。当検証作業においても発生した。

発生時間での M5（二部）でのデータが 1 件しか発生しなかったため、異常値が検出された。その対応として検討した結果、大きな影響が認められなかったため M6（終了）のデータと結合して分析を行った。

表 4.1 多変量解析データ（部分）

作業者			作業人数			作業経験年数				製造品目				製造頻度				交替手順書				確認手段				熟練度			発生時間						発生曜日						事故
A1	A2	A3	B1	B2	B3	C1	C2	C3	C4	D1	D2	E1	E2	E3	E4	F1	G1	H1	I1	J1	K1	K2	K3	L1	L2	L3	M1	M2	M3	M4	M5	M6	N1	N2	N3	N4	N5	N6	事故		
1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1		
0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	1	1	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	2	
1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	2	
1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	
1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	
1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	
1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	
1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	1	1	1	1	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	2	
1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	1	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	
1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	1	1	1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	2	
1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	
1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	

## 2. SPSS 出力データからの影響度分析

SPSS での分析結果を以下に示す（表 4.2 参照）。

デシジョンテーブルで設定した項目 14 個での回帰分析結果である。回帰係数での異常値（1.0 以上）は見られないので、正当性はあると判断する。

表 4.3 は回帰分析結果から影響度分析を行った表である。また、図 4.1 は影響度分析した結果をグラフ化したものである。

影響度分析結果のグラフ（図 4.1 参照）から、事故に影響の大きい項目は

- ①作業経験年数
- ②発生曜日
- ③発生時間
- ④熟練度
- ⑤製造頻度

の順である。

以上の結果はいたって常識的見解であり、誰もが常識的に感じている要因ではないかと思われる。

表 4.2 SPSS での回帰分析結果

モデル要約						
モデル	R	R2 乗	調整済み R2 乗	標準偏差推定値の誤差		
1	.789 <sup>a</sup>	.622	-.054	.509		
RU	0.254836					
分散分析 <sup>b</sup>						
モデル		平方和 (分散成分)	自由度	平均平方	F 値	有意確率
1	回帰	8.111	34	.239	.920	.596 <sup>a</sup>
	残差 (分散分析)	4.926	19	.259		
	合計 (ピボットテーブル)	13.037	53			
係数 <sup>a</sup>						
モデル		標準化されていない係数 B	標準偏差誤差	標準化係数 ベータ	t 値	有意確率
1	(定数)	1.942	.690		2.816	.011
	作業者A2	-.099	.504	-.068	-.196	.846
	A3	-.005	.539	-.003	-.009	.993
	作業人数B2	-.128	.262	-.105	-.489	.631
	B3	-.099	.372	-.072	-.267	.792
	B4	-.003	.366	-.002	-.009	.993
	作業経験年数C1	.899	.671	.247	1.340	.196
	C2	.456	.647	.175	.706	.489
	C3	-.167	.418	-.132	-.400	.693
	C4	-.453	.543	-.211	-.835	.414
	製造品目D1	-.107	.231	-.108	-.462	.650
	D3	-.227	.385	-.134	-.589	.563
	作業頻度E3	-.031	.492	-.012	-.063	.950
	E4	-.466	.272	-.474	-1.713	.103
	E5	-.209	.273	-.182	-.767	.452
	交替F1	-.106	.268	-.087	-.396	.696
	手順書G2	.147	.336	.149	.438	.666
	チェック表H2	.177	.696	.175	.254	.802
	チェック者I1	-.115	.488	-.110	-.235	.817
	作業者判断J2	-.248	.373	-.210	-.664	.515
	確認手段K1	.031	.380	.018	.082	.935
	K3	-.290	.362	-.269	-.800	.434
	熟練度L1	-.008	.401	-.005	-.020	.985
	L3	.050	.402	.030	.125	.902
	L4	.591	.438	.315	1.347	.194
	発生時間M1	-.286	.425	-.183	-.674	.509
	M2	.103	.389	.066	.265	.794
	M3	.488	.574	.188	.851	.406
	M4	.332	.490	.155	.678	.506
	M6	-.237	.299	-.188	-.794	.437
	発生曜日N3	.869	.378	.628	2.299	.033
	N4	.217	.340	.178	.640	.530
	N5	.119	.366	.086	.324	.749
	N6	.576	.318	.514	1.812	.086
	N7	.043	.533	.017	.081	.937

表 4.3 影響度分析

項目番号	項目識別名	項目名	影響度	MAX	MIN
1	A	1.作業者	0.099032	.000	-.099
2	B	2.作業人数	0.128090	.000	-.128
3	C	3.作業経験年数	1.066296	.899	-.167
4	D	4.製造品目	0.226810	.000	-.227
5	E	5.製造頻度	0.466416	.000	-.466
6	F	6.交替	0.106362	.000	-.106
7	G	7.手順書	0.147203	.147	.000
8	H	8.チェック表	0.176535	.177	.000
9	I	9.チェック者	0.114639	.000	-.115
10	J	10.作業者判断	0.247686	.248	.000
11	K	11.確認手段	0.320835	.031	-.290
12	L	12.熟練度	0.598404	.591	-.008
13	M	13.発生時間	0.774476	.488	-.286
14	N	14.発生曜日	0.868738	.869	.000

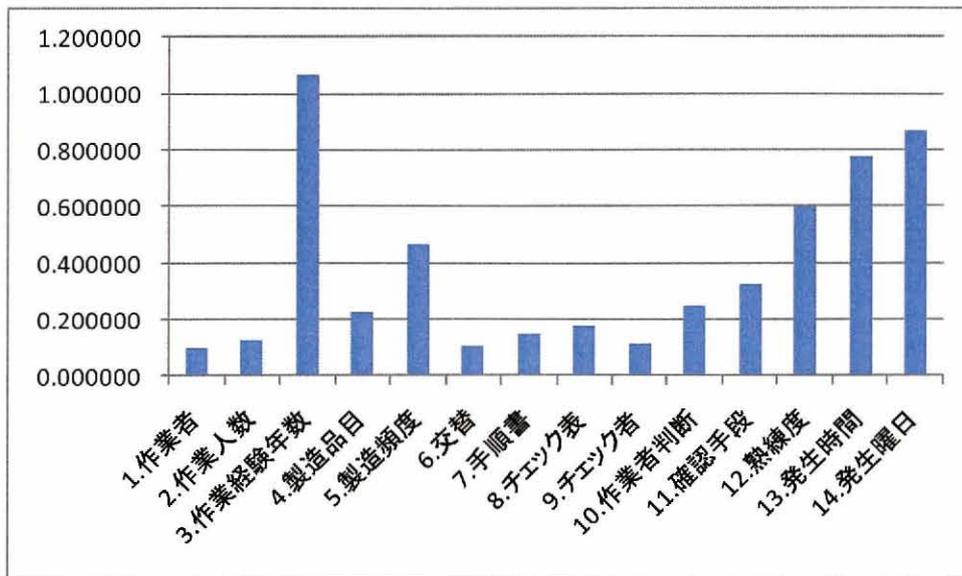


図 4.1 影響度グラフ

### 3. 説明変数選択基準 (Ru) で最適な項目の選択

この分析における最適なアイテム (説明変数) グループを見つけるために、説明変数選択基準 (Ru) を適用した。

説明変数選択基準 (Ru) を求める数式は 45 ページ参照。

表 4.4 は説明変数選択基準 (Ru) 評価表である。

最適な項目の組み合わせは Ru が正の数で最大となるときの組み合わせである。

パターン別の項目の削除基準はカテゴリスコアの有意確率（P 値）の最大値を含む項目から順次削除し Ru を算出している。その結果 Ru が最大となる項目の組み合わせはパターン No. 3 の項目数 12 個となった。情報収集には少ないデータで最大の効果を出すことが大切である。その意味でも 12 個に限定できたことは、製造現場での情報収集に威力を発揮するであろう。

また、この項目群はデシジョンテーブルに反映することになる。

以下にその項目群名とその意味を示す。

- ①作業人数・・・作業にかかわる作業者の人数
- ②作業経験年数・・・事象を起こした作業者の作業経験年数
- ③製造品目・・・事象を起こした時の製造品目
- ④製造頻度・・・事象を起こした製造品目の製造頻度数
- ⑤交替・・・2 交替制の有無
- ⑥手順書・・・作業手順書の有無
- ⑦チェック表・・・作業進捗チェック表の有無
- ⑧作業判断・・・事象が起きた時の判断が作業責任の有無
- ⑨確認手段・・・作業の確認の方法
- ⑩熟練度・・・作業の熟練度要件
- ⑪発生時間・・・事象の発生時間
- ⑫発生曜日・・・事象の発生曜日

事象とは、事故またはヒヤリハット情報の発生時の作業および行動を指す。

表 4.4 説明変数選択基準（Ru）評価表

項目番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
項目 識別名	A	B	C	D	E	F	G	H	I	J	K	L	M	N		
パターン No.	作業者	作業人数	作業経験 年数	製造品目	製造頻度	交替	手順書	チェック 表	チェック 者	作業者 判断	確認手段	熟練度	発生時間	発生曜日	RU	R2
1	○	○	○	○	○	○	○	○	○	○	○	○	○	○	0.254836	0.622
2		○	○	○	○	○	○	○	○	○	○	○	○	○	0.282440	0.621
3		○	○	○	○	○	○	○	○	○	○	○	○	○	0.295209	0.620
4		○	○	○	○	○	○	○		○	○	○	○		0.048513	0.432
5		○		○	○	○	○	○		○	○	○	○		-0.026403	0.337
6		○		○	○	○	○	○		○		○	○		-0.002458	0.326
7				○	○	○	○	○		○		○	○		0.044120	0.319
8				○	○	○	○			○		○	○		0.062242	0.319
9				○	○	○	○					○	○		0.077058	0.317
10				○	○	○						○	○		0.091188	0.315
11				○		○						○	○		0.106404	0.297
12				○		○						○			0.057615	0.173
13				○		○									0.005361	0.077
14				○											-0.005874	0.049

以上の項目群での回帰分析を再度行い、評価の対象とした。

## 4.1.2 回帰分析の結果と考察

### (1) 全体影響度分析

回帰分析で計算された回帰係数より全体の影響度を求めた。その分析結果を以下に示す（表 4.5 参照）。

表 4.5 最適な項目での SPSS 回帰分析結果

モデル要約						
モデル	R	R2 乗	調整済み R2 乗	標準偏差推定値の誤差		
1	.787 <sup>a</sup>	.620	.084	.475		
RU	0.295209					
分散分析 <sup>b</sup>						
モデル		平方和 (分散成分)	自由度	平均平方	F 値	有意確率
1	回帰	8.079	31	.261	1.157	.366 <sup>a</sup>
	残差 (分散分析)	4.958	22	.225		
	合計 (ピボットテーブル)	13.037	53			
係数 <sup>c</sup>						
モデル		標準化されていない係数		標準化係数	t 値	有意確率
		B	標準偏差誤差	ベータ		
1	(定数)	1.868	.501		3.725	.001
	作業人数B2	-.147	.238	-.121	-.617	.543
	B3	-.152	.301	-.110	-.507	.617
	B4	-.034	.315	-.023	-.107	.916
	作業経験年数C1	.884	.604	.243	1.464	.157
	C2	.375	.417	.144	.899	.378
	C3	-.206	.247	-.162	-.832	.415
	C4	-.507	.479	-.236	-1.058	.302
	製造品目D1	-.116	.214	-.117	-.539	.595
	D3	-.259	.348	-.153	-.743	.465
	作業頻度E3	-.004	.449	-.001	-.008	.994
	E4	-.443	.246	-.450	-1.801	.085
	E5	-.195	.251	-.170	-.779	.444
	交替F1	-.121	.245	-.099	-.494	.626
	手順書G2	.119	.276	.121	.433	.669
	チェック表	.292	.374	.290	.780	.444
	H2					
	作業者判断	-.236	.307	-.200	-.768	.451
	J2					
	確認手段K1	.051	.338	.030	.151	.881
	K3	-.301	.305	-.279	-.986	.335
	熟練度L1	-.054	.318	-.034	-.169	.868
	L3	.075	.369	.044	.202	.842
	L4	.591	.401	.315	1.475	.154
	発生時間M1	-.317	.366	-.203	-.867	.395
	M2	.109	.332	.070	.329	.745
	M3	.448	.486	.172	.921	.367
	M4	.263	.388	.122	.677	.505
	M6	-.248	.270	-.196	-.918	.369
	発生曜日N3	.869	.339	.629	2.562	.018
	N4	.206	.297	.169	.694	.495
	N5	.115	.330	.083	.348	.731
	N6	.562	.269	.501	2.084	.049
	N7	-.003	.480	-.001	-.006	.995

表 4.6 最適な項目での影響度分析

項目番号	項目記号	項目名	影響度	MAX	MIN
2	B	2.作業人数	0.152	.000	-.152
3	C	3.作業経験年数	1.391	.884	-.507
4	D	4.製造品目	0.259	.000	-.259
5	E	5.製造頻度	0.443	.000	-.443
6	F	6.交替	0.121	.000	-.121
7	G	7.手順書	0.119	.119	.000
8	H	8.チェック表	0.292	.292	.000
10	J	10.作業者判断	0.236	.000	-.236
11	K	11.確認手段	0.352	.051	-.301
12	L	12.熟練度	0.645	.591	-.054
13	M	13.発生時間	0.765	.448	-.317
14	N	14.発生曜日	0.872	.869	-.003

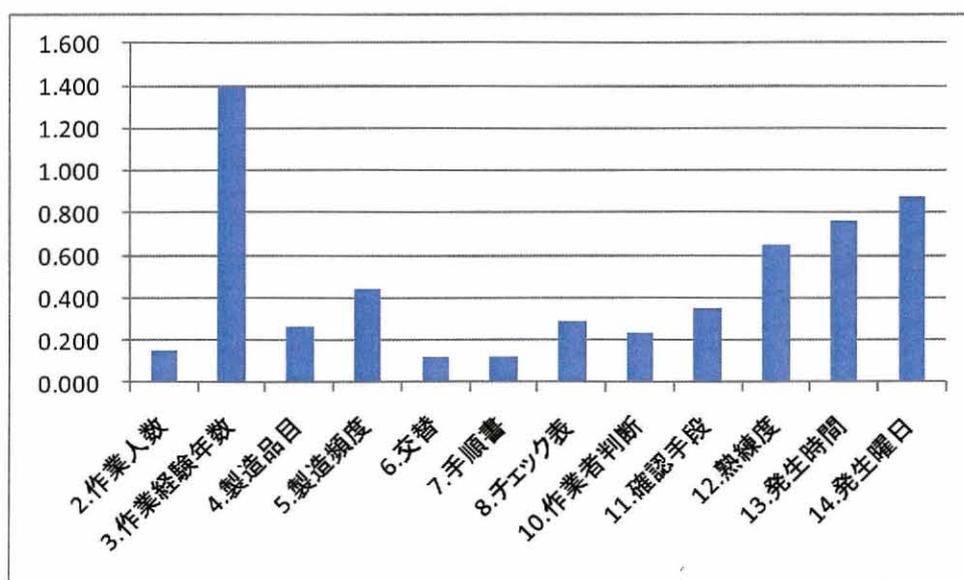


図 4.2 最適な項目での影響度グラフ

影響度分析結果のグラフ（図 4.2 参照）から、事故に影響の大きい項目は

- ①作業経験年数
- ②発生曜日
- ③発生時間
- ④熟練度
- ⑤製造頻度

の順である。

## (2) 各カテゴリの影響度分析

(1)で影響度の大きかった項目のうち3項目について、カテゴリの影響度を計算してみた。以下の表・図(表4.7～図4.5)となる。

### ①作業経験年数での影響度

表4.7および図4.3より経験年数での影響度を見てみると、作業従事半年以内の作業者が事故を起こしている可能性が大である。また、2番目に多いのは経験年数3年以上5年以内の作業者である。これは、作業の慣れによる気の緩みが原因かと考えられる。その対策としては、再教育の時期は作業従事後3年が目安とも思われる。

表 4.7 経験年数別影響度

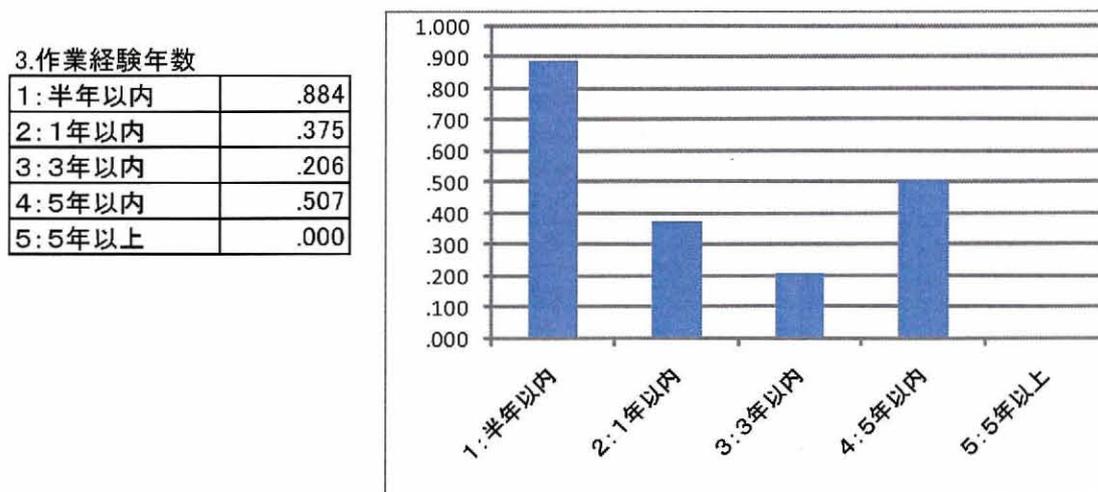


図 4.3 経験年数別影響度グラフ

### ②発生曜日での影響度

表4.8および図4.4より発生曜日での影響度を見てみると、火曜日、金曜日に事故を起こしている可能性が大である。火曜日は休み明けの月曜日の緊張からの解放による気の緩みからの事故である可能性がある。また、金曜日は休日前日という気の緩みが考えられる。また、一週間の疲れが蓄積した結果、集中度が欠落し事故に繋がっているとも解釈できる。これらの事故対策としては朝礼等での注意を喚起する行動が事故抑制に必要となる。

表 4.8 発生曜日別影響度

14.発生曜日	
1:日曜日	.000
2:月曜日	.000
3:火曜日	.869
4:水曜日	.206
5:木曜日	.115
6:金曜日	.562
7:土曜日	.003

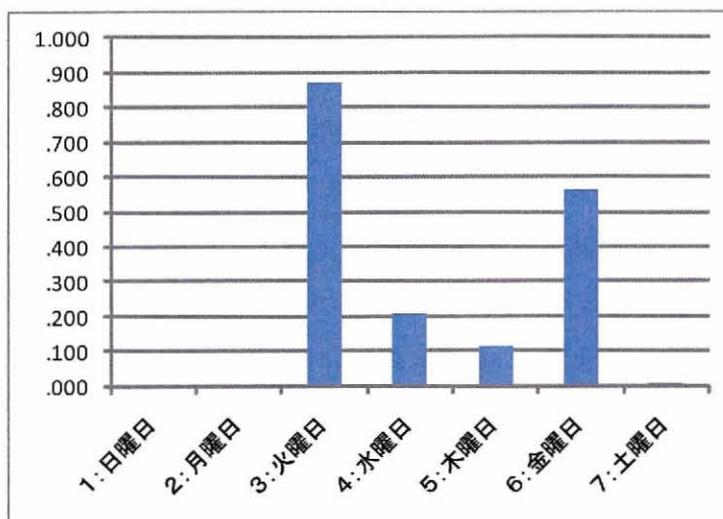


図 4.4 発生曜日別影響度グラフ

### ③発生時間での影響度

表 4.9 および図 4.5 より発生時間での影響度を見てみると、昼連での事故が多い。昼連とは昼食なしの連続作業である。この作業は急ぎの納期などで発生するが、空腹による集中度の欠如が原因と考えられる。また、あさのスタート時にも多い、このことは、あさの緊張感不足が何らかの関連で影響しているものと思われる。

表 4.9 発生時間別影響度

13.発生時間	
1:スタート	.317
2:午前	.109
3:昼連	.448
4:午後	.263
6:二部・終了	.248
不明	.000

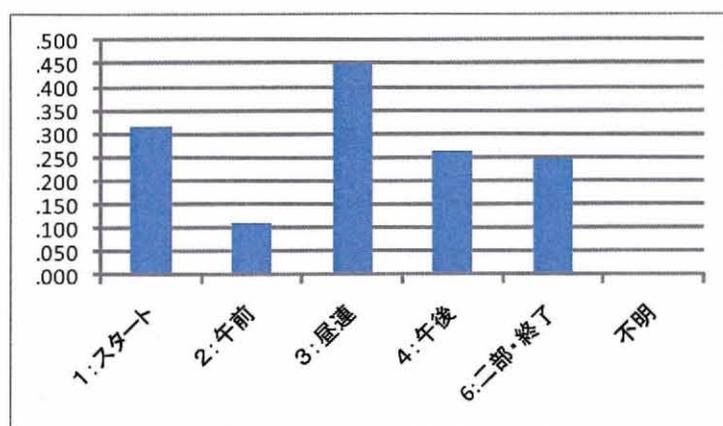


図 4.5 発生時間別影響度グラフ

## 最適な項目での最適な回帰式

説明変数選択基準 (Ru) で選択されたアイテムとカテゴリから最適な回帰式を求めることができる。表 4.10 は求められた回帰式である。この回帰式から以下のことが考察できる。被説明変数  $y$  は事故を意味しており、回帰式に実際のデータを代入して計算させた結果は  $y = 1$  に近ければ、事故誘発のマイナス要因のカテゴリ群と考えられる。表 4.11 では、アイテムのカテゴリに下線のある項目群が事故なし、つまり事故誘発のマイナス要因群である。

作業人数アイテムでは 3 人がマイナスカテゴリとなっているので、作業人数が 3 人の場合は事故を引き起こす要因としては低いと解釈できる。以下各アイテムの事故なし時のカテゴリである。

作業経験年数では 5 年以内の経験者。製造品目では一日の製造品目が 3 アイテム以上時である。製造頻度は月 1 回の頻度。作業交替はなしの場合。手順書はなしの場合。手順書なしでも出来る作業ということか。チェック表はなしの場合。

作業判断はありの場合。作業確認手段は機械・分析で行う場合。熟練度は低い場合。熟練度をあまり要求されていない作業とも解釈できる。発生時間はスタート時である。スタート時は緊張しての作業といえる。発生曜日は土曜日となっている。

逆に事故への影響が大きいカテゴリをみしてみる。

作業人数では 1 人, 5 人, 5 人以上である。少なすぎても, 多すぎても事故の誘発要因となる。作業経験年数では半年, 1 年以内が該当する。未熟さが起こすということだろう。製造品目では一日の製造品目が 2 アイテム時である。製造頻度は週 1, 週 2 回以上の場合である。作業交替はありの場合。手順書はありの場合。手順書に問題があるのか, 手順書がないと作業が出来ないものなのか疑問である。チェック表はありの場合で係数が大きいのは原因の調査が必要である。作業判断はなしの場合。作業確認手段はなし場合。熟練度は高い場合。熟練度の高い作業ということでリスクが高いとも考えられる。また, 作業従事者も限定されるといえる。発生時間は昼連と午後である。昼連は昼食抜きでの連続作業で集中力を欠くこととなり事故要因である。午後は昼食後のリラックスが緊張緩和となり, 集中力を失うことで事故を引き起こす要因となる。などが考察できる。

表 4.10 最適な回帰式

y (事故) = 1.868 (定数)

<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">作業人数</td></tr> <tr><td>B1 (1人)</td><td>0</td></tr> <tr><td>B2 (2人)</td><td>-0.147</td></tr> <tr><td><u>B3 (3人)</u></td><td><u>-0.152</u></td></tr> <tr><td>B4 (4人)</td><td>-0.034</td></tr> <tr><td>B5 (5人)</td><td>0</td></tr> <tr><td>B6 (5人以上)</td><td>0</td></tr> </table>	作業人数		B1 (1人)	0	B2 (2人)	-0.147	<u>B3 (3人)</u>	<u>-0.152</u>	B4 (4人)	-0.034	B5 (5人)	0	B6 (5人以上)	0	<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">作業経験年数</td></tr> <tr><td>C1 (半年以内)</td><td>0.884</td></tr> <tr><td>C2 (1年以内)</td><td>0.375</td></tr> <tr><td>C3 (3年以内)</td><td>-0.206</td></tr> <tr><td><u>C4 (5年以内)</u></td><td><u>-0.507</u></td></tr> <tr><td>C5 (5年以上)</td><td>0</td></tr> </table>	作業経験年数		C1 (半年以内)	0.884	C2 (1年以内)	0.375	C3 (3年以内)	-0.206	<u>C4 (5年以内)</u>	<u>-0.507</u>	C5 (5年以上)	0				
作業人数																															
B1 (1人)	0																														
B2 (2人)	-0.147																														
<u>B3 (3人)</u>	<u>-0.152</u>																														
B4 (4人)	-0.034																														
B5 (5人)	0																														
B6 (5人以上)	0																														
作業経験年数																															
C1 (半年以内)	0.884																														
C2 (1年以内)	0.375																														
C3 (3年以内)	-0.206																														
<u>C4 (5年以内)</u>	<u>-0.507</u>																														
C5 (5年以上)	0																														
<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">製造品目</td></tr> <tr><td>D1 (1アイテム)</td><td>-0.116</td></tr> <tr><td>D2 (2アイテム)</td><td>0</td></tr> <tr><td><u>D3 (3アイテム以上)</u></td><td><u>-0.259</u></td></tr> </table>	製造品目		D1 (1アイテム)	-0.116	D2 (2アイテム)	0	<u>D3 (3アイテム以上)</u>	<u>-0.259</u>	<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">作業頻度</td></tr> <tr><td>E1 (週1)</td><td>0</td></tr> <tr><td>E2 (週2以上)</td><td>0</td></tr> <tr><td>E3 (半月1)</td><td>-0.04</td></tr> <tr><td><u>E4 (月1)</u></td><td><u>-0.443</u></td></tr> <tr><td>E5 (半年1)</td><td>-0.195</td></tr> </table>	作業頻度		E1 (週1)	0	E2 (週2以上)	0	E3 (半月1)	-0.04	<u>E4 (月1)</u>	<u>-0.443</u>	E5 (半年1)	-0.195										
製造品目																															
D1 (1アイテム)	-0.116																														
D2 (2アイテム)	0																														
<u>D3 (3アイテム以上)</u>	<u>-0.259</u>																														
作業頻度																															
E1 (週1)	0																														
E2 (週2以上)	0																														
E3 (半月1)	-0.04																														
<u>E4 (月1)</u>	<u>-0.443</u>																														
E5 (半年1)	-0.195																														
<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">交替</td></tr> <tr><td><u>F1 (なし)</u></td><td><u>-0.121</u></td></tr> <tr><td>F2 (あり)</td><td>0</td></tr> </table>	交替		<u>F1 (なし)</u>	<u>-0.121</u>	F2 (あり)	0	<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">手順書</td></tr> <tr><td><u>G1 (なし)</u></td><td><u>0</u></td></tr> <tr><td>G2 (あり)</td><td>0.119</td></tr> </table>	手順書		<u>G1 (なし)</u>	<u>0</u>	G2 (あり)	0.119																		
交替																															
<u>F1 (なし)</u>	<u>-0.121</u>																														
F2 (あり)	0																														
手順書																															
<u>G1 (なし)</u>	<u>0</u>																														
G2 (あり)	0.119																														
<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">チェック表</td></tr> <tr><td><u>H1 (なし)</u></td><td><u>0</u></td></tr> <tr><td>H2 (あり)</td><td>0.292</td></tr> </table>	チェック表		<u>H1 (なし)</u>	<u>0</u>	H2 (あり)	0.292	<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">作業者判断</td></tr> <tr><td>J1 (なし)</td><td>0</td></tr> <tr><td><u>J2 (あり)</u></td><td><u>-0.236</u></td></tr> </table>	作業者判断		J1 (なし)	0	<u>J2 (あり)</u>	<u>-0.236</u>																		
チェック表																															
<u>H1 (なし)</u>	<u>0</u>																														
H2 (あり)	0.292																														
作業者判断																															
J1 (なし)	0																														
<u>J2 (あり)</u>	<u>-0.236</u>																														
<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">作業確認手順</td></tr> <tr><td>K1 (なし)</td><td>0.051</td></tr> <tr><td>K2 (人)</td><td>0</td></tr> <tr><td><u>K3 (機械・分析)</u></td><td><u>-0.301</u></td></tr> <tr><td>K4 (不明)</td><td>0</td></tr> </table>	作業確認手順		K1 (なし)	0.051	K2 (人)	0	<u>K3 (機械・分析)</u>	<u>-0.301</u>	K4 (不明)	0	<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">熟練度</td></tr> <tr><td><u>L1 (低い)</u></td><td><u>-0.054</u></td></tr> <tr><td>L2 (標準)</td><td>0</td></tr> <tr><td>L3 (高い)</td><td>0.075</td></tr> <tr><td>L4 (不明)</td><td>0.591</td></tr> </table>	熟練度		<u>L1 (低い)</u>	<u>-0.054</u>	L2 (標準)	0	L3 (高い)	0.075	L4 (不明)	0.591										
作業確認手順																															
K1 (なし)	0.051																														
K2 (人)	0																														
<u>K3 (機械・分析)</u>	<u>-0.301</u>																														
K4 (不明)	0																														
熟練度																															
<u>L1 (低い)</u>	<u>-0.054</u>																														
L2 (標準)	0																														
L3 (高い)	0.075																														
L4 (不明)	0.591																														
<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">発生時間</td></tr> <tr><td><u>M1 (スタート)</u></td><td><u>-0.317</u></td></tr> <tr><td>M2 (午前)</td><td>0.109</td></tr> <tr><td>M3 (昼連)</td><td>0.448</td></tr> <tr><td>M4 (午後)</td><td>0.263</td></tr> <tr><td>M6 (二部・終了)</td><td>-0.248</td></tr> <tr><td>M7 (不明)</td><td>0</td></tr> </table>	発生時間		<u>M1 (スタート)</u>	<u>-0.317</u>	M2 (午前)	0.109	M3 (昼連)	0.448	M4 (午後)	0.263	M6 (二部・終了)	-0.248	M7 (不明)	0	<p>+</p> <table border="1" style="border-collapse: collapse; width: 150px;"> <tr><td colspan="2">発生曜日</td></tr> <tr><td>N1 (日曜日)</td><td>0</td></tr> <tr><td>N2 (月曜日)</td><td>0</td></tr> <tr><td>N3 (火曜日)</td><td>0.869</td></tr> <tr><td>N4 (水曜日)</td><td>0.206</td></tr> <tr><td>N5 (木曜日)</td><td>0.115</td></tr> <tr><td>N6 (金曜日)</td><td>0.562</td></tr> <tr><td><u>N7 (土曜日)</u></td><td><u>-0.03</u></td></tr> </table>	発生曜日		N1 (日曜日)	0	N2 (月曜日)	0	N3 (火曜日)	0.869	N4 (水曜日)	0.206	N5 (木曜日)	0.115	N6 (金曜日)	0.562	<u>N7 (土曜日)</u>	<u>-0.03</u>
発生時間																															
<u>M1 (スタート)</u>	<u>-0.317</u>																														
M2 (午前)	0.109																														
M3 (昼連)	0.448																														
M4 (午後)	0.263																														
M6 (二部・終了)	-0.248																														
M7 (不明)	0																														
発生曜日																															
N1 (日曜日)	0																														
N2 (月曜日)	0																														
N3 (火曜日)	0.869																														
N4 (水曜日)	0.206																														
N5 (木曜日)	0.115																														
N6 (金曜日)	0.562																														
<u>N7 (土曜日)</u>	<u>-0.03</u>																														



表 4.12 は提案ラフ集合での出力を解析するために記載しているデシジョンテーブルである。

表 4.12 分析解読用デシジョンテーブル

1	作業者	1: 自社	2: 興産	3: 協力	4: アルバイト			
2	作業人数	1: 1人	2: 2人	3: 3人	4: 4人	5: 5人	6: 5人以上	
3	作業 経験年数	1: 半年以内	2: 1年以内	3: 3年以内	4: 5年以内	5: 5年以上		
4	製造品目	1: 17アイテム	2: 27アイテム	3: 37アイテム以上				
5	製造頻度	1: 週1	2: 週2以上	3: 半月1	4: 月1	5: 半年1		
6	作業交替	1: なし	2: あり					
7	手順書	1: なし	2: あり					
8	チェック表	1: なし	2: あり					
9	チェック者	1: なし	2: あり					
10	作業者 判断	1: なし	2: あり					
11	作業 確認手段	1: なし	2: 人	3: 機械・分析	4: 不明			
12	熟練度	1: 低い	2: 標準	3: 高い	4: 不明			
13	発生時間	1: スタート	2: 午前	3: 昼連	4: 午後	5: 二部	6: 終了	7: 不明
14	発生曜日	1: 日曜日	2: 月曜日	3: 火曜日	4: 水曜日	5: 木曜日	6: 金曜日	7: 土曜日

## 2. 出力データ

プログラムからの結果はファイルに出力される。以下はその結果の中で、ルールのみを取り出したものである（表 4.13～表 4.16 参照）。当データでの出力パターンは事故なしのときで 2 種類（表 4.13, 4.14）事故ありのときで 2 種類（表 4.15, 4.16）である。

表 4.13 決定ルール結論 1（事故なし①）の時の出力リスト

```

Rule(0.962963)_10332_1
1) IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
2) IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 1 0 0 THEN (1) = 1
3) IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
4) IF (3) = 0 0 1 0 0 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 1
5) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
6) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
7) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 0 1 THEN (1) = 1
8) IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
9) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
10) IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 0 1 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 1 0 0 THEN (1) = 1
11) IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 1
12) IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 1 0 0 0 0 0 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
13) IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
14) IF (3) = 0 0 0 1 0 (4) = 0 1 0 (5) = 0 0 1 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
15) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 1 0 0 0 0 0 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
16) IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
17) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 1 0 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 1
18) IF (3) = 0 0 0 1 0 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 1 0 0 0 0 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 1
19) IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
20) IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
21) IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 1 0 0 0 0 0 0 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
=====> IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 1 0 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 1 0 0 0 0 THEN (1) = 1

```

表 4.14 決定ルール結論 1 (事故なし②) の時の出力リスト

Rule(0.962963)_10524_1	
1)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
2)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 1 0 0 THEN (1) = 1
3)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
4)	IF (3) = 0 0 1 0 0 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 1
5)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
6)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
7)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 0 1 THEN (1) = 1
8)	IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
9)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
10)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 0 1 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 1 0 0 THEN (1) = 1
11)	IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 1
12)	IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 1 0 0 0 0 0 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
13)	IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
14)	IF (3) = 0 0 0 1 0 (4) = 0 1 0 (5) = 0 0 1 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
15)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 1 0 0 0 0 0 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
16)	IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
17)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 1 0 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 1
18)	IF (3) = 0 0 0 1 0 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 1 0 0 0 0 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 1
19)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
20)	IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 1
21)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 1 0 0 0 0 0 0 (14) = 0 0 0 1 0 0 0 THEN (1) = 1
=====>	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 1 0 0 (9) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 1 0 0 0 0 THEN (1) = 1

表 4.15 決定ルール結論 2 (事故あり①) の時の出力リスト

Rule(0.962963)_10332_2	
1)	IF (3) = 0 0 1 0 0 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 0 1 (13) = 0 1 0 0 0 0 0 (14) = 0 0 0 1 0 0 0 THEN (1) = 2
2)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 0 1 0 0 0 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
3)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 0 0 0 0 1 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
4)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 0 1 (13) = 0 0 1 0 0 0 0 (14) = 0 0 0 0 0 0 1 THEN (1) = 2
5)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 1 0 0 0 0 0 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
6)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 1 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
7)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 1 0 0 0 0 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
8)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
9)	IF (3) = 1 0 0 0 0 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 2
10)	IF (3) = 0 0 1 0 0 (4) = 1 0 0 (5) = 0 0 0 0 1 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
11)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 1 0 0 0 0 0 0 (14) = 0 0 0 0 1 0 0 THEN (1) = 2
12)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 0 1 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 1 0 0 THEN (1) = 2
13)	IF (3) = 0 0 0 0 1 (4) = 0 0 1 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 1 0 0 0 (14) = 0 0 0 1 0 0 0 THEN (1) = 2
14)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 0 1 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 2
15)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 1 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
16)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 1 0 0 0 0 0 (14) = 0 1 0 0 0 0 0 THEN (1) = 2
17)	IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 1 0 0 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
18)	IF (3) = 0 0 0 1 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 2
19)	IF (3) = 0 0 1 0 0 (4) = 0 1 0 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
20)	IF (3) = 0 0 1 0 0 (4) = 0 0 1 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
21)	IF (3) = 0 0 1 0 0 (4) = 1 0 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
22)	IF (3) = 0 1 0 0 0 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
23)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 0 1 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
24)	IF (3) = 0 1 0 0 0 (4) = 1 0 0 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
25)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
26)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 1 0 0 0 THEN (1) = 2
27)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 1 0 (13) = 0 0 0 0 0 1 0 (14) = 0 0 0 0 1 0 0 THEN (1) = 2
28)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 0 1 (7) = 1 0 (13) = 0 0 0 0 0 1 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 2
29)	IF (3) = 0 0 0 0 1 (4) = 0 1 0 (5) = 0 0 0 1 0 (7) = 0 1 (13) = 1 0 0 0 0 0 0 (14) = 0 0 0 0 0 1 0 THEN (1) = 2
30)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 0 0 0 1 0 0 THEN (1) = 2
31)	IF (3) = 0 0 0 0 1 (4) = 1 0 0 (5) = 1 0 0 0 0 (7) = 0 1 (13) = 0 0 0 0 0 0 1 (14) = 0 1 0 0 0 0 0 THEN (1) = 2
=====>	IF (3) = 1 0 0 0 0 (4) = 0 1 0 (5) = 0 0 1 0 0 (7) = 0 1 (13) = 0 0 0 1 0 0 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 2

表 4.16 決定ルール結論 2 (事故あり②) の時の出力リスト

Rule(0.962963)_10524_2	
1)	IF (3) = 00100 (4) = 100 (5) = 10000 (7) = 01 (13) = 0100000 (14) = 0001000 THEN (1) = 2
2)	IF (3) = 00001 (4) = 100 (5) = 00010 (7) = 10 (13) = 0010000 (14) = 0000010 THEN (1) = 2
3)	IF (3) = 00001 (4) = 010 (5) = 00010 (7) = 10 (13) = 0000010 (14) = 0010000 THEN (1) = 2
4)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 01 (13) = 0010000 (14) = 0000001 THEN (1) = 2
5)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 10 (13) = 1000000 (14) = 0000010 THEN (1) = 2
6)	IF (3) = 00001 (4) = 100 (5) = 00100 (7) = 10 (13) = 0000001 (14) = 0000010 THEN (1) = 2
7)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 10 (13) = 0100000 (14) = 0000010 THEN (1) = 2
8)	IF (3) = 00001 (4) = 010 (5) = 00010 (7) = 01 (13) = 0000001 (14) = 0010000 THEN (1) = 2
9)	IF (3) = 10000 (4) = 100 (5) = 00010 (7) = 01 (13) = 0000001 (14) = 0100000 THEN (1) = 2
10)	IF (3) = 00100 (4) = 100 (5) = 00001 (7) = 01 (13) = 0000001 (14) = 0010000 THEN (1) = 2
11)	IF (3) = 00001 (4) = 010 (5) = 00010 (7) = 01 (13) = 1000000 (14) = 0000100 THEN (1) = 2
12)	IF (3) = 00001 (4) = 010 (5) = 00001 (7) = 01 (13) = 0000001 (14) = 0000100 THEN (1) = 2
13)	IF (3) = 00001 (4) = 001 (5) = 00010 (7) = 01 (13) = 0001000 (14) = 0001000 THEN (1) = 2
14)	IF (3) = 00001 (4) = 100 (5) = 00001 (7) = 01 (13) = 0000001 (14) = 0001000 THEN (1) = 2
15)	IF (3) = 00001 (4) = 100 (5) = 00010 (7) = 01 (13) = 0000010 (14) = 0000010 THEN (1) = 2
16)	IF (3) = 00001 (4) = 010 (5) = 00010 (7) = 01 (13) = 0100000 (14) = 0100000 THEN (1) = 2
17)	IF (3) = 00100 (4) = 010 (5) = 00010 (7) = 01 (13) = 0001000 (14) = 0010000 THEN (1) = 2
18)	IF (3) = 00010 (4) = 010 (5) = 00010 (7) = 01 (13) = 0000001 (14) = 0001000 THEN (1) = 2
19)	IF (3) = 00100 (4) = 010 (5) = 00001 (7) = 10 (13) = 0000001 (14) = 0000010 THEN (1) = 2
20)	IF (3) = 00100 (4) = 001 (5) = 00001 (7) = 10 (13) = 0000001 (14) = 0000010 THEN (1) = 2
21)	IF (3) = 00100 (4) = 100 (5) = 00010 (7) = 10 (13) = 0000001 (14) = 0010000 THEN (1) = 2
22)	IF (3) = 01000 (4) = 010 (5) = 00010 (7) = 10 (13) = 0000001 (14) = 0000010 THEN (1) = 2
23)	IF (3) = 00001 (4) = 010 (5) = 00001 (7) = 01 (13) = 0000001 (14) = 0000010 THEN (1) = 2
24)	IF (3) = 01000 (4) = 100 (5) = 00001 (7) = 10 (13) = 0000001 (14) = 0010000 THEN (1) = 2
25)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 01 (13) = 0000001 (14) = 0000010 THEN (1) = 2
26)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 10 (13) = 0000001 (14) = 0001000 THEN (1) = 2
27)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 10 (13) = 0000010 (14) = 0000100 THEN (1) = 2
28)	IF (3) = 00001 (4) = 010 (5) = 00001 (7) = 10 (13) = 0000010 (14) = 0010000 THEN (1) = 2
29)	IF (3) = 00001 (4) = 010 (5) = 00010 (7) = 01 (13) = 1000000 (14) = 0000010 THEN (1) = 2
30)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 01 (13) = 0000001 (14) = 0000100 THEN (1) = 2
31)	IF (3) = 00001 (4) = 100 (5) = 10000 (7) = 01 (13) = 0000001 (14) = 0100000 THEN (1) = 2
=====>	IF (3) = 10000 (4) = 010 (5) = 00100 (9) = 01 (13) = 0001000 (14) = 0010000 THEN (1) = 2

### 提案ラフ集合分析の結果と考察

表 4.13 から 4.16 まで 4 種類のリストは提案ラフ集合のプログラムにて出力されたものである。

表 4.17 は提案ラフ集合分析の出力例である。この表を用いて表の見方を述べる。先頭の Rule (0.962963) はヒットの確率を表示しておりこの表示では約 96.3%を意味する。その後ろの数字 10332\_1 は識別番号で枝番はデータの内容を表示している。この場合は事故なしを表している。

また、最終行に出力されるのがルール抽出である。この見方は IF(3)の(3)はデシジョンテーブル項目の 3 番目の作業経験年数を示している。その変換のために表 4.12 の分析解読用デシジョンテーブルを参照すると理解し易い。

また、IF(3) = 00001 は作業経験年数のカテゴリが 5 個あって 1 の数字の場所、つまり 5 番目 : 5 年以上が該当することになる。その上で、表 4.17 のデータ

で見ると①作業経験年数②製造品目③製造頻度④チェック表⑤チェック者⑥発生時間⑦発生曜日の7つの項目を示している。

表 4.17 提案ラフ集合分析の出力例

```

Rule(0.962963)_10332_1
1) IF (3)=00001(4)=100(5)=00010(7)=01(13)=0000001(14)=0100000 THEN (1)=1
2) IF (3)=00001(4)=100(5)=10000(7)=10(13)=0000001(14)=0000100 THEN (1)=1
3) IF (3)=00001(4)=100(5)=10000(7)=10(13)=0000001(14)=0100000 THEN (1)=1
4) IF (3)=00100(4)=100(5)=00010(7)=01(13)=0000001(14)=0000010 THEN (1)=1
5) IF (3)=00001(4)=010(5)=00010(7)=01(13)=0000001(14)=0100000 THEN (1)=1
6) IF (3)=00001(4)=010(5)=00010(7)=01(13)=0000001(14)=0100000 THEN (1)=1
.
=====> IF (3)=00001(4)=010(5)=00100(7)=01(13)=0000001(14)=0010000 THE ←ルール抽出
    
```

表 4.18 は事故なしのルールパターン表である。2つのルールが抽出された。2つのルールは作業経験年数、製造品目、製造頻度、発生時間、発生曜日では共通内容になっているが、チェック表、チェック者では別々の出現となっている。この意味はチェック表での管理の徹底とチェック者の存在が事故発生を抑制していると思われる。作業経験年数では5年以上の経験者、製造品目は日に、2アイテムと製造頻度は半月に1度の割で製造されている品目で、発生時間は不明、これは時間認証ができなかったデータが多く存在した関係である。発生曜日は火曜日となっている。

表 4.18 事故なしのルールパターン

3. 作業経験 数年	4. 製造品目	5. 製造頻度	7. チェック表	9. チェック者	13. 発生時間	14. 発生曜日
5:5年以上	2:2アイテム	3:半月1	2:あり		7:不明	3:火曜日
5:5年以上	2:2アイテム	3:半月1		2:あり	7:不明	3:火曜日

表 4.19 は事故ありのルールパターン表である。2つのルールが抽出された。2つのルールは作業経験年数、製造品目、製造頻度、発生時間、発生曜日では共通内容になっているが、チェック表、チェック者では別々の出現となっている。この意味はチェック表での管理の徹底とチェック者の存在が事故発生を抑

制していると思われるがグループでの作業となるとメンバーの質が課題となる。その表れとして作業経験年数で半年以内の経験者は要注意ということになる。製造品目は日に、2アイテムと製造頻度は半月に1度の割で製造されている品目で、発生時間は午後、これは昼食後の満腹感による精神の集中度欠如が原因であろう。発生曜日は火曜日となっていて回帰分析の結果を裏付けるものである。

表 4.19 事故ありのルールパターン

IF (3) = 1 0 0 0 0 (4) = 0 1 0 (5) = 0 0 1 0 0 (7) = 0 1 (13) = 0 0 0 1 0 0 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 2						
IF (3) = 1 0 0 0 0 (4) = 0 1 0 (5) = 0 0 1 0 0 (9) = 0 1 (13) = 0 0 0 1 0 0 0 (14) = 0 0 1 0 0 0 0 THEN (1) = 2						
3. 作業経験 数年	4. 製造品目	5. 製造頻度	7. チェック表	9. チェック者	13. 発生時間	14. 発生曜日
1:半年以内	2:2アイテム	3:半月1	2:あり		4:午後	3:火曜日
1:半年以内	2:2アイテム	4:半月1		2:あり	4:午後	3:火曜日

表 4.20 は提案ラフ集合分析により選択された項目と、説明変数選択基準 (Ru) で選ばれた項目との対比表である。项目的には、①作業経験年数②製造品目③製造頻度④チェック表⑤発生時間⑥発生曜日の6個の項目に共通性が認められる。以上のことから回帰分析と提案ラフ集合分析の間には整合性があると判断できる。

表 4.20 提案ラフ集合分析により選択された項目

項目番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	RU	R2
項目名	作業者	作業人数	作業経験年数	製造品目	製造頻度	交替	手順書	チェック表	チェック者	作業者判断	確認手段	熟練度	発生時間	発生曜日		
RU最大値の組み合わせ		○	○	○	○	○	○	○		○	○	○	○	○	0.295209	0.620
提案ラフ集合の項目			○	○	○			○	○				○	○	0.208699	0.458



#### 4.1.5 分析手法と結果の比較評価

図 4.7 に分析手法の評価を示す。A 食品会社での検証にて 3 種類の分析手法を適用したが、結果は各々を補完できることが判明した。総合評価に記したように 2 つ以上の分析手法を組み合わせることでよりの確な分析結果が得られるものとする。

分析方法	回帰分析	提案ラフ集合	自己組織化マップ(SOM)
優位点	少ないデータでも結果が出せる、操作が簡単なので習熟度が高い。また、データ間の関連性から「ルール」を見出すことができる	確率的に高い「ルール」が抽出できるので、ポイントが明確となり、対応策が立てやすい	ビジュアルで表現できるので利用者にはわかりやすい
要検討	項目数、カテゴリ数が多くなると別途ソフトが必要となる	アイテム数、カテゴリ数が多くなると準じてデータ数も多く必要となる。 データ数 ≥ アイテム数 × カテゴリ数	カテゴリの評価(重みづけなどで正規化)付けが必要
総合評価	総花的ではあるが常識的で納得できる結果が得られる	デシジョンテーブルの適正化で有効である	回帰分析との組み合わせ、ラフ集合との組み合わせで適用することで、説得力のある分析手法となる

図 4.7 分析手法の評価

#### A 食品会社における評価と成果

A 食品会社では、分析結果より

- ① 事故原因が目で見える形になったこと
- ② デシジョンテーブルを工夫することで現場に適合した品質事故の要因分析ができること
- ③ 原因が明らかになれば対応策の策定が可能なことが理解、確認できた。そのことによりデシジョンテーブルの設定アイテム、カテゴリの見直しを現場主

導で開始された。

また、

④事故再発防止報告書の不備の改訂

⑤各工場への新たな取り組みとしての波及効果が大きいと評価している。

また、本部としての新たな取り組みとして、現場教育の重要性が顕著となり e-ラーニングを適用した教育のためのテキスト作りや品質事故の事例と対応策のデータベース化が順次実行されている。事例データベースは既に立ち上がり情報の共有化が進行している。まさに製造部門のイノベーションがはじまったのである。この間 2 年間に要し、現場部門への説明会は延べ十数回におよんだ。製造現場は絶対なる技術への自信を持ち、他人の口の差し挟む隙がない、ましてや外部からの事故についての言及は法度であったに違いない。

今回の取り組みでは具体的に事故構成要因が明確化されたことで、現場担当者の積年の知識が暗黙知から形式知へ表出化された瞬間でもあった。

## 4.2 小型飛行機事故分析

2006年10月の航空法改正で各航空会社は運行には影響ないものの、別の要因が重なった場合に事故につながりかねないトラブル情報（ヒヤリハット）も国交省に報告することが義務付けられた。2008年になってやっと特に事故を誘発する可能性のある要注意事例（重大インシデント情報）として30件を、ホームページに原因や背景を掲載した。この章では公開された要注意事例や事故データから筆者が提案している課題発見プロセスを適用して知見を得る試みを行った。その結果得た知見の評価を述べる。

今後、事故原因分析において原因の要因分析が重要となると考えられる。

テキストマイニングの対象として国交省に報告された特に事故を誘発する可能性のある要注意事例（以下重大インシデントと記す）と事故データを取り上げるがその背景について以下に述べる。

米国の航空機メーカー「ボーイング社」[68]のデータでは、1959年から2005年までの航空機の事故率は、航空機の技術的進歩や保安施設の充実で、70年ごろまでは急減している。しかし、70年以降、大きな変化はない。現状では離陸100万回当たり全損事故は1.5件程度だが、航空需要が増す中、事故率が変わらなければ事故数そのものは増えることになる。

事故原因の66%は、乗務員が原因だとされている従って、事故の発生を減少させるには、人間に関する要因を取り除かなければならない。

1つの重大な事故の裏には、事故には至らなかったものの重大なトラブルが発生したインシデントが29あり、さらにその裏には結果として起こらなかったが当事者が危険を感じたもの（ヒヤリハットと呼ぶ）が300あるといわれている[69]。

（ハインリッヒの法則と呼ばれる）（図4.8）

しかしながら、1992年から2001年の10年間で航空機の出発回数が約30%増加していることを考慮すると、現在の事故率のままでは、将来、世界中で事故が頻発に起こり、航空輸送が到底世の中に受け入れられないものになってしまうのではないかと危惧されている。

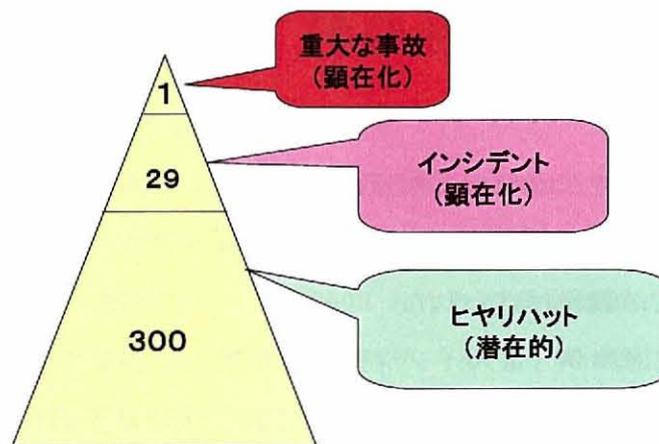


図 4.8 ハイリッヒの法則

一方，日本国内では遅まきながら航空機の計器故障や鳥の衝突，落雷など重大事故につながりかねない安全上のトラブルを分析し，事故防止に役立てようと，国土交通省は2007年7月2日，航空専門家らでつくる「航空安全情報分析委員会」を新設した。

また，事故や事故に準じる重大トラブルは同省航空・鉄道事故調査委員会が担当するが，航空安全情報分析委員会では，事故一步手前の「ヒヤリハット」も含めたトラブル情報をデータベース化するとの方針である。

表 4.22 は航空・鉄道事故調査委員会の調査した事故の件数を1995年からの発生年別事故件数として調査報告書とともに公開された情報である。表によると1995年から2009年までの15年間で発生した飛行機の大型機の事故は143件，小型機の事故は342件と小型機が圧倒的に多い。小型機は個人所有が多く運用管理を含めても，車のような手軽さではないことにもよるものと仮定できる。また，回転翼航空機であるヘリコプターの事故は403件と全体の33%である。ハイリッヒの法則に従えば航空機の大型機では143件の事故に対してインシデント（顕在化している事故数）は29倍の4,247件となりヒヤリハット（潜在的）は実に42,900件と級数的に拡大する。その意味でヒヤリハット情報を如何に収集し，分析することで如何に事故の未然化が図れるかが課題となる。一方，ヒヤリハット情報の提供は，もしかすると懲罰の域を超越しない限り困難が考えられる。つまり，報告者は会社の就業規則や懲罰規定に抵触しないかを考慮すると，自己保全のために報告をやめる場合が想定される。航空会社の規約整備も必要条件である。

表 4.22 発生年別事故件数 (2009/7/31 現在)

下記の件数は、航空・鉄道事故調査委員会の調査した事故の件数である。[101]

年	飛行機			回転翼航空機		滑空機	飛行船	計
	大型機	小型機	超軽量 動力機	ヘリコプ ター	ジャイロ ブレーン			
1995	4	7	10	6	0	1	0	28
1996	8	11	5	8	0	4	0	36
1997	3	11	3	8	2	3	0	30
1998	4	14	5	6	1	6	0	36
1999	1	9	5	7	1	5	0	28
2000	1	5	5	11	1	5	0	28
2001	2	5	2	8	0	4	0	21
2002	4	4	5	15	0	7	0	35
2003	2	10	3	1	0	2	0	18
2004	4	11	2	6	1	3	0	27
2005	1	8	0	7	0	7	0	23
2006	3	3	4	2	1	5	0	18
2007	5	3	4	7	0	4	0	23
2008	3	6	2	3	0	3	0	17
2009	3	0	1	2	0	1	0	7
合計	143	342	150	403	23	176	2	1239

#### 4.2.1 回帰分析の適用 [71,72]

この章では、小型飛行機の事故の分析結果について述べる。今回は近年発生した事故のうち165件を対象とした。航空・鉄道事故調査委員会がそのホームページで情報を公開しているのでそのデータを適用した。事故調査報告書を見ると事故原因の不明データがかなり含んでいる。例えば墜落事故で乗員全員死亡または不明では事故の原因確認が不可能となっている。また、時代とともに調査方法が上達しており、古いデータとの整合性が取りにくい点も考慮が必要であった。また、事故調査報告書は報告者により報告様式が異なる点と報告事項が長々と記述されているため解釈に苦勞する場合があった。それらを克服するのに、デシジョンテーブルは威力を発揮し、その効用が再確認できた。表4.23は小型飛行機用に作成したデシジョンテーブルの項目群である。項目数(アイテム)は11個となり、アイテムに対するカテゴリは多いもので12個にもなっている。データの収集は航空・鉄道事故調査委員会のホームページ[70]よりPDFで登録されている報告書からデシジョンテーブルに手動で変換して分析に使用した。表4.24はその回帰分析の結果である。

表 4.23 小型飛行機のデシジョンテーブル項目

発生月	1月 7月	2月 8月	3月 9月	4月 10月	5月 11月	6月 12月	
発生曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日	日曜日
発生時間	午前	午後	夜間				
天候	晴	曇	雨	雪・雹(ひょう)	もや	不明	
風	なし	静風	弱風	強風	台風接近	突風	
発生状態	離陸時	着陸時	飛行中	準備中			
発生場所	滑走路	空中	その他				
空港区分	飛行場外該当	第1種空港	第2種空港	第3種空港	その他空港		
事故内容	滑走路逸脱	胴体着陸	墜落	その他	不時着(着水)	かく座	接触
事故要因	経験不足	操作の遅れまたは操作量の不足	機材故障	燃料不足	その他	整備不良	
主たる原因	練習生	操縦者(教官)	機材故障	整備不良	その他	気象	
死傷者	なし	あり					
危害金額	なし	あり					
事故の大きさ	評価なし	小破	中破	大破	不明		

表 4.24 の回帰分析より影響度分析を行った結果が表 4.25 とその結果をグラフ化したのが図 4.9 である。このグラフからわかることは事故への影響度が高い項目(アイテム)は①風、②天候、③発生状況の順である。このアイテムの最適化をはかるために説明変数選択基準(Ru)を算出し、評価したのが表 4.26 である。その結果、図 4.9 の 11 アイテムのうち影響度の高い 9 アイテムのグループの Ru 値が最大となったために選ばれたことになる。選ばれなかったアイテムは発生時間と空港区分である。選ばれなかった項目についてのデータを精査すると、小型飛行機においては、ほとんど夜間飛行は行われぬ。午前、午後での飛行でも事故に影響を起こすほどではなかったため。また、空港区分では設備等の要因での事故を想定したが、影響度が小さかったためと考察できる。

選ばれたアイテム(項目)群は、発生月、発生曜日、天候、風、発生状態、発生場所、事故内容、事故要因、主たる原因の 9 アイテムである。

9 アイテムに絞られたメリットはデータ収集において、少ないデータ項目であることにより、発生部署への協力も取り付けやすくなる。また、データ収集に迅速に対応できることで取り組みへの一歩が約束されることになる。

経営者や管理者は最小のデータで最大の効果を望むものである。この選ばれた 9 アイテムはその希望に答えるものであり、新たなデシジョンテーブルのアイテムとして採用されることになる。

表 4.24 小型飛行機の回帰分析結果

モデル要約						
モデル	R	R2 乗	調整済み R2 乗	標準偏差推定値の誤差		
1	.738 <sup>a</sup>	.545	.341	.512		
RU	0.1256714					
分散分析 <sup>b</sup>						
モデル		平方和 (分散成分)	自由度	平均平方	F 値	有意確率
1	回帰	35.819	51	.702	2.674	.000 <sup>a</sup>
	残差 (分散分析)	29.940	114	.263		
	合計 (ピボットテーブル)	65.759	165			
係数 <sup>a</sup>						
モデル		標準化されていない係数		標準化係数	t 値	有意確率
		B	標準偏差誤差	ベータ		
1	(定数)	2.215	.976		2.269	.025
	A1	-.200	.236	-.082	-.845	.400
	A2	-.011	.246	-.004	-.043	.966
	A3	.027	.256	.011	.106	.916
	A4	-.251	.233	-.114	-1.075	.285
	A5	-.057	.214	-.028	-.267	.790
	A6	-.412	.232	-.156	-1.773	.079
	A7	-.128	.221	-.057	-.581	.562
	A8	-.245	.246	-.101	-.996	.321
	A9	-.438	.213	-.205	-2.055	.042
	A10	-.182	.193	-.105	-.946	.346
	A11	-.156	.263	-.056	-.593	.554
	B1	.109	.188	.050	.578	.564
	B2	.238	.173	.114	1.375	.172
	B3	-.020	.201	-.010	-.098	.922
	B4	.190	.158	.110	1.202	.232
	B5	.350	.156	.229	2.244	.027
	B6	.280	.177	.148	1.577	.118
	C1	.100	.278	.079	.359	.720
	C2	.160	.284	.127	.564	.574
	D1	.722	.619	.538	1.167	.246
	D2	.749	.615	.495	1.219	.225
	D3	.696	.633	.263	1.101	.273
	D4	.525	.734	.111	.715	.476
	D5	.999	.759	.211	1.316	.191
	E1	.312	.447	.076	.696	.488
	E2	-.011	.354	-.009	-.031	.975
	E3	.077	.351	.061	.219	.827
	E4	.378	.383	.149	.987	.326
	E5	-.843	.709	-.104	-1.189	.237
	F1	.203	.308	.087	.658	.512
	F2	.169	.289	.133	.586	.559
	F3	.686	.397	.543	1.726	.087
	G1	-.182	.265	-.144	-.684	.495
	G2	-.388	.323	-.306	-1.201	.232
H1	-.302	.214	-.239	-1.415	.160	
H3	-.268	.147	-.162	-1.821	.071	
H4	-.162	.169	-.084	-.958	.340	
I2	.200	.217	.110	.922	.359	
I3	-.004	.253	-.002	-.017	.986	
I4	.813	.223	.561	3.641	.000	
I5	-.054	.211	-.033	-.257	.798	
I6	.300	.233	.181	1.291	.199	
I7	.241	.239	.127	1.008	.316	
J2	.445	.273	.132	1.634	.105	
J5	.221	.672	.027	.329	.743	
J6	.130	.327	.089	.398	.691	
L1	.370	.888	.045	.416	.678	
L2	-.026	.538	-.021	-.048	.962	
L3	-.053	.536	-.033	-.098	.922	
L4	.034	.594	.013	.057	.955	
L5	-.034	.389	-.023	-.086	.932	

表 4.25 小型飛行機の影響度分析結果

項目記号	項目名	影響度	MAX	MIN
A	発生月	0.465	.027	-.438
B	発生曜日	0.370	.350	-.020
C	発生時間	0.160	.160	.000
D	天候	0.999	.999	.000
E	風	1.221	.378	-.843
F	発生状態	0.686	.686	.000
G	発生場所	0.388	.000	-.388
H	空港区分	0.302	.000	-.302
I	事故内容	0.355	.300	-.054
J	事故要因	0.445	.445	.000
L	主たる原因	0.422	.370	-.053

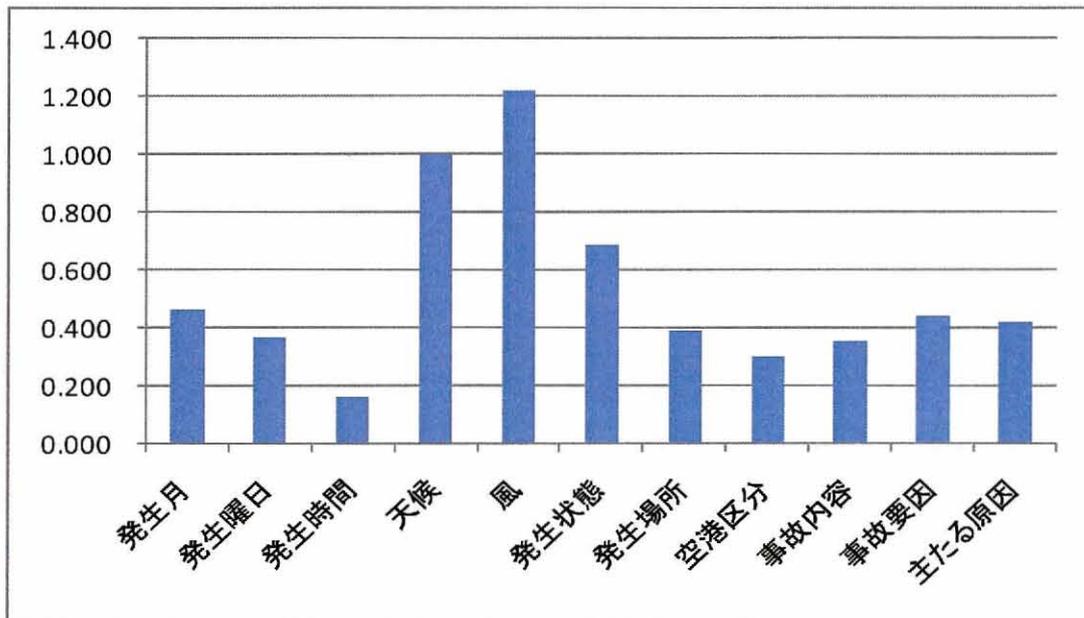


図 4.9 小型飛行機の影響度グラフ

表 4.26 小型飛行機の説明変数基準 (Ru) 評価表

	発生月	発生曜日	発生時間	天候	風	発生状態	発生場所	空港区分	事故内容	事故要因	主たる原因	RU	R2
全項目	0	0	0	0	0	0	0	0	0	0	0	0.125671	0.544704
上位9個	0	0		0	0	0	0		0	0	0	0.149077	0.526373

選ばれた 9 アイテムでの回帰分析を試みた。表 4.27 が回帰分析の結果である。

また、影響度分析の結果が表 4.28 である。この表をグラフ化したのが図 4.10 である。

表 4.27 説明変数選択基準 (Ru) で選ばれた項目での回帰分析結果

モデル要約						
モデル	R	R2 乗	調整済み R2 乗	標準偏差推定値の誤差		
1	.726 <sup>a</sup>	.526	.343	.512		
RU 0.1490767						
分散分析 <sup>b</sup>						
モデル		平方和 (分散成分)	自由度	平均平方	F 値	有意確率
1	回帰	34.614	46	.752	2.875	.000 <sup>a</sup>
	残差 (分散分析)	31.145	119	.262		
	合計 (ピボットテーブル)	65.759	165			
係数 <sup>a</sup>						
モデル		標準化されていない係数		標準化係数	t 値	有意確率
		B	標準偏差誤差	ベータ		
1	(定数)	2.388	.933		2.559	.012
	A1	-.310	.228	-.128	-1.361	.176
	A2	-.028	.245	-.010	-.112	.911
	A3	.037	.251	.015	.148	.883
	A4	-.298	.227	-.136	-1.313	.192
	A5	-.123	.210	-.061	-.588	.558
	A6	-.434	.228	-.164	-1.901	.060
	A7	-.048	.215	-.021	-.221	.826
	A8	-.356	.238	-.146	-1.495	.138
	A9	-.430	.211	-.202	-2.035	.044
	A10	-.189	.191	-.109	-.990	.324
	A11	-.227	.258	-.082	-.878	.382
	B1	.030	.182	.014	.167	.868
	B2	.180	.170	.087	1.059	.292
	B3	-.031	.186	-.015	-.166	.868
	B4	.214	.156	.124	1.371	.173
	B5	.300	.147	.196	2.040	.044
	B6	.248	.173	.131	1.430	.155
	D1	.607	.613	.452	.990	.324
	D2	.600	.608	.396	.986	.326
	D3	.564	.628	.213	.897	.371
	D4	.475	.728	.101	.653	.515
	D5	.840	.750	.178	1.121	.265
	E1	.376	.440	.092	.856	.394
	E2	.059	.344	.046	.172	.864
	E3	.169	.338	.134	.500	.618
	E4	.442	.376	.175	1.175	.242
	E5	-.899	.696	-.111	-1.292	.199
	F1	.124	.303	.053	.409	.683
	F2	.126	.284	.099	.443	.659
	F3	.602	.391	.477	1.540	.126
	G1	-.102	.242	-.081	-.419	.676
	G2	-.393	.317	-.309	-1.238	.218
	I2	.209	.216	.115	.969	.335
	I3	-.048	.250	-.019	-.191	.849
	I4	.755	.219	.522	3.444	.001
	I5	-.071	.208	-.043	-.344	.732
	I6	.244	.228	.147	1.070	.287
	I7	.225	.237	.119	.951	.344
	J2	.359	.268	.107	1.342	.182
	J5	.399	.643	.049	.621	.536
	J6	.057	.310	.038	.182	.856
	L1	.459	.872	.056	.526	.600
	L2	-.101	.513	-.080	-.197	.844
	L3	-.165	.514	-.104	-.320	.749
	L4	-.030	.568	-.012	-.054	.957
	L5	-.090	.373	-.061	-.242	.809

項目記号	項目名	影響度	MAX	MIN
A	発生月	0.471	0.037	-0.434
B	発生曜日	0.331	0.300	-0.031
D	天候	0.840	0.840	0.000
E	風	1.341	0.442	-0.899
F	発生状態	0.602	0.602	0.000
G	発生場所	0.393	0.000	-0.393
I	事故内容	0.827	0.755	-0.071
J	事故要因	0.399	0.399	0.000
L	主たる原因	0.624	0.459	-0.165

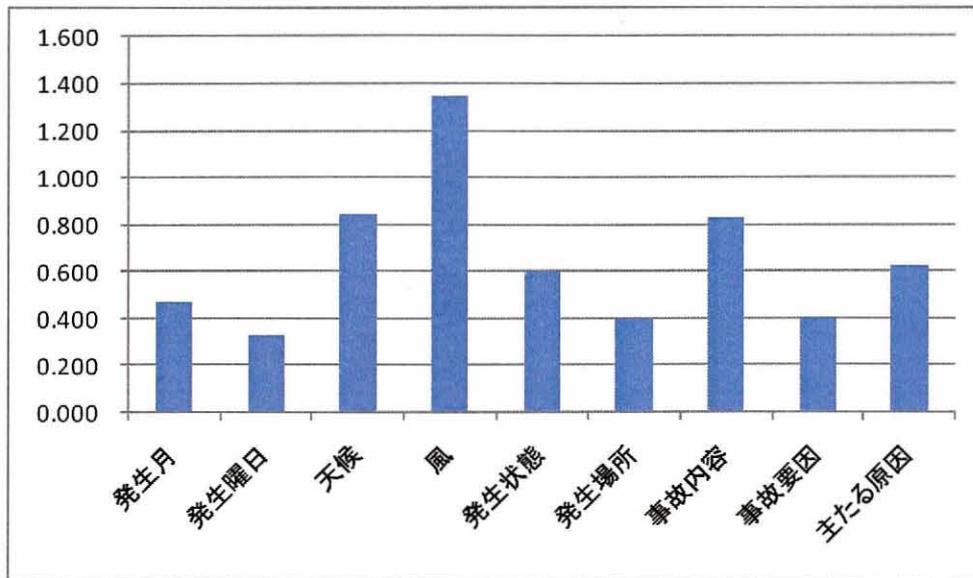


図 4.10 Ru で選ばれた項目での影響度グラフ

これらの一連の回帰分析の結果は図 4.10 で明らかとなった。小型機における事故への影響度の順位は1番が風である。2番が天候である。3番が事故内容であり、以下の順番は④主たる原因⑤発生月⑥事故要因⑦発生場所⑧発生曜日である。これらの影響度の高かった3番までのカテゴリでの影響度を調べたのが表 4.29 から 4.31 である。また、それらをグラフ化したのが図 4.11 から 4.13 である。

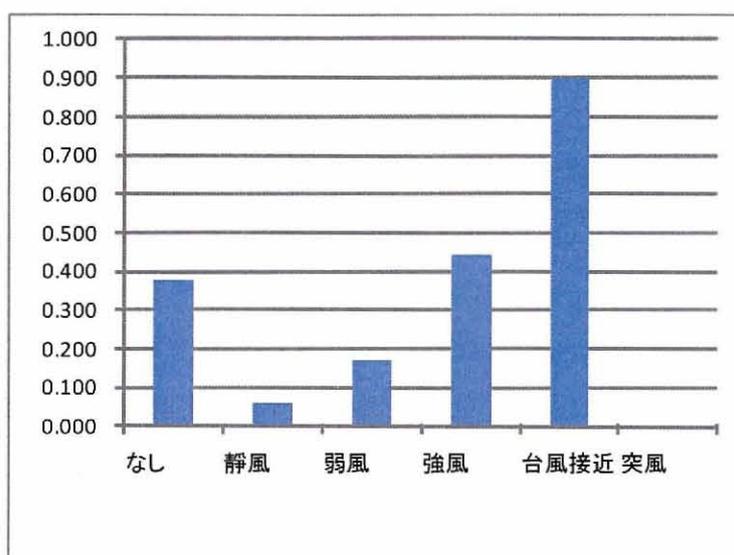
#### 4.2.2 回帰分析の結果と考察

小型飛行機の事故を分析した結果は、事故への影響度で表している。この章はじめに記したように事故原因の 66%は、乗務員が原因だとされている。しかし、小型飛行機においては自然現象からのダメージはこのほか大きいものと推察で

小型飛行機においては自然現象からのダメージはことのほか大きいものと推察できる。今回のデータの分析（図 4.10 参照）結果からは、風が 1 番の影響度となっている。風の中でも台風接近（図 4.11 参照）時および強風の影響度の高さが読み取れる。これは軽量であるために、まともに風と向き合えば、そこから脱出できるエンジンの馬力にも限界があることで事故誘発原因となる。一方、無風状態（風なし）での影響度も高い。これは無風状態での事故を示しており、他に原因があるものと思われる。先に述べた乗務員の事故へのかかわりがうかがえる。

**表 4.29 風の影響度**

風	
なし	0.376
静風	0.059
弱風	0.169
強風	0.442
台風接近	0.899
突風	0.000



**図 4.11 風の影響度グラフ**

次に影響度の高かったのが天候である。天候のカテゴリの影響度（図 4.12）を見てみるともやが 1 番となっている。もや（靄）は国語辞典（大辞林）によると空気中に小さい水滴や吸湿性の粒子などが浮遊し、遠方のものが灰色にかすんで見える状態。視程は 1 キロメートルを超え、霧よりは見通しがよいとある。つまり視界が悪く遠方がかすんで見えない状態となると、広い空中といえども事故を誘発することになる。大型機のような自動ナビゲーションシステムが搭載されている小型飛行機は少ないであろうと仮定できるからである。また、曇りや雨時にも事故への影響は高い。一方晴れにおいても影響度が高いのは、風の影響度の無

風状態と同じ原因に起因するものと推察する。当分析の被説明変数は死亡事故としているために、図 4.13 で見てみると墜落が群れを抜いて一番である。墜落事故での死亡率が高いことを意味している。不時着、かく座での影響度は墜落に比べるとその比ではなく、事故調査報告書でも死亡例は少ない。

表 4.30 天候の影響度

天候	
晴	0.607
曇	0.600
雨	0.564
雪・雹(ひょう)	0.475
もや	0.840
不明	0.000

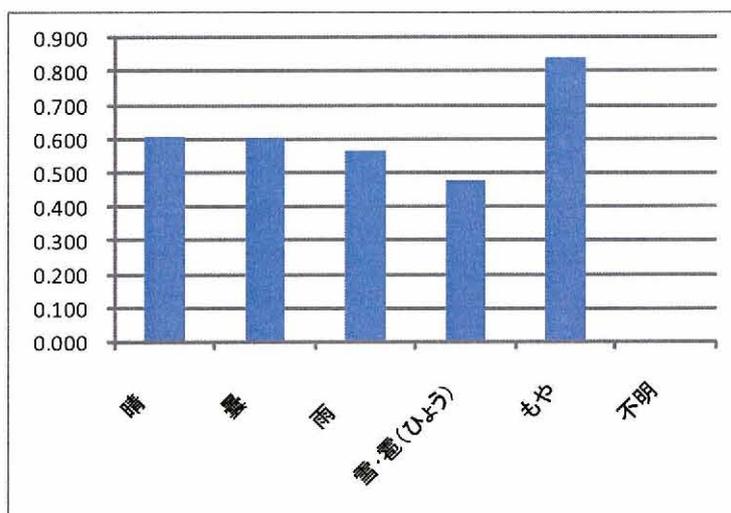


図 4.12 天候の影響度グラフ

表 4.31 事故内容の影響度

事故内容	
滑走路逸脱	0.209
胴体着陸	0.048
墜落	0.755
その他	0.071
不時着(着水)	0.244
かく座	0.225
接触	0.000

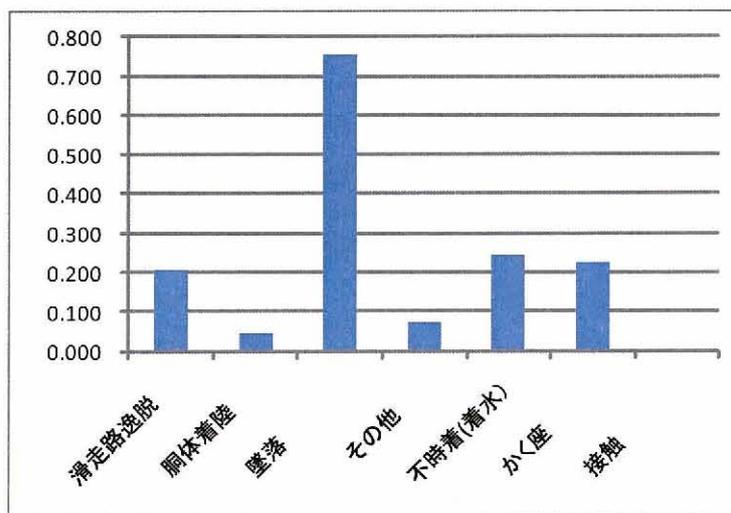


図 4.13 事故内容の影響度グラフ

以上の回帰分析結果からは事故原因の66%は、乗務員が原因の結論には至らなかった。小型飛行機は自家用機が多く、商業機と違って専門の整備士がいるわけでもなく、自由気ままに飛行を楽しむといったことでの事故が割合から見みると多いようである。また、小型飛行機の手軽さが、アクシデント（天候の急変やマシントラブルなど）に遭遇すると技術の未熟さが歴然と表出し、その対応の未熟さが事故の誘因となって、墜落などを引き起こす原因となっているのではないかと推察できる。以上のことから航空機事故は1つの要因のみが事故起因の原因ではなく、人的ミスはもちろんのこと、外的要因で特に天候や風の自然現象などの複合要因において事故に至ると考察できる。しかし、事故調査報告書にも意外と多いのが、事故原因の特定ができずに原因不明や調査中という表現である。報告内容からもうかがえるのは、調査自身が困難な作業であることに起因していることと、一概に特定が困難であるからと思われる。

### 4.2.3 提案ラフ集合分析の適用

表 4.33 は  $R_u$  で選ばれた項目群と提案するラフ集合での項目群対比表である。この表は  $R_u$  値の最大グループと提案するラフ集合分析結果項目群の最小公倍数を回帰分析し  $R_u$  を求めて記述した表でもある。この2つの項目群が同じであることが読み取れる。その意味はこの9つの項目が求める被説明変数には有効であるという証しであると判断できる。

表 4.32 は提案するラフ集合で求められたルール表である。多くは説明しないが、回帰分析結果でも明らかなように風の影響が大きく影響大は明白である。死傷者あり、なしのに関しては事故内容が大きく関連している。運よく胴体着陸の場合は死傷者なしで、運悪くまた、回避能力不足による墜落が死者を招いている。複合的要因が墜落という最悪の結果を引き起こしていることも読み取ることができる。提案するラフ集合分析では項目（アイテム）のカテゴリまで含んだ形でルールが抽出されるのでどの行為がよくって、どの行動が悪い結果を招くかを特定できることで改善へ結び付けることが可能となる。その意味では航空業界でのイノベーションの端緒となれば幸いである。

表 4.32 提案するラフ集合分析結果

	発生月	発生曜日	発生時間	天候	風	発生状態	発生場所	空港区分	事故内容	事故要因	主たる原因
死傷者なし	7月	火曜日		晴	強風				墜落		
	7月	火曜日			強風		滑走路		墜落		
	7月	火曜日			強風				胴体着陸	その他	
	7月	火曜日				飛行中			胴体着陸	その他	
	6月				強風	飛行中			胴体着陸	その他	
	7月				晴	強風			墜落		その他
死傷者あり	3月	木曜日		不明	強風				その他		
	7月	火曜日		もや	強風				墜落		
	7月	火曜日			強風		空中		墜落		
	7月	火曜日			強風				墜落	整備不良	
	7月	火曜日				飛行中			墜落	整備不良	
	7月				強風	飛行中			墜落	整備不良	
	7月				もや	強風			墜落		その他

表 4.33 Ru で選ばれた項目群と提案するラフ集合での項目群対比表

	発生月	発生曜日	発生時間	天候	風	発生状態	発生場所	空港区分	事故内容	事故要因	主たる原因	RU	R2
上位9個	0	0		0	0	0	0		0	0	0	0.149077	0.526373
Rough	0	0		0	0	0	0		0	0	0	0.149077	0.526373

## 4.3 医療事故・・・ヒヤリハット情報の分析

日本ではどのくらいの医療事故が起きているか認識しているであろうか。米国の研究による推定死亡数を基に計算すると、日本では年間 20,600 人から 46,000 人(1998)が医療過誤で命を失っている可能性があり、不慮の事故、胃・大腸・乳がんによる死亡者数より多いことがわかった。国内の八尾総合病院の事故発生率を全国の一般病床数に掛けても、年間 260 万件の事故報告書が出ていることになる。患者に重度障害あるいは死亡をもたらす事故の発生率は不明であるが、それを生み出す裾野はこれほどまでに広いのである[73]。

医療事故調査会では 1995 年以來、400 件の医師が起こした医療事故を医学的に調べ、296 件(74%)が医師の過誤と判定され、うち 188 件(64%)は患者が死亡している。原因の 90%以上が「医師の医療知識・技術の未成熟性、独善性」であった。日本では欧米などに比べ医療基盤が弱いことが医療事故の背景にもなっている。

厚生労働省では、2001 年 10 月から、ヒヤリハット事例を収集・分析し、その改善方策等医療安全に資する情報を提供する「医療安全対策ネットワーク整備事業(ヒヤリハット事例収集事業)」を開始した。事業開始当初、医薬品副作用被害救済・研究振興調査機構(現(独)医薬品医療機器総合機構)が参加医療機関からヒヤリハット事例を収集したのち厚生労働省へ報告し、厚生労働省の研究班が集計・分析を行う枠組みとなっていた。この枠組みに従って第 1 回から第 10 回までのヒヤリハット事例収集が行われ、厚生労働省より集計結果の概要を公表する等、収集したヒヤリハット事例に基づく情報提供が行われた[74]。

2004 年度からは、当機構が医薬品副作用被害救済・研究振興調査機構(現(独)医薬品医療機器総合機構)よりヒヤリハット事例の収集事業を引き継ぎ、第 11 回以降のヒヤリハット事例収集を行っている。集計結果や分析は、当機構のホームページにおいて公表している[75]。

一方、医療事故情報収集については 2002 年 4 月、厚生労働省が設置した医療安全対策検討会議が「医療安全推進総合対策」[75]を取りまとめ公表した。同報告

書は、2001年10月から既に開始された医療安全対策ネットワーク整備事業（ヒヤリハット事例収集事業）に関し、「事例分析的な内容については、今後より多くの施設から、よりの確な分析・検討結果と改善方策の分析・検討結果を収集する体制を検討する必要がある。」と述べるとともに、医療事故事例に関してもその収集・分析による活用や強制的な調査・報告の制度化を求める意見を紹介しつつ、医療事故の報告に伴う法的な問題も含めてさらに検討する必要があると述べた。

医療事故とは、医療にかかわる場所で医療の全過程において発生する人身事故一切を含むもので、廊下での転倒や医療従事者が被害者である場合も医療事故に当たる。これに対して、医療過誤は、医療事故の発生原因として医療機関・医療従事者に過失がある場合をいう。医療事故の類語としては、「アクシデント」、「インシデント」があり、アクシデントは通常、医療事故に相当する用語として用いられる。インシデントは、日常医療の場で、誤った医療行為などが患者に実施される前に発見されたもの、あるいは誤った医療行為などが実施されたが、結果として患者に影響をおよぼすにいたらなかったものをいい、「ヒヤリハット事例」がこれに当たる[75]。

報告の対象となるヒヤリハット事例[76]

- ① 誤った医療行為等が、患者に実施される前に発見された事例。
- ② 誤った医療行為等が実施されたが、結果として患者に影響をおよぼすに至らなかった事例。
- ③ 誤った医療行為等が実施され、その結果、軽微な処置・治療を要した事例などである。

厚生労働省では、平成13年10月から医療機関の協力により、ヒヤリハット事例を収集・分析し、その改善方策等医療機関に資する情報を提供する医療安全対策ネットワーク整備事業（ヒヤリハット事例収集事業）を実施している[77]。

平成16年4月からは、対象を全国医療機関に拡大するとともに、事例の収集実施機関を財団法人日本医療機能評価機構に変更した。

データベースの公開が原則となっていたことが、現在（2009年）医療関係者の登録者のみID管理のもと、データの提供をインターネットでおこなっている。しかし、インターネットでのデータベース公開は行われていない。関係部署に問い合わせた結果、出版物で公開するとの意向である。そんな事情もあって、公開デ

ータの分析への挑戦はまだ実施の段階ではない。

しかし、いずれデータが公開されることを頼みに、データ公開に先立ち、当論文で提案している IT ソリューションである事故報告システムでデシジョンテーブルを作成してみた。そのサンプルは表 4.34 である。ヒヤリハットの場合被説明変数が明確にならない場合が多かったが今回は、影響度ということで、その内容をもし、この事故が起きたらどうであったかの観点で4つの影響度を設定してみた。このデシジョンテーブルが実用性あるかは定かでないので、公開されるサンプルデータで確認したいものである。

表 4.34 事故報告システムで作成した医療向け提案するデシジョンテーブル

項目	値1	値2	値3	値4	値5	値6	値7	値8
発生月(前期)	1月	2月	3月	4月	5月	6月		
発生月(後期)	7月	8月	9月	10月	11月	12月	不明	
発生曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日	日曜日	不明
曜日区分	平日	休日(祝祭日も含む)	不明					
発生時間帯	早朝(5~7時台)	朝(8~10時台)	午前(11~12時台)	午後(13~16時台)	夕方(17~19時台)	夜(20~22時台)	深夜(23~4時台)	
発生場所	病室	処置室	診察室	その他				
患者の性別	男性	女性	患者複数	不明				
患者の年齢	0から19歳	20から39歳	40から59歳	60から79歳	80から150歳	患者複数	不明	
患者の心身状態	意識障害	精神障害	痴呆・健忘	不明	その他			
発見者	当事者本人	同職種者	他職種者	患者本人	家族・付添	他患者	不明	その他
当事者の職種	医師	看護師	准看護師	看護助手	不明	その他		
当事者の職種従事年数	半年以内	1年以内	2~4年	5~7年	8~10年	10年以上		
当事者の部署配属年数	半年以内	1年以内	2~4年	5~7年	8~10年	10年以上		
発生事象	定時見回り時	コール時	処置時	食事時	安静時			
要因1 確認	不十分であった	その他						
要因2 観察	不十分であった	その他						
要因3 判断	判断に誤りがあった	その他						
要因4 知識	不足していた	誤りがあった	その他					
要因5 技術(手技)	未熟だった	誤った	その他					
報告等	忘れた	不十分であった	間違いであった	不適切であった	その他			
身体的状況	寝不足だった	体調が不良だった	眠くなる薬を飲んでいました	その他				
心理的状況	慌てていた	イライラしていた	緊張していた	他のことに気を取られていた	思い込んでいた	無意識だった	その他	
影響度	小さい(処置不要)	中等度(処置が必要)	大きい(生命に影響しうる)	影響なし				

## 第5章 実システムの開発

本研究で述べてきた事柄を，実業界で広く適用・活用してもらうためにはソリューションとしての提供が最もふさわしいと判断した。

ここでいうソリューションとは問題解決のために I T (Information Technology) を用いたシステムであり，筆者が長年従事してきた仕事での知識が生かされる利点がある。しかし，こういうシステムで一番の難題は，ターゲットをどこにあわせるかにある。

当システムのターゲットは，人材も投資資金も弱小な中小企業や，製造部門で事故という悩みを抱えながら，忙しさのために取り組めない管理職等とした。

### 5.1 事故報告システム

当システムの重要な部分は，データの確保である。事故発生部署に報告義務を強いても，タイミングを逸したりするケースが多い。これらを解決するための IT ソリューションを提供する。

機能的には 2 系統のシステムとする。

1 系統は最も大事なシステムでデシジョンテーブル作成支援システムとする。(図 5.1 参照)

初期においては報告書のテキストデータのサンプルデータを使って形態素解析後キーワード抽出および KeyGraph 出力を行い，そのビジュアルデータを参考にしデシジョンテーブルを作成する。最初の一回目で完成品は出来ないので，初期に作成したデシジョンテーブルに実際の報告データをプロットして，使い勝手や，項目，カテゴリの整合性を確認する。デシジョンテーブルは項目のカテゴリの修正等は簡略にできるシステムとしている。また，旧のデシジョンテーブルから新たなデシジョンテーブルの作成，および項目追加なども簡易なオペレーシ

ョンで可能としている。また、デシジョンテーブル作成には専門家とのディスカッションをもって、参加者全員の納得のできるデシジョンテーブルの作成が肝心である。

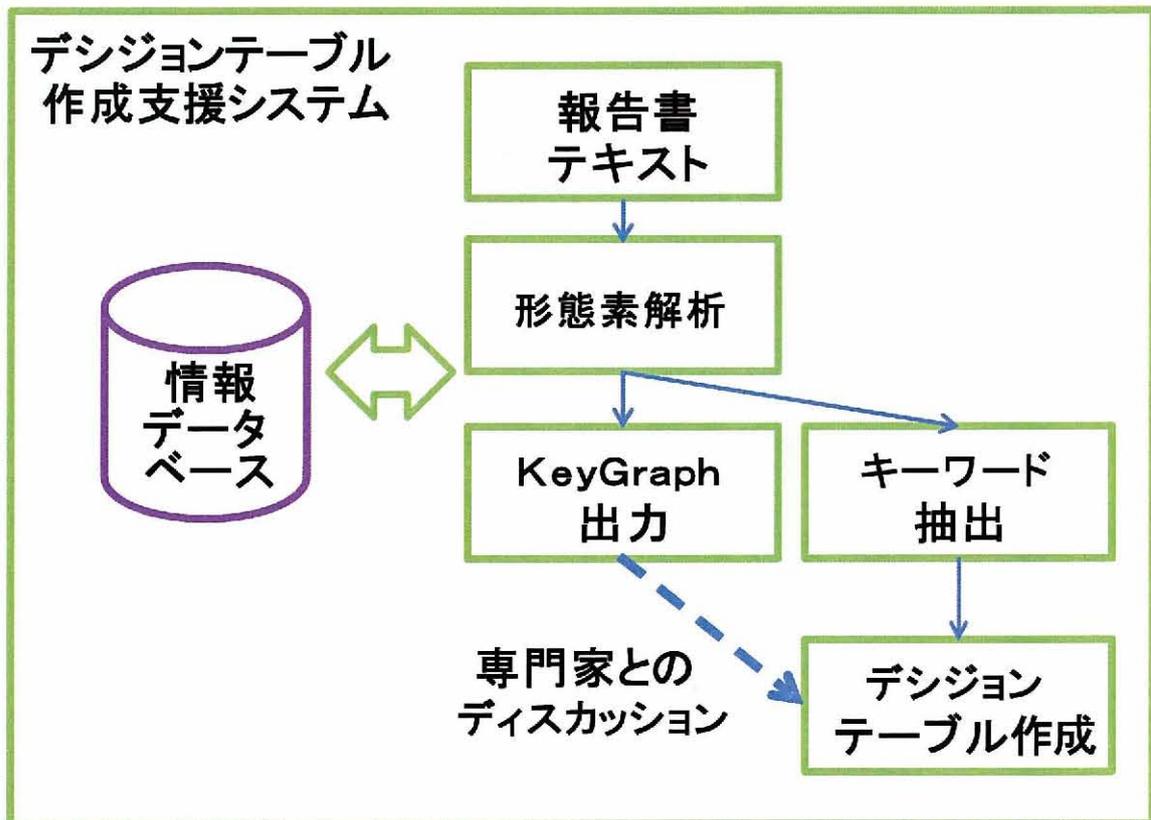


図 5.1 デシジョンテーブル作成支援システムの概要

第 2 の系統はリスク情報登録および検索システムである(図 5.2 参照)

このシステムは現場で実際に作業しているみなさんに直接入力してもらうシステムである。そのために ID を与えパスワードでデータの漏えい保護を行っている。また、入力されたデータは同じ部署なり、登録者に対してのみ公開を行い、必要なキーにて、多方面での検索を可能としている。

データ入力時に適応するデシジョンテーブルを必要に応じて選択することで、その選ばれたデシジョンテーブルに沿ったデータ入力が可能である。

極端には、一人の担当者がいくつもの部署のデータ入力も可能となる。

次の機能はデータ抽出および分析用データの加工工程である。

標準化のもっとも困難な部分であるので、初期は CSV ファイルへの書き出しを行う。書き出されたデータはまだ分析には不十分なので加工が必要となるので注意のこと。

当システムの第 2 版では可能な限り自動化をめざす。

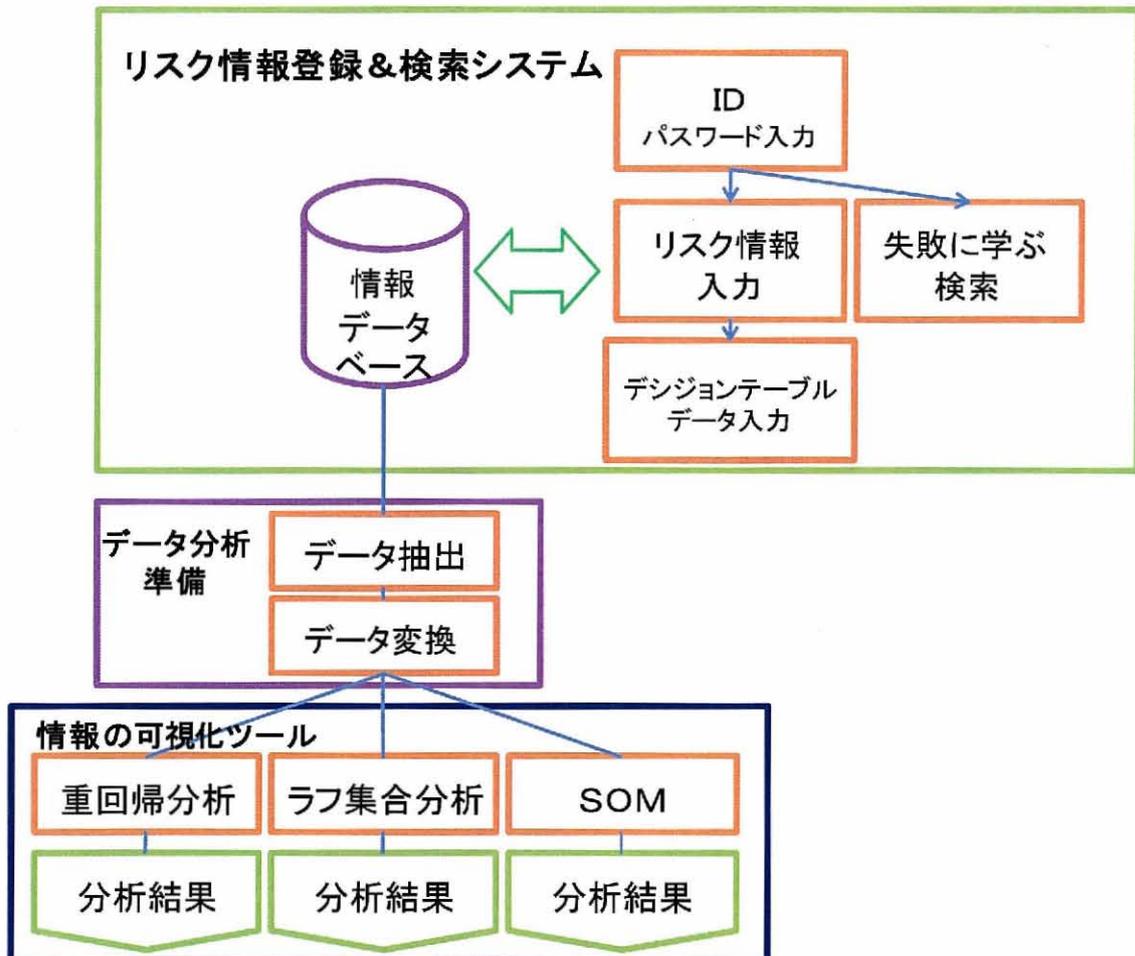


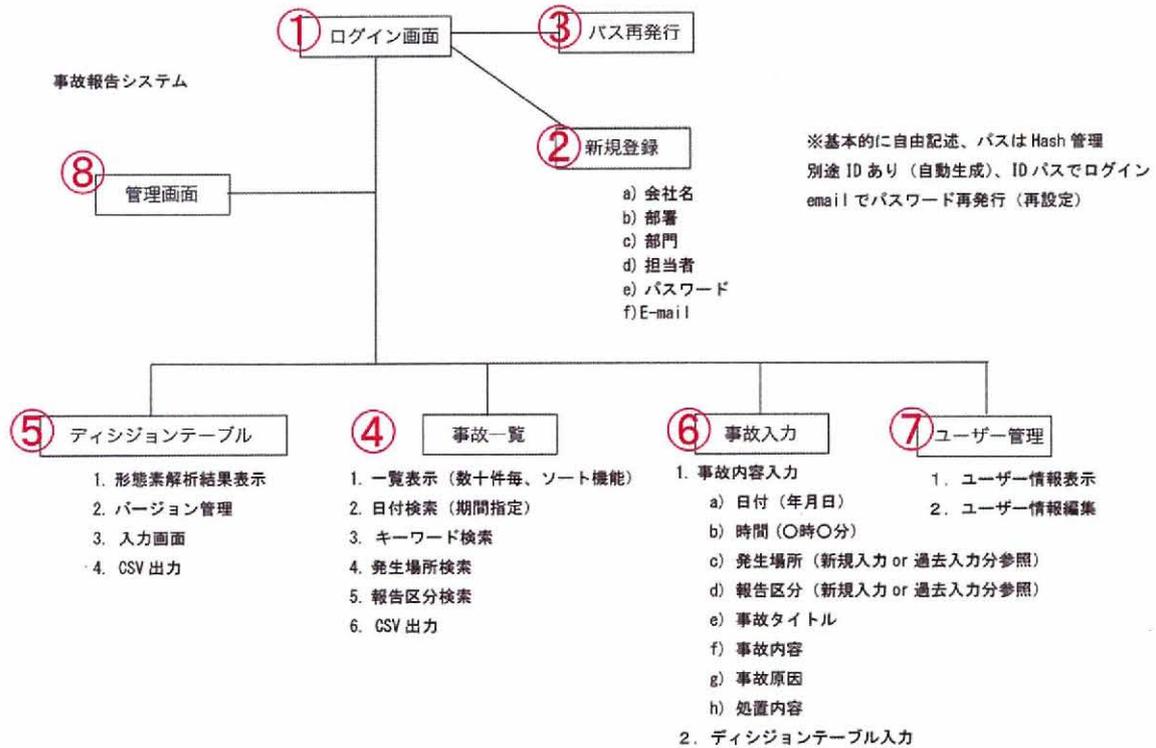
図 5.2 リスク情報登録および検索システムの概要

- ・重回帰分析は，項目数およびカテゴリの数により Excel が使用可能であるが制限がある．説明変数が 16 個である．最近アドオンソフトで説明変数 100 個まで可能なものも廉価で入手できるので検討が必要である．

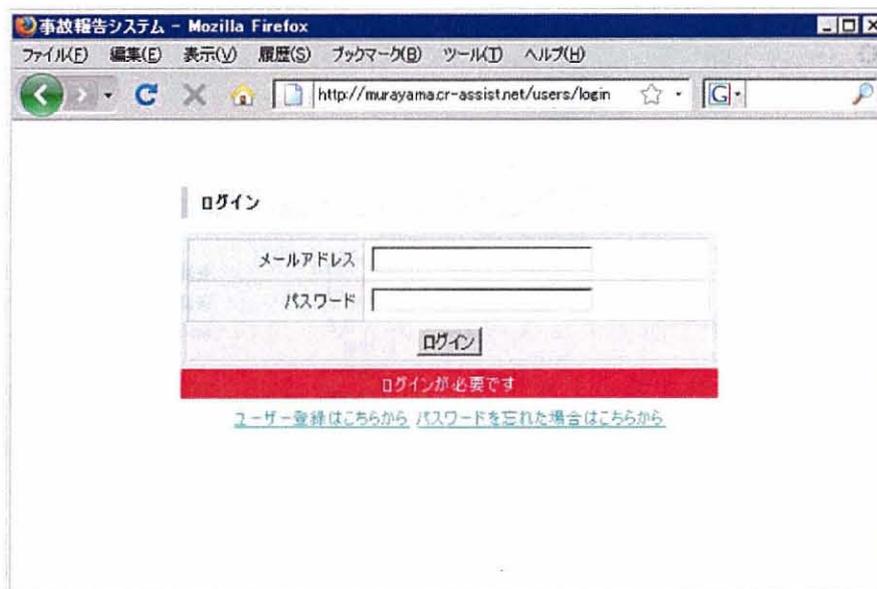
- ・ラフ集合については現在 Linux での C++言語で作成されているので Windows 版への移行を計画中である．現在はデータを預かって処理を支援する．

- ・SOM については前者とのデータ互換がないので中間加工が必要となる．また，プログラムも多岐にわたり，標準ライブラリーも公開されているので利用は可能である．

## 5.2 事故報告システムの操作マニュアル



## ①ログイン画面



ログイン

メールアドレス

パスワード

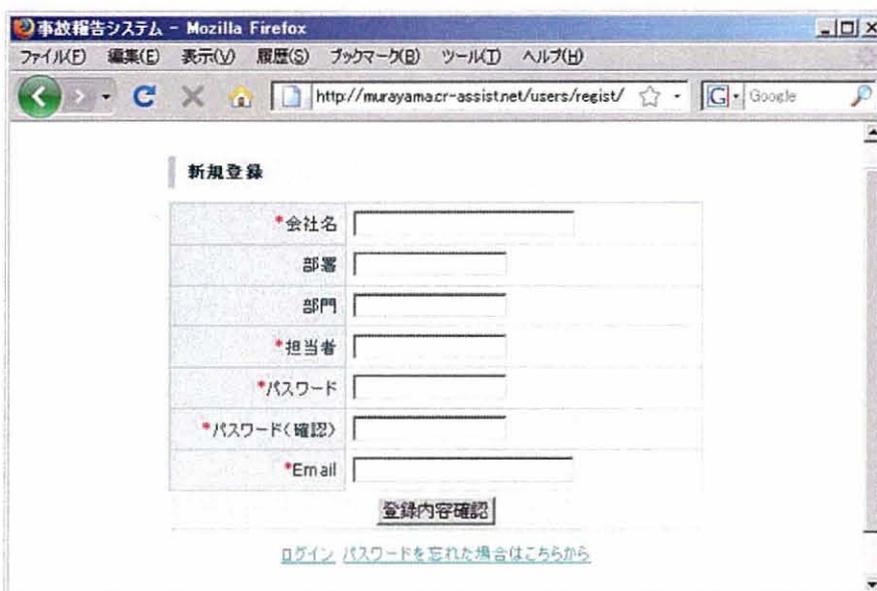
ログイン

ログインが必要です

[ユーザー登録はこちら](#) [パスワードを忘れた場合はこちらから](#)

メールアドレスとパスワードを入力してログインできます。  
最初に、ユーザーを作成してください。

## ②新規ユーザー登録



新規登録

\*会社名

部署

部門

\*担当者

\*パスワード

\*パスワード(確認)

\*Email

登録内容確認

[ログインパスワードを忘れた場合はこちらから](#)

それぞれの項目を埋めてアカウント作成します。  
メールにてアカウント情報が送信され、ログインできます。

### ③パスワード再発行



パスワードの再設定画面(ランダムパスワード生成し、設定します)  
生成されたパスワードがメールで送信されます。  
ログイン後、必要に応じてパスワードの再設定を行ってください。

### ④事故一覧画面 (ログイン直後の初期画面)



期間、発生場所、報告区分、キーワードで検索ができます  
また、ディビジョンテーブル以外の情報をCSVで出力できます。

## ⑤ディシジョンテーブル一覧(設定)



新しいディシジョンテーブルをつくれます。  
 ここで発生期間などを指定した上で形態素解析が利用できるようになります(未実装)

## ⑤ディシジョンテーブル一覧



利用するディシジョンテーブルの指定  
 ディシジョンテーブルのバージョン毎のデータcsv出力ができます。

## ⑥事故入力画面

最低ひとつ、ディシジョンテーブルを設定しないと入力できない形です。  
項目1となっているのは、ディシジョンテーブルの項目選択となります。  
発生場所・事故区分については、過去に入れた項目名を参照できます。

## ⑦ユーザー情報管理画面

ユーザー情報の表示および、編集を行います

## ⑧管理者モード

事故報告システム

ユーザー: 上垣浩二(管理者モード)

事故管理 | ディビジョンテーブル管理 | ユーザー情報管理 | ログアウト

ユーザー情報TOP | ユーザー情報編集

ユーザー一覧

1 of 1 前へ 次へ

ユーザーID	会社名	部署	部門	担当者名	メールアドレス	詳細	使用	使用変更
2	上垣会社	営業部	営業課	上垣浩二	uegaki@airwork.jp	<a href="#">ユーザー情報TOP</a>	○	<a href="#">変更</a>
3	上垣商事	制作	制作	中野	ko2uegaki@gmail.com	<a href="#">ユーザー情報TOP</a>	○	<a href="#">変更</a>
4	地域・研究アシスト事務所	部署名	部門名	四井真介	yotsui@cr-assist.net	<a href="#">ユーザー情報TOP</a>	○	<a href="#">変更</a>
5	テスト会社	テスト	テスト	テスト太郎	yotsui@s4.dion.ne.jp	<a href="#">ユーザー情報TOP</a>	○	<a href="#">変更</a>
6	有限会社 地域・研究アシスト事務所			四井真介	yotsui@sannet.ne.jp	<a href="#">ユーザー情報TOP</a>	○	<a href="#">変更</a>

管理者モードでログインして、各ユーザーのユーザー情報TOPをクリックすると各ユーザーになりきることが出来ます。

管理者(User: [system@nextairs.co.jp](mailto:system@nextairs.co.jp) Pass: qwerty)

## システム情報

- Linux(CentOS5)
- PHP5.25
- Mysql 5.1
- Apache2.2(mod\_rewrite)

## 第6章 結論

### 6.1 まとめ

本論文では、イノベーションを支援するリスク分析ソリューションの研究をテーマとして、自然言語処理技術を利用したテキストマイニング手法を提案した。また、研究の実証のための実践先としてA食品会社の事故再発防止報告書へのルール発見・課題発見プロセスを提案し実証を試みた。提案手法を確実なものにするために、近年、情報の開示が義務付けられた最もリスクの高い航空機業界と人命にかかわる医療現場での事故にチャレンジした。航空機業界ではデータの充実している小型航空機の事故分析を、医療機関ではヒヤリハットデータの分析に適用を試み成果を得ることができた。

テキストマイニングは、前処理でテキストを構造化されたデータに変換する必要がある。その処理のなかで、同じ意味のもので異なる表現がされているものを同定する必要がある。そのような処理として、企業独自の固有語登録、マイニングに不要なストップワード（削除語）などの登録を提案した。また、課題発見ではチャンス発見プロセスを適用し、デシジョンテーブル作成の知を得ることができた。各章で得られた研究成果をまとめると、次の通りである。

第1章では、本研究の背景と目的を次のように述べた。

本研究の目的は、工場の現場で発生した事故（今回は人為的事故）の報告書である日報や事故再発防止報告書のテキストデータを基に、テキストマイニングを適用して、事故報告者の主張（原因、防止策など）を的確に抽出し、その結果を基に得られた知見、課題と現実にたどった結果との比較実証を行い、課題発見のプロセスを見出すことが目的であると宣言した。

第2章では、リスク・マネジメントとテキスト・マイニングについて一般的な考え方について言及した。リスク・マネジメントの適切な実行は、企業の存続にかかわるだけでなく、経営者や従業員の安全、さらには社会全体にもおおきなダメージを回避できる可能性があること。このような意味でも、リスク・マネジメントの適切な実行は、企業の社会的責任であると言っても過言ではないことをあらためて実感できた。

また、テキストマイニングの技術的裏づけとその方法について述べ、著者の主張が明確で、大量な不定形式の文書情報に対して、有効性があると報告されてい

る「共起グラフによるキーワード抽出法キークラフ」を検討した。

第3章では本論である課題発見のプロセスについて、実証を踏まえたプロセスが規定できた。そのプロセスの根幹として、キークラフを適用し、近年ビジネス界において様々な成功を導いているチャンス発見プロセスを検討し提案した。その適用によりデシジョンテーブル作成に導き、そのなかから、解決策のルールの抽出法を提案した。デシジョンテーブルにプロットされたデータ分析手法として、従来の分析手段である回帰分析を、また、量的データに対して属性間の序数性を考慮した提案ラフ集合および自己組織化マップ(SOM)を提案した。

第4章では、連携先であるA食品会社の事故再発防止報告書をデシジョンテーブルにプロットをし、第3章で提案した分析手法に沿って分析し解析を行い、結果と考察を述べ評価を行った。また、課題発見のプロセスを他の実業界への適用として航空機業界および医療事故に試みた。

第5章では提案した課題発見のプロセスを広く実業界に提案するために、簡便に導入できるソリューションをめざした実システムとして開発を試み公開する。ターゲットは中小企業におけるの需要をめざした。その理由は事故分析のニーズはあっても、人手不足や金銭投資の優先順位外であるために導入に踏みきれない中小企業の多いことは自明であるからだ。当システムの導入による事故分析の結果を踏まえて少しでも事故が防げたら、損益にも必ずやプラスになる要因であると確信しているからでもある。

## 6.2 今後の課題

今後検討すべき課題について以下に述べる。本研究の連携先であるA食品会社での実証はこれからも継続しつつ、もっと現場に密着して磨きあげたいと考えている。

また、今後、広く産業界で利用できるためには、考え方、アルゴリズムだけでは無理があるので、ソリューションとしてのシステム(ツール化)が必要である。

### 継続課題として

- ・A食品会社での実証検分を継続し、プロセス等の完成度を高める。
- ・最適な分析手法の研究

### 短期課題として

- ・ソリューションとしてのシステムの完成度を高める。
- ・他の業種での適用推進を行うために、業種別デシジョンテーブルの標準テン

## プレートの作成と提供

候補として

### ①医療業界

医療事故は後を絶たない、命をあずかる医療機関では、医療事故は起こしてはならないことである。その意味では大変やりがいのある業界と考える。

### ②建設業界

もう古い話題になったが、世間を騒がせたのが耐震偽装問題である。

以前から建築業界では、いろんな不祥事を起こしている。また、現場では同じ事故の繰り返しが発生している。(例クレーン車の転倒事故など)

その原因を解明し、対策のために適用を推進したい。

願いは産業界への恩返しを早い時期に行いたいということである。これを、おおきな課題だと認識し、これからは、そのために努力していきたい。

# 謝 辞

本研究を進めるにあたり、多くの方からご指導・ご鞭撻を賜りました。

大変ご多用のなか、終始、懇篤なご指導を賜りました大阪大学大学院、石井 博昭教授、ならびに、適宜明快なご指摘を賜りました同大学院、奥原 浩之准教授、都田 艶子助教に深く感謝の意を表します。奥原先生には県立広島大学での修士課程から5年間もの長きにわたりお世話になりました。

論文審査では、主査の石井 博昭教授、委員の谷田 純教授、沼尾 正行教授、奥原 浩之准教授に感謝申し上げます。

また、本研究に関して有益なご助言、ご助力をいただきました神戸学院大学、柴田 淳子准教授に深く感謝申し上げます。

広島経済大学の同僚でもある木之下 恒雄教授には、海外論文投稿でのご支援ならびに叱咤激励に感謝申し上げます。

本研究の実証の場を提供していただきました、アヲハタ株式会社 常務取締役 廿日出郁夫様（当時）、また、貴重な時間と、資料・データを担当していただいた、品質管理部部長 木村修三様、ご担当の金山晋一様、新畑晶子様、安部尚美様に深く感謝申し上げます。

年の離れた弟のような、石井研の皆様にはなにかとご鞭撻いただきましたこと感謝申し上げます。

最後になりましたが、大学院進学へ快く送り出してくれた、妻貴美子に感謝するとともに、応援してくれた家族のみんなに感謝します。

本論文の一部は広島経済大学共同研究費助成（06-A）によるものです。

## 参考文献

- [1] 上田隆穂, 他 : テキストマイニングによるマーケティング調査, 講談社, (2005)
- [2] 豊田裕貴 : テキストマイニングによるドキュメントデータ分析, 情報と科学と技術, Vol. 53, No. 1, pp.22-27 (2003)
- [3] 空閑正浩 : クレームの真因に迫る, チャンス発見のための情報技術セミナー, 発表資料 pp. 8 (2005)
- [4] 高橋智弘 : リスク・マネジメント入門, 日経文庫 (1997)
- [5] 瀧澤正雄 編著 : 製造業のリスク・マネジメント, 税理経理協会 (1998)
- [6] ジョエル・バーカー : パラダイムの魔力, 日経 BP 出版センター (1995)
- [7] J. A. シュムペーター : 経済発展の理論—企業者利潤・資本・信用・利子および景気の回転に関する一研究〈上・下〉, 岩波文庫
- [8] P. F. ドラッカー : イノベーションと企業家精神 (ドラッカー名著集), ダイヤモンド社 (2007)
- [9] A・G・ラリー , ラム・チャラン : ゲームの変革者 (The Game-Changer), 日本経済新聞出版社 (2009)
- [10] 一橋大学イノベーション研究センター編 : イノベーション・マネジメント入門, 日本経済新聞社 (2001)
- [11] 木村達也 : 流通業イノベーションの発生要因, 白桃書房 (2008)
- [12] 渡辺昌彦編 : 広島ものづくり先進企業, 溪水社 (2009)
- [13] Scott D. Anthony , Mark W. Johnson, Joseph V. Sinfield , Elizabeth J. Altman : The Innovator's Guide to Growth, Harvard Business School Pr; illustrated edition (2008)

- [14] Clayton M. Christensen : The Innovator's Dilemma, Harvard Business School Pr; illustrated edition (1997)
- [15] 市村由美, 長谷川隆明, 渡辺勇, 佐藤光弘 : テキストマイニングー事例紹介, 人工知能学会誌, Vol.16, No.2, pp.192-200 (2001)
- [16] 市村由美, 鈴木優 : テキストマイニング技術と応用, 東芝レビュー, Vol.56, No.5, pp. 19-22 (2001)
- [17] 河野浩之, 川原稔 : WEB 検索におけるテキストマイニング, 人工知能学会誌, Vol. 16, No. 2, pp. 212-218 (2001)
- [18] 那須川哲哉 : コールセンターにおけるテキストマイニング, 人工知能学会誌, Vol. 16, No. 2, pp. 219-225 (2001)
- [19] 那須川哲哉, 河野浩之, 有村博之 : テキストマイニング基盤技術, 人工知能学会誌, Vol. 16, No. 2, pp. 201-211 (2001)
- [20] 渡部勇 : ビジュアルテキストマイニング, 人工知能学会誌, Vol. 16, No. 2, pp. 226-232 (2001)
- [21] Bing Liu : Web Data Mining、Springer (2007)
- [22] Soumen Chakrabarti : Mining the Web、Morgan Kaufmann (2002)
- [23] Ronen Feldman, James Sanger : The Text Mining Handbook, Cambridge University Press (2007)
- [24] 淵上 美喜,末吉 正成,高山 泰博,今村 誠 : 事例で学ぶテキストマイニング, 共立出版 (2008)
- [25] Diane J. Cook, Lawrence B. Holder : Mining Graph Data、Wiley-Interscience (2006)
- [26] 大澤幸生, ネルソ・ベンソン, 谷内田正彦 : KeyGraph 単語共起グラフの分割・統合によるキーワード抽出, 電子情報通信学会論文集誌, j82-D1, No2, pp.391-400 (1999)
- [27] 奥村学, 難波英嗣 : テキスト自動要約に関する研究動向 (巻頭言に代えて) , 自然言語処理, 「テキスト要約のための言語処理」特集号, Vol. 6, No. 6, pp. 1-26 (1999)

(1999)

- [28] 西尾章治郎, 他 : マルチメディア情報学, 第 8 卷, 情報表現, 岩波書店 (2000)
- [29] 長尾真, 他 : 言語の科学, 第 9 卷, 言語情報処理, 岩波書店 (1998)
- [30] 大澤幸生 (編者) , 他 : チャンス発見の情報技術, 東京電機大学出版 (2003)
- [31] 大澤幸生 : ビジネスチャンス発見の技術, 岩波アクティブ新書 (2004)
- [32] 大澤幸生 : チャンス発見のデータ分析, 東京電機大学出版 (2006)
- [33] 松本祐治, 他 : 言語の科学, 第 3 卷, 単語と辞書, 岩波書店 (1998)
- [34] 株式会社 言語工学研究所 ホームページ : <http://www.gengokk.co.jp/yoigo.htm>
- [35] 横田正一, 他 : マルチメディア情報学, 第 3 卷, 情報表現, 岩波書店 (2000)
- [36] 長尾真 : 岩波講座, ソフトウェア科学 15, 自然言語処理 (1996)
- [37] 松本祐治 : 形態素解析システム『茶釜』 Version2.3.3 使用説明書 (2003)
- [38] 奥村学 : 自然言語処理関連ツール, 情報処理学会誌, Vol. 41, No. 11, pp. 1203-1207 (2000)
- [39] 松本祐治 : 形態素解析システム「茶釜」 情報処理学会誌, Vol. 41, No. 11, pp. 1208-1214 (2000)
- [40] 藤井美和, 他 : 福祉・心理・看護のためのテキストマイニング入門, 中央法規 (2005)
- [41] 林俊克 : EXCEL で学ぶテキストマイニング入門, オーム社 (2002)
- [42] 砂山渡, 大澤幸生, 谷内田正彦 : KeyGraph, 発見科学とデータマイニング, pp.45-53, 共立出版 (2000)
- [43] 株式会社 構造計画研究所 : KeyGraph 手法によるキーワード抽出アプリケーション (無償版) 操作マニュアル 第 1.4 版 (2003)

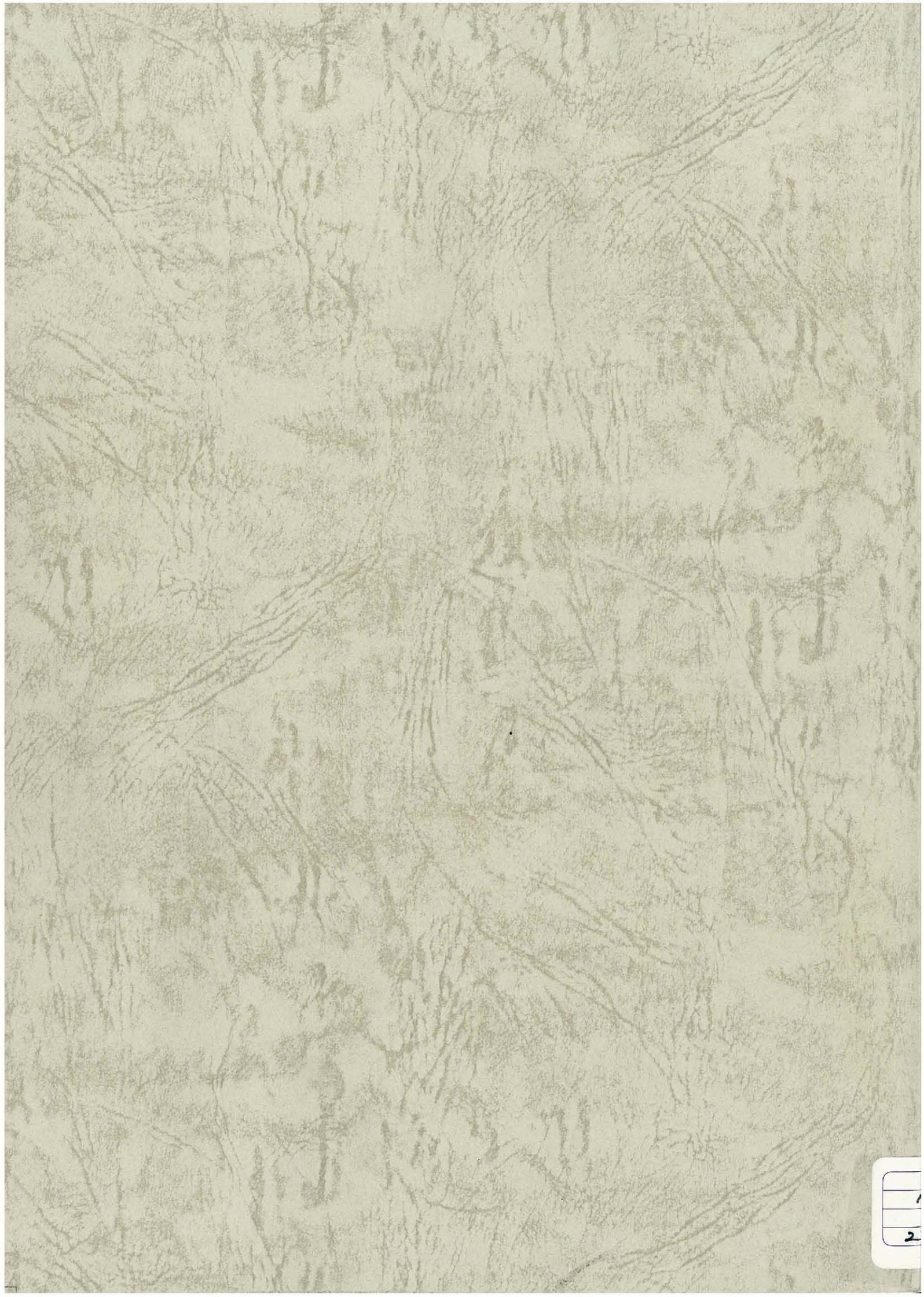
- [44] 岡崎直観, 大澤幸生, 石塚満 : チャンス発見のための統合型データマイニングツール Polaris, 第 53 回人工知能基礎論 (SIG-FAI53) 研究会, 東京大学先端科学技術研究センター (2004)
- [45] 朝日秀眞, 大澤幸生 : 360 度評価における自由回答・選択式回答の混合データからの人事評価尺度発見, 人工知能学会誌, Vol. 20, No. 3, pp.167-176(2005)
- [46] 大沢幸生, 福田寿, 高山美和 : 二重螺旋モデルを用いたスーパーの顧客行動変化の予兆発見, 知能と複雑系, 128-31, pp. 169-174 (2002)
- [47] 斉藤孝広, 渡部勇 : 障害情報からのマイニング, 情報学基礎, 61-20, 自然言語処理, 142-20 (2001)
- [48] ナレッシュ・K・マルホトラ : マーケティング・リサーチの理論と実践 技術編~, 日本マーケティング・リサーチ協会監修, 同文館, (2007)
- [49] 上田拓治 : マーケティングリサーチの論理と技法 第3版, 日本評論社, (2008)
- [50] 上田太郎 : 相関があるかを見つける簡便法, オペレーションズ・リサーチ, (1997. 7), p. 493~496
- [51] 上田 太郎 : Excel でできるデータマイニング入門, 同文館 (2001)
- [52] 上田 太郎 : データマイニング実践集, 共立出版 (1999)
- [53] 上田 太郎 : データマイニングの極意, 共立出版 (2002)
- [54] 上田 太郎, 小林 真紀, 淵上 美喜 : Excel で学ぶ回帰分析入門, オーム社 (2004)
- [55] 赤池弘次, 甘利俊一, 北川源四郎, 他 : 赤池情報量基準 AIC, 共立出版 (2007)
- [56] 和光隆光 : 回帰分析, 朝倉書店 (2002)
- [57] 奥原浩之, 松原行宏, 杉原一臣, 石井博昭 : 感性評価のための属性の序数性を考慮したラフ集合によるルール抽出, 電気情報通信学会論文誌, Vol. J87-A, No. 7, pp.1045-1053 (2004).
- [58] Koji Okuhara, Shujiro Murayama and Hiroaki Ishii, Expansive Rule Construction for Text

Mining with Large Scale Data, Proceedings of the 11th Asia Pacific Management Conference, in CD-R, pp. 3D4-1-3D4-4, Tainan, Taiwan, (November 18-20, 2005).

- [59] 森典彦, 田中英夫, 井上勝雄 : ラフ集合と感性, 海文堂出版株式会社 (2004)
- [60] マイケル J・A・ベリー, ゴードン・リノフ : データマイニング手法 (2 訂版), 海文堂書店 (2005)
- [61] T. コホネン (著), 得高平蔵, 大藪又茂, 堀尾恵一他監修 : 自己組織化マップ (改定版), シュプリンガー・ジャパン (2007)
- [62] 得高平蔵, 大北正昭, 藤村喜久郎 : 自己組織化マップとその応用, シュプリンガー・ジャパン (2007)
- [63] Shujiro Murayama ,Koji Okuhara and Hiroaki Ishii : 「Problem Discovery Framework by Text Mining from Accident Reports at Manufacturing Site」, Proceedings of the 11th Annual International Conference on Industrial engineering-Theory,Applications and Practice Nagoya,Japan(October 24-27, 2006) P.835~840 (2006)
- [64] Shujiro Murayama, Koji Okuhara and Hiroaki Ishii, Risk Management by Problem Discovery Framework from Accident Report, Proceedings of the 12th Asia Pacific Management Conference, in CD-R(ISBN: 974-8257-30-4), pp. 259-262,Bangkok, Thailand (November 17-19, 2006).
- [65] Shujiro Murayama, Koji Okuhara and Hiroaki Ishii, Innovation in Manufacturing Premise by New Finding Obtained From Accident Relapse Prevention Report, Proceedings of the 13th Asia Pacific Management Conference, in CD-R, pp. 1124-1129, Melbourne, Australia (November 18-20, 2007).
- [66] Shujiro Murayama, Junko Shibata, Koji Okuhara, Hiroaki Ishii, The Innovation of Manufacturing Premise based on Knowledge obtained by Quality Accident Report, in CD-R,ISDA2008, pp. 593-598, Kaohsiung, Taiwan (November 26-28, 2008)
- [67] Shujiro Murayama, Koji Okuhara and Hiroaki Ishii, Framework of Risk Elimination from Reports on Unsafe Acts, Mechanical and Physical Hazards, International Journal of Japan Association for Management Systems(IJAMS), 2010 掲載決定
- [68] Boeing 社の資料より  
(<http://www.boeing.com/>)

- [69] Heinrich, H.W., Peterson, D and Roos, N. : Industrial Accident Prevention(5thed.),(1980).  
(財)総合安全高額研究約 : ハインリッヒ産業災害防止論, 海文堂, (1982)
- [70] 運輸安全委員会ホームページ : 航空事故インフォメーション  
(<http://jtsb.assistmicro.co.jp/jtsb/aircraft/new/air.asp>)
- [71] Shujiro Murayama, Koji Okuhara, Hiroaki Ishii, THE MINISTRY OF LAND, INFRASTRUCTURE AND TRANSPORT OF JAPAN AND FINDING FROM SERIOUS, INCIDENT REPORT CONSIDERATION OF ACCIDENT RELAPSE PREVENTION PLAN DECISION MODEL-, The International Conference on Business & Information 2008, pp. 153-160, Honolulu, Hawaii(2008.9.2-3)
- [72] Shujiro Murayama, Koji Okuhara and Hiroaki Ishii, Knowledge Discovery for Secure Traffic Management Based on Risk Analysis in Transportation Industry, Proceedings of 5th International Conference on Planning and Design, pp. 135-139, Tainan, Taiwan (May 25-29, 2009)
- [73] 堺秀人 : 医療事故の全国的発生頻度に関する研究, 平成 15 年度~平成 17 年度総合研究報告書 (厚生労働省科学研究費補助金医療技術評価総合研究事業), (2006)
- [74] 厚生労働省ホームページ「医療安全対策について」  
(<http://www.mhlw.go.jp/topics/bukyoku/isei/i-anzen/index.html>) 参照.
- [75] 濃沼 信夫 : 医療安全用語辞典, 日本病院管理学会, エンゼビア・ジャパン(2004)
- [76] 厚生労働省ホームページ (医療安全対策のページにおける「報告書等」のページ)  
(<http://www.mhlw.go.jp/topics/bukyoku/isei/i-anzen/houkoku/index.html>) 参照.
- [77] 財団法人日本医療機能評価機構「医療事故情報収集・分析・提供事業」に関するホームページ (<http://www.jcqh.or.jp/html/accident.htm#med-safe>) 参照.
- [78] T. ベッドフォード、R. クック : 確率論的リスク解析、シュプリンガー・ジャパン (2006)
- [79] Fabrice Guillet, Howard J.Hamilton : Quality Measures in Data Mining, Springer (2007)
- [80] インターリスク総研編 : リスクマネジメント, 経済法令 (2007)
- [81] James Lam : Enterprise Risk Management, Wiley (2003)

- [82] John Wang : Encyclopedia Of Data Warehousing And Mining、 IGI Global; illustrated edition (2005)
- [83] Daniel M. Kammen, David M. Hassenzahl: Should We Risk It?, Princeton Univ Pr; Reprint (2001)
- [84] James R. Evan , David Louis Olson : Introduction to Simulation and Risk Analysis, Prentice Hall College Div; Har/Cdr (1998)
- [85] Ian H. Witten Eibe Frank : Data Mining、 Morgan Kaufmann; second edition (2005)
- [86] Tim Bedford , Roger Cooke : Probabilistic Risk Analysis, Cambridge University Press; illustrated edition (2001)
- [87] Galit Shmueli, Nitin R. Patel, Peter C. Bruce : Data Mining for Business Intelligence、 Wiley-Interscience (2006)
- [88] 橋木俊詔, 益永茂樹編 : リスク学入門 1, 岩波書店 (2007)
- [89] 益永茂樹編 : リスク学入門 5, 岩波書店 (2007)
- [90] Shujiro Murayama, Junko Shibata, Koji Okuhara, and Hiroaki Ishii, Data Mining for Hazard Elimination Through Text Information in Accident Report, Asia Pacific Management Review (APMR), 2010 掲載決定



2