

# 自然言語からのルール発見と体系化

## Rule extraction from natural language and manufacturing knowledge systematization

藤原玲子\*<sup>1</sup>  
Reiko Fujiwara

北村章\*<sup>2</sup>  
Akira Kitamura

\*<sup>1</sup> ベリングポイント株式会社  
BearingPoint Co., Ltd.

\*<sup>2</sup> 鳥取大学 工学研究科  
Tottori University

Abstract— Due to the ‘Year 2007 problem’ (2007 is the beginning year when the baby-boomers reach the age of retirement), the need to plan for skill succession is now widely recognized in Japan. [1] [2] In addition, merge and acquisition (M&A) and overseas manufacturing are accelerating a necessity of enterprise knowledge management with “Common Language“. We are studying an intelligent information platform in the realm of production, named “SKIP” (Self Knowledge & Information Expansion). This paper describes the overview of the “SKIP” and focus on natural language processing for unstructured manufacturing information. Case study is an analysis of production reports in a metal press factory..

### 1. はじめに

2007年の熟練者の大量退職の始まりに対応して、効果的に知識の共有・技術の継承を行っていくにはどのように情報の体系化を行っていくのがよいのか、また、従来では人間にしか解釈できない大量の非構造の情報を、いかに意味のある情報として技術的に処理し知識化することができるかが大きな課題となっている。この課題に対し、本研究では、「ものづくり」の領域において、日々蓄積されていく情報を扱うプラットフォームとしてSKIP (Self Knowledge & Information Expansion) を構築し、様々な要素技術を試みている。

本稿では、このSKIPの概要を示すとともに、特に、テキストや画像を含んだ非構造化情報の処理についていくつかの試みを紹介する。また、その中でも、熟練技術者の隠れた技術やスキルを、テキスト文章から抽出し、検索し、可視化する技術について考察する。

構築したシステム利用の事例としては、金属プロセス加工の工場で行った実験を取り上げた。金属プレス加工の不具合報告書を対象に、マイニングを利用したルールの抽出とその体系化、可視化を試みた事例を述べる。

### 2. インテリジェント製造プラットフォーム－SKIP

#### 2.1 SKIPのコンセプト

SKIPは、文字どおり「自ら拡張する」ことをコンセプトとしている。そしてそれは単に情報の保管庫というだけでなく、知識を獲得し、体系化し、分析するプラットフォームである。同時に情報の当事者と異なる視点で関連付けられた情報が、人間に新たな視点をもたらす、単純なインストラクションではない深い理解を生み、創造の支援となることを想定している。

日々収集し、蓄積した情報から抽出した新たな知識を、体系化する。そして可視化し、検索し活用/分析し、さらにこれを新たな創造や改善に活かす。こうしてみると、SKIPの目指すところは、エキスパートシステムではなく、“カイゼン”を支援するプラットフォーム、もしくは新たな気づきを生む“セレンディピティマシン”を目指している。[4]

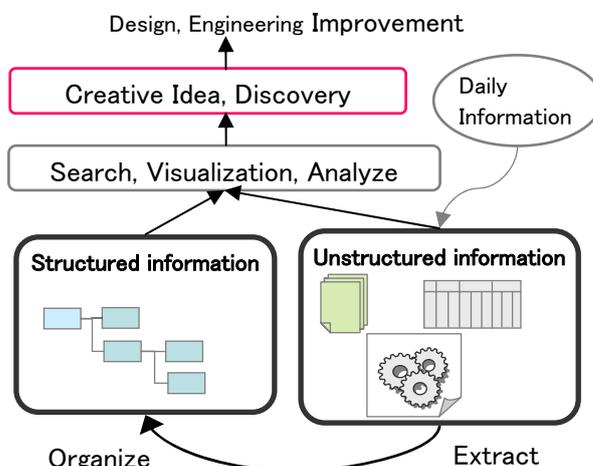


図1. SKIPのコンセプト

Fig.1: Concept of SKIP

現状のSKIPプラットフォームでは、利用可能な要素技術を統合して利用している

#### 2.2 SKIPの構成

工場内に存在している情報を調査してみると、膨大な量のテキストドキュメントや蓄積されたデータが、様々なところに分散していることに気づく。共有すべきものは、個別のファイル形式でサーバーのファイルシステムに保管されているか、業務アプリケーションシステムで収集されたものであれば、特定のデータベースに保管されているものの、全体としては、これらを分類する共通なフレームワークを持たない。そこで、われわれは、まず、これらの情報を“体系化情報”と“非構造化情報”の2つのカテゴリーに分けて考えることにした。体系化されている方を、「現場ナビゲーション」、非構造化のほうを、「ミッシュマッシュ」と名づけた。

これらのコンテンツを横断的にみつける方策としては、現在、Google Enterprise Search[10]を利用している。これにより、アクセス権やセキュリティを担保した上で、キーワードによる検索により、特定のデータベースに保存されていなくても、幅広い形式の情報にアクセスすることができる。

図2に、SKIPのトップページを示す

連絡先: 藤原玲子, ベリングポイント株式会社, 東京都千代田区丸の内 1-11-1 パシフィックセンチュリープレイス丸の内23F, reiko.fujiwara@bearingpoint.com





以上の実験の結果から、自然文による製造情報の、テキストマイニングによる解析は、定型データの定量的な分析では読み取れない関連性を見つけるのに有効であることが分かった。すなわち、コンピュータによる高度情報処理技術を利用することで、人間では不可能な大量の自然文を処理することができる。ただし、マイニングはそれ自体で、自動的に知識を発見するものではないので、必ず熟練者の知識と経験を入れ込み、これを使ってマイニングの専門家(ナレッジエンジニア)が処理することが必要である。

## 5. 知識の探索

### 5.1 知識と知識の関連

さらにもうひとつ、断片的な情報を関連付け、新たな知識を見つたり、気づきを促したりするひとつの方法を提案したい。

図 8 は、ことばネットワークから Topic Map[7]へプログラムを通して変換した例である。ひとつの不具合現象を中心とした複数の関連する知識を表している。

ことばネットワークでは表現されていない、トピックとトピックの信頼度(関係の強さ)を表しているほか、関連するトピックへは外部の情報へリンクが可能である。

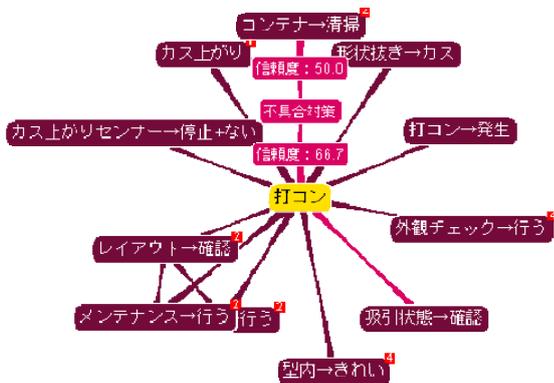


図 8. ことばネットワークから変換されたトピックマップ  
Fig. 8: Transformed Topic Map

### 5.2 ルート探索

Topic Map では、利用者が、中心となるトピック(始点)から、関連するトピックを辿って関連する知識(終点)に辿り着く。例えば表出した不具合現象の対策を探すような利用の仕方である。

これに対して、このトピック間の信頼度を利用して、一番可能性の高いルートを示すようにした。利用者が気付かなかった新たな気づきを利用者に与える事ができる可能性がある。

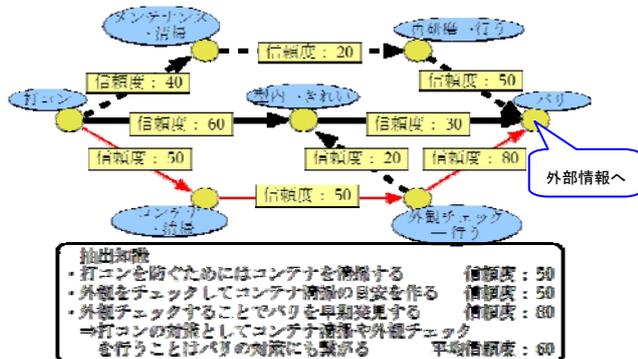


図9. トピックマップによるルート探索  
Fig. 9: Route Navigation by Topic Map

## 6. 考察

本論稿では、知的製造のプラットフォームとして SKIP とその概要を紹介した。SKIP は、セマンティック Web のコンセプトに基づいた、技術の組み合わせを試行している。知識ベースであるだけでなく、利用者とのインタラクティブなやり取りを通して、「自ら拡張する」プラットフォームを目指している。

以下に、本論稿の考察をいくつか挙げておく。

- ・企業内ビジネスプロセスマネジメントのフレームワークは、製造情報の知識共有と非構造化情報管理にも有用である。
- ・製造情報の自然文解析は、“やわらかい情報”及び暗黙知を引き出す手がかりとして有用である。
- ・熟練技術者の語彙による辞書は、熟練者の視点からその評価軸を表したもので、自然文から知識を抽出するのに必要である。

今後の研究課題としては、こうした熟練者のみを知るようなルールを、解析のパラメータとして実装し、特定の事象を複数の視点から評価することである。

SKIP の非構造化情報から、抽出した解析結果を形式知化する中で、「複数視点」の情報を関連付けるような体系化は、知識の組織内共有の目的だけでなく、人間に気づきを与え新たな創造を支援するのに有用であるかもしれない。複数視点による気づきのネットワーク化とその実証実験に取り組んでいきたい

## 参考文献

- [1] S. Sano, ““Year 2007 problem” and manufacturing industry. - imagine the short/medium/long term “Retirement-bubble” countermeasures - CIAJ J Comun Inf Netw Assoc Japan, 2005.
- [2] 小橋康章訳, 「創造的認知」, 森北出版株式会社, 1999年.
- [3] 田浦 俊春, 「技術知の位相—プロセス知の視点から (新工学知)」, 東京大学出版会 (1997/04)
- [4] David Green, “The Serendipity Machine: A Voyage of Discovery Through the Unexpected World of Computers,” Allen & Unwin (May 1, 2005)
- [5] Scheer, A.-W. : “Architecture of Integrated Information Systems”, Berlin, Springer, 1992.
- [6] H. Ando : “Applying Text Mining Techniques to Machinery Failure Report Analysis”, Journal of the Society of Naval Architects of Japan, No.192(20021200) pp. 475-483
- [7] 内藤 求: 「トピックマップ入門」, 東京電機大学出版局 (2006/12)
- [8] J. Kikuchi, “Introduction to Text Mining Tool: Text Mining Studio”, Bulletin of the Computational Statistics of Japan, Vol.18, No.1(20060630) pp. 45-49
- [9] Google Enterprise Solutions, Mini Search Appliance, Available, <http://www.google.com/enterprise/mini/>