

# 〈研究発表〉

## テキストマイニング手法による技術継承支援システムの検討

上野 健郎<sup>1)</sup>, 秋山 浩秀<sup>2)</sup>

メタウォーター株式会社（〒105-6029東京都港区虎ノ門4-3-1, E-mail: ueno-takeo@metawater.co.jp）<sup>1)</sup>

メタウォーター株式会社（〒105-6029東京都港区虎ノ門4-3-1, E-mail: akiyama-hirohide@metawater.co.jp）<sup>2)</sup>

### 概要

運転維持管理現場のノウハウは、OJTにより個人から個人へと暗黙知で継承され、組織的な表出化・蓄積・スキルアップとはなりにくい。そこで、本研究では直引継ぎ記録文書にテキストマイニング手法を適用してノウハウを抽出・可視化する実験を行った。その結果、暗黙的ノウハウを抽出可能であることが確認できた。本手法を、新人従事者などがブログ的に単純記録・蓄積した文章群に適用することで、熟練技術者の暗黙知や未知の新規ノウハウを抽出する際の支援が可能であり、技術継承に資する可能性が高いものであることを確認した。

キーワード: 技術継承、運転支援、テキストマイニング、ノウハウ収集、キーングラフ

### 1. はじめに

水道施設における運転維持管理は、「人」が行う作業の積み重ねで成り立っている。一般的に、現場業務ノウハウは先輩から後輩へOJTにより受け継がれていくものである。この場合、ノウハウは個人に蓄積され、組織的なスキルアップや蓄積がされにくい状況にある。野中郁次郎らは知識創造におけるSECIモデル<sup>1)</sup>で、「共同化 (Socialization)」→「表出化 (Externalization)」→「連結化 (Combination)」→「内面化 (Internalization)」のサイクルをまわすことにより、組織としての学習が進み、ノウハウは組織のものとなり知識創造へと発展すると述べている。この考えを水道の運転維持管理におけるノウハウ取得・知識創造に適用したものが図1であり、組織的なナレッジ定着化を図ると同時に、そこから業務支援ツールの開発が可能となると考えられる。<sup>2)</sup>そこで筆者らは、この考えに基づき、受託していた某市水道施設の運転維持管理現場においてノウハウ収集・蓄積実験を実施し考察を行ったものを報告している。<sup>3)</sup>この知見に基づき、今回、実用化に向けた検証試験を別の受託現場を対象として実施した。本研

究発表では、ノウハウ抽出の概念を説明し、予備実験を概観したうえで、実用化検証の結果と実用化システムの検討結果を報告する。

### 2. ノウハウ抽出の概念

ノウハウ抽出には、大澤幸生の提唱するKeyGraph<sup>4)</sup>の概念を適用した。以下にその概要を示しながら、水道現場への適用についての考察を加える。

#### 2.1 ノウハウの構造

ノウハウは、コンテキストとコンテンツからなる。コンテキストはシナリオを構成する。シナリオは、状態と行動からなり、それらの連続的なつながりがシナリオとなる。その様子を図2に示す。ノウハウ集を作成する場合、コンテキストを明示しなければそのコンテンツは意味をなさないため、コンテキストの解説にかなりの部分が裂かれる。特に読み手として新人を対象に含める場合は、コンテキストがないとそのノウハウを誤解される恐れがある。

一方、熟練者などノウハウを公開する側では、コンテキストは暗黙知であり、それをあらためて記述するにはかなりの労力を要する。

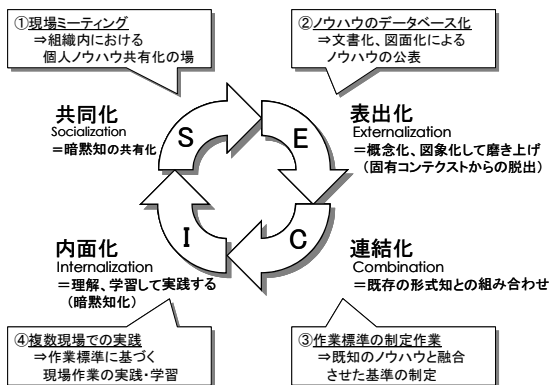


Fig.1: SECI model application for the waterworks

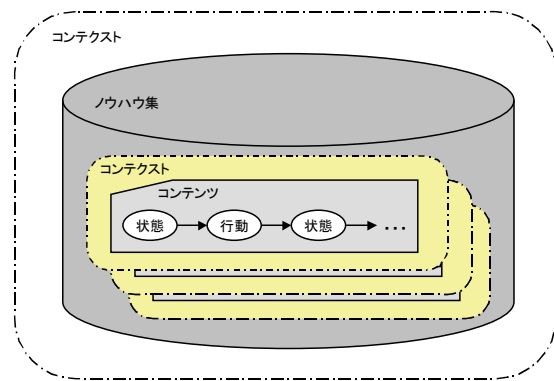


Fig.2: Structure of know how

## 2.2 テキストマイニング処理の概念

そこで、自由記述文書を形態素解析し、そこに出現した単語の共起度によりシナリオマップ表示を行う、KeyGraphの概念をノウハウの取得に適用した。

形態素解析とは、文書を日本語の意味を持つ最小の単位(品詞)に分解することを言う。形態素解析された文書は、単語の集合体とも言えるものである。

分析の元になる文章は、話題の要素単位(1つの情報のまとめ、トピック)を1単位(1レコード)として記録される。例えば、「本日、○○設備の△号□□設備で××故障が発生した。原因は、◇◇弁の誤動作であった。◇◇弁を清掃することで正常に復帰した。」のように、□□設備の××故障についての記事で1レコードを形成することになる。

このような文書を形態素解析したものを1単位とし、これらをまとめた文書集合にたいして、全ての単語について共起頻度を計算する。出現頻度、共起頻度、複数単語間での共起頻度などにより、シナリオ抽出に適した一定のアルゴリズムでグラフ作成(シナリオマップ作成)が行われる。(アルゴリズムの詳細は参考文献4)を参照されたい)

図3にシナリオマップの出力例を示す。

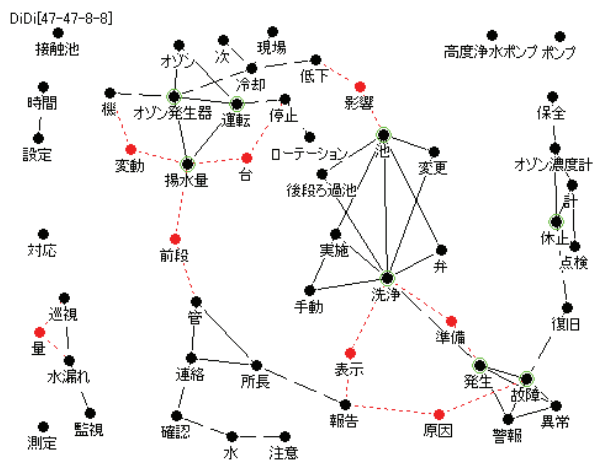


Fig.3: Example of the Scenario-map

## 2.3 シナリオマップによるノウハウ抽出手順

ノウハウ抽出は、シナリオマップの作成と、階層の読み出し、コンテキストの推定、上位シナリオの作成などからなる。シナリオマップはあくまでもヒントであり、そのヒントを元に議論することで、ノウハウが徐々に可視化され、実体が形成される。

シナリオマップからノウハウを抽出する過程は、状態と行動の想定であり、シナリオ仮説を立てて検証することになる。その際、記述されていないコンテキストを補いながら検証する。つまり、現場の事情をある程度知った上で、想像力豊かに仮説を立てていくことが必要となる。

## 3. ノウハウ収集のための予備実験

### 3.1 運転操作記録によるノウハウ抽出

水道施設において運転管理状況を観察していると、「特定の条件で同じ操作を行う」、「工事や異常時対応などで同じ操作を行う」というようなことがしばしば見られる。これは特に熟練した運転操作者によくみられ、プラントの状態を瞬時に判断し、特に意識することがなくても、プラントの状態を目標に近づけようとする暗黙のノウハウがあると考えられる。これら熟練操作者に暗黙的に蓄積されている運転管理ノウハウをどのように効率良く蓄積し、未熟練者に技術継承することができるのかという課題に対し、筆者らはプラント運転操作履歴とその根拠を記録蓄積し、その分析を試みた。<sup>3)</sup>

プラント運転操作履歴は通常業務においても日誌に記録をとっている。実験ではさらにその操作の根拠や理由の記録を実施した。

その結果、以下のような特徴が判明した。

- ① 運転操作はルーチン的であり、特別なノウハウを必要とする特殊操作は少ない
- ② 早期に暗黙知化されやすい
- ③ ノウハウとなるような操作は複数の操作が絡みやすく、単純ではない
- ④ 背景となる情報(操作の発端)も一義的ではなく階層化されていることが多い

以上により、運転操作記録の解析は、規定外操作の整理と規定化には役立つものの、新規ノウハウ発見となるような要素は出にくいとの結論に達した。

### 3.2 引継ぎ記録によるノウハウ抽出

次に、日勤・夜勤間の業務引継ぎの記録を電子化して蓄積し解析を試みた。引継ぎの記録は、特に形式や文体を指定せず、フリーフォーマットのテキスト形式で入力した。入力した文章は、KeyGraphを使用して解析を行った。

ノウハウの抽出は、マップ表示結果を見て関係者が議論し、元のテキストに戻った上で表出化させるという手順をとった。

マップ表示による議論には、水道の運転管理をまったく知らないソフトウェア設計技術者を試験的に参画させ、何らかのノウハウを得られるかどうかの実験を行った。その結果、このソフトウェア技術者はマップ表示を基に探索を行い水道施設における残塩管理のノウハウを抽出した。これは、水道技術者にとってはごく常識的なものであったが、この技術者にとってはまったく新規のノウハウであった。

このことから、引継ぎ記録のような何らかのまとめを持ったテキストデータを用い、KeyGraph概念による構文解析したシナリオマップ表示を手がかりにすれば、そ

の中に含まれる探索者にとって未知のノウハウを抽出できる可能性があることが示された。

## 4. 実用化に向けた検証実験

### 4.1 引継ぎ記録の仕組み

今回、引継ぎ記録の解析を適用した現場は、約20名の従事者がいる都市部の浄水場である。引継ぎの記録は、引継ぎ打合せで話された事項から、管理者が重要であると認識し、当日非番の者にも伝達する必要がある項目を指示する事で作成した。それ以外にも、各班責任者クラスには、独自の判断で情報を記入するようにした。

記載ルールは特に設けず、自由記述としたが、記入に当たっては、必ず施設区分と装置名の入力を求めた。記入には、現場イントラWebの仕組みを使い、現場LAN内のいずれのパソコンからも常時入力・参照可能とした。図4に入力シートの例を示す。

その結果、1年間で411件の書込みがあり、文字数は最大500文字程度、最小40文字程度で、平均文字数は140文字であった。

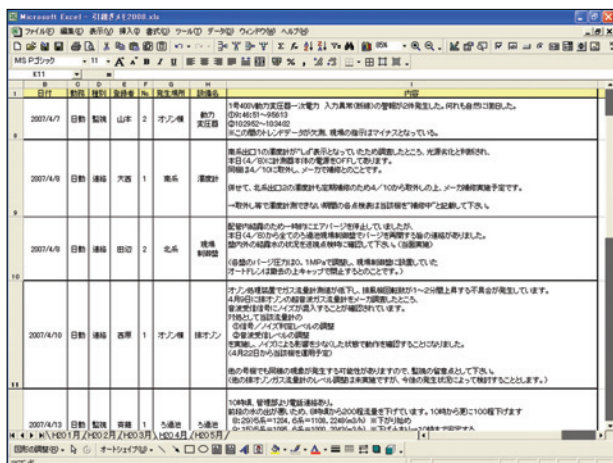


Fig.4: Example of the input sheet

### 4.2 ノウハウ抽出の手順

ノウハウ抽出は、3段階で実施した。

#### ①カテゴリー分類(フィルタリング)

- ・施設単位、設備単位、系列単位などで引継ぎ記録記載時にカテゴリー情報を付加する
- ・ノウハウ取得対象単位を区切り、何について書かれたブログかを明らかにする。(選択式にする)
- ・文章記述は、主語と動詞を明確にする。(最低限のルール)

#### ②KeyGraph解析(キーワードの可視化と関連付け)

- ・設備でフィルタリングし、分析単位とする
- ・形態素解析は、原則として名詞と動詞を対象にする。
- ・KeyGraphでは、パラメータ調整を細かく実施して、図を何種類か用意する

- ・特定のノウハウ関連単語に注目する(原因、理由、ため、から、注意、留意、故障、異常 etc.)

#### ③ナレッジ抽出検討会での議論

- ・可視化されたシナリオマップ図を基に、シナリオ(連続した状態と動作や出来事のつながり)を見つける
- ・シナリオ展開を検討し合い、コンテキストを踏まえたノウハウ仮説を立てる
- ・ノウハウシナリオにまとめる
- ・シナリオとは、「Aというときに、BがCしてDとなる。この理由はEであり、Fという方法でGをすれば良い」というふうに5W1Hを明確にしてまとめる。

## 4.3 ノウハウ抽出検証実験の結果

検証試験により次のような知見が得られ、本手法によりノウハウを抽出することが可能であることが確認できた。

- ①KeyGraphでは、原則として1回に1個のシナリオ(またはシナリオ群)が抽出される。
- ②そのため、元となる引継ぎ記録のトピックは、施設・設備区分をあらかじめ決めた上で、どの分類について記述しているのかを明確にする。
- ③その上で、シナリオマップ表示する際には、前処理として施設・設備区分によりフィルタリングをおこなって、特定設備について分析を行う。
- ④記録は、1レコードは1トピックについてのみとし、複数のトピックは記載しない。
- ⑤1トピックについて複数の人が自由に複数レコードを作成したほうが、ノウハウは得やすい。
- ⑥主観的な記述があるほうが、ノウハウが得やすい。

例えば、前出の「○○設備の△号□□設備◇◇弁動作」の例では、「この弁は、構造がやや特殊であるため詰まりや引掛掛かりが発生しやすいようである。過去にも誤動作が発生していたため、点検周期を短くしていたが、日常的にもっとチェックする必要があるかもしれないと思う。」というような、記述者の考えや情報を記入することで、コンテキストの部分や、ノウハウの核心に触れる部分が得られる。記述者には無意識に書き込んでもらい、自慢話も記入してもらうことがノウハウ獲得につながると思われる。

## 5. 実用化システムの検討

### 5.1 実用化するために必要な条件

引継ぎ記録は、最低限のルールのみを設定し、ブログ形式で記述する。最低限のルールは次の通りである。

- ①5W1H形式で記述する。
- ②主語述語を明確にする。
- ③事実と共に、できるだけ記述者の主観(何故そうするのか等の考え)も記載してもらう。

④記述の対象が何であるかを施設・設備、機器を統一名称で記録する。

ブログの記入は、最低でも1人は実施するようにするが、担当にならなくても自身の覚書として多数でのブログ利用を推奨する。

## 5.2 システム構成

図5にシステム構成を示す。

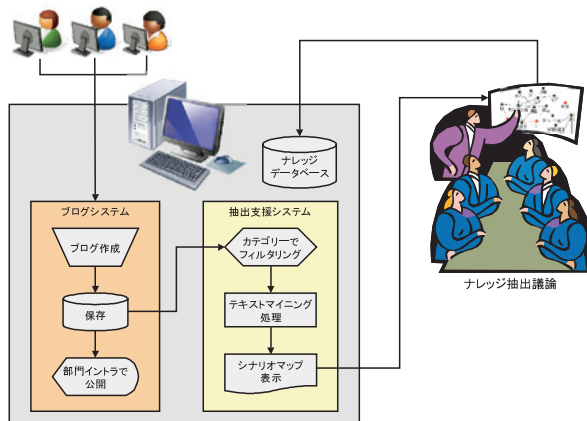


Fig.5: System configuration

## 6. まとめ

熟練者からなる成熟した組織体においては、暗黙知を前提とした会話がなされるため、そこからノウハウ表出化をすることは難しいと考えられていた。

しかしながら、新規受託現場の立上げ(官から民への引継ぎ)などにおいては、受け取った情報を単純にブログに記録していき、ノウハウ抽出支援ツールを用いることで、後日、ノウハウを可視化できることがわかった。こ

の手法は、成熟した組織においても、新人を投入しOJTで学んだことを何でも記録・電子化させることで適用可能である。さらに、シナリオマップ表示アルゴリズムをチューニングすることにより、気づいていない更なる未知のノウハウ、特に複合的な操作や、系列の関連など、わかりにくいものを抽出できる可能性も出てくるものと期待される。

今後は、監視システムの運転操作記録、故障警報記録との連動や、さらにはプロセス計測値との連動により詳細なノウハウ取得システムを目指したい。

最後に、本実験にご協力いただいた現場従事者諸氏に感謝申し上げます。

## [参考文献]

- 1) 野中郁次郎, 竹内弘高: 知識創造企業, 東洋経済新報社, 1996
- 2) 上野健郎, 桐野秀明, 秋山浩秀: 「GENESEED service」水管理サービスの基幹技術, 富士時報 Vol.79, No.6, P.486-489, 2006
- 3) 上野健郎, 桐野秀明, 牧内淑実: 技術継承にむけた運転管理ノウハウ可視化支援ツールの検討, 第59回全国水道研究発表会講演集, P.386-387, 2008
- 4) 大澤幸生, チャンス発見のデータ分析, 東京電機大学出版局, 2006