

自然言語処理技術による情報マネジメントの実際

企業における情報共有の促進事例

企業を取り巻く環境の変化

インターネット時代になり、社内外から膨大な情報を入手できるようになっている現在、日々蓄積される情報の中から、ナレッジを抽出し、共有・活用することが企業にとって重要な課題といえる。こうした状況で、ジャストシステムはナレッジ・マネジメントを実現する「ConceptBase Search（以下、CB Search）」という製品を提供し、6年以上が過ぎた。CB Searchは企業内の不定形な情報を自由に情報検索できる概念類似検索エンジンである。概念類似検索には、(1) 言語処理に必須となる形態素解析と、(2) 類似度判定に必須となるベクトル空間モデルを利用した統計処理により実現されており、この2つの基礎技術により、高精度かつ高速なテキスト情報処理が実現される。なお、ここではCB Searchの技術的な解説ではなく、実際にCB Searchを活用し、ナレッジ・マネジメントを実践している大林組の取り組みについて述べる。

(株) ジャストシステム
 浮川 和宣
 k-ukigawa@justsystem.co.jp

(株) ジャストシステム
 植松 直也
 naoya_uematsu@justsystem.co.jp

(株) ジャストシステム
 鈴木 高
 Takashi_Suzuki@justsystem.co.jp

(株) 大林組
 太田 洋行
 ota.hiroyuki@obayashi.co.jp

ゼネコンのものづくりとナレッジ・マネジメント

「ゼネコン」のものづくりについて少し紹介する。奇異に思われるかもしれないが、一般にゼネコンと称される総合請負業（GENeral CONtractor）は、「非製造業」に分類される。建設作業を行っている現場では、いわゆる「現場監督」と呼ばれるゼネコンの現場担当者が全体的な工事管理を行い、実際の作業は専門工事業者が行う。見方を変えれば、ゼネコンは専門工事業者の建設作業に従事するための「情報を加工し、マネジメントする」のが業務といえる。

ゼネコンは、企画、設計、保全管理、技術開発という段階で建物や構築物に関するさまざまな「ナレッジ」を組織的に取り扱うが、直接的なものづくりである施工段階は現場事務所が担当する。工場のラインで生産される自動車や電化製品とは異なり、建造物は諸条件に応じて

注文で1つだけ作る「特注品」である。同じように見えるビルでも、予算や工事期間などクライアントの意向が異なり、さらに気候条件や敷地の持つ特性、周辺環境などもまちまちである。これらを総合的に踏まえた上で、ゼネコンはプロダクトとプロセスの両面から各段階で最適解を導き出すことが求められる。

このような特注品を生み出す最前線基地が現場事務所であり、大林組では土木部門、建築部門を合わせると国内外に1,000カ所以上も点在する。これらは、工事期間だけ一時的に設置されるオフィスであり、一般的には「工事事務所」とか「作業所」と呼ばれる。工事規模によって、数十人のスタッフが詰める事務所から1人ですべてを切り回す現場までさまざまである。さらに、同業者と共同企業体を構成する場合も少なくない。

施工管理を行う現場担当者は、建造物の品質を確保しつつ、効率的に建設作業を進めなければ企業としての収益を生み出せない。同じ建物でも工事の進め方によって品質や収益に雲泥の差が生じることになる。たとえば通常はA、Bという順序で行う2つの作業を、何らかの工夫により同時並行で実施できれば、工事期間を短縮できる。その結果、竣工時期を早め、工事中に利用する足場などの仮設資材や建設機器のリース料を減額することができ、ひいてはビジネススピードが加速化している現在では顧客満足度の向上につなげることができる。

現場担当者はよくオーケストラの指揮者に例えられる。いろいろな作業を分担する専門工事業者の能力を「いかに引き出すか」が腕の見せどころとなる。実際に手を動かす作業員から「あの会社の現場は、段取りがいいから仕事がしやすい」といわれるように、効率よく作業できる環境を整えることも現場担当者の使命である。このように情報を集めて加工し、プロダクトとプロセスの両面から工夫して、複数の専門工事業者が連携できるように段取りしていくことが、ゼネコンにおける現場のナレッジ・マネジメントといえる。

もちろん、このようなものづくりは、ゼネコンと専門工事業者の間で完結しているわけではなく、巷で指摘されることの多い建造物に関する生産性の低さを向上させようと、多岐にわたる関係者もそれぞれの立場で取り組んでいる。公共工事に対して中央官庁や地方自治体では、電子入札や電子調達が進められており、その透明性を高めるために垂直方向の標準化が進んでいる。また、クライアント、設計事務所、共同企業体を構成する同業建設業者、専門工事業者が水平方向で、生産性を向上させるための標準化や電子化が進められている。産業全体のすそ野の広がりを考えると急激な変化は難しいが、生

産性を向上させる改革は着実に進行していると確信している。

インフラ整備と情報提供手段の見直し

限られた人員でさまざまな課題を解決し、トラブルに対処しなければならない現場事務所に対して、それを支援する本社や支社のオフィスを「常設部門」と大林組では呼んでいる。

直接、クライアントにサービスを提供する営業部門や設計部門もあるが、常設部門はさまざまに現場事務所をバックアップしている。こうした部門は、施工技術や安全管理にかかわる参考資料を配布したり、技術発表会を開催したりして、具体的な施策を通じて現場事務所をサポートしている。このうち、参考資料は配布対象や改訂頻度は異なるが、冊子としてまとめられ定期的に配布される場合が多かった。

具体的な活用例として、阪神・淡路大震災後に某鉄道高架橋工事の設計変更への対応を紹介する。この工事では、橋桁やそれを支える柱の鉄筋量を当初設計時よりも増やすように求められ、その結果として過密配筋にコンクリートをどう充填したらいいのかという技術的な課題が生じた。当時は、技術資料である「工法と技術」により社内で保有する技術の概要を網羅的に閲覧でき、担当する現場係員は、これを頼りに「高流動コンクリート」を見つけて、課題解決を行えた。ただ、この冊子は2年に一度しか配布されず、6～7年前まで全社的な情報ネットワークがまだ整っていないため、情報提供がタイムリーに行えていなかった。1998年には国内現場事務所から一部で社内ネットワークにアクセスできるようになり、従来の情報提供と活用の枠組みが大きく変わり始めた。

建設市場の縮小に伴う競争激化のために、業務の効率化とアウトプットの均質化が一層求められるようになっており、配布される冊子による技術情報と個人あるいは現場事務所内の対応に頼る従来手法では、効率化を向上させることは限界に達していた。

もともと情報の収集と整理、配布を担当する常設部門が存在し、利用者にアウトプットを提供する仕組みも確立していた。情報インフラとして整いつつあったコンピュータ・ネットワーク上で流通する情報に対して「イントラネットにおける技術情報の蓄積と活用」を施策の1つとして推進し、よりタイムリーで費用対効果が高い技術情報を全社的に提供する仕組みへと昇華させようと

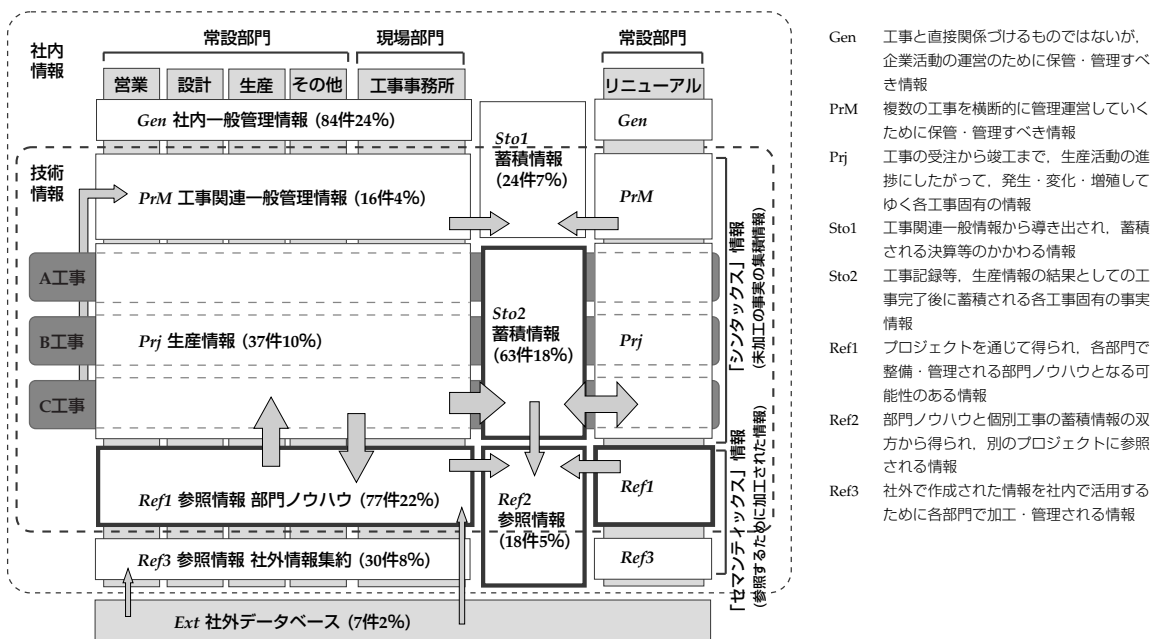


図-1 技術情報マップ

という動きが始まった。

実態調査と技術情報マップの作成

このような時期に情報システム部門は、情報の共有化や迅速な伝達を実現する「情報保管管理システム」のガイドライン策定に着手した。まず、導入対象となるユーザ部門にアンケート調査を実施して、「どのような技術情報がどこにあるのか」を洗い出し、「技術情報マップ」を作成することから始めた。

およそ技術情報を保有すると考えられる本社内の所轄部門を業務分野別に選び、情報内容、利用目的、作成部門、保管部門、保管状況、メンテナンス状況、利用部門とその頻度、情報容量、保管媒体などの調査項目を一覧表にしたアンケートを送付した。

その結果、有効回答が41部門から寄せられた。特定技術をスプレッド・シートにまとめて利用する施工実績一覧表から月次工事進捗報告書まで、ボリュームに大小の差はあるものの、実数で356種類もの情報を全社で保有していることが判明した。

内容を詳細に調べると、技術情報はその実情に即した媒体でそれぞれ蓄積して、グループウェアなどを活用した情報の2次利用や相互利用の萌芽も認められることが判明した。その一方で、同じ内容の情報を複数部門で重複して作成・保管している事例も存在していた。

アンケート結果を整理するために、横軸方向に業務フローに合わせて部門を配して、「その情報がどの段階で、どのように利用されているか」を考察してマッピングしたのが、「技術情報マップ」である(図-1)。

内容的にはおおまかに

- ・生産情報-プロジェクトの進捗に伴い作成される企画書や設計図、打合せ記録などの情報
- ・蓄積情報-プロジェクト完了後に保管される記録情報
- ・参照情報-生産情報や蓄積情報から加工された普遍性のあるノウハウとして活用される可能性がある情報に大別でき、蓄積情報および参照情報が大きな比重を占めることが確認できた。

さらに内容にはこだわらず、電子化されているかどうか、インデックスが付与されているかどうか、そして電子化されている場合には、コンピュータ・ネットワークを介した利用ができるかどうか、という観点からも分類を試みた。これらは、その情報の検索性と関係が大きく、「インデックス付きの電子化情報」がネットワーク環境で利用できる最も望ましい姿であることは言うまでもない(図-2)。

情報蓄積技術の要素シイヤ

また、技術情報に関する実態調査の分析と並行して、それを具体的に実装するための情報蓄積に関するコンピ

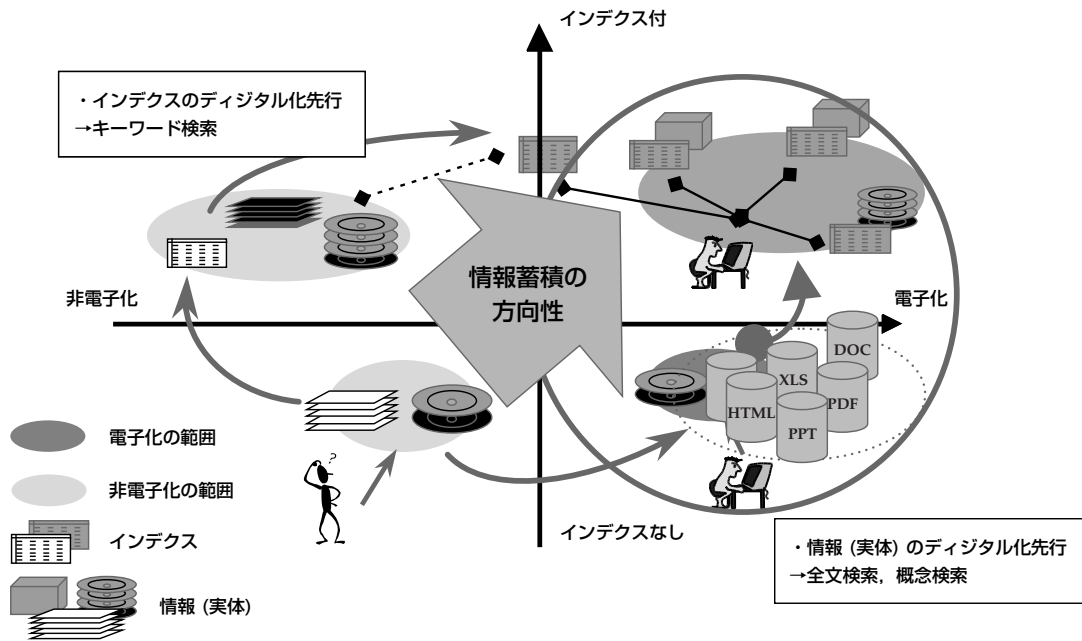


図-2 情報蓄積の方向性

ユータの利用技術についても調査した。長期的に広く通用するであろう要素技術を選定することは、このような情報の時間軸に対する相互運用性を維持する上でも肝要なことであると考えたからである。最終的にはコード化、フォーマット、メディア、アプリケーション、システムという5層構造のレイヤに分解して要素技術を捉え、レイヤ間の独立性や高い市場占有性による開発費投資の期待、デファクトおよびモジュールの標準化動向、既成の社内要素技術との相互運用性などをもとに、個々の技術を評価した。

5つのレイヤのうちメディアとフォーマットを中心に調査したが、重要なのはこのレイヤの独立性を確保できれば、システムやアプリケーションがたとえ陳腐化しても、その部分を取り替えるのが比較的容易だということである。

さらに、静的な技術情報のうち、蓄積情報は容量が大きくなりがちで、そのままでは2次利用の有効性が少なくなる。現場事務所で技術情報を活用するためには、常設部門で蓄積情報を分析・評価してその価値を判断するか、あるいは意味付けをして参照情報に抽出・昇華させる必要があると判明した。

現場への技術情報の提供が課題の1つであることは先に述べた。イントラネットにかかわる技術は未知の部分が多く、当初は全社的な利用展開を進めていたグループウェアの補完的な存在でしかなかった。その利用も常設部門の草の根活動として位置付けられ、たとえばブーリアン検索による全文検索エンジンも試用段階にとどまっていた。

大林組のこのような時期にジャストシステムは、製品化直前のCB Searchの概要説明とデモンストレーションを常設部門の情報化推進担当者を対象に実施した。CB Searchの製品コンセプトが「ユーザ企業に受け入れられるかどうか」、「ユーザーニーズに合った技術であるかどうか」という市場調査の意味合いも含んでいた。

開催後に実施したCB Searchに関するアンケートでは、参照情報への適用ニーズが高い、社内標準OA環境で利用できるOAアプリケーションによる非定型の文書ファイルへの利用ニーズが高い、という回答が数多く寄せられた。

大林組の情報システム部門では、個人向けに開発された「CB Search20」を導入し、1997年12月～1998年1月にかけて機能検証を実施した。この検証でコンテンツ面からのユーザーニーズとシステム面からの要件を満足できる見通しが立ったので、イントラネットからの操作性やメンテナンスの簡便性を考慮して、Web版「CB Search1000」を導入し、技術情報に対する適用を図った。その後、新規に専用サーバを導入して1998年9月に常

CB Searchの適用と展開

コンピュータ・ネットワークが整備されつつある中で、

設部門に試用公開した。この時のコンテンツは、技術研究所の「研究ダイジェスト」だけであった。現場事務所向けのネットワーク環境はまだ構築途上で、セキュリティ面の問題が解決されておらず、全社で一様なサービスが行える状態ではなかったため、当面は常設部門をユーザ対象とした。このような導入の経緯を経て、当初の主要ユーザとして想定していた現場事務所に公開できたのは、Webシステムとしてネットワーク環境でセキュアに設定できた2000年6月になってからである。

利用状況と新たな課題

導入後5年あまり経過しているこのシステムには、現在、人日計算で1月に3,000件から4,000件のアクセスがある。つまり、1日200人から300人は何かしら検索をしている状況である。また、このうち7割ぐらいが工事事務所からの利用で占められるようになってきている。データベースの数と文書数は13部門、74データベース、約10万2,000文書と増え、そのうち5万文書が技術情報である。

CB Searchの技術情報への適用に関しては、ユーザそれぞれの立場でメリットを示すと次のように示すことができる。

- ・エンドユーザ：複数ある文書データベースに対して横断的に検索できる
- ・情報提供部門：日常業務で利用するOA文書ファイルを作成する延長で、文書データベースへ移行できる
- ・情報システム部門：個別に検索システムを開発する必要がなく、文書データベースの追加が容易に行える

ただ、エンドユーザが、どのようなコンテンツにアクセスしているのか、それが有用なのかどうか、現場事務所が「情報を加工し、マネジメントする」ことで段取りをうまくつけているかどうか、については、把握できているとは言い難い状況である。

ナレッジ・マネジメントの観点からすれば、エンドユーザのアクセス状況を分析して、どのようなキーワードで検索しているのか、需要と供給のミスマッチが起こっていないか、まで踏み込んで検証しなければならない。利用状況を分析した結果を、情報提供部門にフィードバックし、情報をさらに利用しやすいかたちで提供できないか模索している。

ジャストシステムの新しい知識創造への取り組み

このように大林組で活用されているCB Searchは、技術情報のマネジメントに一定の成果を上げている。大林組ではさらに利用しやすい情報提供を目指しているが、検索条件と検索結果というログ情報だけではエンドユーザの利用実態を把握しきれない状況に至っている。これに対して、ジャストシステムでは、「そのドキュメントが利用された前後の関係が重要であり、その背景情報の「気化」が進むと、本来、情報が持っている有用性が薄れていく」という仮説に基づいて新たな取り組みを進めている。たとえばAという情報に対し、それを誘発した元の情報が何であるかということに辿りつけないと、その情報に至るまでの因果関係が見失われ、成功に至るまでのシナリオが見えなくなる。情報を「気化」させることなく「固形化」させ、ナレッジとして活用することが最重要課題である。

そこで、次のステップとして、高度な自然言語処理を基本に、情報をとりまく前後・背後関係（情報コンテキスト）をあつかえるような統合ナレッジソリューション「GrowVision（グロービジョン）」の開発を進めている。

GrowVisionでは、情報の入力、蓄積、分析・再利用といった情報ライフサイクルをトータルに支援・管理するとともに、ワークスペースと呼ぶ仮想的なコラボレーションの場を提供することで、情報コンテキストを自動抽出・管理できるようにしている。このような新しい知識管理の実現において中心的役割を果たすのは高度な自然言語処理技術であり、また、人と情報の関係性や情報コンテキストの研究とシステム化である。ジャストシステムはこれらコア技術を用いた新しい知識創造プラットフォームを今後も提供することで、企業における知的情報マネジメントおよび知識創造を支援したいと考えている。

参考文献

- 1) 福田，太田，網脇ら：建設会社における技術情報の蓄積に関する考察
(1) -実態調査に基づく技術情報の類型化と情報蓄積の方向性の検証-，日本建築学会，第20回情報システム利用技術シンポジウム論文集，pp.241-6（1997）。
- 2) 堀内，大友，太田ら：建設会社における技術情報の蓄積に関する考察
(2) -適用事例による要素技術レイヤーの有効性の実証と知識創造への役割-，日本建築学会，第22回情報システム利用技術シンポジウム論文集，pp.199-204（1999）。
- 3) 太田洋行：技術情報マップ作成し，CB Searchで全社KMシステムを構築，日経BP社技術研究部 InfoTech シリーズ，実践ナレッジ活用法 眠る文書情報を価値に変える，pp.160-172（2001）。
(平成15年9月4日受付)

