

自然言語処理技術による情報マネジメントの実際

# 次世代の企業間協業・連携への支援事例

## －文書ライフサイクルとコラボレーションへの自然言語処理の適用－

今後の企業競争力の強化にとっては、異業種を含む企業間協業・連携が重要となる。その実現には、企業間での文書の共有、文書類が生成された背景情報の共有が必要であり、それらを促進する上で、自然言語処理を用いた文書検索や文書要約が重要な役割を果たす。本稿では文書を情報・知識の伝達媒体とした協業・連携機能の実現に、自然言語処理技術を活用した事例を報告する。

### 業務プロセスの変化と企業間連携

従来、日本の製造業は、商品企画から保守・サービスまでのすべての業務プロセスを1つの企業で行う、垂直統合型の事業体制が多かった。しかし、今後は、企業本来の能力とスピードを活かして競争力を強化するという観点から、複数の企業がブロードバンド・ネットワークを活用して業務プロセスを協同で行う、水平統合型の事業体制が増えてくると考えられる。

### 文書ライフサイクルとコラボレーション

文書ライフサイクルから見たコラボレーション領域を図-1に示す。創造的な革新は、多くの場合、収集した各種情報を昇華させ、新たなアイデア、方法論を生み出す場面で行われる。個人の頭の中でもややもやしていた仮説や断片的な考えが、ある瞬間に結合したひらめきとなり、頭の中で文書化されてさらに明瞭になり、新たな発見に繋がると考えられる。その契機となる要因はさま

富士ゼロックス（株）  
横田 元  
hajime.yokota@fujixerox.co.jp

富士ゼロックス（株）  
田邊 栄一  
eiichi.tanabe@fujixerox.co.jp

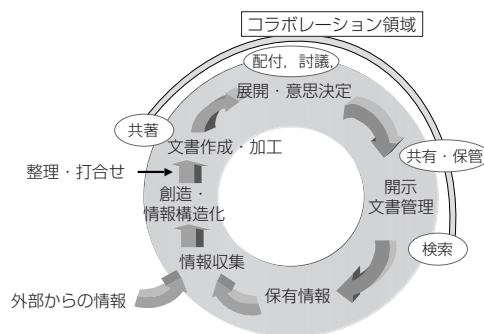


図-1 文書のライフサイクルにおけるコラボレーション領域

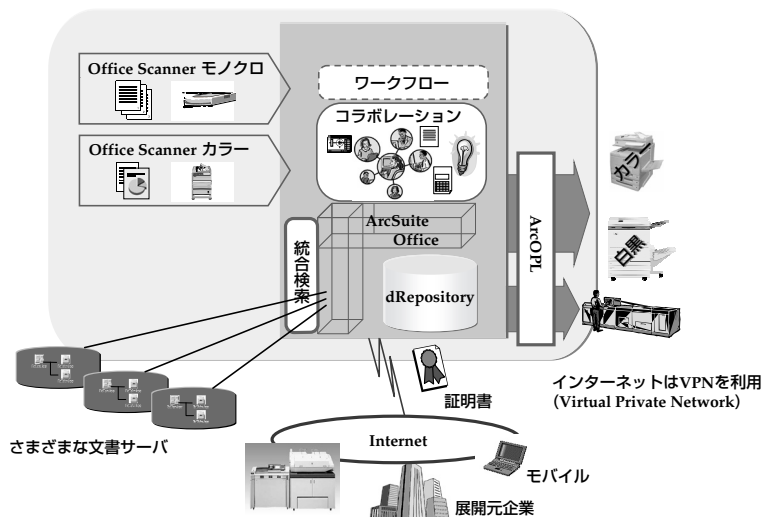


図-2 「Arc」システムイメージ

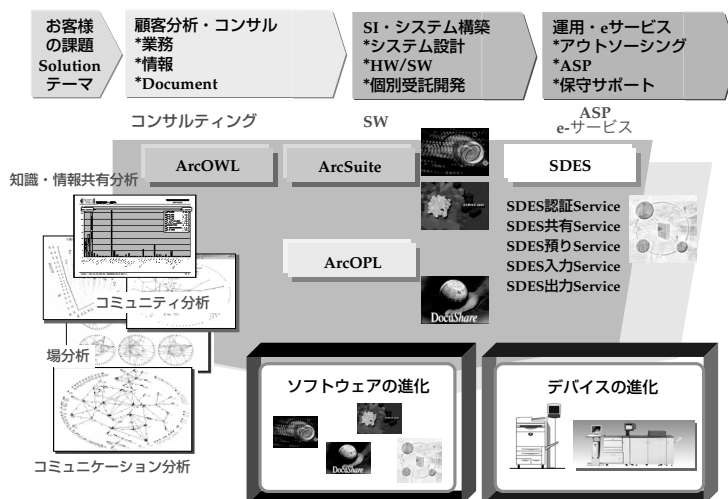


図-3 「Arc」コンセプト

さまざま考えられるが、思索であったり、人との議論であったり<sup>1)</sup>、異分野の知識との出会いだったりする。その意味で非定型で自由闊達な意見交換や、さまざまな有用な情報への簡単なアクセス、開かれた議論ができる場などが、新たな発想を生み出す機会を増す可能性が高い。すなわち図-1の文書ライフサイクルで、形式知化された文書を通じた企業間コラボレーションにより、競争優位な企業力を構築できる可能性がある。この形式知化された文書を媒介としたコラボレーションに、文書検索技術や文書要約が役立つ。

### 事例「Arc」システムイメージとコンセプト

以下に「Arc」システム事例を紹介する。システムの主な機能は2つあり、第1は文書管理である(図-2のArcSuite Office部分)が相当する)。文書/図書の属性・全文・関連(類似)文書検索<sup>2)</sup>、文書分類、文書要約表示<sup>3)</sup>文書間関係表示や文書単位のセキュリティ管理ができる。また、統合検索としてロボット型文書収集と文字情報抽出、形態素解析からインデクス生成が行える。第2は場の提供によるコラボレーション機能であり、タスク単位に登録したメンバー間のコミュニケーション支援と文書管理を行う(図-2のコラボレーション部分)。ソリューション全体をコンセプトとして(図-3)に示す。

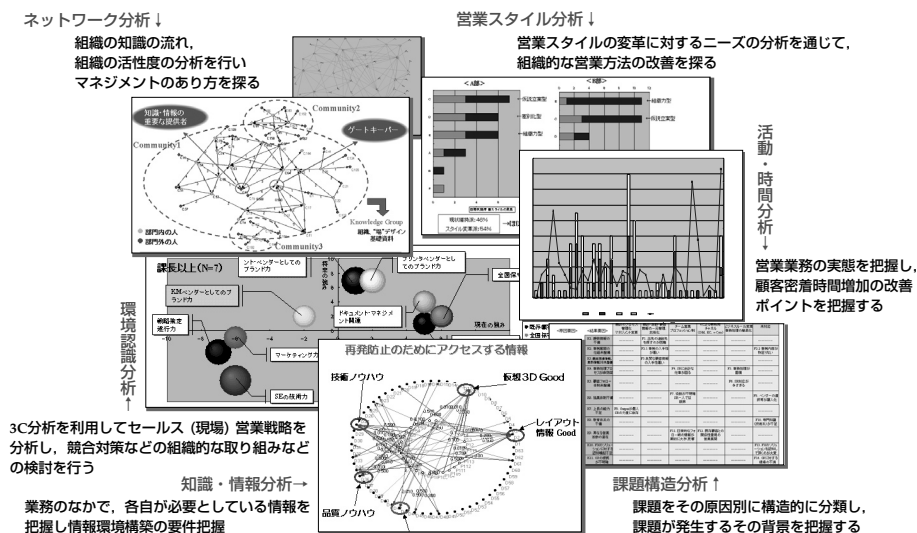


図-4 顧客分析結果の一例（可視化・定量化）

顧客の業務プロセスを、文書内容、文書の使われ方の視点で分析を行う。併せて業務にかかわるメンバのコミュニケーション（フォーマル／インフォーマル）のされ方、内容を分析し、キーとなるプロセスとそのプロセスで使われる文書内容を特定し、文書管理の仕方、使い方を提案する。その後ソフトウェアパッケージを組み合わせ、システム構築を行う。図-3に示す ArcOWL が業務プロセス・文書分析（顧客分析・コンサル）に相当する。ArcSuite は文書管理の基本パッケージで、オプションとしてコラボレーション機能を有する。ArcOPL はプリント系のソフトウェアで、ここでは説明を省略する。SDES はパブリック・キー・インフラのセキュリティ技術であるが、ArcSuite と今後接続される。

以下、ArcOWL による業務分析機能、文書管理機能、コラボレーション機能について述べる。

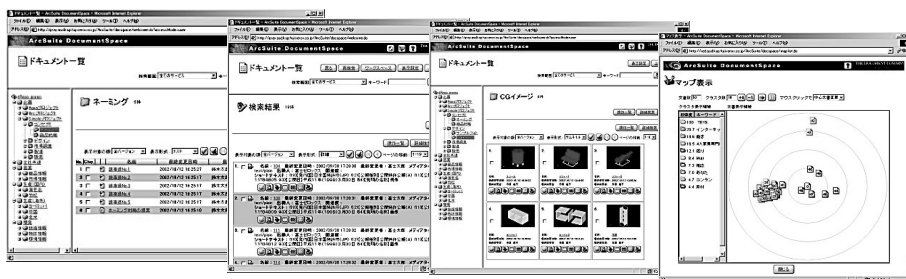
### ArcOWL による業務分析

ArcSuite をより効率的に業務に適用するには、顧客の業務分析を最初に行う。この分析では、文書がどのように使われているか、どんな情報（言葉）が人々の間を流れているかを分析する。また実際の業務にかかわるフォーマル／インフォーマルなコミュニケーションで、どんな情報が交換されるかを調査し、業務プロセス上の課題を抽出する。ここでは種々のツールを使い、言語情報を含めた業務とコミュニケーションの関係を分析する。

### 顧客分析例

最初に顧客それぞれにヒアリングを実施し、以降の分析ステップ（ネットワーク分析、業務スタイル分析、環境認識分析、知識・情報分析、活動・時間分析、課題構造分析）でどの範囲のプロセス、人員、文書の分析を行うかを決定する。ネットワーク分析では、実際に誰が情報を発信し、その情報がどのように流れていくかをチャートに表す。図-4 ネットワーク分析では「ゲートキーパー」に情報が一度集まり、情報を再配信している様子が示され、どの役割の人が多くの情報を発信しているか分かる。業務（営業）スタイル分析では、個人単位で一匹狼的にやるとか、組織的にやるなど業務のスタイルを調べる。業務メンバがどのぐらいの構成でどのように存在しているか、仕事のやり方を分析していく。環境認識分析では、職場のメンバがどういう意識を持って仕事をしているかを調査する。たとえば自分の会社の強み弱みはどこにあるかをグラフィカルに表していく。知識・情報分析は、どういう情報に多くアクセスしているかを表している。これは、仕事をやっていく上でキーとなる文書を分析するための図となる。活動・時間分析は、実際の業務に対し、何にどのぐらい時間を使っているか等が、色分けされる。縦軸と横軸はさまざまな取り方をするが、こうしたグラフで仕事を可視化し分析を行う。最終的には、課題構造分析において業務実施上の課題を表にまとめ、「顧客に文書化とプロセスの改善」を提案する。分析の結果得られた文書体系に、フォルダやキャビネットを定義し、文書を構造化して文書管理システムを構築する。

◆ 文書の体系化管理



◆ 統合検索による広範囲な情報提供サービス

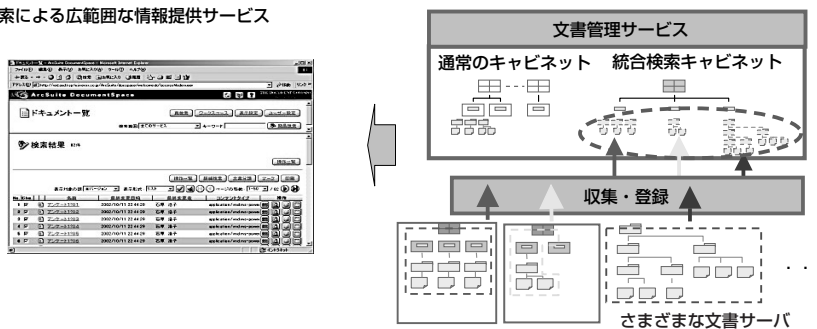


図-5 文書管理と統合検索

文書管理と統合検索

ArcSuiteの文書管理機能では、属性検索、全文検索、関連文書検索ができ、さらに文書に含まれた単語の類似度をベクトル計算により分類することができる。図-5に実際に文書を検索した結果を示す。文書タイトル表示、要約文表示、サムネール表示を選択できる。文書分類の結果は、文書間の距離を抽象的に表現した2次元マップに表示できる。統合検索機能は、文書管理システムの一機能であり、ロボット型文書収集機能にて（他の）複数のサーバに分散して格納されている文書を文書管理システム内へ取り込み、検索用の情報を生成する機能を提供する。

文書管理システムは企業レベルで使えるスケールを狙い、以下を実現している。

- ・ 1,000 万文書程度の格納、検索
- ・ 文書単位のアクセス管理  
人名もしくは役割に対して参照／書き込み等のアクセス権を与える。
- ・ 文書の版管理  
文書変更時は文書をロックし、他人が文書に対する書き込みを禁止する。

コラボレーション、その見せ方

コラボレーションに関してはタスクの概念を入れ、タスク単位でのメッセージ、文書の格納管理と表示機能を有している。タスクメンバが文書を格納すると、あらかじめ指定したメンバへ、メールにより、新規格納があったことが通知がされる。こちらの文書管理は時系列的な文書の格納を行っている。タスク定義においては、親子関係のタスク定義、独立したタスク定義が行える。関係しているタスクの一覧表示、タスク間の構造表示やタスクの進捗状況に応じた色表示ができる（図-6）。

「ArcOWL」での分析結果に対する顧客の声

What's New 新たな情報周知の重要性を再認識

今まで約 20 社を分析した結果からは、社内にある新しい情報が社員に周知されていないケースが多かった。実際は社内にあるにもかかわらず、それが社員にうまく伝わっておらず、社員は新しい情報を得ようと一生懸命社外にアクセスしていた。

Know Where, Know Who による効率化を実感

「どこに知識があるか」、「だれがそれを知っているか」の情報が生産性改善に繋がった。

◆ コラボレーション

- ▶ タスク単位でのメッセージ、文書の格納、表示
- ▶ 複数タスクのネットワーク上での場の実現

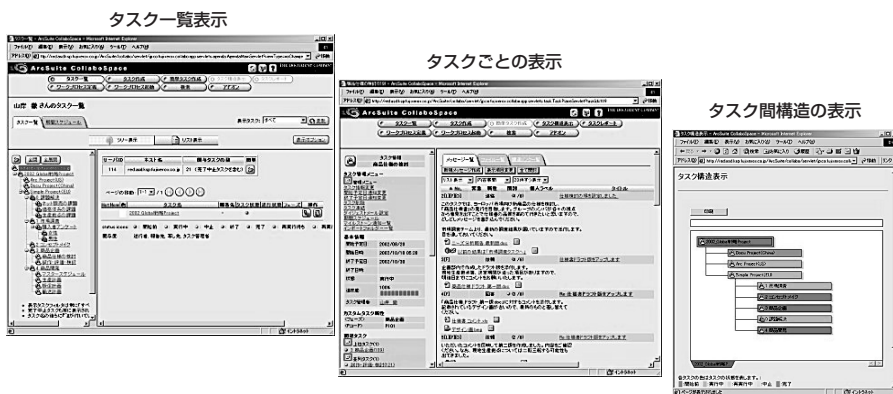


図-6 コラボレーション, その見せ方

コミュニティ作りの有効性を発見

コミュニティをうまく作ると効果があった。たとえば開発段階で営業の人がコミュニティに入っていると、市場導入される製品に対するリスクが低減されるというケースがあった。この他にも、突発的な事故などに組織横断的に対応するプロジェクトでの、密なコラボレーションなどが期待できる。



Arc の事例は、実業務に即した文書の体系的な管理と、非定型的な組織間のコラボレーションを統合したコンセプトにより、以下の効果を訴求している。

- ・ 主幹業務における文書管理体制の改善→経営効率向上
- ・ アドホックなコラボレーション（タスク運営）支援
- ・ 主幹業務（プロセス）と臨機応変な業務（プラクティス）の相互作用の促進

今後の自然言語処理への期待

文書格納時の性能向上が必要である。前述したように、全文検索、関連文書検索を行うために、文書からの文字情報の抽出、形態素解析、文書固有の特徴語の抽出、インデックスの生成において、性能の向上が必要である。特に形態素解析に処理時間を要している。現状では一晩に数万文書程度を格納できるが、新規導入時にシステムの移行に時間を要する場合がある。1桁くらいの性能向上を期待したい。さらには自然言語処理自体の性能を向上させることにより、意味解析に基づく要約作成機能への拡充、そして意味解析をさらに進めた、自動文書生成機能などによる新しい“気づき”への支援機能などが期待される。

参考文献

- 1) Schrange, M.: Shared Minds, 邦題: マインドネットワーク, プレジデント社.
- 2) 永峰, 梅基他: 関連文書検索システムの開発 (1~4), 情報処理学会第58回全国大会 (H11年).
- 3) 岡, 小山, 宮内: 句表現要約手法に基づく要約システム, 99/03, 言語処理学会.

(平成 15 年 9 月 4 日受付)

