

ナレッジマネジメントを支援する文書オントロジーを 利用した類似文書検索システムの研究

The research of similar document retrieval system by Document Ontology to support
Knowledge Management

井口 勝¹ 小谷 善行^{1,2}

Masaru Iguchi¹, Yoshiyuki Kotani^{1,2}

¹東京農工大学 工学府 電子情報工学専攻

¹Department of Electronic and Information Engineering

Graduate School of Engineering, Tokyo University of Agriculture and Technology

²東京農工大学 工学研究院 先端情報科学部門

²Division of Advanced Information Technology & Computer Science

Abstract: 企業はそれ自身の拡大のため、また合併吸収などによって、それら企業内の業務体制が変わるとともに、企業の知識を表現した多くの文書の管理体制も変わることが多い。特に過去の文書情報を確認したい場合、現在のそれとの関連がないと、過去の事例の存在を知らない検索者にとって、その検索は不可能に近い。本研究は開発部門を例に挙げ、文書ドメインオントロジーを利用して、上記の環境下でも一連の文書を検索可能なシステムを提案するものである。

1. はじめに

企業内では、その歴史の長さに応じて過去の業務によって多くの蓄積された文書（ナレッジ）が存在している。特に製品開発部門では、このナレッジを効果的に活用することで、生産性や企業価値を高めることができる[1]と共に、その固有技術を次世代の技術者に継承できると考える。ここで本稿で述べるナレッジとは、企業活動特に製品開発部門で必要となる知識や方法論(Know-How)を指す。しかし、企業がそれ自身の規模拡大や、昨今の業界再編に伴い文書の管理体制が変化することが多々あり、結果様々な形式（例えば管理番号体系）に基づき大量の文書が蓄積されることになり、それらが活用しやすい状態に整理されず、一方でベテラン技術者によって文書で記述されるべき求めるナレッジが彼ら自身で保有するメモであることもあり、すぐに取り出せないといった問題がある。そのため、過去に文書やメモとして蓄積されているにも関わらず、その中に潜む例えば製品の不具合解消などの教訓が活かされ

にくく、担当者の転勤や退職に伴い、業務が停滞するといった問題が発生する[3, 4]。

上記問題を解消するために、ナレッジマネジメント(以下、KMとする)が注目されている[1, 2, 5, 6]。KMとは、企業内のナレッジを管理する営みを指す。KMを通じ、ナレッジを共有することで生産性や企業価値の向上等が期待されている。現状多くの企業では、KMを実現するため様々なツールが利用されているが、著者らは、上記ツールのみではナレッジを十分に活用できないと考える。それは結局はKMを運用するのは人であり、そこで蓄積されたナレッジそのものが、次世代に継承するための教育ツールであるべきと考えるからである。本件研究では、様々な形式によって関連性を失った文書を文書オントロジーによって一連の集合としてまとめ、ベテランによる技術継承を支援し、且つ業務経験が浅い人物が求めているナレッジを検索しやすくするという目的別に応じた文書検索を可能とする、システムを提案することを目的とする。本稿では研究の初期段階であるため、提案システムの概要をまとめる。

まず2節において関連研究をまとめ、3節でKMの一般論を記し、4節で筆者のうち1人が勤務する企業の開発部門を例に、文書管理の体制とその問題をあげ、5節で提案システムの概要を記す。

2. 関連研究

一般的に企業における文書を形式知として整理し、知識の共有化や明確化を図るために、AI技術の適用が研究されている[5]。最近では、オントロジーを用い、対象とするドメインの知識を整理することで、KMに活用する試みがいくつか存在する。例えば、オントロジーを活用した技術の組織的蓄積・継承を目指す研究プロジェクトがあり[4, 7]、その技術の種類を踏まえたオントロジーによる技術継承に取り組んでいる。また、人間が行う問題解決活動にある原理的知識を多面的に明らかにするため、問題解決の対象となるドメイン知識と、またそれに対するタスクとそれに基づく状態変化を明示的に記述するタスクオントロジーに関する研究も行われている[8]。また業務メールに蓄積された暗黙知に着目し、メールから自動抽出した Know-How, Know-Who を可視化し、低負担かつ不案内な利用者にも活用可能な企業内情報

共有システムが提案されている[9]

これらの研究の共通的な課題として、オントロジーの生成・管理に時間や手間がかかることが想定されているので、それらオントロジーの運用を容易にすることについても検討されている[10, 11, 12, 13]。

しかし、これらの取り組みではナレッジ抽出の精度向上については検討されているが、本論で提案する検索者のスキルや目的に応じた文書検索の方法は論じられていない。

3. ナレッジの形成

ナレッジ形成について、KMの活用モデルとして、SECIが提案されている [1][2]。SECIのモデルを図1に示す。同モデルでは、表出化・連結化と呼ばれるプロセスにより暗黙知を形式知に変換すること(以降、形式化とする)で、人間が自身の知識として、ナレッジを取り込むことが可能になるとしている。図2では、筆者らの経験に基づき、製品開発部門において業務経験が浅い技術者(例えば若手設計者)がどのようにして知識を得るかを示している。彼らは大抵外部からの要求によって、業務を開始する。彼らはまだ市場の要求を直接理解することはないので、一度社内の部門(例えば営業部門)を介し、部門の要求として情報を得る。その先、どのようにしてその要求にこたえるか(業務を開始するか)は、ベテラン技術

者の指南を受けることによって決定する。その指南を理解することによって、それを自らの知識として得る。それはすなわち「知識の創造」として自分自身に蓄積される。しかしながら、筆者らは、長く続いた国内の不況などの経営環境の変化によって人員の整理などがあり、残された少数のベテラン社員の業務負担は増加したと考える。そのような状況では、若手技術者への指南の十分な工数を得ることは困難であると考えられる。さらに前述文書(知識)管理が不十分であるならば、若手技術者が自ら当該開発における過去の事実を知ろうと思っても、不可能に近い。

本システムではそのようなベテラン社員や若手技術者がそれぞれ異なる目的で文書を検索できるように、例えば ISO 9001 で規定された文書番号体系で集合化された文書を、文書オントロジーによって検索者の目的別にそれら文書を抽出するシステムを手案するものである。



図1: SECI モデル

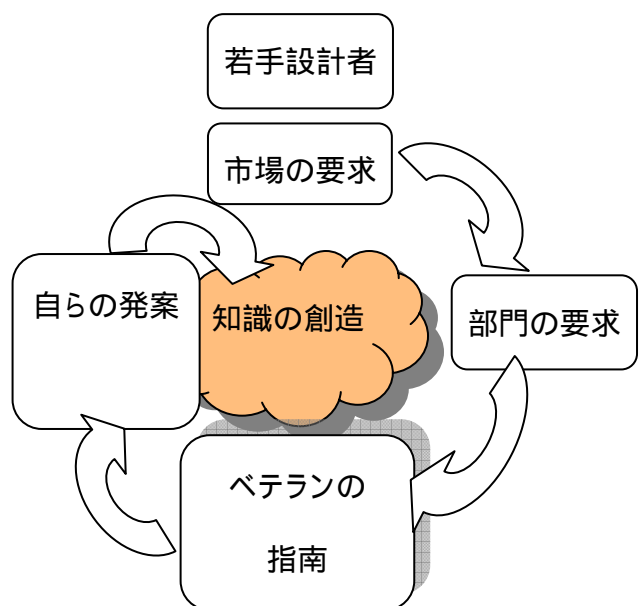


図2: 若手設計者の知識創造プロセス

4. 文書管理体制とその問題点

図3にて、筆者のうちの1人が勤務する企業の開発部門における、文書作成のフローを示す。作成された文書はその時点で暗黙的文書となる。文書管理体制に特段の変更のない場合、ISO9001に基づく社内規定によって体系化された番号台帳で管理番号を採番する。

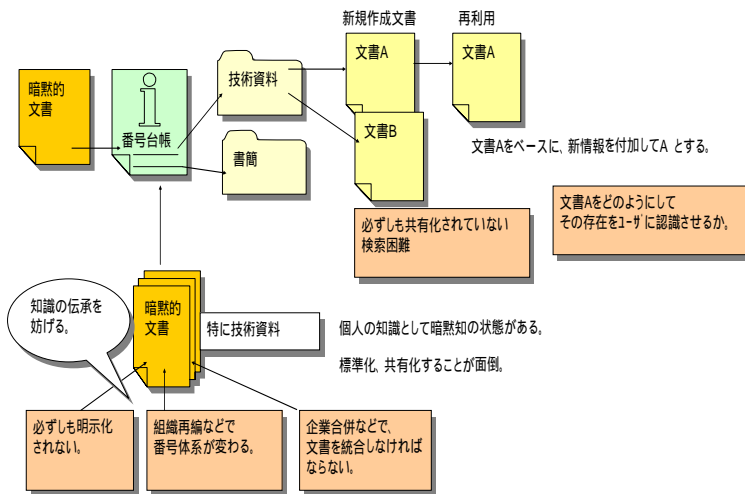


図3: 文書作成のフロー

管理番号が付与された文書は、PCサーバにファイリングされ技術部門のスタッフが閲覧可能(形式知)となる。この過程において、正式文書となるためには、管理者の印または手書きのサインを必要とするが、本論ではそれを省略する。しかし、これまでのフローでは、管理番号の採番や形式知となった文書のサーバへのファイリングなど、ほとんどが人手によって行われているので、その恣意性によって一貫性がなくなる恐れがある。また前述のように、管理体制が変わることによっても同じ状況となることが考えられる。その際、旧管理体制における文書は、管理下でない文書とみなされ「暗黙的文書」に分類されると考える。その他にもここでは示されていないが、製品開発の工程における、品質評価の文書のように、独立した複数の文書の集合体を「1つの文書」として管理番号が付与されることも考えられる。その際は、検索者によっては各文書を検索したり、また文書集合体を検索するものもある。よって、独立した文書や、ある開発中の製品と同クラスの過去の製品の開発記録としての文書を検索するなどの目的別の検索に応じるためにも、管理番号体制が変わっても一貫性を持たせることは重要な課題と考える。

5. 提案システムの概要

図4に提案システムの概観を示す。本論作成現在、筆者のうちの1人は、勤務先の企業にて80人の設計または試験研究の管理者及び担当者に対して、アンケートを実施し機能別・実務経験別に検索者が必要とする「目的別の検索」を抽出し概念化を図ろうとしている。

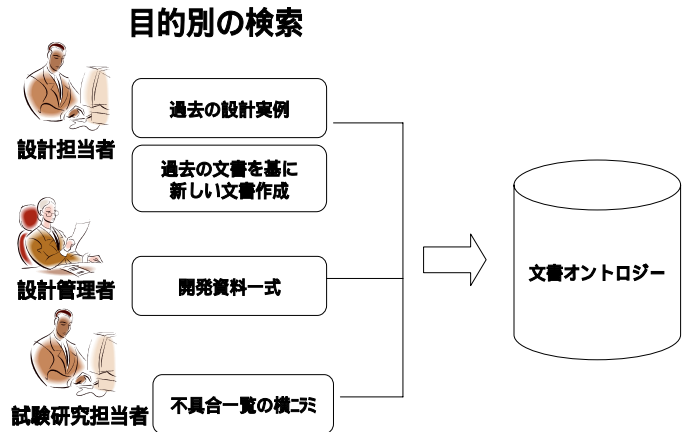


図4: 提案システムの概観

筆者らの目論見では、まず検索者を「設計担当者」(様々な実務経験者を含む)と「設計管理者」、また実機を評価する「試験研究担当者」に分けた。それらの役割上、どのような目的で文書の検索を行うかをここでは、仮に図4に示す4つに分類した。それらの目的別の検索に応じて答えを出すべく、文書が登録されたデータベースに「文書オントロジー」を予め構築する。

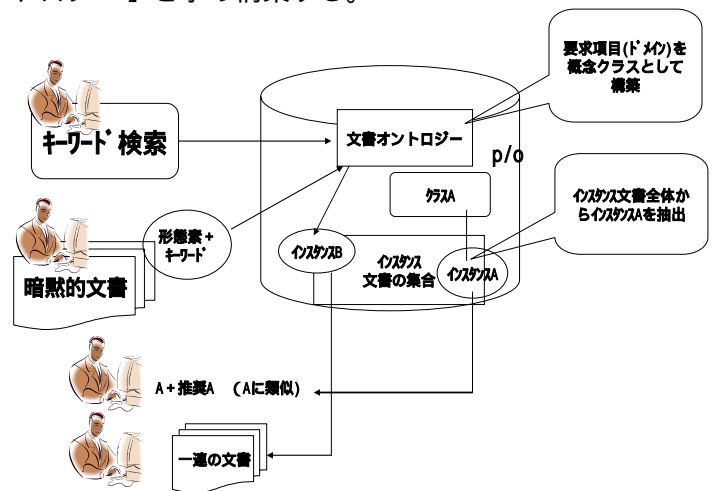


図5 提案システムの構成(概要)

文書オントロジーは現在構築中なので、詳しくは記さないが、各クラスはドメインのオントロジーとして文書を概念化する。また下位概念のクラスでは、前述の検索目的に合うように設定され、例えばそのクラスの内容をISO9001の規定された定義によって自然言語で記され、その特

長語によって、図中のインスタンスの集合から、目的に応じた文書を抽出する。

そして、検索者のスキルやニーズに応じて彼らの要求する文書を、提供する。特に設計管理者(またはベテランの設計者)のように、ある開発の一連の文書を要求する場合、単一の文書を提供するだけではなく、関連するであろうすべての文書を提供することができる。その検索方法についても、実務経験の浅い設計担当者や、ある特定の試験データに関する文書を要する検索者のように、管理番号や機種名、プロジェクト名のようなキーワードで検索する場合と、暗黙的文書の記述内容から自動的に検索するような場合に対応するシステムを構築する予定である。

の開発 集合知としてのオントロジー構築を目指して , JAWS-2008 (2008)

参考文献

- [1] 野中郁次郎, 竹中弘高, 梅本勝博訳: 知識創造企業, 東洋経済新報社(1996)
- [2] ナレッジクリエーション研究会編: ナレッジクリエーション, 工業調査会(2003)
- [3] ナレッジ・ワーカーへの投資に関するパラドックス, Gartner research(2002)
- [4] 岡部雅夫: オントロジーによる技術・技能の組織的蓄積のサポートについて, 情報システム学会第 2 回研究発表大会, SWS-05 (2006)
- [5] 山口高平: 知識マネジメントと AI 技術, 人工知能学会, Vol. 22, No. 4, pp. 461{466 (2007)
- [6] Nancy M. Dixon ,梅本勝博訳 ,遠藤温訳 ,末永聡訳: ナレッジ・マネジメント 5 つの方法, 生産性出版(2003)
- [7] 岩間貴史, 立花浩, 山崎浩志, 岡部雅夫, 黒川利明, 小林圭堂, 加藤美穂, 吉岡亜紀子, 山口高平: 業務知識の組織的蓄積・活用を支援するオントロジーに対する一考察, 情報システム学会第 3 回研究発表大会, D1-1 (2007)
- [8] 堀雅洋, 瀬田和久: 問題解決タスクのためのオントロジー, 人工知能学会, 人工知能学会誌, Vol. 19, No. 2, pp. 159{165(2004)
- [9] 間所峻洋, 中辻真, 岡本賢一郎, 宮崎純生, 原田剛, 第 19 回セマンティックウェブとオントロジー研究会資料, SIG-SWO-A802-04
- [10] 廣田健, 古崎晃司, 溝口理一郎: オントロジー俯瞰のための概念マップ生成ツールの開発, JSAI2008, 2E3-02 (2008)
- [11] 古崎晃司, 來村徳信, 池田満, 溝口理一郎: 「ロール」および「関係」に関する基礎的考察に基づくオントロジー記述環境の開発, 人工知能学会論文誌, Vol. 17 (3), pp.196-208 (2002)
- [12] 古崎晃司, 笹島宗彦, 來村徳信, 溝口理一郎: オントロジー構築入門, オーム社(2006)
- [13] 沈偉, 川村隆浩, 中川博之, 中山健, 田原康之, 大須賀昭彦: オントロジー構築サービス ONTOMO