

ナレッジマネジメントツール：ビジネス情報ナビゲーター

Knowledge Management Using Business Information Navigator

あらまし

インターネットやイントラネットの普及により、コンピュータが扱うデータ量は増加の一途である。従来は、RDB (Relational Database) のような体系的に管理・運用されたデータが主であったが、ネットワークの発展で日常的に発生するメールやWebなどテキストで書かれている非構造化データが大幅に増え続けている。これらの非構造化データの活用は、ユビキタス社会において企業競争力の決め手と考える。非構造化データは、断片的な情報として散在している場合が多く、知識として活用するためには、それらを統合して全体を俯瞰できるようにしなければならない。情報統合のコストを縮小し、断片的な情報を関係付けることにより、有機的な知識としての活用を低コストで可能とする技術が必要となる。

本稿では、断片的な情報を統合して全体を俯瞰するビジネス情報ナビゲーターについて、人検索システム (KnowWho) と金融情報システムを例に紹介する。

Abstract

The spread of the Internet and intranets is bringing a huge increase in the amount of electronic data in the world. In the past, such data control systems as Relational Database (RDB) were employed to systematically structure and manage most electronic data. However, because of the expansion of the network community, the volume of unstructured data such as e-mails and Web pages is rapidly increasing. As a result, better techniques for managing unstructured data have become critical requirements in the ubiquitous society. Most unstructured data is fragmentary, and to make the most of unstructured data as knowledge, we have to integrate it and view it as a whole from various points of view. Two major issues in knowledge management are how to reduce the cost of data integration and how to make connections among fragmentary data. In this paper, we introduce two Fujitsu products for managing unstructured data: the KnowWho Human Knowledge Navigator part of Fujitsu's Business Information Navigator and a financial information management system.



松井くにお (まつい くにお)
ナレッジ研究センター 所属
現在、言語・知識処理関連の研究に従事。



津田 宏 (つだ ひろし)
広報IR室 所属
現在、富士通研究所関係の広報活動に従事。



片山佳則 (かたやま よしのり)
ビジネスインキュベーション研究所
企画・マーケティング推進室 所属
現在、セマンティックWeb関連の研究、およびBI企画活動に従事。

まえがき

ナレッジマネジメントにおける知識創造の原則は、SECIモデル⁽¹⁾を回して、暗黙知と形式知をスパイラルに向上させるものである。現状をかんがみると、事例の蓄積による暗黙知から形式知への変換は盛んである。ただし^{たくみ}匠の技の伝達のように、暗黙知そのものの伝承や、事例の再利用による形式知の活用はなされているが、形式知から暗黙知へ内面化のプロセスが十分ではない。散在する大量の形式知には、重要な情報が埋もれて見えなくなっており、関係付けのされていない断片情報では、「知識」とは言えない「ゴミ」が蓄積されている場合もある。

情報を知識として活用するためには、再利用しやすいようにきちんとした形式に整理して蓄積する方法と、断片的な情報を統合して全体を俯瞰できるようにする方法がある。前者は、再利用頻度が高く、価値の高い情報であるならばコストに見合う効果は十分に期待できるが、情報量が多いほどその作成コストは大きくなる。後者は、情報統合のためのコストは大きいですが、分野を限定すれば自動化が可能となり、今までは「ゴミ」扱いされていたメモ的な情報が貴重な知識へと大化けする可能性を秘めている。

富士通研究所では、情報統合の基盤として次世代Webの技術をいち早く取り入れ、後者の有機的な知識の活用を実現させている。

本稿では、断片的な情報を統合して全体を俯瞰する方法を、人検索システム（KnowWho）と金融情報システムを例に紹介する。

情報統合のアプローチ

情報統合を行うに当たり、既存する膨大で多様な情報そのものを統一的なものに作り変えるのは現実的ではない。このため、多様な情報源に含まれる情報のエッセンスである「メタ情報」を共通化することが考えられる。ここで必要なのは、メタデータの形式や用語の意味（以下、スキーマ）を共通化しておくことである。

例えば後述のKnowWhoでは、様々な文書やシステムなどのオフィス業務活動を構成する情報に共通して現れる「人」や「モノ」に注目し、それらに関する基本情報と、互いの関係をメタ情報として共通に記述することで情報に新たな価値を与えている。

スキーマやメタデータの表現形式には、次世代Web技術として注目されているセマンティックWebの中核仕様であるRDF（Resource Description Framework）やOWL（Web Ontology Language）を用いた⁽²⁾ 今後、セマンティックWebの発展とともにRDF、OWLに準拠したメタ情報がWebや社内システム上に増えてくれば、将来それらとの情報交換が容易になる。

メタ情報をどう作るかは非常に重要な課題である。現実的には、自然言語処理などの技術を用いて、文書コンテンツから自動作成するなど、人手工数の削減が望まれる。

ビジネス情報ナビゲーター

断片的情報から組織化知識を導くシステムとしてビジネス情報ナビゲーターを開発した。この核は、情報統合と「見える化」である。これらの適用例として、社内でスキルを持った人を検索するKnowWho⁽³⁾と、金融機関が抱える顧客の関連情報を視覚化する金融情報システムを紹介する。

人検索システム（KnowWho）

人や人脈（コミュニティ）の検索は、「業務の情報の50～75%は人から直接得る」⁽⁴⁾や「検索の問題の一つは人を探せないこと」{Google CEO Eric Schmidtのコメント（2004.3.22）}のような理由のため、企業活動にとって重要である。これまで多くの人検索システムでは、各人が情報を人手で登録してきた。このため、情報更新コストが高かったり、信頼度がまちまちだったりして現実的にはうまく機能しないことが多い。

そこで、ここに示すKnowWhoでは、社員の日常作業（文書やメールの作成、打合せなど）から、スキルや人脈といった人のメタ情報を自動で抽出して検索可能にする。これにより、特別な情報の準備が不要となる。以下、研究者KnowWhoプロトタイプをもとに本検索システムの概要を説明する。利用シーンとして、営業・SEが顧客から問合せを受け、その技術に詳しい研究者やそのグループを見つけ出し、商談機会を逃さない、というケースを考える。

(1) 技術マップ

「セキュリティ」というキーワードを入力すると、社内の関連部署と関連キーワードの関係が技術マップとして得られる{図-1(a)}。セキュリティに関

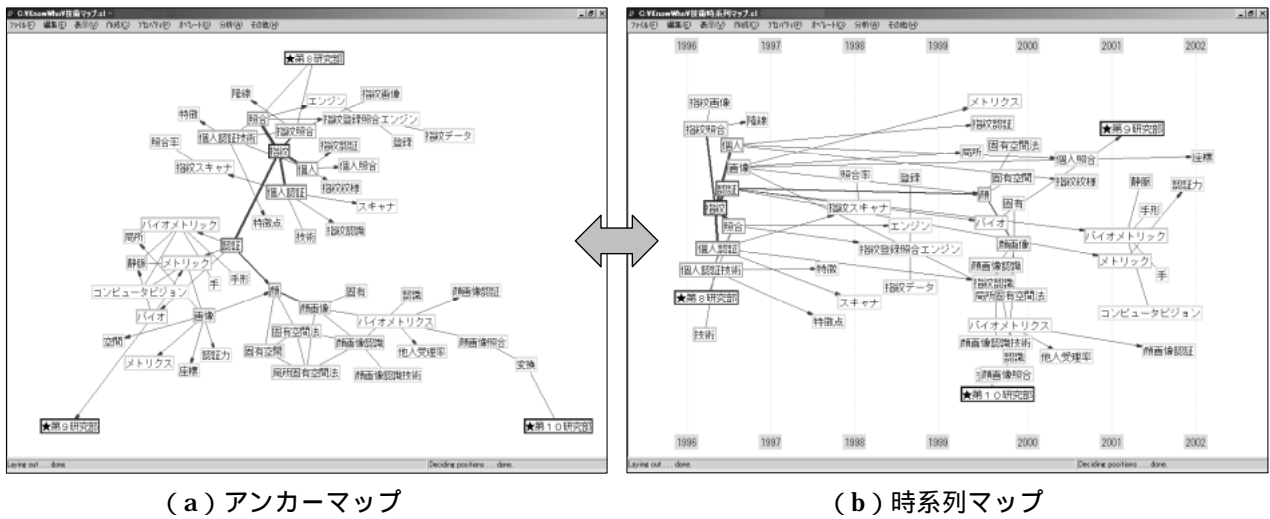


図-1 技術マップ
Fig.1-Technical term map.

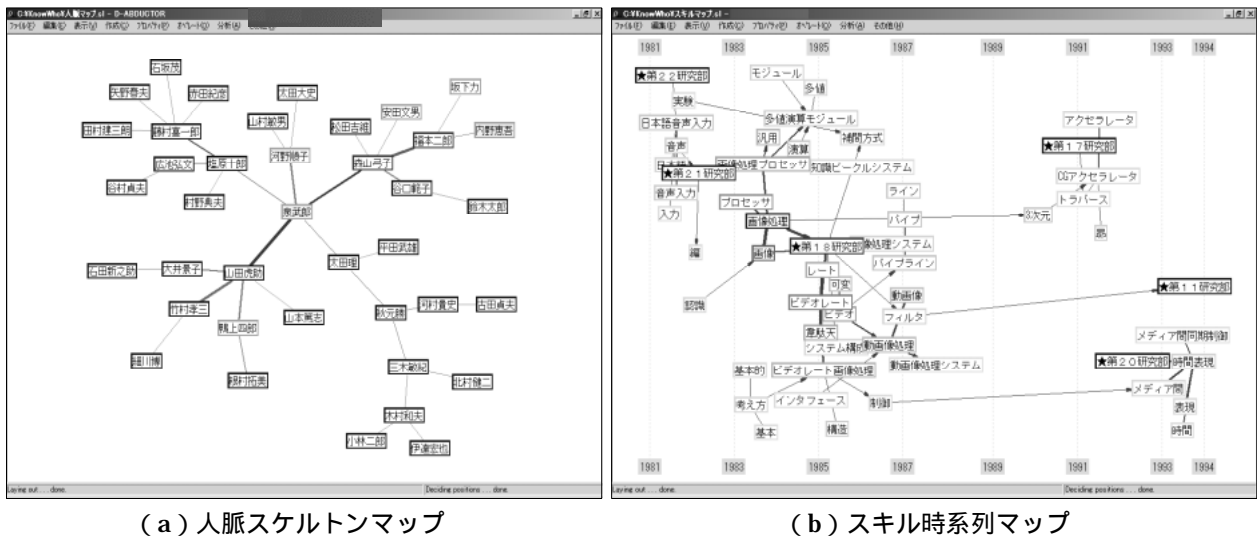


図-2 人脈マップとスキルマップ
Fig.2-Human network map and personal skill history map.

連深い（その用語を含む文書が多い）三つの部署が三角形に固定して配置され、関連するキーワードが近くに配置される。関連の強い線のみを表示させることで、例えばセキュリティに関して、第8研究部は「指紋」に関係し、第9研究部は「静脈」、第10研究部は「顔画像」に関係していることが分かる。同じ関連キーワードを時系列的に並べ替えることもできる{図-1(b)}。これによりどの研究部が何をいつ始めたかや、トレンドなキーワードも簡単に見つけることができる。

ここで、キーワードのような人のメタ情報は、研

究者が作成した、報告書や論文などのテキストから自動で抽出している。

(2) 人脈マップ

技術マップの各キーワード（例えば「静脈」）から、そのキーワードに関係の深い研究者とその人脈も表示できる{図-2(a)}。ここで人脈は、同じミーティングに参加している度合いで人の近さをモデル化し、線の太さで表示している。また、人がどのくらいキーワードを文書などで使用しているかの相対的な度合いで人のスキルをモデル化し、枠の色で表示している。これにより、図から「静脈」「セ

キュリティ」に関して泉さんを中心としたいくつかのグループがあることが見て取れる。ここでも、ミーティングのログや、テキスト内容から自動でキーワードなどのメタ情報を抽出している。

(3) スキルマップ

人脈マップで泉さんを指定し、図-2 (b) のようにその人のスキルの時系列推移を見ることが出来る。キーワードは、その人の書いた文書の中で初めて使われた日時に配置される。

以上のマップは、研究員の日常業務で作成される文書やグループウェアの情報から自動的に作成されることに着目されたい。人のスキルは固定したものではない。自動化によって常に最新の情報を取り出したり、逆に3年前の自動車業界に詳しい人を探したりというような利用シーンが考えられる。

また、基となる情報も文書やグループウェアのような意識的に作成するもの以外に、RFIDタグと組み合わせることで部屋の入出やプロジェクトでの資料説明などを記録したのも対象と考えられる⁵⁾

金融情報システム

一般に金融機関は、基幹系、法人営業推進系、融資支援系など、それぞれ異なるアプリケーションで顧客（取引先）に関する情報を、様々な形式で蓄積している。「リレーションシップバンキング（顧客とのビジネスで蓄積された情報をもとに各種の金融サービスを進めること）」というビジネスモデルでは、これらの情報を統合的に活用することが推奨される。

この金融情報システムは、これらのアプリケーションに分散した断片的な情報に含まれる、数値やテキスト情報から、資金の流れやビジネス関係構造、ビジネスチェーンを明らかにし「見える化」することで、ビジネスマッチング、商品セールスの効果的な手続き、内部統制などに活用しようというアイデアである。

例えば、通常担当者は、自分の担当顧客がどこ取引しているかは把握していても、その先がどうなっているかという関係は見えにくい。ところが、複数の担当者の情報を統合して「見える化」することで、顧客とその仕入先や販売先の幅広い関連を一目で把握できる。また、顧客間の関係も、お金の流れや仕入れ先関係など様々な種類をナビゲートしながら検索するインタフェースにより、リレーション

シップバンキングにおいて、新たなビジネスの切り口を見つけ出せる。

情報統合と見える化の要素技術

情報統合は、異種の情報源すべてを統合した新たな統合DBを作成するのではなく、元情報のエッセンスである共通の枠組みによるメタ情報を作成して統合を行う。このメタ情報ではオブジェクト間の関連情報が重要であり、「見える化」においては、関連情報を分かりやすくマップ化し、その上で情報のナビゲーションを行うことが必要である。

これらの情報統合と「見える化」を実現するためには、以下に述べる情報抽出、関係付け、可視化の要素技術が必要となる。

情報抽出技術

情報抽出（Information Extraction）とは、テキストに含まれるオブジェクト間の関係を、自動的に高精度で取り出す技術である。自然言語処理分野の技術が利用される。その中のNamed Entity（NE）抽出は、テキスト中の人や場所などの固有名詞とその種類を同定するもので、表-1のような抽出精度が得られている⁶⁾

前述の研究者KnowWhoでは、技術報告書やミーティング資料などから抽出精度の高い「日付」、「人」、「部署」などをNE技術で抽出している。

関係付け技術

KnowWhoでは、人のスキル検索が目的であるため、スキルに関する「技術キーワード」と「人名」を抽出対象として「技術キーワードの共起関係」、「ミーティングやメールなどのグループによる人の関係」、「人に依存したキーワードの系列関係」の提示をゴールとして、情報統合を実現している（図-3）

表-1 NE抽出ツール（boosting）の抽出精度

Named Entity (NE)	抽出精度 (%)
ARTIFACT (人工物)	28.31
DATE	93.53
LOCATION	81.25
MONEY	100.00
ORGANIZATION	73.07
PERCENT	90.00
PERSON	86.44
TIME	95.72
平均	81.53

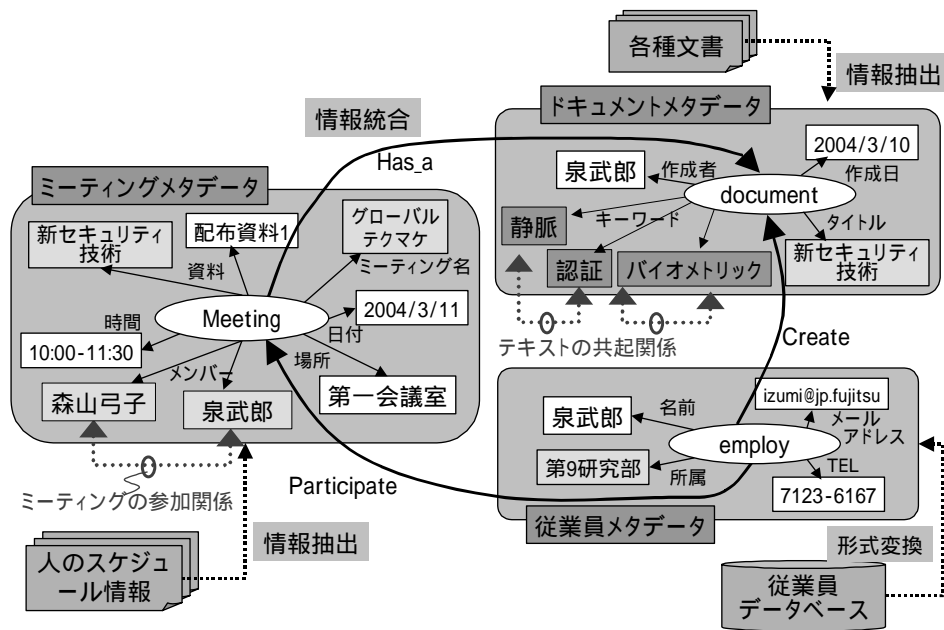


図-3 メタ情報間の関連付けによる情報統合
Fig.3-Information integration by relating metadata.

このための関係付けには、

- (1) 異なる表現でも同一と判定する名寄せ、
 - (2) 事象に基づく情報の関連付け
- の二つがある。図-3では、泉さん、森山さんが2004/3/11の同じミーティングに参加し、そこでは、第9研究部に所属の泉さんが作成した「新セキュリティ技術」という文書（静脈や認証、バイオメトリックなどのキーワードを含んでいる）が資料となっていることなどの関係が表されている。

可視化技術

可視化は、特許分析システムで構築されたグラフ表現技術⁹⁾を活用し、検索結果の図的な分かりやすい提供と、簡単操作での関連情報へのナビゲーションを実現している。

KnowWhoでは、関連部署とキーワードの引っ張り合った結果を表示するアンカーマップ{図-1(a)}や、関連の特性を強い関係を残して表示するスケルトンマップ{図-2(a)}、時系列マップ表示{図-2(b)}などを用い、各表示ノードからはメニュー選択で関連情報へナビゲートできるようにしている。

金融情報システムでは、キャッシュフロー、取引関係、融資状況などをナビゲートし、顧客や営業店などの観点で情報提供している。

む す び

ネットワークの広がりによりメールやWebなども含めて、多くの組織では膨大で多様なテキスト文書が蓄積されている。これらに含まれる断片情報を、戦略を持って、組織的に知識化し、ビジネスに活用させ、新たなナレッジマネジメントのツールとしようというねらいでビジネス情報ナビゲーターを開発した。

情報統合技術により、社内の断片情報を活用して人検索システム(KnowWho)を実現できること、また、金融情報システムでは、複数の担当者の情報を統合して、これまで見ていなかった顧客間の取引の全貌などを把握できることが分かった。いずれも、散在していた大量の形式知に埋もれていた情報を新たに知識として活用できた事例と言える。

後者については現在、研究所と営業・SE、コンサルがチームとなり、先進的な金融機関に蓄積されている実際のデータでの共同研究を行いながら、技術の実証を進めている。本技術は、さらに、流通・製造分野などへ広く拡販を進める予定である。

参考文献

- (1) 野中郁次郎ほか：知識創造の方法論～ナレッジワーカーの作法～．東洋経済新報社，2003．

- (2) 井形伸之ほか：セマンティックグループウェア：RDFを用いたKnowWhoの実現．第5回セマンティックウェブとオントロジー研究会資料，2004.3，A303-05，p.05-01-05-08．
- (3) 津田 宏：ヒューマンナレッジナビゲーター～セマンティックWebによる次世代KnowWho～．INTAPセマンティックWebコンファレンス，2003.11，p.7-1-7-5．
- (4) Gartner research：The Knowledge Worker Investment Paradox (2002)．
- (5) 津田 宏：ミーティング情報マネジメントシステム．INTAPセマンティックWebコンファレンス2005 (2005.2)．
<http://www.net.intap.or.jp/INTAP/s-web/data/conference2005>
- (6) 岩倉友哉ほか：テキストからのメタデータ情報抽出とKnowWho検索への応用．第18回人工知能学会全国大会，3H3-04，(2004.6)．
- (7) 片山佳則ほか：セマンティックグループウェア WorkWare++とKnow Who検索への応用．情報処理学会研究報告 情報学基礎，No.071，(2003.5)．
- (8) 渡部 勇：ビジュアルテキストマイニング技術 - 特許検索・分析支援への適用 - ．FUJITSU，Vol.56，No.4，p.371-377 (2005)．

