

ビジネス・インテリジェンスと人工知能技術 04

武田浩一（日本アイ・ビー・エム(株) 東京基礎研究所）：takedasu@jp.ibm.com

04

ビジネス・インテリジェンスと人工知能技術

ビジネス・インテリジェンス（BI）とは、企業が持つ社内外の情報を活用し、ビジネスにおけるよりよい意思決定を迅速・確実に行うための活動や手法に関する包括的な概念（conceptual umbrella）である。BIという用語は、Howard Dresner が1989年に命名したとされており、最近のインタビュー^{☆1}で、Dresner はBIが技術的な内容にフォーカスしたのではなく、企業内の人材とプロセスを統合してビジネス目的を達成することが主眼であると明確に述べている。Dresner はさらに、BIが情報の民主化（information democracy）といえる方向に進んでおり、少数のアナリストによる高度な情報分析から、ビジネスの現場で誰もが意思決定や行動につながる知見（insight）の獲得のためのインフラとして普及すると予見している。このためには、たとえば重要業績評価指標（key performance indicator, KPI）の目標達成や改善を可能にするBIや、BIツールの標準化と理解、業務プロセスへの統合といった諸問題について言及されなければならないが、ここでは技術的な側面を中心に、人工知能技術を応用した最近のBIについて解説する。特にテキストマイニングを利用して、従来のデータベースに含まれる数値・コード化された項目を中心とする構造化情報と、フリーテキストで記述される非構造化情報を統合することで可能になったBIの応用について詳しく紹介する。

情報系データベースの発展

BIは1990年代のデータウェアハウス¹⁾、OLAP(Online Analytical Processing)²⁾、データマイニング³⁾といった概念の実用化によって大きな発展をとげた。それまでのデータベース・システムは、主にオンライン・トランザクションの処理を効率化・頑健化することが主目的であったが、蓄積された情報を知識化し、企業の意思決定に利用するというデータベースの役割が脚光を集めるようになった。前者は基幹系データベース、後者は情報系データベースとして区別されている。情報系データベースの大きな特徴は、基本的に過去の履歴データを蓄積し、読み取り専用の情報として時系列分析や統計的分析、集計・レポート生成などの用途に使われることである。また、多次元分析という手法では、従来の基幹系データベースでは属性値としてフラットに表現されていた情報（たとえば支店名）を、次元（dimension）と呼ばれる分析軸に構造化することで（たとえば支店別、エリア別、支店規模別などの）意味的な分類や階層関係を利用した分析が容易になっている。さらに、このような分析軸を利用した詳細化（ドリルダウン）、特定の側面の抽出（ス

ライジング）といった操作が知られており、BIの標準的な対話分析手法を確立している。世界最大の流通業者であるウォールマート社では、データウェアハウスを店舗運営や商品販売の分析⁴⁾に活用しており、この分野での顕著な成功事例の1つにあげられている。

データマイニングは、BIの分野で最も人工知能技術の応用が進んでおり、統計的手法、発見的手法、機械学習やニューラルネットワークの手法などによりさまざまな規則性やパターンの獲得を可能にしている。特に、大量のデータを類似したグループに分類するクラスタリング、高頻度で同時に現れる項目の相関規則の抽出、ある項目の値（目的変数）を予測する決定木の獲得などがよく知られている。これまでに購買履歴分析、顧客管理、リスク管理、不正アクセス検知、遺伝子発現解析など、多数の応用事例が報告されている。顧客情報の分析により、キャンペーンに適した顧客の最適化を図り、ダイレクトメールの量を削減しつつ、収益を改善したといった事例が特に有名である。また、データマイニングは現在でも非常に活発に研究されており、著名な国際会議の1つであるKDD(Knowledge Discovery and Data Mining)の2005年の研究論文投稿数は465件（うち採択論文数はポスター発表を除き40件）であった。近年の技術的な特徴として、木やグラフなどの構造を持つデータのマイニング、Webを対象としたコンテンツやコミュニ

☆1 <http://www.bizintelligencepipeline.com/181501967>

ティのマイニング、センサ情報や時系列情報の実時間データマイニング、医療や国土安全保障といった新しい応用分野の拡大、などが指摘できる。

データウェアハウスやデータマイニングについては、すでに豊富な解説記事や入門書が存在するため、ここでは詳しい説明は省略している。

テキストマイニング

テキストマイニングとは、大量のテキスト情報から新しい知見を獲得する技術の総称である。2000年頃を境にBIの分野ではテキスト情報の利用が盛んになり、筆者の所属する研究グループでも、年間50万件以上の社内のコールセンタでの電話対応の内容を、よくある質問（FAQ）の作成や新製品の初期不良の早期発見といった業務に活かせないかという課題に取り組むようになった⁵⁾。それまでは、テキスト情報はBIユーザにとっては参考情報または情報検索の対象項目にはなっても、データマイニングにおける分析の対象項目には含まれず、死蔵された情報といってよかった。また、当時商用化されていたテキストマイニング技術では、文書やフリーテキストを、そこに含まれる単語とその出現頻度や出現文書数に基づく重みの対からなるベクトルで表現し、文書概念検索、分類、クラスタリングなどを行うものが主体であった。このため、テキスト中の構文構造を利用した分析やFAQ候補を効率的に抽出するような作業は（仮に可能であったとしても）相当な人手を要するものであった。

問題の本質は、テキストには種々の粒度の情報が含まれており、BIの用途によってテキストから抽出する情報や分析手法を変えるとともに、実世界の概念的な体系や分類によって情報を整理しないといけないという点であった。表-1にテキストから抽出可能な情報の粒度と

その典型的な用途を示す。情報の粒度は上から下に向かって細かく（より詳細性が高く）なっており、一般に荒い粒度の情報ほど潜在的な異なり数（個別の要素数）が少なくなるため、それを含む文書数は増加する。したがって、ある単語を含む文書を検索し、その結果を分析すると一般的な傾向は分かるかもしれないが、（検索結果には多くの文書が含まれるため）そこからFAQを効率的に作成するためには、疑問文や、疑問文に高頻度で現れる係り受けの情報（「資料を入手できるか？」など）が必要となる。

これらの異なる粒度の情報をすべて抽出し、構造化情報としてデータベースで表現することは可能であるが、オントロジーと呼ばれる概念的な体系、同義語や上位・下位語の関係、肯定・否定表現など、自然言語特有の知識や構造が扱えなければならない。このような情報は、テキスト中に表現されている情報とは別に、情報の検索や分析時に動的に必要とされることが多い。たとえば、「アドバイス」と「助言」はほぼ常に同義語として、パターンの抽出や傾向の分析においてひとまとまりの表現として扱いたい。場合によってはこれに「意見」や「コメント」といった表現もまとめて要件件数を集計したり、「<参考意見>を＋ください」といった頻出パターンとして集約することがある。このように意味的・階層的な情報を伴った分析がテキストマイニングの大きな特徴である。

テキストマイニングをBIに利用することで、これまでの数値的な分析に、その文脈や背景を記述したテキスト情報をコンテンツとして取り込めるとともに、以下のような新しい応用が可能になった。

FAQ作成とアラート発信

顧客関係管理（CRM）では、コールセンタ業務の改善や効率化が大きな課題の1つである。新製品の発売や初期不良などに起因して新たな質問が急増する現象がよく知られており、問合せ窓口でこのような質問を新たなFAQとして管理することで、窓口での適切な回答を実現するとともに、企業の運営するWebサイトでFAQを掲載することで電話問合せの削減を図ることができる。FAQの頻度の計算は単純ではない。日本語の場合には、同義語以外にも、語順の自由度や助詞・丁寧表現などの多様性によって、同じ内容の質問が多くの異なる表層表現に対応するため、文単位の頻度でそのままFAQの判定をすると個別の文を計数するのみで精度がよくない。逆に、文に含まれるキーワード単位の頻度は質問と関係ないキーワードを過剰に計数しているため、やはり精度が悪い。このため、データマイニングの相関規則の計算

情報の粒度	情報抽出の手段	典型的な用途
文字（列）	n-gram(隣接n文字)の抽出など	全文検索
単語	形態素解析	キーワード検索、概念検索、クラスタリング、文書カテゴリ付与
固有表現（人名、住所、組織名など）	固有表現抽出	個人情報抽出
係り受け（主語＋述語など）	構文解析	FAQ作成、評判分析
文	文分割	要約（重要文抽出）
パラグラフやテキスト全体	XMLやHTMLに含まれる文書構造の抽出	原文表示

表-1 テキストに含まれる情報の粒度とその主な用途

手法を文節の係り受け構造に適用し、各述語とその修飾句から頻出パターンを求める手法が提案されている。このようにして求められた係り受けパターンは、たとえば、「サイト(で)」+「Wiki(を)」+「立ち上げられるか?」といった可読性をそなえつつ、表現の多様性をかなり吸収しており、FAQ候補の頻出パターン計算手法として有望である。

アラート発信は、緊急の問題発生や問合せの急増などを検知し担当者に通知する機能である。前述のようなFAQ作成の手法と同様に、テキストに含まれる問題表現や質問表現を抽出し、新規の問題や一定の期間に出現頻度が通常の範囲を超えるものを通知する。個別の表現の出現頻度の変化をもとにアラートを発生させるだけでなく、製品群中の特定の製品に関する問題のような突出した相関関係が検知された場合にも対応できる。従来の固定された問題分類コードや製品名といった構造化情報では、新規の問題に対して適切な分類コードがなかったり、「その他」という分類コードに未整理問題がすべて含まれるといった状況がよく見かけられたが、テキスト情報を併用することで、このようなケースでも問題記述の特定により対応できるようになった。

情報抽出

テキストマイニングの大きな機能の1つは、テキスト中に出現する特定の種類の情報を抽出することである。キーワード検索用に索引を生成するための単語抽出(これに、表記のゆれの統一や、不要語の除去なども含める)は典型的な情報抽出である。このうち、固有表現(named entity)と呼ばれる、人名、地名、組織名、日付といった意味的に分類された表現を抽出する機能が特に重要である。人にかかわる固有表現抽出は、最近の個人情報保護や後述するようなマスタデータ管理においてよく使われる技術である。人名や地名に関する辞書を利用する手法や、「～様」といった敬称との共起パターンから人名などを推定するルール・ベース手法、すでに人名をタグ付けしたコーパスから機械学習によってタグ付けに必要なパラメータを獲得する手法などが知られている。固有表現抽出は、さらに質問応答(question answering)という応用にも使われている。質問応答では、たとえば「モーツアルトの生誕年はいつか?」といった質問に対して、テキスト中から回答にあたる部分を抽出して提示する。従来のテキスト検索手法であれば、「モーツアルト」および「生誕年」という単語を含むテキストを検索するだけであるが、質問応答では、「モーツアルトは1756年にザルツブルグで生まれた」といった表現か

ら、1756年という日付に関する固有表現を認識し、生誕年に合致する可能性があるものとして回答候補にする。質問応答技術は、大量のテキスト情報を知識源とし、個別の事実や関係を検索する手段として有望な技術である。このようなレベルでの情報抽出には、構文的なパターンもよく用いられる。上記の例であれば、「<人>は<日付>に<場所>で生まれた」というパターンによる情報抽出を行えば、生誕年や出身に関する質問に回答することができる。

生命科学の分野では、生医学文献抄録あるいは全文に現れるタンパク質、遺伝子、疾患名といった固有表現の抽出を利用して、相互関係や代謝にかかわる知識の獲得が精力的に研究されている。このような情報は、たとえば創薬の支援や疾患関連遺伝子の探索などに利用されている。この分野では、MEDLINEという約1,300万件の生医学文献抄録や、MeSH(Medical Subject Heading)という文献の主題を表す約23,000語の記述子などが米国立医学図書館から公開されており、研究目的であれば無料で使用ライセンスを取得できる。このようにコーパスやオントロジーの事実上の標準化ができていることもあり、言語資源やマイニング手法の相互運用性が高いことが大きな特徴になっている。図-1に、我々が開発したテキストマイニング・システムの画面例を示す。ここでは1992年から2004年までのMEDLINE文献抄録約500万件から、“leukemia(白血病)”という単語を含む約2万件の文献抄録を検索し、テキスト中に現れるタンパク質名(縦軸)と疾患名(横軸)の相関関係を分析している。図中で濃く表示されたセルほど相関関係の強いものであり、2つの遺伝子GCNT2とCD4が、ともに移植片対宿主病(GVHD)や主要組織適合性複合体(MHC)といった疾患や免疫の関連用語に対して類似した相関の強さを示していることが分かる。画面左側には分析のためのオントロジー選択メニューが表示されており、MeSHを含む多数の体系やテキストに現れた固有表現の種類などが分析軸として選択できる。

メールやブログから人と人との関係を中心としたコミュニティの抽出や分析を行う研究も多数報告されるようになってきた。人をノード、情報の発信・受信の関係や友人関係をリンクとしてグラフ化した構造のうえで、コミュニティや情報伝達の特徴を分析する。メールやブログに含まれるテキスト情報を併用すると、話題別の人間関係分析や本文に現れる専門用語などをもとに専門家検索(expertise locator)といった応用が可能になる。

感情分析・評判分析

感情分析(sentiment analysisまたはaffect analysis)は、

Database: MEDLINE_500M
Number of documents: 5466968

KEYWORD SEARCH: leukemia 20726 docs

Search Criteria Searched documents: 20726 documents
KEYWORD: leukemia 20726 docs DELETE

Document count: 文献抄録の絞込みのための検索条件 (leukemiaを含む20,726件を選択)

Category Tree: SORT Name, Vertical Category: GeneSymbol(LocusLink), Horizontal Category: Disease(List)

Vertical category: GeneSymbol(LocusLink)
List: subcategories keywords
Sort: Frequency Alphabet
Max lines to display: 10 20 50 100

Horizontal category: Disease(List)
List: subcategories keywords
Sort: Frequency Alphabet
Max lines to display: 10 20 50 100

Subcategories/keywords	leukemia 20726	disease 3512	cancer 1678	acute lymphoblastic leukemia 1636	disorder 796	GVHD 515	chronic myeloid leukemia 530	syndrome 511	CR 453	BREAST CANCER 426	acute promyelocytic leukemia 362	conditioning 334	multiple myeloma 317	neutropenia 272	APL 271	LUNG CANCER 239	MHC 235
t	2972	530	85	352	118	81	95	64	58	23	85	38	25	17	78	7	56
LIPC	1038	35	50	18	3	1	8	1	3	31	41	2	3	1	17	12	4
GCNT2	870	194	73	70	20	123	30	19	45	22	7	37	17	19	6	8	79
p	638	161	69	90	24	32	14	11	49	12	8	20	11	16	7	13	9
CD34	606	119	15	72	30	51	38	20	1	1	1	1	1	1	1	1	4
BCR	583	115	15	104	23	11	12	12	1	1	1	1	1	1	1	1	3
RA	565	95	28	10	26	2	8	10	0.8	0.1	15.0	0.2	0.2	0.9	16.9	0.1	3
CD4	564	112	10	34	27	58	15	13	9	3	3	9	8	6	2	1	49

分析のための項目 (MeSHなどのオントロジー、遺伝子名、タンパク質名などの固有表現、係り受けといった情報を選択可能)

遺伝子名(縦軸)と疾患名(横軸)の相関関係の分析。GCNT2とCD4がともに移植片対宿主病(GVHD)や、主要組織適合性複合体(MHC)に対して類似した相関の強さを示している。

図-1 生医学文献抄録に対するテキストマイニング

発話やテキストにおいて話者（著者）の意見や感情を分析する手法である。これまでBIでは客観的な情報の分析を主に扱ってきたが、テキストに記述された対象や事実に関する主観的な態度に基づいて、それらの対象や事実を分類あるいは重み付けすることが試みられるようになった。多くの人間の支持を受ける意見や、ある事実に対する強い疑念といった情報は、Webページや特にブログのような個人の情報発信が大規模に行われる環境において重要な意味を持つと考えられ、新しい知見の獲得手法として注目を集めている。

評判分析 (reputation analysis) は、そのような感情表現のなかでも、特に対象についての好評と不評に関する表現を抽出し、その対象がどう評価されているかを判定する手法である。インターネットで入手可能な口コミ情報によって消費者の購買行動が影響を受ける昨今の状況にあっては、評判分析はマーケティング活動において

欠かせない手段になりつつある。筆者の所属する東京基礎研究所でも、社内外の口コミ情報やアンケート回答などのさまざまな情報とその分析業務・改善活動などに対して、評判分析の有効性を確認している⁶⁾。

評判分析の技術的な課題と面白さについて若干補足すると、以下のようにまとめることができる。

- (1) 「～が好きだ」、「～が悪い」といった直接的で、特定の分野にあまり依存しない好評・不評表現が存在し、基本的な評判分析の知識として利用できる。さらに、肯定形的好评・不評表現を全否定したものは一般に評価が反転する。
- (2) 好評・不評表現を含む構文のレベルにおいて、好評・不評という分類が変化することがある。たとえば、「変更できない」というのは一般に不評表現であっても、「管理者しか変更できない」という文脈であれば不評表現とはいえないことがある。同様に、「A

よりも B がよい」という表現は、B については好評表現であるが、A については不評表現といえることがある。ほかにも慣用句的な表現や、条件節、「悪いとは思わない」といった述語の入れ子などによって好評・不評表現が変化することが分かっている。

- (3) 分野に特有な好評・不評表現といえるものが存在する。映画評論では「泣ける」というのは好評表現であり、「先が読める」のは不評表現である。洗濯に関する記述で、「色がおちる」というのは不評表現であるが、「汚れがおちる」というのは好評表現である。このような表現の分野依存性や多様性が、評判分析のための知識ベース構築を困難にしている。

最近の研究によって、テキスト中で文単位に評判表現が含まれているかどうかの自動判定や、少数の手がかりとなる好評・不評表現をもとに、コーパス中の文脈情報から新たな好評・不評表現の自動獲得を行う手法が報告されており、上記 (3) の問題も解決されつつある。このように、BI で取り扱う分析手法も、テキストマイニングの活用によって次第に意味処理を取り込めるようになりつつあるといえる。

評判分析技術は、平成 15 年度の IPA（情報処理推進機構）未踏ソフトウェア創造プロジェクトで、東京工業大学精密工学研究所奥村研究室によって開発された blogWatcher という、ブログの収集と分析を行うツールに含まれたことで身近なものとなった。最近では、インターネット検索エンジンで提供されるものも現れており、多様化する検索サービスを形成しつつある。

データベース・マーケティング

テキストマイニングの初期の応用は顧客からの問合せ分析が中心であったが、顧客に対して商品の紹介や勧誘を行うセールス・マーケティング活動にも有効であることが分かってきた。戸谷が指摘する⁷⁾ように、顧客情報を基幹系から情報系のデータベースに移動しただけでは、分析に必要な情報の有無や情報の鮮度・更新頻度に問題があることが多く、期待した効果が出にくいことがあった。ところが、顧客に対する電話等のアプローチで会話した内容を記録し、テキストマイニングによって分析することにより、顧客の好み、特定の商品に関する興味、購入予定といった情報をタイムリーに取得できるようになった。たとえば、「ハードディスク容量が足りない」、「BS デジタル・チューナー内蔵機種を重視する」、「ボーナス時期に購入したい」といった表現に、パターンを利用した情報抽出を適用することで、それまでの構造化情報からは獲得しにくかった情報が分析に利用できるようになった。このような情報抽出結果を特定の形

式で構造化することにより、従来の構造化情報に基づくデータマイニングやレポート作成のためのツールがそのまま適用できる。

また、商品への興味や購入予定に関する情報を販売促進のキャンペーン活動に連携させることで、どのような顧客にアプローチすればより効果があるかという因果関係を検証しつつ情報抽出パターンの修正を行い、情報抽出精度を漸次改善することができる。

BIの新たな展開

（データウェアハウスを含む）BI ツールの市場規模は 2006 年で約 25 億ドル（国内では約 468 億円）に達し、2009 年までに年率 7.3% で世界的に堅実な成長を遂げると予想されている。BI の成長を支えているのは、CRM や ERP（enterprise resource planning）などの企業内アプリケーションの普及によるデータの増大であり、このようなデータを全社的に業務に活用することが、企業にとっての最優先課題となっているためである。とりわけ、以下に示すマスタデータ管理や身元分析といった新しい技術が、これまでの BI の役割を大きく変える機動力として注目されている。

マスタデータ管理 (MDM)

データウェアハウスの構築においては、各業務部門で運営されているシステムの多様な情報源から ETL（Extract, Transfer, and Load）と呼ばれるプロセスによってデータを抽出・変換・ロードし、全社的なレベルの情報に統合している。これにより、部門に散在する情報をとりまとめ、より包括的で確度の高い意思決定を可能にしている。この際に問題となるのが、個別の情報源で採用されている項目名やコード体系の統合である。これまでの ETL ツールでは、大量のデータを処理する高いスループットの実現やリアルタイム性が重視されていたが、データの統合にかかる問題は十分に解決されていなかった。たとえば、日本では平成の市町村合併により、住所表記が大きく変わってきている。新・旧の住所表記を正しく対応づけてデータの標準化を行わないと、データの整合性が失われ、同一の顧客に新・旧の住所で 2 通のダイレクト・メールを発送するといった不都合が生じやすい。標準化されたデータは再利用性が高いため、データウェアハウスのみでなく、社内のさまざまなアプリケーションでの共有が望める。このような背景から提案されるようになったのが、マスタデータ管理 (MDM) という考え方である。MDM では、顧客、商品、社員といった、企業にとってきわめて重要な実体についての標準的

なデータの表現形式やコード化を定義し、全社的な情報の共通化・透明化を実現する。これは個人情報保護法など法令順守の観点からも重要であり、データに基づく活動や意思決定における矛盾や不整合を排除できる。ETL ツール上にこのようなデータ統合のロジックや処理フローを容易に開発できるように改良が行われ、MDM ツールとして商品化されている。

MDM は、スキーマ統合のようなデータベース的課題も含んでいるが、意味的な実体についてのデータを標準化するため、意味処理を扱う人工知能的な手法とも整合性が高い。前述の住所表記の例でも、所在地についての体系化や、「近傍」の概念など、空間データマイニングで扱われる手法が利用できる。また、テキストマイニングの手法により、テキストから個人情報を検出したり、名前+住所+職業といった、個人に関するまとまった情報の抽出によって MDM を支援できる。意味的な観点によるデータの統合 (semantic integration) は、最近では特集記事としても取り上げられており⁸⁾、オントロジーやデータの統合技術が紹介されている。

身元分析 (identity resolution)

BI における最近の人工知能的な分析の顕著な例として、身元分析をあげることができる。身元分析は Jeff Jonas (現在 IBM Distinguished Engineer) によって実用化された技術であり、偽名を使っている犯罪者を特定することで注目を集めた。身元分析では、個人がいくつかの名前を使っていることをさまざまな情報をもとに認識し、同一人物である可能性が高いものを特定する。さらに、関係分析 (relationship resolution) という手法を適

用することで、犯罪者とテロ組織との関係などを明らかにすることができる。これらの分析においては、データの匿名化によって個人情報に関するプライバシーを保護しており、情報漏洩などのリスクに対処している。

身元分析は、当初の応用の範囲を超え、金融機関における本人確認や CRM での顧客情報の整備に利用されつつある。身元分析を構成するのは、複数の情報源に含まれる実体の認識、同一性判定、匿名化、関係の認識、といった技術であり、このような処理の多くは人や実世界に関する知識や常識を利用しており、同様のことが商品、組織、住所といった実体にも適用可能であると考えられる。それまでは人の持つ知識に頼っていた業務が、このような分析によって高速化、大規模化しつつシステムに置き換えられることで、BI の日常的な業務における重要性を大いに高めることが期待されている。

参考文献

- 1) Inmon, W. H. (藤本康秀, 小畑喜一共監訳) : データウェアハウス 一構築編一, オーム社 (1997).
- 2) Codd, E. F., Codd, S. F. and Salley, C. T. : Providing OLAP to User-Analysts: An IT-Mandate, Arbor Software Corporation (1993). 現在以下の URL から入手可能 http://dev.hyperion.com/resource_library/white_papers/providing_olap_to_user_analysts.pdf
- 3) Fayyad, U. M. 他 : Advances in Knowledge Discovery and Data Mining, MIT Press (1996).
- 4) Westerm, P. (須藤晶子, 平田真理共訳) : ウォルマートに学ぶデータ・ウェアハウジング—流通業界“巨人”の躍進を支える情報基盤の全貌, 翔泳社 (2003).
- 5) 那須川哲哉 : コールセンターにおけるテキストマイニング, 人工知能学会誌, Vol.16, No.2, pp.219-225 (2001).
- 6) Nasukawa, T. and Yi, J. : Sentiment Analysis : Capturing Favorability using Natural Language Processing, Proc. 2nd Intl. Conf. on Knowledge Capture (K-CAP), pp.70-77 (2003).
- 7) 戸谷圭子 : 金融サービス・マーケティングにおけるデータマイニングの諸問題, 人工知能学会誌, Vol.19, No.5, pp.607-609 (2004).
- 8) Special Issue of Semantic Integration, AI Magazine, Vol.26, No.1 (2005).

(平成 18 年 6 月 5 日受付)