

ビジネス・インテリジェンス・システムにおける 情報要求の抽出手法

関根 純^{†1} 末永 高志^{†1} 矢野 順子^{†1}
中川 慶一郎^{†1} 山本 修一郎^{†1}

ビジネス・インテリジェンス (BI) が、売上げ増加やコスト削減などの企業目標の達成に役立つ情報を提供するコンセプトとして注目を集めている。本稿では、それを支援する BI システムが満たすべき情報要求を、実際のデータを用いてその有効性を検証しながら抽出する手法を提案する。まず、企業目標の達成に必要な仮説を抽出し、次にその仮説を実現する施策を明らかにする BI の業務手順を提案する。この手順で生成される情報が BI システムの情報要求となる。また、多様な BI システムをレベル分けする 2 つの観点を提案し、各レベルが業務手順のサブセットを支援することを示す。さらに BI システムでは、仮説や施策の網羅性を企業の経営層に納得させることが重要であり、そのためのフレームワークを与えるモデルを提案し、これが BI システムで扱う情報の整理にも活用できることを示す。実例により、本手法の適用性を示した。

A Method of Extracting Information Requirements for Business Intelligence Systems

JUN SEKINE,^{†1} TAKASHI SUENAGA,^{†1} JUNKO YANO,^{†1}
KEI-ICHIRO NAKAGAWA^{†1} and SHUICHIRO YAMAMOTO^{†1}

Business intelligence (BI) is becoming a key means of providing information necessary for achieving business goals such as improving profits or reducing business costs. This paper proposes a method of extracting information requirements for BI systems by validating the effects of the information. The validation processes include extraction and validation of hypotheses for achieving business goals and execution of actions needed to implement the hypotheses. Then, information created in each process is defined. Two viewpoints for classifying various BI systems are also proposed, and each class is mapped to a subset of the validation processes. Finally, a modeling method that gives a framework for clarifying and organizing the hypotheses, actions, and the information is

proposed. Studies of real cases have shown the validity of the method.

1. はじめに

売上げ増加やコスト削減などの企業目標の達成に必要な情報を提供する、ビジネス・インテリジェンス (BI) の概念が広まっており、これを支援する BI システムの構築もさかんになってきた。BI システム構築のための要件の抽出にあたっては、企業目標の達成に役立つ情報をいかに提供するかが重要であり、この情報要求の抽出と整理を体系的に行う手法が必要とされている。

この手法の検討にあたっては、BI システムに特徴的な以下の 3 点の考慮が必要である。まず BI システムでは、通常、企業の業務を支える基幹システムからの業務データを用いて計算処理を行い、企業目標の達成に役立つ情報を生成するが、それが実際に役立つかどうかは実データを用いて検証してみる必要がある。次に BI システムには、企業目標を数値化した KPI (Key Performance Indicator) や関連情報を OLAP (On-Line Analytical Processing)¹⁾ ツールを用いて可視化するレベルから、企業目標を達成する仮説の立案をデータマイニング技術を用いて支援するレベル、さらには仮説を実現する施策の実行管理を支援するレベルなど多様なものがあり、手法はこれらをカバーする必要がある。最後に、BI システムの構築を企業の意思決定者が納得するためには、企業目標を達成する仮説やそれを具体化する施策が合理的かつ網羅的に検討され、BI システムに投資すればその達成に役立つ情報が提供されることの説明が必要であり、これらの仮説や施策、そして BI システムが提供する情報の全体像を見せるフレームワークが必要である。これら 3 点を考慮した情報要求の抽出と整理の手法が必要とされている。

これまで、情報要求を抽出し整理する以下の手法が提案されている。データモデリング手法^{2),3)} は、データベース構造の設計だけではなく、情報要求の抽出と整理にも活用できるが、企業目標との関係を整理することはできない。バランススコアカード⁴⁾ は、企業目標を 4 つの観点で整理し、それを定量的数値目標である KPI に関係付ける手法であり、目標の達成度を表す結果系の KPI に加えて、目標の達成に必要な原因系の KPI を関連

^{†1} 株式会社 NTT データ
NTT DATA CORPORATION

付けて管理する．原因系の KPI を制御することにより，企業目標を達成することを想定している．たとえば，売上げが達成したい結果系 KPI，顧客あたりの単価が原因系 KPI である場合，顧客あたりの単価を上げることにより売上げの増加を狙うことが企業の戦略となる．このバランススコアカードを補完する手法である FBCM (Fact Based Collaboration Modeling)⁵⁾ は，課題に網羅性や現実性があるかどうかを現場観察により検証し，課題を解決する業務プロセスと情報システムの要件を体系的に整理する．この手法は，BI システムにも適用可能な汎用的な手法であるが，BI システムで重要な，情報要求の抽出と整理に特化した手法ではない．同様に，ゴール指向分析⁶⁾⁻⁸⁾ も企業目標とシステム要件の関係を整理する汎用的な手法であり，情報要求の抽出と整理に特化した手法ではない．データ分析フレームワーク⁹⁾ は，BI システムの中でもデータマイニングに関する分析手順を体系化したものであり，データ分析が達成する目標を 9 種類に分類しているが，情報の整理については言及していない．Sekine ら¹⁰⁾ は，BI のビジネスプロセスと様々なレベルの BI システムの対応付けを提案しているが，情報要求をどう抽出するかについては具体化していない．

本稿では，企業目標とその達成に役立つ情報を整理し，その情報が実際に役立つかどうかを実データを利用して検証し，検証が成功した場合には，その過程で得られる情報を BI システムの情報要求とする手法を提案する．そのためまず，企業目標の達成に必要な情報を洗い出す BI の業務手順を提案する．検証はその業務手順のサブセットを試行的に実施することにより行う．また BI システムには可視化のレベルから施策の実行管理のレベルまで多様なレベルの BI システムが存在することから，これらを分類する観点を提案し，各レベルが上記の BI の業務手順の一部を支援することを示す．最後に，仮説や施策を整理するフレームワークとなるモデルを提案する．

2 章では，BI の業務手順とそこから抽出される情報要求を提案するとともに，多様な BI システムが 2 つの観点で分類でき，これらが提案した業務手順の一部を支援することを示す．3 章では，仮説や施策を整理する 3 種類のモデルを提案する．4 章では，モデルの適用性を定量的に評価するとともに，筆者らの実施例によりどのような企業目標に対してどのような KPI が抽出され，どのようにモデルが構築されるか，また BI の業務手順のどの部分を支援するのかを示す．5 章で考察し，6 章でまとめる．

2. BI の業務手順と検証手順

本章では，BI の業務手順を提案する．BI では，生成される情報が企業目標の達成に向けて役立つ情報であることを事前に検証する必要があることから，提案した業務手順は BI シ

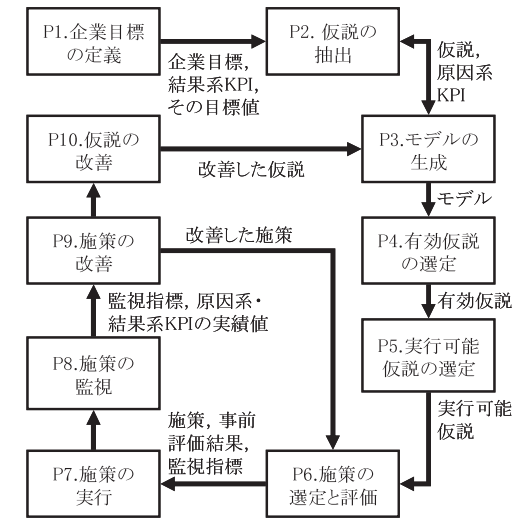


図 1 BI の業務手順と生成情報

Fig. 1 Business processes and information of BI.

ステムのシステム要件抽出のための検証手順ともなる．2.1 節では，まず BI の業務手順を提案し，検証手順との関係を示す．次に，2.2 節では，BI システムのレベルを分類し，それと BI の業務手順との関係を示す．

2.1 BI の業務手順

業務手順の概要は，次のようになる．BI では，1 章で述べたように，企業目標の達成に役立つ情報の提供が目的であることから，まずは企業目標そのものを整理し，その目標の達成度を定量的に表す結果系の KPI を洗い出す．次にその結果系 KPI の達成を可能とする仮説を抽出し，その仮説の達成を可能にする原因系の KPI を求める．さらに，この仮説から抽出された原因系の KPI を制御して結果系の KPI の目標を達成するためには，何らかの施策を実行する必要があることから，その施策を洗い出し実行する．その施策の達成に時間がかかり，達成状況の管理が必要になる場合には，その実行状況を表す監視指標を設定し，それを用いて監視する．施策を実行しても企業目標が達成されないなら，仮説や施策を見直す．

以下では，BI の業務手順を構成する 1 つ 1 つの作業をプロセスと呼ぶことにし，全体で 10 個からなるプロセスを提案する．各プロセスとそこで抽出される情報要求の全体像を図 1 に示す．

(P1) 企業目標の定義

企業目標を定義し、その目標の達成度を定量化する 1 つ以上の結果系の KPI を選定し、その目標値を決定する。このためにバランススコアカードや FBCM などの手法を活用する。たとえば、顧客満足度の向上が企業目標である場合、目標の達成度を表す結果系 KPI は、顧客満足度である。結果系 KPI の測定方法や、具体的な目標値の設定方法については、それぞれの企業の状況に応じて異なるので、個別に検討する必要がある。たとえば顧客満足度であれば、顧客へのアンケートの各項目の平均点により定量化し、目標値は 10 段階評価で 9 点以上というように設定できる。

BI システムの構築にあたって、必ずしも本プロセスを経ず OALP ツールにより企業全体の状況のレポート機能を実現する場合があるが、レポートの目標が明確でないと BI システムの投資効果を示すことができないことから、本手法では、本プロセスと次の「仮説の抽出」プロセスを設けることにした。

(P2) 仮説の抽出

企業目標の達成に必要な仮説を抽出する。仮説とは、結果系 KPI と、それを向上するために効果が高いと BI の実施者が認識する要因を書き下したものであり、その要因には原因系 KPI が 1 つ以上含まれる必要がある。原因系 KPI の制御により、結果系 KPI の向上が達成できることを想定する。たとえば、企業目標である顧客満足度の向上に対して、「顧客満足度を上げるためには社員満足度を上げる必要がある」が仮説となり、結果系 KPI は顧客満足度、原因系 KPI は社員満足度となる。仮説は、この例のように、原因系 KPI をどう達成すべきかまで言及するものではなく、これは後に述べる施策の選定と評価のプロセスで具体化される。別の例でいうと、「商品の売上げの改善のためには、市場で強みのある商品 A に重点化すべきだ」も 1 つの仮説であり、結果系 KPI が全体の売上げであるのに対し、原因系の KPI は商品 A の売上げとなる。

なお、次のプロセスであるモデルの生成と関連するが、仮説の抽出は、過去の筆者らの分析経験から、2 つの方法に整理した。1 つの方法では、企業目標を達成するための課題を明確にした後、関係者に対するインタビューなどを通じてその課題を解決する仮説を仮決めする。そして、これだけでは仮説の網羅性が担保できないことから、仮説の全体像を整理するフレームワークを次の「モデルの生成」プロセスで作成する。フレームワークに照らして不足があれば仮説を追加する。このフレームワークを、本稿では「仮説構造化モデル」と呼ぶことにする。一方もう 1 つの方法では、仮説構造化モデルを先に生成し、そのモデルを構成する個々の要素ごとに課題を解決する仮説が考えられるかどうかを網羅的に検討し

抽出する。これら 2 つの方法をカバーするため、図 1 においては「仮説の抽出」プロセスと、「モデルの生成」プロセスを、相互に連携するプロセスとした。いずれの方法においても、仮説を思いつくためには、課題を熟知した企業の業務担当者と BI 技術の専門家の連携が必要であると考えられる。

(P3) モデルの生成

本プロセスでは、仮説、あるいは仮説を具体的に実現する施策を整理するための仮説構造化モデルを、手動で、またはデータマイニング技術¹¹⁾などを用いて自動で生成する。モデルの詳細については 3 章で述べることにし、ここでは一例を示す。

たとえば、顧客宅への訪問サービスを行う企業において、サービス時間を短縮する企業目標が与えられたとする。このためにまず、サービス員が営業店からサービスエリアに移動し、そのエリア内の顧客宅を巡回し、さらには担当している複数のサービスエリアを移動するという業務の流れを整理する。次に各業務の所要時間を計測することにより、どの業務に無駄があるかが分かり、業務ごとに短縮のための仮説を考えることができる。この例では、業務の流れが仮説構造化モデルとなり、これに施策を対応付けて網羅性をチェックすることができる。この場合、サービス時間が結果系 KPI、施策を打つ各業務の所要時間が原因系 KPI となる。

(P4) 有効仮説の選定

前のプロセスで抽出された仮説は、それだけでは事実かどうかは分からない。本プロセスでは、これを検証し、有効な仮説を絞り込む。検証は、通常、統計的な手法やシミュレーションを用いて行う。たとえば、「顧客満足度の向上に社員満足度の向上が役立つ」という仮説が相関分析を通じて検証されることになる。

なお、(P3) や (P4) のプロセスの実施にあたって、検証に必要な業務データが揃わず原因系 KPI が算出できないデータの入手性の問題、業務データは入手できても意味ある原因系 KPI を生成できない導出可能性の問題、および生成された原因系 KPI の品質^{12),13)}の問題が起きる場合がある。このような場合には、仮説に含まれる原因系 KPI を代替する他の原因系 KPI を検討することになる。これが難しい場合には、当該仮説に基づく施策の実行を諦めることになる。

ここで導出可能性とは、本稿で提案する概念であり、入力となる業務データから、期待する出力が計算処理により導出できることを表す。良いアルゴリズムがない場合もあれば、複数のソースからの業務データをうまく統合して処理できない場合もある。後者の例でいえば、業務データの ID や名前に不整合があり突合せできないことが原因であり、この場合に

は名寄せなどの新たな技術が必要になる。

(P5) 実行可能仮説の選定

前プロセスで抽出された仮説は、いくつかの理由により有効ではあっても実行できない可能性がある。本プロセスでは、実行可能な仮説を選定する。実行できない理由としては、その企業の置かれた状況により実施できない、あるいは原因系の KPI を制御できないことがあげられる。たとえば、実施できない例であげられるのは、「顧客満足度の向上のためには、商品の単価を下げる必要がある」という単純な仮説が抽出できたとしても、企業の財務状況から簡単に単価を下げることはできない場合である。また原因系 KPI を制御できない例であげられるのは、「店舗の売上げ向上のためには、駅からの距離の短縮が必要」という仮説が抽出できたとしても、原因系 KPI である駅からの距離は、店舗そのものを物理的に移動しない限り既存の店舗の売上げ向上には役立たないという場合である。このような仮説は、実行可能と見ることはできない。

(P6) 施策の選定と評価

実行可能な仮説を選定したら、それぞれの仮説に対してそれを改善する具体的な施策を選定し、可能ならその施策により改善できるかどうかを事前評価する。施策実行の効果を事前評価するのは、効果が分からないまま実行すると、業務の現場を混乱させるからである。事前評価のために、業務プロセスのシミュレーション手法が利用されることもある。

また施策の実施に時間がかかる場合、施策の実施の程度を表す監視指標を明らかにする。たとえば、原因系 KPI である社員満足度を向上させるための施策として、残業時間の削減を施策として考えたとしても、その企業の置かれた状況から実現には時間がかかるかもしれない、平均残業時間が監視指標となる。

(P7) 施策の実行

選定し事前評価した施策を実行する。施策の実行に BI の持つプロアクティブな機能が活用される場合もある。たとえば顧客あたりの平均単価を上げることが原因系 KPI であるとする、顧客を層別し、特定顧客が所属する層でよく買われている商品を宣伝するリコメンデーションの機能は、それを実現する施策と見ることができる。書籍販売では、Amazon^{14),15)} のリコメンデーション機能がよく知られている。また、金融機関である Capital One¹⁵⁾ では、コールセンタでの受付を、顧客にパーソナライズしたサービス宣伝の場に変えて業績を上げている。

(P8) 施策の監視

施策の実施状況を監視指標を用いて監視し、それが部分的にでも原因系 KPI および結果

系 KPI の実績値の向上につながっているかどうかを確認する。

(P9) 施策の改善

施策の実施が原因系 KPI および結果系 KPI の向上につながっていない場合には、施策を見直す。

(P10) 仮説の改善

それでも改善されない場合には、仮説そのものを見直す。

以上、10 個のプロセスは改善のサイクルを構成し、これを用いて企業目標の達成に向けた施策の改善を繰り返すことになる。また情報要求は、結果系 KPI、原因系 KPI、監視指標および、これらを生成するのに必要な業務データからなり、これらを体系化するために仮説構造化モデル、企業目標、仮説、施策が存在する。

なお、提案した BI の業務手順は、BI システムを用いて BI の業務を実施する場合のフルセットの手順であるが、これをすべて検証の段階で実施する必要はなく、BI システム構築の効果が説明できる範囲でかまわない。したがって、検証手順は業務手順のサブセットとなりうる。どこまで検証するかは企業の状況によって異なると思われる。たとえば、小規模でも試行的に施策を実行してみる場合もあれば、施策の選定とシミュレーションによる事前評価にとどめる場合もある。検証のためには、少なくとも「施策の選定と評価」プロセスまでは実施する必要がある。

2.2 BI システムのレベルと業務手順の関係

BI システムには、様々なレベルがあると考えられるが、本節ではそれを分類するとともに、それがどのように BI の業務手順を支援するかを示す。多様な BI システムが存在する理由は、BI システムが BI の全業務を支援するためには、企業が BI そのものの効果を納得し情報に基づく企業経営の考え方を浸透させている必要があり、そこまで効果を感じていない企業では、まずは企業の経営状態の可視化からスタートすることが多いためである。このような観点で見ると、BI システムを分類する軸の 1 つは、BI システムの支援範囲が可視化レベルなのか施策の実施レベルなのかになると考えられる。もう 1 つの軸は、BI システムの提供する機能であり、単に情報を管理するレベルなのか、データマイニングなどの BI の高度な技術を使って役立つ情報を生成するレベルなのかである。これらを組み合わせると BI システムは 2 軸の 4 レベルに分類できる(表 1)。それぞれのレベルを以下で説明する。
レベル 1: KPI の可視化

企業の経営状況を可視化するなど、結果系 KPI、原因系 KPI、およびそれに関係する業務データを OLAP ツールなどを用いて可視化して見せる。BI のプロセスでいうと (P1) ~

表 1 BI システムのレベル
Table 1 BI system levels.

機能 支援 の範囲	情報の管理	情報の生成
可視化	レベル1 KPIの可視化	レベル2 データ マイニング
施策の実施	レベル3 施策の管理と 評価	レベル4 サービス実現

(P4) が対応する。

なお、このレベルにとどまる場合には、企業の経営状況など企業を取り巻く状況の数値的な理解のみが可能である。この結果、仮説も必ずしも原因系 KPI を明確にするものではなくなる可能性がある。たとえば、「サービス A の利用者は、20 代男性の若者であり、時間的な制約がないことが評価されているポイントである」といった事実認識に関する仮説となり、サービス売上げ増加のための顧客理解には使えるが、直接的に原因系 KPI を指定するものではなくなる。

レベル 2：データマイニング

カバーする業務プロセスは、レベル 1 と同様であるが、データマイニング技術を用いて、仮説やモデルを生成したり検証したりする。

レベル 3：施策の管理と評価

企業が実施する施策を管理するなど、原因系の KPI とそれを改善する施策の監視指標を可視化する。また、シミュレーション技術などを活用して、施策が確かに効果を上げることが事前に検証する機能を提供する場合もある。BI のプロセスの (P5) ~ (P9) が対応する。

レベル 4：サービス実現

上記のレベル 1 からレベル 3 とは異なり、本レベルでは施策そのものを実現する。企業の業務を改善するというよりは、社内あるいは社外のユーザに対して企業を根本的に変革するような高度なサービスを、最適化やリコメンデーションなどの BI の技術を用いて実現する。前節の (P7) であげたりリコメンデーションの事例がこれに対応する。BI のプロセスでいうと (P7) ~ (P9) が対応する。

3. 仮説構造化モデル

本章では、仮説構造化モデルの考え方、およびそのモデル化の手法として 3 種類を提案する。まず 3.1 節では、モデル化の考え方を提案する。次に 3.2 節から 3.4 節で、それぞれのモデル化手法の考え方と考案するに至った理由、仮説や情報要求の整理にどのように使われるのか、およびどのように仮説構造化モデルを抽出するかを提案する。

3.1 モデル化の考え方

本稿で提案する仮説構造化モデルは、企業目標を改善する仮説を対応付けるものであることから、企業目標を取り巻く業務や業務対象などの要素とそれらの関係を構造化して見せるものでなければならない。そしてこれらの 1 つまたは複数の要素に仮説が対応付けられ、位置づけが理解できるものである必要がある。一方仮説は、1 つの結果系 KPI と 1 つ以上の原因系 KPI からなるから、結果として、仮説構造化モデルは原因系 KPI、結果系 KPI を 1 つ以上の要素にマッピングできる。また、仮説に含まれる原因系 KPI の向上を実現するために施策が存在することから、施策と施策の達成度を表す監視指標もマッピングできる。原因系、結果系 KPI および監視指標は、業務データから計算処理により導出される。これらの関係を図 2 に示す。ここで問題となるのは、仮説構造化モデルを構成する要素として何を選択するべきかであり、それによりモデル化の手法も複数存在すると考える。次節以降では、これを提案する。

3.2 タイプ 1：業務対象のカテゴリ分け

本タイプでは、企業目標の達成の対象である業務対象とそれを分類したカテゴリをモデル化する。業務対象をカテゴリ分けすることにより、注力すべきカテゴリを明確にし、そのカテゴリに属する業務対象に対して特化した仮説や施策を検討することが可能になる。この結果、この仮説に関係する原因系 KPI もこの特定のカテゴリに対応付けられる。

このタイプは、次の 2 つの考え方に基づき導出した。まず、情報の管理を主とするレベル 1 とレベル 3 の BI システムは、一般に OLAP ツールを用いて実装されることが多いが、その際には、企業全体の売上げやコストなどの結果系の KPI と、それに影響を与える地域別、品目別、支店別などの観点で集計した売上げやコストなどの原因系の KPI を対比させて表示し、特定の地域、品目、支店の売上げやコストが全体に影響を与えているかどうかを調べ、問題があればそれに対して仮説を立て施策を打つ。ここで、地域、品目、支店は、仮説やそれに対応する原因系 KPI を分類するカテゴリと見ることができると考えた。またレベル 2 の BI システムでは、売上げ向上のため、たとえばデータマイニング技術の 1 つで

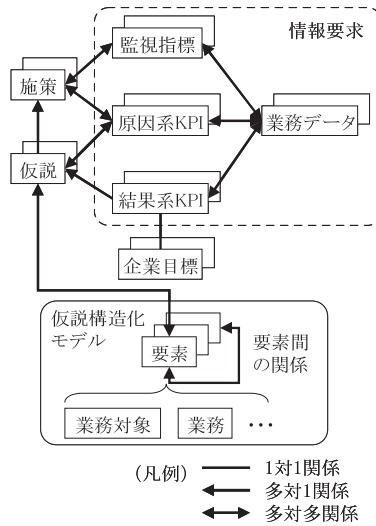


図2 仮説構造化モデルの構成
Fig. 2 Hypotheses structuring model.

あるクラスタリングを用いて顧客を分類し、その特定の分類に絞って売上げ向上のための仮説を立て施策を打つことがある。この場合、各分類はカテゴリに該当することから、本タイプが利用できると考えた。このようにレベル1から3のBIシステムで適用できるモデル化手法と考える。

本タイプにおけるカテゴリは、上記のレベル1やレベル3では、業務対象に付随した品目や支店などの特徴量からBIの実施者の選択により抽出され、レベル2では、クラスタリングなどのデータマイニング技術により抽出される。

このカテゴリの粒度は、施策を打ったときの効果の波及範囲と施策の検討にかかるコストにより適度な数となるように選択する必要があると考える。実施者が特徴量を選択する場合は、特徴量を選んだ時点で粒度は決まってしまうが、細かすぎると施策の検討をしきれなくなるため大分類を追加することが必要となる。たとえば、品目であれば商品カテゴリで大分類を行うし、支店であれば、支店を管轄する地域で大分類を行い、その単位で施策を打つことになる。一方、データマイニング技術を用いる場合には、クラスタリングした結果にBIの実施者が人間に理解できる意味づけをする必要があり、クラスタ数が多すぎるとそれが現

実にどのような特徴的な集団を表しているのかを説明できなくなるという問題から、10個以内には収めるようにするのが現実的と考える。

なお、カテゴリ分けの考え方をさらに発展させたものとして、業務対象が属するカテゴリが時間とともに変わる、カテゴリ間での状態遷移のモデル化がある。この場合には、より望ましいカテゴリに業務対象のカテゴリが変わるような仮説や施策を検討することになり、仮説や施策は遷移前と遷移後の2つのカテゴリの組合せで分類することになる。4章では、その事例を示す。

3.3 タイプ2：業務対象間の関係付け

本タイプでは、企業目標を達成する対象である複数の業務対象とその関係をモデル化する。これにより仮説を立案し施策を打つ業務対象の明確化が可能になる。

このタイプは次の2つの考え方に基いて導出した。まず3.1節に述べたように、仮説を分類するには、その仮説を取り巻く業務対象などの要素とそれらの関係を構造化して見せる必要があると考えた。またレベル2のBIシステムでは、重回帰分析などのデータマイニング技術を用いて、原因系KPIと結果系KPIの因果関係を調べる場合があることから、原因系KPIに対応する業務対象と、結果系KPIに対応する業務対象、およびそれらの間の関係をモデル化する本タイプにより、仮説や、それを構成する原因系・結果系KPIを整理できると考えた。

たとえば、営業担当者あたりの売上げ増が企業目標であるとする、そこに関係する業務対象としては、「顧客」、「営業担当者」の2種類があり、それらに「広告」や「訪問」といった関係がある。このような業務対象とその関係が仮説構造化モデルとして明確になると、売上げを上げるための仮説として、「広告」に着目した、ダイレクトメールなどの広告効果の向上や、「訪問」に着目した、営業担当者による顧客への訪問時の顧客満足度の向上、たとえば訪問回数増加や商品説明の充実、などを検討し分類することが可能になる。

またたとえば、重回帰分析などのデータマイニング技術を用いることにより、上記の結果系KPIである営業担当者あたりの売上げ増に役立つ原因系KPIが、営業担当者の商品説明力と訪問回数であることが導出できたとすると、これから逆に、「顧客」、「営業担当者」などの業務対象や、「訪問」、「商品説明」などの関係を表すモデルを生成することが可能になる。

上記の事例から分かるように、本タイプで必要とする業務対象やその関係は、重回帰分析などのデータマイニング技術により抽出される場合もあるし、BI実施者により手動で抽出される場合もある。

3.4 タイプ3：業務プロセスの記述

本タイプでは、企業目標を達成する業務対象に関する業務プロセスをモデル化する。企業目標は、企業活動のパフォーマンスを改善することが普通であることから、その企業活動を業務プロセスに分解すれば、改善のための仮説や施策を業務プロセスにより分類することが可能になると考えた。業務プロセスによる分類は、タイプ1のモデルである「業務対象のカテゴリ分け」と類似しているが、業務対象の特徴を用いたカテゴリ分けではないため、別タイプとした。

なお、BIの技術、たとえば予測、シミュレーション、あるいは最適化の技術などを使って業務を変革するようなサービスを実現する、レベル4のBIシステムの場合には、業務プロセスを詳細に分析して仮説や施策を検討する場合がある。4章に事例を示す。

本タイプは、業務プロセスモデリングを実施することにより抽出される。

4. 実例での検証

本章では、本手法の手順に沿って筆者らが実施した4つの実例を説明するとともに、3章で提案した仮説構造化モデルがどの程度、現実の分析事例をカバーできるのかを検証する。

4.1 適用事例

本節では、本手法に基づいて筆者らが実施した4つの事例を示す。その中で企業目標に対する、仮説、KPI、BIシステムのレベル、仮説構造化モデルのタイプを示す。表2では、4つの事例について、BIシステムのレベル、仮説構造化モデルのタイプをまとめる。各事例について、どのプロセスでどの業務が実施され、情報が導き出せたかを括弧内に示した。

事例1：プロジェクトの利益向上

本事例では、プロジェクト制の企業において、各プロジェクトの利益の向上が企業目標として設定され、結果系KPIはプロジェクト利益であった(P1)。これを達成する仮説として、「社員満足度が向上すれば利益が上がる」、「要員の必要に応じた流動性確保がコストを下げる」などがプロジェクトマネージャからのヒアリングにより抽出された(P2)。これらの体系付けのため、筆者らは図3に示すタイプ2の仮説構造化モデルを作成した(P3)。これらの仮説の検証のため、各プロジェクトのコストと利益、および所属する社員の社員満足度の平均値を収集し、因果関係について分析を行った結果、最初の仮説は正しいことが確認され、社員満足度が原因系KPIであることが確認できた(P4)。一方、2つ目の仮説は、要員の流動性に関するデータが入手できず確認ができなかった(P4)。このような状況から施策

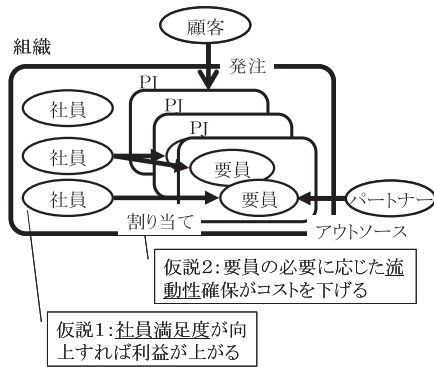
表2 事例のBIシステムレベルとモデルタイプ
Table 2 BI system levels and model types.

事例番号		1	2	3	4
BIシステムのレベル	1) KPIの可視化	○	-	○	○
	2) データマイニング	-	○	-	-
	3) 施策の管理と評価	-	-	-	○
	4) サービス実現	-	-	○	-
仮説構造化モデル	1) カテゴリ分け	-	○	-	-
	2) 業務対象間の関係付け	○	-	-	-
	3) 業務プロセスの記述	-	-	○	○

の検討はなされたが実施には至らず、BIシステムのレベルは、「KPIの可視化」にとどまった。なお、本事例の場合には、組織単位はあっても、プロジェクト単位の業務データがタイムリーに収集できない、組織によって業務データの品質が異なるなどの問題が顕在化し原因系KPIの抽出が一部困難であった。

事例2：医療費削減

本事例では、健康保険組合での医療費削減が目標であり、医療費が結果系KPIであった(P1)。これを達成するためデータマイニングにより、図4に示す疾病間の状態遷移を表すタイプ1の仮説構造化モデルが作成された(P3)。状態遷移図では、疾病ごとの平均医療費と次の疾病への移行確率も算出されており、これを元に「医療費が高む疾病になる前の状態で次への移行を阻止すれば、少ない施策費で医療費削減ができる」という仮説が立案された(P2)。この仮説に基づき効果の高い疾病を特定してシミュレーションを行い具体的な改善効果を算出した(P4)。医療費が高む疾病の直前の疾病の患者数が原因系KPIとなる。元となる業務データとしては、疾病名、処置内容、医療費が含まれるレセプトデータを用いた。施策の実施はBIシステムの範囲外となったためBIシステムのレベルとしては、「データマイニング」レベルにとどまった。



(注) 結果系KPI:PJ利益
原因系KPI:社員満足度, 流動性

図3 プロジェクトの対象

Fig. 3 Business objects of projects.

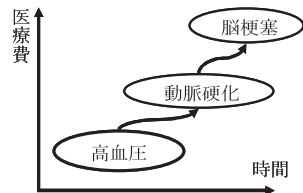


図4 疾病間の状態遷移

Fig. 4 State transition diagram of diseases.

事例3：在庫の削減

本事例では、ある小売企業における商品在庫の削減が目標であり、結果系 KPI は在庫金額であった (P1)。これを達成する仮説として、「人手に任せている商品の需要予測の精度向上」があげられ、需要予測と実際の需要との差が原因系 KPI となった (P2)。この差を少なくするため、販売データを用いた需要の自動予測を試み、人手を超える精度を確保する目処がたった (P5, P6)。本事例の場合には、事前の検証だけでなく、需要予測精度の向上に向けて予測ロジックの継続的な改善の取り組みを行った (P7~P9)。この結果、BI システムのレベルは、「KPI の可視化」と「サービス実現」レベルとなり、仮説構造化モデルの種類は、需要の予測とそれに基づく発注のための業務プロセスの検討を行ったという意味で、

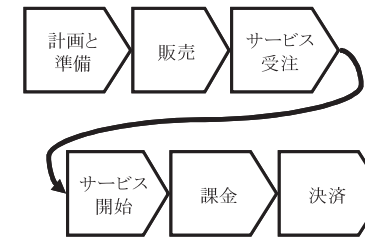


図5 サービス代理店の業務プロセス

Fig. 5 Business processes of service agents.

タイプ3となった。

事例4：コールセンタへの問合せ削減

本事例では、あるサービス提供企業においてサービスの代理店からの問合せ件数を削減し、代理店の満足度を向上させることが目標であり、問合せの件数が結果系 KPI であった (P1)。この達成のためサービス代理店での業務プロセスをタイプ3の仮説構造化モデルによりモデル化し (P3)、問合せがどの業務プロセスで多く発生するかを調べた (図5)。その結果として、特定の業務プロセスで多くの問合せが発生することが分かり、仮説として、「特定の業務プロセス A の問合せを削減すれば全体の問合せも削減できる」があがった (P2)。特定業務プロセス A の問合せ数が原因系 KPI となる。さらに詳しく分析した結果、真の原因はサービス説明書の不備にあることが分かり (P4, P5)、説明書の改善を実施し、効果を確認できた (P6, P7)。この結果から、定期的に問合せ内容を分析して施策の実施と管理を行う、「KPI の可視化」と「施策の管理と評価」レベルの BI システムを構築した (P8, P9)。本事例の場合には、問合せ内容が膨大でありかつ日本語であるため、テキスト解析技術を用いて問合せの分類を行っている。

表2から分かるように、様々な BI システムがあり、複数のレベルをカバーするものもある。

4.2 仮説構造化モデルの適用性評価

仮説構造化モデルの適用性を検証するため、筆者らが所属する分析チームにて実施した過去8年の80件の分析事例について、3つのモデル化手法のどれを使用していたかを評価した。その結果、タイプ1は30%で24件、タイプ2は38%で30件、タイプ3は25%で20件となり、カバーできなかったものは7%の6件であった。この結果から、93%は3種類の

仮説構造化モデルでカバーできることが分かる。カバーできなかった6件の内容は、次のようなものであった。2件は、レコメンデーションなどBIの技術を用いて新規サービスを行うものであり、BIシステムのレベル4に相当する。この件では、課題の解決を狙うものではなかったため、仮説や仮説構造化モデルを検討することがなかった。1件は、通常と異なるアクセスログを早期に発見することを狙いとしてデータマイニングを行ったため、特定の課題は明確であったがモデルを構築することはなかった。通常と異なるアクセスログを検出した後で、それを分類し施策を打つところまで検討できていたなら、タイプ1の仮説構造化モデルに分類された件であったと想定される。残り3件は、将来予測のためのシミュレーションなど特定の技術課題を企業から依頼されてそれに応えた案件であり、企業側では課題の全体像や仮説構造化モデルは存在したかもしれないが、筆者らは知ることができなかった。

5. 考 察

本稿で提案したBIの業務手順と仮説構造化モデルについて考察する。

まず業務手順であるが、これは筆者が所属する分析チームで実践してきた様々な分野での111件の分析の、属人的なノウハウを今回体系化したものであり、4章に示した本手法の実案件への適用例により実用性は示せたと考える。その一方、3つの限界があると考えられる。1つ目は、現実のBIシステムにおいて、最初からBIの業務手順をすべてカバーすることは難しいことである。適用にあたっては、まず可視化のレベルから入って、企業意思決定者の理解を得てから、次のレベルに到達する現実的なアプローチが必要と考える。ただし、「サービス実現」レベルについては他のレベルの延長にあるものではなく、革新的なサービスを提供しようとする企業の意思により実現されると思われる。2つ目は、本手法が前提とする課題の解決が必要とされていない場合があることである。課題の解決のためにBIシステムを構築するのではなく、今持っている情報を最大限に活用して新たにサービスを考えたいという企業からの依頼に応える場合にこのような状況が発生する。このような場合には、企業目標の達成からBIシステムの情報要求を検討する手法は使えず、より発見的な手法で検討してゆく必要があると考える。たとえば、通常と異なるアクセスログのパターンを発見することが目的であった4.2節の例もこの場合に該当する。3つ目は、BIシステムの構築にあたって、まず企業内の業務データを統合するデータウェアハウスを作り、後は経営層の必要に応じてレポートをアドホックに生成したいという考え方をとる場合である。この考え方に従うと、情報要求はシステム構築時に抽出できずレポート設計時に抽出することになる。この場合には、必要な情報がデータウェアハウスにあるかどうかの保証がなく、

BIシステムの構築時点での投資効果も明確に示せないという問題が残される。

次に仮説構造化モデルについて考察する。提案した3種類のモデル化手法の適用範囲が広いことは前章で定量的に示せたと考える。適用できなかった6件の例についても、そもそもモデル化のプロセスを経ていない事例であり、モデル化手法そのものには問題があるとは考えていない。また、この3種類でどの程度の事例をカバーできるかについては、次のようにも考えられる。Suenagaら¹⁶⁾は、多くのデータ分析事例の整理の結果、データ分析のシナリオが業界によらず4つの目的に分類できることを示している。このうち2つの目的は、原因と結果の因果関係を洗い出し、原因の制御により結果を改善するものであり、この場合にはタイプ2が適用できると考える。またもう1つの目的は、施策を実施する対象を分類して絞り込むものであり、この場合にはタイプ1またはタイプ3の分類が適用できると考える。一方最後の目的は、検出した状況が通常状態から外れていることを早期に発見するものであり、発見が目的のため、目標に対してそれを解決する仮説を検討する仮説構造化モデルを対応付けることは困難である。4.2節で述べた、モデル化手法を適用できないアクセスログの場合がこれに相当する。このようなことから、目標に対してそれを解決する仮説を検討する上記の3つの目的の場合には、本提案の3種類のモデル化手法が適用できると考える。

6. まとめと今後の課題

本稿では、企業目標の達成に役立つ情報を提供するという観点から重要である、BIシステムの情報要求の抽出手法を提案した。提案した手法では、まず企業目標の達成に貢献する有用な情報が抽出できるかどうかを試行的に実施し、その検証結果が良ければ情報要求の結果をBIシステムにて実現するアプローチをとる。そのためのBIの業務手順と情報の抽出手法を提案した。また、BIシステムには様々なレベルがあることから、これを4レベルに分類し、それぞれが業務手順のサブセットに対応付けられることを示した。さらに、3種類の仮説構造化モデルにより、企業目標を達成する仮説や施策を整理、分類できることを示した。

今後の課題について3つあげる。まず本稿では、情報要求の抽出と整理に特化した手法を提案したが、より汎用的なゴール指向分析手法とどう関係付け、一体となった手法にするかが課題である。ゴール指向分析手法と情報要求の抽出手法の関係の明確化や組合せ手法の検討が必要になると想定される。またそのためには、本稿で提案した各プロセスとその入出力情報をより精緻に定義し、そのうえで一体化することが必要である。今後、詳細化をはかりたい。次に本稿では、3種類の仮説構造化モデルを用いて、仮説や施策、そしてそれに関係

する情報を分類・整理できることを示したが、このような分類で十分なのかについて検討の継続が必要である。特にタイプ2については、業務対象とその関係を整理するものであるが、それをどう抽出するかについては明確な指針が示せていない。一般論での検討は難しいが、業界などの状況に特化した手法を検討するなど、より精緻な手法が今後必要と考える。最後に、本稿で提案した仮説構造化モデルは、統計処理やデータベースの分野で提案されている統計モデルやデータモデルと必ずしも無関係ではなく、統計モデルの一部が本稿で提案する仮説構造化モデルにつながり、また本稿の仮説構造化モデルの一部がデータモデルにつながるなどが事例などを通じて明らかになってきている。今後、これらの関係を整理してゆきたい。

参 考 文 献

- 1) Golfarelli, M., Rizzi, S. and Cella, I.: Beyond Data Warehousing: What's Next in Business Intelligence?, *7th International Workshop on Data Warehousing and OLAP (DOLAP 2004)*, pp.1-6 (2004).
- 2) Simson, G.C. and Witt, G.C.: *Data Modeling Essentials*, Morgan Kaufmann (2004).
- 3) Navathe, S.B.: Evolution of Data Modeling for Databases, *Comm. ACM*, Vol.35, Issue 9, pp.112-123 (1992).
- 4) Kaplan, R.S. and Norton, D.P.: *The Balanced Scorecard Translating Strategy into Action*, Harvard Business School Press (1996).
- 5) Kokune, A., Mizuno, M., Kadoya, K. and Yamamoto, S.: FBCM: Strategy Modeling Method for the Validation of Software Requirements, *Journal of Systems and Software*, Vol.80, Issue 3, pp.314-327 (2007).
- 6) 山本修一郎: ゴール指向によるシステム要求管理技法, ソフト・リサーチ・センター (2007).
- 7) Yamamoto, S., Kaiya, H., Cox, K. and Bleistein, S.: Goal Oriented Requirements Engineering - Trends and Issues, *IEICE*, No.11, pp.2701-2711 (2006).
- 8) 山本修一郎: 非機能要求とゴール指向要求定義, 情報処理学会誌, Vol.49, No.4, pp.371-379 (2008).
- 9) 末永高志, 中川慶一郎, 関根 純, 山中啓之, 高橋彰子: 業務データ分析のためのデータ分析フレームワークの開発, 情報処理学会論文誌 データベース (TODS), Vol.1, No.2, pp.15-25 (2008).
- 10) Sekine, J., Suenaga, T., Yano, J., Nakagawa, K. and Yamamoto, S.: A Business Process-IT Alignment Method for Business Intelligence, *10th Workshop on Business Process Modeling, Development and Support (BPMDS'09)*, pp.46-57 (2009).
- 11) Berry, M.J.A. and Linoff, G.S.: *Data Mining Techniques: for Marketing, Sales and*

Customer Relationship Management, 2nd Edition, Wiley (2004).

- 12) Kahn, B., Strong, D. and Wang, R.: Information Quality Benchmarks: Product and Service Performance, *Comm. ACM*, pp.184-192 (Apr. 2002).
- 13) Wang, R., Kon, H. and Madnick, S.: Data Quality Requirements Analysis and Modelling, *9th International Conference on Data Engineering*, pp.670-677 (1993).
- 14) Recommendations: Amazon.com. <http://www.amazon.com/gp/help/customer/display.html?nodeId=13316081>
- 15) Davenport, T.H. and Harris, J.G.: *Competing on Analytics: The New Science of Winning*, Harvard Business School Press (2007).
- 16) Suenaga, T., Takahashi, S., Saji, M., Yano, J., Nakagawa, K. and Sekine, J.: A Framework for Business Data Analysis, *Workshop on Business Intelligence Methodologies and Applications (BIMA'08)*, pp.703-708 (2008).

(平成 21 年 3 月 27 日受付)

(平成 21 年 9 月 11 日採録)



関根 純 (正会員)

昭和 57 年東京大学大学院工学系研究科計数工学専攻修士課程修了。同年日本電信電話公社。平成 17 年 NTT データ技術開発本部副本部長。博士 (工学)。データベース, BI の研究開発に従事。日本データベース学会, ACM 各会員。



末永 高志

平成 9 年早稲田大学理工学部経営システム工学科卒業。平成 11 年同大学大学院理工学研究科修士課程修了。同年 NTT データ入社。パターン認識, データ分析技術の実用化研究に従事。電子情報通信学会, 人工知能学会各会員。



矢野 順子

平成 15 年東京理科大学大学院工学研究科経営工学専攻修士課程修了。同年 NTT データに入社。データ分析技術の実用化研究に従事。日本 OR 学会会員。



中川慶一郎

平成 4 年早稲田大学大学院理工学研究科修士課程修了。同年 NTT データ入社。早稲田大学大学院理工学研究科博士課程満期退学。NTT データ技術開発本部主幹研究員。データ分析技術の実用化研究に従事。博士(工学)。日本 OR 学会会員。



山本修一郎(正会員)

昭和 52 年名古屋工業大学情報工学科卒業。昭和 54 年名古屋大学大学院工学研究科情報工学専攻修了。同年日本電信電話公社入社。平成 2 年 NTT データ技術開発本部副本部長。平成 19 年同社初代フェロー。システム科学研究所所長。ソフトウェア工学、ユビキタスコンピューティング、知識創造デザインの研究に従事。情報処理学会業績賞、逓信協会前島賞等受賞。博士(工学)。人工知能学会知識流通ネットワーク研究会主査(平成 19 年~)。電子情報通信学会、日本ソフトウェア科学会、人工知能学会、ACM、IEEE 各会員。