

東京C.P.A.

公認会計士業務資料集 VI

I T 委員会研究報告書

「ビッグデータ分析の経営などへの活用事例」

日本公認会計士協会東京会

平成 27 年 6 月 5 日

日本公認会計士協会東京会  
会長 柳澤義一 殿

日本公認会計士協会東京会  
I T 委員会  
委員長 森山直輝

## 研究報告書

下記研究内容につきまして、ここに報告いたします。

記

<研究内容>  
「ビッグデータ分析の経営などへの活用事例」

教材コード	K090003
研修コード	9106
履修単位	2 単位

# 研 究 報 告 書 目 次

はじめに .....	1
第1章 ビッグデータとは .....	2
1. ビッグデータの概念整理 .....	2
2. ビッグデータを活用することのメリット .....	3
第2章 ビッグデータを利活用する上での留意点 .....	5
1. パーソナルデータの利活用の留意点 .....	5
(1) パーソナルデータの利活用の基本理念及び原則 .....	5
(2) 保護されるパーソナルデータの範囲 .....	6
(3) パーソナルデータの利活用のルールの内容のあり方 .....	7
(4) パーソナルデータの保護のための関連技術の活用 .....	8
(5) 情報漏洩リスク対応 .....	8
2. 制度改正とビッグデータの活用の留意点について .....	8
(1) マイナンバー制度 .....	8
(2) 個人情報保護法及びマイナンバー法改正 .....	9
第3章 データを活用する上での前提 .....	15
1. データ環境 .....	15
2. ツール環境 .....	15
3. 情報リテラシー .....	16
第4章 ビッグデータの経営面の活用 .....	17
1. 経営にかかるデータ活用の可能性事例 .....	17
(1) ビッグデータの経営面の活用事例 .....	17
(2) ビッグデータを用いた財務データ分析の拡充 .....	19
2. 監査にかかるデータ活用の可能性 - 巨大容量仕訳データの分析 - .....	26
(1) 監査における仕訳データの分析（仕訳テスト） .....	26
(2) 仕訳データ量 .....	26
(3) 巨大容量の仕訳データ分析方法 .....	27
(4) BIツールを利用した巨大容量の仕訳データ分析 .....	28
(5) 仕訳テスト以外の監査の領域へのBIツールの適用の可能性 .....	30
第5章 ビッグデータ分析ツールの特徴 .....	31
1. ビッグデータ活用例とツールの役割等 .....	31
(1) ビッグデータと分析ツールの関係 .....	31
(2) 分析ツールの役割 .....	32
2. ビッグデータ分析ツールの特徴 - 経営管理 .....	33
(1) 分析ツール大別 .....	33
(2) 独立BIの機能 .....	34
(3) 活用例 .....	35
(4) 導入等 .....	35
(5) 監査等での活用 .....	36

3.	ビッグデータ分析ツールの特徴 － 予知保全	36
(1)	予知保全分野におけるビッグデータ分析とは	36
(2)	予知保全分野におけるビッグデータの診断プロセス	36
(3)	予知保全分野におけるビッグデータ分析の留意点	37
4.	予知保全分野におけるビッグデータ分析の監査への応用	37
(1)	内部統制の観点から	37
(2)	不正との関係	38
	おわりに	40

## はじめに

情報通信技術の進展によって、これまで電子データ化できなかつたような種類の情報が電子データ化できるようになった。また、新しい情報処理技術の開発によって、これまで処理できなかつた大量のデータを処理する事が可能となつてきている。このような、多様な目的に利用される、大量のデータの集合体を一般的にビッグデータと呼ばれている。

ビッグデータは、様々な分野で活用することが見込まれるものである。近年のコンピュータの性能向上や分析用ツールの普及により、ビッグデータを取り扱うための技術的・コスト的な制約が取り払われつつある。特に企業は経営活動の判断を行うために様々な情報（データ）を分析し、意思決定に役立てていることから、ビッグデータを経営に活用することの関心が高まっている。

ビッグデータは様々な分野で活用できることが見込まれるため、我々公認会計士にとつても今後関心が高まるものであると考えられる。しかし、ITの専門家ではない公認会計士にとって、ビッグデータとはどのようなものなのか、実際にどのように活用されているか、と言うことについての知識は必ずしも十分ではないと思われる。

そこで本委員会では、ビッグデータとはそもそもどのようなものであるか（第1章）、及びビッグデータを活用する上で留意しなければならないこと（第2章）、更にはビッグデータを活用する上での前提となる事項（第3章）、といったビッグデータの理論的側面を整理し、ビッグデータを活用する上での全体像を押さえた上で、具体的な経営・監査の場面におけるビッグデータの活用事例について持ち寄って議論を行つた（第4章）。

更に、ビッグデータを効果的かつ効率的に実施する上でツールの活用も欠かせないため、ビッグデータ分析ツールの機能やメリット、活用例等についても取りまとめを行つた（第5章）。

少しでも会員の皆様のご参考になれば幸いである。

### （利用上の注意）

当研究報告書については、会員・準会員各位自らの責任においてご活用ください。

当研究報告書の利用に伴い、結果的に発生しました不利益については、原因の如何を問わず、当会では一切の責任を負いかねますのであらかじめご承知おきください。

なお、当研究報告書は、日本公認会計士協会の公式見解ではなく、あくまでも当委員会の研究成果を取りまとめたものです。

# 第1章 ビッグデータとは

## 1. ビッグデータの概念整理

昨今、情報技術の飛躍的な進化にあいまって、ストレージ(データを永続的に記録する装置)の大容量化、データ処理速度の向上が実現され、データの保存・分析コストは低下してきている。

ビッグデータを「事業に役立つ知見を導出するためのデータ」とし、ビッグデータビジネスの目的について、「ビッグデータを用いて社会・経済の問題解決や、業務の付加価値向上を行う、あるいは支援する事業」と目的に着眼して定義している例がある。また、ビッグデータとは様々な情報の集まりであり、その巨大なデータ群全体の総称として用いられるものである。その特徴として、リアルタイム性をもった情報であり、日々大量に生成される情報であることが挙げられる。

ビッグデータに統一的な定義の見解はないが、本研究報告書上、ビッグデータとは「多様な目的に利用される、大量のデータの集合体」と定義し、論旨を展開することとする。

本研究報告書においては、下記のようなデータの情報量が大量になった場合、ビッグデータに該当するものとしている。

例えば、購買履歴情報、位置情報、ソーシャルネットワークサービス（以下「SNS」という。）における投稿、銀行取引情報、財務報告データ等が挙げられる。

これまで、データの収集・分析には、財務データ・販売データ・購買データ等、一定の基礎に基づき整理されたデータである構造化データが主に用いられてきた。そのため、大量の非構造化データの収集・分析には多大なコストが必要であったが、情報技術の進歩とストレージの低コスト化により、非構造化データの管理技術が実用化され、クラウドコンピューティング等の大量データの分散処理技術も普及し、大量データの活用が可能となってきた。

ビッグデータは、どの程度のデータ規模かという量的側面だけでなく、どのようなデータから構成されるか、あるいはそのデータがどのように利用されるかという質的側面において、従来のデータとは違いがあると考えられる。

また、ビッグデータは全てのデータを対象としているという点でも特徴的である。従来は、情報が膨大になる場合は無作為抽出した標本に基づく分析に頼らざるを得なかつたが、情報技術の進歩に伴い、このような制約が取り払われ、これまで把握・分析することが困難であった相関関係等を見出すことが出来ると考えられる。

歴史的には、データ収集は、例えば国勢調査のように、大規模なものは国家が実施するものであった。人口調査等は人手によるものであり、膨大なコストと時間が必要とされていた。その後、統計学的手法である無作為抽出により、一定数の標本があれば高い精度で全体の状況を把握することが可能となった。この無作為抽出という考え方により、例えば経済調査や人口動態調査を全件調査ではなく、標本による調査で行うことが出来るようになった。

ただし、標本調査による場合、無作為性を確保することが困難な場合もあり、データ収集に偏りがあった場合においては、誤った推定計算となってしまうこともある。標本調査

は全体の動向を、特定の目的から推定するための手法であり、全く別の視点での分析等を行うことは困難である。

ビッグデータは必要に応じて部分的に情報収集することから、全体的な情報を積極的に収集する方向へシフトしたものであり、可能な限り全ての情報を電子化し、データとして収集することを目指す発想（データフィケーション）である。検討対象となる情報が全体となることにより、標本からの母集団の推定ではなく、様々な角度からの分析、個々のデータの異常値の発見等、膨大な情報に埋もれていた物事を発見するような分析も可能となると考えられる。

## 2. ビッグデータを活用することのメリット

情報技術の進歩等により、これまで不可能であった大量データの蓄積・分析が可能となったことで、全ての情報を電子化しデータとしてすることで、そのデータの活用を検討することが可能となった。ビッグデータの活用・分析は発展途上であり、今後、新たな市場・サービスの創出の可能性があると考えられる。例えば、小売、自動車、医療、農業、エネルギー、インフラ等の分野における活用が挙げられる。

小売分野においては、以前からPOSデータ等を使用して行われていた売れ筋商品の把握や、顧客単価を上昇させるための商品や店員の最適配置の検討に加えて、年齢・性別・属性・趣向、ウェブサイト閲覧履歴、位置情報や商品購買履歴などを組み合わせ、顧客毎にカスタマイズされた商品・サービスの提供、ウェブサイトにおける広告表示等を行うといった取組が考えられる。

自動車分野においては、走行車両におけるセンサからデータを収集し、最適ルートの案内、急ブレーキが多発する危険な地帯を分析し運転者への注意喚起、走行距離に応じた自動車保険料の設定、車両盗難時における車両追跡を行う等のサービス等が考えられる。

医療分野においては、医療機器がアナログからデジタルへ移行されてきていることや、電子カルテの普及に伴い、CTやMRI等の先進医療機器から大量のデータが得られるようになってきている。これらを分析して個々の患者に最適な治療を提供することが可能となると考えられ、医療機関が個別に管理している患者データを共有することが出来れば、データの収集・分析から、医師の診断に至るまでのプロセスの効率化、介護等の医療サービスへの活用、新たな医療機器の開発へつながる可能性もあると考えられている。

農業分野においては、気温・日射量・雨量や農作業の内容をセンサやモバイル端末で収集し、生産計画から収集・出荷までのプロセスをデータで管理することにより、農作業の効率化や作物毎の収益率の把握、収穫量増加やノウハウの共有化を図ることが可能となると考えられる。農業従事者は高齢者が多いことから、高齢者でも操作が容易である安価なITシステムの開発や、流通・地域・消費者をつなぐ食のバリューチェーン確立による農業の生産性向上が可能となっていくと考えられる。

エネルギー分野においては、スマートメータやHEMS（Home Energy Management System）、ビル内のBEMS（Building Energy Management System）等を通じて収集された電力データを分析し、電力需要に応じた発電や送配電、最適な運転状態への自動制御を行うことにより、効率的なエネルギーマネジメント等が可能となると考えられる。

インフラ分野においては、道路や港湾等のインフラが老朽化するなか、政府は情報技術

を活用してインフラ管理の安全性、信頼性、効率性の向上を目指す方針であり、センサは非破壊検査技術等を利用し、膨大な数のインフラ維持管理データを収集・分析して維持管理情報の仕組みを構築することでトータルコスト削減や更新費の平準化が可能となっていくと考えられる。

#### 【参考文献】

- ・ 総務省「情報通信白書 平成 24 年版」
- ・ ビクター・マイヤー=ショーンベルガー、ケネス・クキエ「ビッグデータの正体」講談社、2013 年 5 月

## 第2章 ビッグデータを利活用する上での留意点

ビッグデータの利活用は産業面でのニーズの高まりが期待されている一方、世界的には、ビッグデータの範囲内に含まれるパーソナルデータ（経済産業省の定義では、個人情報保護法に規定する「個人情報」に限らず、位置情報や購買履歴等広く個人に関する個人識別性のない情報も含むデータ）の利活用に関する議論が行われている。

パーソナルデータの利活用については、多くの可能性が期待されている一方、プライバシーの保護等の観点から様々な課題が指摘されている。パーソナルデータの利活用に関する課題の多くは、パーソナルデータの利活用のルールが明確でないため、企業にとっては、どのような利活用であれば適正といえるかを判断することが困難であること、消費者にとっては、自己のパーソナルデータが適正に取り扱われ、プライバシー等が適切に保護されているかが不明確になっており、懸念が生じていることから発生している。したがって、パーソナルデータを利活用する際には、適正性の判断や社会一般が想定するプライバシーに留意する必要があると考えられる。

なお、我が国のパーソナルデータの保護に関する法律としては、個人情報保護法、行政機関個人情報保護法、独立行政法人等個人情報保護法が挙げられる。また、パーソナルデータの利活用については、統計法、電気通信事業法による通信の秘密の保護、知的財産権の保護、情報公開法による不開示情報の保護なども関連する。

上記のようにパーソナルデータは法律でも保護されており、また法律に違反していなくても社会一般が想定するプライバシーに反していると、社会的な批判を受ける可能性があるため非常に重要である。そのため、ビッグデータを活用する上での留意点として、特にパーソナルデータに焦点を絞って記述する。

また、現在個人情報に関する法律の改正が検討されているため、法律の変更により、パーソナルデータの利活用の留意点が変わる可能性があることにも注意する必要がある。今後の法改正では、個人情報の定義の明確化や、匿名化の加工方法や、個人情報保護の方針の作成が要求されるようになる可能性があり、法律が改正されればこれらに対応する必要が出てくると考えられる。そのため、今後の法律の改正の動向には十分留意する必要がある。

### 1. パーソナルデータの利活用の留意点

#### (1) パーソナルデータの利活用の基本理念及び原則

パーソナルデータの利活用に関する研究報告書として、「パーソナルデータの利用・流通に関する研究会報告書（総務省・平成25年6月12日）」がある。「パーソナルデータの利用・流通に関する研究会報告書」では、パーソナルデータの保護の目的を明らかにするという観点から、パーソナルデータの利活用の基本理念として、以下の2点を挙げている。

- ① 個人情報を含むパーソナルデータの保護は、主としてプライバシー保護のために行うものである。

- ② プライバシーの保護は、絶対的な価値ではなく、表現の自由、営業の自由などの他の価値との関係で相対的に判断されるべきものである。
- その上で、上記のパーソナルデータの利活用の基本理念を具体化するものとして、次の7項目をパーソナルデータ利活用の原則として提示している。

透明性の確保	パーソナルデータの利用に関し、本人が必要な情報に容易にアクセスする機会を提供すること。
本人の関与の機会の確保	パーソナルデータの本人が、パーソナルデータをどのように利用されるかについて関与する機会を確保すること。
取得の際の経緯（コンテキスト）の尊重	パーソナルデータの利用は、本人がパーソナルデータを提供した際の経緯（コンテキスト）に沿って、本人の期待と合致する形態で行うこと。
必要最小限の取得	パーソナルデータの取得は、パーソナルデータの利用目的の実現のため必要最小限のものとすること。
適正な手段による取得	パーソナルデータの取得は、適正な手段によるものとすること。
適切な安全管理措置	パーソナルデータは、パーソナルデータの性質に沿って適切な安全管理措置をとること。
プライバシー・バイ・デザイン	パーソナルデータを利用する者は、商品開発時などそのビジネスサイクルの全般にわたって、プライバシーの保護をデザインとしてあらかじめ組み込んでおくこと。

## (2) 保護されるパーソナルデータの範囲

「パーソナルデータの利用・流通に関する研究会報告書」では、保護されるパーソナルデータの範囲については、プライバシーの保護というパーソナルデータの利活用の基本理念を踏まえて考えるべきであるとしている。その際、現行の「個人情報」の範囲や、諸外国や国際機関等で保護の対象とされているパーソナルデータの範囲等を踏まえて、保護されるパーソナルデータの範囲を画定する必要があるとされている。

以下の表は現行の「個人情報」とパーソナルデータの例を示したものである。

個人情報となるもの	個人情報か不明確なもの (パーソナルデータ)	
<ul style="list-style-type: none"> <li>・氏名</li> <li>・住所</li> <li>・顔の画像</li> </ul>	<ul style="list-style-type: none"> <li>・端末番号</li> <li>・会員番号</li> <li>・体重</li> <li>・指紋</li> <li>・パスポート番号</li> </ul>	<ul style="list-style-type: none"> <li>・位置情報</li> <li>・クレジット番号</li> <li>・身長</li> <li>・血液型</li> </ul>

一般的に、パーソナルデータの利活用に関し、プライバシーが問題になるのは、当該パーソナルデータと特定の個人の結びつきが強い場合である。したがって、保護される

パーソナルデータの範囲については、現行の個人情報保護法と同様に、個人識別性を有するものとすることが、基本的には妥当であるとされている。

なお、取得等の際に特定の個人が識別されなかつたとしても、他のパーソナルデータと併せて分析されること等により、特定の個人が識別される可能性があることについて、十分に配慮する必要があるとされている。

取得の際に特定の個人が識別されなくても、実質的個人識別性の要件を満たすので保護されるパーソナルデータになりうる具体例としては、①個人のPCやスマートフォン等の識別情報（端末ID等）、②継続的に収集される購買・貸出履歴、視聴履歴、位置情報等、③IPアドレス、クッキーなどが例示されている。これらは、直接個人を特定できるものではないが、他の情報と組み合わせると特定の個人を識別できるようになる可能性が高く、実質的個人識別性の要件を満たすと考えられるものである。

保護されるパーソナルデータに当たらないと考えられる例としては、一般に公開されている国の統計情報など再識別化を不可能または十分に困難にしたといえるものが挙げられている。これは実質的個人識別性はないといえることから、保護されるパーソナルデータには当たらないとされている。

なお、他の情報との連結等により再識別化の可能性がある匿名化されたパーソナルデータについても、適切なセーフガードを設定すれば、実質的個人識別性は無くすることが出来るため、保護されるパーソナルデータに当たらないとして、利活用を行うことが可能になると考えられている。

### (3) パーソナルデータの利活用のルールの内容のあり方

「パーソナルデータの利用・流通に関する研究会報告書」では、パーソナルデータの利活用について以下のようにルールの内容のありかたを示している。保護されるパーソナルデータの中には、氏名などの通常公にされている情報から、健康に関する情報など人に知られたくない情報まで様々な性質のものがある。このため、保護されるパーソナルデータを一律に取り扱うのではなく、そのプライバシー性の高低に応じて適正に取り扱うことが必要である。

また、パーソナルデータの取扱いについては、取得の際の経緯（コンテキスト）に沿った取扱いである場合と、それ以外のコンテキストに沿わない取扱いの場合に分けて、適切なあり方を考えるべきである。

保護されるパーソナルデータは、そのプライバシー性の高低により、次の3類型に分類し、それぞれの類型に応じて適正に取り扱うべきであるとされている。

類型	例
一般パーソナルデータ（保護されるパーソナルデータのうちプライバシー性が高くないもの）	<ul style="list-style-type: none"><li>・氏名など本人を識別する目的などで一般に公にされている情報</li><li>・本人の明確な意図で一般に公開された情報</li><li>・名刺に記載されている情報など企業取引に関連して提供される情報（ビジネス関連情報）</li></ul>

<p>慎重な取扱いが求められるパーソナルデータ（保護されるパーソナルデータのうちプライバシー性が高いもの。一般パーソナルデータ又はセンシティブデータ以外の保護されるパーソナルデータ）</p>	<ul style="list-style-type: none"> <li>○スマートフォンやタブレット端末など移動体端末に蓄積される以下のようなパーソナルデータ <ul style="list-style-type: none"> <li>・電話帳情報</li> <li>・G P Sなどの位置情報</li> <li>・通信内容・履歴、メール内容・送受信履歴等の通信履歴</li> <li>・アプリケーションの利用履歴、写真・動画</li> <li>・契約者・端末固有 I D</li> </ul> </li> <li>○継続的に収集される購買・貸出履歴、視聴履歴、位置情報等</li> </ul>
<p>センシティブデータ（保護されるパーソナルデータのうちプライバシー性が極めて高いもの）</p>	<ul style="list-style-type: none"> <li>・思想、信条及び宗教に関する情報</li> <li>・人種、民族、門地、身体・精神障害、犯罪歴、病歴その他の社会的差別の原因となるおそれのある事項に関する情報</li> <li>・勤労者の団結権、団体交渉その他団体行動に関する情報</li> <li>・集団示威行為への参加、請願権の行使その他の政治的権利の行使に関する情報</li> <li>・健康又は性生活に関する情報</li> </ul>

#### (4) パーソナルデータの保護のための関連技術の活用

パーソナルデータの適正な利活用の促進のためには、プライバシーを保護するために利用可能な技術（プライバシー強化技術： Privacy Enhancing Technologies (P E T s) ）を最大限に有効活用することも必要と考えられる。

プライバシー強化技術には様々なやり方やツールが存在するが、一例としては暗号化技術や匿名化技術といったものが挙げられる。

暗号化技術の活用例としては、データの内容毎に情報漏洩した場合の影響を考慮して、それぞれ違った取扱いにするよう整理し、暗号化するといったことが考えられる。

匿名化技術の活用例としては、匿名化により非識別化されたデータと元の識別可能なデータ（連結可能匿名化における対応表を含む。）の双方を保持・使用する場合は、これらのデータは別々に保管するようルール決めするなどが考えられる。

#### (5) 情報漏洩リスク対応

ここまででは、パーソナルデータとしての留意点を述べてきたが、パーソナルデータも含めたビッグデータは企業の競争力の源泉となる営業秘密になるものもある。したがって、営業秘密管理のガイドラインを参考に、情報漏洩リスク対策を行うことが不可欠である。その際、個人情報保護マネジメントシステム（JIS Q 15001）または情報セキュリティマネジメントシステム（ISO 27001）等の認証制度の活用も有効となる。

## 2. 制度改正とビッグデータの活用の留意点について

### (1) マイナンバー制度

行政手続における特定の個人を識別するための番号の利用等に関する法律（マイナン

バー法）が 2016 年 1 月から運用されるが、個人番号の利用範囲が制限され、原則として役所に書面を提出する場面以外では利用できない（以下の表参照）。

なお、特定個人情報とは個人情報保護法の個人情報とマイナンバー法の個人番号を合わせたものであり、複数の特定個人情報を検索できるように保存すると「特定個人情報ファイル」となる。

		個人情報保護法	マイナンバー法
適用除外		個人の数が過去 6 か月以内に 5,000 を超えない者	なし
利 用	利用範囲	定めなし（企業側で自由に設定）	厳しく制限（法定）
	利用目的の変更・目的外利用	ある程度認められている	厳しく制限
提 供	第三者提供	ある程度認められている	厳しく制限
	オプトアウト方式による第三者提供	認められる	認められない
	第三者への委託	認められる	
収集・保管		「利用」と同じ	「提供」と同じ
データベースの作成		制限なし	厳しく制限
安全管理措置		あり（生存者の個人情報が対象）	あり（死者の個人番号も対象）
行政（または第三者機関）の監督権限		立入検査権なし	立入検査権あり
罰則		直罰規定なし	直罰規定あり

なお、特定個人情報（個人番号）はビッグデータとして活用できないが、法人番号は民間で自由に利用してよいので、企業の与信管理等に活用することが今後増えると思われる。

## (2) 個人情報保護法及びマイナンバー法改正

「個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律案」が第 189 回国会で審議されており、以下のような改正項目が挙がっている。

改正のポイント	内容
情報の定義の明確化	<ul style="list-style-type: none"> <li>個人情報の定義の明確化（身体的特徴等が該当）</li> <li>要配慮個人情報（いわゆる機微情報）に関する規定の整備</li> </ul>
適切な規律の下で個人情報等の有用性を確保	<ul style="list-style-type: none"> <li>匿名加工情報に関する加工方法や取扱い等の規定の整備</li> <li>個人情報保護指針の作成や届出、公表等の規定の整備</li> </ul>
個人情報の保護を強化	<ul style="list-style-type: none"> <li>トレーサビリティの確保（第三者提供に係る確認及び記録の作成義務）</li> </ul>

	<ul style="list-style-type: none"> <li>不正な利益を図る目的による個人情報データベース等提供罪の新設</li> </ul>
個人情報保護委員会の新設及び権限	<ul style="list-style-type: none"> <li>個人情報保護委員会を新設し、現行の主務大臣の権限を一元化</li> </ul>
個人情報の取扱いのグローバル化	<ul style="list-style-type: none"> <li>国境を越えた適用と外国執行当局への情報提供に関する規定の整備</li> <li>外国にある第三者への個人データの提供に関する規定の整備</li> </ul>
その他	<ul style="list-style-type: none"> <li>本人同意を得ない第三者提供（オプトアウト規定）の届出、公表等厳格化</li> <li>利用目的の変更を可能とする規定の整備</li> <li>取り扱う個人情報が5,000人以下の小規模取扱事業者への対応</li> </ul>

上記のように、法改正によって留意するべき点が変化していくため、その時々の改正にも注意を払うべきである。

### (3) 特定個人情報に関する安全管理措置と公認会計士の責務について

マイナンバー制度の整備に伴い、個人番号・特定個人情報の漏えいを防ぐための制度も整備されている。個人番号・特定個人情報の漏えい、滅失又は毀損の防止その他の適切な管理のために、必要かつ適切な安全管理措置を講じる必要がある。この取り扱いを定めたガイドラインとして、「特定個人情報の適正な取扱いに関するガイドライン」が規定されている。ここで規定されているガイドラインの内容は以下のとおりである。

No	規定の分類	安全管理措置の内容（本則）	（参考）中小規模事業者における対応方法 記載ない部分は本則に従う
1	基本方針の策定	特定個人情報等の適正な取扱いの確保について組織として取り組むために、基本方針を策定し、従業者に周知徹底することが重要である。	
2	取扱規程等の策定	事務の流れを整理し、特定個人情報等の具体的な取扱いを定める取扱規程等を策定しなければならない。	特定個人情報等の取扱等を明確化する。 事務取扱担当者が変更となった場合、確実な引継ぎを行い、責任ある立場の者が確認する。
3	組織的安全管理措置	事業者は、特定個人情報等の適正な取扱いのために、次に掲げる組織的安全管理措置を講じなければならない。 a 組織体制の整備 安全管理措置を講ずるための組織体制を整備する。	a 組織体制の整備 事務取扱担当者が複数いる場合、責任者と事務取扱担当者を区分することが望ましい。

	b 取扱規程等に基づく運用 取扱規程等に基づく運用状況を確認するため、システムログ又は利用実績を記録する。	b 取扱規程等に基づく運用 特定個人情報等の取扱状況の分かる記録を保存する。	
	c 取扱状況を確認する手段の整備 特定個人情報ファイルの取扱状況を確認するための手段を整備する。なお、取扱状況を確認するための記録等には、特定個人情報等は含めない。	c 取扱状況を確認する手段の整備 特定個人情報等の取扱状況の分かる記録を保存する。	
	d 情報漏えい等事案に対応する体制の整備  情報漏えい等の事案の発生又は兆候を把握した場合に、適切かつ迅速に対応するための対応体制を整備する。情報漏えい等事案が発生した場合、二次被害の防止、類似事案の発生防止等の観点から、事案に応じて、事実関係及び再発防止策等を早急に公表することが重要である。	d 情報漏えい等事案に対応する体制の整備  情報漏えい等事案の発生等に備え、従業者から責任ある立場の者に対する報告連絡体制等をあらかじめ確認しておく。	
	e 取扱状況の把握及び安全管理措置の見直し  特定個人情報等の取扱状況を把握し、安全管理措置の評価、見直し及び改善に取り組む。	e 取扱状況の把握及び安全管理措置の見直し  責任ある立場の者が、特定個人情報等の取扱状況について、定期的に点検を行う。	
4	人的安全管理措置	事業者は、特定個人情報等の適正な取扱いのために、次に掲げる人的安全管理措置を講じなければならない。  a 事務取扱担当者の監督 事業者は、特定個人情報等が取扱規程等に基づき適正に取り扱われるよう、事務取扱担当者に対して必要かつ適切な監督を行う。	
	b 事務取扱担当者の教育 事業者は、事務取扱担当者に特定個人情報等の適正な取扱いを周知徹底するとともに適切な教育を行う。		
5	物理的安全管理措置	事業者は、特定個人情報等の適正な取扱いのために、次に掲げる物理的 安全管理措置を講じなければならない。  a 特定個人情報等を取り扱う区域の管理	

		<p>特定個人情報等の情報漏えい等を防止するために、特定個人情報ファイルを取り扱う情報システムを管理する区域（以下「管理区域」という。）及び特定個人情報等を取り扱う事務を実施する区域（以下「取扱区域」という。）を明確にし、物理的な安全管理措置を講じる。</p>	
	b	<p>機器及び電子媒体等の盗難等の防止</p> <p>管理区域及び取扱区域における特定個人情報等を取り扱う機器、電子媒体及び書類等の盗難又は紛失等を防止するために、物理的な安全管理措置を講じる。</p>	
	c	<p>電子媒体等を持ち出す場合の漏えい等の防止</p> <p>特定個人情報等が記録された電子媒体又は書類等を持ち出す場合、容易に個人番号が判明しない措置の実施、追跡可能な移送手段の利用等、安全な方策を講じる。「持出し」とは、特定個人情報等を、管理区域又は取扱区域の外へ移動させることをいい、事業所内での移動等であっても、紛失・盗難等に留意する必要がある。</p>	<p>c 電子媒体等を持ち出す場合の漏えい等の防止</p> <p>特定個人情報等が記録された電子媒体又は書類等の移送に当たっては、安全な方策を講じる。</p>
	d	<p>個人番号の削除、機器及び電子媒体等の廃棄</p> <p>個人番号若しくは特定個人情報ファイルを削除した場合、又は電子媒体等を廃棄した場合には、削除又は廃棄した記録を保存する。また、これらの作業を委託する場合には、委託先が確実に削除又は廃棄したことについて、証明書等により確認する。</p>	<p>d 個人番号の削除、機器及び電子媒体等の廃棄</p> <p>特定個人情報等を削除・廃棄したことを確認する。</p>
6	技術的 安全管理措置	<p>事業者は、特定個人情報等の適正な取扱いのために、次に掲げる技術的 安全管理措置を講じなければならない。</p> <p>a アクセス制御</p>	<p>a アクセス制御</p>

	<p>情報システムを使用して個人番号関係事務又は個人番号利用事務を行う場合、事務取扱担当者及び当該事務で取り扱う特定個人情報ファイルの範囲を限定するために、適切なアクセス制御を行う。</p>	<p>特定個人情報等を取り扱う機器を特定し、その機器を取り扱う事務取扱担当者を限定することが望ましい。</p>
	<p>b アクセス者の識別と認証</p> <p>特定個人情報等を取り扱う情報システムは、事務取扱担当者が正当なアクセス権を有する者であることを、識別した結果に基づき認証する。</p>	<p>機器に標準装備されているユーザー制御機能（ユーザーアカウント制御）により、情報システムを取り扱う事務取扱担当者を限定することが望ましい。</p>
	<p>c 外部からの不正アクセス等の防止</p> <p>情報システムを外部からの不正アクセス又は不正ソフトウェアから保護する仕組みを導入し、適切に運用する。</p>	<p>機器に標準装備されているユーザー制御機能（ユーザーアカウント制御）により、情報システムを取り扱う事務取扱担当者を限定することが望ましい。</p>
	<p>d 情報漏えい等の防止</p> <p>特定個人情報等をインターネット等により外部に送信する場合、通信経路における情報漏えい等を防止するための措置を講じる。</p>	

このガイドラインでは、特定個人情報を利用する安全管理措置の本則および、中小規模事業者における対応方法が規定されている。ここで、中小規模事業者とは事業者のうち従業員の数が100人以下の事業者であって、一定の事業者を除く事業者を言うが、公認会計士は大規模事業者としての対応が求められることに留意が必要である。

## 【参考文献】

- ・ 総務省「パーソナルデータの利用・流通に関する研究会報告書～パーソナルデータの適正な利用・流通の促進に向けた方策～」2013年6月12日
- ・ 内閣官房情報通信技術（IT）総合戦略室「個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律案（概要）」
- ・ 影島広泰「実務におけるマイナンバー制度対応」（セミナー資料）
- ・ 特定個人情報保護委員会事務局「マイナンバーガイドライン入門」

## 第3章 データを活用する上での前提

ビッグデータを活用する上で、様々な制約や障害が発生することが想定されるが、基本的にデータ環境、ツール環境、情報リテラシーの3点が必要な要素となる。この3点について以下に説明する。

### 1. データ環境

データ環境とは対象となるデータの質の問題である。ビッグデータ分析は、多様かつ大量のデータを分析することにより、未知の知見を得ようとするものであるが故に、データの種類も多様なものになる可能性が高い。しかしながら、この多様なデータ群については網羅的かつ正確なデータが収集されていることが前提となる。

例えば、会計監査において不正な仕訳について新たな関係を見つけるためビッグデータ分析を実施する場合を想定する。従来、仕訳テストにおいては、不正な仕訳の計上が不自然な時間に行われている、承認がない、あるいは切りのいい金額であるといった特徴を持っているという過去の監査実務により蓄積された経験により、このような特徴のある仕訳をデータから抽出する方法が浸透していると思われる。これに対してビッグデータ分析を実施することで、過去の不正仕訳と様々なデータの間に新たな関係を発見することにより、従来は考慮しなかったような仕訳抽出の基準となるフィルタを発見できる可能性がある。例えば従業員の家族構成や借入金の状況、従業員の移動データ、出退勤情報、建物内の消費電力情報なども分析の対象に加え、統計解析を実施する。この時に、従業員の移動データが捏造されていたり、特定の曜日のデータが欠損している場合、統計解析自体が無意味なものになる可能性がある。このようにビッグデータ分析においては使用するデータが多様なため、データ群についてどの程度正確で網羅的なものなのかを把握し、それがどのように入手可能であるかを入念に検討する必要がある。

### 2. ツール環境

ツール環境とは、ビッグデータ分析を実施するために必要なツールを用意できるかどうかの問題である。

まずソフトウェアであるが、クエリツール、統計解析ツール、各種分析ツールなどが必要となる。ビッグデータ分析においてはデータが多様であることが特徴であるために使用するソフトウェアの選定を誤ると、目的の達成は非常に困難なものとなる可能性がある。一般的には統計解析ツールは比較的万能に使用できるが、市販されているビッグデータ分析ツールには目的や取り込めるデータが限定された分析ツールも多い。先の例では多変量解析を実施できる統計解析ツールは有用であろうが、従来から会計監査で広く利用されているIDEA<sup>1</sup>やACL<sup>2</sup>などのC A A T専用ツールは目的に適合しない可能性がある。

---

<sup>1</sup> IDEAとは世界有数のデータ分析・監査ツールの一つである。開発・発売元であるCaseWare International Inc. はカナダに本拠地を構える。

次にハードウェアであるが、処理対象となるデータが大量であるという特徴を考慮すると、必然的に高性能な処理スペックを持つハードウェアの準備が必要となる。公認会計士の業務領域を想定する場合、スーパーコンピュータのようなものを準備することは極端であるが、ノートPC 1台で目的が達成可能な場合も限定されると思われる。

### 3. 情報リテラシー

情報リテラシーとは、情報活用能力のことであり、ここではビッグデータ分析を行うに当たって前提となる知識などについて記述する。ビッグデータ分析においては、統計学の理解が必要になる場合がある。ビッグデータ分析においては、解析対象となるデータが多様かつ大量であるために、解析のための変数も必然的に多くなり、単純な相関分析や因子分析などではなく多変量解析の技術など、高度な統計学的解析手法が用いられる場合もあるためである。たとえ分析結果のみを活用する場合でも、データの本質を読み解くためには、どのような分析が行われたかをある程度理解しておく必要があり、この理解のために統計学の理解が必要となるのである。このデータ解析の過程で、生成された解析結果から元のデータの本質を読み取り、それを次のステップに繋げることができなければ、ビッグデータに隠された有用な情報も見えてくることはないであろう。但し、市販の分析ツールの中には、上述のような解析プロセスを経て、解析結果が理解しやすいよう標準化した分析ツールも販売されている。このような分析ツールは同業種のビジネスには横展開しやすく、統計的な知識の側面での負担も軽減されるというメリットはあるが、全く未知の結果が導き出される可能性も低くなるというデメリットは理解しておく必要がある。

また、ビッグデータ分析においては、全く新しい知見が導き出される可能性があるものの、データ解析の結果、「何もなかった」という結果が導きだされる可能性もある。このような状況を周囲が理解していなければ、ビッグデータ分析を実施することは困難になるであろう。極端な例えではあるが、先の例では、従業員の移動データを解析すると、「不正仕訳を計上する前には短い距離の歩行を繰り返す」という相関関係が存在すれば、仕訳を抽出する際の新たなフィルタとして利用できるかもしれないが、このこと自体をばかばかいとする企業風土があるようでは、新しい知見の導出は困難なものになるであろう。新しい知見の導出には新規の設備投資も必要になることが多いが、成果も不透明であり、企業の研究開発と同様に、ビッグデータ分析という情報価値をいかに重要視するかという企業風土は重要なものとなる。

最後に、ビッグデータを活用する上ではプライバシーの問題が発生する可能性がある。取り扱うデータが多様あるという特徴から、第2章に詳細に記述されているようにビッグデータ分析の過程でプライバシーに関する情報が特定されてしまう可能性もある。常に注意すべき事項である。

---

<sup>2</sup>ACLとは世界有数のデータ分析・監査ツールの一つである。開発・発売元である ACL Services Ltd. はカナダに本拠地を構える。

## 第4章 ビッグデータの経営面の活用

### 1. 経営にかかるデータ活用の可能性事例

ビッグデータは、既に経営面、すなわち経営者の決断について、様々に活用されている。

この章では、ビッグデータの経営面での活用について、(1)として、ビッグデータの収集、分析、活用のフェーズ毎に、日本の金融サービス会社の取組事例を紹介し、また、収集・分析・活用にトータルで取り組んで成果を挙げた日本のメーカーの事例を紹介し、それに加え、(2)としてこれらの可能性を示唆する事例、取組として「ビッグデータを用いた財務データ分析の拡充」を紹介する。

#### (1) ビッグデータの経営面の活用事例

##### ① ビッグデータの収集事例、分析事例、活用事例

###### A. ビッグデータの収集事例 — A証券会社

###### 【概要】

A証券会社はホールディングス体制であるが、その特徴は、ビッグデータの収集はホールディングスが集中して実施する点である。

その情報収集までの流れは、まず消費者自身がサイト上で、利用金融機関のインターネットサービスのIDとパスワードを入力・登録し、各金融機関のサイトからの資産情報の取得操作を行うことから始まる。次に、それらの取得された情報を、個人が特定できない形で統計処理し、自社のグループ企業約30社の中にある「ビッグデータ室」に情報展開する形で、各社の情報収集が実現されている。

(この収集データを各社が、自らの業務に活かす形で分析・活用することが期待されている。)

###### 【特徴】

この事例に見られるように、企業内部に消費者自身が情報生成するプラットフォームを保有していることや、そこで蓄積された情報を個人が特定されないように情報展開できる統計処理の技術・ノウハウを保有していること、更に、情報の受け皿となる体制を整備している点が特徴として挙げられる。

###### B. ビッグデータの分析事例 — B証券会社

###### 【概要】

B証券会社では、1日当たり約900万件のTwitterやFacebookの情報を収集した上で、その膨大な情報を銘柄毎に定めたキーワードで絞り込みを行い、株価データや顧客データとの相関を分析している。端的に言うと、株価を予測することである。

銘柄を46社、そのキーワードを約43,000と設定するところからスタートして、株価データの変化や顧客属性毎の売買行動との相関関係の強いキーワードを厳選して絞り込んでいくことで、分析の精度を高めている。

同社のビッグデータ分析には、フリーソフトウェアである分散ソフトをベースにした分析ソフトとテキストマイニングツールを導入し、ビッグデータの活用基盤を開発して

いる。

#### 【特徴】

この事例のように、膨大なビッグデータから不要な情報を排除し価値ある情報を絞り込むキーワードを厳選していること、株価データ等、相関関係を分析する対象を明確にしていること、分析に必要なシステム基盤を構築していることは、最近のビッグデータの分析における典型的な方法である。

### C. ビッグデータの活用事例 ー C銀行

#### 【概要】

C銀行では2013年4月より、振込取引、ATM取引、インターネットバンキング取引等から年間数億件以上も発生している内部ビッグデータをデータベース上で統合し、不正取引や事務ミスが起こるパターン、事務量が増加するパターン等との相関関係等を分析し、内部組織の業務改善やリスク管理高度化に活用している。

#### 【特徴】

ビッグデータの情報活用はソーシャル情報等の外部情報だけに依存すると思われがちであるがそうではなく、内部に潜む取引情報・事務情報を内部ビッグデータとして利用し、ビッグデータの分析結果を内部の各部署へ還元し、活かすことが出来ると考えられる。具体的には以下の2例がある。

#### ア. 不正取引監視の強化による銀行口座の犯罪利用防止

C銀行では、勘定取引、振込取引、ATM取引、インターネット取引など年間で数億件の取引が発生するが、外部不正監視、内部不正監視などの機能により、あらかじめ登録したパターンとの相違をシステム上で自動チェックすることで、異常な取引パターンを検知することにより、不正取引などを監視する。これにより、顧客の銀行口座の不正利用を発見もしくは牽制する。

#### イ. 事務品質管理機能により行内事務プロセスを効率的に改善

銀行の営業店では、勘定取引や照会処理、顧客との面談といった、様々な手続きにまつわる事務作業が数多く存在する。従来は、これらの事務プロセス改善を人手により部分的に対応してきたが、それを、各手続きの事務量や事務事故や苦情などのリスク情報を包括的にシステム上で管理し、関連性を分析し、この分析結果をもとに、効率的かつリスクを低減する事務プロセスを導き出し、営業店・本部へフィードバックすることで事務品質を向上させる。

### ② ビッグデータの収集事例、分析事例、活用事例 ー D社（建設機械大手企業）

#### 【ビッグデータの内容】

自社が製造販売した建設機械の所在地、車両状態（画像データ等）、稼動状況

#### 【ビッグデータの収集方法】

従業員等に製品個々の製品にGPSやセンサを取り付けたり、協力会社の担当者が撮った画像データを、通信衛星経由でデータをサーバに収集する。

### 【ビッグデータの分析と活用】

- ア. 建設機械の画像データを元に、配車計画や作業計画の作成支援、最適時期の点検や部品交換など、製品毎や顧客毎に分析することにより、保守・運用サービスの質が向上した。
- イ. 建設機械の稼動状況から、製品の需要予測の分析を行い、生産計画・販売計画に応用した。

### 【ビッグデータの活用について】

売上の拡大とともに「自社建設機械システム」の導入台数は増加し、地域別や機種別の需要動向予測を可能にしている。これにより精度の高い販売・生産計画を作成することが出来、高い業績を実現する原動力になっている。すなわち、稼動状況が高い地域では、営業を強化することにより販売増を狙うことが出来る。一方、稼動状況が低くなったら、早めに生産を絞り、在庫調整に入ることが可能になった。

### (2) ビッグデータを用いた財務データ分析の拡充

ビッグデータは「多様な目的に利用される、大量のデータの集合体」として本報告書で定義している。実際に企業経営にビッグデータを活用する際、詳細な個々のデータをそのまま利用して分析を行う場合と一定の性質をもつデータを集計した結果（以下「統計情報」という。）を利用して分析を行う場合が考えられる。本研究報告書ではビッグデータ分析は、単に大容量データを高速で処理された結果ではなく、高速で作成された結果から将来の傾向やパターンを予測し、最適な意思決定を行うことにあると考えられる。そのため、従来から存在する公表済みの統計情報や比率情報等のデータ等もビックデータに含まれると考えられる。現在でも、社内に蓄積された財務データと市場環境の推移等の情報を組み合わせた活用は実施されているが、それらを更に効果的かつ効率的に実施することで、経営者の投資意思決定や内部監査に活用することが可能になるとと考えられる。具体的には事業セグメント毎の将来の運転資本分析に市場での競争優位、市場規模を反映した分析を行うことで、多種多様な投資計画を立案することができるとともに、投下資本及び資金繰り計画の立案にも活用することが可能であると考えられる。グローバルに事業を展開する企業では、地域別の市場動向や成長性、インフレ率を加味するとともに、カントリーリスクや税制（税負担率等）を加味した製造拠点の設置、立地や研究人員の拡充の可能性を加味した研究開発機関の設置等を検討するといったことも考えられる。

このような分析の事例として、下記の財務情報・比率情報を参考として、ある事業会社Eを使用した地域別の市場動向や成長性、インフレ率を分析した事業戦略の立案事例を例示する。

(事例)

### E社の内部的な財務情報

E社は「自動車ビジネス」「金融ビジネス」「その他ビジネス」をグローバルに展開する事業会社である。

証券取引所に上場を行い、外部金融機関からの借入も行っている。

単位：百万円

	2012年	2013年	増減率
売上高	18, 583, 000	22, 064, 000	18. 7%
営業利益	355, 000	1, 320, 000	271. 8%
営業利益率	1. 91%	5. 98%	
税前利益	432, 000	1, 403, 000	224. 8%
営業利益率	2. 32%	6. 36%	
税後利益	283, 000	962, 000	239. 9%
営業利益率	1. 52%	4. 36%	
株主資本	10, 550, 000	12, 148, 000	15. 1%
純資産	11, 066, 000	12, 772, 000	15. 4%
総資産	30, 650, 000	35, 483, 000	15. 8%
純資産比率	36. 10%	35. 99%	
株主資本利益率	2. 68%	7. 92%	
従業員数	325, 905	333, 498	2. 3%
従業員1人あたり売上高	57	66	

現預金	1, 679, 200	1, 718, 297	2. 3%
総資産比	5. 48%	4. 84%	
研究開発費	779, 806	807, 454	3. 5%
売上高比	4. 20%	3. 66%	

### 各セグメントのビジネスの状況

E社は、連結売上高の90%以上を自動車ビジネスが占めているため、自動車セグメントにフォーカスした事業戦略の分析を実施することとし、自動車セグメントの状況のみを示す。

### E社の自動車セグメントの状況

単位：百万円

	2012年	2013年	増減率
売上高	16, 994, 546	20, 419, 100	20%
営業利益	21, 683	944, 704	4257%
営業利益率	0. 13%	4. 63%	
総資産	12, 261, 814	13, 179, 741	7. 5%

### E社の地域別の販売台数

単位：万台

	2012年	2013年	増減率
日本	207	228	10.0%
北アメリカ	187	247	31.9%
欧州	80	80	0.1%
アジア	133	168	26.9%
その他	128	164	27.7%
合計	735	887	20.7%

#### 〈ビジネスの状況〉

会社の業績は日本及び海外の販売台数の増加に伴い売上高が増加している。

なお、商品の販売価額の大幅な変動は無かったものとする。

#### 〈事業リスク〉

E社は、公表している資料で以下の事項を事業リスクとして認識しているとする。

- ・自動車市場の競争激化と需要変動

世界の自動車産業のグローバル化に伴う競争激化

各市場での法規制や自然災害等の発生による事業運営の停滞

原材料価格や燃料価格の大幅な変動

- ・ブランドイメージの維持、発展

革新的で価格競争力のある新商品の投入

効果的な販売促進の実施

ブランドイメージの維持・発展

デジタル情報技術への依存による情報流出リスク

#### E社の内部的な財務情報から読み取れること

E社は2012年と比較して2013年の全体の売上高が約18%増加しているが、この主要な理由は自動車セグメントの売上高の増加によるものである。

地域別の販売台数では、2013年の日本およびアメリカの販売台数475万台は、全体の販売台数887万台の約53%を占めており、日本とアメリカが主要な市場であると考えられる。ただし2012年と2013年の販売台数の増加率としては、アメリカに次いでアジアが2番目となっており、かつ2013年の販売台数でも第3位であることから、アジア市場の重要性も高くなっていると考えられる。

## 公表されている統計・比率情報

### ① 2050 年までの地域別の人団推移

単位：百万人

出典：総務省統計局 2-2 人口上位 20 か国の人団推移から上位と日本を抜粋

	1950 年	2013 年	2030 年	2050 年
日本	84	127	117	97
インド	376	1, 252	1, 476	1, 620
中国	544	1, 386	1, 453	1, 385
アメリカ	158	320	363	401
インドネシア	73	250	293	321

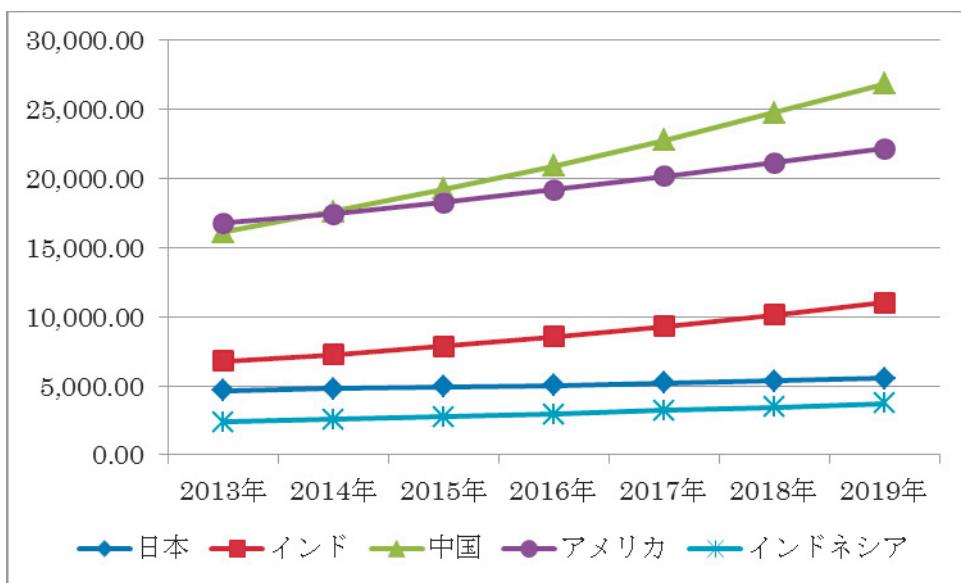
※2030 年以降は予想データである

### ② 人口推移の検討で対象とした国の人団 2019 年までの購買力平価 G D P の推移

単位：10 億米ドル

出典：世界の経済・統計 情報サイト IMF - World Economic Outlook Databases

	2013 年	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年
日本	4, 667. 55	4, 788. 03	4, 916. 76	5, 051. 65	5, 202. 41	5, 361. 35	5, 527. 62
インド	6, 775. 97	7, 277. 28	7, 882. 59	8, 550. 04	9, 304. 36	10, 139. 14	11, 044. 01
中国	16, 149. 09	17, 632. 01	19, 230. 29	20, 933. 27	22, 780. 30	24, 756. 23	26, 867. 43
アメリカ	16, 768. 05	17, 416. 25	18, 286. 69	19, 196. 55	20, 168. 83	21, 157. 73	22, 147. 84
インドネシア	2, 388. 95	2, 554. 31	2, 744. 41	2, 958. 29	3, 200. 22	3, 463. 97	3, 747. 61

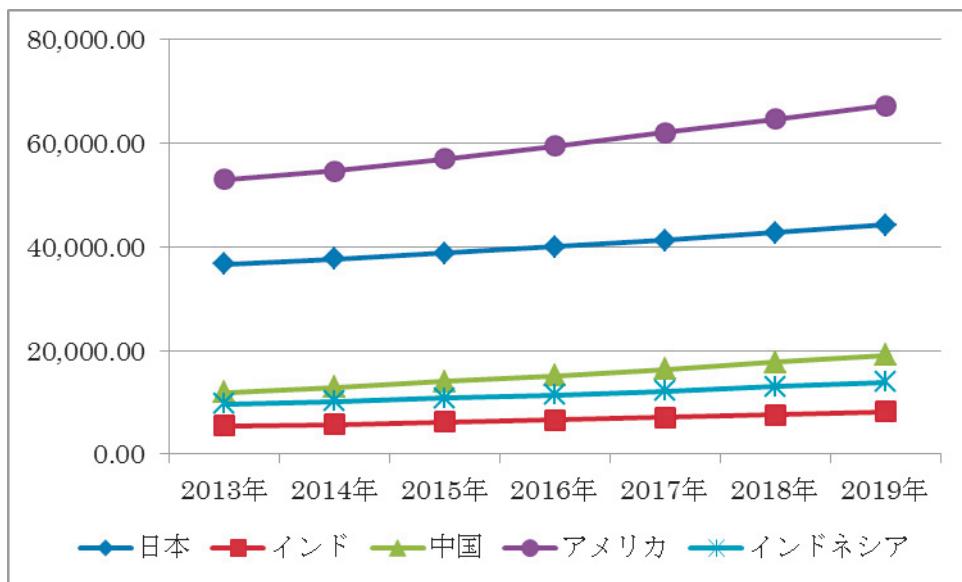


③ 人口推移の検討で対象とした国の2019年までの人口1人当たりの購買力平価GDPの推移

単位：米ドル

出典：世界の経済・統計情報サイト IMF - World Economic Outlook Databases

	2013年	2014年	2015年	2016年	2017年	2018年	2019年
日本	36,653.92	37,683.05	38,797.39	39,983.05	41,318.77	42,745.66	44,258.58
インド	5,449.82	5,777.02	6,176.29	6,612.26	7,102.19	7,638.90	8,212.59
中国	11,867.67	12,893.43	13,992.69	15,156.58	16,412.42	17,747.89	19,166.25
アメリカ	53,000.97	54,678.17	57,045.46	59,502.61	62,118.42	64,749.37	67,348.00
インドネシア	9,634.66	10,156.71	10,759.18	11,434.59	12,195.79	13,015.29	13,883.05

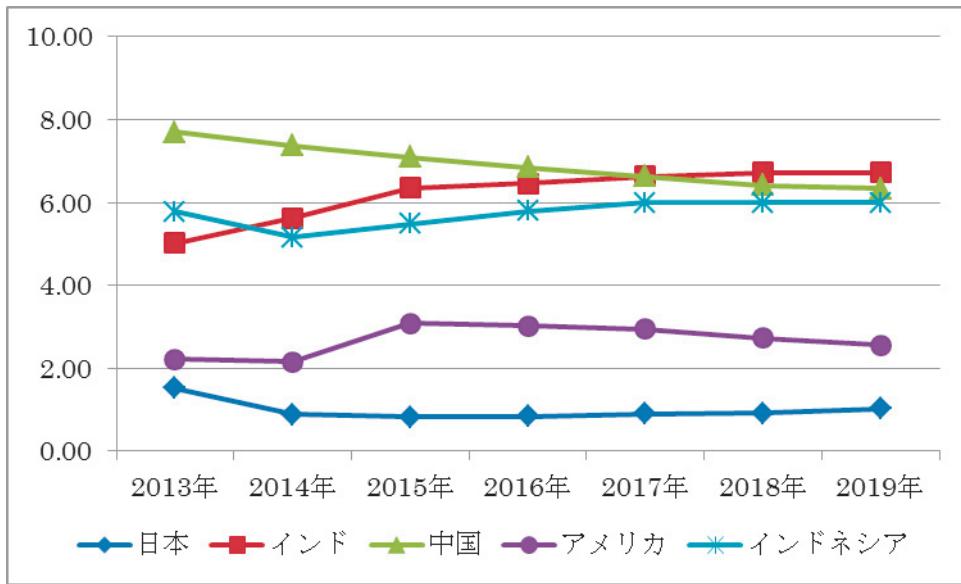


④ 人口推移の検討で対象とした国の2019年までの経済成長率(実質GDPベース)

単位：前年度比の変動率

出典：世界の経済・統計情報サイト IMF - World Economic Outlook Databases

	2013年	2014年	2015年	2016年	2017年	2018年	2019年
日本	1.52	0.89	0.83	0.84	0.91	0.92	1.02
インド	5.02	5.63	6.35	6.46	6.63	6.72	6.72
中国	7.7	7.38	7.09	6.84	6.63	6.42	6.33
アメリカ	2.22	2.15	3.09	3.03	2.95	2.73	2.56
インドネシア	5.78	5.16	5.49	5.80	6.00	6.00	6.00



※ 購買力平価GDPとは、異なる通貨で計算されるGDPから為替変動の影響を除いて同じ基準で比較出来るように算定されたものである。一般的にGDPは経済力を表す指標として使用されており、当研究報告書では、購買力平価GDPを、同じ基準で経済力を比較するために使用している。

#### 公表されている統計・比率情報から読み取れること

基本的に、自動車の使用者である人口が増加すればするほど市場が拡大すると考えられる。資料①から、今後の人口の増加は、中国、インドをはじめとしたアジアが上位になると予想される。

各国の購買力平価GDPを見ると、資料②から中国が非常に伸びていくよう見えるが、資料③から人口1人当たりの経済力はアメリカ、および日本のはうが高い。自動車は食料品や衣料品などに比べれば価格が高く、ある程度の経済力がないと購入できないと考えられることから、アメリカや日本のはうが自動車を購入出来る層が多い可能性がある。

#### ビッグデータから試るE社ビジネスの今後のテーマ

##### ①日本・北アメリカでのマーケットシェアの維持・拡大

日本・北アメリカでのブランド力・販売力の維持・拡大を基礎とした、研究開発活動の強化及び今後のマーケット拡大が見込まれるアジア市場へどのようにブランド力・販売力を拡大していくかがポイントとなる。

日本・北アメリカでは爆発的な人口増加が見込まれない状況ではあるが、資料③から、高い経済力を有しているため、製品の単価を維持するとともに、獲得した対価を基に技術革新を繰り返していくことが重要となる。

## ②アジア地域への事業の拡大

E社の日本・北アメリカでのブランド力・販売力を人口増加に伴う購買力の拡大が見込まれるアジア地域にどのように拡大していくのかが重要となってくるが、アジア地域の中で大きなウェイトを占める中国・インドでは日本・アメリカ等に比べ購買力平価GDPは低く、日本・アメリカと同様の販売価格では販売台数が増加しないことが予想される。その為、品質を維持したうえでのコスト削減及び現地生産・現地販売の地盤の確立が重要課題となると考えられる。

中国・インドでは不透明な政策運営や、中央・地方政府の不統一性、経済法規制度の未整備、運輸・電力等のインフラ整備、品質管理の困難性、代金回収リスク、優秀な人材の確保の困難性等の種々の問題があるが、中国・インドの1人当たり購買力平価GDPは2013年から2019年にかけ150%超の増加が見込まれており、資料④からも今後の市場規模の拡大が確実に見込まれる。

短期的なリスクを負ったとしても長期的な投資を行う価値は多分にあると考えられる。

### 【参考文献】

- 高安美佐子(著、編集)、田村光太郎(著)、三浦航(著)「学生・技術者のためのビッグデータ解析入門」日本評論社、2014年6月10日
- 朝野熙彦「ビッグデータの使い方・活かし方—マーケティングにおける活用事例」東京図書、2014年1月

## 2. 監査にかかるデータ活用の可能性 - 巨大容量仕訳データの分析-

### (1) 監査における仕訳データの分析（仕訳テスト）

経営者による内部統制の無効化から生じる不正リスクを識別・評価し、対応手続を実施するために、仕訳入力に関して、プロセスの理解、内部統制の整備状況の評価、重要な虚偽表示リスクの評価、内部統制の運用状況の評価、及び実証手続を実施することが求められている。特に仕訳に係る実証手続を実施する際に、網羅的な仕訳の詳細テスト（仕訳テスト）が重視されてきており、この仕訳テストを実施する際に、被監査会社から仕訳データを入手し、C A A T (Computer Assisted Audit Techniques、コンピュータ利用監査技法) を用いて仕訳テストを実施することが今日では一般的となっている。仕訳テストは、経営者等による内部統制の無効化による不正な仕訳投入がないことを確認するために実施されるため、母集団に対して網羅的に実施することが重要であるが、大量の仕訳に対して手作業で不正な仕訳の有無を検証することは極めて困難であり、仕訳データを入手してコンピュータを利用して分析することによりはじめて母集団に対して網羅的に仕訳の検証を行うことが可能となる。仕訳テストの実施方法としては、例えば、1年間の全仕訳データを入手の上、勘定科目毎に集計して試算表と照合して網羅性を確認した上で、その全仕訳データの中から、手作業で入力された仕訳か自動登録された仕訳かの区別、仕訳の入力された日付や時刻、仕訳の入力者及び承認者、勘定科目のパターン、金額、伝票の種類、仕訳入力画面の種類、等の切り口で不正な仕訳のシナリオを検討し、特定の不正入力の可能性がある仕訳を抽出の上、検証するような方法がある。

### (2) 仕訳データ量

小規模の企業であれば、年間の仕訳データ量がさほど大きくないことが多い、そのような場合には、汎用的ビジネスソフトウェアである Access や Excel 等を使用して分析を実施することも可能である。中～大規模企業になるとデータ量が大きくなることが多く、数十万～数百万レコードの仕訳明細数になると Excel や Access 等による分析は困難となり、C A A T 専用ツール（A C L、I D E A といったC A A Tのために開発・販売されているソフトウェア）を利用するが一般的である。

しかし、巨大企業になると仕訳の明細数が年間合計で1億レコードを超えるような会社もあり、データ容量も 20GB を超えるような場合もある。このような巨大容量データになると、C A A T 専用ツールでも分析が困難となってくることがある。データ量が巨大になると、C A A T 専用ツールを使用しても一つの分析を行うのに数時間～数十時間以上かかるようになることがあり、複数のロジックを組み合わせてまとまった処理を実行しようとすると、翌日になっても、あるいは数日たっても処理が完了しないこともある。

監査における仕訳データ分析は、はじめから厳密な分析・抽出の条件が確定していないことも多く、試行錯誤を伴うことが多い。例えば、ある特定の抽出条件でデータを抽出した際に、あまりにも多くの件数が抽出されてしまう場合がある。このような場合、抽出されたものすべてに対して監査手続を実施することは困難であり、別の抽出条件により再抽出したり別の抽出条件を追加したりして、よりリスクが高いと考えられる取引を抽出することになる。

このような試行錯誤を繰り返す場合、一つ一つの処理が短時間で終わらないと、分析の

際の思考が妨げられ、有効な分析が行えなくなる恐れがある。ある処理を実行し、結果が翌日になって出て、更に分析条件を少し変えて実行してその結果がまたその翌日にならないと出ないということになると、試行錯誤しながらあまり多くの分析条件を試すのは極めて困難である。もしこのような処理が数秒～数十秒で終われば、実行結果を見ながら監査人の思いつくままに様々な分析を実行し、その中から有効な分析を選択することが可能になる。大容量のデータを短時間で分析することは、時間短縮の効率性だけでなく、分析の有効性を高めることにもつながることになるのである。また、そもそも処理が現実的な時間内に完了しない場合には、事実上監査における仕訳データの分析は不可能となる。

### (3) 巨大容量の仕訳データ分析方法

前述のとおり、巨大なデータ量の仕訳データの場合は、汎用ビジネスソフトウェアやC A A T 専用ツールでは分析が困難となることがある。このような巨大容量データに対する分析の場合には、①D B M S (Database Management System、データベース管理システム)を使用する方法や、②B I (Business Intelligence、ビジネスインテリジェンス)ツールを使用する方法等が考えられる。

#### ① D B M S (データベース管理システム)を使用する方法

D B M S とは、情報システムのデータベースを構築・運用する際に使用されるソフトウェアであり、「S Q L」と呼ばれるデータベース操作言語を使用して操作するのが通常である。例えば、企業の会計システム、販売システム、在庫購買管理システムといった業務システム等においては、D B M S を使用してデータを管理していることが通常である<sup>3</sup>。D B M S は大容量のデータを記録・保存・照会することを前提としているため、D B M S を使用することにより巨大容量の仕訳データの分析も可能となる。ただし、D B M S は前述したS Q Lというデータベース操作言語を使用してデータの分析や抽出を行う必要があるため、S Q Lに精通した者でないと分析は困難である。また、D B M S であっても巨大容量のデータになると、分析が完了するまでに一定の時間がかかる場合が多い。

(なお、D B M S を使用する分析の場合、(1)業務システムのD B M S とは別のD B M S を準備して分析する方法と、(2)業務システムのD B M S を直接操作して分析する方法、がある。(2)業務システムのD B M S を直接操作して分析する方法の場合、誤った操作によりデータを破損してしまうリスクや、負荷の高い処理を実行した際に業務システムの処理速度が低下して業務運用に支障を来すリスクがあるため、留意が必要である。)

#### ② B I ツールを使用する方法

B I ツールとは、大容量のデータ分析のために特化したソフトウェアであり、ビッグデータを高速で処理することが出来るツールである。「第5章 ビッグデータ分析ツールの特徴」で紹介されているツールがB I ツールの例である。一般的には、B

<sup>3</sup>業務システムを使用してデータを登録・加工・照会する際に、業務システムのアプリケーションプログラム内に記述されたS Q Lが裏側でデータベースを操作する。

B I ツールは D B M S よりも更に高速で分析することができ、かつ S Q L やプログラミング等の知識がなくとも分析が可能である。今日においては、B I ツールは様々な会社から販売されており、比較的安価で購入できる小規模のものから、大規模かつ非常に高額なものまで様々である。B I ツールを使用すると巨大容量の仕訳データの分析も可能となるが、B I ツールを購入するための投資が必要となる。

また、B I ツールに限った話ではないが、ツールを使用して分析を行う場合、処理速度はソフトウェアだけでなくハードウェアの性能にも大きく依存する。特に近年増えてきたインメモリ(高速なメモリ上にすべてのデータを読み込ませて配置して分析を実施し、メモリより速度の遅いハードディスク等の外部記憶装置を使用しない)のB I ツール製品の場合には、メモリの容量等に大きく依存するため、高速な分析を実施するためにはハードウェアに対する多額の投資も必要となる場合がある。

#### データ分析ツールの比較

ツールの種類	対象データ容量	操作	費用
汎用ビジネスソフトウェア (Excel、Access 等)	小容量の場合に有効。扱えるデータ容量に制限あり。	分析の際には関数やマクロを使用するため、これらの知識が必要。	◎ (比較的安価)
C A A T 専用ツール (A C L、I D E A 等)	大容量で高速な分析が可能。 ただし、巨大容量の場合には分析は困難となる。	あらかじめ監査で使用する分析機能が組み込まれているため、複雑な関数やマクロの知識がなくとも分析可能。ただし、ある程度高度な分析を実施するためには関数やマクロの知識が必要となる。	○
D B M S	大容量でより高速な分析が可能。	S Q L の知識が必要。	△
B I ツール	巨大容量でも高速な分析が可能。	ツール毎の操作方法を習得する必要がある。	△～× (比較的高価)

#### (4) B I ツールを利用した巨大容量の仕訳データ分析

実際にB I ツールを使用して巨大容量の仕訳データを分析した事例を紹介する。

ある被監査会社では、年間合計で 1 億レコードを超す仕訳明細があり、同一システム内の自動仕訳や他システムからのインターフェースにより自動登録される仕訳を除いた、システムに手入力されたマニュアル仕訳だけでもおよそ 1,500 万レコード以上あった。巨大容量データのため、1 年間の仕訳データを一度に会社の会計システムから抽出して分析することが困難であり、被監査会社に監査用の仕訳データ抽出プログラムを作成してもらい、この抽出プログラムを使用して月次決算が締まった後に毎月 1 か月分のデータを抽出していた。それでも 1 か月のデータ容量は 2 GB 程度もあり、C A A T 専用

ツールでの分析は極めて処理に時間がかかるため、不可能ではないものの非常に困難であった。

このため監査チームは、データをD V D - Rにコピーして監査事務所内のデータ分析専門チームに渡し、四半期毎に分析を依頼していた。データ分析専門チームでは、多くの被監査会社に共通的に適用できる定型的な分析のための専用プログラムをD B M S上に構築して使用していたが、定型的な分析以外の各監査チーム固有の分析を依頼されることもあり、その場合にはD B M S上でS Q Lを使用して分析を実施していた。

高性能なサーバ上でD B M Sを使用しているため比較的高速に分析が可能であったが、当該被監査会社の仕訳データ量は膨大であったため、四半期毎に分割した分析でもかなりの時間がかかっており、監査チームがデータ分析専門チームにデータを渡してから、分析を実施して結果合を報告書としてまとめて、その結果の報告を受け取るまで1か月以上かかっていた。分析結果を確認し、分析や抽出のシナリオを精緻化若しくは変更して再分析を行いたい場合も多かったが、データ量が大きすぎたため一度実施した分析の調整と再実行することは極めて困難であり、なかなか監査チームが思うような分析を行うことができないでいた。

そのため、巨大容量のデータの分析が可能なB Iツールを導入し、監査チーム自身で分析を実施することを検討した。B Iツールの中には比較的安価なものもあり、データ分析チームに支払っている金額よりも安価な金額で購入できるB Iツールもあったため、それを購入して分析を実施することとした。B Iツールを購入し、監査法人内で現在使用していないサーバを安価で調達し、分析環境を構築した。分析の実施に当たっては、当該被監査会社の監査に関与しているI T専門家がB Iツールを使用する環境の構築とツールの操作を担当し、監査チームとI T専門家が協業してどのような分析を行うかの分析のシナリオを検討し、分析を実施した。

分析のシナリオ検討に当たっては、実際にB Iツールを操作しながら意見を出し合い、その場で思いついた切り口や内容で操作を実施して、結果を確認しながらシナリオを精緻化していく。数か月分の数千万レコードの仕訳明細であっても、分析を実行して数秒～数分で結果が返ってくるため、その場で試行錯誤の繰り返しができた。

例えば、仕訳データの「摘要欄」に、ある特定の用語や記号が含まれている仕訳を抽出するためのテキスト検索を実施する場合、半年分の約5千万レコードのデータを対象としても数秒～数十秒で結果が表示されるため、いろいろなキーワードを気軽に試すことが可能であった。このようなテキスト検索を同様のデータ量に対してC A A T専用ツールを使用して実施した際には、数時間が経過しても結果が返って来ないこともあるため、いろいろなキーワードを試すことは困難であった。

### 実施した分析の例

分析項目	分析内容
キーワード検索	マニュアル入力仕訳について、仕訳ヘッダと明細データの「摘要欄」に、特定のキーワードや記号が使用されているものを抽出した。どのようなキーワードや記号に不正の兆候があるのかが事前にはわからなかつた

	ため、他社における実際の不正事例や過去の監査の経験等に基づいて様々なキーワードや記号を試した。
伝票登録者毎の傾向分析	マニュアル入力仕訳について、伝票登録者毎に、仕訳の入力件数、金額、勘定科目、入力時間、入力日、反対仕訳の入力数、伝票種別、その他様々な切り口で傾向を分析し、異常値の有無を検討した。分析の切り口を決める際に試行を繰り返した。
通例でない仕訳パターンの分析	マニュアル入力仕訳について、仕訳の貸借の勘定科目の組合せのパターンを分析し、組合せの中から年間の件数が少ないものや、通常は想定されない貸借の組合せの仕訳を試行を繰り返しながら抽出した。
通例でないサイトでの入金・支払の分析	通例でないサイトでの入金や支払には循環取引やその他不正の可能性があると考え、債権及び債務の計上日、計上入力日、入金・支払予定日、実際の入金・支払日、入金・支払入力日付等の様々な日付の相関関係を試行しながら分析し、異常性の有無を確認した。

ただし、単純な検索等であれば比較的短時間で結果が返って来たが、1年間の自動仕訳等も含めた1億レコードを超える全仕訳データに対してある程度まとまった処理を行おうとすると、なかなか結果が返って来ないことがあった。分析のために準備したサーバがそれほど高性能でなかったこともあり、より大容量なデータに対して複雑な分析を実施する場合には、より高性能なハードウェアの導入が必要であることが分かった。

#### (5) 仕訳テスト以外の監査の領域へのBIツールの適用の可能性

今回は「仕訳データ」のみを分析対象としたが、企業の業務システムの中には他にも様々なデータ（受注データ、出荷データ、請求データ、債権データ、発注依頼データ、発注データ、入庫データ、債務データ、在庫データ、原価データ、その他）が含まれている。これらのデータは仕訳データ以上に膨大であるため、データの入手は技術的にも簡単ではなく、どのようにデータを入手するかが課題となる。しかし、いったんデータさえ入手できれば、高性能なBIツール等を使用して比較的高速に分析が可能であり、監査の有効性をはるかに高めることが可能になると考えられる。

また、企業の業務システムの中にあるデータのみならず、その他の企業内外の様々なデータが利用できれば、更に監査の可能性は広がると思われる。

どのようにしてこれら企業の内外のデータを入手し、どのようなシナリオに基づいて分析するのかが今後ますます重要になると考えられるが、その前提として、監査においても最新のデータ分析のテクノロジーを積極的に採用していくことが望まれる。

なお、最新のテクノロジーを利用する際には、分析用ソフトウェアやそのソフトウェアを高速に動作させるためのハードウェアが必要になるため、これらに対する理解と投資が必要となってくる点が課題である。

## 第5章 ビッグデータ分析ツールの特徴

ビジネス上「ビッグデータ」という言葉が使われ始めてから久しい。ただし、ビッグデータという言葉を聞くものの自社でいかにビッグデータ分析し活用すべきかに迷う企業が少なくないと考える。また、私たち公認会計士についても、ビッグデータが監査や経営指導等の中でいかに役立つか、活かすべきかを考えている会員は少数派と思われる。

このように、ビッグデータに対する理解等が十分でないと考えられる状況もあるが、「第4章 ビッグデータの経営面の活用」で紹介したとおり、企業がビッグデータを活用して経営や顧客サービスに活かす事例が増えている。また、多数の企業での利用が期待される領域においてはビッグデータ分析ツールが開発されITベンダ等から提供されている。本委員会では、これら市販されているツールの用途や特徴を理解することはビッグデータ分析をより深く理解する助けになる、会員が行うビッグデータ活用に関する経営指導やビッグデータ活用監査等参考になると考えた。そこで、販売されている複数のビッグデータ分析ツールの用途や特徴を調査することとした。本章では、調査したツールについて紹介する。

また、BIツールと設備保全のためのツールについては、詳細な紹介を行うこととする。BIツールは企業の事業活動で得られた販売・購買・人事・財務等の情報などを分析し、グラフ等を使い経営意思決定を支援するもので、利用する企業が多くなっている。BIツールを理解することは、企業の経営分析のための基盤作りを支援する際に役立つと考えるためである。また、監査においてBIツールから情報を取得し証拠とすることがあるため監査を考えた場合にも理解することが有用と考えたためである。設備保全のためのツールは設備の故障を設備に取り付けたセンサ等から得られる振動等情報を使って予測し、設備故障による経済的損失を軽減するものである。設備保全分野の業務に役立つものであるが、ツール活用の前提や考え方は監査や不正調査の場面において参考になる点もあると考えたので紹介する。

### 1. ビッグデータ活用例とツールの役割等

#### (1) ビッグデータと分析ツールの関係

今日の経済社会におけるビッグデータビジネスの発展は、データ分析技術の発展すなわち分析ツールの発展と言えることが出来る。

1990年代、科学分野や気象分野において「ビッグ」データは存在していた。すなわち、当時のスーパーコンピュータが数年を要しても処理できないほどの膨大なデータが蓄積されていたのである。

これが2000年代からのIT技術の進歩により徐々に分析ツールの処理可能なデータ量が拡大し、処理スピードも格段に速まった。ビッグデータ活用の障壁が下がったことで、研究施設や企業において広くビッグデータ分析が行われるようになった。

例えば、小売店において顧客の購入データ、来店データ、性別・年齢などの個人情報データを組み合わせて消費動向を分析し、将来予測のもと商品の仕入れを行うことは現在では珍しいことではない。情報を収集する媒体には、ポイントカードがよく用いられ

る。

企業だけが恩恵を受けているだけではない。気象庁の天気予報には、スーパーコンピュータによるビッグデータ分析が用いられている。分析対象は膨大な観測データである。今まででも観測データの分析自体は行われてきたが、近年ではより詳細な分析が可能となり、予報の的中率が上がることで我々利用者も恩恵を受けている。

## (2) 分析ツールの役割

ビッグデータ分析には様々な種類があり、同じようにビッグデータ分析ツールも多種多様に存在する。様々な形態の分析ツールがある中で共通点となるのは、「ビッグデータ分析ツールは、ビッグデータを有効に活用するために、データを適時に処理・分析することが不可欠である。そして、その処理・分析は、ビッグデータの活用目的によって最適化されている。」ということである。分析ツールは手段に過ぎず、活用目的を果たすのに有用な情報を提供するのが役割である。反対にいえば、どれだけ大量データかつ高精度の分析を行っても、活用することができなければそのツールの価値は無いに等しいであろう。

上述の例にあるとおり、ビッグデータの活用事例は幅広い分野にわたっている。その中でどのような活用目的と分析ツールがあるのか紹介したい。一例をまとめると次表のとおりである。

主な利用組織	活用目的	ツールの役割
事業会社全般	経営管理	抽出の難しいE R P <sup>4</sup> の膨大なデータを解析し、ユーザが分析を容易に行えるようシステム画面上でレポーティングを行う。 ※「2. ビッグデータ分析ツールの特徴－経営管理」で詳しく紹介する。
製造業	予知保全	工場設備に付けたセンサにより収集される膨大な稼動監視データを用いて、設備の故障予測を立てる。 ※「3. ビッグデータ分析ツールの特徴－予知保全」で詳しく紹介する。
商業施設	マーケティング	商業施設に訪れる客の持つスマートフォンアプリにより収集したログ（行動履歴や購買履歴、位置情報、外部環境）を分析し、消費者のニーズを推定する。
教育機関	学習指導	問題集などの教材を問題単位でデジタル化し、生徒の学習の記録や行動履歴データを分析した結果、生徒の学習傾向を明らかにし、個人毎に適した教育を行う。

<sup>4</sup>E R Pとは、製造・物流・販売・調達・人事・財務管理会計などの基幹業務に関する統合型（業務横断型）基幹業務システムである。

事業会社全般	顧客満足度の向上	サイトに訪れたユーザのページ遷移や検索キーワード、滞在時間などを分析し、リアルタイムでユーザ毎に最適な情報を提供する。
小売業	製品開発	顧客データやPOSデータなどのマーケティング情報、受発注情報や財務情報などを分析し、顧客ニーズを推定する。

いずれもそれぞれの活用目的に合わせてビッグデータ分析ツールが開発されている。次節から、2つの分析ツールを詳しく説明する。

### 【参考文献】

牧野武文「進撃のビッグデータ」マイナビ、2014年6月

新日本有限責任監査法人「ビッグデータ 企業の競争と業務に変革を起こす」2014年

## 2. ビッグデータ分析ツールの特徴 ー 経営管理

今日多くの企業は、他社開発のERPパッケージを中心とし、それに必要な設定と修正を加え、更にPOSや顧客満足度を向上させるCRMなどの必要なシステムを組み合わせた全体システムにて事業活動を維持・運営している。子会社展開や新事業展開の際、既存システムを適用せず実情に即したシステム構築を行う場合や、M&Aによる事業拡大を行う場合には、複数のシステムが混在する複雑な状況が生じることとなる。

この全体システムには、日々刻々と事業活動のデータが蓄積され、事業規模の拡大・多角化・国際化と共にその量は爆発的に増大する。

このデータを活用し、適時適切な経営意思決定を可能とさせる分析ツールを検討することとする。

### (1) 分析ツール大別

データを、収集・蓄積・分析・報告し経営意思決定に役立てるものがBI(ビジネスインテリジェンス)である。BIは、提供者により2種類のものがある。一つは、ERPパッケージ提供者がそのERPに同梱若しくは追加導入することができるものとして提供するものである。他は、ERPパッケージの提供元とは異なる企業が開発・販売するものである。

後者は、報告作成を容易なものとすることに主眼を置いている。その反面、収集・蓄積されるデータの信頼性確保やデータ保全には相当の努力を要し、更にトラブルが生じた際、その解決のため複数のITベンダと相談・交渉などを行わなければならないこととなる場合がある。

複数のERPや他のシステムが混在する複雑な状況下では、後者のBI(本章では以下「独立BI」という。)を活用することが、費用対効果・弹力的運用で優れていると考えられるので、以下ではこれについて記述する。

## (2) 独立B Iの機能

独立B Iは、データを収集・変換し独自の保管場所に保管・蓄積し報告を作成する3つの機能から概ね構成されている。保管・蓄積をまず理解し、その後収集、次いで報告の順に進めることができ理解しやすいであろうと思われる所以、その順で説明する。

### ① 保管・蓄積

独立B Iは、多種多様な分析を高速で効率的に行えるようにするために多次元・多階層の構造（以下「キューブ」という。）をつくりデータを保管する。データ配列・内容は、データ抽出元のERPパッケージ、POS、CRM等とは必然的に異なるものである。

このキューブに対するアクセス権は、データ注入とその使用の2側面から検討する必要がある。まず、データ注入については、キューブはERPとは独立の存在となることから、ここに直接データを注入することも可能となる。データの改ざんや変更・削除は、報告される情報に致命的な誤りを与える恐れがあり、独立B Iの存在意義そのものを失わせるものとなる。このため、データ入力・変更・削除のアクセス権設定は重要な問題となる。独立B I選定に際しては、慎重な検討を要し且つその運用面での監視も必要となる。

次にその使用に関するアクセス権には、外部に対してはもちろんのこと企業階層により開示される対象が制限されるべきものが多く含まれていることから、データ利用アクセス権の設定の可否は、独立B I選択の重要な点の一つとなる。多くの独立B Iは、ユーザ毎にこのアクセス権設定を可能としている。

### ② 収集・変換

大企業に多く採用されているERPパッケージでは、例えば顧客情報を幾つかのテーブルに分けて保管する、一つの会計仕訳情報を分けるなどと情報保管場所は複雑に細分化されている。更に、保存されたデータそのものを抽出すると意味の無い数字の羅列にしか見えず、多通貨対応のため金額は小数点以下2桁で記録され、日本語などは独自の言語変換が行われているなどの特徴を有している。

データ収集に当たっては、まず必要なデータが何処にあるのかを特定する必要がある。この作業は、極めて煩雑で労力を要するものである。独立B Iの中には、特定のERPパッケージに対し有効なデータ収集設定をあらかじめ組み込んでいるものがある。

次に収集したデータを、キューブに注入するため変換（並び替え、結合・分離、読み替え）する必要がある。前述の独立B Iは、この変換機能を併せ持っている。

他方、データ収集に際してはどの期間のデータを対象とするかが問題となる。データ収集量が多くなるほど抽出エラーを生じる可能性が高くなる。このため、前回抽出した以降に生成されたデータのみを収集対象とする独立B Iもある。

更に、複数のERPパッケージ、POS、CRMなどから抽出されたデータを紐付けられなければ、一連のものとして分析し報告することが出来ない。データの欠落や瑕疵も同様の結果を生む。独立B Iの中には、この点に着目し、その補完や修正を容易にする機能を搭載したものもある。

独立B I導入に先立ち若しくは導入時に、データ抽出元の諸マスタや重要取引データの整合性や過不足のないことを十分に検討し、独立B Iにより経営意思決定に資する報告が得られる素地を整えることが重要である。

### ③報告機能

報告は、保管・蓄積された膨大なデータからなにを読み取るかである。多くのユーザが必要とするであろう報告形式が、あらかじめ200以上組み込まれている独立B Iもあるので、ユーザは特段の習熟無しに容易に報告を得ることが出来るツールも存在する。更に、報告からその基礎となる個別データまで遡る機能を有しているものもある。他方で、ウェブやExcelでの報告入手を可能とするもの、Excelなどにて保管・蓄積したデータを直接対象として分析を行い、独自の報告を得ることも可能とするものがある。

## (3) 活用例

ここで、ヒアリングした内容に基づき2つの活用事例を挙げる。

### ① A社

10億件以上の明細データを蓄積するに至り、ERP提供者から購入したツールでの分析は処理速度が急速に劣化した。更に同ツールは操作が複雑なため、多くのユーザが使いこなせずその利用も低下した。他方、POSデータなどのデータとも併せて分析したいとの要請があったが、これに対応することが出来なかった。

独立B Iの導入により、操作が簡単で処理が早いことからユーザが拡大し、POSなどのデータとも併せて分析・報告できるようになり各経営階層の情報量が増加し、適切な経営意思決定が迅速に行えるようになった。

### ② B社

日本の他世界4拠点でそれぞれ稼動しているERPは、拠点毎に独立しているため全体を把握するには多くの時間を要していた。

B Iの導入により、拠点毎のERPデータを統合一元化し、一体のものとして分析することが可能となった。

東日本大震災時、重要サプライヤの被災によりパーツ供給が止まった際、B Iにてグローバルでのパーツ在庫及び所要量を確認、在庫に余裕のある拠点から他拠点への融通を行うなど、強力な情報基盤として活用された。

## (4) 導入等

最適なB Iの選定、機器の調達、設定、ユーザ教育等が導入に係る仕事となる。同時に、ERP等のデータを整理整頓する必要がある場合、B Iから有用な報告が得られるまでに相当の期間を要する場合もある。コストはその提供者により様々なので、初期コストは定額か、そのユーザ数等に応じ増加するものか、またメンテナンス・フィーはどのように計算されるのか等、十分な比較検討が必要となる。

## (5) 監査等での活用

B I は、比較損益計算書、部門別・品目別損益計算書、比較貸借対照表、貸借対照表残高・増減推移、債権債務年齢調べ、在庫受払、諸回転率やK P I と呼ばれる主要業績評価指標を報告する機能をあらかじめ有しているものもあり、監査等に有効な手段となる。しかし、B I は外部監査に活用するためには、まずそこに蓄積されているデータとE R P に蓄積されているデータが同一であることを確認する必要があり、I T 内部統制の検討や合計試算表などによる金額確認などを実施する必要がある点に留意する必要がある。

## 3. ビッグデータ分析ツールの特徴 － 予知保全

製品や生産設備の故障発生を予測し、稼動停止による経済的損失を未然に防止する予知保全分野においてビッグデータを活用しているビッグデータツールベンダーへのヒアリングを行った。以下、(1) から(3) まではヒアリングした内容を取りまとめたものである。

### (1) 予知保全分野におけるビッグデータ分析とは

ビッグデータ分析の各種分野の中でも、製品や生産設備の故障発生を予測し、稼動停止による経済的損失を未然に防止する「予知保全」分野は、近年、設備の劣化状態や性能の状態を基準として修理や交換の時期を決定する保全形態として、多くの企業に着目されている。

従来、製品や生産設備は、故障やトラブルなどの有事が生じた場合に「事後保全」ということでアフターサービスマーケットなど幅広く対応されてきた。そして、この「事後保全」を更に事前に予防するための施策として「予防保全」ということで定期メンテナンス等の措置を講じている会社も多い。ここで着目すべきことは有事が生じた場合の「事後保全」、有事に繋がらないための「予防保全」に先駆けて故障やトラブルを予知する「予知保全」という分野に注目が集まっているということである。

この「予知保全」分野においては、製品や生産設備に設置したセンサ等により、膨大な稼動監視データを集める収集し分析することにより、故障やトラブルの予兆を検出するビッグデータの新しい活用分野である。



### (2) 予知保全分野におけるビッグデータの診断プロセス

故障予測の予知保全分野においては、まず故障するとインパクトの大きい製品・設備に関して製品や生産設備にセンサを設置し、データの収集を行う。次に、収集したデータを解析することで、製品や機械の劣化や変調を発見する。このデータ解析段階においては膨大なセンサデータに独自のアルゴリズム<sup>5</sup>を加味し、ビッグデータを分析していくことが特

<sup>5</sup> アルゴリズムとは、数学、コンピューティング、言語学、あるいは関連する分野において、問題を解くための手順を定式化した形で表現したものを使う。

に重要となってくる。このアルゴリズムの設定には過去の経験や統計的パターン等の専門的パラメーターが考慮されることになる。

そして、製品や機械の状態を分析・評価するとともに、故障やトラブルといった有事が生じる可能性が高くなつた場合にはリスクを認識し、警告を発するようとする。そして最終的に製品や設備のユーザはこのリスク警告を十分判断した上で意思決定を行うことになる。

### (3) 予知保全分野におけるビッグデータ分析の留意点

故障予測の予知保全分野においては、設置したセンサからの情報の収集により故障の対応策を講じることが出来るが、この予知保全の分野は過去の有事の事象を評価分析したものであり、過去の有事のパターンに当てはまらないような事故やトラブルに対しては予測ができないという点で留意が必要である。また、予知保全の分野で事故等の回避のための対抗策が講じられた結果、それが障害となり次の有事の検出が遅れるという予期せぬ事象が発生する可能性が副次的に生じてしまうという点も留意が必要である。例えば機械にある一定の負荷をかけると故障が生じやすくなるというアルゴリズムがある場合に、これを回避するためにコンピュータ制御した場合には、また別の予期せぬ事象が副次的に生じてしまう可能性があるということである。

そのため、故障予測の予知保全分野においては、アルゴリズムの見直しを常に行い、フィードバックしていくことで発展させていくことが非常に重要である。

## 4. 予知保全分野におけるビッグデータ分析の監査への応用

### (1) 内部統制の観点から

先にも述べたとおり、予知保全分野においては、有事の際の「事後保全」、定期メンテナンスなどの「予防保全」、そしてそれらを予測する「予知保全」にも着目されてきており、これは内部統制における「予防的統制」、「発見的統制」という統制行為に応用できるのではないかと考えられる。例えば内部統制上の不備が生じたパターンを日付、時間、G P S、メールログ、出退勤記録といったデータに、過去の分析結果や経験等の独自の専門的判断というアルゴリズムを加味して、「予知的統制」という予防統制の前の予測が可能となる可能性がある。

この予知的統制で検証する内部統制の範囲やデータの例としては、以下のようなものがあると思われる。

#### ① 人的要因

残業時間、欠勤率、離職率、人事ローテーションが未実施であること、社内コミュニケーションの頻度等。

#### ② 組織的要因

事業部としてキャッシュフローがマイナスであるが、営業利益が生じている、業界全体の景気が悪化しているのに、なぜか売上が安定している、発注担当者、調達担当者、検収担当者の組合せがいつも同じ等。

#### ③ 取引上の特徴の要因

異常な時間帯のシステム入力、発注から検収までの期間が内容と比較して著しく短期間である等。

上記のような要因を内部統制の検証の対象範囲に含め、その分析により予防統制の前の予測が可能になるのではないか、と考えられる。

更に、このようにして監査対象とするべき内部統制の対象範囲を拡張することによって、例えば急に金回りが良くなる、毎晩飲み歩いており遅刻が増えた等の「個人的要因」がビッグデータ分析により早期発見できるようになると、より効果的な余地ができるようになる可能性があると考えられる。



## (2) 不正との関係

予知保全分野のビッグデータは不正との関係に応用できるとも考えられる。内部統制には限界があるため不正の発生を完全に排除することは難しいが、不正リスクを低減することに役立つものである。不正はひとたび起これば企業の信用・ブランドの毀損となる点で製品や生産設備の故障やトラブルと共通するところがある。不正は「不正のトライアングル」（機会、動機、正当化）が重なり合うことで発生する可能性が高まるという点でこの3要素を事前に予測できれば不正の発生を予知することも可能と考えられる。

例えばプロジェクト別の損益を担当者別に管理し、それに対するノルマがある場合や、それに対する業績運動型賞与が支給されるような会社においてはプロジェクト損益をよく見せたいという不正の動機がある。更に休日においてもプロジェクト管理システムの修正が可能である場合、人目に触れにくいという点で不正の機会が平日の場合よりも高まる可能性がある。また、プロジェクト損益の付け替えが恒常的に行われるような風土がある場合には「他の人も行っているので少しくらい大丈夫だ」という正当化理由があることになる。このような事象に対してはG P Sデータの分析や出退勤記録の分析、プロジェクト別損益データの分析が有効な可能性がある。不正発生という有事に対して、監査人の経験値といった独自のアルゴリズム（例えば、潔癖な人は不正を起こさない等）を加味することが出来るようになれば、ビッグデータを利用した監査がより有用なものとなってくると期待される。

このような分析の参考となる手法のひとつに、デジタルフォレンジックがある。デジタルフォレンジックとは「犯罪の立証のための電磁的記録の解析技術及びその手続」[警視庁ホームページ平成21年警察白書より引用]と定義され、近年犯罪捜査に使用される技術である。デジタルフォレンジックの具体的な手法の例としては、コンピュータや携帯電話のデータを作業用に複製し、メール履歴、S N Sの履歴、検索履歴、アクセスログ、U S Bデバイスの接続履歴、インストールされたソフトウェアの履歴、復元された削除データといったデータを検証することで、状況証拠を積み上げたり、直接証拠を突き止めること等がある。監査においても同様の手法を用いて、データ検証を行い不正の兆候を見つけることに応用できれば、不正の早期発見に利用できる可能性があると考

えられる。

【参考文献】

- 株式会社 KPMG FAS フォレンジック部門「企業不正の調査実務」中央経済社、2012年12月

## おわりに

本年度の委員会活動は、特に先進的と考えられるテーマを採用した、とてもチャレンジングなものであった。その中で、2点印象に残ったことがある。

1点目は、ビッグデータの活用は難しい理論が必要なわけではないと言う点である。今回の報告書からは割愛されているが、設備保全のツールベンダにヒアリングした際に、機械の設置場所に目をつけた分析を行っているという話を伺った。機械が密集している部分は熱がこもるので同じような壊れ方をするが、端のほうにある機械は左右に空間があるので、熱が逃げるため違う壊れ方をすることである。そのため、分類として、密集している部分にある機械と端にある機械という切り口で分け、それぞれ異なる分析を行うと、新たな発見ができたとのことである。よく考えれば、分析対象を性質に応じて分類して分析すると言うことは当然のことであり、監査の分析的手続でも階層化や母集団の切り分けなどの手法で行われているものである。書籍などでは複雑なデータ分析手法が紹介されていることが多く、それらが必要になる場面もあるが、実際に分析し、活用する場面で重要なのは、対象を観察し、理解することであり、これは監査にも通ずることであるという点が印象に残った。

2点目は、活用方法を考え出すのは常に人間であるという点である。先に挙げたヒアリングの中でもツールベンダの方が強調されていた点であるが、活用方法を考え出すツールは存在しないのである。言い換えれば、いくら大量のデータが存在し、分析するためのツールが充実していても、その内容を理解し、そこから何かを気づき、ひらめくのは人間にしかできないことということである。これは監査でも同様であるのではないかと思う。いかに、詳細なデータや監査証拠が存在し、それを分析するツールが用意されていても、不正、誤謬を発見するのは監査人のスキルにかかっている。今後は、ビッグデータやその分析ツールを監査に利用することがあるかもしれないが、そういった場合であっても、ビッグデータやそのツールが不正・誤謬そのものの発見方法を教えてくれるわけではない。その発見方法は監査人が自ら考え出さなければならないのだと感じた。

最後に、本研究報告書を取りまとめるに当たり、貴重なお時間を割いてヒアリングに応じていただいたツールベンダの皆様に、心から感謝を申し上げたい。

日本公認会計士協会東京会  
(平成 26 年度) I T 委員会

委 員 長	森 山 直 輝
副 委 員 長	曾 根 正 則
同	新 出 谷 崇
委 員	荒 井 覚
同	伊 勢 田 英 之
同	魚 橋 直 子
同	嶋 田 恵 介
同	嶋 矢 剛
同	高 野 英 之
同	高 橋 孝 治
同	寺 村 航
同	松 本 裕 治
オブザーバー	神 崎 時 男

(以上 13 名)

担当副会長	淺 井 万 富
担当常任幹事	田之倉 敦 司
担当幹事	倉 林 正 憲

---

東京 C.P.A.  
**公認会計士業務資料集 第55号 VI**

平成27年12月21日 印刷

平成27年12月25日 発行

編集 日本公認会計士協会東京会 調査研究部  
発行 日本公認会計士協会東京会  
〒102-8264 東京都千代田区九段南 4-4-1  
公認会計士会館  
電話 03(3515)1183  
調査課ダイヤル・イン

印刷所 大和メディアクリエイティブ株式会社  
〒162-0824 東京都新宿区揚場町2-20 大和ビル  
電話 03(3269)6351(代)

---

本書の無断複写・複製・転載を禁じます。