

データ分析の事業性の評価を行うプラットフォームの提案と考察

小栗 秀暢^{†1} 松井 くにお^{†1}

近年ビッグデータ分析市場の拡大が叫ばれているが、現実的には国内における事業会社での活用事例は少ない。ビッグデータ分析の現状の課題点を整理し、事業会社の参入を促すために、経済性を評価できるプラットフォームを提案する。

Proposal and discussion platform for the assessment of business data analysis

HIDENOBU OGURI^{†1} KUNIO MATSUI^{†1}

Although the market of big data analysis is expanded in recent years, there is few entry of companies. In this paper, I classify the subject point of big data analysis. And I propose a Bigdata assessment platform that can facilitate the entry of companies.

1. はじめに

近年のビッグデータの分析市場の拡大に合わせて、各社より分析結果による改善事例やマーケティング活用事例が発表されるようになった。

だが@ITの調査によると、ビッグデータデータについての『その意味や具体的な取り組み内容』となると、システム運用現場のITプロを含めても、まだ半数以上が「聞いたことがない」「言葉は知っているが意味は知らない」と答えた。[1]

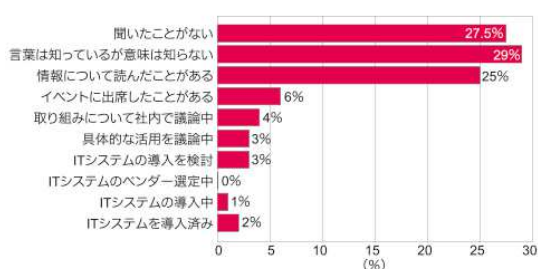


図1 @ITのアンケート結果より

実際に業務に『導入している』と答えたのは3%に過ぎない。データ分析による事業改善への活用は、ビッグデータの活用が求められる以前からサービス科学の分野で課題になっていたが、矢野経済研究所の調査によると、ビッグデータの市場規模は2020年度には現在の市場規模の5倍以上である1兆円を超えると予想されている。[2]

だが、この市場規模は『ソフトウェアライセンス、システム構築、ハードウェア、保守』等の市場規模であり、ビ

ッグデータの解析によって得られる事業の広がりや収益性とは結びつかない。

2012年に発表された総務省のActive JAPAN ICT戦略では「データの利用事業者及びその支援事業者からなる市場において、今後、少なくとも10兆円規模の付加価値創出及び12~15兆円規模の社会的コスト削減の効果がある」[3]としており潜在的な市場としては認識されているが、国内におけるデータ利用ベンチャー企業の上場など、顕著な影響は未だ見えておらず、データ分析による事業が活発に開始されているとは考え難い。

そのため、本稿ではビッグデータの利活用を活発化するため、多くの企業がデータ分析への参入し、データの流通を促すようなプラットフォームの構造について提案する。

2. ビッグデータ分析推進の課題

ビッグデータの今後の成長性は多くのレポートにて報告されているが、同時にその普及を妨げる要素や課題についても多くの点が指摘されている。

情報通信審議会は、図2のような形で問題点と採るべき施策をまとめている。これは政府から見た視点であるため、網羅的に問題点が整理されているが、実際のデータを活用する企業としての視点については、具体的な問題として言及されていない。

^{†1} ニフティ株式会社
NIFTY Corporation

- ◎情報通信審議会「知識情報社会の実現に向けた情報通信政策の在り方」
1. 官民のデータのオープン化、各種データを横断的に利活用できる環境の整備
 2. 多種多量なデータをリアルタイムに収集・伝送・解析等する技術やデータ秘匿化技術等の研究開発・標準化
 3. データサイエンティストの育成
 4. ビッグデータビジネスの創出に寄与するM2Mの普及促進
 5. ビッグデータの活用に関するICTの利活用を阻む規制・制度改革の促進
 6. 異業種・産学官の連携によるデータ資源の蓄積等ビッグデータの活用に関する推進体制の整備
 7. 外国政府等とのビッグデータの活用に関する対話の強化や、ビッグデータの活用に関する計測手法の確立

図2 情報通信審議会がまとめた問題点

また、城田[4]は、ビジネス上での著書の中での問題点として

1. プライバシー問題と国毎に異なる法的基準
2. データのオープン化/マーケットプレイス化に伴うデータ形式/語彙の不統一の影響
3. データ・サイエンティストの不足

同様に鈴木[5]はビッグデータ市場に関する問題点を以下のようにまとめている。

1. 人材不足
2. プライバシ・機密情報への抵触
3. データの精度の悪さや誤用、不適切利用

堀内[6]は以下のように語っている

1. 期待する成果が何かを見極め、どのようなタイミングで、どのようなデータが必要になるのかを議論する。
2. データの入手コストに見合う成果を自社が期待しているのかどうかを検討する。

これらの問題点は、それぞれの著者の観点から語られたものであるため、これらを大きく分類し、フレームワーク化したものを図3に示す。

これら問題点は、それぞれの著者の観点から語られたものであるため、これらを大きく分類し、フレームワーク化したものを図3に示す。

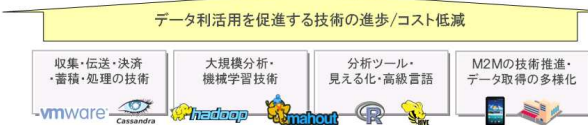
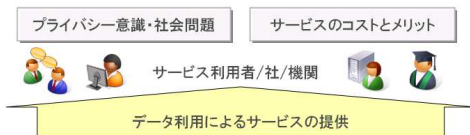


図3 ビッグデータ利活用推進のフレームワーク

ビッグデータの利活用推進に関する意見発信者を大きく区分すると、法的側面、経済的側面、公的な側面、という3つの論点から指摘されているものが多いため、その3

点を 安全性(法・モラル)/効率性(企業・収益)/有益性(利便・公益)という3つの軸とした。また、縦軸はそれぞれの 技術(手法)/サービス(流通)/企画(事業化)に対して発生している問題点として整理した。上と下には、それを実現するための関連技術と、実現すべき生活スタイルを配置した。

3. ビッグデータ分析の論点調査

このフレームワーク内での論点/問題点についての文章がどのように流通しているか、定量的な目安を作るべく調査を行った。

図4は、Googleの検索エンジンを用いて、上記の利活用推進のフレームワークについての検索結果数の量と、全体からの出現率について、日本語と英語の両方で調査したものである。

○Google検索エンジンにて課題ワードを検索した結果(2012年9月6日実施)

検索ワード	出現数(約)	日/英比
ビッグデータ	18,200,000	1.1%

検索ワード	出現数(約)	出現率	日/英比
1. ビッグデータ 社会的責任	6,570,000	40.6%	127.41%
2. ビッグデータ 法規制	3,770,000	23.3%	506.7%
3. ビッグデータ 価値 評価	2,850,000	17.6%	11.09%
4. ビッグデータ 事例	1,820,000	9.4%	5.9%
5. ビッグデータ データの精度	1,300,000	8.0%	42.4%
6. ビッグデータ リスク算出	1,180,000	7.3%	27.36%
7. ビッグデータ コスト算出	1,070,000	6.6%	12.21%
8. ビッグデータ 成功事例	626,000	3.9%	9.9%
9. ビッグデータ セキュリティ	601,000	3.7%	1.7%
10. ビッグデータ 経済性評価	547,000	3.4%	11.58%
11. ビッグデータ 技術者の育成	548,000	3.4%	20.1%
12. ビッグデータ オープンガバメント	48,300	0.3%	3%

図4 検索ワード調査結果(日本語)

あくまでも google の検索結果のインデックス数の比較であるため、具体的な内容の成否については問わない、という前提だが、この結果と、図5の英語による同内容の検索結果を比較することで、日本のビッグデータ分析の議論がどのように進むかの仮説を作ることができる。

○Google検索エンジンにて課題ワード(英語)を検索した結果(2012年9月6日実施)

検索ワード(英文)	出現数(約)	英/日比
Big data	1,420,000,000.0	87.65432

検索ワード	出現数(約)	出現率	英/日比
1. Big data Security	303,000,000.0	21.3%	575%
2. Big data example	225,000,000.0	15.8%	27%
3. Big data Open Government	142,000,000.0	10.0%	335%
4. Big data Success stories	55,400,000.0	3.9%	101%
5. Big data regulation	29,800,000.0	2.1%	9%
6. Big data accuracy	26,900,000.0	1.9%	2.4%
7. Big data engineer training	23,900,000.0	1.7%	50%
8. Big data cost evaluation	7,680,000.0	0.5%	8%
9. Big data value evaluation	4,930,000.0	0.3%	2%
10. Big data CSR	4,520,000.0	0.3%	1%
11. Big data economic evaluation	4,140,000.0	0.3%	9%
12. Big data risk evaluation	3,780,000.0	0.3%	4%

図5 検索ワード調査結果(英語)

【今後のビッグデータ議論の傾向の仮説】

①安全性に関する議論は、現状の法制度の問題から、具体的なセキュリティ確保の問題に移行していく。

②効率性に関する議論は、現状のデータに関する抽象的な価値から、具体的な利活用事例の発表に移行していく。

③有益性に関する議論は、社会的責任の議論から、オープンガバメントを活用する議論に移行していく。

▽日本語での検索結果

	技術	サービス	企画
安全性	3.7%	23.3%	7.3%
効率性	8.0%	9.4%	17.6%
有益性	3.4%	0.3%	40.6%

▽英語での検索結果

	技術	サービス	企画
安全性	21.3%	2.1%	0.5%
効率性	1.9%	15.8%	0.3%
有益性	1.7%	10.0%	0.3%

図6 日本語/英語の検索結果の比較

有益性に関する検索結果を確認すると、企業が今後ビッグデータ領域に進む意義として「社会への貢献」のため、と論じたものが多く存在した。日本におけるビッグデータの分析市場は、まだプライバシーや法規制への抵触についてのリスクが高く感じられ、成功事例が流通しておらず、イメージ戦略的な意味合いで議論をしていることが多い、という状況が見えてくる。

現状のフレームワークをベースとして、各要素を備えたプラットフォームの構造の大枠について、図7に示す。

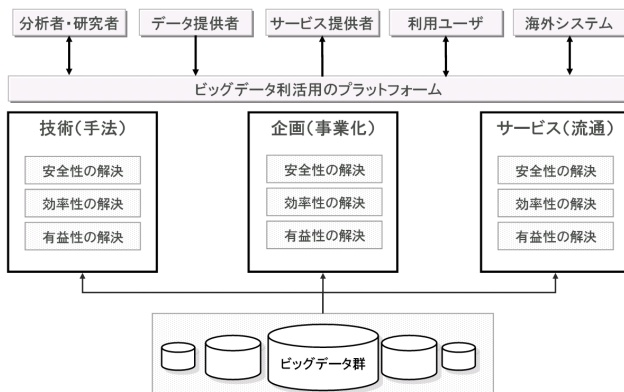


図7 課題要素を反映したプラットフォーム案

技術(手法)、流通(サービス)、企画(事業化)の全ての要素について課題となっている箇所は、政府や企業などの複数の組織や構造的な問題と関連しており、単体によって解決されるものではない。

これらの軸は、それぞれが独立したプラットフォームとして存在するのではなく、互いに関係しあった内容である上、異なる評価軸を設定する必要がある、定義が難しい。

曾根原[12]は、ライフログデータの利用における評価軸の関係を「ライフログの提供・非提供が一種の公共財ゲームである」と定義している。安全性と有益性は必ずしも両立せず、効率性とも軸が異なることから、総合的な評価を難しくしている。

本フレームワークでは、企画(事業化)の評価軸として

- ・安全性は 違法→不適切(脱法)→合法(適切)
- ・効率性は 負債→現状維持→収益(経済性)
- ・有益性は 無益→無害→有益・有用

の形で設定した。

この評価軸を曾根原[12]の作成した「保護と活用の両立モデル図」と合わせて、縦軸を[安全性:合法⇔違法]の軸とし、横軸を[有益性:有益⇔無益]の軸と設定したのが図8である。

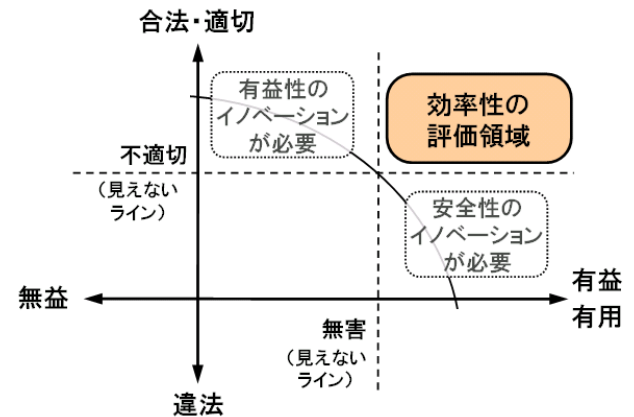


図8 安全性と有益性の軸と効率性の関係

合法で適切な手段で作られたデータである程、暗号化や秘匿処理などによってデータの有益性は減少する。逆に違法や不適切な目的でデータを利用した場合は、他社にないデータ利用であることから有益性は増すが、摘発されるリスクが高まる。

実際に企業がビッグデータ分析を実施するには、そのデータを活用することが「合法であり、かつ有益であること」が最低条件となる。

だが、実際に事業の検討を行うと、有益・有用の軸には「無害」という見えない線があり、これを超えないとデータ分析を行う意味がないことが解る。言葉の定義としては

無益：利益のないこと。むだなこと。

無害：害がないこと。悪い影響を及ぼさないこと。
(「無益」には、ベースとなる事業に対して悪影響を及ぼすニュアンスが含まれるのに対して、「無害」には他に被害を及ぼさない、という意味が含まれる。)

ビッグデータ分析の影響範囲は、利用についての法的見解、データの授受作業、社内手続き、IR的な観点なども含まれ、会社組織に関わる全てのステークホルダーに対して無害であることを証明することは、「悪魔の証明」とも似て非常に難しい。ここに定義が難しい「見えないライン」が存在する。

又、安全性の評価に関しても、法令や規制だけでなく、「不適切」という見えないラインが存在する。

この「不適切のライン」は更に複雑であり、個人差や時代による変遷、更には他社の問題も含まれる。

例え昨日まで問題のないデータ分析であったとしても、他社が不祥事などを起こしたことでユーザーの目が厳しくなり、利用に制限がかかるなどの事例もある。(例：2012年8月、Tポイントツールパーの個人行動履歴の収集方法に非難が集まり、他の企業が同様の事業を終了した事例。)

これは「不適切のライン」が明確でないために発生する連鎖的な問題であり、多くの企業がデータ分析に感じるリスクの問題と直結している。

本来的に効率性や経済性を評価するためには、これら見えない2つの軸「無害」と「不適切」のラインを超えていることをある程度保証する制度が必要となる。そのラインを超えた位置にあるデータについて、ようやく効率性、経済性、事業性などを検討することが可能となる。

また、逆に考えればこの「合法だが不適切」「有益だが有害」の部分は多くの事業会社は検討しない。だからこそ、リスクを負ってでも、イノベーションや利活用を行うことで、他社にない強みを作ることも可能となる。

だが、「不適切なイノベーション」で発生したリスクは一社が負うのではなく、社会全体に対して影響を及ぼし、市民運動や法改正などの動きにつながっていく場合もある。私企業が自社の都合によって、自由に「不適切のライン」を引くことは、他の企業にとっては新たなリスクとなる。

このような不適切な利用に対して、ポリシーの状況を詳らかにし、一律に変更・適用できる機能がプラットフォームには求められる。

図9は、上記の問題点のプラットフォームから、企画(事業化)の評価部分にフォーカスした構造案である。

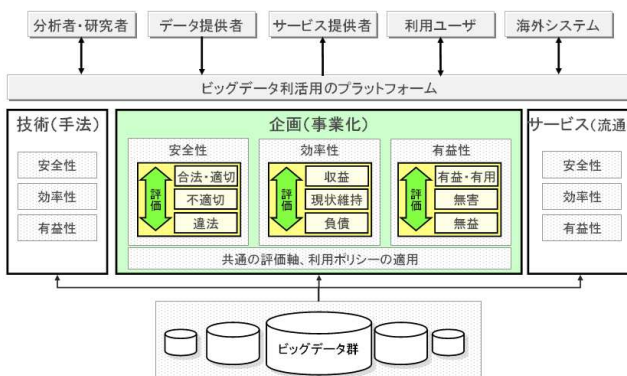


図9 企画(事業化)評価にフォーカスした構造案

本稿では、上記の要素の全てをカバーすることは難しいため、今後の検討する項目を事業者の立場から考え、企画(事業化)に伴う要素として必要な効率性の評価を「経済性」にフォーカスして検討し、プラットフォームに必要な要素についてまとめることとした。

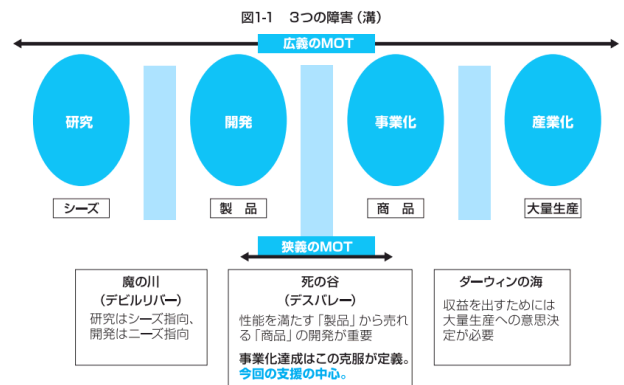
4. 経済性評価の過去研究について

経済性の評価についての過去研究は、他の分野については多く存在する

小松[7][8]による企業情報システムや医薬開発プロジェクトの経済性を評価したものや、技術経営(MOT)における研究開発から実際の商品への転換を行う「死の谷」の考え方などが参考となる。

だが、それらはいくまでも研究開発に閉じたものであり、

データ分析に対する経済性の分析として完全に適用することは出来ない。



出典) 出川通著、「技術経営の考え方 MOTと開発ベンチャーの現場から」¹⁾に加筆

図10 事業化の3つの障害[9]

ビッグデータ分析/データ流通の独自性としては、技術経営の目的が、「死の谷」を乗り越えた先にある事業化/商品化以降の展開であるのに対して、ビッグデータ分析は、目的に応じて研究段階から、開発、マーケティング、事業化、大量生産までも含めた、様々な目的に適用できる点にある。

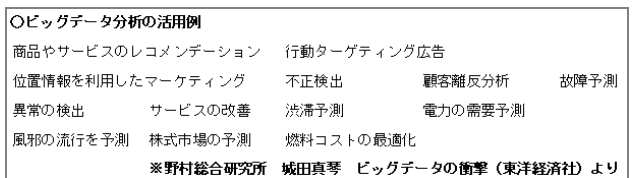


図11 ビッグデータの活用例

活用例を見ても、「行動ターゲティング広告」などは、ビッグデータから直接的に大量生産レベルの商品を開発している作業となり、不正検出や故障予測は、既存のシステムのコストダウン手段の提供。予測やマーケティングはサイズから製品への適用段階、と考えることができる。

過去の「死の谷」を越えるための手段が、個人の仮説であったり、サンプリングされたデータによるリサーチだったのに対して、ビッグデータは顧客の全ての姿を分析する点において、事業可能性は高いはずだが、具体的な事業への適用を行うには効果が不明瞭であるためコストが高く見積もられてしまう。

松岡功は「気になるのは、BI市場が日本で長らく膨らまなかったことだ。その主因は、BIの必要性は認識されつつもコスト面で割高感が強かったことにある。」[10]と語っている。

日本における分析者の数は少ないとされているが、分析の知識に加えて技術経営の知識を持っている人間は更に少ない。ビッグデータの分析に対して多くの企業が参入しない要因の1つは、技術経営を実行する側が納得性を持つ形で経済性とリスクを評価し、事業判断が可能な仕組みが存在しないことと考える。

5. リスクと経済性の算出について

もう1つの特徴として、データ分析には、常に鮮度が劣化する上に、蓄積量は増加するという点がある。

通常の研究開発の場合、時間と共に商品の数が多くなり、そこから価値や売上が増えるという計画を建てる事ができるが、データ分析の場合、分析を行った瞬間からその鮮度が落ちていき、常に辞書を更新するなどの対応を行わないと価値を保てない場合がある。また、その対象となるデータは加速度的に増えていき、蓄積費用と分析時間は常に増加する傾向にある。

対費用効果で検討したとしても、将来価値を単純に算出すると、経済性として不適合になる可能性が高い。

そのようなデータ分析プロジェクトを現状に合わせて評価する方法を検討する。

○経済性指標の整理

テクノロジーマネジメントの分野ではピーター・ボイアー[11]が研究開発の評価方法として正味現在価値の算出方法を紹介している。

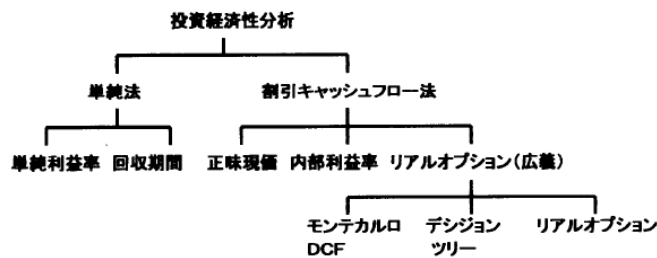


図12 経済性指標の一般的な分類[7]

一般的に、正味現在価値型の評価を行うと基礎的な研究開発型の事業については評価が低くなる傾向にある。データ分析もこの例に漏れない。

このような価値が減少する事業への価値算出の場合、割引キャッシュフロー（DCF）法による価値の算出が一般的に行われる。

○割引キャッシュフロー法にて経済性を計算する際に必要な変数		
システム投資額	投資効果額	システム維持費
法定耐用年数	法人所得税	資本コスト（利息）
評価期間	投資効果分布	投資効果標準偏差

図13 DCF法で必要な変数

上記の「システム投資額」「投資効果分布」「投資効果標準偏差」は、通常のプロジェクトならば一律の値を入れ込むことができる。

だが、未だ技術の進歩が続き、技術者の数も少ないビッグデータ分析の市場では、革新的な手法の発見や、分析者の知見等によって、成果や工数に大きな違いが出てくる可能性がある。また、クラウドソーシングの普及により、必要なサーバリソースやミドルウェアの価格は年々下がっており、業者変更によって収益が大きく変更する可能性もあ

る。

その反面、クラウド事業者によるシステムトラブルの発生事例も多く報告されており、リスク面でも未知の現象が発生する。

これらのリスクとベネフィットの発生確率を予測するためには、モンテカルロシミュレーションなどを用いて変数の発生確率を求めるのが一般的である。

○モンテカルロシミュレーションを行う変数の例

- 機会：技術革新によるコストダウン（メリット）
- 期間：必要な日数の幅（メリット/リスク）
- 成果：成果物の効果の幅（メリット/リスク）
- 破損：システムの破損/リカバリ等（リスク）
- 漏洩：重要情報のハッキング・漏洩など（リスク）

モンテカルロシミュレーションは、上記のような変動する可能性が高い変数を定め、信頼性の高い確率分布にて大量の計算を行い、発生確率の分布を調べるものである。

具体的には、データ分析にかかったコスト、試行回数、手法による成果の違い、トラブル発生回数などを統計的に算出することが出来れば、予測の信頼性を高めることが可能となる。

6. 経済性算出事例

このような、割引キャッシュフロー法+モンテカルロシミュレーションによって、実際のデータ分析の価値算出が可能かを判断するため、擬似的なコスト計画を設定したプロジェクトについて、経済性の観点から検討した。

又、その価値を正確に算定するために、有益性の観点から「データセンターのコストダウン」の可能性を取り上げ、各社のクラウドサーバの価格分布を調査し、使用した。

安全性の観点からは、総務省が発表している電気通信事業者の「事故の発生確率」について調査し、割引キャッシュフローに変換してシミュレーションに組み入れた。

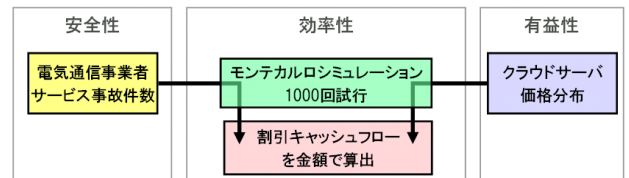


図14 シミュレーションのイメージ

▽経済性の評価事例

・ニフティ株式会社が、初期費用として21,8百万円のビッグデータ分析とシステム開発を実施する。

- ・それによる人件費削減効果は年あたり20百万円。
- ・データ規模は現在1テラ。毎年500G増加予定。
- ・減価償却、及び評価期間を5年に設定。

まず、この数字をベースにして事業計画書を作成し、5年後までの事業価値の算出を行った。算出に必要な加重平均資本コストなどの数字は、一般的なファイナンス情報から得られた数字を利用している。

▽5年後までのフリー・キャッシュフロー予測

費目	1年後	2年後	3年後	4年後	5年後
フリー・キャッシュフロー	0.9	-0.3	-1.5	-2.7	-3.9

加重平均資本コスト(WACC) : 3.16% 単位:百万円

▽割引キャッシュフロー法による事業価値の算出

$$\text{事業価値} = \sum_{n=1}^5 \left\{ \frac{n\text{年度フリー・キャッシュフロー}}{(1 + \text{WACC})^n} \right\}$$

$$= \frac{0.9}{(1 + 0.0316)^1} + \frac{-0.3}{(1 + 0.0316)^2} + \frac{-1.5}{(1 + 0.0316)^3} + \frac{-2.7}{(1 + 0.0316)^4} + \frac{-3.9}{(1 + 0.0316)^5} = -6.3 \text{ (百万円)}$$

※現在の案では投資額として不適格

図 15 事業価値算出

図 15 のように、一般的なDCF法によって算出すると、本プロジェクトは2年後からキャッシュフローが悪化することから投資不適格となり、プロジェクトの見直しを要求される。

現在のシステム費用計画では不適格となるが、今後の技術の進歩とサーバ利用料金の低減を鑑みて、収支としてペイラインに乗る確率を求めてみる。

会社	値段(月額)	H社比
A社	¥5,760	0.18
B社	¥18,144	0.57
C社	¥945	0.03
D社	¥5,300	0.17
E社	¥2,500	0.08
F社	¥2,400	0.08
G社	¥6,825	0.21
H社	¥31,815	1.00

図 16 クラウドソーシングの価格分布調査結果

検索エンジンにて検索した、上位8社のサーバ標準価格について調査したところ、標準偏差0.311で最大値1.0~最小値0.2の間に分布していることが判明した。

上記の割引キャッシュフローの計算の中から、サーバ利用料金に該当する部分について、上記の分布で金額を変化させ、シミュレーションの変数として加えた。

次にリスクの発生可能性の変数を加える。

サービスに対してリスクは常に発生するが、その一般的な発生率として適当なものは少ない。そのため、今回は一般的な指標として、総務省の登録電気通信事業者の数と、総務省に申請された事故の発生数を利用した。

一般的に登録電気通信事業者はその他の会社よりもセキュリティに対して厳密に運用されており、通常のWEB

サービスなどの管理レベルと比べても、問題が発生する件数は少ないはずである。

だが、電気通信事業者はインフラ的な音声通信サービスや回線サービスを営んでいるものも多く、一般的なWEBサービス事業者等と比べて利用頻度が高く、影響が発生した人数が多くなっていると考えられる。

	登録事業者	届出事業者	合計
電気通信事業者数(H22年度)	329	15339	15668

総務省:電気通信事業者数の推移より

事故種類	定義	事業者数	発生件数	1社平均	事業者数ベースの事故発生率	発生件数ベースの事故発生率
重大事故発生件数	3万人以上かつ2時間以上	8	17	2.1	0.05%	0.11%
詳細な様式による報告	500人以上又は2時間未満	127	8773	69.1	0.81%	55.99%

総務省:電気通信サービスの事故発生状況(平成23年度)より

図 17 総務省に届けられたサービス事故件数

これらの数字は発生件数であり、それによる最終収支への影響などは正確に算出できない。そのため、最終的な事業価値に対して、以下の金額を減ずることで対応した。

- ・重大事故の発生=1000万円の収支減
- ・詳細な様式による報告の発生=100万円の収支減

割引キャッシュフローの結果に対して、現在のサーバ費用の分布と事故発生率の確率をモンテカルロシミュレーションに加え、5年間の事業価値を算出した結果が図18となる。

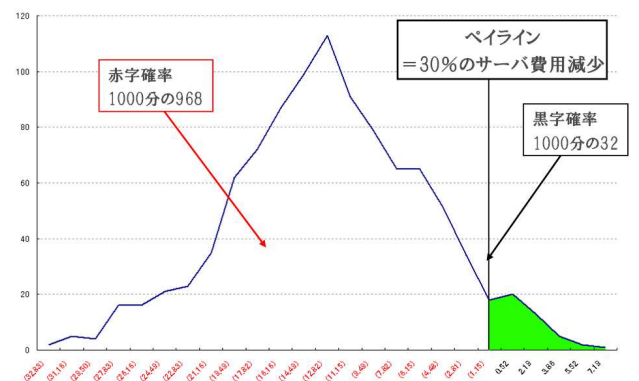


図 18 モンテカルロシミュレーション (1000回試行) 結果

本試行により、サーバシステムの価格が現状よりも約30%値下げになった場合、事業収支がプラスとなることが判明した。現状の価格分布では、黒字が発生する確率は3%程度となる。(同試験を10回実施した平均値=3.09%)

また、登録/届出電気通信事業者のレベルでセキュリティを保っている場合、事故発生率は年間1%以下に抑えられ、その影響が重大事故であっても1000万円以下で済むのならば、全体に対する収支の変化には余り影響がないと考えられる。

これにより、本プロジェクトの予算・効果の目標値が定められる上に、外部要因でサーバ費用の価格改定が発生した場合、本計画を変更し、収支予測として再利用できる。

これ以外にもリスクとして発生する要因や、個人の知見によるプロジェクトの成功確率の変化などは存在するが、定量データが非常に少ない。手法の共通化とプラットフォーム

ーム化によってこれらのデータを整備することで、予測の精度が向上し、情報収集のための労力の削減が可能となる。

7. まとめ プラットフォームの構造案

これまでの内容をまとめると以下ようになる。

日本においては法規制などの問題から、多くの企業が実データの分析ではなく、CSR的な側面からデータ分析事業に参入を検討している。

事業社が参入する上で検討する部分は企画（事業化）の要素が大きい。特に効率性（本稿ではその中でも経済性の算出にフォーカス）が不明瞭である点は、経営者との情報共有を遅らせ、利活用が進まない一因となる。

明確な数字で示される経済性の算出は可能だが、データ分析の事業価値は低くなりがちであり、かつ、安全性と有益性という定性的な要素と連動している部分が多い。

そのため、経済性を正確に算出するには、有益性や安全性の評価データを利用したり、分析手法/分析者毎の効率性の評価などを蓄積/統計化できるプラットフォームが必要となる。

本稿では、そのようなプラットフォームの検討のため、経済性の算出を単純な割引キャッシュフロー法で算出するだけでなく、安全性と有益性の定性的なデータを発生確率に落としこみ、シミュレーションを行った。

なるべく多くの事例をプラットフォーム側で収集し、統計的にメリット/リスク発生確率を含めたシミュレーション結果を、経営者、研究者、データ利用者などが共有することで、結果的に多くの事業社の参入を促すことが可能となると考える。

この考えをもとにして検討したプラットフォーム構造の案を図19に示す。

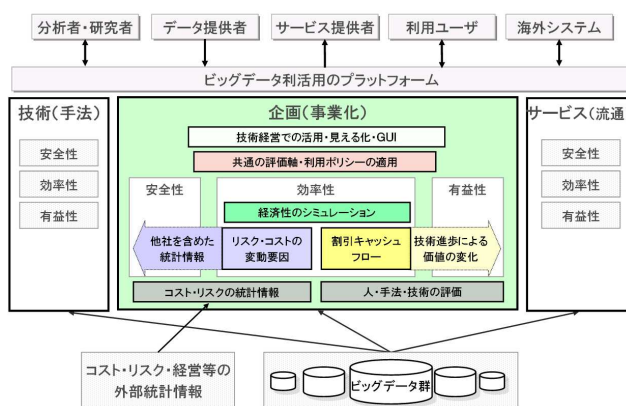


図19 プラットフォーム構造案

本構造では、まず安全性の評価として、外部のオープンデータや検索エンジンなどの結果を統計情報化し、リスクやコストの算出のための基礎データとする機能を有する。

逆に、有益性の評価は内部で分析を行っている技術者やそこで利用されているソフトウェア/サーバリソースの評

価を蓄積し、そこから導き出す構造となる。

ビッグデータの利活用は統一されたポリシーに基づいて行われ、問題発生リスクを社会に拡散するリスクを排除する役割も負う。上記の要素は共通の評価軸にて経済性の評価がされ、技術経営が求める形での見える化が実現できる。

このような構造のプラットフォームによって、ビッグデータの利活用が推進されるのではないかと考える。

8. 課題

まず、本稿の検討範囲はビッグデータの課題についての企画（事業化）の効率性の評価の部分に閉じており、今後これ以外の部分についての検討が必要となる。

安全性と有益性の定性的な指標を完全に定義することは非常に難しいと考えている。

また、経済性の部分だけで考えた場合も、制度的な課題と技術的な課題が存在する。

▽制度的な課題

多くの事業調査がそうであるように、相対的な成功率や期待収益などは事業社ごとに秘匿され、他社に対して情報を提供することは通常行わない。

また、プラットフォーム事業者は、通常はクラウド事業者などと同一であることから、分析内容や演算試行回数などを無許可で盗聴するような行為は当然行ってはならない。現状の事業社の登録情報と、分析担当者、サーバ契約期間などを用いて、可能な限りの分析を行うことで、参加事業社の認知を向上させていくことが現実的だろう。

総務省による事故発生原因は、人為要因、設備要因、外的要因の3種に区分されているが、外的要因は通信事業者による広範囲な調査によって定量化が可能になる。ディープ・パケット・インスペクションのような、ISP事業者がネット行動を捕捉するような技術も、「不適切な見えないライン」によって制限されている状況下では、データの収集は難しい。

▽技術的な課題

多くの分析事例を集め、その中でどのような箇所が上ブレ/下ブレとなった要因なのかを、様々な経営指標から調査し、自由にシミュレーションに組み入れる技術が必要となる。

様々な変数のシミュレーション結果を共有し、その分野のデータ分析に必要な要素を特定、見える化していく仕組みを取り入れることで、プラットフォームの利活用を推進することができる。

クラウド事業者や分析技術者が情報を提供し、システム開発会社やサーバ事業者などがリスクを含めた総合的な提案が可能になることでビッグデータの利活用は推進されると考える。

参考文献

- 1) 内野宏信@IT 情報マネジメント編集部,ビッグデータが"バズワード"から脱せない背景,@IT 情報マネジメント,2012-08-06
- 2) 矢野経済研究所,2012 ビッグデータ市場ー将来性と参入企業の戦略ー,2012-04-20
- 3) 情報通信審議会,知識情報社会の実現に向けた情報通信政策の在り方～Active Japan ICT 戦略～,2012-06-12
- 4) 城田真琴,ビッグデータの衝撃——巨大なデータが戦略を決める,東洋経済新報社,2012-6-29
- 5) 鈴木良介,ビッグデータ ビジネスの時代,翔泳社,2011-11-09
- 6) 堀内秀明,ビッグデータで留意すべき 4 つのポイントとその取り組み方,ソフトバンクビジネス IT+,2011-08-09
- 7) 小松 昭英,企業情報システムの経済性評価フレームワーク : リアルオプションの適用,経営情報学会誌 13(1),pp.95-118, 2004-06
- 8) 小松 昭英,医薬開発プロジェクトの経済性評価,国際プロジェクト・プログラムマネジメント学会誌 2(2), 89-101, 2008-03-14
- 9) 経済産業省 九州経済産業局,事業化支援ガイドブック 覚えておきたい儲けのコツ
- 10) 松岡功,老舗が反撃に出た激戦区・データ分析市場の行方,ITmedia,2012-05-28
- 11) Peter Boer,F.,The Valuation of TechnologyBusiness and Financial Issues in R&D,John Willey & Sons, 1999(宮正義監訳,技術価値評価-R&D が生み出す経済的価値を予測する,日本経済新聞社,2004)
- 12) 曾根原 登,クライシスに強い社会・生活基盤の創生-ライフログデータ駆動の情報循環とプライバシー保護-,信学技報 vol. 112, no. 226, EMM2012-60, pp. 27-32, 2012 年 10 月.