

# アジャイルに取り組む ――データ活用の基盤づくりの勘所



執筆者： 櫻村清尊

株式会社シグマシス システム・シェルパ プリンシパル

ソフトウェアハウス、独立系 SIer を経て、2009 年1月にシグマシスに参画。カットオーバーをシステムのスタートラインと位置づけたインフラ設計、データ設計、運用設計に強みを持つ。製造流通から金融、商社の幅広い業界の基幹・個別システムの導入の実績を持つ。データベース講師の経験も有する。

前回まではデータ活用の現状、業務部門、情報システム部門の協業の重要性について、マーケティングよりの目線で解説してきた。ここからは少し角度を変え、企業内システムにおけるデータ活用の位置づけと、分析対象データについて触れた後、求められるプロセスについてひも解いてみたい。

なお、あくまで全体の関係性や流れを重視するため、テクニカルに詳細な内容までは踏み込まないことを予めご容赦頂きたい。

## ポートフォリオにおいて 「データ活用」の位置づけは

近年に限った話ではないが、「システムポートフォリオ」という考え方が昨今は重要視されている。データ活用論に入る前に、まずこのポートフォリオの考え方について、企業として踏まえておくべき論点を押さえておきたい。

ポートフォリオの軸としては、まず利益貢献度がある。これは、もう一段掘り下げると、コスト削減への貢献と売り上げ向上への貢献に分類できる。例えば、マイナンバー対応のように、利益貢献度が低くとも法規制対応のために必須、というシステムはあるものの企業が利益を追求する以上、ここに貢献しないシステムは開発／導入すべきではない、というの

が基本的な考え方だ。

もうひとつの軸として、要求変更度がある。こちらも一段掘り下げると、ICT 技術の進化に伴い市場が変化するのに対応するケースと、新市場の開拓など、自社の競争力に直結する領域を独自で切り開いていくケースに分かれる。

前者は、要求変更度が比較的安く標準化が可能なため、パッケージ導入や SaaS 利用のウォーターフォール型開発、導入が可能だ。一方、要求変更度が高い後者は、スクラッチ開発かつ、仮説とともに最小限の開発を繰り返し、反応をフィードバックする“アジャイルリーン型”とし、市場のトップランナーを目指すか、せめて市場から遅れない程度のスピード感は持続したい。

ちなみに、会社のホームページやメール、ファイルサーバなどは利益貢献度で計りにくいながらも会社インフラとして必須であるため、サービス利用を前提として、運用までを含めてコストを最小化するのが妥当だと考える。

さて、そんな中でデータ活用の位置づけを改めて考えると、自社の競争力に直結するという意味で、要求変更度が高い部類に入る。しかし、「売り上げへの貢献」という意味で漠然とした大きな期待がある一方、「最終的に利益に貢献するかは分からない」という側面も持っている。

要は、「やってみなければわからない」という世界だから、投資対効果で慎重にポートフォリオを管理する企業にとっては、積極的な導入に踏み切れない領域だ——これが、データ活用に対する企業の本音だろうと筆者はみている。

## 「データ活用」は「駄目なら捨てる」覚悟で臨む

そうなると、仮に「データ活用」に取り組むにしても、導入するアプリケーションは、導入事例が豊富なツールで極端なハズレがないものを選ぶ必要がある。データ分析パターンも自社独自のものをチャレンジするのではなく、導入済みの各社が実施しているものを採用することになる。

これでは、データの活用を通じて、市場のトップランナーを目指すどころか、なんとか市場から遅れずにいられるかいられないか、のラインにとどまってしまう。

また、多くのツールは分析データが表やキューブ(データセット)になっているのが一般的であり、このデータセットの設計は、専門知識と試行錯誤が必要となる。結果的に、そのまま導入しようとする、サーバやツールの導入を含めて多くの時間と人手、そしてコストが費やされてしまうことになる。

基盤導入やデータセットの設計などの時間を短縮する方法は、「駄目なら捨てる」という意識を持ち、「駄目なら捨てられる」という環境を作ることだ。データ活用においては特に、基盤にクラウドを活用し、試行錯誤と相性のいいアジャイル

リーン型で取り組むのが適している。

例えば、クラウド上で Hadoop などオープンソースの大量データの製品やサービスを利用し、グラフ作成まで実施する仕組みを構築することで、必要なタイミングで複数のパターンを検証できる。これで、データセットの作成を設計、開発するのではなく、入力データを変化させながらグラフ作成し、その有効性を検証するトライ&エラーを実施できる。

データ活用は、投資対効果が分からないからこそ、アジャイルリーン型で取り組むのだ。結果として、時間とコストを削減するとともに、不要な資産化を回避できる。

## データは内部を確保してから外部を検討

では、どのようなデータを使ってデータのトライ&エラーに臨めばいいのだろうか。

データ分析というと、外部データの活用の議論が最近では活発になっているが、その前にやるべきことがある。企業規模が大きくなるほど、業務部門の立場ではどこにどのようなデータが社内是否存在するか見えない。よって、情シス部門と業務部門が積極的に協業し、企業内データの全体を俯瞰することが第1歩になる。外部データに目を向けて分析の視野を広げるのは、それからだ。

外部データは、「自分で作るデータ」「購入するデータ」に大別できる。

### 自分で作るデータ

- ワンボードコンピュータとセンサを利用しデータを作成する
- ウェブデータをクロールしデータを作成する

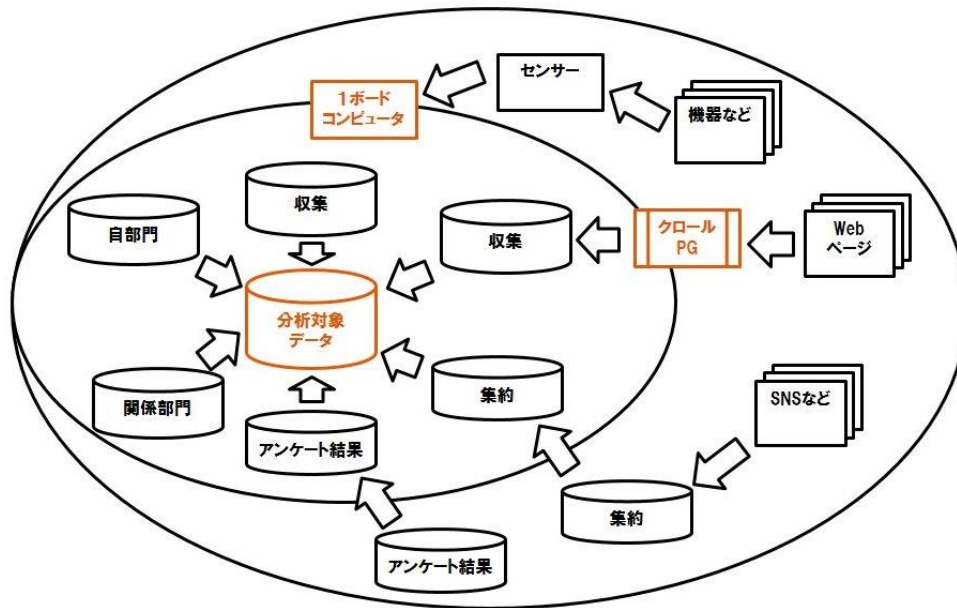
データ作成をサポートするサービスも存在するので、実際の気軽にトライすることができる。たとえば、シングルボードコンピュータでネットワーク中継すれば、植木の温度などをリモートで計測できるし、また、「募集」や「求人」をキーワード

にウェブ検索することで求人情報データを作成できる。

### 「購入するデータ」

- アンケートサービスから購入する
- ソーシャルデータ販売サービスから購入する

アンケートサービスは、ほしいデータの属性を明確にして依頼すれば、アンケート作成から実施、結果をとりまとめた上で、整理したデータを提供してくれる。Facebook や Twitter などからデータを抽出し提供してくれるサービスもある。



業界部門をまたがるデータ連携＋外部データを積極的に活用

## データは種類に加えて「状態」が鍵

企業内外のデータの入手方法と同様に意識したいのが、それら各種データの「状態」だ。同じデータであっても、状況によって状態が変化するので、活用するにはそれを踏まえて使いこなす必要がある。ここでは、「固定」「流動」という視点で、簡単な例をあげながらそれぞれの状態をみてみたい。

### 流動→固定の例

大きな事件があると連日ニュースで報道されるものだが、そのニュースの情報は報道されている間(解決される前)はリアルタイムデータ、すなわち流動データだが、いったん解決すると、歴史的事実としてヒストリカルデータ、すなわち固定データに変化する。例えば、データ活用で使うソーシャル

データも、タイムリーな内容で賑わいを見せている間は流動データだが、落ち着いてくると、まとまった固定データとして扱われるようになる。

### 固定→流動の例

旅先の写真は、瞬間を切り取った「写真」という意味ではヒストリカルデータ(固定データ)だが、それを SNS に投稿したり友人とのメッセージ交換に添付した瞬間に、リアルタイムデータ(流動データ)に変化する。

ここでは「流動」と「固定」、「リアルタイム」と「ヒストリカル」という表現を使ったが、これ以外にも「フロー」と「ストック」という視点もある。このように状態のイメージを膨らませると、データの使い方はさまざまな広がりを見せる。

逆に一つのデータを「どのような状態として扱っているのか」あるいは「どのような状態として分析したいのか」を明確にしておかないと、正確なデータ活用が行われず、ミスリードな結果に振り回されることになるから要注意だ。

## 価値創造のスピードが データ活用の成否を決める

このように、企業の内外データを「状態」を意識しながらインプットして、トライ&エラーで相関関係を見出しながらデータセットを設計すれば、効率良くトライ&エラーのサイクルを回して適切な分析結果に近づいていくことができる。

だが、本当に肝心なのは、データ活用の結果、ビジネスにおけるアクションにつなげていくためのシステム構築(もしくは、機能追加、修正など)だ。第3回目の連載でも述べたように、アクションにつながらなければ業績の向上にはつながらない。しかも、データ活用においては、アクションの「スピード」が重要だ。

企業である以上、費用対効果の観点から慎重に判断する姿勢が必要ではあるが、慎重さだけを重視していると、せっかくデータ分析で新しいマーケットやビジネス機会を見つけても、他社に先を越されてしまう。

システム構築においても、企業生命に影響がでない程度でスピードに賭け、「駄目なら諦める」というようなリスクの取り方も必要となっているのではないだろうか。データ分析に取り組む以上、システム構築のプロセスにも圧倒的なスピードが求められることは、もはや避けられない。

例えば、データ分析の結果、なんらかの新しいサービスを始めようとするならば、その因果関係を検証している間にもサービス立ち上げの準備をスタートさせ、市場や顧客からのフィードバックを得ながらチューニングしていくようなプロセスでなければ間に合わない。

さらに、その場合のプログラミングを含む開発は、あくまで

もリリースまでのスピードを重視して考えることになる。アジャイルリーンのプロセスを採用した上で、ビジネスがうまく行かなければ捨ててしまうという、覚悟も必要となる。

再利用される保証がないシステム開発である、という前提に立ち、抽象度が低かったり、再利用性が低かったりすることに拘泥することなく、「一日でも早く稼働するシステム」を開発することに注力するのだ。

リファクタリングは、ビジネスが拡大する過程の中で実施すれば十分なのである。せっかくのデータ分析の結果は、アクションのスピードにサポートされなければ、すべてが無駄になってしまうのだから。

なお、データ分析の能力の確保はどの企業にとっても大きなテーマになりつつあるが、能力強化に向けた筆者なりの考え方を最後に述べておきたい。

社員トレーニングや、協力会社との協業は大事だが、社員が自身で作成したデータセットまたはグラフを広く公開し、分析やディスカッションに参加する人を広く求めることで、知見を補強することができる環境が、ネット上には既にある。データ公開にはセキュリティの問題が絡むが、公開の方法や範囲をコントロールすることで、可能性は確実に広がる。

トライ&エラーとあわせて、オープンなアプローチを試みながら、スピードを高めていく姿勢を大切にしたい。

第5回は、データ活用を支える組織のあり方、ワークスタイルについて解説する。

※この論文はZDnet Japan(2015年6月~9月)に掲載されたコラムの再掲載となります。