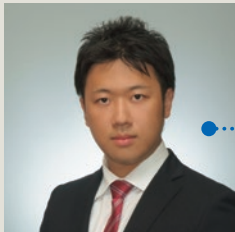


データサイエンティストの活用戦略

— データ分析を継続活用するために —

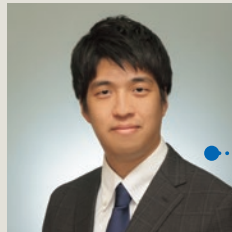
ビジネスにおけるデータ分析の重要性が高まるなか、データサイエンティストが不足しており、各社が人材確保に力を入れている。業務にさまざまな恩恵をもたらすデータ分析を継続するためにも、データサイエンティストの特性を理解し、活用することが急務であり、育成も含めた人材確保に向けた戦略を考えるべきである。



野村総合研究所
システムコンサルティング事業本部
システムデザインコンサルティング部
副主任データサイエンティスト

ひろた そういちろう
廣田 壮一郎

専門は業務改革支援、データ活用・分析支援、先端技術の調査・適用検討



野村総合研究所
システムコンサルティング事業本部
システムデザインコンサルティング部
コンサルタント

さくやま かつや
作山 克也

専門はデータ活用・分析支援

データサイエンティストの希少性、重要性

データを基に企業の経営戦略策定や業務の高度化・自動化を行う、ビジネスアナリティクスへの取り組みが進んでいる。これまで現場の経験と勘に頼っていた業務も、データによって定量化し、推進できるようになりつつある。ビジネスに寄与する情報をデータから抽出する人材は、データサイエンティストと呼ばれている。企業はビジネスアナリティクスを推進するため、データサイエンティストの確保に向けた活動に日々取り組んでいる。

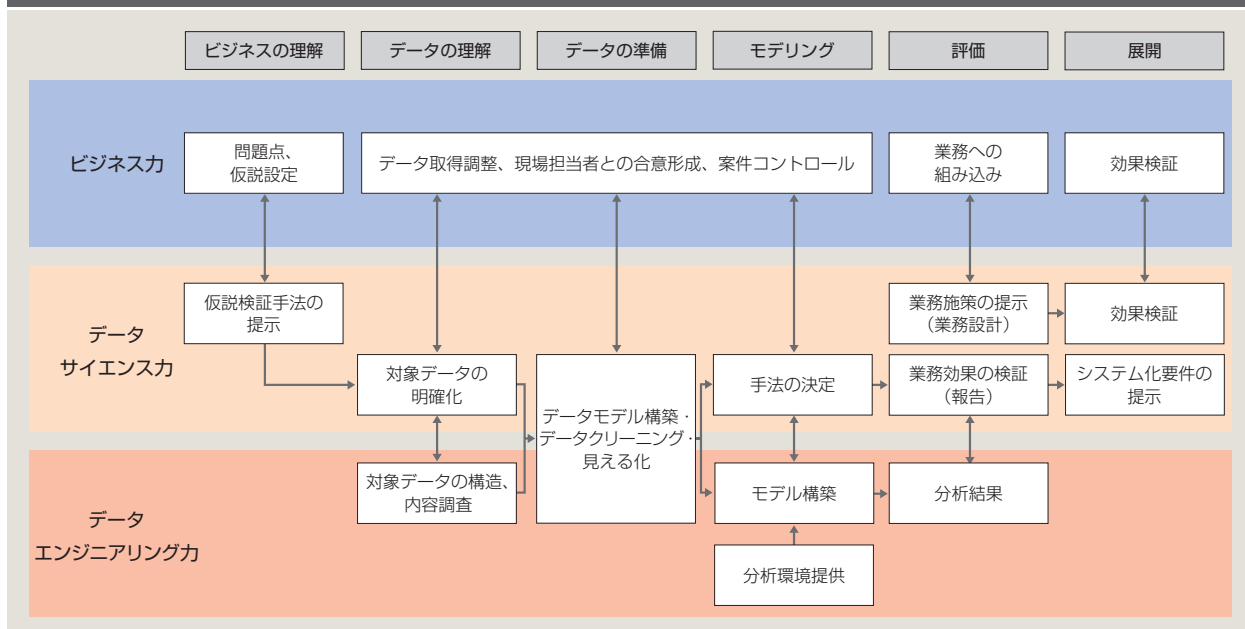
IT人材の中でも、特にデータサイエンティストは供給不足の状態にある。希少性からデータサイエンティストを特別な待遇で迎える企業も多く、人材争奪戦が起きている。2013年の米国O'Reilly Media社の調査によると、アメリカにおけるデータサイエンティストの初任給は年額1,000万円を上回る企業もあり、ここからも人材の希少性、重要性が伺える (<https://www.oreilly.com/data/free/files/stratasurvey.pdf>参照)。

そもそもなぜデータサイエンティストは、これほどまでの特別待遇がとられるほど希少な存在なのか。理由の1つとして、必要なスキルの幅広さがある。一般社団法人データサイエンティスト協会では、次の3つの力をスキルセットとして定義している。1つは、ビジネス課題を整理し、解決する力である「ビジネス力」。もう1つは、情報処理や統計学などの情報科学系の知恵を理解し、使う力である「データサイエンス力」。そして、データサイエンスを意味のある形に、実装、運用できるようにする「データエンジニア力」である。これらのスキルをバランスよく保有している人材は希少なため、どのように育成し確保するのか、戦略的に考える必要がある。そこで本論では、データサイエンティストの育成と確保に必要な考え方について論じる。

データサイエンティストの育成

データサイエンティスト育成においては、必要なスキルのうち、どのスキルを重点的に習得させるのか意識しておくべきである。全

図1 アナリティクス系プロジェクトのプロセスモデルとスキル、タスクの関係



てのスキルを一気に習得するのは難しい。特に「ビジネス力」を習得する段階と、「データサイエンス力」や「データエンジニアリング力」を習得する段階の間に壁があると考えられる。これはデータサイエンティストの業務をイメージすると分かりやすい。人と相対するなかで培われるスキルと、データと相対するなかで培われるスキルの差である。

一般的には、分析作業を担当しているメンバーが、そこで習得した「データサイエンス力」と「データエンジニアリング力」をベースに、ビジネスサイドと対話を重ねて「ビジネス力」を習得していくのが、基本的な育成シナリオになると思われる。なぜならデータサイエンティストとしての「ビジネス力」を発揮する工程を担当する際には、「データサイエンス力」や「データエンジニアリング力」に対する理解が必要となるからである。「ビジネス力」のみしか保有していないメンバーが設定する問題設定は、技術的な実現可

能性を考慮できていないことも多い。

CRISP-DM (Cross-Industry Standard Process for Data Mining) というコンソーシアムが提唱している、同名のデータ分析プロジェクトの標準プロセスに、各スキルを主に保有するメンバーと担当するタスクについて加筆したものを図1に示す。段階的に育成していく場合は、初めに「データサイエンス力」や「データエンジニアリング力」に定義されているようなタスクの経験を積みせると良いだろう。その際は、プロジェクトの推進に、時間的にも体制的にも余裕を持たせるように心掛けたい。ビジネス上の課題解決には複数の手法が存在する。試行錯誤しつつ、さまざまな観点からデータを観測することで、初めて新たな知見を得られる。しかし成果を急ぐあまり、試行錯誤の回数を減らしてしまったり、報告書作成に注力したりするケースがある。データサイエンティストに試行錯誤する時間を与えることで、プロジェクトの

生産性は一時的に低下するかもしれない。しかし、ここで得られるのは、次に生かせるノウハウである。試行錯誤の機会を奪うことは、データサイエンティストとしての学習機会を奪っていることに他ならない。

次に、データサイエンティストの育成候補となる人材に欠かせない素養について述べる。彼らの多くは好奇心旺盛であり、他人が気づかない事象をデータから見つけることを好み、最新技術習得のための努力もいとわない。膨大な知識やスキルを必要とするデータサイエンティストとして、変化の早いこの領域を生き抜くためにも、これらは必要な素養といえる。こうした素養は、育成の施策を打つことが難しい。そこで、素養のある人材を見いだすための仕組みが必要となる。ここでは、データ分析に興味のある人材が自発的に集うようなコミュニティを提供することが望ましい。例えば、社内の誰でも参加できるような、データサイエンスに関連する勉強会やコンテストを企画し、集まったメンバーから候補を見いだすといったことである。その際には、予算確保やマネジメント層に理解を得るなど、データサイエンティスト育成を組織的にバックアップする仕組みも併せて考える必要があるだろう。一見すると効率が悪い手法に思えるかもしれないが、適性のある人材を見だし、データサイエンティストとして育てていくには、効率的な方法だと筆者は考える。

データサイエンティストが活躍できる組織づくり

優秀なデータサイエンティストを確保した

にもかかわらず、成果が出ないと感じている企業や組織も多いのではないだろうか。データサイエンティストに対する期待は過度に高まっており、データさえあれば全ての課題が解決できると考える人も少なくない。しかし、データ分析は魔法のつえではない。データサイエンティストに任せるだけで、企業のビジネス課題を解決できるわけではない。データ分析が力を発揮する領域の見極めや、解決すべきビジネス課題の明確化が不可欠である。さらに、データ分析を意味のあるものにするため、改善の対象となる業務や組織のあり方、分析に必要なデータの理解など、さまざまな知識や準備も必要となる。

システム開発同様、データ分析の活動は多岐にわたる。そのため、プロジェクトの推進に必要な知識や経験を備え、1人でカバーできる人材を見つけるのは難しい。そこで、メンバーが互いにスキルを補い合うようなチームビルディングが理想だが、十分な体制を組めないケースも多い。チーム一丸となって取り組む体制を取るべきだが、データサイエンティスト1人に全てを任せる体制になってしまっているケースが散見される。

後者の状態に陥る原因の多くは、データ分析技術に対するマネジメント層の理解不足である。データサイエンティストは、プロジェクトメンバーとして活動することが多く、マネジメント層に就くケースはほぼないため、この層ではデータサイエンスに関するスキルや技術への知識が不足しがちである。その場合、データ分析の結果に対するレビューをうまく実施できないため、データサイエンティストの能力を引き出しスキルを生かすことが

できないケースが多々ある。

マネジメント層に求められる能力で最も重要なのは、ビジネスとデータサイエンスをつなげる能力である。必ずしも高度なデータ分析技術を有している必要はないが、データサイエンスの基礎的な技術の知識や、分析結果をレビューできる能力が必要だろう。

こうなると、データサイエンティストを育成し活躍させる組織を立ち上げるためにデータサイエンティストが必要という、鶏が先か卵が先かのようなジレンマが生じてくる。そこで、データサイエンティスト育成を社内で行うことにこだわるのではなく、外部の優秀なデータサイエンティストを、マネジメント層と対等に交渉できるポジションで迎えることも視野に入れるべきだろう。その際は「ビジネス力」、「データサイエンス力」、「データエンジニアリング力」を満遍なく兼ね備えた人材を募集すべきである。要求される給与は高くなるかもしれないが、これは投資であり、惜しむべきではないと筆者は考える。

外部の優秀なデータサイエンティストを雇用するに当たり、いくつか課題がある。特に難しいのは、既存の人事制度や給与体系では、優秀なデータ分析人材を確保することができない点である。現状のキャリア制度では希望年収に応えられず、評価制度やキャリアパスも不明確で、専門外の業務を担当させられる可能性もあるなど、データサイエンティストから見ると積極的に契約できない内容となっているケースも散見される。そのままの状態では、優秀なデータサイエンティストを採用することはできないだろう。

こうした状況を打開する施策の1つに、

データサイエンティストを有する組織を、既存の組織とは別に独立した組織として用意する方法がある。この組織の中では、データサイエンティストは個人事業主のように自分の案件を創出していくこととなる。その際、事業部との距離が遠くなる課題が出てくる。この課題に対処するため、事業部と強いパイプラインを持っている人材も組織に加えることを検討したい。

データ分析に対する投資対効果

データ分析を活用したいと考えて、PoC (Proof of Concept : 概念実証) を実施する企業が増えている。しかしPoCで十分な結果が得られたとしても、その後の活動がうまく進まないことがある。要因の1つとして、データ分析の技術的な要素を理解し、継続的に運用する人材が社内に存在しない点が挙げられる。例えば、需要予測モデルを構築したとしても、市場は時間とともに変化するため、モデルの予測精度は落ちて陳腐化する。モデルの精度を監視しメンテナンスする人材がいなければ、事業における活用には不安が残る。

データサイエンティストを社内に確保するのは投資と同じであり、一足飛びに効果を実感できるものではない。データサイエンティストが活躍できる組織を立ち上げるには時間を要する。早くから挑戦を続けてきた企業は、いち早くデータ分析の恩恵を享受することができるだろう。社内にデータサイエンティストが1人いるだけでは解決できない。全社で包括的に、データサイエンティストを支援する仕組みを検討すべきである。 ■