

データドリブン経営を目指す企業に必要な データサイエンス

中林 紀彦

データドリブン経営は客観的な指標に基づいた経営資源の配置や意思決定、事業に関わるさまざまな施策を行うものであり、多くの企業がデータドリブン経営への取り組みを進めている。本稿では、データドリブン経営を成功に導くために必要不可欠な「データサイエンス」を企業経営の中に組み込み実践する方法を、ヤマトグループの取り組みを例に解説する。

キーワード：データサイエンス、データサイエンティスト、デジタルトランスフォーメーション、ビッグデータ、AI（人工知能）

1. はじめに

2020年、創業101年目を迎えるヤマトグループは大きな変革を開始した。変革の方向性を示した経営構造改革プラン「YAMATO NEXT100」[1]を発表し、社会インフラを担う一員としてこれからも社会の課題に正面から向き合い、お客さま、社会のニーズに応える新たな物流のエコシステムを創出することで、次の時代も豊かな社会の創造に持続的な貢献を果たす企業へ変貌を遂げようとしている。この「YAMATO NEXT100」の三つの基本戦略は次に示すものであり、データに基づいた経営に転換する「データドリブン経営」を戦略の中核に据えている。

(1) お客さま、社会のニーズに正面から向き合う経営へ転換する

機能単位の組織を、リテール・地域法人・グローバル法人・ECの四つの顧客セグメント単位（事業本部）に再編し、経営と事業の距離を縮め、「お客さまの立場で考え、スピーディーに応える」ヤマトを取り戻す。

(2) データに基づいた経営へ転換する

デジタルトランスフォーメーション（DX）による物流オペレーションの効率化、標準化に加え、データ分析に基づく業務量予測、経営資源の適正配置、プライシングを上位レイヤーで迅速に意思決定するデータドリブン経営へ転換し、第一線がお客さまにしっかりと向き合える「全員経営」を復活する。

(3) 共創により、物流のエコシステムを創出する経営へ転換する

自前主義にこだわらず、お客さま、物流事業者をはじめとするさまざまなステークホルダーとオープンな物流インフラを共創し、共に成長する企業へ進化する。

データドリブン経営は客観的な指標に基づいた経営資源の配置や意思決定、事業に関わるさまざまな施策を行うものであり、多くの企業が注力して取り組むテーマの一つである。このデータドリブン経営を成功させるうえで必要不可欠なものが「データサイエンス」であり、本稿ではこの企業経営の中に「データサイエンス」を組み込んでデータドリブン経営を成功に導くために重要となる要素をヤマトグループの取り組みを例に解説する。

2. データサイエンスとは

そもそも企業の中で求められる「データサイエンス」とはどのようなものか？一般的にデータサイエンスは「データを用いて新たな科学のおよび社会に有益な知見を引き出そうとするアプローチのことであり、その中でデータを扱う手法である情報科学、統計学、アルゴリズムなどを横断的に扱う。」[2]という言葉で定義されているが、ヤマトグループでは周辺環境も含めてもっと広義なものとして捉えている。まずデータサイエンスは手法の一つであり目的は経営課題や事業課題を解決するために用いられるものであるということが大前提である。データサイエンスに取り組む多くの企業は方法論から入りがちで後述する「データサイエンティスト」がデータサイエンスを実現する人材であり、データサイエンティストを採用したり育成すれば解決するという幻想を抱いている。しかしデータがなければデータサイエンスを行うことはできないし、またデータを処理する環境も必要である。さらにデータ

なかばやし のりひこ
ヤマトホールディングス株式会社
〒104-0031 東京都中央区京橋 2-2-1 京橋エドグラン 19階

サイエンティストが導き出した分析結果やアルゴリズムを実際のビジネスの現場で利用しなければ意味はなく、アウトプットを事業の中に組み込んで運用する仕組みや体制も必要となる。このように「データサイエンス」とは企業の中でデータを活用して価値を生み出すための一連のプロセスであり、企業経営・事業運営に深く関わるものである。

3. データサイエンスが企業にもたらす価値

ではこの「データサイエンス」は企業にどのような価値をもたらすのか？企業の取り組みや課題に応じて効果やインパクトの大きさもさまざまであるが、ヤマトグループでは収益（売上）とコスト削減（利益）への貢献を大きな目的としている。たとえば収益増への貢献としては適切なプライシング（価格設定）を行うことによる新規顧客の拡大や既存顧客とのビジネスの拡大の効果を狙い、コスト削減効果としては予測にもとづく適切な経営資源の配置を行うことでコスト効率を最大化することを目指している。具体的には全国各地数千の宅急便センターとよばれる拠点に集まる荷物の量を予測し、それに応じた人員配置や車の手配などを事前に最適化することでコストを最適化し利益の最大化を図っている。予測期間に関しては勤務のシフトを1ヶ月以上前に確定させる必要があるため3ヶ月先までの日次での予測が必要であり、それなりの効果を出すためには誤差も数%以内で予測できるような高度なモデルが求められる。定量的な効果についてはここでは詳しく述べることはできないが、「YAMATO NEXT100」では営業収益を1.6兆円から2兆円へ、営業利益を1,200億円以上を数値ターゲットとしデータドリブン経営に関わるデータサイエンス活用からの貢献もこの中に含まれるものである。

4. 事業会社が必要とするデータサイエンティスト

前述のように価値をもたらすデータサイエンスを実行するための「データサイエンティスト」はどのようなスキルをもっている必要があるのか、特にヤマトグループのような顧客にサービスを提供するような事業会社ではどのようなスキルが必要とされるのかをデータサイエンティストの役割を細分化しながら解説する。図1はデータサイエンスのフレームワークを示したものであるが、データサイエンティストに必要なスキルはビジネスの理解からアルゴリズムのシステムへの組み込みまで多岐にわたる幅広い知識とスキルが必要で

1. ビジネス（社会）の課題を理解する
 - 解決すべき課題と費用対効果
2. データ・分析課題へ落とし込む
3. オペレーションの理解
 - データの理解
 - 足りないデータの取得
4. データ収集・蓄積
5. データ加工
6. データ分析・モデル作成
 - 機械学習
 - 深層学習（Deep Learning）
7. 結果の評価
 - 施策を提案する
8. オペレーションへの組み込み

図1 データサイエンスのフレームワーク

ある。

ヤマトグループではデータサイエンティストのスキル要素を図2のように大きく四つに分類してプロジェクトテーマに応じてそれぞれのスキルをもったメンバーをチームとして組成してデータサイエンスに取り組んでいる。

以下に四つの分類をデータサイエンスのフレームワークに照らしながら詳しく解説する。

4.1 ビジネス・スキル

フレームワークの1から3（1. ビジネスの課題を理解する、2. データ・分析課題へ落とし込む、3. オペレーションの理解）を担当する役割の人材であり、データサイエンス・プロジェクト全体のリーダーを担う。対象とするビジネスのビジネスモデルや収益構造、事業課題などを理解するのはもちろんであるが、後述する「データサイエンス」、「データエンジニアリング」、「システムエンジニアリング」とデータサイエンス全体に関する俯瞰的な知識と経験が必要となる。事業部門のメンバーと共にビジネスの課題を定義しそれをデータサイエンスの課題に落とし込んでチームを組成し問題解決に当たる。また現場のオペレーションにも精通している必要があり、利用するデータの発生過程やデータサイエンスのアウトプットを適用する際のオペレーションの設計にも関わる必要がある。

4.2 データエンジニアリング・スキル

フレームワークの4から5（4. データ収集・蓄積、5. データ加工）を担当し、データ分析に必要なデータセットを収集したり必要な形式に加工する役割を担う。データセットは不完全なものもあるので必要に応じて欠損値を補完したり分析可能なデータ形式への変換を行う。必要なスキルとしては、データの整形やクレン

ビジネスの経験

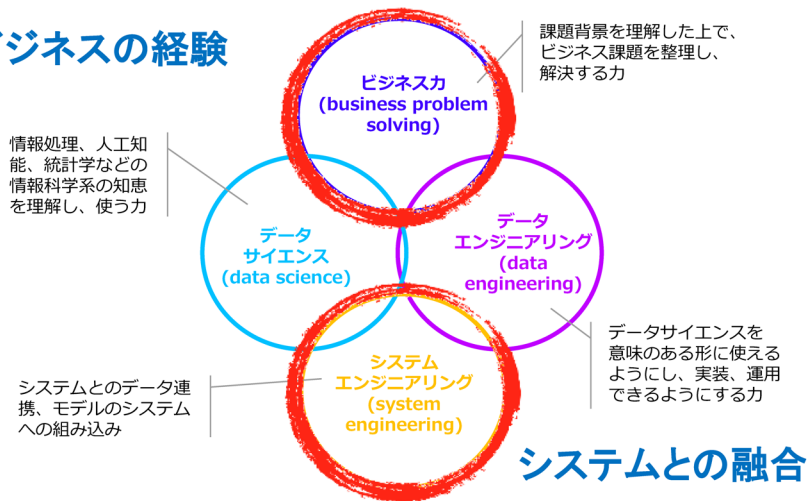


図2 事業会社が必要とするデータサイエンティストのスキル

ジングのための Python や専用ツールを扱う技術、分析に利用されるデータベースに関するスキルが挙げられる。パソコンでは扱えないようなビッグデータを扱う必要もあるのでクラウド上でのデータハンドリングの知識や経験も必須である。

4.3 データサイエンス・スキル

フレームワークの6から7(6. データ分析・モデル作成, 7. 結果の評価)を担当し、集計処理など単純な統計的なデータ処理から機械学習を用いた高度な予測モデル、さらには深層学習による画像解析や音声認識など対応する課題によって用いるアプローチも多岐にわたる。

たとえば機械学習は表1のように学習のタイプによって大きく三つに分類される。教師あり学習は名前のとおり正解を教えて学習するタイプのアプローチで、株価の予測や売上の予測など定量的な数値を予測するのに用いられる。教師あり学習のもう一つの代表的な問題はデータを分類するというもので、たとえば何百万人もの顧客を属性や購買パターンによって分類し広告などのターゲットを絞るのに利用されたり、サービスを解約しそうなグループをあらかじめ予測して事前施策につなげる活用などがある。

また正解データが与えられない教師なし学習の代表的なタスクはクラスタリングとよばれるもので、データをいくつかのグループに分類する手法である。このアプローチは顧客や商品を属性ごとに分類して、マーケティング施策のダイレクトメールや商品のレコメンドに利用されたりしている。

さらに近年では強化学習とよばれるアプローチも研

表1 機械学習の種類

タイプ	問題の種類	アルゴリズムの例
教師あり学習	回帰	線形回帰
		ロジスティック回帰
	分類	決定木 ランダムフォレスト
教師なし学習	クラスタリング	k平均法
強化学習	ゲーム	Q学習

究開発が進められている。強化学習はエージェントとよばれるものを定義してある条件下でエージェントが最も多くの報酬を得られるような方法を学習する手法である。囲碁のAlphaGoやカーレースのゲーム(自動運転)などゲーム分野での応用が有名であるが、ビジネスの世界では金融分野でも金融商品のポートフォリオ管理などへの応用に「複利強化学習」の研究が進んでいる。

ここで必要となるスキルはPythonのscikit-learnなどの機械学習フレームワークや深層学習のフレームワークを扱ってモデルを構築すること、また最近ではクラウドで簡単に機械学習モデルを構築できるSaaS系サービスが増えてきているのでそれを活用することなども必要なスキルセットの一つとなっている。

4.4 システムエンジニアリング・スキル

フレームワークの8(8. オペレーションへの組み込み)を担当し、データ分析の結果や構築されたモデルを実際のビジネスの現場で活用するプロセスでデータサイエンスの効果を出すために非常に重要な役割を担う。データサイエンスの成果物というとグラフや表が並ぶ報告書をイメージされがちであるが、データサイ

AI-Readyな企業に向け、経営層、専門家、従業員、システム・データに関して、5段階のレベル分けを提示

	経営・マネジメント層	専門家	従業員	システムレベル・データ
レベル5	AI-Powered企業として確立・影響力発揮 全ての事業・企業がAI×データ化し、業界そのものの本質的な刷新 (disruption) を仕掛けている。			
	<ul style="list-style-type: none"> AI×データを理解するCXOが全社、業界の刷新の中心を担う 業界全体、他社との連携を推進 	<ul style="list-style-type: none"> 全技術者が領域×AI知識を持つ AI×データ活用技術、研究高面の最先端の人材、経験を持つ 	<ul style="list-style-type: none"> 皆が理数・AI×データ素養を所持 社内外の専門家と共同で活用 ミドル層は資本、人脈で貢献 	<ul style="list-style-type: none"> リアル空間も含めて全てがデータ化、リアルタイム活用 協調領域では、個別領域のAI機能、API提供、共通PF化 競争領域では、独自機能のAI開発、サービス化
レベル4	AI-Ready化からAI-Powered化へ展開 AI×データによって企業価値を向上。コア事業における価値を生むドライバーとしてAIを活用。			
	<ul style="list-style-type: none"> AI×データを理解し事業活用する人材を経営層に配置 AI-Readyになるまで投資継続 	<ul style="list-style-type: none"> AI×データ活用技術開発、研究高面で最先端テーマの取組み開始 	<ul style="list-style-type: none"> 過半数が高いAIリテラシーを所持 データ・倫理課題を整理・遵守 AI×データによる業務刷新が推進 	<ul style="list-style-type: none"> 業務システムと分析システムがシームレスに連携 大半の業務データがリアルタイムに近い形で分析可能
レベル3	AI-Ready化を進行 既存の業務フローのAI×データ化による自動化に目途がつく。戦略的なAI活用も開始する。			
	<ul style="list-style-type: none"> 経営戦略にAI活用を組み込み AIへの投資をコミットメント 幹部社員へのAI教育を実施 	<ul style="list-style-type: none"> 相当数のAI分析・実装要員を持つ 独自のAI開発・事業展開が可能 	<ul style="list-style-type: none"> 実務へのAI活用が徹底 そのための手順やツールも整備 社員へのAI教育を開始 	<ul style="list-style-type: none"> 業務フロー、事業モデルがデータ化 業務系に加え分析系のデータ基盤も整備開始 領域特性に応じてAI化、RPA適用等を使い分け
レベル2	AI-Ready化の初期段階 AI活用についてスモールスタートで経験を積む。一部の簡易業務のAI化も専門家力を借りつつ着手開始。			
	<ul style="list-style-type: none"> AIの可能性を理解し方向性を発信 具体的な戦略化は未着手 データ・倫理課題は未整理 	<ul style="list-style-type: none"> 少数がAI・データを理解 外部と協力的、既存技術を活用 	<ul style="list-style-type: none"> 一部AI基礎の理解 AI×データ素養を持つ社員も存在 AI人材の採用を開始 	<ul style="list-style-type: none"> 一部業務でAI機能の本格適用を実施 一部データが分析・活用可能な形で取得可能に 顧客行動、環境、リアル空間のデータ化は未着手
レベル1	AI-Ready化着手前 AIの方法論の議論が先行し、AI×データを活用した事業運営・刷新・創造は未着手。			
	<ul style="list-style-type: none"> AIへの理解がない AIが業界や自社の企業経営に与える影響の認識も不十分 	<ul style="list-style-type: none"> システムは外部委託中心 IT部門はIT企業とのつなぎ役 	<ul style="list-style-type: none"> 経験、勘、属人的対応が中心 課題も人員、工数をかけて対応 理文分離型の採用 	<ul style="list-style-type: none"> レガシーシステムが肥大化 データの収集、取り出し、統合に年単位の時間が必要 データの意味や示唆の理解も不十分

図3 AI-Readyな企業 (文献 [3] より引用)

エンズは意思決定や業務に反映されてこそ効果を発揮するので、ビジネスプロセスやITシステムに組み込まれていくことが重要となる。データを可視化するものはダッシュボードとしてPCはもちろん、タブレットなどで意思決定者がいつでもどこでも確認し意思決定に利用できるシステムが必要となり、モデルを構築した場合には「ML Ops」ように業務システムに組み込んで利用するシステムが必要となる。「ML Ops」は英語ではDevOps for Machine Learning や Machine Learning & DevOps と表現される考え方で、昨今は機械学習のモデルをシステムに組み込むアプローチが一般的になってきている。

ここでは、発生したデータをRDBやNoSQLデータベースに蓄積するデータストアに関するスキル、XML、JSON、画像、動画、音声、テキストなどデータフォーマットに関するスキル、さらにはモデルから推論を返すML Opsシステム構築に関するスキルが必要となる。

5. データドリブン経営を成功に導くために必要な五つの要素

前節までは企業の中でのデータサイエンスの位置づけや価値、データサイエンスを実行するための人材であるデータサイエンティストについて述べたが、ここからは視野を広げて企業の中でデータサイエンスを

用してデータドリブン経営を成功に導くために必要な要素を五つのパートに分けて解説する。

図3は経団連の提言「AI活用戦略～AI-Ready社会の実現に向けて」[3]の中でまとめられているAI-Ready化(AIを活用するための準備)を進めるためのガイドラインからの抜粋で、AI-Readyな企業に向け、経営層、専門家、従業員、システム・データに関して、5段階のレベル分けを提示したものである。この部分は著者も中心メンバーとなってまとめたものであるが、「AI」は特定の技術を指すのではなく、広く「高度に複雑な情報システム一般」を指すものとして議論しているため「データサイエンス」と置きかえても問題ないとする。ここでもデータサイエンティストのような「専門家」だけではなく「経営・マネジメント」、「従業員」と企業に関わる人すべてがレベルを上げないと活用は成功しないと述べられている。また「システム・データ」に関しても、旧来からの逐次処理を多用したレガシーなITシステムでは課題が多く、業務システムと分析システムがリアルタイムにシームレスに連携する仕組みの構築が求められている。

以下は「AI-Readyな企業」の内容に著者の視点を加えて、データサイエンスを活用してデータドリブン経営を成功に導くために必要な要素を五つの観点で整理したものである。以降それぞれについて解説を加え

ていく。

1. 経営戦略の中で“データ”や“データサイエンス”を議論する
2. データサイエンスの結果を事業の中に組み込んで運用する
3. 必要な組織を立ち上げ人材配置する
4. アーキテクチャを描いて環境構築する
5. 事業会社で戦力になる人材教育を理解し実践する

5.1 経営戦略の中で“データ”や“データサイエンス”を議論する

昨今データは石油のような資源であるとなえられ、企業の中でも重要な経営資源として位置づけられている。しかし「ビッグデータに取り組む」や「AIを活用する」のように手段が目的化してしまっているケースも散見される。ヒト・モノ・カネ・データを経営資源としてどう企業経営、事業運営に組み込むかを経営陣が議論し経営戦略や中期経営計画に組み込んで実行する必要がある。前述の AI-Ready な企業の経営・マネジメントはレベル 3 以上でなければいけない。さらに AI やデータサイエンスを理解して実装できる人材を複数人経営メンバーに配することで、企業価値を向上させコア事業における価値を生む源泉としてデータを活用しデータサイエンスを駆使することができる。

5.2 データサイエンスの結果を事業の中に組み込んで運用する

前節のシステムエンジニアリング・スキルの部分で述べたように、データサイエンスの効果を出すためにはデータ分析の結果や構築されたモデルを実際のビジネスの現場で活用するプロセスが非常に重要である。単なる数値やグラフではなくデータサイエンスの結果に血を通わせる必要がある。現場のオペレーションに組み込むためには「ML Ops」のようなシステムを構築するだけではなく、仕事のやり方や規則、さらには働き方などにも踏み込んで変革するケースもある。データサイエンティストの職域を超える活動ではあるが、必要なステークホルダーにデータサイエンスの内容や効果を理解してもらいチームとしてプロセス構築と運用に当たる必要がある。また一度構築した機械学習のモデルは永久に利用できるものではなく、定期的に精度を監視しながら必要に応じてモデルの再構築やプロセスの改善なども必要となる。このような活動を定期的に行うには従業員のレベルもレベル 3 やレベル 4 以上である必要がある。

5.3 必要な組織を立ち上げ人材配置する

データサイエンティストは個人で活動するのではな

厨房だけではなくフロアスタッフや食材の調達も必用
各事業会社をもつレストランをグループで統合



図4 【集中と分散の例】チェーンレストラン方式の組織

く組織化して機能させる必要がある。ヤマトホールディングスでは図4のように考えて組織を配置している。データサイエンスの活動をチェーンレストランにたとえてデータサイエンティストは料理人であると仮定する。本部のセントラルキッチンでメニュー開発や食材の集中購買などを行う中央の機能と、各店舗のキッチンでお客さまに提供する料理を調理する分散でもつ機能が必要であり、集中させるリソースと分散でもつリソースのバランスが重要である。またこの組織には権限と予算を与えてさまざまなデータサイエンスの施策をスピード感をもって実行していく体制が必要である。

5.4 アーキテクチャを描いて環境構築する

データサイエンス成功のためには IT システム特にアーキテクチャも非常に重要な要素である。アーキテクチャは元来、建築の言葉で、建築物を設計する際に共通となる設計基準および規制を設けることで、建物の耐久性、利便性、安全性、コストの最適化を実現する考え方で、特に大規模な IT の全体設計をする場合、重要な概念であり、都度発生するビジネス要件に対し基準や規制なく開発を進めると、業務全体の利便性・効率性・ROI・拡張性に欠け、システムとしての行き詰まりが生じる。都市設計にたとえると、未秩序に継ぎ足して建ててしまっていた温泉街は個別最適に建築したため、利便性、効率性、安全性が低く、全体最適というアーキテクチャを追求したスマートシティは、都市の抱える諸課題に対し、新技術を活用し、マネジメント（計画、整備、管理・運営など）が行われ、全体最適化が図られる持続可能な都市となる。図5はヤマトホールディングスが策定したアーキテクチャであり、現状ではシステム間連携がスパゲッティ状態であるため、リアルタイムに情報連携できるように IT 基盤を抜本的に見直し、機能の最適配置を実現するものとなっている。機械学習のモデルをシステムに組み込

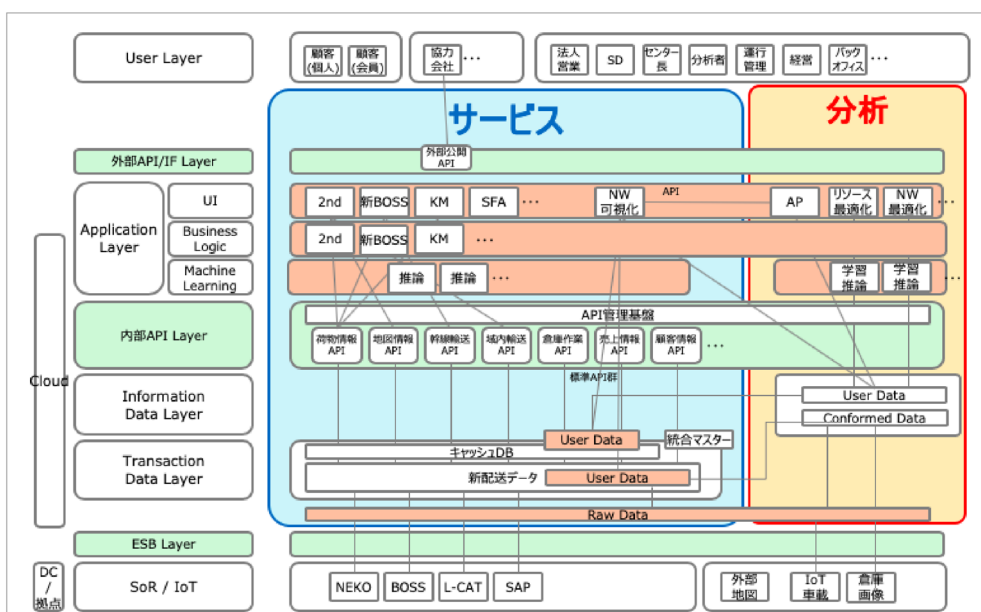


図5 ヤマトホールディングスにおけるアーキテクチャの例

む「ML Ops」を実現するためにさまざまなサービスから発生するデータを分析基盤とシームレスな連携を取りリアルタイムなデータ活用を実現している。

5.5 事業会社で戦力になる人材教育を理解し実践する

最後は人材育成である。図2で示した人材特にビジネスとデータサイエンスの両方を理解する人材を育成するのは一朝一夕でなし得ることでない。中長期的な計画を立案し実行していく必要がある。昨今ではデータサイエンティスト育成研修のようなものが多く出ているが、レベル3以上の専門家人材を育成するためには自社で自社のデータを使った育成プログラムを開発し運営することが必要となる。基礎的な部分は外部講習でカバーできるが、自社にとっての優秀な人材を育成するための汎用的なカリキュラムは存在しないし、バックグラウンドや特性も個人差が大きいので、実践の中でOJT方式で身に付けるのが最善である。

6. おわりに

企業経営の中に「データサイエンス」を組み込んでデータドリブン経営を成功に導くために重要となる要素をヤマトグループの取り組みを例に解説してきた。データサイエンスを成功に導くにはデータサイエンティストのような「専門家」だけではなく「経営・マネジメント」、「従業員」と企業に関わる人すべてがレベルを上げることが重要である。このような活動を継続することで数年後にはデータサイエンスが「民主化」され特別な組織や人材がいなくてもPCのオフィスソフトのように従業員だれでもが活用している時代が訪れるであろう。

参考文献

- [1] ヤマトホールディングス、「経営構造改革プラン『YAMATO NEXT100』を策定」, <https://www.yamato-hd.co.jp/news/2019/20200123.html> (2020年8月27日閲覧)
- [2] Wikipedia, <https://ja.wikipedia.org/wiki/データサイエンス> (2020年8月27日閲覧)
- [3] 週刊経団連タイムズ、「提言『AI活用戦略～AI-Ready社会の実現に向けて』を公表」(2019年2月21日 No. 3397), https://www.keidanren.or.jp/journal/times/2019/0221_02.html (2020年8月27日閲覧)