

特別対談

データサイエンス教育の
これから

東日本旅客鉄道株式会社
渋谷 直正 様

【渋谷様PROFILE】

2002年、日本航空 (JAL) 入社。月間2億PVに上るホームページのログ解析や顧客情報分析を担当。分析に基づくターゲティングでWebのバナー効果を改善するなど、売上向上に貢献。2014年、日経情報ストラテジー誌の「データサイエンティスト・オブ・ザ・イヤー」受賞。その後、デジタルガレージを経て、2021年6月、東日本旅客鉄道に入社。マーケティング本部 戦略・プラットフォーム部門 データマーケティングユニット 担当部長。主にSuicaなどのデータを使ったマーケティング施策を実施する他、社内でのデータマーケティング人材教育なども行う。

滋賀大学 データサイエンス学部
河本 薫 様

【河本様PROFILE】

1991年、大阪ガス入社。1998年から米国ローレンスバークレー国立研究所でエネルギー消費データ分析に従事。帰国後、データ分析による業務改革を推進。2011年からはビジネスアナリシスセンター所長として大阪ガスのデータ分析組織を定着させた。2013年、日経情報ストラテジー誌の初代「データサイエンティスト・オブ・ザ・イヤー」受賞。2018年4月より滋賀大学 データサイエンス学部 教授 兼 データサイエンス・AIイノベーション研究推進センター 副センター長。博士 (工学、経済学)。

「アップスキリング」をテーマに開催された2022年度のユーザーコンファレンス。特別対談には日本を代表するデータサイエンティストのお二人にご登壇いただき、「ビジネスに活かすデータサイエンス教育」の観点で、企業とアカデミック、それぞれの立場からお話いただきました。

(オーガナイザー:NTTデータ数理システム 田辺隆人)

データサイエンス教育と
アップスキリングへの向き合い方

——データサイエンティスト育成のスタンスについて。

河本 私は長年大阪ガスで、データと分析力で自社のビジネスを良くするとともに、分析力を持つ人材の育成に携わってきました。この経験は大学教育にも役立つのではないかと。そう考え、滋賀大学に赴任しました。大学でまず感じたのは「分かる」と「役立つ」の違いです。大学では、学問の醍醐味は世界で初めてのことが「分かる」ことなんだなと感じました。でも元企業人の私にとっては、分かるだけではなく、役立たなければ意味がない。特にデータサイエンスは、データをどう役立てるかを考えることが大事です。データ分析という「理」、考える「文」。その両方が行える人材を育てることを私のミッションと定め、日々学生と向き合っています。

渋谷 私はいくつかの企業を経験しましたが、基本的にはずっとマーケティングの世界にいます。マーケティングに携わるのはどちらかというと文系出身者が多く、もう10年以上、そうした人たちと一緒に仕事をしながら、彼らにデータ分析の文化を広げてきました。その中で自分が見つけた答えは、事業会社においては一般のビジネスパーソン、言わば文系のビ

ジネスパーソンを、データ分析人材としていかにアップスキリングするかが重要だということです。

データマーケティングに数学はいらない!?

——企業で役立つ人を育てるためにアカデミックとして注力していること。また企業にとってシチズンデータサイエンティストが必要な理由とは。

河本 企業のデータサイエンティストは、「課題は何か」を考えるとこころから取り組まなければいけません。それを学生のうちから学んでほしくて、私は企業のリアルな事例・データを使い、通常は社会に出てからしか学べないことを教えています。つまり「分かる」だけではなく、「役立つ」ことを成し遂げられるデータサイエンス人材の育成を、大学教育の場でも目指しています。

渋谷 現在の私の部署の社員は9割が文系です。彼らはマーケティングにおけるデータ分析の重要性を知っていますが、データを扱うことには心理的なハードルがある。そんな彼らに私は「データマーケティングに数学やプログラミングは不要」と言い切ります。本当は深くやろうとすれば数学もプログラミングも必要です。でも参入する時点では不要だし、それよりもハードルを取り除き、安心して成果を上げてほしい。シチズンデータサイエンティスト*としてアップスキリングしてほしいのです。ただし大学では理論や分析の背景が分かる人を育てていただきたい。車で言えば運転できるだけではなく、動くメカニズムやブレーキの構造も分かる、ということです。

*一般のビジネスパーソンのうち、Excelなどの分析ツールを駆使して分析する人材。専門人材の対比として、あるいはデータ分析の民主化の文脈で使われる。

河本 私もマーケティングのデータサイエンティストは社内、担当内で育てる必要があると思っています。そうでないとすべてが外部委託になり、大事な部分がブラックボックスになってしまうからです。本来、経営や事業の側で責任を持つべき部分が丸投げでは、データサイエンスという手段が本来の目的を見失わせてしまいかねません。できればシチズンデータサイエンティストを育てる一方で、少人数でも一定レベルのデータサイエンスの知識を持つ人を置くことを、企業の方々にはお勧めしたいと思います。それによってデータサイエンスと経営・事業をしっかりとつなげることができます。

——**渋谷様のところは、マーケティング担当の方が作った仮説を協力会社がモデル化するというように、うまくコラボレーションしていらっしゃる。**

渋谷 原則的には協力会社さんともコラボしながら、自前でやっています。ただ、中にはマーケティングで出した分析のアイデアを社外のプロにモデル化してもらうこともあります。河本さんが言われた「シチズンデータサイエンティストとは別にデータサイエンスが分かる人も必要」という話はその通りで、当社ではそれを私が担い、社外のプロからあがってくるモデルを精査・検証しています。他方、コラボレーションということでは、協力会社以外に大学との連携もあります。先日、ベイズ統計を使った柔軟なモデルを作る際は、自分たちだけでは難しかったのでモデリングを一部依頼しました。案件の9割はある程度定型でアプローチでき、1割はオーダーメイドが必要と思います。

なぜ「理」も「文」も必要か

——**渋谷様のところでは、モデル化する前の分析のアイデアや説明変数をどのように見つけているか。**

渋谷 特徴量を見つけ出すホワイトボックス型のAIを使ってデータから見つけることもあります。基本的には担当者が考えます。私たちは需要予測のために時系列モデルをつくることが多いのですが、重要なのは、実はパラメーターを一生懸命チューニングすることよりも、オリジナルの先行指標や要因を見つけることだったりします。例えば鉄道の利用客数を予測するとき、最近なら新型コロナの感染者数を勘案しますが、機械的に感染人数を入れても高い精度になりません。流行り始めた当時の感染者数1500人と、感染状況に慣れた今の1500人で、人の行動は異なるからです。そうした説明変数の加工には、一般常識や感覚、業界知識も必要で、機械で選ばせるよりも、人がアイデアを持って考える方が納得のいくモデルになります。

河本 予測に役立つ説明変数を考えるとき、人は頭の中で因果関係をたどりま。これこそが人の能力。一方で、アカデミアの世界では、説明変数をどう考え出すかより、モデリングの



ほうが重視されます。しかし私は、データサイエンスは実務であり、実務家を育てたいと考えているので、そこもしっかりと教えたい。そのためには、数学力だけでなく論理的思考力も求められる。論理的思考を表現するのは言語です。だから、私は、数学教育と日本語教育の両方を大事にしなければならぬと思っています。

——**リクルーティングで渋谷様が重視する人物像は。**

渋谷 私の立場からは数学やプログラミングのスキルよりも、コミュニケーション能力や、問題に対して興味を持ち、考える力を発揮できる人を求めます。問題の背景を考える。分からなくても考え抜く。どうしても分からなければ人に聞く、といったように、行動とコミュニケーションがとれる人を採用したいですね。

データサイエンス教育に数理システムの力を

——**これからNTTデータ数理システムに期待すること。**

河本 数理システムとは2008年からの付き合いで、データサイエンス系の相談をしたいときは真っ先に声をかけます。理由は信頼できるから。仕事をお願いすると、まず「現場の方に合わせてください」と言われます。単に「データ分析します」ではなく、課題を一緒に考え、何とか解決しようとしてくれます。数理システムが掲げる「データ分析の専門力を活かしながら、社会の問題を解決する専門家」という言葉通りの会社。そこはこれからも変わらずにいてほしいですね。

渋谷 数理システムにはブラックボックスなところがない。数理システムが発行している冊子『読本』シリーズでは、難しいとされている最適化という分野でも、数学を使わずに分かりやすく説明していますが、その姿勢は冊子の中だけに限りません。本当に分かっている人が、誰でも分かるように説明しようとしている姿勢がひしひしと伝わってくる。シチズンデータサイエンティストを増やしたいと考える者にとって、こんな心強いことはありません。このスタンスをぜひ守り続けてほしいです。