

トレンド・ワード

データサイエンス

生徒の学びや進路選択、そしてその後の人生に影響を与えるような革新的な技術や価値観を「社会のトレンド」として、「働く」「暮らす」「学ぶ」の観点から解説する本コーナー。今回は「データサイエンス」を取り上げる。行政や企業などで、意思決定をする際に活用されている「データサイエンス」。それを専門に学べる学部・学科がここ数年相次いで大学に新設され、人材育成が強化されている。なぜ、データ分析を基にした意思決定が重視されるようになったのか。データサイエンスではどのような資質・能力が求められるのか。企業のデータ分析及び分析人材の育成に携わる統計家の西内啓氏に話を聞いた。



解説者

株式会社データビークル
 取締役副社長
 CPO（最高製品責任者）

西内 啓

にしうち・ひろむ

東京大学助教、大病院医療情報ネットワーク研究センター副センター長等を経て、現職。企業向け分析ツールの開発、官民のデータ活用プロジェクトの支援に従事。著書に、『統計学が最強の学問である』（ダイヤモンド社）など。

サマリー

データを分析して根拠を導き、 経験がなくても価値ある意思決定を行う

統計学などを用いて、データから意味ある情報を見いだす

「データサイエンス」は、統計学や情報科学などを応用してデータを活用し、そこから何らかの価値を生み出す技術である。大きく分けると、2つのアプローチがある。

1つは、人間が社会や自然の複雑な挙動を理解するために、データを

収集・分析し、そこから意思決定や問題解決などを行うものだ。例えば、企業が消費動向の分析結果を商品開発に生かしたり、自治体が気象情報や被災記録などの分析結果を基に防災施策を立案したりといった活用がある。

もう1つは、人間を介さずに済むように、作業の自動化や効率化を図るものだ。いわゆるAI開発と呼ば

れるものがそれにあたり、人間の行動に関するデータから「人間だったらどのように処理するか」を高精度に予測できるようにし、機械がその行動を自動的に行えるようにする。

人や社会に関するデータを分析するデータサイエンスでは、統計学や情報科学の知識・技能に加え、人や社会への関心・理解が重要だ。統計家で株式会社データビークル取締役

副社長の西内啓氏は、こう説明する。

「データサイエンスに求められるのは、データから規則性や関連性などを見いだす統計的視点、データを適切に処理する計算機科学的視点、データの活用方法の発想や分析結果の解釈に必要な人間的視点です（図1）。それら3つの視点からのアプローチによって、データを社会に役立つ情報にすることができるようです」

エビデンス重視の意思決定と ITの進化の両面から発展

データサイエンスが活用されるようになった背景には、「エビデンス

図1 データサイエンスの3つの視点

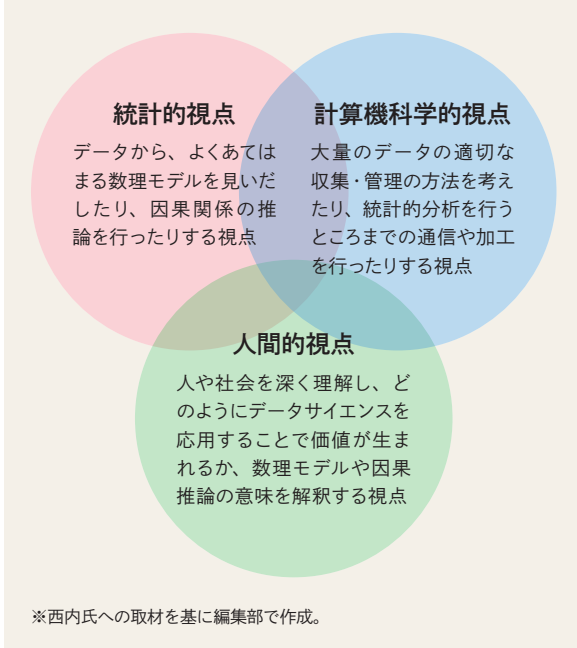
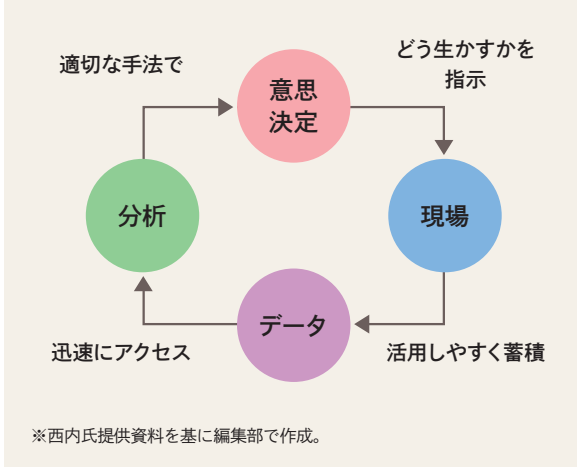


図2 データサイエンスを活用した意思決定のサイクル



ベースド」という考え方が、医療分野を中心に1990年代前半から広まったことが挙げられる。

「人の命を預かる医療では、統計学などの手法を用いてより効果的で効率的な治療法を見だし、科学的根拠（エビデンス）と医師の経験に基づき治療法を判断することが一般的になっていきました。その考え方は、効率的に大きな利益を目指す企業でも活用されるようになりました」

また、ITの進化によって、社会の隅々にまでネットワークが構築され、多様なデータを収集できるよう

になるとともに、膨大なデータも容易に計算することが可能になった。

「科学的根拠を基にした意思決定を重視する考え方も、膨大なデータを活用して有益な情報を見いだせるようになった環境整備が相まって、データサイエンスを活用して意思決定の精度を高め、より大きな成果に結びつける（図2）という動きが広がっていったのです」

データサイエンスは、誰もが身につけるべき素養に

データサイエンスの利点は、経験

がないことでも、データの分析結果を根拠として、最適と思われる意思決定ができることだ。例えば、複数の優秀な販売員の経験をデータ化してよりよい接客方法を見いだせれば、接客が未経験の販売員でも、商品の購入に結びつく接客ができるようになることが期待できる。

「データ分析により、これまで気づかなかつた傾向や想定外の関連性などが示されることがあります。それを基によりよい方法を生み出すこともできるのです。そうしたことから、データサイエンスの価値があ

ると言えるでしょう」

どのような業種や職種にも意思決定をする場面があることを考えると、データサイエンスの活用は、行政や企業にとどまらず、スポーツや芸術、教育など、あらゆる分野に広がっていくと考えられる。

また、技術の進歩によって高度なデータ分析の自動化が進めば、専門家でなくても、意思決定や問題解決などにデータを活用する「市民データサイエンティスト」(*)が活躍するようになると思われる。

「統計的視点や計算機科学的視点をITの技術でサポートできるようになれば、データサイエンスを活用するハードルはぐっと下がります。そうならば、仕事や生活のあらゆる場面で活用できるデータサイエンスは、すべての人に求められる素養となっていくのではないかと考えています」

●次ページからは、「働く」「暮らす」「学ぶ」の3つの切り口で、データサイエンスによる社会や生活の変化を具体的に見ていく。

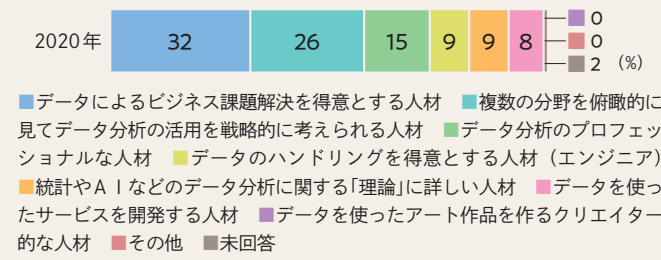
* アメリカのリサーチ&アドバイザリ企業「ガートナー」が提唱した概念。

働く

専門性×データサイエンスで仕事の幅が広がる

企業では、利益に直結する商品開発や販売戦略などに活用できる手法として、データサイエンスが広まった。現在では、社員の勤怠や評価などのデータを分析して、労働環境を改善したり、人員配置や採用を行う

図3 今後3年間で採用・育成したいデータサイエンティストの人材像



※一般社団法人データサイエンティスト協会「データサイエンティストの採用に関するアンケート」(2020年調査)を基に編集部で作成。

たりするなど、企業経営全般に積極的に取り入れられつつある。

それに伴い、複数の分野や全体を俯瞰できるデータサイエンティストの採用・育成のニーズが高まっている(図3)。

「日本では元々、統計学を駆使して、品質の高い製品を生み出し、製造工程を改善したりしてきました。それを強みとして、日本の経済は発展してきたと言えます。科学的根拠を基にした意思決定の文化は、製造の業務では定着していますから、マーケティングや営業、人事などの業務にもその文化が浸透すれば、日本全体の経済成長が期待できるでしょう」

データサイエンスの専門家でなくても、自分のやりたい仕事の専門性に加えてデータサイエンスの素養も身につけていれば、仕事の幅を広げられるだろう。

「例えば、研究者になれば、実験や調査の結果を分析するので、デー

タサイエンスが必要です。また、会社員でも、管理職になれば、部下の過去の業務実績などを基に適性を分析し、誰にどの仕事を任せられるかを判断することが求められます。自分

の専門性とデータサイエンスによる分析力のかけ合わせは、より適切な意思決定や問題解決を可能にし、仕事を上で大きな強みになります」

暮らす

根拠を求める姿勢が、生活の質向上につながる

住民・消費者としてデータサイエンスの視点を持つことは、生活の質向上のためにも大切だと、西内氏は語る。

「日常生活でも、データサイエンスの視点を持つと、自治体や企業から十分な説明がないままに、非効率な施策やサービスが漫然と行われていたら、それに気づけるかもしれません。そこで住民・消費者として、根拠を求める姿勢をさせば、自治体や企業は効果的な施策やサービスの提供に努めるようになるでしょう。多くの人がそうした姿勢を持てば、社会全体がよりよくなっていくのではないのでしょうか」

新型コロナウイルスの感染拡大防

止においてもデータサイエンスは大いに活用されており、それは暮らしに直結している。例えば、いわゆる「3密」を回避する方策が示されたのは、保健所の疫学調査の結果、感染者と非感染者の違いとして、「密閉・密室・密集」の3つがリスクファクターとして見いだされたからだ。

日常生活や人生においても意思決定を行う場面は多々あり、データサイエンスの視点を生かすと有効な判断を下せる場合もある。例えば、様々な学習方法の中から、研究で効果的とされている方法を探し、試してみ、自分に合う方法を見いだすようにすれば、迷いなく効率的に学習を進められるだろう。

新型コロナウイルスの感染拡大防

学ぶ

データサイエンスの素養を育む教育が活性化

高校では、数学や情報で統計学の学習内容が拡充

社会におけるデータサイエンスの活用拡大を受け、大学はデータサイエンスの育成に力を入れていく。17年度、滋賀大学が日本初のデータサイエンス学部を設置。以降、データサイエンスに関する学部・学科の新設が相次いでいる。

19年度、国が策定した「AI戦略2019」では、AIに対応した人材育成の一環として、すべての大学生・高等専門学校生が初級レベルの数理・データサイエンス・AIを履修するための教育課程の整備が打ち出された。21年度には、そのうちの優れた教育プログラムを認定する制度(図4)がスタートした。高校では、新学習指導要領において、数学科や情報科で統計学やデータ活用などに関する学習内容が拡充される。

て、西内氏は次のように語る。

「数学では、幾何学以外の分野は統計学の基礎となるので、高校時代にしっかり習得しておくことよいでしよう。地理や現代社会では、人口や産業などのデータを用いて学びますから、グラフの表示の仕方やデータの読み取り方など、データ活用を実践的に学べる科目として重要です」

科学的根拠を基に考察する学習としては、仮説を立てた上でデータを集めて分析したり、条件の違いの統計的有意性を実験で調べたりといった活動を行う理科や探究学習が挙げられる。それらの学習で重要なのは、懐疑的な思考力を持つことだという。「つい素直に納得できそうなことでも、あえて立ち止まって、疑問を持つようにすると、理屈の穴を見つれたり、別の方法に気づいたりすることがあります。日頃から、そうした意識で物事を見るようにするといでしょう」

図4 「数理・データサイエンス・AI教育プログラム認定制度 (リテラシーレベル)」概要

種類	主な要件	認定数
認定教育プログラム	<ul style="list-style-type: none"> 大学、短期大学、高等専門学校の正規の課程 学生に広く実施される教育プログラム(全学開講) 具体的な計画の策定、公表 学生の関心を高め、かつ、必要な知識及び技術を体系的に修得 学生に対し、履修を促す取り組みの実施 自己点検・評価(履修率、学修成果、進路等)の実施、公表 当該教育プログラムを実施した実績のあること(人文・社会学等を含む複数学部等からの履修) 	大学 66校 短期大学 2校 高等専門学校 10校
認定教育プログラムプラス	<ul style="list-style-type: none"> 上記認定要件を満たすこと 学生の履修率が一定割合以上(全学生の50%以上(3年以内に達成見込みも可)) 大学等の特性に応じた特色ある取り組みが実施されていること 	大学 10校 高等専門学校 1校

認定数は、2021年8月時点。
※文部科学省「数理・データサイエンス・AI教育プログラム認定制度(リテラシーレベル)」を基に編集部で作成。

個別最適な学習をより効果的に行えるように

データサイエンスは、学習方法にも変化をもたらしている。その一例は、学習者の理解度や弱点などに応じた教材を自動的に提供するアプリケーションだ。

「データサイエンスを活用したアプリケーションによって、学習者が1人で自分の課題に応じた学習を効率的に取り組めるようになりまし。個別最適な学習の実現は、そうしたアプリケーションの活用が1つの鍵になると考えています」

学習改善や授業改善にも、データサイエンスは有効だ。例えば、学習者を2つのグループにランダムに分け、それぞれ違う方法で学習した結果、どのような違いが生じるかを比較すれば、効果的な学習方法や指導方法を見いだすことにつながる。

「科学的根拠を基にした学習改善や授業改善は、データを活用すれば物事を効果的に改善できると、生徒が実感できる身近な具体例になるでしょう。加えて、社会をよりよくするために既存のことは見直す重要性を、生徒に伝える機会にもなるのではないでしょう」