

特集 データサイエンス教育を知る

データ駆動型超スマート社会を支える
データサイエンス教育

慶應義塾大学大学院
健康マネジメント研究科教授・ 渡辺 美智子
放送大学客員教授

1. 第4次産業革命とデータサイエンス
人材育成

IoT、ビッグデータ、人工知能、ロボットに代表される第4次産業革命技術革新が産業や社会のあり方を大きく変えようとしている現在、その変革を担うデータサイエンス・データアナリティクス系人材の不足は世界各国で大きな課題となっています。

すでに欧米諸国や中国では、ビッグデータ・オープンデータ活用が本格化した2010年以降から、コンピュータサイエンスと応用統計科学を融合したデータサイエンス学部（学科）や研究科、学位の創設を行ってきていますが、それと並行して初等中等教育における共通コアカリキュラムでのICTを活用した実践的統計内容の強化、高校におけるデータサイエンスコースの新設など、この領域の教育体系の構築と拡大を質と量の双方で急ピッチに進めています。

その一方で、従来、統計を専門とする学部や研究科が大学になく、また、統計の授業といえば数理統計的な理論が主であった日本においては、データ分析系人材の欠如は、欧米・中国等の海外諸国とは比べものにならないほど大きいと言えます。そのため、科学技術イノベーション戦略（平成27年6月閣議決定）において、「我が国は欧米等と比較し、データ分析のスキルを有する人材や統計科学を専攻する人材が極めて少なく、我が国の多くの民間企業が情報通信分野の人材不足を感じており、危機的な状況にある」が示され、続く第5期科学技術基本計画（平成28年1月閣議決定）において、「データ解析やプログラミング等の基本的知識を持ちつつビッグデータやAI等の基盤技術を新しい課題の発見・解決に活用できる人材などの強化を図る。」方針が明記されています。現在各府省では、この方針の具体化に向けた政策が進められているところです。

文部科学省は、「第4次産業革命に向けた人材育成総合イニシアティブ～未来社会を創造するAI/IoT/ビッグデータ等を牽引する人材育成総合プログラム～」（平成28年4月）において、「次代を拓くために必要な情報を活用して新たな価値を創造していくために必要な力や課題の発見・解決

にICTを活用できる力を発達の段階に応じて育成」として、第4次産業革命技術の基盤となる数理・情報・データサイエンス等の人材育成・確保に資する施策を初等中等教育、高等教育から研究者レベルでの包括的な人材育成総合プログラムとして体系的に実施することを示しています。

実際に、初等中等教育における学習指導要領の改訂、大学における「数理及びデータサイエンスに係る教育強化」の拠点校の選定、滋賀大学、横浜市立大学データサイエンス学部および立教大学データサイエンス副専攻の設置などが実現していますが、今後も初等中等教育における統計教育等の体系化と充実、大学での専門学部の新設、大学院等データサイエンスに係る教育・研究機能の強化、研究・教育における企業と大学との連携、実務においてデータ分析を担う専門家の職掌の把握とキャリアパスの確立、社会人再教育の仕組みの構築などが進んでいくものと思われます。

2. 高大接続と学習指導要領の改訂

データサイエンス教育自身は、ビッグデータを基盤に大規模な社会問題の解決に向けた高度なデータアナリティクス系人材の育成だけを目指すためのものではなく、基本的な日常の意思決定や判断において、データを客観的なエビデンスとして科学的な思考および統計的な思考で対峙する能力養成自身を指し、国民全体に涵養すべきキーコンピテンシーとしても位置づけられるものです。

特に、データに基づく科学的意思決定力の強化は、21世紀型スキル教育やグローバルリーダーシップ教育の中でも最重要視されている内容であり、日常の事象に関する観察からの気づきを確率的現象として統計数理モデルで捉え、現象間の関連性のパターンをモデルで評価するというデータ思考力（Think with Data）の涵養が基本です。その上で、将来の各分野のデータサイエンス専門職能の育成も見据えた教育、この双方の視点が重要です。そのため、国際的には学会等が示したガイドラインに沿って、学校教育の早い段階から身の回りの課題発見と問題解決にデータに基づく確率的思考で判断したり、統計グラフ等の活用して意思決定するプロジェクトベースの教育が実施され

ています。

日本でも現行の学習指導要領に従って、2011年度より小学校、2012年度に中学校(前倒し実施)、2014年度に高校と主に算数・数学教科の中で、活用を重視したデータの分析に関連した内容領域が30年ぶりに必修として設定されています。とくに数学Iのセンター試験では、ヒストグラム、箱ひげ図、散布図、相関係数などが現実の文脈ベースの解釈を中心に出题されていることから、この領域に対する学校現場の関心も高まっています。さらに、2020年度から順次完全実施、移行期間中の先行実施も可能とした次期の学習指導要領(2016・2017年度告示)では、小学校からのプログラミング教育導入と併せ小中高校を通じて統計内容の更なる拡充が図られています。数学教科では記述統計に加え推測統計までを文系理系を問わず多くの高校生が学習し、必修化される情報Iの中に「データの活用」や情報II(選択)の中に「データサイエンス」の単元が含まれます。入試改革と併せ、科学的探求・問題解決・意思決定のプロセスを通してデータとデータ分析に基づく思考力・判断力・表現力が主体的・実践的アクティブラーニングによって育成されようとしています。しかし、これまでなかった新しい内容領域であるため、教員養成や教員研修等に果たす大学の役割も大きくなっていると思われます。

3. データサイエンススキルと価値創造

データサイエンスは、数理・統計に関する分析スキルおよび分析的思考力、実際のデータおよびデータベースに対するコンピューティングスキル、データの背景に関する知識と経験知(エキスパート)の融合から、データの文脈に沿って価値を生み出す学際的な領域と考えられています(図1)。そこで使用される分析の方法は、取り扱うデータ系の進化に伴って、従来の統計学で分析手法も扱っていた手法から新しいアルゴリズム的手法へと高度化しています(図2)。

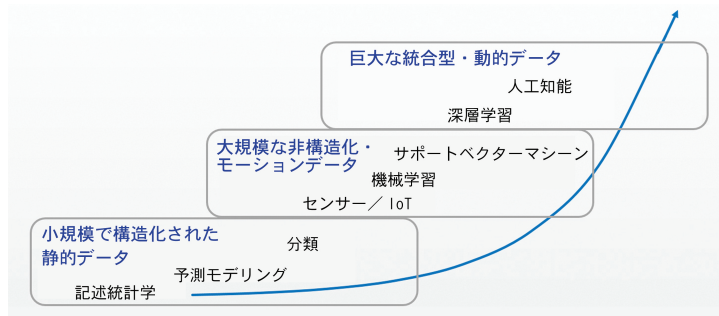


図2 データ体系と分析技術の進化 (* 著者和訳)

出典 J. L.s Priestley : The Challenge of Transdisciplinary Science The Case of Data Science., SAS GLOBAL FORUM2018

一般に、分析の価値の高度化は、記述的分析(何が起きているのか?)から要因探索的分析(なぜ起こったのか?)、そして予測的分析(何が起こるのか?)から、予測モデルによる制御・最適化(どうやって望ましい状態にしていくのか?)の順に実現します。ただし、問題解決の成否は、データ分析技術が高度であればよいということではなく、分析の前段階である、現象への理解と仮説の形成および分析の後段階である、現象の文脈に沿った結果の解釈とそのエビデンスに基づく現象への意思決定の適切性で決まってくるので、データサイエンススキルの育成に関しては、ICTのスキルや数理的・統計的分析手法などの個別の方法論の習得と並行して、問題解決と価値創造の一連のプロセスへの理解を経験的に積上げることが重要と考えられています。

4. 放送大学データサイエンスエキスパートプラン

データサイエンスを大学学部教育で行う場合、統計科学、情報コンピュータサイエンス、数理科学、問題解決・コミュニケーションを含むビジネスおよびマーケティング科学の複合領域として科目構成されています。例えば、放送大学では、学校教育基本法改正による科目群履修認証制度エキスパートプランを実施していますが、今年度秋よりその中に「データサイエンス」プランを新規に開設し、上記の関連科目群の中から必要単位数を取得した学生に「データサイエンス」の認証を与える準備をしています。

5. おわりに

実世界のデジタル化とデータ流通が急速に拡大していく中で、データの処理と解析結果を実世界へフィードバックする能力育成をいかに学生に図っていくのか、その体制作りが今大学に求められ、喫緊の課題になっています。この新しい学際領域に対して産学官の連携による教育体系が期待されています。

データサイエンティストに求められる能力

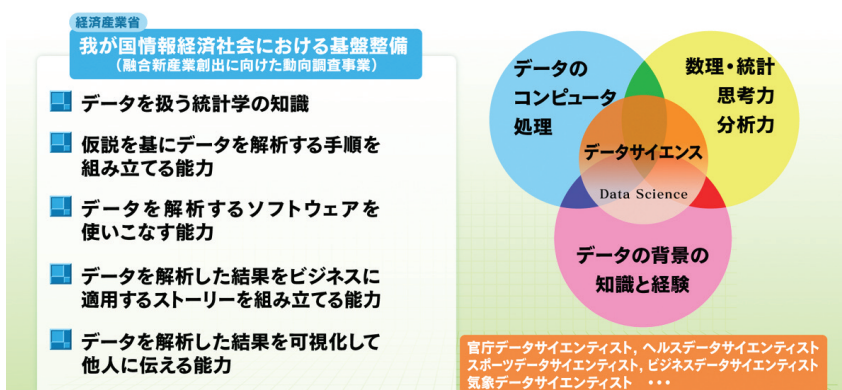


図1 データサイエンススキル (出典 放送大学TV「身近な統計」)