

特集にあたって

椿 広計 (統計数理研究所)

日本で最初に「データサイエンス教育」の名で全学部の講義を実施したのは、1997年慶應義塾大学SFCである。私は、その立ち上げに参加した。文理を問わず全新生のモバイルPCに統計ソフトウェアJMP学生版をプレインストールし、「知的創造のための思考の工具箱」を標榜した。新生にデータとは何か、どう収集するか、データの質とは何かといった概念的教育を行ったうえで、ソフトウェアに支援された統計グラフィックスから始めて、重回帰分析や主成分分析などの「データ分析」の面白さを経験させた。興味が生じたら次学期に選択科目で最尤法などを学ぶという逆転の発想である。学生自ら集めたデータを分析し、レポートを書き、優秀なレポートを表彰する。翌年は授業を聴いた2年生がTAとなり、教員が書いた初年度教科書をわかりやすく手直しするという、なかなか得難い盛り上がりを経験した。自身は、筑波大学社会人大学院で第2世代人工知能を含む探索的データ解析と統計モデルに基づく仮説検証型データ分析教育を2006年頃まで実施していた。自身のその後の統計教育が完全に変わった、楽しい思い出である。

それから20年以上経ち、2019年9月から本年3月まで、「数理・データサイエンス・AI教育プログラム認定制度検討会議」が内閣府などを事務局とし、国立大学協会・永田恭介会長（筑波大学学長）を座長として行われた。わが国50万人の大学生に文理を問わず数理・データサイエンス・AIのリテラシープログラムを提供し、そのGood Practiceを認証するという取り組みである。並行して、2020年4月拠点大学（北海道大学、東京大学、滋賀大学、京都大学、大阪大学、九州大学）を中心とする「数理・データサイエンス教育拠点強化コンソーシアム」が、「数理・データサイエンス・AI（リテラシーレベル）モデルカリキュラム～デー

タ思考の涵養～」も発表した。このリテラシーレベルにつぐレベルの教育認定に関わる検討会議も2020年9月上旬からスタートした。

近年データサイエンス概念が二つの意味で変化した。一つは2012年頃、経営学者ダベンポートが主唱した、ビッグデータハンドリングのための情報処理能力、「データエンジニアリング」側面の強調である。もう一つは、「現象のモデリング」の後の意思決定という2ステップアプローチとは異なる、一段階最適化のデータアナリティックスが急成長したことである。数理最適化と統計的機械学習を中核数理とする第3世代人工知能が急進したのである。この状況の中で、OR学会誌から「データサイエンス教育の潮流」という特集企画を持ち掛けられた。今回の対談では滋賀大学、広島大学、北海道大学のデータサイエンス教育拠点がどのようなポリシーで人材を育成しようとしているのかを語っていただいた。また、大阪大学、広島大学、早稲田大学におけるデータサイエンス教育を紹介いただいた。実は、討論会においても敢えてデータサイエンスとは何かということにあまり立ち入らなかった。データサイエンスをどうとらえるかは、各大学の判断に任せたのである。頂戴した原稿や意見はある意味で価値最適化というデータサイエンスの目的に対してOR的側面が極めて明瞭に出たものとなり、興味深い。各大学は、データサイエンスの統計側面、情報処理側面、数理最適化側面、マネジメントサイエンス側面をどのようなバランスで教育に実装するか、育成すべき人材像を明確にしたうえで、特徴ある教育システムを構想する必要がある。本特集に協力いただいた先生方に敬意と謝意を表したい。本特集がわが国のデータサイエンス高等教育のデザインの一助となれば幸いである。