

機関誌「データ科学」発刊にあたって

早稲田大学データ科学センター所長 理工学術院教授・松嶋敏泰

「現在の国勢を詳明せざれば、政府すなわち施政の便を失う。過去施政の結果を鑑照せざれば、政府その政策の利弊を知るに由なし。」

これは、早稲田大学の創設者である大隈重信が130年以上前に残した言葉で、現在の国の情勢を詳細に明らかにしなければ、政府は政治を執り行うことができない。また、過去の施政の結果と比較してみなければ、政府はその政策のよしあしを知ることができない、と統計（データ）からの意思決定の重要性を既に述べている。

データからの意思決定は、従来は、理工学などのデータが実験等で取得でき数学等で論理的に結論を導出できる（科学的方法と呼ばれる知の創造プロセスを用いた）科学分野や、データ取得や分析に多額の費用がかけられる一部の重要な意思決定でのみ行われていた。ところが、近年、情報・通信インフラの発展と、データを分析する理論や技術の進歩により誕生してきたデータ科学（データサイエンス）により、以前は不可能と思われていた広範な科学・学問領域や様々な意思決定問題においても、データを用いた科学的方法による意思決定が可能となってきている。

このデータ科学により、我々の意思決定や知の創造のプロセスが大きく変わり、社会に大きな変革が起こると捉えられている。知の創造と継承を担う拠点である大学は、この大きな流れを先導する核となることが期待されており、「早稲田大学データ科学センター」は大隈の意を源流として、2017年12月に全学的な組織として設置された（設立時の名称はデータ科学総合研究教育センター）。データ科学センターでは、教育面では様々な分野でデータを活用した意思決定ができる有用な人材の育成を目指し、研究面ではデータを基にした知の創造プロセスや意思決定プロセスの進化・深化により様々な分野における革新的研究や問題解決を目指してきた。

データ科学は、分野を問わないデータからの意思決定の科学と捉えられるメタ科学であり、学問としては体系化の途上にある発展中の学問でもあるため、上記の目的を目指したデータ科学センターの活動には様々な課題が存在している。例えば、専門の違う様々な学生に未だ寄せ集め的で学際的なデータ科学をどのように身に付けさせるか、社会における多様なデータからの様々な意思決定を大学のみで扱えるのか等の課題である。センターではそれらの課題に対して新たな試みの活動で解決を図ってきた。

まず、センターでは、全学の学生を対象として、データ科学のコアとなる考えかたを、統計学や機械学習やAI等の寄せ集め的な知識ではなく、統一した視点から身につけてもらうことを目的とした独自のデータ科学教育プログラムを構築した。この教育プログラムはオンデマンド授業をベースとした学習システムにより提供され、それぞれの学生の専門性、基礎力、目指すデータ科学の活用分野に合わせて、柔軟なカリキュラムが選択できるようにしている。さらに、その履修状況によりリテラシー級から上級までの認定も受けられる制度まで整えることができた。

センターの活動はこのような正規科目の教育プログラムの提供に限らず、テーマ、解析目的、利用データ等の条件を様々に変えて毎年全学規模で開催しているデータサイエンスコンペティション、データを用いた様々な分野の専門研究について個別相談が受けられるデータ科学研究相談など多岐に渡っている。このような正規科目提供以外の活動も加わることで、総合的なデータ科学の考え方や実践力が身に付き、様々な研究や実務の分野で活躍できる人材が育成されると考えている。

また、社会における多様なデータを用いた様々な意思決定についての教育と研究を行うためには、企業と大学の連携協力は欠くことができない。まず大学では入手できない多様なデータ、特に公的機関や企業の貴重なデータを教育研究に利用できることは大変ありがたい。このようなデータを利用可能とするため、企業と大学間の秘密保持契約のような法律的な面の対応から、研究者や学生の利用者の利便性を重要視しながらもセキュリティを保証する情報システムまで含めたトータルの仕組みとして「Waseda Integrated Research Platform (WIRP)」を構築している。

社会での様々な意思決定におけるデータ科学の活用については、学生が企業での実務体験としてそれらを学ぶ教育的インターシップも行われている。またその逆の方向として、センター独自のデータ科学教育プログラムは、専門や基礎知識のばらつきを吸収する柔軟な教育プログラムとなっているため、既に業務で培われた専門性に加えデータを活用するスキルをもった人材の企業での育成にも利用可能となると考えている。さらに社会人が実務におけるデータ活用を通じてデータ科学を実践的に学ぶプログラムも始まっており、センターの提供するこれらの社会人教育プログラムにより、社会全体を対象としてデータ科学を活用できる人材の育成も視野に入れている。

このように企業と大学が一体となった教育と研究における連携により、社会全体におけるデータ科学の浸透と発展が期待できると考えている。その連携の仕組みの一つとして、産官学連携部会、社会人教育部会、キャリア教育部会の3つの部会で構成される「早稲田大学データ科学研究教育コンソーシアム」の活動も行っている。

これらのデータ科学センターの教育と研究の定常的活動や、研究プロジェクトから様々な成果があがってきているものの、それらを蓄積し発信する仕組みは現在まで整っていなかった。正規科目などでは、オンデマンドコンテンツや教科書として、個別の研究プロジェクトにおいても学会発表や論文としてその成果が蓄積され発信されてきてはいるものの、その他の多岐にわたる活動の内容や成果についての情報の蓄積と発信は十分ではなかった。

分野を問わず社会全体に大きな影響を及ぼし、さらに発展し続けているデータ科学の教育と研究においては、ここまで述べてきたような多岐にわたる総合的で新たな試みが必要であり、データ科学センターが試みてきているそれらの情報を整理蓄積し発信することは大変有意義であると考えられる。その目的のため、早稲田大学データ科学センターでは機関誌「データ科学」をここに発刊する運びとなった。この機関誌により、まずはセンターの活動の一部の成果の発信からスタートすることになるが、徐々にセンターの多岐にわたる活動の多面的な成果を整理し発信できるよう発展していくことも考えていきたい。皆様のご協力のもと、この機関誌が直接的に間接的にデータ科学の発展に少しでも寄与できるような仕組みとして育っていくことを期待している。