

数理・データサイエンス教育強化拠点コンソーシアムが目指すもの



数理・データサイエンス教育強化拠点コンソーシアム議長
北川 源四郎
(東京大学数理・情報教育研究センター 特任教授)

コンソーシアムの設立とその背景 標準カリキュラム・教材を協働して作成 他大学への普及方策を検討実施

人間社会の発展の歩みは一律ではない。P. F. ドラッカーが指摘したように、人間社会では数百年に一度、あるとき急激な変化が始まり、世界観も価値観も社会構造も政治構造も技術や芸術も変えて、やがてそれまでは想像できなかったような新しい世界が拓ける^[1]。われわれは今まさにそのような歴史の転換期のまっただなかにいる。言うまでもなく、この転換の原動力は情報通信及びセンサー技術の飛躍的発展によって出現したビッグデータである。社会の隅々に配置されたセンサーからは日々大量のデータが生み出され、ネット社会には人間社会のデジタルデータが大量に蓄積している。その結果、自動運転、画像認識、医療診断等の先端技術が急速に発展するだけでなく、通信・情報発信、情報検索、販売・流通、コンピュータゲーム等個人の生活に密着した社会の隅々に至るまで、これまでの社会常識は一変している。かつて蒸気機関が産業革命を引き起こし工業化社会と資本主義社会を生み出したように、ビッグデータの出現とその利用技術の飛躍的発展が新しいデータ駆動型社会への転換をもたらそうとしている。

データから知識を獲得する方法としての統計学とデータサイエンスは、生命が長い進化の過程で獲得した知的情報

処理の能力をサイバー世界において模倣したものに過ぎないともいえる。しかし、近年の情報機器とAI等の情報処理方式の飛躍的発展は人類を知的情報処理のスケールの限界から解き放ち、あたかも量が質に転化したような様相を呈している。また、これまで個人が習得した知識や経験を他者に転移する方法が極めて限定されてきたのに対して、サイバー世界に構築された知識や情報処理方式はほとんど一瞬のうちに大きなコストなしに移転し、共有できることも発展の速度を上げている要因である。

このようなデータ駆動型社会への転換にあたっては、データに基づく知識獲得や価値創出が実現の鍵となる。これまでの経済発展を支えてきたのは、土地、資本、労働であったとすれば、今後のデータ駆動型社会の発展の原動力は、データそのものとデータからの知識創出の方法である。そのため、理論科学、実験科学、計算科学に加えデータサイエンスが第4の科学的方法論として重要になっている。今後は文系の社会人にとっても科学・技術研究に携わる理系の研究者にとってもデータサイエンスは不可欠な素養となる。

このような歴史的变化をいち早く認識した欧米では21世紀初頭から統計教育の強化が始まり、さらに2010年代に入るとデータサイエンスの教育プログラムが急速に立ち上がってきた。こうした世界的潮流の中で、遅ればせながらわが国においても、データ駆動型の社会実現の鍵として数理・データサ

イエンス教育強化の取り組みが急速に進められるようになった。

2014年の日本学術会議提言^[2]ではデータ科学を専門とする教育組織の設置、基幹的研究組織における恒久的なデータ解析部門の設置が必要であることを指摘するとともに、当面の対応策として、日本版のインサイト・プログラムの早急な設置を提言している。また、これを受けて2015年の産学官懇談会報告^[3]では、毎年500人規模の棟梁レベル育成及び全国50万人のリテラシーレベルや独り立ちレベルの大学教育を加速させるために、主要10大学程

度で人材育成をスタートするとともにMOOC等のオンライン教材を整備し、全国への波及効果を狙うことを提言している。

2017年度には文部科学省の数理・データサイエンス教育強化方針に基づき、北海道大学、東京大学、滋賀大学、京都大学、大阪大学、九州大学の6校が拠点校として採択され、各大学における全学数理・データサイエンス教育の実施に向けた取り組みを開始するとともに、数理・データサイエンス教育強化拠点コンソーシアム(以下、コンソーシアム)を形成して全国普及に向けた活動が開始された^[4]。このコンソーシアムの主要な役割は全国的なモデルとなる標準カリキュラム・教材を協働して作成するとともに、他大学への普及方策を検討し実施することである。2019年度には全国全大学への普及を加速するための20の国立大学が協力校として採択され、さらに2020年度には協力校3校、特定分野協力校7校が新たに加わった。

コンソーシアムとしての取り組み内容 カリキュラム、教材、教育用データベースに 関する3分科会を設置

コンソーシアムはその主要なミッションである全国的モデルとなる数理・データサイエンスの標準カリキュラムや教材を作成し普及するために、カリキュラム、教材、教育用データベースに関する3分科会を設置し全国普及に向けた活動を行ってきた。2019年10月にはAI戦略2019に対応して特別委員会を設置した。

図1:モデルカリキュラム(リテラシーレベル)の構成([5]より転載)

導入	1. 社会におけるデータ・AI利活用		
	1-1. 社会で起きている変化	1-2. 社会で活用されているデータ	
	1-3. データ・AIの活用領域	1-4. データ・AI利活用のための技術	
	1-5. データ・AI利活用の現場	1-6. データ・AI利活用の最新動向	
基礎	2. データリテラシー		
	2-1. データを読む	2-2. データを説明する	
	2-3. データを扱う		
心得	3. データ・AI利活用における留意事項		
	3-1. データ・AIを扱う上での留意事項	3-2. データを守る上での留意事項	
選択	4. オプション		
	4-1. 統計および数理基礎	4-2. アルゴリズム基礎	
	4-3. データ構造とプログラミング基礎	4-4. 時系列データ解析	
	4-5. テキスト解析	4-6. 画像解析	
	4-7. データハンドリング	4-8. データ活用実践(教師あり学習)	
	4-9. データ活用実践(教師なし学習)		

(1) カリキュラム分科会

この分科会は、全国的なモデルとなる標準カリキュラムを協働して作成し普及に取り組むことを目的としている。ただし、全国800近くの多様な大学に同一のカリキュラムを設定することは現実的でも妥当でもないため、当分科会では数理・データサイエンス教育のためのスキルセットと参照基準を作成し、各大学の自発的なカリキュラム開発を支援することを目指している。

スキルセットとしては海外の動向等も考慮して以下のような大分類を想定している。

- データの法規・倫理
- データの記述・可視化
- データの取得・管理・加工
- 統計基礎
- 数学基礎
- 計算基礎
- モデリングと評価

これに導入部としての「データサイエンスを学ぶ意義」を加えた8つの大分類で構成されている。各大分類ではリテラシーレベル、初級、中級、上級の4段階のスキルを設定するとともに、それぞれのスキルの学修目標を提示して、各大学の特色や事情によって選択できるようにする予定である。

(2) モデルカリキュラム(リテラシーレベル)の全国展開に関する特別委員会

国のAI戦略2019に対応して、数理・データサイエンス・AIのリテラシーレベルのモデルカリキュラムを策定する必要が生じたため、2019年度に産業界、公私立大学の有識者を中心に特別委員会を設置して策定にあたった。意見募集を経て2020年4月15日には「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～」^⑤を公開した。このモデルカリキュラムは今後のデジタル社会において、数理・データサイエンス・AIを使いこなすための基礎的素養を身につけ、人間中心の適切な判断・説明・活用ができ、AI等の恩恵を享受できるようになることを目標としており、図1のように導入、基礎、心得、選択の4つのカテゴリーから構成されている。モデルカリキュラムを参考に各大学の事情に応じて、2単位程度の科目を新設したり、既存のコースに埋め込んで実装することが期待される。

(3) 教材分科会

この分科会は全国的なモデルとなる教材を協働して作成し普及に取り組むことを目的として設置されたもので、

- データサイエンス教科書シリーズの企画編纂
- eラーニング教材、講義動画等の統合的配信方法の検討
- 普及方法やその他の可能性の検討

をミッションとしている。すでにデータサイエンスの入門教科書シリーズを企画・編集し、現在までに9冊が刊行され、近日中に全10巻が完結する予定である。これに加えてリテラシーレベルのモデルカリキュラムを発表したことに対応し、数理・データサイエンス・AIのリテラシーレベルの入門書の出版を企画中である。

私立文系を含め全大学等での数理・データサイエンス・AIのリテラシー教育を普及するためには、産学官懇談会の提言でも指摘されたようにMOOC等のオンライン教材の整備が必要である。各拠点校では既に教材の作成が進んでいるので、教材分科会では教材ポータルサイトを設置して、コンソーシアム所属の大学へのリンクを張り、滋賀大学のMOOC教材等のコンテンツにアクセスできるようにしている。これに加え、リテラシーレベルモデルカリキュラムの発表に対応して、カリキュラムのキーワードに対応するオンライン教材(動画、

スライド)を開発し、公開の予定である。

この活動の一環として、東京大学ではスライド教材「数理・データサイエンス・AIリテラシーレベル教材」^⑥を開発し、公開している。この教材は、モデルカリキュラムに完全準拠してオプションも含め全てのキーワードをカバーしていることと、教育目的に関してはCC BY-NC-SAライセンスのもとで、ページ単位で自由に利用できることが特徴である。

(4) 教育用データベース分科会

データサイエンスの教育では現実世界の実課題と実データに触れることが重要である。本分科会はデータサイエンス教育用のデータを収集し、各大学が使用できる環境を整備することを目的として設置されており、

- 教育用各種データの収集・公開に関する検討
- 既存の公開データベース情報に関する検討
- オープンソース等の情報に関する検討

をミッションとしている。

既にアカデミック及び官庁データのうちカリキュラムとの関係も考慮しつつ、データサイエンス教育に適切と考えられる27のデータを選定している。また教育用データ提供システム(データ提供ポータル)の仕様について検討し、主査大学の北海道大学において実装し、試験運用してレビューに基づく改善を実施している。

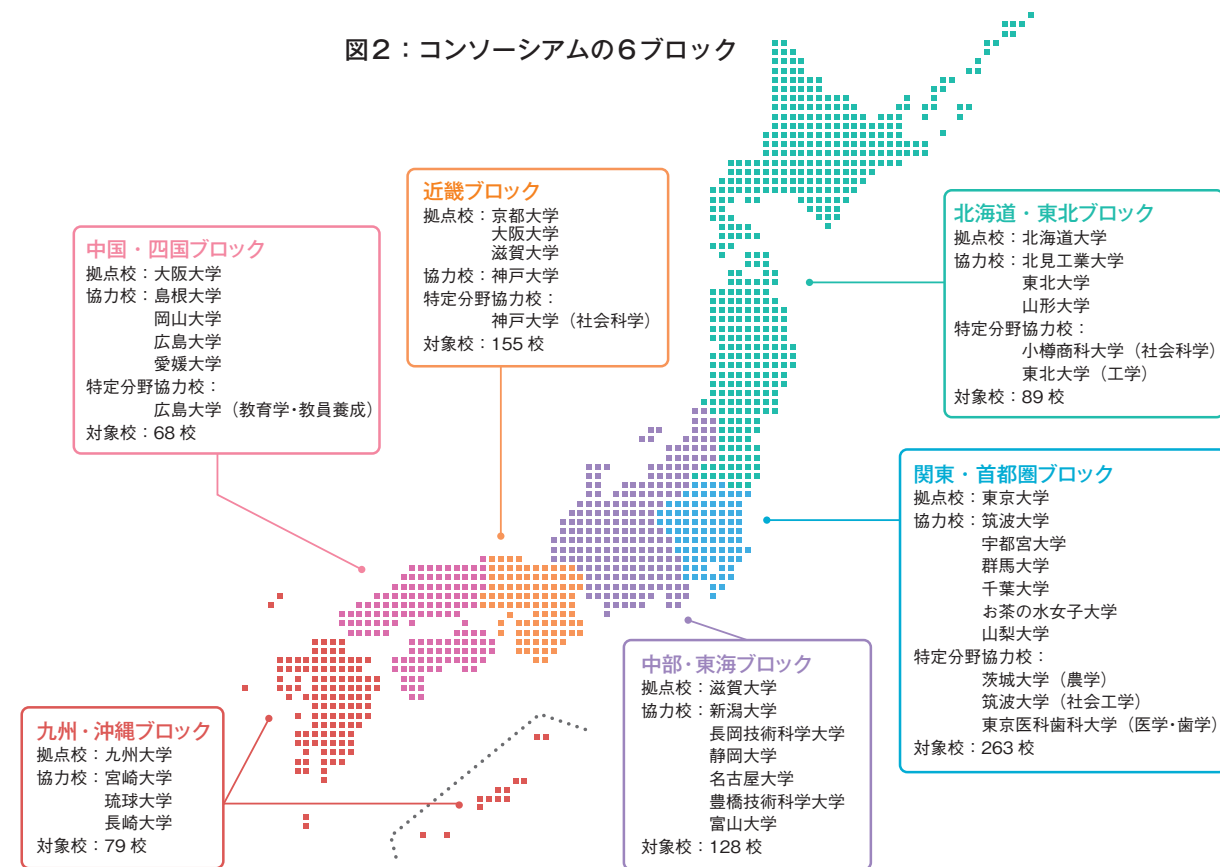
さらにモデルカリキュラムの公開に対応して、産業界の協力を得てビジネス系の実課題と実データ収集・整備を進めている。

●6ブロックにおける活動

コンソーシアムでは2019年度の協力校の参加を機に全国展開を加速するために全国を6ブロックに分けて、それぞれ地域に対応した普及活動を開始している(図2)。各ブロックでは、

- 1) ブロック会議を開催して、数理・データサイエンス教育の実施状況を把握するとともに、他大学への展開計画・状況を共有する
- 2) ブロック別ワークショップを開催して、数理・データサイエンス教育の普及を目指し、数理・データサイエンス教育に関するFD(Faculty Development)活動や具体的

図2: コンソーシアムの6ブロック



な教育の実装方法に関する意見交換を行う

3) FD活動を通じ、数理・データサイエンスを教えられる人材ネットワークを拡大する

等の活動を行っている。

今後に向けた課題と展開の方向性

コンソーシアムでは今後、本年4月に公表したリテラシーレベルのモデルカリキュラムの全国展開及び上位のカリキュラムの策定を重要な課題と考えている。特に前者の実現に向けて、教材(スライド、動画、教科書)の充実、授業で活用できる実課題や実データの収集・公開を行う予定である。また、全国全ての大学、高専での実施においては教員や教育コンテンツ等を直ちに完備することが困難な大学が多いと考えられる。そのためオンライン教材を活用した講義や成績評価の方法等に関して、FDの実施やノウハウの交換が必要である。特に、モデルカリキュラムの理念の共有や個々の教員レベルに落とし込んだ教材等の活用方法、評価方法、TIPS紹介等を盛り込

んだワークショップやFDをカスケード展開するシステムの確立が必要である。そのためには、これらを担えるエバンジェリスト(仮称)の養成を検討している。

また、このような全国展開にあたっては、今後コンソーシアムへの公私立大学の参画が必要であるため、本年度より、コンソーシアム活動に参加していただける公私立大学を募集中である。さらに今年度は、昨年度のリテラシーレベルに続き応用基礎レベルのモデルカリキュラムの検討を進める。そのために新しい検討体制を立ち上げ、本年度中にモデルカリキュラムを策定する予定である。

〈参考文献〉

- [1] P.F. Drucker「ポスト資本主義社会」(1993)ダイヤモンド社
- [2] 日本学術会議提言「ビッグデータ時代に対応する人材の育成」、情報学委員会 E-サイエンス・データ中心科学分科会(2014)
- [3] 情報・システム研究機構「ビッグデータの利活用のための専門人材育成について」、ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会(2015)
- [4] 数理・データサイエンス教育強化拠点コンソーシアム ホームページ <http://www.mi.u-tokyo.ac.jp/consortium/>
- [5] 数理・データサイエンス教育強化拠点コンソーシアム「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～」 http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf
- [6] 東京大学数理・情報教育研究センター http://www.mi.u-tokyo.ac.jp/6university_consortium.html