

# 社会におけるデータサイエンス教育の 体系化と必修化の背景

渡辺 美智子<sup>1)</sup>

## 1. AI人材育成と学習指導要領改訂

新しい学習指導要領は、既に小学校で全国実施、今年4月からは中学校でも実施、2022年度から高等学校でも全国実施となる。とくに、高等学校では、新しく、共通必修履修科目として「情報Ⅰ」が新設され、大学入試センター共通新テストでの採用が予定されている。大学入試センターからは、情報Ⅰの「試作検討問題」が諸関連学会・組織等に送付され意見公募が行われたところである。

今回の改訂の経緯には、これから迎える人工知能型社会に向けて、子供たちが身に着けるべき力が次のように述べられている。

『今の子供たちやこれから誕生する子供たちが、成人して社会で活躍する頃には、(中略)グローバル化の進展や絶え間ない技術革新等により、社会構造や雇用環境は大きく、また急速に変化しており、予測が困難な時代となっている。(中略)こうした変化の一つとして、人工知能(AI)の飛躍的な進化を挙げることができる。人工知能が自ら知識を概念的に理解し、思考し始めているとも言われ、雇用の在り方や学校において獲得する知識の意味にも大きな変化をもたらすのではないかとこの予測も示されている。(中略)このような変革のなかで、生徒一人ひとりが将来の社会の担い手として、個人と社会の成長につながる新たな知識や価値を生み出す力をもつことが期待されている。』

つまり、既存の知識や技能を単に理解させ記憶させるだけではなく、児童生徒自らが新しい知識を獲得する探究力、それを個人や社会の価値に結びつける創造力を身に着けることが意図されている。そのための改訂の大きな柱として、プログラミング教育・統計教育の充実が謳われている。

プログラミング教育の充実では、小学校においてプログラミング教育が必修化され、算数、理科、総合的

な学習の時間など各教科等において、プログラミングを体験しながら、コンピュータに意図した処理を行わせるための論理的思考力を身に付ける学習活動が実施される。また中学校でも、技術・家庭科(技術分野)においてプログラミングに関する内容が拡充され、更に高等学校では、先述のように、共通必修履修科目として新しく「情報Ⅰ」、選択科目として「情報Ⅱ」が設置された。「情報Ⅰ」の中には、『データの活用』、「情報Ⅱ」の中には『データサイエンス』の名称が使用される単元が含まれており、全ての生徒がデータサイエンスと関連したプログラミング教育、ネットワークやデータベース教育を系統的に学ぶことになる。

統計教育の充実に関しては算数・数学科において、小学校1年か「データの活用」領域が新設され、中学校の現行の「資料の活用」領域も「データの活用」に名称が変更された。このことから、小中学校を通して、「データの活用」の系統性が強調され、高校の数学Ⅰ「データの分析」および情報Ⅰの「情報通信ネットワークとデータの活用」、数学Ⅱ「統計的推測」および情報Ⅱ「情報とデータサイエンス」の単元に繋がっていく。

高等学校の情報科の学習指導要領解説では、とくに上記の単元で、中学校数学科、高等学校数学Ⅰ、数学Ⅱの内容との連携をカリキュラムマネジメントを通して相互補完的に行うことも示されている。数学科の改訂では、これまで中学で取り扱われていた代表値の取り扱いが小学校に移行、高校で扱われていた四分位数、箱ひげ図の内容が中学校に移行、「数学Ⅰ」の「データの分析」の単元で、現行では扱われていなかった統計的検定の考え方が入り、「数学Ⅱ」で、統計的検定の理論も含めた統計的推測の単元の重要度が高まり、文系・理系の進路を問わず大学の入試内容に反映される形で配置される。

統計・データ活用、ひいてはデータサイエンス教育に関しては、数学科と情報科の連携が重要であり、数学科同様、共通内容で必修化される「情報Ⅰ」は冒頭で述べたように、センター共通新テストに含まれる方向で、既に準備・検討が進められており、大学入試で、「データの活用」力が数学と情報の二つの科目で観点の

2021年1月31日受付 2021年2月10日受理

1) 慶應義塾大学健康マネジメント研究科

違いはあれ問われることになる。

今回の改訂内容は、生徒の知識獲得に向かうデータを活用した探求のための思考や推論・判断の仕方を、段階を追って高度化するための内容であり、かつ現在の人工知能技術を支える数理・データサイエンス教育の基礎に繋がる。同時に、人工知能がデータから知識(判断・行動選択のルール)を獲得するアルゴリズムを基礎的に理解する素地ともなる。

## 2. データサイエンス教育の体系化

「データの活用」が、小・中・高等学校教育を通じ、数学の教科だけではなく情報科も含めて、教科をまたがって拡充される背景には、世界中を大きく、しかも急激に変革させているデジタル革命がある。世界の企業の時価総額ランキングでは、2021年1月時点で、トップ10には、米国のビッグ5と呼ばれる、アップル、アマゾン、アルファベット(グーグルの親会社)、マイクロソフト、フェイスブック、中国のIT企業のテンセント、アリババ等で占められている。日本企業は、50位内にも入っていない。平成元年(1989年)時点でトップ5を含め50位中に日本企業は32社入っていたことを言えば、隔世の感がある。

ビッグ5は知られているように、膨大なデータをベースに事業展開しており、グーグル経営関係者の著書How Google Works(私たちの働き方とマネジメント) [1]には、21世紀はビッグデータの時代である、ビッグデータはそれを解析するデータサイエンティストを必要としている、データを上手く活用し分析できる人が世界をリードすることが述べられている。また最近の経済誌は再三、20世紀の経済資源は石油であったが、21世紀はデータが重要な経済資源となったことを「データは21世紀の石油」として取り上げており、デジタル革

命が本格化していることは紛れもない事実である。

この状況の中で日本政府の危機感も大きく、今やデジタル革命が世界の潮流であり、経済のみではなく社会活動への影響力がモノから情報・データとそのサービスへ移行したこと、デジタル市場では、データの重要性が高まり、その利活用に向けた制度改革が重要であるとの認識から、「官民データ活用推進基本法」という法律を2016年12月に急ぎ制定・施行している。法律では、社会における官民が保有するデータを流通・活用し、自立的で個性豊かな地域社会の形成、新事業の創出、国際競争力の強化などを目指すために、データを利活用し社会の価値に結びつける人材の育成及び確保として、第十七条「国は、官民データ活用に係る専門的な知識又は技術を有する人材を育成し、及び確保するために必要な措置を講ずるものとする。」および、第十八条「国は、国民が広く官民データ活用に関する関心と理解を深めるよう、官民データ活用に関する教育及び学習の振興、啓発及び知識の普及その他の必要な措置を講ずるものとする。」が設けられている。この法律により、データ活用に関する教育・普及は文部科学省始め政府全体の取り組みとして、地方行政との連携の下に推進されることになる。

また、政府は未来投資戦略2017―「Society 5.0」の実現に向けた改革―(2017年6月)、「未来投資戦略2018―「Society 5.0」「データ駆動型社会」への変革―(2018年6月)」で日本の舵取りの方向性を示しているが、そこでも、膨大なデータの集積(ビッグデータ)がデータサイエンス・AI技術によって個別化した最適なサービス価値を社会に提供することおよびその可能性があらゆる領域に広がっていること、日本は今後、データ分析系・データサイエンス系人材、AI人材の育成に力点を置くことが経済再生と成長戦略の柱となっている。

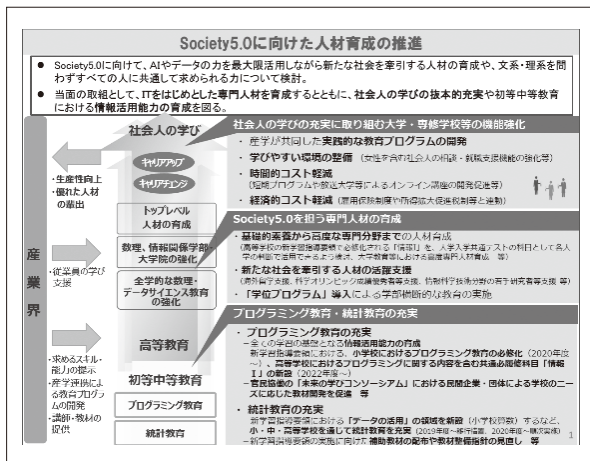


図1 第16回未来投資会議配付資料(2018.05.17)

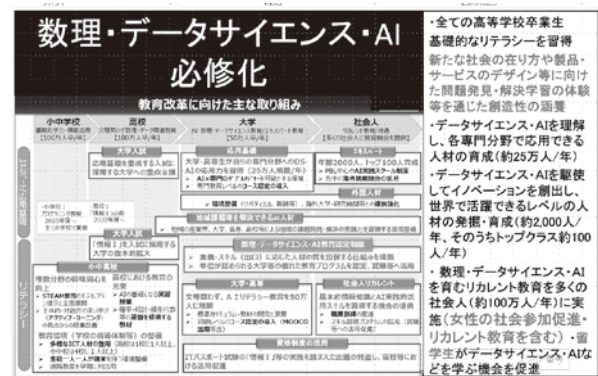


図2 AI 戦略 2019より筆者作成

未来投資戦略が議論された政府の未来投資会議(2018年5月17日)で、「AI時代の人材育成」が取り上げられており、その際、文部科学大臣が提出した「Society5.0に向けた人材育成の推進」資料の一部が、**図1**である。また、**図2**は、2019年6月の「AI戦略2019」の内容である。

今回の学習指導要領での「データの活用」等の統計教育の拡充が、単に初等中等教育の段階に限られたことではなく、初等中等から大学・大学院、社会人のリカレント教育に至る体系的な教育改革の枠組みの中で一斉に推進される方向が示されている。プログラミング教育の目的もデータと統計処理を意識した情報活用能力の育成が目的であることもわかる。

### 3. 第4次産業革命の始動と Society5.0

デジタル革命とともに、現在、言われているのが第4次産業革命である。これまでに、産業の有り様を一変させ人間の働き方そのものを変革させた産業革命は3回ある。それは、18世紀後半の工業化の黎明期を語る第1次産業革命で、このとき蒸気機関による自動化が推進された。次いで、19世紀後半の大量生産と文明化を語る第2次産業革命が、電気による自動化によってもたらされた。そして、20世紀後半に、電子化による製品・生産設備システムが進化し、いわゆるコンピュータによる自動化によって第3次産業革命が起こっている。

現在、これらの産業革命に次ぐ第4次産業革命の時代に入っているというのがグローバルな認識である。第4次産業革命技術は、IOT (Internet of Things,モノとモノがインターネットで繋がり、互いにセンサー等で検知したデータ、もしくはデータから判断した知識を更新し、更に新しい判断ルールを構築、それに基

づくレコメンドや行動選択をモノ(機械)が自動で行う仕組み)、ビッグデータ、ロボット、人工知能等である(**図3**)。

日本政府も、2018年6月、総理大臣によって、第4次産業革命に入っていることを踏まえた人材育成改革の政府方針が下記のように表明された。

「デジタル革命が急速に進展する中で、価値を生み出すデータや人材をめぐる熾烈な争奪戦が世界で繰り返されています。このまま手をこまねいてはならない。Society5.0に向かって、我が国こそが、世界をリードしていかなければなりません。正に、この数年が我が国にとって勝負です。本年を第四次産業革命元年とする。生産性革命の実現に向けて、あらゆる分野で、その社会実装を進めていきます。そのために、自動運転、ヘルスケア、デジタルガバメントなどの重点分野について、産官協議会を設けて官民の叡智を結集し、変革の牽引力となるフラッグシップ・プロジェクトを推進します。加えて、こうした社会変革を実現するための基盤となる、大胆な規制改革に挑戦するとともに、AI人材の育成を始めとした教育システムの改革、大学改革などイノベーションを生み出すエコシステムづくりを進めます。」

ここで言う Society5.0とは、狩猟社会(Society1.0)、農耕社会(Society2.0)、工業社会(Society3.0)、情報社会(Society4.0)に続き、人類が初めて経験する第5番目の新しい社会、「データ駆動型超スマート社会」とされている。この新しい社会は、先述のIoT、ビッグデータ、人工知能、ロボット等の第4次産業革命技術であるデータ駆動型技術が基盤となって、多様なサービスが結合してより最適なサービスの自動化が進行する社会である(**図4**)。

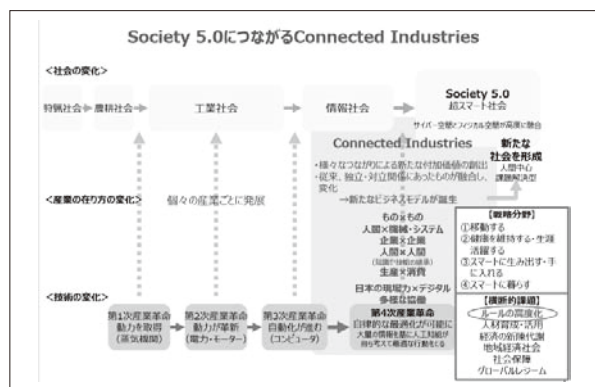


図3 平成29年 経済産業省 産業構造審議会資料

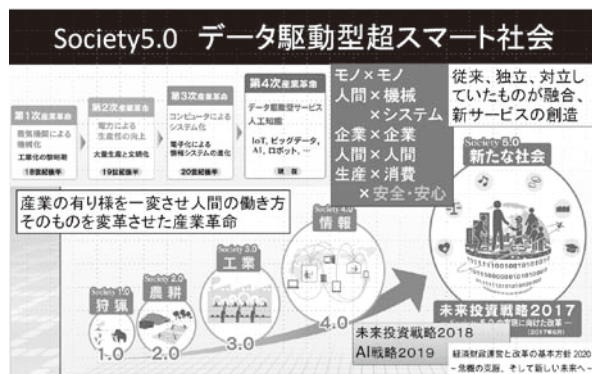


図4 Society5.0の概念図



#### 4. サイバー・フィジカル・システム (CPS) と データ駆動超スマート型社会

既に、Googleの人工知能搭載のコンピュータ棋士(アルファ碁)が名人イ・セドル九段に勝利したことで、人工知能技術に関して一般の認知度も上がっている。また、米国や中国における自動運転車の社会実装化も注目を集めているところである。日本においても2020年までの社会実装を目指して、高齢化が進行する中山間地域における人流・物流の確保のため、「道の駅」等を拠点とした自動運転サービスの実証実験が全国で実施されている。その他にも、自動翻訳、介護ロボット、身体機能の部分麻痺に対して装着者の微弱な「生体電位信号」を皮膚に貼ったセンサーで検出し、装着者の意思に従って身体動作を補助するサイバニクス技術が駆使されたロボットスーツHAL、リストバンドで高齢者の生体ログデータが病院の医療機器に自動送信され、瞬時に疾病リスクが予測されるシステム、普通の医師の診断精度を超える皮膚がんの人工知能による画像診断、レジのない自動決済店舗 Amazon GO、会話を学習し自身で調べて回答や依頼事項に回答する Amazon や Google のスマートスピーカー等、これまでSFの世界とされていたことが、急速に現実化している。

経済産業省は、このような新技術が結集し実現されるデータ駆動型超スマート社会をサイバー・フィジカル・システムCPS(Cyber Physical System)による社会として、図3のように示している。そこでは、現実の断面(フィジカル、実体、現実世界)から小型化・高性能化したセンサー等計測機器でデータを収集し、そのデータが示すサイバー(デジタルな)空間上で成立する行動選択や判断の高度化・最適化、自動制御のための法則(ルール)を推測し、それを現実世界にサービスとして

適用する、さらにデータを収集しルールの精度を上げていく・・・を巡回させ、生活や社会システムサービスの最適化がどんどん進んでいく社会が描かれている。

サイバー・フィジカル・システムにおける循環の3ステップは、以下である(図5)。

- ①データ収集：リアルな(現実の)対象からデータを収集、対象を多次元のデータを要素とするプロファイルで捉える。(例)交通事故が対象であれば、状況や車種、運転車の属性、事故の種類、事故の程度：死亡者を出したかどうかなどが、データ項目となり、1件1件の事故のプロファイルとなる。
- ②データの蓄積・解析：データ化されたプロファイルの蓄積(プロファイルの大量観察)から関連性の解析を通して、社会課題の解決に資する知識(智慧、戦略)を抽出(例)車種や状況の違いから事故の程度を説明するモデルや予測するモデルを構築し、人工知能としてアルゴリズム化。
- ③解析成果を現実世界へフィードバック：②のモデルを自動制御や自動レコメンド機能としてICT機器に搭載

#### 5. データからルールを推測する統計的機械学習と 学習指導要領

人工知能技術が一気に実用レベルに至った背景には、前節で示したように、状況を示す多次元かつ大量のデータから次々と予測や判断のためのルールを推測していく統計学の数理モデリングの技法とそれを演算可能にしているコンピュータの能力拡大がある。データから予測や判断のためのルール(数理モデル)を推測していくことを学習といい、とくにビッグデータを背景に、モデル選択の自由度や誤差に関する評価を含めてアルゴリズム化させ、コンピュータ内で実行させる手法を統計的機械学習(マシン・ラーニング)と言っている。機械学習によって得られたルールが機械を最適な方向へ制御し動作させている。

この技術は、すでに、メールのスパム検知、クレジットカード請求や保険請求等の不正検知、数字や顔画像の認識、Amazonの書籍や商品の自動レコメンドーション、医療診断、信用リスク予測、自然言語処理などで既に多く活用され、ビジネス用途での成功を機に爆発的に普及が進み、今日の人工知能の社会実装化を本格化させている。また、このような限定された範囲と目的での機械学習(統計的な数理モデル)の活用(データアナリティクスの機能)を、IOTによるモノとモノとがデータで繋がる技術を用いることで広範囲に広げ、これまでにない社会全体のスマート化が図るのが、CPS

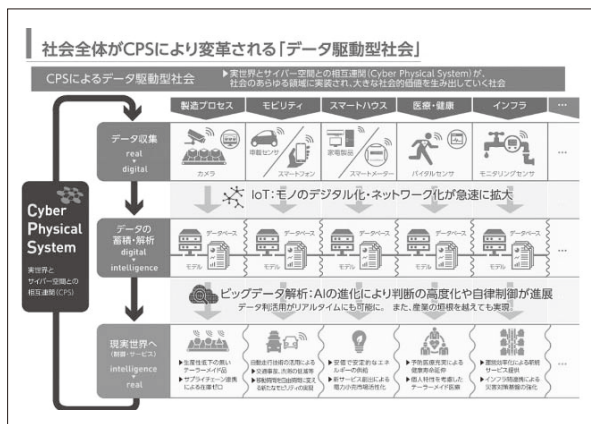


図5 「データ駆動型社会」情報経済小委員会中間取りまとめ

による Society5.0 である。

いずれにしても従来、専門知識を有した個人や組織の判断を伴う仕事の多くが、これからはセンサー等データ取得機能とデータアナリティクス機能が同時に備わった、いわゆる人工知能機能搭載の小型のコンピュータデバイスに代替される社会が到来してくる。現在、その社会に向けた変革が進められている。

今回の学習指導要領改訂は、この変革を担う人材育成の一環として位置付けられている。したがって、プログラミング教育も単なる規定の手順のコードが組めることだけが目的ではなく、最終的には、データ処理・解析処理を組み込んだ AI 型のプログラミング教育が目指されている。そのため、高等学校では、情報科で、比較や関連性の分析、予測や判断・分類のためのデータ活用や機械学習の内容が含まれ、数学科では、その基礎となる散布図やモデリング、統計的推測の内容が含まれている。

また、予測モデルの構築のためのデータ構造は、データを対象のプロファイル(ベクトル)が繰り返し測定されたもの(行列)として取り扱う。ベクトルや行列を抽象的に学習する前に、その表現の身近な例として、データ行列とデータ分析を先に学習する流れは、数理・データサイエンスを全学部で強化する大学の数学教育およびデータサイエンス教育の最新の米国のガイドライン [2] でも示されているところである。

## 6. 大学における数理・データサイエンス教育強化

図2に示したように、Society5.0に向けた人材育成の推進と教育改革は、初等中等から大学等高等教育、社会人のリカレント教育に至るすべてのレベルで同時に着手され強力に推進されている。大学においても文部科学省は、Society5.0の実現に向け、あらゆる産業で

より多くの人材が人工知能、IoT、ビッグデータ、セキュリティといった第4次産業革命技術を使いこなす能力を必要としていることから、その能力のベースとなる、データをとらえ、その本質を見極め、問題を発見・整理分析・解決する力、数学を活用した論理的思考力などを持つ人材の育成教育を文系理系に係わらず全ての大学学部生を対象に計画的かつ着実に全国的規模で推進するため、2016年度に図6にある6大学を数理及びデータサイエンスに係わる拠点校として選定しており、2017年度より、学部の枠を超えた全学的な教育が開始されている。同時に、東京大学を幹事校とする拠点大学コンソーシアムが標準カリキュラムの作成や教材の開発等を進めており、今後全国の大学へ普及・展開が図られることになる。

### おわりに

今回の学習指導要領改訂に際して、中央教育審議会は、子どもたちが判断の根拠や理由を明確に示しながら自分の考えを述べることに課題があること、学ぶことの意義が実感できているかどうか、自分の判断や行動がよりよい社会づくりにつながるという意識が国際的に低く、学ぶことと自分の人生や社会とのつながりを実感しながら、自分の能力を引き出し、学習したことを社会の課題解決に活かしていくことに課題あることを指摘し、その上で、子どもたちがこれから迎える複雑で予測困難な時代を前向きに受け止め、社会や人生をより豊かなものにするができるようになることを目指して、「社会に開かれた教育課程」を理念として答申している。

自らの知識や技能を生活や社会の中で出会う課題の解決に主体的に活用することや、その際に、関連する限られた情報を的確に理解し、自分の考えの形成に生かし、探究や思考のプロセス、判断の根拠を他者に伝え、提案に説明責任を果たすことは大人でも難しいことである。とくにこれまで、目的的に情報やデータ活用を行う統計教育が長く指向されてこなかった日本では、ビッグデータやデジタル革命云々の以前に、社会におけるデータ活用能力の欠如は深刻である。

人口知能型社会を担うデータサイエンス教育自身は、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、基本的な日常的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自身を指し、既にデータに基づく科学的意決定力の強化は、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る。初中等教育

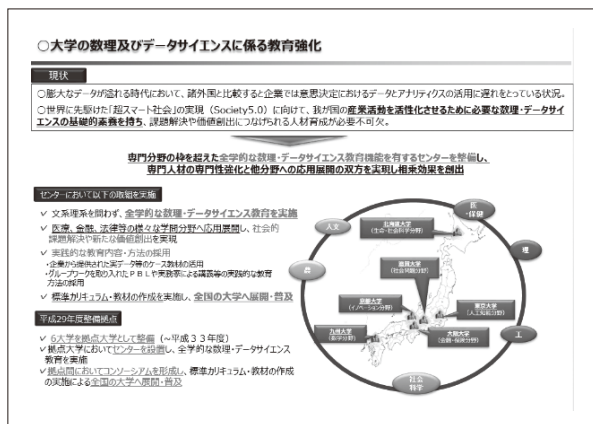


図6 大学の数理・データサイエンス教育強化に向けた取り組み 文部科学省高等教育局 [3]

段階においてその充実を図ることが、来たるべき Society5.0におけるデータサイエンス専門職能の育成にも繋がっていく。

現在の子どもたちは、人工知能ネイティブ世代である。機械的で定型的な思考は、将来、間違いなく人工知能に代替される。人工知能は与えられた目的の中で処理しかできないが、人間は、社会や生活の未来を創造的にデザインし、自ら目的を設定し、目的に応じて必要な情報を見だし、情報を基に深く現象を理解して考えをまとめることができる。人工知能に代替されない人間本来の探究的かつ創造的な思考力を新しい「データの活用」領域で育むことが望まれている。

## 参考文献

- [1] エリック・シュミット他(2017): How Google Works 私たちの働き方とマネジメント, 日経ビジネス人文庫.
- [2] Richard D. De Veaux, et.al.(2017): Curriculum Guidelines for Undergraduate Programs in Data Science, Annu. Rev. Stat. Appl., 4, 2.1-2.16.
- [3] 文部科学省高等教育局(2018): 大学の数理・データサイエンス教育強化に向けた取り組み, JUCE Journal No.1, p2