

大東文化大学におけるデータサイエンス教育構想

永田 清

大東文化大学 経営学部

nagata@ic.daito.ac.jp

キーワード 全学共通カリキュラム, オムニバス講義, 実例中心教育

1 はじめに

近年盛んに論じられるようになったデータサイエンス教育は、理工系の学部生のみならず、経営学、経済学といった社会科学系学生、および法律、政治、文学、語学などを学ぶ学生にも必要なものとされている。本学は、2023年度時点で8学部20学科の学生11,153人の内、一部に理系と分類される1,258人以外は、ほとんどが文系学生の大学である。多くの文系大学生がそうであるように、数学の基礎知識やその思考方法に関して必ずしも十分なレベルにあるとは言えないのが現状である。

このような多くの学生に対し、データサイエンス教育をどのようにすべきかの検討が2021年度後期から行われ、2022年度にその方向性を決め、2024年度開講を目的にカリキュラムを作成することとなった。本稿では、まず一般的なデータサイエンス教育の背景と経緯を見直し、本学ではどのような経緯を経て具体化が行われるかを述べる。

2 データサイエンス教育の背景と経緯

2000年に当時の内閣によって発表されたe-Japan戦略では、IT革命を知識創発型社会への移行であり、産業革命に匹敵する歴史的な大転換を社会にもたらすものとし、新しい国家基盤の必要性を、主に法制度や情報通信インフラなどの国家基盤を中心に述べている。また、目指すべき社会は「すべての国民が情報リテラシーを備え、豊富な知識と情報を交流し、競争原理に基づき、常に多様で効率的な経済構造に向けた改革が推進される社会であり、知識創発型社会の地球規模での発展に向けて積極的な国際貢献を行う。人財の育成についても触れており、実際にIT教育・学習の振興と人材の育成に関しては、2002年に初等中等教育学校で、2003年に高等学校の学習指導要領が改定され、普通教科としての情報教育が始まった。

e-Japan戦略はいくつかの改定を経ており、2004年の新重点5分野には旧重点5分野に対する横断的課題とし

て“研究開発の推進”、“国際的な協調・貢献の推進”、“雇用問題等への対応”、“デジタル・ディバイドの是正”、および“国民の理解を得られる措置”が提起された。その中でも、健全な高度情報通信ネットワーク社会の発展には欠かせない事柄とされている“デジタル・ディバイドの是正”は、雇用との関係もあり大きなテーマとされた。近年ではAIの発展やビッグデータなどのデータ分析能力差などによる雇用への影響もこの分野の問題であると考えられる。

2009年のi-Japan戦略2015における重点分野は、“電子政府・電子自治体分野”、“医療・健康分野”、“教育・人財分野”である。教育・人財分野では、客観的な効果測定の下、子どもの学習意欲や学力と情報活用能力の向上、高度デジタル人財のミスマッチが生じない安定的・継続的な仕組みの確立、大学等における情報教育、デジタル基盤、遠隔教育等の充実などが述べられている。ここで、高度デジタル人財として以下のような条件が提示されている。

- 新しいテクノロジーやイノベーションを創造できる人財
- デジタル技術のみならず、経営や業務改革など幅広い知識と知見を有する人財
- 情報システム・ソフトウェアを構築するためのアーキテクチャやシステム設計力を有する人財
- 難度の高い情報システム・ソフトウェアを使いやすく、高信頼なものとして実現に導くプロジェクトマネジメント能力を有する人財
- 高度なソフトウェアエンジニアリング能力を有する人財
- 高度な知識を持った情報セキュリティ人財
- デジタル技術と業務の両方に精通し、新しい事業・サービスを創造できる人財

ここでも理工系の人財といった視点ではなく、経営やその戦略に関する知見が明記されており、データサイエンス教育の方向性を観ることが出来る。

このような流れの中で「AI戦略2019」では、“我が国が、人口比ベースで、世界で最もAI時代に対応した人材の育成を行い、世界から人材を呼び込む国となり、それを持続的に実現するための仕組みが構築されること”を戦略目標として、教育改革大目標において以下のような目標が設定された。

- 全ての高等学校卒業生が、「理数・データサイエンス・AI」に関する基礎的なリテラシーを習得。また、新たな社会の在り方や製品・サービスのデザイン等に向けた問題発見・解決学習の体験等を通じた創造性の涵養
- データサイエンス・AIを理解し、各専門分野で応用できる人材を育成（約25万人/年）
- データサイエンス・AIを駆使してイノベーションを創出し、世界で活躍できるレベルの人材の発掘・育成（約2,000人/年、そのうちトップクラス約100人/年）
- 数理・データサイエンス・AIを育むリカレント教育を多くの社会人（約100万人/年）に実施（女性の社会参加を促進するリカレント教育を含む）
- 留学生がデータサイエンス・AIなどを学ぶ機会を促進

AI戦略2019において、数理・データサイエンス・AIに関する人材については、「リテラシーレベル」、「応用基礎レベル」、「エキスパートレベル」の3つのレベルに区分するとともに、それぞれの人材育成方策が示されている。

「リテラシーレベル」 全ての大学・高専生（約50万人卒/年）が身につけるべき「デジタル社会の基礎知識（いわゆる「読み・書き・そろばん」的な素養）

「応用基礎レベル」 一定規模の大学・高専生（約25万人卒/年）が自らの専門分野への数理・データサイエンス・AIの応用基礎力を習得し、様々な専門分野へ応用・活用することができる能力

「エキスパートレベル」 実課題にAIを活用してイノベーション創出に取り組む能力を有するものとして、大学院生や若手研究者の育成

大学等におけるリテラシーレベルの数理・データサイエンス・AI教育についてあるべき姿は、以下のようにまとめられている。

- 全ての大学等が数理・データサイエンス・AIの必要性・可能性を十分に理解し、自らの特色を活かした教育に取り組むこと
- 全ての大学等の学生が、専門分野や数理的な習熟度等に合った適切な内容・方法の教育を受けられること
- 全ての大学等の学生が、社会で必須とされる数理・データサイエンス・AIの基礎的知識・スキルを身につけ、将来の選択肢を十分に確保できるようにすること
- 一部の学生については、基礎的知識・スキルの学びを通じて、数理・データサイエンス・AIへの興味や才能を開花させ、より高度な人材に成長していく苗床となること

応用基礎レベルの数理・データサイエンス・AI教育についてあるべき姿は、以下のようにまとめられている。

- 自身の専門分野での研究や卒業後の就業に際して、数理・データサイエンス・AIを活用し、課題を解決できるようになること
- 身につけていく素養
 - － 数理・データサイエンス・AIが社会でどのように活用され、どのように新たな価値を生んでいるのかを理解する
 - － AI・データにはバイアスが生じうる特性があり、使い方によって公正性に関する問題があること、プライバシー保護・セキュリティに関わる課題等があることを理解する
 - － 実際に取得できるデータは、そのままデータ分析をできる状態にはなく、データ分析をできるようにするための工程（データクレンジングなど）に多大な労力が必要となることを体得する
- 全ての大学等が数理・データサイエンス・AI教育の必要性・可能性を十分に理解し、自らの特色を活かして取り組むリテラシーレベルの当該教育の基盤の上に、文理を問わず多くの大学、学部、学科等において自らの専門分野や数理的な習熟度等に合っ

た、より発展した内容の教育が適切な方法で受けられること

- より高度な人材に成長していく基礎となること

数理・データサイエンス・AI教育の基盤として、新学習指導要領では、小・中学校、高等学校を通じて統計教育やプログラミング教育の充実が図られることとなっている。また、大学等においても関連領域の学部等が次々と設置されるとともに、「数理・データサイエンス・AI教育の全国展開」等の事業によってその教育が進展しつつある。しかし、教員リソースの確保困難、ノウハウの蓄積不足等といった問題を依然として抱えており、全ての大学等の学生に対してリテラシーレベルの数理・データサイエンス・AI教育を提供できる状態にまでは至っていない、との認識である。

3 本学におけるデータサイエンス教育検討

データサイエンス教育レベルは、「リテラシーレベル」、「応用基礎レベル」、「エキスパートレベル」の3つのレベルに区分されているが、本学ではまず「リテラシーレベル」教育の検討を行うこととした。MDASH認定も視野に入れておくために、“数理・データサイエンス教育強化拠点コンソーシアムによるモデルカリキュラム ([3, p. 8]) や、他大などにおいて実施されているカリキュラムを検討することとした。

3.1 モデルカリキュラム

「リテラシーレベル」のカリキュラム構成は、“導入”、“基礎”、“心得”、“選択”の4段階に分かれているが、これらの順序は固定されたものでないといったことが指摘されている。

“導入”では、社会におけるデータ・AIの利活用が主題となっており、様々な分野におけるデータ活用事例や関連する技術の概要などが教授されると考えられる。

“基礎”では、データを読み、ある程度の扱いができることが主題となる。統計の基礎は“選択”段階でのオプションに分類されており、データ処理・活用の基本が教授される。

“心得”は、データ・AIの利活用における留意事項であり、データに関連した情報セキュリティとAI活用において留意しておくべき事柄などが教授される。

“選択”では、統計や数理基礎、アルゴリズム基礎、データ構造とプログラム基礎、時系列データ解析、テキスト解析、画像解析、機械学習へのデータ活用などを扱う。

モデルカリキュラムでは、これらの段階をどのように配置し、科目に落とし込んでいくかを3つのケースで示している。ケース1は、上記4段階を独立した科目に落とし込んで、1科目または複数科目の履修によりリテラシーレベルとしている。ケース2では、複数科目に関連を持たせた構成で、全体でリテラシーレベルを修得する。ケース3は、入門、応用・基礎、発展などといった段階を独自に設定し、各段階に科目を配置する。

3.2 本学における授業構想と経緯

2022年2月に、データサイエンス教育の検討部会WGの打ち合わせが、各学部で情報教育を担っている教員有志によって行われた。その際の留意事項として、私学助成を受けるための条件、本学におけるWiFi環境と情報教室の制約などが話し合われた。各学部が開講している情報リテラシに若干の変更加える程度では、私学助成の条件を満たさないであろうこと、およびリテラシ教育対応教員の多くが非常勤講師であることから、内容の統一を徹底させ難いことなどが挙げられた。また、全ての学生が履修可能なものとするには、対面授業のみではコマ数と教室数制約条件を解決できず、コロナ禍で普及したリアルタイムオンライン授業もWiFi接続に不安があることなどが指摘された。結果として、全学共通科目としてオンデマンド教材を使った授業が最も可能性が高いということになり、他の関係教員からも意見を聞くこととした。

2022年3月には学長室のもとに、学長室長、学園情報センター所長と事務員、図書館長、東松山校舎事務長、および学務課員で構成される「データサイエンス基礎教育に関する検討部会」を設置することとし、主な検討事項として、全学共通科目内にデータサイエンス系科目を開設する可能性(短期的課題)、短期的課題に対応する制約への対処、体系的データサイエンス教育プログラム、などを長期的課題として検討することとなった。

2022年度に複数回行われた検討部会の結果、本学独自の教育モジュールを反映して、以下のような基本的な方針が提案された。

[概要] 全学共通の自由科目として、データサイエンスをめぐる現状、およびデータを扱う上での基礎的な内容についての基礎科目を、全学生が履修可能な講義形式科目(1コマ)を設置する。

[授業内容と形式]

- － 主担当教員1名のもと、社会においてデータがどのように活用されているか、またその意義

- などを複数教員によるオムニバス方式で行う。
- 1回60分程度の動画を各教員が作成し、Google Drive上に置いてmanaba(本学LMS)からリンクを張る。ただし、動画のダウンロードは不可とする。
- 各回で課題(レポート等)の提出を求め、学生による感想やフィードバックも検討する。

[実施時期] 2024年4月から開講

[内容検討・更新] 前期(4月から7月)と後期(9月から1月)は同様のものとし、翌年度のものは内容の検討、更新を行う

[想定される課題]

- 受講者数の予想がつかず、多人数(500人を超えるような場合)レポートなどの評価の負担が大きい。
- 実例講義の内容にEXCELやデータ分析の基礎知識を想定することが難しい。

3.3 講義内容

2023年4月時点で、15回分の講義内容案を以下に示す。ただし、今後検討が行われ、内容に変更が加えられることも考えらる。

1. オリエンテーション
2. 概要・手法： 記述統計(データの見方、分析方法)
3. 概要・手法： 推測統計
4. 概要・手法： 情報検索、自然言語処理、画像処理、ニューラルネットワークなどの諸技術
5. ケーススタディ： 健康情報科学
6. ケーススタディ： スポーツ・バイオメカニズムとデータ測定・評価・分析
7. ケーススタディ： 産業関連分析
8. ケーススタディ： 物流・交通データ分析
9. ケーススタディ： 日本語データ処理
10. ケーススタディ： ウェブシステムとデータ収集・加工・管理・運用
11. ケーススタディ： 漢字データの編集・考証・分析

12. ケーススタディ： 意思決定モデルと避難行動のシミュレーション
13. ケーススタディ： ビックデータ解析(出産育児サイトにおけるSNS分析)
14. セキュリティ・倫理： データ・AIの利活用における留意事項
15. まとめ

上記講義案からも分かるように、この講義では数理的な基礎知識の仮定を最小限とし、データサイエンスに関連する事柄を身近に感じられるような構成になっている。全学共通科目の多くは、対象学生を1年生としており、高校までの共通知識である程度の理解が得られような授業を目指している。

4 考察と今後の課題

2000年以降の我が国におけるIT環境の発展やその方向性を概観し、現在盛んに謳われているデータサイエンス教育をその流れの中で把握した。これらの状況を理解したうえで、本学における対応の経緯と現状での案を示した。

実際の開講は2024年度4月からであり、受講学生や担当教員からのフィードバックを得て、より良い授業となることを期待している。

参考文献

- [1] 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、「数理・データサイエンス・AI教育プログラム認定制度(リテラシーレベル)」の創設について、2020年3月、<https://www8.cao.go.jp/cstp/stmain/image/ninteisousetu.pdf>, 2023年6月9日参照
- [2] 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、「数理・データサイエンス・AI教育プログラム認定制度(応用基礎レベル)」の創設について、2021年3月、<https://www8.cao.go.jp/cstp/ai/suuri/ouyoukiso.pdf>, 2023年6月9日参照
- [3] 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～」, 2020年4月, http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf, 2023年6月9日参照
- [4] 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(応用基礎レベル)モデルカリキュラム～AI×データ活用の実践～」, 2021年3月, http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_ouyoukiso.pdf, 2023年6月9日参照