

産学コンソーシアムによるデータサイエンス人材育成の実践

松原 茂樹^{1,a)} 中岩 浩巳¹ 駒水 孝裕¹ 鈴木 優² 井手 一郎¹ 西村 訓弘³ 速水 悟² 武田 一哉¹

概要：データサイエンス人材育成を産学連携で実践する教育プログラムの設計と運用について述べる。本プログラムは、大学院生および社会人を対象とし、データサイエンティストに必要な能力である「実世界データ知識」「ツール活用スキル」「異分野協業マインド」を涵養することを目的とする。企業や地方公共団体が提供するデータを用いてグループワークにより課題を解決する「実世界データ演習」を中核とするプログラムを設計した。複数の大学が産業界と連携して教育するためのガイドラインを整備することで、課題、データ、ツール、メンタなどの教育資源を大学間で共用することを可能としている。本プログラムを開講した2019年度から2021年度までの3年間で165名の修了生を輩出している。

キーワード：データサイエンティスト育成、実世界データ、教育ガイドライン、リカレント教育、大学連携

1. はじめに

超スマート社会では、データを数理的に分析し、活用する能力を備えた人材が強く求められる。このような社会要請に応えるべく、大学などの高等教育機関でデータサイエンス教育の強化が進められている。大学のカリキュラムで展開するデータサイエンス教育にも様々なレベルがある。内閣府の「AI戦略2019」では、数理・データサイエンス・AI教育を、「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段階に区分している[1]。このうち、「リテラシーレベル」「応用基礎レベル」については、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」が作成されたこと[2],[3]、また、文部科学省の「数理・データサイエンス・AI教育プログラム認定制度」が開始されたことなどもあり、全学教育や学部教育のカリキュラムにおけるデータサイエンス教育の導入が進んでいる。

一方、「エキスパートレベル」に対応したデータサイエンス教育を大学が提供する試みは多くない。このレベルは、大学院生や社会人を対象としたデータサイエンティストの育成に対応している。今後、大学におけるデータサイエンス人材育成の取り組みが拡大していくために、関連する事

例を蓄積し整理することで、大学院教育やリカレント教育としての育成プログラムのあり方を定めていくことが望まれる状況にある*1。

本稿では、データサイエンス人材育成のための教育プログラムの設計と運用について述べる。本プログラムは、名古屋大学、岐阜大学、三重大学、広島大学による大学連携の事業として推進している*2。大学院生および社会人を育成の対象とし、データサイエンティストに必要な能力である「実世界データ知識」「ツール活用スキル」「異分野協業マインド」を涵養することを目的としている。

本事業では、企業や地方公共団体から提供されるデータを用いて、グループワークで課題を解決する「実世界データ演習」を中核とした教育プログラムを設計し、運用している。また、実世界データとその取扱いに関する講義を、企業の経営者や技術者が担当するなど、連携大学と企業とでコンソーシアムを形成し、産学連携によるデータサイエンス人材の育成を実践している。

データサイエンス人材育成の取り組みの普及を目指し、複数大学が連携して実世界データ演習を共有するための共通ガイドラインを策定している。ガイドラインに準拠してプログラムを設計することにより、課題、データ、ツール、メンタなどの教育資源を共用することができ、効率化を実現している。本プログラムを2019年度に開始し、2021年度までの3年間で165名の修了生を輩出している。

¹ 名古屋大学数理・データ科学教育研究センター
Nagoya University

² 岐阜大学工学部
Gifu University

³ 三重大学大学院地域イノベーション学研究所
Mie University

a) matsubara.shigeki.z8@f.mail.nagoya-u.ac.jp

*1 AI戦略2019では、エキスパートレベルのデータ人材を年間2000人育成することを2025年の目標として設定している[1]。

*2 文部科学省 未来価値創造人材育成プログラム「超スマート社会の実現に向けたデータサイエンティスト育成事業」の補助のもと本事業を推進している。

表 1 データサイエンス人材育成プログラムの概要

	R2P2ガイドラインとの対応	本事業の履修コース	
		修士レベル履修コース	社会人履修コース
名称		先進データ科学履修プログラム	産業データ科学履修プログラム
育成人材像		デジタル技術を活用した課題解決手順・グループワーキングの技法を習得し、自らの業務や学術の展開に活かせる人材、Society5.0超スマート社会に向けた産業構造の変革を先導できる人材	
履修資格	Prerequisite	大学院修士課程に在籍し、選抜試験に合格した学生	大学卒相当で、データサイエンスの活用の必要性、または、キャリアアップの希望を有する社会人
知識・技術・技能・能力	Prerequisite/Classes/Practicum	統計数理知識、データ解析知識、最先端デジタル技術、グループワーク技能、プレゼンテーション能力、コミュニケーション能力	
教育内容・方法	Prerequisite/Classes/Practicum	講義6科目以上 実世界データ演習	講義2科目以上 実世界データ演習
指導体制	Mentoring	プログラム指導教育 実世界データ演習メンター	
学習時間	Prerequisite/Classes/Practicum	講義90コマ以上 実世界データ演習150時間	講義30コマ以上 実世界データ演習100時間
修了要件	Qualification	必要履修単位の取得、実世界データ演習の評価（DSセンター教員の評価、課題提供企業・団体の評価、グループメンバー相互評価）	

本稿の構成は以下の通りである。続く2章でデータサイエンティストが備える能力について論じる。3章では、本事業で設計した教育プログラムについて述べる。4章では、事業の推進体制と運用実績を示す。5章では、教育プログラムの共通ガイドラインについて述べる。

2. データサイエンティストが備える能力

超スマート社会では、産業を飛躍的に効率化させる構造変革が不可欠であり、その先導を担うデータサイエンティストには、以下の3つの能力を兼ね備えることが求められる。

- (1) **実世界データ知識** 汚れたデータや、取得条件の異なる少量データの活用方法、データへのタグ付け方法、プライバシーの保護などに関する知識
- (2) **ツール活用スキル** 適切なデータ処理方法と処理ツールを多くのツールから「見出す」スキル
- (3) **異分野協業マインド** データを介して、異なる分野の専門家とコミュニケーションし、課題解決に向けて知識や技術を持ち寄り協業するマインド

これら3能力を養うためには、実世界の具体的なデータと最新のデータ処理ツールを活用し、グループワークを通じて多様な知識を統合しつつ、実世界の課題を解決する実践的教育が必要となる。

3. データサイエンス人材育成プログラム

名古屋大学、岐阜大学、三重大学、広島大学の連携事業により、データサイエンティストに必要な3能力（実世界データ知識、ツールの活用スキル、異分野との協働マインド）を兼ね備えた人材を育成するための教育プログラムを設計、運用している。本事業で設計した教育プログラムの概要を表1に示す。以下に示す2種類のプログラムを設けている。

- **先進データ科学履修プログラム** 大学院生を対象とする。名古屋大学、及び、岐阜大学それぞれプログラムを開講する。大学院生は所属する大学のプログラムを履修する。三重大学、広島大学の大学院生は、名古屋大学のプログラムを履修する。
- **産業データ科学履修プログラム** データサイエンスの活用を必要とする、または、キャリアアップの希望を有する社会人を対象とする。東海国立大学機構でプログラムを開講する。

履修には、選抜試験に合格する必要がある。

履修プログラムは、「実世界データ演習」を前提とした教育プログラムの共通ガイドライン **R2P2** (Requirements for Real-world data Practicum Programs) に準拠して定めている (R2P2の詳細は5章で述べる)。

- 実世界データ演習 (practicum)

を中心に、

- 事前科目 (prerequisite)
- 講義科目 (classes)
- メンタリング (mentoring)
- 修了評価 (qualification)

を履修プログラムごとに規定し、実世界データ演習のデータと課題をプログラム間で共用している。

本章の以下では、本事業で設計したプログラムの内容及び履修の流れについて述べる。大学院生向け(名古屋大学)と社会人向けプログラムにおける履修スケジュールを図1に示す。プログラムによって開講時期は異なるものの、修了時期は同一とし、合同で修了式を開催している。開講科目とその内容、R2P2区分との対応を図2に示す。

3.1 事前科目

選抜試験に合格したプログラム履修生が、実世界データ演習に参加できるレベルに到達するための導入科目であ

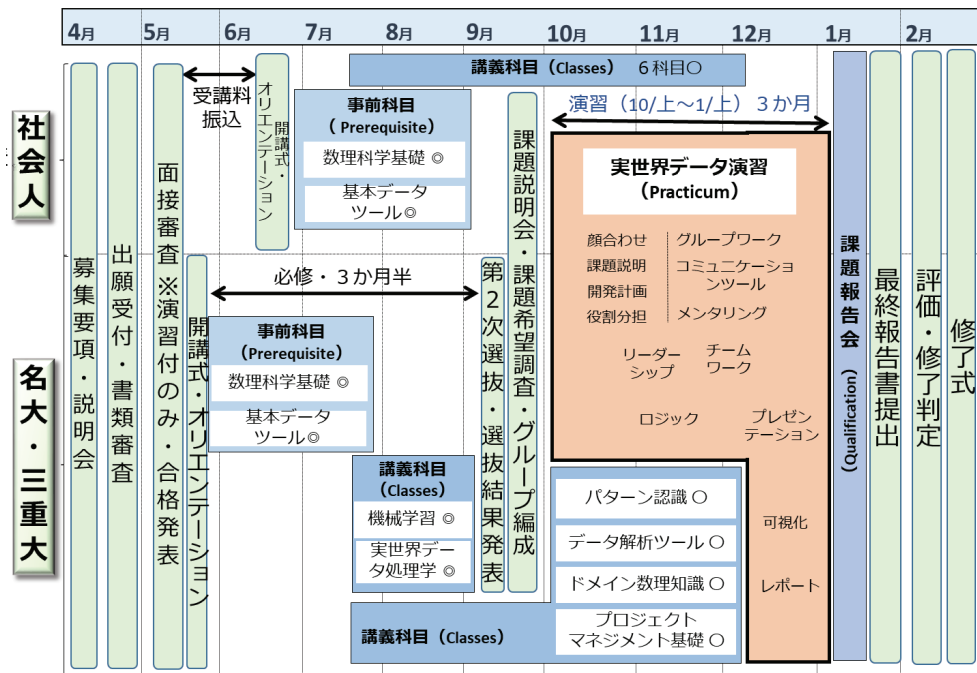


図 1 履修スケジュール（大学院生向け（名古屋大学開講）と社会人向け履修プログラム）

る。「基本データツール」「数理科学基礎」の2科目を設け、全履修プログラムにおいて必修科目としている。

基本データツールでは、初歩的なデータ解析ツールの利用スキルを習得する。自身のPCに演習環境を整え、演習を交えて幅広く学ぶ。数理科学基礎は、統計学やデータ解析学の基礎を講義と演習により学修する。

3.2 講義科目

実世界データ演習の内容を深く理解し、課題を解決するための方法を着想できる能力を養うための科目である。全体で8科目を設けており、そのうち、「機械学習」「実世界データ処理学」の2科目を必修としている。

機械学習では、講義形式で理論を、また、実習形式でディープラーニング等の機械学習の実施法を学ぶ。実世界データ処理学では、大学や国研等の教員や研究員、あるいは、企業の経営者や技術者による講義がオムニバス形式で展開される。実世界データ処理学のテーマを表3に示す。すべて異なる講師が担当している。学術界の講師による9講義、産業界の講師による8講義から構成されており、産学連携のもと実世界データに関わる幅広いテーマを提供している。

3.3 実世界データ演習

実世界データ演習は、自学・自習とグループワークに基づいて進められる。取組み期間は約3か月としている。企業や地方公共団体から提示された課題候補に対し、データ量や必要な作業量を考慮し、実世界データ演習の課題群を選定する。2019～2021年度の3年間の選定課題を表4に

示す。

履修生の所属組織、専門分野、実務経験などの多様性と履修生の希望を考慮し、5名程度からなるグループを編成する。原則、1つの課題に複数のグループを割り当てる。各グループの履修生は、課題とデータを提供する企業や団体との間で機密保持契約を結ぶ。演習の開始後、課題提供企業・団体から、課題の意味と提供されるデータの収集方法について詳細な説明を受け、演習で想定するゴールを課題提供企業との間で共有する。

3.4 メンタリング

本事業では、連携大学の教員、課題提供企業・団体の担当者のほか、QTA (Qualified Teaching Assitant) が履修プログラムのメンタを担当する。実世界データ演習では、以下のメンタリングの機会を設けている。

- **中間報告会** グループごとに取組みの進捗を報告し、課題提供企業・団体の担当者や連携大学の教員などからアドバイスを受ける。
- **課題報告会リハーサル** 下記の課題報告会に向けて実施する。成果の発表方法に関して連携大学の教員からアドバイスを受ける。

QTAには、名古屋大学または岐阜大学の博士課程の大学院生のうち、データサイエンスに関する実務経験を有する学生が認定の上、採用されている。QTAは、

- 事前科目や講義科目の補助教員として、資料の作成や履修生の質問対応を行う。
- 実世界データ演習のグループに割り当てられ、質問対応やアドバイスを行う。

表 2 履修プログラムの開講科目（◎は履修プログラムの必修科目であることを、○は選択科目であることを示す）

R2P2ガイドラインの区分	科目名	科目内容（キーワード）	履修コース		
			修士（名大）	修士（岐大）	社会人
Prerequisite	基本データツール	データツール初歩（10種類）	◎	◎	◎
	数理科学基礎	統計学、データ解析理論	◎	◎	◎
Classes	機械学習	理論と応用、実習、性能評価	◎	◎	○
	実世界データ処理学	データの実践的処理技法	◎	◎	○
	データマイニング	統計分析、仮説検定、知識抽出		○	○
	パターン認識	識別、判別、推定、分類	○	○	○
	データツール	高度データツール（10種類）	○		○
	ドメイン数理知識	データの背景知識の数理	○	○	○
	高度プログラミング	クラウドコンピューティング		○	○
Practicum	プロジェクトマネジメント基礎	グループワーク技法、発表技法	○	○	○
	実世界データ実習	課題解決のグループワーク	◎	◎	◎

表 3 実世界データ処理学のテーマと講師
（大学等の教員・研究員による講義には「学」を、民間の経営者・技術者による講義には「産」を明記）

1	データを扱う上での倫理	学
2	情報処理と法律問題	産
3	データを安全に活用するために	学
4	実世界データ処理におけるデータセット・タグ付け	学
5	デジタルイノベーション	産
6	超スマート社会のデータサイエンティスト	産
7	自ら学ぶ力が伸びる学習法の研究	産
8	材料系・生産系データ	学
9	医療情報の倫理や個人情報保護法の観点	学
10	自治体などの行政データの扱い	学
11	データと事業	産
12	E-commerce 企業における機械学習の活用	産
13	宇宙地球環境研究におけるデータ解析と高性能計算	学
14	データ活用と機械学習	産
15	因果推論	学
16	データ解析のポイントと事例	産
17	学術データマネジメント	学

表 4 実世界データ演習で取り組んだ課題

実世界データ演習の課題名		
2019年度	2020年度	2021年度
養鶏事業の生産効率のためのデータ解析	自ら学ぶ力が伸びる学習法-学習履歴データを用いて-	AI配車による業務平準化
画像解析によるトマトの花質と生育関係性	物流データ分析(配送ルートの最適化等)	ポストコロナ時代における公共交通の課題解決
資質・能力の伸びに効果的な学習の在り方	バイク購入を検討するロイヤルカスタマーの探索	個人顧客の外貨売買における属性相場の特徴分析
LLPオペレーション作業予測・要員配置	公共交通の課題解決に向けた検討プロジェクト	化成肥料の生産効率改善
顧客満足度調査に関する分析	亜鉛メッキ製造工程における良品条件の解析	出店立地の構築及び追加投資店舗の選定

まとめる。修了評価は以下の3つの評価を総合して行う。

- 大学教員による評価
- 課題提供企業・団体による評価
- グループメンバー間の相互評価

修了生には連携大学の共通の修了証を授与する。

4. プログラムの推進と実績

4.1 推進体制

4.1.1 大学連携による教育の推進

本事業は、名古屋大学、岐阜大学、三重大学、広島大学の4大学が連携し、米国・ノースカロライナ州立大学（NCSU）の協力のもと推進している*3。4大学は、本事業の開始前から、データサイエンティスト養成に関わる特色ある教育プログラムをそれぞれ実施してきた。本事業では、4大学が有する教育資源*4を結集し、また、企業や地方公共団体とも連携することで、実践力を備えたデータサイエンティスト人材を育成する体制を整備している。

本事業は、名古屋大学数理・データ科学教育研究センターが主管し、事前・講義科目の教材コンテンツ、実世界データ演習の課題とデータ、講義や演習を補助する QTA など

*3 NUSU Institute of Advanced Analytics (IAA) は「実世界データ演習」のプログラムで有数の実績を有している。

*4 例えば名古屋大学では、博士課程リーディングプログラム「実世界データ循環学リーダ人材養成プログラム」[4]を展開しており、そこで形成された講義コンテンツや QTA 人材などの教育資源を、本事業でも活用している。



図 2 実世界データ演習におけるグループワーク

などのメンタリングを担当する。

3.5 修了評価

実世界データ演習の成果を、課題報告会でのプレゼンテーションと提出されたレポートをもとに評価する。課題報告会では、グループの最終的な成果を、課題提供企業・団体に対して発表する。レポートには、グループでの取り組みの成果に加え、自身が果たした役割を履修生個人が取り

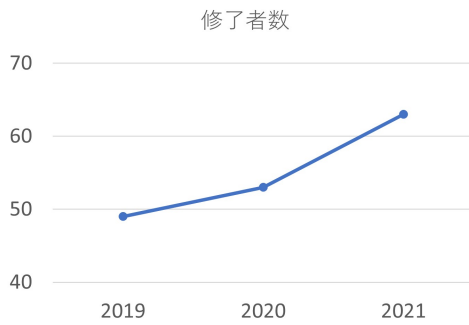


図3 履修プログラムの修了者数の推移 (2019~2021年度)

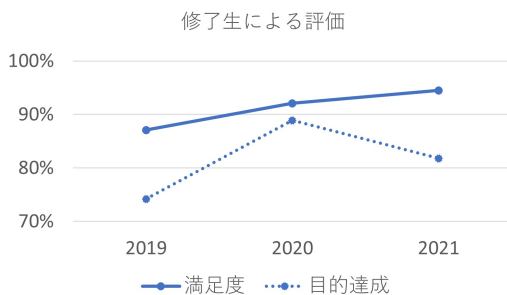


図4 修了者によるプログラムの評価 (2019~2021年度)

の教育資源を一元的に管理し、大学間での共用を可能にしている。事業の推進は、大学を超えたマネジメントの下で行っており、具体的には、4大学の実施担当者間の打ち合わせを毎月開催し、事業の推進に関して検討を重ねている。

4.1.2 産学コンソーシアムの形成

連携する大学が企業や地方公共団体との間で産学コンソーシアムを形成し、「実世界データ演習」を実施している。コンソーシアムに参画する企業や地方公共団体は、以下の役割を担う。

- 課題とデータを提供する。実世界データ演習に提供した課題に取り組むグループに対し、課題やデータの取扱いに関してメンタを担う。
- 実世界データ演習の中間報告会および課題報告会で行われる発表に講評を与える。各グループの成果レポートを査定する。
- データサイエンス人材に対する社会的ニーズに照らし、本事業の育成プログラムが適切に運用されるよう助言する。

4.2 実績と評価

4.2.1 プログラム運用の実績

本事業では、2019年度にプログラムの運用を開始し、2021年度までに3期生が修了している。

履修プログラムの定員は、院生向け(名古屋大学)が20名、院生向け(岐阜大学)が10名、社会人向け(名古屋大学)が20名の計50名である。2019~2021年度の3年間で165名(大学院生92名、社会人73名)が修了してい

る。プログラム修了者数の推移を図3に示す。修了者数は年度ごとに高まっているとともに、いずれの年度も定員と同等もしくは定員を超える修了者数を達成しており、データサイエンス人材育成に対する社会のニーズに一定の対応ができたと考えている。

プログラム運営に関わる実績は以下の通りである。

- プログラムの支援体制として、メンタを担うQTAは、2019年度の開示時は5名体制、2020, 2021年度はそれぞれ9名、8名で実施した。QTAによるメンタリング体制の強化は、品質の高い履修プログラムを提供する上で不可欠である。
- 実世界データ演習の課題数は、各年度5課題を設定している。履修生全体は50名を超えるため、課題あたり2~3グループ編成となっている。
- 2019年度より、全ての事前科目および講義科目をビデオコンテンツ化している。講義はオンデマンドで視聴でき、社会人など履修生は限られた時間内で柔軟に学修できる。

4.2.2 プログラム履修生の評価

プログラムの修了生に対して、課程終了後にアンケートを実施している。履修プログラムへの総合評価として、プログラム全体に対する満足度、及び、プログラムへの参加目的の達成度、をそれぞれ4段階で尋ねたところ、図4のような回答が得られた。概ね良好な評価を得ており、年度ごとのプログラム改善の効果も現れている。

一方で、講義の数量や難易度、メンタリングのあり方に対する問題も顕在化しており、引き続きプログラムを改良していく必要がある。特に、様々な履修生が様々な目的でプログラムに参加していることを鑑み、その多様さに対応したプログラムの設計と運用が課題である。

5. 共通ガイドライン

5.1 ガイドラインに基づく教育プログラムの設計

本事業では、複数大学の連携で実世界データ演習を実施するための共通ガイドライン R2P2 を策定している^{*5}。R2P2に基づく教育プログラムは、**事前科目**(prerequisite)、**講義科目**(classes)、**実世界データ演習**(practicum)、**メンタリング**(mentoring)、**研究活動**(research project)、**修了評価**(qualification)という6つの基本要素(図5参照)を規定することで作成される。このガイドラインを用いることで、実世界データ演習を様々なプログラムで共用することを可能としている。

3章で示した履修プログラムは、この共通ガイドラインに基づき設計している。R2P2に基づく各履修プログラムの構成を図6に示す。「先進データ科学履修プログラム」は、名古屋大学では正規のカリキュラムとは独立している

^{*5} NCSU/IAA (<https://analytics.ncsu.edu/>)における取組事例を参考に策定した。

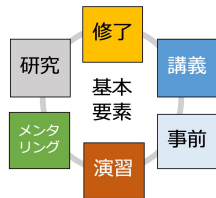


図5 R2P2教育プログラムの基本要素

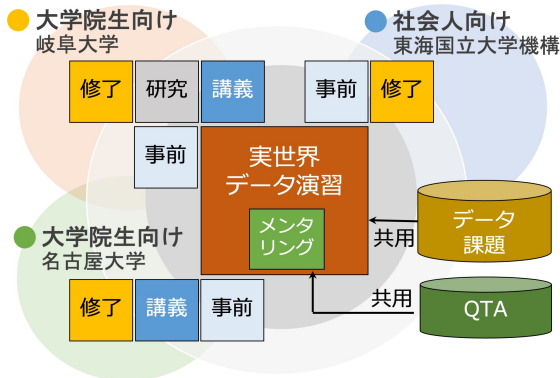


図6 R2P2共通ガイドラインに基づく教育プログラムの構成

のに対し、岐阜大学では大学院のカリキュラムと関連付けられている。両プログラムで開講される講義科目は一部、異なっている*6。「産業データ科学履修プログラム」では、事前科目のみを必須としている。*7

一方、「実世界データ演習」についてはすべての履修プログラムで共通に開講している。データ、課題、ツール、メンタなどをプログラム間で共用することで、実世界データ演習型教育プログラムの効率的な運用を可能としている。

5.2 ガイドラインの拡張と展開

例えば社会人の学び直しのような、新たな教育需要に応える大学機能は、個別の大学ですべてを開発するのではなく、複数の大学が連携し、より広く社会のニーズをとらえ、多様なプログラムとして実装することが持続可能性の点から重要となる。

本事業では、R2P2ガイドラインの拡張により、実世界データ演習型教育プログラムを構成する基本要素を分離し (unbundling)、異なる事業との間で基本要素を共有する (sharing) ことにより、新たな教育プログラムの組成を可能としている。実際、本事業で開発した様々な基本要素は、連携大学が展開するデータサイエンスに関連した12の教育プログラム (本事業のものを含む) で共用されている*8。このような試みの実現性は、本事業の連携大学内に留まるものではなく、R2P2共通ガイドラインに準拠することで、

*6 共通する講義科目は、資料や動画などの講義コンテンツを両大学で共用している。

*7 すなわち、社会人履修生は、講義科目の履修なしでも修了することができる。

*8 三重大学や広島大学では、本事業で設計した教育プログラムの講義科目を大学院教育に活用している [5]。

他の大学の教育プログラムにも展開することができる。

6. おわりに

本稿では、データサイエンス人材育成を産学連携で実践する教育プログラムの設計と運用について述べた。本プログラムの特徴は以下の通りまとめられる。

- 大学院生および社会人を育成の対象とし、データサイエンティストに必要な能力である「実世界データ知識」「ツール活用スキル」「異分野協業マインド」を涵養することを目的としている。
- 企業や地方公共団体から提供されるデータを用いて、グループワークで課題を解決する「実世界データ演習」を中核とする。大学と企業で形成するコンソーシアムのもと推進している。
- 複数の大学が実世界データ演習を共有するための共通ガイドラインを策定している。これに準拠して設計することで、課題、データ、ツール、メンタなどの教育資源をプログラム間で共用できる。

2019年度にプログラムを開講してから3年間で165名の修了生を輩出している。修了生の中にはデータサイエンス関連の企業に就職した学生、また、本プログラムで習得した知識を活かしてベンチャー企業を立ち上げた学生なども含まれている。

今後の課題として、講義コンテンツの拡大、メンタリング体制の強化、履修生の受講ニーズへの対応、共通ガイドラインの普及、が挙げられる。

謝辞 本事業は、文部科学省 Society5.0に対応した高度技術人材育成事業 未来価値創造人材育成プログラム「(a)超スマート社会の実現に向けたデータサイエンティスト育成事業」に採択された取組「実世界データ演習」を用いる価値創造人材教育の大学連携 (事業責任者: 藤巻 朗 (名古屋大学副総長)) において推進している。本事業の推進に多大なご尽力いただき、名古屋大学、岐阜大学、三重大学、広島大学の関係各位に感謝する。

参考文献

- [1] 内閣府統合イノベーション戦略推進会議: AI戦略2019, <https://www8.cao.go.jp/cstp/ai/aistrategy2019.pdf>
- [2] 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). http://www.mi-u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf
- [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 (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). http://www.mi-u-tokyo.ac.jp/consortium/pdf/model_ouyoukiso.pdf
- [4] 名古屋大学博士課程教育リーディングプログラム 実世界データ循環学リーダー人材養成プログラム, <https://www.rwdc.is.nagoya-u.ac.jp/>
- [5] 木島 正明, 土肥 正: 大学における AI・データサイエンス教育—広島大学の取組—, オペレーションズ・リサーチ, Vol. 65, No. 10, pp. 537–543 (2020).