

筑波大学における 全学必修のデータサイエンス教育

和田 耕一, 佐久間 淳, 平田 祥人, 福地 一斗, 青砥 隆仁, 五十嵐 康彦,
今倉 暁, Vasilache Simona Mirela, 海野 広志, 遠藤 結城, 岡 瑞起,
川口 一画, 國廣 昇, 滝沢 穂高, 津川 翔, 三末 和男, 三谷 純

筑波大学では、2019年度より、データサイエンス教育を情報リテラシー教育の一つとして位置付け、人文・生命・理工・情報・医学・体育・芸術など極めて広範な分野にわたる全初年次生を対象として2単位の必修科目「データサイエンス」を開設した。開設にあたっては、筑波大学開学以来実施してきた情報科目も含めて一体的に設計し、情報教育をさらに充実させる形で実装した。本稿では、筑波大学の教育組織、教養教育の運営体制、「データサイエンス」開講に至る経緯、データサイエンス教育の現状、FD活動について述べる。

キーワード：データサイエンス、情報教育、全学必修

1. はじめに

近年の情報技術の発展に伴い、情報を適切に収集・管理し、その分析結果を有効に役立てる技術があらゆる分野において必須な時代となった。学問分野では、大量のデータを収集・分析することにより研究を推進させる第4のパラダイムとしてのデータサイエンスの重要性が広く認識されるようになってきている。また産業分野では、データの効果的な集積と活用による価値の創出が、産業競争力の強化に大きな役割を果たすようになった。さらに急速に発展する人工知能技術は、一部の分野では人間の認識・判断能力を凌駕するレベルに到達しつつある。このような技術的背景に基づいた社会環境の変化は、大学が輩出すべき人材像にも変革を促している。

筑波大学では1973年の開学以来今日に至るまで、基礎的な情報リテラシーとコンピュータの利用技術の修得を目的とした教養教育としての科目「情報」をすべての初年次学生の必修科目として開設してきた。「情報」は、社会の情報環境の進展・変化に対応して、その学修内容を継続的に見直しつつ実施してきたが、データ

サイエンス教育も情報リテラシーの一つとして位置付けることとし、2019年度から全学の初年次生の必修科目として2単位の「データサイエンス」を新たに開講した。

本稿では、筑波大学の教育組織と教養教育の体制、「データサイエンス」開講に至る経緯、データサイエンス教育の現状、ファカルティ・ディベロップメント(FD)活動について述べる。

なお、2020年度より、教育関連の委員会構成など運営体制が改編されており、本稿での委員会名は現在と異なっている場合がある。

2. 教育体制と情報教育

2.1 教育組織

筑波大学には、他大学の学部に対応する組織として学群が置かれており、人文・文化学群、社会・国際学群、人間学群、生命環境学群、理工学群、情報学群、医学群の七つの学群がある。学群には学生が所属する組織として学類が計23置かれている。また、専門的な能力を必要とする体育分野、芸術分野には、専門的な一貫教育を行う専門学群として、体育専門学群、芸術専門学群が置かれている。入学者数は約2,100名である。

2.2 教養教育の運営体制

筑波大学では、教養教育を担う科目の一つとして「共通科目」を開設している。共通科目は、総合科目、国語、外国語、情報、体育、芸術から構成されている[1]。これら共通科目に関して、持続的に教養教育の質を高めていくために、教養教育カリキュラムの企画・実施を中心となって担う全学的組織として教養教育機構が設置されており、その下に共通科目の各科目に対応し

わだ こういち, さくま じゅん, ひらた よしと,
ふくち かずと, あおと たかひと, いがらし やすひこ,
いまくら あきら, ヴァシラケ シモナ ミレラ,
うんの ひろし, えんどう ゆうき, おか みずき,
かわぐち いっかく, くにひろ のぼる, たきざわ ほたか,
つがわ しょう, みすえ かずお, みたに じゅん
筑波大学システム情報系情報工学部
〒305-8573 茨城県つくば市天王台 1-1-1
wada@cs.tsukuba.ac.jp
jun@cs.tsukuba.ac.jp
hirata@cs.tsukuba.ac.jp
fukuchi@cs.tsukuba.ac.jp

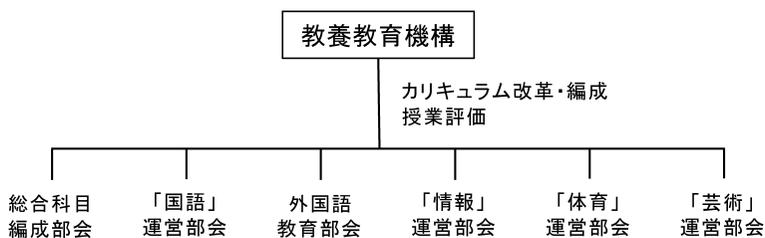


図1 教養教育の運営体制

た部会が置かれている(図1)。図に示されているように、情報教育については、全学の教育組織から選出された委員からなる「情報」運営部会が置かれている。運営部会のミッションは、情報教育の基本方針・計画について審議し、決定することである。さらに、「情報」運営部会の下に情報系教員19名(2020年度現在)から成る「情報」推進室が設置されており、運営部会で決定された基本方針に基づいて、クラス編成やシラバスの策定、授業コンテンツの作成など、授業実施に係るすべてを担っている。

2.3 従来の共通科目「情報」

本節では、まず「データサイエンス」に先立って全学必修で実施してきた共通科目「情報」の実施状況について述べる。「情報」は、「情報」(講義)1単位、「情報」(実習)1単位から成り、計2単位を1年次必修科目として開講していた。約2,100名の1年次生を約50クラスに分け、情報系の常勤教員34名、非常勤講師31名、大学院生によるティーチングアシスタント50名により実施してきた。

「情報(講義)」においては、コンピュータの仕組みやセキュリティなどについて学修させるが、特に近年は情報システムの安全な利用とネットリテラシー、知的財産権と情報倫理、セキュリティとプライバシー保護に関する知識の習得を入学後5週間以内に完了させるようカリキュラムを構成している。

「情報(実習)」では、PCを用いてWindowsの基本動作、Webページ作成、表計算、プレゼンテーションなどについて実習を行ってきた。筑波大学では全学計算機システムと呼ぶ1,100台から成る教育用の計算機システムを運用しており、本実習は、その端末室(以下、サテライト室と記す)で実施してきた。

2.4 「データサイエンス」開講に至る経緯

1節で述べたように、統計的データ解析やデータサイエンスはいまや学問分野を問わず必須のスキルである。また、あらゆる行政・産業分野においてもデータに基づいた客観的な判断や意思決定の必要性は今後ま

すますます高まることが予想され、卒業後にも必要とされるスキルである。このような認識のもと、2017年から共通科目「情報」運営部会において、全学の初年次生を対象としたリテラシー教育としての「データサイエンス」(2単位)開講に向けて議論を開始した。

「データサイエンス」についてもPCを用いた教育が必須と考え、実施形態を演習とすることを前提に、教育内容とともに、クラス分けやサテライト室の割当、必要な人員について試算を重ねた。人員に関しては、毎年の安定した授業運営を考慮すると非常勤講師の大幅な増員に頼ることは難しく、また常勤教員が追加で負担できる教育負荷にも限りがあるため、新たに数名の教員を雇用する必要があることがわかった。新任教員については、2019年度に公募を行い、新規に採用した。

2017年度末の教育組織長が参加する会議で共通科目「情報」を4単位化する方針が承認され、さらに具体的な実装設計を進めた。教育内容についても検討を重ね、データサイエンス関連の学習項目を単に追加するのではなく、従来の「情報」の教育内容も含めて全体を一体的に見直し、カリキュラムを策定した。

2018年6月に「情報」運営部会で実施体制と教育内容が承認され、6月末から7月初旬にかけて全学教育組織長に対して説明会を複数回行い、実施方法や教育内容について説明し、意見交換を行った。それまでの会議で概略を説明していたこともあり、また各教育組織でデータサイエンス教育の重要性に関して認識されていたため、特に反対意見はなかった。その後、2018年7月開催の全学的な教育会議で承認され、2019年度から「データサイエンス」を開講することが決定された。

3. 筑波大学におけるデータサイエンス教育の現状

本節では、「データサイエンス」の設計意図と実施方法、教育内容について述べる。

3.1 「データサイエンス」設計の意図

2.1節で述べたとおり、筑波大学の学生が所属する学

表1 「データサイエンス」シラバス

授業計画	講義概要
第1週～第5週	<p>社会におけるデータサイエンス： データに基づく課題解決、データサイエンスの様々な事例、データに関する法規と倫理</p> <hr/> <p>データの収集： データサイエンスの基本プロセス、データの種類、データの収集、データの前処理、データの再利用性</p> <hr/> <p>データの管理： データ管理の意義と目的、データ収集項目の設計、情報構造と表現の分離、高度なデータ管理とビッグデータ</p>
第6週～第10週	<p>データの可視化： 可視化の意義と目的、視覚的表現の選び方</p> <hr/> <p>データの分析： 離散変数の理解と統計、量的変数の理解と統計、因果と相関、時系列データ、ネットワークデータ、高度なデータ分析と人工知能</p>

群の専門分野は、人文・文化、社会・国際、人間、生命環境、理工、情報、医学、体育、芸術と多岐にわたり、学生がもつ興味や知識、スキルは学群によって大きく異なる。この点を考慮して「データサイエンス」の授業内容は以下の3点を考慮して設計した。

1. データサイエンスを活用するさまざまな専門分野の教員から直接ビデオ講義を通じて語りかけてもらい、学問に対する興味や動機が多様に異なる学生が、データサイエンスを学ぶ必要性を理解し、データサイエンスを学ぶ動機を高められるようにする
2. データサイエンスの中心的スキルであると捉えられている「統計」のみならず、データの収集・管理などデータサイエンスを取り巻く工学的技術、社会におけるデータサイエンスの位置付けや、エビデンスに基づいて意思決定する者全員が理解すべき人権・法制度・プライバシー保護・倫理観の理解など、データを扱うために必要な基礎的教養が習得できるようにする
3. 毎回コンピュータを用いた演習を取り入れ、現実世界のデータや自分たち自身を対象として収集されたデータを扱い、そのデータ処理結果に触れる機会をもつことで、データサイエンスの結果がもたらす効果を実感し、自分ごととして捉えることができるようにする

3.2 「データサイエンス」の実施方法

3.2.1 実施形態

開講科目数は日本語50クラス、英語1クラスである。データサイエンスは、10週2コマ（1コマは75分）の2単位科目として設計され、授業はすべて一人一台のコンピュータが利用可能なサテライト室で実施してい

る。一回あたりの授業時間は150分であり、標準的には講義60分、授業管理システムを用いたクイズ15分、コンピュータを用いた演習75分で構成される。演習課題は一部は授業中に実施し、一部は授業後の課題として課される。各クラス150分のうち75分は、学生による演習の実施を補助するために、ティーチングアシスタントが割り当てられている。

3.2.2 教材構成

さまざまな学群の学生の必修科目であることを念頭に、理学・工学分野だけでなく、文系や体育分野、芸術分野などの非理工系分野まですべての学士課程に対応可能な標準教材（スライド）、演習課題、クイズ（小テスト）を作成した。入門レベルから高度なレベルまで複数の演習課題を用意し、数学的スキルに応じて柔軟に授業内容を構成できるように設計している。データサイエンスの個別分野への応用や、高度なデータ管理、高度なデータ分析などについては、これを専門とする教員によるビデオ講義および演習を作成し、提供している。また筑波大学で実施している英語プログラムに対応するため、標準教材およびビデオ教材については英語の教材も作成した。これらの教材セットはすべての講師に配布され、講師はそれぞれの学群やクラスごとの学生の特性に合わせて内容をカスタマイズしたうえで、授業を行う。

3.3 「データサイエンス」の教育内容

3.1節の設計意図を踏まえ、教育内容は以下のように設定した。シラバスは表1を参照されたい。

3.3.1 学生の学習動機を高める導入ビデオ講義

授業序盤では、筑波大学の人文・社会・国際・生命・理工・人間・情報・医学・体育・芸術などさまざまな分野において、データサイエンスに関わる研究者が提

供する演習付きビデオ講義を視聴する。2019年度は、筑波大学においてデータサイエンスに関連した研究を行う教員に依頼し、以下のビデオ講義を提供いただいた。

- ・生命科学とデータサイエンス (医学医療系 尾崎 遼)
- ・ヒューマンインタラクション (図書館情報メディア系 落合 陽一)
- ・仮説検定入門 (数理物質系 小池 健一)
- ・現代サッカーボールの空力特性 (体育系 洪 性賢)
- ・臨床医学・社会医学とデータサイエンス—ヘルスサービスリサーチの視点から (医学医療系 田宮 菜奈子)
- ・デジタル・ヒューマニティーズ—人文学と情報学の接点が導く新たな知識の世界 (人文社会系 和氣 愛仁)

2020年度にはこれらに加え、サッカー画像解析と戦略立案 (体育系)、海洋における放射能汚染の推移に関するオープンデータ (人文社会系)、AIによる人権侵害と公平性の確保 (システム情報系)、人工衛星データによる森林伐採の安全な飲み水へのアクセスの影響分析 (人文社会系)、シミュレーションによる津波の予測と防災 (システム情報系) などのビデオ講義を提供いただく予定である。動画講義はオープンコースウェア (OCW) として一般公開されており、筑波大学への所属の有無にかかわらず、視聴が可能である [2]。OCW による講義公開以来、総アクセス数は 13,000 件に達しており、大きく注目を集めていることがわかる。

3.3.2 社会におけるデータサイエンスの位置付けの学習

実際にデータ解析を行うためには、検証すべき仮説の設計、データ収集のプランや収集したデータの保管・管理・破棄など、データ解析戦略の立案以外のさまざまな周辺的な作業項目への配慮が必要となる。適切にこれらを設計するためには、遵守すべき法制度の理解や倫理観の習得、人権保護の意識などが必要不可欠である。これらのデータ解析に関する周辺知識の不足は、古くは人権意識に欠ける実験で収集したデータに基づく研究を、近年では個人情報への不適切な利用や差別的な扱いなどを引き起こした。この点を考慮して、データを取り扱うための具体的な技術的方法論に取り組む前に、筑波大学における研究倫理の考え方と研究倫理審査の仕組み、プライバシー保護やデータ匿名化の考え方と個人情報保護法、統計法、AIによる人権侵害やそれに伴う被害などについて、事例を交えながら講義を行う。

3.3.3 データの収集・管理・分析

講義はデータサイエンスの全体像やライフサイクルに触れたうえで、データ工学と統計学の2分野のストーリーに沿って進行する。データ工学分野の話題においては、データそのものについてのマイクロな理解からマクロな理解に展開し、データの収集と管理について議論する。はじめにデータの最もアトミックな単位である属性値がもつ尺度水準とその性質、属性値の集合として定義されるレコード (タプル)、識別子の考え方、テーブルとスキーマの関係、複数のテーブルの扱いと複数のテーブルの結合、関係データベースとスキーマの設計、SQLの基礎を扱う。周辺的な話題として、質問表設計、文書構造と視覚表現の分離、データの再利用性、メタデータ、オープンデータ、ビジネスで利用されるデータなどに触れる。データベースやビッグデータの高度活用、IoT/CPSなど、現代的な大規模データ活用の考え方は、データベースを専門とする筑波大学の教員による60分のビデオ講義 (計算科学研究センター 天笠 俊之) にて扱う。この講義動画もOCWにて公開されている。

統計学分野の話題においては、データ収集の考え方として母集団と標本、全数調査と標本調査、系統誤差と偶然誤差の考え方を説明する。データの前処理として、データの重複、欠損値、外れ値、不整合の考え方とその対処などについて学ぶ。その後データ分析の手法として、質的データの理解と記述統計、量的データの理解と記述統計を学んだ後に、相関関係と因果関係の違い、複雑なデータの分析 (時系列データ、ネットワークデータ) などを扱う。より発展的な話題として、質的データの統計的検定 (カイ 2乗検定)、量的データの統計的検定 (z-検定, t-検定)、線形回帰の教材が用意されており、学生の背景知識やスキルに応じて難易度は調整される。またそれぞれの統計的分析に付随するデータの可視化の考え方や適切な可視化表現の選択について、不適切な可視化表現の実例を交えて学ぶ。機械学習やAIなど、より高度なデータ解析については、AIを専門とする筑波大学の教員による60分のビデオ講義 (システム情報系 馬場 雪乃) にて扱う。この講義動画もOCWにて公開されている。

3.3.4 理解を深める小テストと計算機演習

講義で扱った内容のうち、知識の理解を問う内容は、授業支援システムに準備された小テストを用いてその理解度のチェックを行う。またデータの収集・管理・分析・可視化に関わる内容は、授業後半において、計算機を用いて実際にその処理を体験し、理解を深める

表2 授業評価アンケート質問項目

1. 授業の準備は十分にされていたと思いますか。
2. 教員の説明や授業の進め方は適切でしたか。
3. 授業を通じて、この科目に関連する分野への興味や関心が高まりましたか。
4. 総合的に判断して、この授業を受講してよかったと思いますか。
5. 授業の内容をどの程度理解しましたか。
6. 授業で習ったことは、今後、どの程度役に立つと思いますか。
7. 授業の難易度はどの程度でしたか。
8. 理解すべきポイントを分かりやすく説明していたと思いますか。
9. 適切な教材（スライドなど）を準備していたと思いますか。
10. スライドや板書は見易さに十分配慮していたと思いますか。
11. 声の聞き取りやすさに十分配慮していたと思いますか。
12. 講義開始前、あなたのデータ収集・管理の習熟度は次のどの程度だったと思いますか。
13. また、講義が終わって現在のあなたのデータ収集・管理の習熟度は次のどの程度だと思えますか。
14. 講義開始前、あなたのデータ可視化・分析の習熟度は次のどのレベルだったと思いますか。
15. 講義終了後、現在のあなたのデータ可視化・分析の習熟度は次のどのレベルだと思えますか。
16. 中学・高校の授業におけるデータ解析ソフトの学習内容と今回の講義内容はどの程度重複していましたか。
17. データサイエンスについてより詳しく知りたいと思った項目を上げてください。
18. データサイエンスの授業において、不要だと思った項目を上げてください。
19. 授業全体に関する感想、意見、希望があれば記載してください。

構成としている。

作成した教材は、数理・データサイエンス教育強化拠点コンソーシアムに提供し、数理・データサイエンス教育の普及を図るとともに、他大学に対しても2021年度後半に提供することを予定している。

4. ファカルティ・ディベロップメント (FD)

「データサイエンス」は2019年度から全学必修化された新しい科目であり、また教科書などの教材も整備されていないことから、シラバス作成を含めすべての教材を一から作成する必要があった。このことから、なるべく多くのフィードバックを学生から取得し、これを教材改善に活用する必要があった。筑波大学では、全学的にすべての科目で共通に導入されている授業評価アンケートと、データサイエンス科目新設にあたって導入された教育効果測定との二つの取り組みを中心に教材および授業改善に役立てることとした。

4.1 授業評価アンケート

筑波大学ではFD活動の一環として、全学規模で授業評価アンケートの実施を行っている。すべての開設講義において授業評価アンケートが実施される。授業評価アンケートでは、授業の内容に関しての主観的な評価を回答する。アンケートは共通の質問項目4問と各講義独自のいくつかの質問項目で構成されており、データサイエンスの授業では15項目の独自の質問項目を設けて、全19項目のアンケートを実施した。質問項目の詳細を表2に示す。また、アンケートは全10回

の授業において第9回目の講義の後に、オンライン上のアンケートという形式で行われる。

授業評価アンケートの結果は、集計されたうえで各担当講師にフィードバックされる。また、後に紹介する教育効果測定の個票とアンケートの個票を結合し、さらに詳細な分析を行う予定である。2019年度は、2,175人の受講者に対して1,712人の回答が得られた。したがって、全体の約78.7%の回答が得られたこととなる。アンケートの集計結果から、授業を受ける前と比べて、主観的な習熟度が上がっている様子が確認できた。学生の主観としては、「データサイエンス」の演習を含む実施方式が習熟の助けになったことが見て取れた。一方で、関連分野への関心の高まりが全体に比べて低かった。今後の授業内容の改定で、関心を高めるための工夫が課題となる。

4.2 教育効果測定

2019-2021年度のデータサイエンスの授業では、学生のデータサイエンスに対する意欲や知識を主観的および客観的に調査する教育効果測定を実施している。調査対象は、日本語で「データサイエンス」の授業を受講している学生全員である。効果測定は、授業支援システムを利用して、Web上でのアンケート形式で、授業時間中に時間を確保したうえで行われている。

教育効果測定は、2回に分けて行われる。1回目は、初回（全10回）の講義の冒頭で行われる。1回目の教育効果測定において、学生は、教育効果測定の趣旨を説明する説明書を読んだ後で、教育効果測定への参加

同意書を Web 上のフォームを利用して作成する。その後、教育効果測定に参加に同意した学生は、自身の所属する学群や高校在学時に履修した数学科目など自身の属性に関する設問や、授業への動機に関する設問に答える。また、データサイエンスの専門用語に関して説明できるかどうかを自己申告させる。1 回目の教育効果測定において学生は約 15 分の時間を使って回答する。

2 回目の教育効果測定は第 9 回目の講義の後で行われる。2 回目の教育効果測定は、通常、前述の授業評価アンケート、動機に関する設問、授業で学んだデータサイエンスの専門用語を説明できるか（自己申告による主観評価）、授業で学んだデータサイエンスの項目を理解しているか（ランダムに選択された問題に回答できるかどうかを評価する客観テスト、4 択の問題を 6 問回答）からなる。2 回目の教育効果測定においては学生は 25 分の時間を使って回答する。いずれの調査においても、学生は同意撤回書を提出することで、本調査研究への参加を取りやめることができる。

2019 年度では、受講者 2,175 人に対して、調査研究への協力者が 1,757 人、最終的に与えられた全部の設問に回答した回答者が 1,382 人となった（受講者に対する回答者数の割合約 63.5%）。

教育効果測定の結果は、授業評価アンケートへの回答、専門用語が説明できるかどうかの自己申告や、客観テストが研究参加者の属性や動機によって統計的に説明できるかどうか、などの観点から現在解析を行っており、別の機会に、効果測定の結果を公表する計画である。

5. おわりに

本稿では、筑波大学で 2019 年度より開設した全学必修の授業である「データサイエンス」について、本学の教育組織、教養教育の運営体制、「データサイエン

ス」開講に至る経緯を交えて、データサイエンス教育の現状、FD 活動について述べた。

2 単位の全学必修科目の新規設置については、各教育組織のカリキュラムへの影響も大きく、少なくとも基本的な方針については早期に合意を得る必要があった。そのため、いくつかの会議や説明会を通じてその時点での検討状況を逐一報告して、懸念を払拭し、理解を得るよう努めた。

教育内容の設計の際は、学生の所属する教育組織が広範な分野に渡るため、学生の教育内容に関する興味、学習への意欲を喚起させることに特に注意を払った。

授業の実装については、開学以来の全学必修の情報教育のスキームを踏襲することをまず決めた。実装設計のノウハウが十分に蓄積されていたため、確実な見通しが立てられ、検討すべき事項が早期に絞れたことは実現の容易化につながったと考えている。

今後は授業評価アンケートや教育効果測定の結果分析を進めつつ、教育方法や内容の改善を継続的に行う予定である。

謝辞 「データサイエンス」の教材作成にあたり多数の貴重なアドバイスをいただいた筑波大学教授天笠俊之先生、武蔵野大学助教岡田龍太郎先生に感謝する。本講義「データサイエンス」の一部は、文部科学省「大学における数理・データサイエンス教育の全国展開」事業「数理・データサイエンスの基礎的素養を身につけるための科目新設および全学生への必修化」（筑波大学）の助成を受け、実施している。

参考文献

- [1] 筑波大学教養教育スタンダード, https://www.tsukuba.ac.jp/education/pdf/tstandard/ug/ug_00-3.pdf (2020 年 7 月 10 日閲覧)
- [2] 筑波大学オープンコースウェア, <https://ocw.tsukuba.ac.jp/data-science/> (2020 年 7 月 10 日閲覧)