

論証としてのデータサイエンス

——秋田県立湯沢高校における「デジタル探究」カリキュラム作成と実践——

久富 望 桑川 薫樹

はじめに

本論文は、秋田県立湯沢高校で実践した「デジタル探究」カリキュラムの報告および考察を行い、高校におけるデータサイエンス教育のあり方について提起するものである。

近年、データサイエンス教育の需要の高まりが確認できる。大学ではデータサイエンスを扱う学部の新設、高校では数学分野における確率統計分野の大幅拡充や「情報」必修化、SSH等におけるデータサイエンスの採用の増加等に、その需要は反映されている。

しかし、デジタル時代の「読み・書き・そろばん」である「数理・データサイエンス・AI」の基礎を全ての国民が育む必要があることが指針¹として示される一方で、カリキュラムの具体像を伴った目的論、すなわち「なぜ、どのようにデータサイエンスを学ぶのか」といった議論は不十分である。また、新たな課程を始める際の時間的・環境的制約や、次々と現れる学習コンテンツに過度に左右され、実際の教育現場では新しい手法や技術に接することで精一杯になってしまいがちである。

本論は、統計的な概念やAI等の先端技術ではなく、論証そのものをデータサイエンス教育のカリキュラムの中心に位置付ける可能性について論じるものである。さらに、「データサイエンス教育」と「ICT活用」を峻別することで「なぜ、どのようにデータサイエンスを学ぶのか」の議論に貢献できると主張するものである。行論で示すように、本実践においては「総合的な探究の時間」のような探究学習の中で認識の深まりが起きる過程に、データサイエンスを有意味に学ぶ契機が見出せる。これは、今後の中・高等教育におけるデータサイエンス教育に一定の示唆を与えるものである。

以下では、まず第1節、2節において、本実践の背

景と行われるまでの経緯について論じる。第3節では、カリキュラム及び実践の概要を第1～第3クールに分けて論じる。第4節ではいくつかの成果物を取り上げ、その分析及び考察を行い、教育課程におけるデータサイエンス学習の意義、高大連携やICT活用への示唆を述べる。

1. 「デジタル探究」発足の経緯

1.1 秋田県における「探究」実践の蓄積

秋田県には探究学習を地域や自治体と共同で作ってきた蓄積がある²。湯沢高校でも、秋田県の探究活動等実践モデル校の一つとして、探究活動とカリキュラム・マネジメントの融合を模索し、同校の掲げる「湯高力³」を軸に実践を重ねてきた。また、湯沢市等から提供されたテーマに関して探究活動が行われてきた。例えば、「人口減少に対応し、地域活性化に資するためのまちづくり（若者や若い子育て世代が住みたくするような理想のまちづくりとは）」「豪雪を乗り越えた果樹農家を応援するために（被害を乗り越えた農業生産確保に向けた取組を考える）」等、地域の暮らしや産業に密着した探究活動が蓄積され、生徒が「自分ごととして」これらに取り組む文化が形成されている。このような探究学習の実践の蓄積は、本実践を実施する上での土壌になった。

1.2 秋田県DX推進計画と「デジタル探究」の計画

2022年3月に発表された秋田県DX推進計画における「教育のICT化の推進」の1つとして、デジタル技術等を学びながら探究的学習活動を行うことを目指した「デジタル探究」の実施が掲げられた。これは、秋田県教育委員会高校教育課の「未来へRUNプロジェクト事業」の3つのプロジェクトの1つとして位置付

けられ、半数以上の生徒が大学進学を目指す普通科高校の最初の1校として、湯沢高校が「デジタル探究」の指定校とされた。このため、本実践の着想時からの目標は、普通科高校におけるデジタル人材の育成に定められていた。

2. 「デジタル探究」カリキュラム作成における方針

前節の目標のため、本実践では「デジタル探究」修了時における生徒のイメージ作りからカリキュラム作成を始めた。

2.1 「デジタル探究」の修了時のイメージ

「一人ひとりが、将来よりよく生きるため、一皮むけて成長する機会を、たくさん作ることを修了時のイメージとして掲げた。より具体的なイメージとして、「(秋田市等の) 都会に出る人」にはデジタル技術に触れる近未来の準備になること、「地元に残る人」には自分の将来像にデジタル技術が自然と存在するきっかけとなることを設定した。後者の例として「農業を継ぐが、自分の土地に設置したIoT機器で気象情報等を収集し、親の経験知の確認・発展を主体的に目指す」「地域の輸送業務を担うが、AIを用いた物流・従業員等の最適化を主体的に目指す」等のイメージを設定した。これは、彼らのうち一定数が、将来的に地域の資産を担う人材となりうることを想定したものである。このように、「デジタル人材」として目指す生徒の姿は、なんらかの形で「デジタルツールやデータが自然に(意図を持って)使おうとしている」姿と定義した。これらは「デジタル探究」の初日に生徒たちに一斉講義の形で示した。また、1月に1時間かけ、著者自身の10代、20代における「一皮むけて成長する機会」について紹介した。

2.2 「デジタル探究」で目指した学習経験

湯沢高校が培ってきた「探究」のテーマのうち各自の興味ある問題に対し、統計局 e-Stat・各省庁・秋田県や湯沢市等の自治体によるオープンデータ等から収集し、プログラミングを用いた処理を行い、解釈して発表し、互いに批判的思考を加える事を目指した。

これらは、現行の学習指導要領における算数・数学の「データの分析」の内容と深く関連する。小学6年

生では「身の回りの事象から設定した問題について、目的に応じてデータを収集し、データの特徴や傾向に着目して適切な手法を選択して分析を行い、それらを用いて問題解決したり、解決の過程や結果を批判的に考察したりする力」を、中学3年生では「標本と母集団の関係に着目し、母集団の傾向を推定し判断したり、調査の方法や結果を批判的に考察したりする力」を、「データの分析」において目指すべき資質・能力として求められている。本実践は、探究活動という点を除くと、本質的にはこれらを大きく超えるものではない。なお、プログラミングを用いた処理については、1年次では数学I「データの分析」の範囲内とし、高度な統計的手法の利用は2年次以降に回すこととした。

そのうえで、前述の「修了時のイメージ」を見据えながら、生徒のデジタルリテラシーのバラつきや探究の過程への参加度合い等を踏まえ、指導内容を決定・変更していった。その際、データに基づいた探究活動を目指す本実践は広い意味での科学的態度の滋養につながると考え、批判的思考に類する学習経験を一貫して取り入れた。

3. カリキュラムの概要

以下の活動⁴において、湯沢高校のデジタル探究担当は、1年部主任と担任・副担任1人ずつの合計3人である。授業の進行は筆者が現地やオンラインにて進行し、そのうち半数程度は5クラス同時に実施し、各教室では1年生の担任・副担任等がサポートした⁵。

3.1 第1クール：グラフ化とデータ批判

第1クールでは (a)プログラムを用いてデータからグラフを出力する演習 (b)湯沢市の人口データのデータ形式の変更 (c)自作のグラフをもとに「何かを主張してみる」活動 (d)互いの内容について批判的思考につながるコメントを試みる活動を行った。

(a)の活動では、筆者が用意したPythonのプログラムを生徒がWeb上で動かし、総務省統計局 e-Stat からAPI等を用いて収集された全国の人口データをグラフに表した。これは、誰でも「実行すればすぐ動く」プログラムによって、漠然としたコンピュータへの忌避心を取り払うこと等を狙ったものである。実際に、全ての生徒が複数のグラフを作る作業を体験できた。

(b)では、PDFで公開されている湯沢市のデータ⁶を生徒が手分けしてGoogleスプレッドシートに入力し、プログラムが読み込める⁷データを作成した。これにより、①地元湯沢市のオープンデータに直接触れること②目的に応じて複数のデータ形式がある感覚に触れること、等を狙った。また、(a)、(b)のデータを組み合わせ、日本全国及び湯沢市の過去10年分の男女/年齢別の人口データをグラフ化した⁸。プログラムのうち図1上部のパラメータの部分を生徒は書き換え、様々なグラフを出力することができる。図1上部の状態で行すれば、2015、2021年の年齢別人口構成を全国と湯沢市と重ねた図1下部のようなグラフが出力される。

```

danjo = ["男女計"]##男・女・男女計、が使える。例:["男","女"]
area = ["全国","湯沢"]##全国・湯沢、が使える。
base = "平均値"##縮尺を、そのまま・平均値・中央値・最大値・合計値、から1つだけ選ぶ
years = [2021,2015]##出力年(2012-2021)をここで選ぶ
view_graph(danjo, area, base, years)

```

□ 2021 男女計 全国 平均値 1260.66 千人
 2015 男女計 全国 平均値 1270.32 千人
 2021 男女計 湯沢 平均値 427.29 人
 2015 男女計 湯沢 平均値 482.88 人

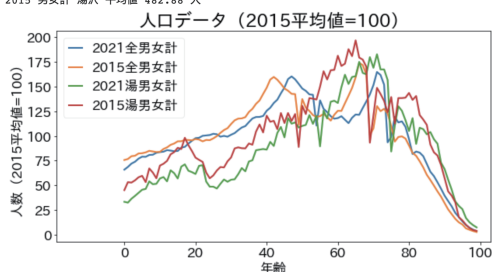
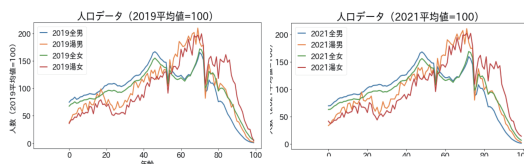


図1：グラフ作成のプログラムの一部

(c)は、作成したグラフから気づいたことを記述してスライドに表現する活動であった。また、気づいたことをよりの確に表現するようグラフのパラメータを各自で変え、新たな気づきを記述したスライドを各生徒が作成した。これは、データから何かを主張することに慣れると共に、主張とグラフの内容を合わせる試行錯誤の経験を意図したものである。「湯沢は60歳から70歳くらいの人口が一番多いが、全国的に見ると40歳くらいと70歳くらいの人口が最も多い」といった比較的具体的なものが見られる一方で、「日本は高齢化社会であることがわかる」といった抽象的な記述も見られた。最終的に、図2のようなスライドを各生徒が3枚以上作成した。

(d)では、(c)で作成したスライドに対して、生徒同士で互いにコメントし合うグループワークを行った。こ

れは、筆者が批判的思考を深めるために複数の学校で実践しているもの⁹の1つであり、グループワークのルールとして「でも...」からコメントを始める、ということを示した。ある主張・ステートメントに対して、別の立場の存在を認識させることを意図して行ったものであるが、「でも...」に続く内容は対立する意見や批判ばかりではなく、話題の付加や感想でもよいと教示し、取り組みやすさに配慮した。それでも、グラフから気づきを書き出す事より、「でも...」を付け加える事に難しさを感じていると分かった¹⁰。



このグラフから全国の男子、女子と湯沢の男子と女子を比べてみると、湯沢は年齢ごとに平均値が異なっていてグラフはギザギザになっているが、全国は比較的なめらかなグラフで、だんだんと上がっていったり下がっていったりと急な増減がない事に気づいたが、2019年の人口データと比べると、20代より少し前の全国女子が上がっているが2021になると20代より少しあとの年の平均人口があがっていることに気づいた

図2：第1クールの生徒の成果物の例

3.2 第2クール：論証の基礎演習「三角ロジック」

第2クールでは、データとその加工・表現を用いて論拠のある主張をすることを到達点とし、題材は前年度の湯沢高校の「探究」のテーマ等から生徒が各自で選んだものとした。具体的には、(a)生徒ごとにグループに分かれてアイデアを出しながらデータを探し、(b)選んだデータについてグラフの出力、(c)見つかったデータを参照した「三角ロジック」の作成を到達点とした。最後に第1クール同様、(d)互いの「三角ロジック」へ批判的思考につながるコメントを試みた。

(a)では、各テーマごとにグループに分かれ、テーマの希望調査時に収集した生徒の意見を参照しながら、オープンデータ等から必要な情報を探し、どんなデータが必要であり、どんな結論があり得るか、等について話し合った。テーマに対する理解、情報収集を目的とした活動であったが、「どのようなデータが必要かわからない」「どのように探せばよいか分からない」というグループが半数程度であった。そこで、各グループが必要な情報やデータについてGoogleフォームで質問を受け付け、回答に対してTAがアイデアやデータの

例を提案する形で支援した。

(b)の活動では、(a)で話しあった結果をもとに、データを収集し、必要に応じてグラフ化する作業を行った。第1クールで行った、プログラムによるグラフ化の技術の活用ができるように意図したが、コードの意味について解説をする時間が不十分だった。このため、グラフ化については、依頼を受けてTAが代わりに行う形で進めた。データを用いて地域のテーマに関して考えることを、プログラミングの体験より優先すべきと判断したためである。

(c)では、ここまでの活動で得られた、漠然とはあるが形成されてきた課題・問いに対する「主張」と、収集したデータ、及びこれを加工したグラフを「三角ロジック」にまとめる演習を行った。自分がここまで行った活動の全体像を可視化し、メタ的に認識させることが主な目的であるが、データを使って言いたいことに応じて表現し、主張を組み立てるという、論証の基本単位の習得を意図して活動を行った。また、グラフ化という行為が、データから意味を取り出し、主張を支持する「論拠 (warrant)」として機能させようとするものであることを強調した。成果物として図3のようなスライドのペアを各生徒が1つ作成した。

最後の(d)では、第1クールのときと異なり個人作業とした。すなわち、グループ内で批判的思考を共有す

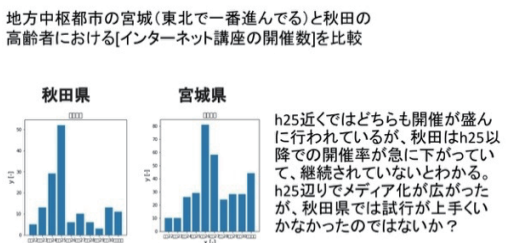
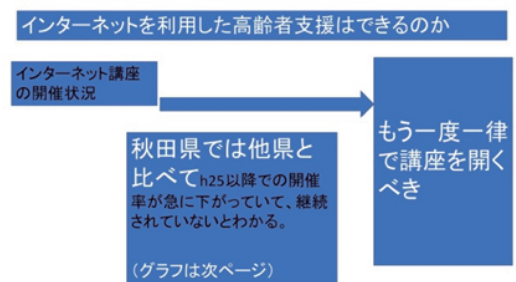


図3：第2クールの生徒の成果物の例

るのではなく、各自が Google フォームに回答した。また、回答者が分からない形で回答内容をフィードバックすることとし、第1クールより一歩踏み込んで、説得力のある発表や論文を育てるための「批判」を積極的に求めた。これは、各個人の主張に対して多くの批判的意見を集め、第3クールに用いるためである。

3.3 第3クール：論証モデルの拡張とデータの活用

最後の第3クールでは、読み手や対立する意見を想定しつつ「データをもとに論拠を持って主張を組み立て、文章として表現する」ことを到達点とした。具体的には、(a)データサイエンスと論証の関係についての講義 (b)「三角ロジック」から「対話型論証モデル」への拡張 (c)論証のモデルを用いたペアワーク、(d)レポートの執筆、(e)生徒間でのフィードバックを行った。

(a)の講義では、「デジタルで探究するということ：データサイエンスは論証である」という題のレジュメを用意し、グラフの表現や代表値の表現が意図を持った warrant として用いられていることへの理解や、「対話型論証モデル」¹¹⁾の紹介等を目的として講義を行った。

(b)では、第2クールの「三角ロジック」から「対話型論証モデル」のワークシートを完成させる作業を各自が行い、図4のような成果物を作成した。これは、「三角ロジック」だけでは足りない「対立意見」「反駁」を、第2クールの最後に集まった批判的意見と、それに対する自らの反論によって埋めたものである。このことで、他者を想定して主張を補強、修正することを身につけることを目標とした。この際、対話型論証モデルの枠組みを広く捉え、「対立意見」の枠に、対立的でない補足的なフィードバックコメントを入れることや、それに対する補助的な論拠を「反駁」の枠に入れることも認めた。これは、適切な批判的意見がない場

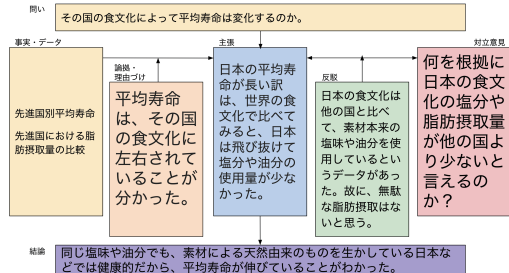


図4：第3クールの生徒のワークシート作成例

合や、優れた批判的意見に対する反駁が見つからない事例を踏まえ、「自分の主張との違いを含む意見」¹²に限らず、まずは任意の「他者」を考慮して主張を作ることを重視したためである。

(c)では、(b)で完成させたワークシートの内容を、他者に伝えられる形で表現する活動を行った。具体的には、一方がワークシートを見せずに口頭で説明し、他方は白紙の対話型論証モデルのワークシートにその説明を埋める作業を行なった。そして、説明をした側と受けた側でワークシートを比較した後、互いの反駁の論拠となるデータや事実を協力して探した。

この活動は、モデルが示す論証の「型」に話型として親しみ慣れることを意図している。外化された「型はめ」に陥らないために、口述言語として一度内化し、自然な言葉で話すよう求めることで、単なるワークシートから思考の「モデル」へと昇華されることを狙った。論証モデルが、論証を組み立てる「型」であると同時に、認識の枠組み（「読む」型）でもある¹³ことを踏まえ、コミュニケーションのコードとして用いたのである。この活動には2時間を充て、1時間目は自分のワークシートに関して、2時間目はペアを変えて1時間目のペアのワークシートに関しても、口頭での伝達を試みた。

(d)では、最終成果物として図5、6のようなレポートを各自が執筆した。(c)における、同時的な「図」の状態の論証から、系列性を持った「論述」へと変換し表現した経験を活かし、到達点として定めた「データをもとに論拠を持って主張を組み立て、文章として表現する」パフォーマンスを行った。クラウド上の文書の同時編集機能を用いて随時添削を行ったが、クラスメイトから「対立意見」をもらってブラッシュアップを図ったり、必要なデータを追加で探しに行ったり、グラフを自ら作成したりと、これまで経験した活動の成果が各所に見られた。また、論拠となるデータに関するグラフ化の支援も引き続き行ったが、筆者の用意したプログラムや、Web上のグラフ作成ツールを用い、自らグラフ化を行う生徒が第2クールの時より増えた。

(e)では2、3人ずつのグループに分かれ、他クラスの3人分のレポートをメモを取りながら各自で読み、グループ内で感想を聞き合い、読んだレポートの「良いところ」「改善すべきところ」を各自でGoogleフォ

ームに回答した。このうち「良いところ」に関する約120人の347件の感想に対し、機械学習¹⁴を用いてクラスタリングを行ったところ、「わかりやすい文章」「はっきりした主張」「根拠となるデータ」「データ・グラフの見やすさ」「ユニークな発想」に該当する感想に分かれた。これらを用いて、全体に対してフィードバックを行い、「いい論証とは何か」という評価クライテリアに関する合意形成、及び共有を狙った。

【1年デジタル探究】

デジタル化を利用した高齢者支援はできるか ーデジタル社会の現状ー

デジタル化を利用した高齢者支援はできるか。この問いに対して、まずは現時点で高齢者にデジタル機器がどれくらい浸透しているかを知るために、情報通信機器の保有状況を調べた。

図1のグラフのオレンジ軸は携帯電話を、黄色軸はスマホを、緑軸はパソコンを示している。なお、ここから「高齢者」を65歳以上と定義する。グラフを見ると、60歳ごろからスマホやパソコンの割合が減少傾向にある。また、80歳以上と高齢者世帯では他と比べて大幅に割合が減少している。高齢者全体としてまだ携帯電話を所持している人が約50%以上と、一定量の利用者がいることがわかった。このグラフを踏まえて、なぜ、こんなにも所持率が減少しているのか、また、高齢者が何か不安に感じていることがあるのではないかと疑問が上がった。そこで、社会のデジタル化に伴ってどのような不安を感じているのかを調べた。

図2のグラフは高齢者世帯（青軸）と非高齢者世帯（赤軸）を対象としたデジタル化に伴う不安を示したグラフである。グラフを見ると、非高齢者世帯は不安を感じていない人が多くデジタルデバイス（情報通信技術の恩恵を受けられる人と受けられない人の格差）の進行は伺えない。しかし、高齢者世帯は「利用する必要がない」、「インターネットについてよく知らない」という項目に大きなピークが見られる。よって高齢者世帯にはデジタルデバイスの進行が読みとれる。また、共通して操作の難しさを感じている人も多いことがわかった。

2つの調査結果を踏まえると、高齢者は「必要性を感じないこと」や、知識不足が原因となって利用者の減少に繋がっていると考えられる。このことから、高齢者支援の一環としてまずは高齢者へのデジタル化を推奨することが必要であると考えられる。

一方で、利用する必要がない高齢者に対してわざわざデジタル化を要求する必要がある



図5：生徒の最終成果物の例（1）

4. 「デジタル探究」の成果と考察

4.1 最終成果物の例：成果と課題

図5の例では、問題提起からデータの説明、データの解釈、そして結論へと論証を複数のデータにまたがって組み立てることができている。「対立意見」として受けた批判や質問を踏まえ、さらに別のデータを用いた論へと発展しており、他者の立場を踏まえつつデータに意味を見出すことが一定程度達成できた例として捉えられる。

図6の例は、健康を維持することと運動についての

【1年デジタル探究】

健康を維持するためには ——健康を維持することと運動は関係あるのか——

人々の健康を維持するにはどうしたらいいか。私は、適度に運動する機会を設けることが必要だと考える。平成27年生命表と28年健康・栄養調査を参考に作られたデータ(図1)を見ると、歩く歩数が多い県でのほとんどで平均寿命が長くなっていることがわかる。また、相関係数のグラフを出すと(図3)のようになった。これは正の相関があるといえる。よって、これらは適度な運動が寿命にもたらしている効果であると裏付けるデータであると考えられる。よって私は、運動は寿命に深く関係のあることであると考える。



歩数が多いのにも関わらず、その県より少ない歩数の県に寿命が下回っている県がいくつかあるという指摘も受けたが、これに関しては、青森県を例に考えてみて、青森県死因割合というデータ(図2)を見ると死因は悪性新生物(がん)の割合が一番多いということがわかり、次に喫煙者数を調べてみた。すると喫煙者数が全国で一番高いことがわかった。喫煙はがんのリスクを高める。このことから、青森県は運動の有無に関わらず沢山の喫煙が寿命を短くしていると考えられる。各県の生活習慣の違いによる多少のデータのばらつきがみられたが、ほとんどの県は歩数が多くなるのに伴って長寿化がみられるので、やはり私は運動をすることと健康を維持することは関係があるのではないかと考える。

図6：生徒の最終成果物の例（2）

レポートであるが、初めは日常生活における歩数と平均寿命の関係について、本文にもあるように「歩く歩数が多い県のほとんどで平均寿命が長くなっている」といった記述のみを行っていた。しかし、「「多い県」とは?」「それで本当に、関係があると言えるのか、はっきりはわからない」といった指摘を受けて、相関係数を持ち出した説明に修正している。ここにおいて、データに基づいた主張をするために特定の統計量を算出するという、データサイエンスの根本的な思考が見出せる。これはデータサイエンスを学ぶ上での自然な動機が生じた典型として位置付けられよう。

4.2 本実践の解釈

今回、「データから主張を作る」ワークに始まり、グラフ化やデータの記述によってデータに解釈を加え、それらを論証として組み立てることに取り組んだ。その中で、論証の手段として統計量の算出が有効となる場面も見られた。

このような、データサイエンス学習に対して切実性を有するような文脈の発生は、地域課題に切実性を持

って向き合うことのできる探究学習の文化的土壌と、「対話型論証」のような認識論的枠組みの共有を前提とし、その文脈の中にデータの処理・探索を位置付けることで達成できたものであろう。

これは、統計学の発展の歴史的経緯とも合致する。大塚淳¹⁵は現代統計学のパラダイムを整理し、帰納推論という普遍的な問題に対するアプローチ群としてデータサイエンスを捉えた。例えば記述統計の多くは、膨大なデータに対して、ある主張につながる、あるいは支持するための「意味」を取り出すアプローチである。また、深層学習等に代表されるモデリングの手法は、高精度での推測ができるモデルならば「真の分布」を良く再現したモデルである、というプラグマティックな正当化¹⁶のアプローチである。いずれにせよ、現在「先端技術」の代表例として頻繁に現れる「データサイエンス」「ビッグデータ活用」「AI」等の概念は、歴史的に分化・発展してきた科学的認識論の枠組みの中で捉えられるものである。

4.3 実践の中で直面した諸課題

本実践を通じ、教育目標に対応できる学習経験や指導の計画を試みることに、単元設計（マイクロな設計）と長期的な指導計画（マクロな設計）を往復しながらカリキュラム全体の改善を図ることは、常に意識された¹⁷。このことは少なくとも2つの点で重要な役割を果たし、公立高校におけるデータサイエンス教育のあり方を考察する材料を提供しようと筆者は考える。

1つ目は、新たな課程を作成する際の時間的・環境的制約が背景にある。少なくとも公立校においては数年ごとの転勤が年度末に決まり、カリキュラムを作成する上で厳しい時間的制約が存在する¹⁸。また、データサイエンス教育であればICT端末の導入に伴う学校環境の変化、情報通信環境のトラブル等の環境的制約が考えられる¹⁹。本実践でも、1年の後半はカリキュラムを作りながら授業実践を行う状況にあった。その中でも、マイクロな設計に偏ることなくカリキュラム作成と実践を進められたのは、最初に生徒のイメージを明確化し、少なくとも教育目標と学習経験・指導を一体として進めた影響が大きい。

2つ目は、教育課程にデータサイエンス学習を組み入れる中で、ICTとの向き合い方についてである。

データサイエンスという、ある種的话题性に富んだ内容をカリキュラムに導入する上で、関連性の深いプログラミング学習やデジタル教材、あるいはメタバースの活用等、真新しい学習コンテンツに過度に着目するあまり、目的論の議論が希薄になるケースが散見される。これは、データサイエンスの習得とデジタル技術の活用を峻別することによって回避すべき問題である。今回であれば、第2クール(b)において、プログラミングによるグラフ化に対し、データから「意味」を取り出す部分を優先してフォーカスし直した点は、ICT活用に固執せず本来の目的論を追った点で示唆的である。すなわち、あくまでデータに基づいた科学的認識論の形成という目標に基づいてカリキュラムを編成することで、単なるICT活用の実践にとどまらない、データサイエンスの見方・考え方を育てるようなカリキュラムの開発を期待できる。本実践は、そのような目標設定の方法論的基盤として、論証モデルや「逆向き設計」の考え方を参照したものである。

4. 4 高大連携の観点から：何を支援すべきか

本実践では大学教員と大学院生である筆者がほぼ毎授業にわたってサポートを行い、カリキュラムの作成にとどまらず、現地やリモートからの授業進行、プログラムの作成、チャット形式による質問への回答、レポートの添削等を行った。このうち、大学教員が最もサポートすべき内容は、論証の過程へのフィードバックであると考えられる。論証の過程に関して「モデル」という形で共通言語を作り、フィードバックや別の視点からのコメントを入れる形での支援は、高大連携の一つの有用な形として捉えられる。

別の視点として、大学の支援を得られにくい高校は全国に多数あり、湯沢高校もその1つであった²⁰。地域の大学生との連携が不可能な制約の元での高大連携は、地域間教育格差の是正という観点から留意されるべきである。本実践においても、オンラインによる授業支援・進行に意義が見出せた一方、現地でのコミュニケーションが不可欠となる場面も多数存在した。

一方、取扱可能な実データの用意には特に留意が必要である。人口データのような地域の実データを用いることは学習の文脈を形作る上で有効である。しかし、大量の数値データがPDF形式で公開されている等、デ

ータサイエンスの教材として取り扱う上でデータ形式が障壁になる場合は多い。その確認や解決にも、企業や大学等の外部機関と技術的な連携は有効に機能する。学校が目標を定め主体的にカリキュラムを計画されていれば、プログラムの作成やデータの前処理の技術的な支援は、円滑に機能することが期待できる²¹。

5. おわりに

本論文は、データサイエンス教育を、論証の方法論の獲得として位置付けたカリキュラムの可能性を論じたものである。実践の内容を踏まえて言えば、そのカリキュラムの意義は、論証を深めようとする活動の中で、自然にデータサイエンスを学ぶ動機が生じる点に見出せる。すなわち、「AI」「ビッグデータ」と言ったキーワードで科学や産業の情勢を理由に語られがちなデータサイエンス教育を、生徒が持つ自然な文脈の中に位置づけるための橋渡しとして、論証の形式論理的なモデルが機能するのである。

今回は、その「自然な文脈」は湯沢高校が培ってきた実践が引き出した。すなわち、これまで行われてきた地域共同型の探究が、生徒たちの生活にとって十分に切実性を持っていたため、そのプロセスを論証モデルという認識の「型」で捉え直しながらデータサイエンスと接続させることで、データサイエンス教育に対しても自然な動機を生むことができた。

このようにデータサイエンスを位置付けると、数理的な統計の知識や、プログラミングの技術の獲得は、場合によっては後に回すことになる。しかし、結果として、自らWebアプリを活用してグラフを作成したり、相関係数等の指標を算出して批判に対して反駁しようとする等、有意味な文脈の中で知識・技能を使用し獲得していく事例が見られた。個別的な知識・技能のみに捉われず、カリキュラムの系統性・整合性を確保する上でもデータサイエンスを学ぶ目的を定めることは有効であろう。このことは、プログラミング等の特定の技術や経験が過度に強調される領域や、高大連携が不可欠な場合において重要な論点である。

各学校が持つ教科あるいは探究学習の土壌に合わせ、データサイエンスを論証の方法論として位置付け組み込んでいくことは、データサイエンスの専門的人材を十分に有しない日本各地の学校が、外部の専門的人材

の力を借りながら、主体的にデータサイエンスのカリキュラムを編成していく上で重要である。

謝辞

本実践は、京都大学国際高等教育院附属データ科学イノベーション教育研究センターから支援を受けた。

註

¹ 文部科学省「数理・データサイエンス・AI教育プログラム認定制度」(2023年3月13日アクセス)。

² 阿部昇「秋田県の子どもの学力が「13年間連続トップクラス」なワケ」Diamond Online, 2021年10月26日(2023年2月22日アクセス)。

³ 湯沢高校生として身につける資質・能力をまとめた「湯高力」と、対応するルーブリックを全教科で作成し、改善を重ねている(秋田県立湯沢高等学校「探究活動等実践モデル校(C型)実践報告書」2021年, <https://yuzawa-hs.net/wp-content/uploads/2021/05/tankyu-model.pdf>, 2023年3月14日アクセス)。本実践でも、筆者は「湯高力」と「デジタル探求」のルーブリックを適宜参照し、指針として用いていた。

⁴ 本実践の1時間毎の授業内容、テンプレート等を以下で公開し、使用したプログラムの更新を続けている。https://github.com/nkutomi/DataScience_for_HighSchool

⁵ このような遠隔授業の形態は、データサイエンス教育に限らず、新しい教育課程の実現にあたり人的資源の確保が難しい地域において有効である。

⁶ 最近5年分より以前のものは、湯沢市市民生活部市民課より提供いただいた。

⁷ PDFや、セルの多くが結合されたような単純でない表形式のデータ等は、プログラムを用いた自動処理に適していない(奥村晴彦「ネ申 Excel」2013年, <https://oku.edu.mie-u.ac.jp/~okumura/SSS2013.pdf>, 2023年3月13日アクセス)。

⁸ GoogleColabを用い、無料で実行可能なクラウド仮想環境の上でプログラムを実行した。

⁹ 久富望、栗栖弘昌、入江克雅「データに対する批判的思考を深めるための授業実践例」『日本デジタル教科書学会発表予稿集』第11巻, 2022年, 19-20。

¹⁰ アンケート調査において、「難しかった」「少し難しかった」の合計が前者は68%、後者は82%であった。

¹¹ 松下佳代『対話型論証モデルによる学びのデザイン』勁草書房, 2021年。

¹² 松下, 同上, p.54。

¹³ 松下, 同上, p.154。

¹⁴ アルゴリズムとしては, Sentence Transformer (Nils Reimers, Iryna Gurevych, Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, arxiv, 2019.)

による埋め込みをもとに混合ガウス分布を用いたクラスタリングを行い, クラスタのタイトルは GPT-3を用いて自動生成した。

¹⁵ 大塚淳『統計学を哲学する』名古屋大学出版会, 2020年。

¹⁶ 大塚, 同上, p.164。

¹⁷ 本実践にあたっては「逆向き設計」(G.ウィギンズ, J.マクタイ(西岡加名恵 訳)『理解をもたらすカリキュラム設計:「逆向き設計」の理論と方法』日本標準, 2012年。)の考えを参照している。

¹⁸ 本実践であれば, 湯沢高校での「デジタル探求」の実施自体が前年末に決まり, 筆者の関わりが始まった4月初旬には「デジタル探求」の目標のみが定まった状態であり, 初回授業を5月末に移した。さらに, 本稿を執筆している3月上旬でも次年度の「デジタル探求」の担当者が確定しないため, 打ち合わせには一定の制約がかかっている。

¹⁹ 本実践でも, 湯沢高校への1人1台の Chromebook の導入が並行して行われたため, 端末への慣れに時間を取られた。また, 1学期はネットワークの切断に毎回のように見舞われ, 授業内容の変更を行った。

²⁰ 湯沢市は約20万人が住む横手盆地の南部に位置し, 約50km 南の山形県新庄市とも繋がりが深い, 横手盆地にも新庄盆地にも大学はない。一番身近な大学は奥羽本線で2時間弱を要する秋田市内の大学である。

²¹ カリキュラムの到達点が明確であれば, 時間的・技術的制約の元で様々な選択が求められる場合も, ねらいの再確認や指導内容の決定は行いやすい。

(教育学研究科・助教)

(修士課程)

受理 2023年 2月27日