

第15講 「データサイエンスの未来と教育への展望」

【学習到達目標】

- ・データサイエンスの未来展望と教育的意義を理解する。
- ・教育現場における役割と必要性を説明できる。
- ・発展に伴う課題とそれに対する教育の対応策を考察できる。

1. データサイエンスとは

データサイエンスは、ビッグデータを収集・分析・解釈し、社会課題の解決や意思決定の支援を行う学問分野です。

データサイエンスは、私たちの社会のさまざまな分野で重要性を増しています。

2. 未来展望

産業界の医療、金融、交通の各分野について見てみましょう。

・医療：個別化医療、早期診断

医療分野では、個別化医療や早期診断を可能にし、特定の遺伝性疾患の早期発見精度を大幅に向上させています。

米国の個別化AI診断システムで、特定の希少遺伝性疾患の早期発見精度が従来の60%から93%に向上した例があります。

・金融：リスク管理、詐欺検出

金融分野では、リスク管理や詐欺検出に利用され、取引データから不正リスクを自動的に評価するシステムが開発されています。

NECが、横浜銀行用にAI不正・リスク検知サービスを開発し提供しました。

この事例では、横浜銀行の取引システムの一次審査を、従来は人手で行っていました。一次審査の初期審査の一部をAIで採点（スコアリング）し、その結果を人間による2次審査に提供し省力化したということです。

・交通：自動運転、交通流の最適化

交通分野では、自動運転や交通流量の最適化に貢献し、AIを活用した無人自動運転タクシーが既に実用化されています。

米国では、Google社の関連会社の「Waymo」（ワイモ）が、世界で最も進んでいる無人の自動運転タクシーでレベル4のロボットタクシーですが、サンフランシスコ、フェニックス、ロサンゼルス の3都市で一般向けに稼働させています（2024年11月末現在）。

今後、ニューヨークでも稼働させるとか、日本のトヨタ自動車もWaymoを導入すると

か報道されています。

3. 教育におけるデータリテラシー

情報化社会に適応するため、システム思考、プログラミングの基礎、統計学の基礎、データ可視化と解釈力が重要なスキルとなります。主体的に意思決定ができる市民を育成するために、概念設計ができ、AI に何をさせるか考えられる能力が必要です。つまり、システム思考が重要になります。

学校教育で必要なスキルとしては、プログラミングの基礎、統計学の基礎、及びデータ可視化と解釈力です。

1) プログラミングの基礎

・「つるかめ算」

例えば、「つるかめ算」のプログラミングでは、まず概念設計シートを作成し、続いてプロンプト（指示文）を生成します。生成したプロンプトは、「貴方は有能なAI教師です。つるかめ算です。足50本で20匹の場合のプログラムを作成してください。」です。

```
Python # 変数を設定します
total_animals = 20 # 全体の生き物の数
total_legs = 50    # 全体の足の数
crane_legs = 2     # 鶴の足の数
turtle_legs = 4    # 亀の足の数
# 解決方法を応用します
# もし全部が鶴だった場合の足の数を計算
assumed_total_legs = total_animals * crane_legs
# 実際の足の数との差を計算
leg_difference = total_legs - assumed_total_legs
# 亀の数を計算（足の差を、鶴と亀の足の数の差で割る）
num_turtles = leg_difference // (turtle_legs - crane_legs)
# 鶴の数を計算
num_cranes = total_animals - num_turtles
# 結果を表示
print(f"鶴の数：{num_cranes}匹")
print(f"亀の数：{num_turtles}匹")
このプログラムを実行すると、以下の結果が表示されます。
<正解> 鶴の数：15 匹、亀の数：5 匹
```

図1 つるかめ算の生成されたプログラムと正解

・「ループリック評価」

1つ目は、学校教育で必要なスキル「プログラミング基礎」です。ループリック評価があります。概念設計とプロンプト処理を2段階で行うような形になります。まず第1段階として、ループリック作成用概念設計シートを用いて、ループリック表を生成します。第2段階として、その結果生成されたループリック表を用いて、実際に児童生徒の作文を評価するという形になります。

具体例を見てみましょう。概念設計シートで「人物：公立小学校で教員。内容：1) 小学5年生の生徒たちを書くことの指導をするために、2) ループリックを作成したい。形式や語調：1) 小学5年生でもわかるような表現、2) ループリックを表形式で。制約：1) 単元名は『文化を受けつぐ』ことについて考えよう、2) 評価内容は『小学校学習指導要領（平成29年告示）』における、該当学年の「書くこと」の指導事項を踏まえて作成。その他：無し」と記載します。これからプロンプトが生成されます。生成AIの「Gemini」が「横浜の公立小学校で教員をされているのですね。5年生の「書くこと」の指導で、「文化を受けつぐ」をテーマにしたループリック作成、素晴らしいです！学習指導要領を踏まえつつ、5年生にも分かりやすい表現で表形式のループリックを作成してみました。」と言って、図2のループリック表件式で作成してくれます。

評価の観点	よくできている (A)	まあまあできている (B)	もう少し頑張ろう (C)
書く内容	伝えたい文化について、具体的な例を挙げて説明している。その文化を受けつぐことの大切さや、受ける。その文化を受けつぐことの大切さや、受けつぐためにできることについて、切さや、受けつぐためにできることについて、自分の考えを理由をつけて分かりやすく書いている。読み手が興味を持つような工夫（例：問いかけ、部分がある。印象的な言葉）が見られる。	伝えたい文化について説明しているが、説明が簡単になっている。文化を受けつぐことの大切さや、受けつぐためにできることについて、自分の考えを書いてや、受けつぐためにできることについて、自分の考えが書けていない部分がある。	伝えたい文化について書こうとしているが、説明が簡単になっている。文化を受けつぐことの大切さや、受けつぐためにできることについて、自分の考えが書けていない部分がある。
構成	始め・中・終わりの構成がはっきりとできている。それぞれの部分で書くことがまとまっている。考えのながれがスムーズで、読みやすい文になっている。	始め・中・終わりの構成はできているが、それぞれの部分で書くことが十分にまとまっていないところがある。考えのながれが分かりにくい部分がある。	始め・中・終わりの構成が意識できていない。書くことの順序がばらばらで、考えのながれが分かりにくい。
表現	読んだ人が情景を思い浮かべることができるよう、言葉を選んで詳しく書いている。自分の気持ちや考えを表す言葉が豊かで、適切な使われている。句読点や段落分けなどが正しく使われていて、読みやすい工夫がされている。	読んだ人が内容を理解できるように、言葉を選んで書いている。自分の気持ちや考えを表す言葉が適切に使われている。句読点や段落分けなどの使い方がおおむね正しい。	使う言葉が簡単で、説明が不足しているところがある。自分の気持ちや考えを表す言葉が少ない。句読点や段落分けなどの使い方が間違っているところがあり、読みにくい。

図2 概念設計シートから生成されたループリック表

次に第2段階の評価用の概念設計シートを見てみましょう。概念設計シートでは「人物：公立小学校で教員/小学校3年生。内容：1) 以下の作品を今回作成したループリックで評価、2) 点数もつけて、3) 児童生徒が今後どのように学習をすればいいのか、自分自身で自己調整ができるようなアドバイス。制約：1) 作成したループリックに従って、2) 評価内容

が、ループリック表上でわかるように、3) 満点は 10 点、4) アドバイスの文章は児童生徒を励ますように」してくださいというような形です。児童生徒の作品は『昔からつづいている京都のちえ』（橋口諒太郎、京都市立御所南小学校 3 年）で、京都広告協会主催作文コンクールの中中学校部門で最優秀賞を取ったものです。この作品では「ぼくは、京都について学習して京都のみ力を知るまでは京都はふつうにある町だと思っていました。それから京都についてあまり考えたことがありませんでした」というような感じで始まります。

この文章を、ループリック評価してみると、総合評価は点数 10 点満点中 8 点です。評価のコメントは「『文化を受けつぐ』というテーマで、京都の和菓子作り、竜王園、京町家、京料理、竹田つけものなど、具体的な事例を挙げて『ちえ』というキーワードで一貫して説明されており、非常に分かりやすいです。小学 5 年生の作品としては、内容の深掘りがよくできています」と。そして「構成、表現と目的」というような感じで評価を記し、コメントをつけています。

橋口諒太郎さんへのアドバイスは「橋口さん、こんにちは！『昔からつづいている京都のちえ』を読ませてもらいました。『京都はふつうの町だと思っていた』という最初の気持ちから、『ちえ』というキーワードを見つけて、たくさんのすごい発見をしたこと、本当に素晴らしいですね！」。ということで下の 5 行「次の学習では、橋口さんが見つけた『ちえ』が、具体的にどんなふうに「昔からつづいて」きたのか、もっと調べてみるのはどうでしょうか？例えば、昔の人はどんな道具を使って、どんな工夫をして『ちえ』を生み出してきたのか、絵や写真と一緒にまとめてみると、もっともっと『ちえ』のすごさが伝わる作品になると思いますよ。橋口さんの学びは、これからもっともっと広がっていくはずですよ。応援しています！」というような形になるわけです。

2) 統計学の基礎

2 つ目は、学校教育で必要なスキル「統計学の基礎」です。文部科学省は「小学校では統計的に分析するための知識や技能を理解し、身近な生活の場面の問題を解決するためにデータを集めて表やグラフに表し、統計量を求めることで、現状や分布の傾向を把握したり、2 つ以上の集団を比較したり意思決定をする。統計的手法を用いて出された結果を、批判的に考察する」と。この「批判的に考察する」というのは小中高一貫して言われてるわけです。

3) データ可視化と解釈力

3 つ目は、学校教育で必要なスキル「データの可視化と解釈力」です。データの可視化として、解釈力のためには、ヒストグラム、円グラフ、二項分布グラフ、箱ひげ図、散布図や、地図上での可視化があります。テキストの可視化として、例えば「走れメロス」の冒頭の部分に関して、文字頻度はグラフ化すると棒グラフになります。文字頻度をワードクラウドにしてみると、よりデータがはっきりします。「メロス」という言葉が非常に大きく、次に大きいのは「村」ということが容易にわかります。重要なキーワードをワードクラウドでは、可視化できるというわけです。

さらに、主体的意思決定と市民教育の観点から、データを根拠に考える行動する力を育成

すること、社会を支える市民性の涵養ということが大切です。そしてシステム思考を実践するということが大事だということがあります。

4. 倫理的課題

倫理的課題には、プライバシーの保護、データバイアスによるリスク及び、AI の透明性と説明責任があります。データバイアスによるリスクは、以前第 6 講にもバイアスとバリエーションという話がありました。バイアスは、データ不足によるリスクです。AI の透明性とは、AI の仕組みや判断過程を公開し信頼できるようにすることです。AI の説明責任とは、なぜそのような結果に至ったのかを説明できることです。

法的・社会的課題には、データ利用に関する法整備、社会制度との調和と、自動化による職業や労働環境の変化というものがあります。日本の AI 戦略として最初の頃に出たものですが、すけれども「データにはノイズや偏ったデータがあって、AI が信頼できる結果をできない可能性があります。AI を安全安心に社会実装するためには、信頼できる品質のデータにより AI 製品・サービスの信頼性を担保する仕組みが必要である」と。例えばどのような問題が起こったかということ、人事採用や就職の採用の場合に、過去の採用者が男性が多く学習データ男性が多いと、男性がいいんだというようなバイアスデータを受け、女性の応募者を不採用にして、男性の方を採用したしまった。アメリカで起こったのは、黒人と白人の応募者があった場合に、白人は採用されるが、黒人を落とされることが多いということで、問題になったわけです。それで、学習データを増やさなきゃいけない。女性のデータや有色人種のデータを増やすし、バランスを取っていくというようなことがありました。別件ですが、少しのノイズで AI が誤認識することがあります。「STOP」の標識文字にテープを貼っただけで、速度制限に読み間違えたという事例があります。

政策として、AI 製品・サービスの信頼を担保する仕組みとして、データ品質の指標とその測定方法の策定及び国際標準化と、第三者評価プロセスそれにガイドラインを整備していく必要があります。

教育現場での対応策として、まず倫理教育の強化があります。「こういうことはやったらいけないよ。こういうことは大丈夫」という線引きをきちんと抑えていくことが大切です。それから法的知識の習得、特に著作権などは法的な問題が結構あります。最後に、柔軟な思考と継続的な学習を育成することが必要です。

5. 未来の教育像

未来の教育としては、データ駆動型の教育プログラム、社会全体のデータリテラシー向上及び、データサイエンス教育が社会全体のデータリテラシー向上と持続可能な発展に貢献することを理解してもらうことが大切です。

データ駆動型の教育プログラムの例として、現在実証実験中の「作文・論文作成支援」の事例があります。これには、人間が作成する「概念設計シート」が一番大事です。

「作文・論文作成 AI エージェント 1」は、作文や論文生成を概念設計シート 1 に基づいて行います。概念設計シート 1 には、具体的に「どんな実践研究か？内容の骨子を教える」などが書かれています。

続く、3 種の「採点・査読 AI エージェント 2」は、概念設計シート 2 に記載の「作成論文を採用して良いか」などに基づき評価します。採点・査読 AI エージェント 2 は、例えば新規性 20 点、妥当性とか信頼性 20 点等という評価結果を人間に知らせます。

最後の「研究倫理処理 AI エージェント 3」は、概念設計で例えば「18 歳未満の児童生徒を対象にしてる？その場合の親の許諾書はある？」と書いておけば、AI エージェント 3 が研究倫理処理してくれます。他にも、教育上不適切な言葉が使われてないか？とかをチェックします。この一例では、概念設計シートという形で、人間が与えたデータによって、AI エージェントが駆動し、データ駆動型の教育プログラムが実現するわけです。

6. まとめ

データサイエンスは、社会教育の両面で重要性を増してきます。このため、教育者は未来を見据えた教育を実践する必要があります。未来の教育は単なる知識の伝達ではなく、客観的なデータを活用した問題解決能力や倫理感を育むことが求められているわけです。

【課題】

- ・データサイエンスの未来において、社会や産業界で期待される役割と、その教育的意義について述べなさい。
- ・今後の教育現場において、データリテラシー教育を推進するために必要な取り組みや内容について具体的に述べなさい。
- ・データサイエンスの発展に伴う倫理的・社会的課題を挙げ、それに対して教育現場でどのような対策や教育内容を取り入れるべきか、あなたの考えを述べなさい。