

信州大学のデータサイエンスリテラシー教育

鈴木治郎[†]

[†] 信州大学 全学教育センター

[†]szkjiro@shinshu-u.ac.jp

キーワード データサイエンスリテラシー, AI 戦略 2019, MDASH 認証

1 はじめに

文部科学省による AI 戦略 2019 推進に伴い、全国の大学では MDASH (数理・データサイエンス・AI 教育プログラム) 認定制度 [1] への認証に向けて、自大学におけるデータサイエンス教育の整備を進めており、すでに 200 校を超える認証校が誕生している。そうした中で信州大学では、オンデマンド授業による 1 単位科目「データサイエンスリテラシー」[2] を設けることにし、2023 年 4 月の 1 年生から履修科目として開講した。この授業の実施にもとづき 2023 年 5 月に MDASH への認証申請を済ませた。

以下ではそこにいたる学内決定過程および教育内容整備について概説する。なお、ここでとりあげた大学などの実施する各コースは、実際の検討期間 (2020 年から 2023 年まで) の各時点にもとづくものであり、2023 年 6 月現在でも継続的に実施していることは意味しないので注意されたい。

2 ロードマップを立てる

信州大学では新入生向けの共通教育を担う学内組織・全学教育機構^{*1}が 2006 年に発足し、学内の共通教育カリキュラムに対して継続的に、検討および実施を進めてきている。

そのため MDASH 認証に向けても全学教育機構が

1. 原案を各学部へ提示・説明して了解を得る
2. 学内全体会議で承認

とする手順で進めていく通常課題の一つとして扱った。

ところで MDASH 認証申請のための科目新設を行うならば

1. 認証申請のための教育活動は前年度の活動にもとづき整備する
2. 必修化した教育活動など申請年のものは、修正申請することで認可を受ける

と進めることを事実上要する。この進め方に合わせる時、学内での教育活動実績は

1. 申請年度に必修化する科目は MDASH 認証に合致するものとして構築
2. 前年度の申請科目は、必修化する科目へと自然にアップデートされるものを用意する

と設計することが自然であると考え進めた。

そのための教育内容整備は、全学教育機構教員の一人が数理リテラシー科目 (2 単位)「数を読む技術」として実施してきたものがあるので、

2020 「データサイエンスのための数をつかむ技術」と改称のみで実施

2021 データサイエンス全般を指向した「データサイエンス概論」に、内容を拡充および取捨選択により実施

2022 MDASH 認証のための科目雛形を授業の一部にもつ 2 単位科目「データサイエンス入門 (以下、入門と略す)」に内容を整理して実施

2023 「データサイエンス入門」から、MDASH 認証を目指す 1 単位科目を取り出し、「データサイエンスリテラシー (以下、リテラシーと略す)」として実施

と年を追って進めた。

^{*1} 学生教育関連を担う組織を再編し、2023 年 4 月に教育・学生支援機構を設けたことに伴い、そこに属する組織の一つ全学教育センターに、全学教育機構は改称した。

3 教材展開モデルに検討した事例

文部科学省が MDASH 認証制度を設ける以前から、MOOC (Massive Open Online Course) など、動画講座をもとに修了テストに合格すれば、修了証を発行してきた講座がある。この講座および各大学の先行事例をもとに、どのような教材展開のモデルを設けることが現実的であるかを検討した。

3.1 gacco の例

JMOOC (Japanese MOOC) として社会的浸透も進む gacco で 8 回授業からなる場合、7 回までは多肢選択テストを実施し、最終回 (第 8 回) は受講生間相互評価を行うことで修了証を発行する。相互評価のシステムには、機能的に Open edX をもとにしたと思える LMS (Learning Management System) が使われている。

3.2 香川大学の例

香川大学ではベネッセ提供の動画とテストからなる教材を利用している。7 回に多肢選択問題のテストが提供され、最終回 (第 8 回) にはレポート (記述問題) を課している。私たちはベネッセに、この講座の資料説明をしてもらった。記述問題の採点に当たっては、そのための非常勤講師の確保など苦労話があるという話も聞いた。

3.3 共通教材で成績評価テストを独自実施する例

岐阜大学と名古屋大学は国立大学法人として現在同一であり、教材を共通利用している。しかし、成績評価テストは大学ごとに別個に設けることにより、大学の独自性を保っている。

リテラシー科目の成績評価問題を、学部・学科等ごとに学部担当者が採点するなど独自運用の可能性を検討する上で参考となった。

3.4 信州大学の選択

多くの大学では、統計学や情報処理基礎科目をいくつか統合して認証されるべき科目のメニューを作り上げている。信州大学がその方向を選ばなかった消極的理由に次がある。

1. 統計学系および情報処理基礎系の科目を共通教育科目として新入生全員が受講可能とできる開講数確保が難しいこと
2. 全学教育機構教員以外も開講する科目群の一つ一つをすべて MDASH 認証を満たすよう申請書作成することは作業負担も大きい

私たちは gacco の評価方法をモデルとしてリテラシー科目の構築を進めることにした。学部・学科等ごとの運用は、第 2 節に述べた各学部説明会で授業担当者を誰が担うかの意見を聴取する中で、学部・学科等が独自に成績評価を行う可能性は消えた。

信州大学で利用してきた LMS (Moodle) にはワークショップという相互評価可能なモジュールを備えている。この方法を選ぶ場合、相互評価の信頼性をどう確保するかは課題として残される。

4 オンデマンド授業の準備

4.1 授業担当者の準備

信州大学では新入生約 2000 名を抱える。オンライン科目では履修を途中であきらめる受講生がもともと多い。とくに必修科目であり、途中であきらめる受講生を減らしたい。そこで受講生の学習状況を管理しコミュニケーションをとれるよう、受講生全体を学部・学科単位を基本とする 100 名前後で約 20 クラスを設けることにした。必修化初年度である 2023 年度にデータサイエンス担当教員は 1 名しかいないため、リテラシー科目担当を担う教員 2 名を 2022 年 1 月から採用する計画を立てた。募集の一部は何度か繰り返した結果、1 月に 1 名、9 月に 1 名を確保できた。

これにより 2022 年度前期に 2 名、後期に 3 名で、2023 年度からの必修化に向けた授業の準備作業を行うことができた。

4.2 授業内容の準備

1 単位科目のリテラシー実施に当たり、最大の検討課題は授業内容の水準をどのように設定するかであった。授業水準の設定とは、合否判定テストの水準をどのように設定するかである。本授業は必履修科目として設置されるため、高校までの事前知識に関わらず誰でも一定の学習を積み合格できる水準に設ける必要がある。一方

で、誰もが合格できる水準を目標としては、多くの受講生の学習意欲を減ずるとともに、履修することを苦役と感じる受講生が増える可能性もある。

これらの問題意識のもとに

1. 各教材には理解度確認テスト（成績評価せず）を設ける
2. 確認テストの水準に相当する知識を教材から読み取れていれば、合格点 60 % を確実にとれる水準に成績評価テストを設ける（全体 100 点中 90 点）
3. 成績評価を伴わない高水準の学習内容を設け、それを学習した受講生が成績評価に加点できるよう、「良い質問」を評価する仕組みを設けた。良い質問は自由記述であり、他受講生も閲覧（質問内容および教員による評価）できるよう、LMS のフォーラムモジュールを用いた。この良い質問へも加点した（全体 100 点中 20 点を上限）。

を基本設計とした。上記の素点合計は最大 110 点になるが、100 点を超える受講生でも上限を 100 点とすることを授業シラバスに明記した。

2021 年度実施の「データサイエンス概論」を、2022 年採用のデータサイエンス担当教員を交えて検討し、データサイエンス入門（2022 年度）授業では、その前半 8 回がリテラシー科目が試行実施となるよう改善作業を続けた。

参考までに入門科目における相互評価問題は「DX 化ということばは DX に「化」を付したものである。この用法は不適切であることわかる多肢選択問題を作成せよ」である。相互評価にあたっては、受講生の作問を評価する視点をいくつか提供し、Yes/No で回答できるようにした。

4.3 高水準の内容

ここでいう高水準とは、ビジネス系なら日本経済新聞、科学系なら日経サイエンスなどに書かれる、ある程度専門性をもつ解説記事を拾い読みできる水準を想定した。簡単な導入記事を教材として提供した上で、さらに

1. オンライン上の事例や用語解説記事または動画への参照
2. R または Python による簡単なサンプルプログラムの提供

を提供した。サンプルプログラムの提供を R または Python としたのは、たとえば多くの受講生にとって身近な表計算ソフト（Excel など）は以下の欠点を抱えるからである [3][4]。

1. 処理の再現性がよくない
2. 処理の際にコピー汚染（コピー&ペースト関連操作に伴い、元データ破壊や関係ないデータの不慮の挿入など）が起こりやすい
3. セルの値表示において内在する数式と計算値の二重性はスキル不足の受講生に混乱を招く

なお第 3 点は [3] で書いたが、[4] では強調していない。

5 今後の課題

私たちが構想した

1. 教材の提示
2. 確認テストの受験と学習
3. 成績確認テストによる評価
4. 受講生間相互評価による記述問題

にもとづく学習サイクルにおける問題には、たとえば受講生にとって学習しやすさはどうか、受講生にとって許容可能な事前知識で書かれているか、相互評価問題は成績評価試験として妥当であるか、などがある。また、2025 年度入試では新課程の高校情報 I を履修した学生たちが入学してくる。それに対して適切な水準で教材が提供されているかも今後を検討を行う必要がある。

参考文献

- [1] 文部科学省、『MDASH 認定制度（リテラシーレベル）』、https://www.mext.go.jp/a_menu/koutou/suuri_datascience_ai/00002.htm
- [2] 信州大学共通教育、『データサイエンスリテラシー』<https://campus-3.shinshu-u.ac.jp/syllabusj/Display?NENDO=2023&BUKYOKU=G&CODE=G0A20003>
- [3] 鈴木治郎、『Excel で実験する統計学』、ピアソン・エデュケーション、1998 年（第 1 版）、2000 年（第 2 版）
- [4] 鈴木治郎、『表計算でデータサイエンス教育はできるのか』、第 6 回国際 ICT 利用研究会基調講演、2021 年、<https://iiar.org/iars/event/012.php>