

# 生成 AI を用いた DX リテラシー標準準拠スキルサーベイの構築と検証

坂本昌宏<sup>1</sup> 松原行宏<sup>1</sup> 毛利考佑<sup>1</sup>

**概要:** 少子高齢化社会を迎えつつある現在、地方自治体においても「消滅可能性」という言葉が取り上げられるようになり、デジタル技術を活用した行政の効率化は待ったなしの状況である。こういった状況を踏まえて、広島市職員の DX リテラシー向上を目指し、研修コンテンツを開発してきた。研修の効果として広島市役所における DX 推進件数の増加は見られたものの、教材の網羅性及び難易度が今後の課題として浮上した。そこで、研修コンテンツの作成に先立ち、経済産業省・IPA (= Information-technology Promotion Agency, Japan:独立行政法人情報処理推進機構) が発表した「デジタルスキル標準」に準拠した、職員のスキルレベルを測るためのサーベイシステムを開発することにした。その際、問題 DB の作成が極めて困難であったことから、生成 AI を活用して問題を作成し、研修の一環としてサーベイを実施した結果、「データを扱う」「データによって判断する」スキルが低いことが判明した。あわせて、生成 AI を用いた作問の妥当性及び注意点について、識別率測定や、異なる AI を用いた問題検証を行い、問題の質を向上させる方法について検討した。

**キーワード:** 地方自治体, DX, 人材育成, DX リテラシー, スキルサーベイ, eラーニング

## Construction and validation of a DX literacy standards-based skills survey using generative AI

MASAHIRO SAKAMOTO<sup>†1</sup> YUKIHIRO MATSUBARA<sup>†2</sup>  
KOUSUKE MOURI<sup>†3</sup>

**Abstract:** With the declining birthrate and aging society, the term "possibility of extinction" has come to be used by local governments, and there is no time to spare to improve administrative efficiency through the use of digital technology. In light of this situation, we have developed training contents with the aim of improving DX literacy among Hiroshima City officials. Although the training program has been effective in increasing the number of DX promotions in Hiroshima City Hall, the comprehensiveness and difficulty of the training materials have emerged as issues to be addressed in the future. Therefore, prior to the development of the training contents, we decided to develop a survey system to measure the skill level of the employees in accordance with the "Digital Skill Standards" announced by the Ministry of Economy, Trade and Industry and the IPA. Since it was extremely difficult to create a question DB, we created questions using a generative AI and conducted a survey as a part of training. In addition, we examined the validity and caution in creating questions using a generative AI, measured the discrimination rate, verified questions using different AIs, and examined ways to improve the quality of the questions.

**Keywords:** Local government, DX, human resource development, DX literacy, skills survey, e-learning

### 1. はじめに

筆者は 2016 年、広島市に民間経験者採用（行政職）として入庁し、入庁後は情報セキュリティ研修の委託業務を担当した。その後、自らが職員講師となつての研修コンテンツの作成などに従事している。

こういった自らの担当業務から感じた職場課題である「情報セキュリティ事故の防止」「効率的な情報システムの導入」といった ICT 利活用に関する職員訓練の体系化を目指して、2020 年熊本大学大学院社会文化科学教育部教授システム学専攻博士前期課程に入学し、インストラクショナルデザインについて研究を行った。

丁度、入学直後から新型コロナウイルス対策として集合型研修の全てが禁止されたことをうけて、Moodle を用いた

DX 人材育成を目的としたコンテンツ開発に取り組み、Keller の提唱した ARCS モデル（A：注意（Attention），R：関連性（Relevance），C：自信（Confidence），S：満足感（Satisfaction））[1]のなかでも、特に地方自治体業務の Relevance を重視した eラーニングコースの開発と、その改善を修士研究として実践した[2]。

この実践において、広島市職員・広島市広域都市圏職員から概ねのべ 2,000 名程度の受講者が「手挙げ」で参加し、当初の 2020 年度（第 1 期）・2021 年度（第 2 期）・2022 年度（第 3 期）の 3 年間にかけて、受講者による形成的評価をもとにした情報セキュリティ+DX 人材育成のコース開発に取り組んだ。

その結果、第 1 期では途中離脱が数多く生じたコンテンツの受講データを分析し、たとえば 1 コース 30 分程度だ

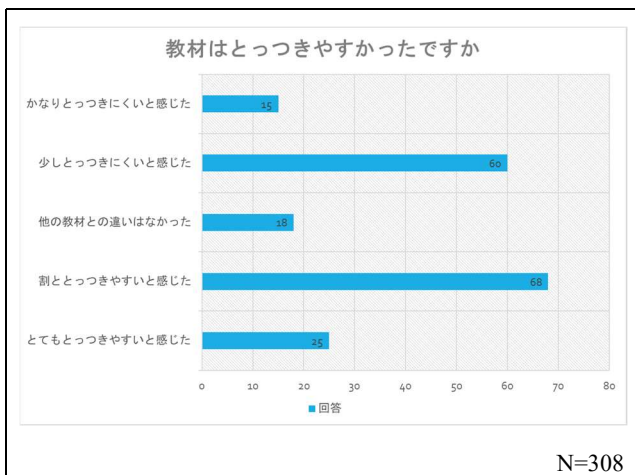
<sup>1</sup> 広島市立大学大学院情報科学研究科  
Hiroshima City University Graduate School of Information Sciences  
and Technology, Asaminami, Hiroshima 731-3166, Japan

った動画について視聴離脱データを元に「地方公務員が業務の合間に最適に閲覧できる15分を最長とする動画コース」に編集したり、業務の手が空いた時間に一気に視聴したいという要望に応じて「倍速再生機能の搭載」をしたり、窓口業務職員ではそもそも音を出しながら視聴できないという要望に対応するため「全て文字起こしを行い、字幕が入ったコンテンツへ変更する」という機能面での改善を行った。

更に、「わからないIT用語がある」「難易度が高すぎる」といった声が受講者の8割から寄せられたことや、コース内の各章テストでの正解率にばらつきがあったことから、教材の課題について小テスト結果データ及び受講者からの形成的評価を元に課題を把握し、第2期開発においては「とっつきやすいコース」となるように改善した。

そうしたところ、第2期においては「とっつきやすかった」と回答した職員が半数となり、しかし残りの半数は「とっつきづらい」と完全に二極化する結果となってしまった(表1)。

表1 第2期における「とっつきやすさ」の2極化



なお、学習目標であった「DXを推進したいと思うようになる」という点については、半数の受講者にとって、とっつきづらいものであったにも関わらず、受講後アンケートにおいて完走者の9割が「DXを受講前よりも身近に感じるようになった」とする積極的な回答であったことから、学習目標は達成できたものと判断した。

こういった状況を踏まえて「地方自治体職員の全員がとっつきやすいDX教材を作成する」という第3期テーマを掲げてコース開発に取り組み、最終的には2極化を収束させ「全ての受講者がとっつきやすい」と考え、「全ての受講者がDXの推進に興味がある」と考えるコースの開発に成功した。

また、受講効果としての行動変容であるが、カークパトリックの4段階モデル[3]で評価したところ、レベル3である「行動」の件数について、情報政策部門で把握している業務改善におけるICT活用支援件数が研修開始年度から連続して倍増していることから(表2)、十分な効果が出てお

り、レベル4の結果についても、広島市のDX事例として例えば「大型ごみを写メで撮影すると必要な処分手数料が表示され、そのままオンライン決済できる」というシステムを導入することで、それまでは電話の問い合わせに対してリストを見ながらコールセンターで対応していた業務が24時間365日対応可能となるといった、コスト削減や市民の利便性向上につながった事例や、地域団体の活性化に向けて、Microsoft 365を活用した情報共有機能や地域活動補助金申請業務支援スマートフォンアプリをローコードツール(Microsoft PowerApps)を用いて市職員が内製で開発したことなど、他都市の範となるDX推進事例も毎年度次々と生まれている状況である。

表2 支援件数の推移

2018年度	28件	
2019年度	30件	
2020年度	23件	
2021年度	49件	☞第1期研修の成果
2022年度	104件	☞第2期研修の成果

## 2. 教材開発における課題

第1章だけでみると良いことづくめであるが、次の課題に直面している。

その課題とは「全員のとっつきやすさを追求したため、教材の難易度が大幅に易化してしまい、簡単な内容のみでコースが構成されている」ことである。

また、筆者が接する職員との交流を通じて「これを学んでほしい」と考えている内容で構成されており、DX推進に必要な学習事項の網羅性にも問題がある。

加えて、広島市職員のDX水準を高め、組織の各部門においてある程度自律的なDX推進を実施するといった組織目標の達成のためには、更に多くの職員が効率的に受講できる環境を提供する必要があると考えた。

この課題に対して、たとえばU-demyといった外部の有料コンテンツを受講させる取り組みなども行っているが、コスト面で全ての職員に行き渡させることは困難であるし、民間事業者を対象としたコースが圧倒的多数であることから、「学んだことを業務に活かす具体的な方法がわからない」といった受講者の声が挙げられており、外部コンテンツのみでは地方自治体におけるDX人材育成の推進は困難であり、地方自治体の業務関連性(=Relevance)と網羅性を考慮したコンテンツ開発が必要であるとの結論に行き着いた。

また、総務省もこの状況には危機感を有しており、(2020年(令和2年)12月策定が初版であった「自治体デジタル・トランスフォーメーション(DX)推進計画」[4]はすでに第3版にまで改められており、早くも最新は2024年4月24日の第3版となっており、それまでの「外部人材の確保」に軸足を置いた内容から、外部人材の確保は民間でもDX

人材不足の状況下で困難なことであり、民間経験者を支援者として送り込んでも「地方自治体業務との Relevance」という点で思ったほど機能していなかったということをおまえ、版を重ねるごとに「内部人材育成」に重点が移りつつある状況である。こういったことを受けて、地方自治体職員に対する効率的な DX 人材育成の仕組みを構築することが必須の情勢である。

一方で、「とはいうものの DX 人材を構成する必要なスキルとは何か」という点についても議論がわかれてしまい、何から手を付けたらよいかかわからない、というも従前から課題であった。

しかし、この「DX 推進に必要な学習パスをはっきりさせてほしい」という主に産業界からの要求に従い、経済産業省・IPAが「デジタルスキル標準」を2022年12月に発表した。

このデジタルスキル標準は、DSS-P と DSS-L の2体系から構成されており、前者がより実践的な (Practical) なデジタル人材を、後者がより基本的 (Literacy) なデジタル人材を定義している[5]。

DSS-P については実際にプログラミング等の実装業務に従事することを想定したスキルセットとなっていることから、DX の要素技術を知り、活用方法を理解することが求められる地方自治体職員に対しては DSS-L に従って人材育成を行うことが最適であると考え、これらの学習目標を網羅するコースを作成することを考えた。

しかし、従来筆者が行ってきた研究では、筆者が重要だと考えた「学習目標」に対してコース設計をしていたことから、DSS-L に対する網羅性が乏しかった。加えて、標準化されたスキルでサーベイする仕組みも存在していないことから、地方自治体職員が現在どの程度の知識を持っていることを確認し、その上で、その知識に対応したコース開発を行うことの必要性に気が付くことができた。

従前より様々な民間事業者等が「DX 時代の教材」は提供しているが、サーベイできないスキルを効率的に育成することは困難であると考え、コース構築に先立ってスキルサーベイのシステムを開発することを思い立った。

その際、民間事業者が提供している研修コースやスキルチェックについても調査したのだが、まだこれらのスキル標準が登場して日が浅いことから「[学習目標]について聞いたことはありますか」といった単純な問題ばかりであり、この種の設問では市内で実施されている様々な研修 (内部統制に関する自治体法務研修、環境問題に関する研修など) において8割以上の職員がその内容に関係なく「どちらともいえない」と回答することを把握しており、テストとして有効に機能していないことを踏まえて、各学習目標に対して具体的な設問を作成することとした。

### 3. 生成 AI によるサーベイ問題 DB の作成

DSS-L は 16 分野から構成されている。これをそれぞれ 10 問ずつ作問しても 160 問必要となるし、さまざまな分野に広がっている DSS-L の各学習領域について優良な問題を作成する方法を考え出すことができなかった。

表 3 DSS-L の 16 分野

分野 1-1	Why	Why-社会の変化
分野 1-2		Why-競争環境の変化
分野 1-3		Why-顧客価値の変化
分野 2-1	What	What-社会におけるデータ
分野 2-2		What-データを読む・説明する
分野 2-3		What-データを扱う
分野 2-4		What-データによって判断する
分野 2-5		What-デジタル技術-AI
分野 2-6		What-デジタル技術-クラウド
分野 2-7		What-デジタル技術-ハードウェア・ソフトウェア
分野 2-8		What-デジタル技術-ネットワーク
分野 3-1	How	How-データ・デジタル技術の活用事例
分野 3-2		How-ツール活用
分野 3-3		How-セキュリティ
分野 3-4		How-モラル
分野 3-5		How-コンプライアンス

出典 デジタルスキル標準 (経済産業省・IPA) [5]

当初は、ITパスポートの問題を抽出したり、DX 検定を参考にしたり、デジタルスキル標準に関するキーワードを検索して問題を自分で作成しようとしていたが、10 問作成するのに1週間以上を要し、DSS-L が策定された直後でもあることから、これに設定された学習目標に対する知識レベルを測定する問題をどのように作ったらよいか皆目見当がつかない状況であった。ならば、より少ない問題数で効果的に測定すればよいだろうという考えもあるが、16 分野の中には学習項目が多い領域もあり、単純に 250 問程度の問題 DB が存在しないとサーベイとして成り立たないという考えに至った。

そこで、生成 AI (Microsoft Bing:現在は Copilot) に「[対象分野]に関する 5 択問題を 10 問作成して」というスクリプト 16 分野分を[対象分野]へ次々と入力して問題を作成した。

Bing については作問根拠が表示されるので、作問根拠を確認し、明らかに的外れなもの以外は、筆者が解いて解釈に悩む問題であっても、筆者自身も DSS-L の全分野について精通しているとは言えないわけであるから、そもそも一度は全て問題として採用し、広島市職員に解答させることで IRT-2PL[6]を用いて問題 DB の「識別率」を測定し、良問・悪問をふるいにかけ、最終的に優良な問題 DB を残すことを想定した。

なお、生成 AI を利用した作問に関する試みとして、プログラミング演習問題自動生成手法[7]、情報教育分野における設問生成[8]、対話型生成 AI を活用した TOEIC 試験対策の一手法[9]、といった事例が見られる。

これらの先行研究は、「生成 AI を用いた作問」という点

においては共通しているが、作問の妥当性に関する評価について、いずれもが「根拠となるテキスト」又は「既存の問題事例」が存在しており、それらを用いて作問について「根拠に対する有効性」について検討されている。つまり、「これを満たすことができれば適切」と想定するコンテンツと対比することができる研究であると言える。

それに対して、本研究において作問の対象とした DX スキル標準については、それを測定するためのサーベイ問題は存在しておらず、なおかつ DX スキル標準の「学習目標」は定めてられているが、具体的な教材例も存在していない。

こういった点を踏まえて、作問の妥当性については今回のスキルサーベイの目的である「地方自治体職員の DX スキルを測定する」ことに焦点を当て「スキルが正しく測定できている設問は、その識別率が高い」という考え方を採用し、なおかつ 300 名超の受験結果データが得られていることから、統計的処理による識別率測定は可能であり、受験結果データを IRT-2PL で測定することで、生成 AI を用いた作問の妥当性を測定する事に取り組んだ。

なお、Microsoft Bing においては「より創造的に」「よりバランスよく」「より厳密に」の生成に関する 3 つのモードがあり、この中で「より創造的に」モードのみ円滑な問題生成が実現できたことから、16 分野すべて「より創造的に」モードで生成した。

図 1 Bing の動作モード

## 4. サーベイ実施環境と結果

### 4.1 実施環境

サーベイは、広島市職員として採用後「中堅職員」として位置づけられる入庁 4～6 年目職員を対象として実施される「中堅職員研修」の 1 つのメニューである「情報セキュリティ・ICT 活用研修」において実施した。広島市においては「行政職・技師職・保健師・消防士・保育士」という大きく 5 区分の職種が存在しており、当該研修については全ての職種において受講するものと位置付けられている。

受講生は広島市企画総務局人事部研修センターが対象者をリストアップし、指定して受講させるため受講者から見た場合は「強制参加型」の研修となっている。

同研修はこれにより指定された中堅職員が年間を通じて各種の研修に参加することになるのだが、筆者が担当している主に「内部不正」に関する理解を学習目標とした「情報セキュリティ研修」について、コロナ以前は集合型研修で実施していたが、コロナにより集合型研修が全廃された令和元年度以後、例年 7 月～8 月の 2 か月間、e ラーニングにより実施することとなっている。

コロナ感染拡大防止のため e ラーニング化された際は「従来の研修を動画で撮影し、庁内ファイルサーバーで共有した動画ファイルを閲覧する」方式で統一されていたが、本研修においては筆者が構築した Moodle 環境へ移行し、同環境の「小テスト」機能を使い、生成 AI が作成した問題を利用したサーベイを実施した。

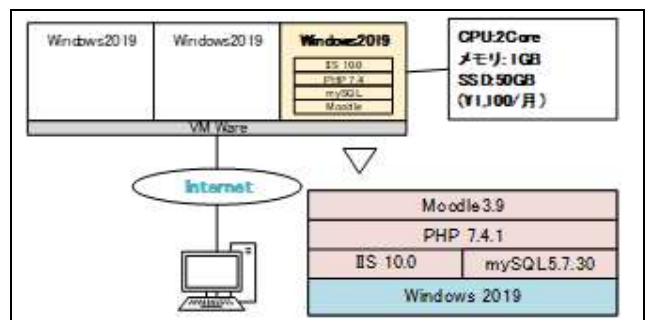


図 2 サーベイ実施環境

## 4.2 実施結果

Moodle のテスト機能を利用し、広島市中堅職員研修（対象者約 400 名）の情報セキュリティ研修の一環として 2023 年 7 月～9 月に実施した。

なお、eラーニング研修においては「模範回答集」が出回ってしまうということや「検索エンジンで回答を調べる」といった不正行為の発生に対応するため、50 分の制限時間を設け、問題母数 250 問から 100 問をランダム出題とする方法とした。

得点分布のばらつきも確認できたことから（図 3）、懸念した不正行為はなかったと考えている。

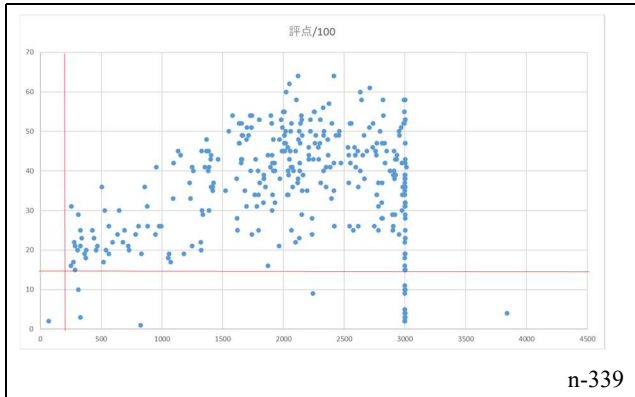


図 3 得点分布

なお、受験結果であるが DSS-L のうち「データを扱う」「データによって判断する」が極めて低い数値であった。減少する地方リソースを有効に活用するため、EBPM(=Evidence Based Policy Making)の推進が強く求められている地方自治体職員においてこの結果は課題であり、データ分析に関する教材の開発から着手すべきという結論を得ることができた（表 4）。

また、筆者が 2022 年度に提供した DX 研修の受講完了者について見たところ、同研修のターゲットとしていた「分野 1-1 Why-社会の変化」の数値が未受講者と比べて良好であり、形成的評価を経て完成させたコースの学習成果について確認できた。

## 5. 生成 AI 作問の評価

### 5.1 識別率の測定

生成 AI による作問は受験者の能力を正しく測定しているのかどうか確認するため、IRT-2PL を用いた「識別率」の測定を行った。

その結果、識別率がマイナスとなる「受験者の能力が識別できていない（不適切）」問題出現分野に偏りがあることが判明した（表 4）。しかし、識別率がプラスの問題のみを抽出して集計した結果と大きな差分は生じていないことから、不適切率の高さが第 4 章で言及した「EBPM 分野の理解度が低い」という評価そのものに与える影響は少ないものと判断できる（表 5）。

表 4 サーベイ結果と識別率

項番	分野	適切のみ	全体	差分
1-1	Why-社会の変化	32.45%	20.28%	12.16%
1-2	Why-競争環境の変化	11.87%	10.19%	1.68%
1-3	Why-顧客価値の変化	11.39%	12.05%	-0.66%
2-1	What-社会におけるデータ	18.50%	18.50%	0.00%
2-2	What-データを読む・説明する	22.91%	18.48%	4.44%
2-3	What-データを扱う	19.32%	17.36%	1.96%
2-4	What-データによって判断する	24.83%	20.48%	4.35%
2-5	What-デジタル技術-AI	18.46%	17.64%	0.82%
2-6	What-デジタル技術-クラウド	11.87%	10.26%	1.61%
2-7	What-デジタル技術-ハードウェア・ソフトウェア	13.01%	11.34%	1.67%
2-8	What-デジタル技術-ネットワーク	15.55%	15.07%	0.48%
3-1	How-データ・デジタル技術の活用事例	21.70%	21.70%	0.00%
3-2	How-ツール活用	23.43%	23.43%	0.00%
3-3	How-セキュリティ	14.25%	13.32%	0.93%
3-4	How-モラル	17.61%	17.61%	0.00%
3-5	How-コンプライアンス	27.68%	25.90%	1.77%
合計		18.06%	16.37%	1.69%

表 5 識別率を考慮した正答率

項番	分野	問題数	適切	不適切	不適切率	正答率
1-1	Why-社会の変化	10	6	4	40.00%	50.22%
1-2	Why-競争環境の変化	10	7	3	30.00%	24.85%
1-3	Why-顧客価値の変化	7	6	1	14.29%	21.30%
2-1	What-社会におけるデータ	6	6	0	0.00%	35.72%
2-2	What-データを読む・説明する	16	11	5	31.25%	37.06%
2-3	What-データを扱う	24	14	10	41.67%	25.95%
2-4	What-データによって判断する	15	9	6	40.00%	23.79%
2-5	What-デジタル技術-AI	38	35	3	7.89%	43.60%
2-6	What-デジタル技術-クラウド	19	15	4	21.05%	36.92%
2-7	What-デジタル技術-ハードウェア・ソフトウェア	33	26	7	21.21%	35.88%
2-8	What-デジタル技術-ネットワーク	14	13	1	7.14%	33.23%
3-1	How-データ・デジタル技術の活用事例	9	9	0	0.00%	36.92%
3-2	How-ツール活用	11	11	0	0.00%	48.64%
3-3	How-セキュリティ	14	13	1	7.14%	44.01%
3-4	How-モラル	11	11	0	0.00%	60.95%
3-5	How-コンプライアンス	8	7	1	12.50%	41.86%
合計		245	199	46	18.78%	37.53%

また、識別率の低い問題は「出題の意図があいまい」なものに多い傾向を見いだせた。たとえば「1-2 Why-競争環境の変化」において識別率が低い問題として「DX によって変化する顧客価値に対応するために、個人が行うべきことは何でしょうか？」という問いが挙げられる（図 4）。

DXによって変化する顧客価値に対応するために、個人が行うべきことは何でしょうか？

A：データやデジタル技術に関する最新情報を常に収集し、学習すること  
 B：データやデジタル技術に関する自身の強みや弱みを把握し、改善すること  
 C：データやデジタル技術に関する自身の目標や計画を立て、実行すること  
 D：データやデジタル技術に関する自身の成果や学びを振り返り、共有すること  
 E：データやデジタル技術に関する他者の意見や視点を聞き、理解すること

ANSWER: E

図 4 Bing により生成された問題例

このことを踏まえ、回答が一意に定まらない問題を発見するため「異なる AI で問題を解かせる」ことに取り組むこととした。

### 5.2 異なる AI を用いた確認

今回の問題は Microsoft Bing を用いて生成したもののだが、この問題を Google Gemini に回答させて、一致しない問題の傾向を更に詳細につかむこととした。

その結果、「Why-競争環境の変化」「What-社会におけるデータ」「How-ツール活用」「How-コンプライアンス」といった分野において特に高い不一致率が出現した（表 6）。

この詳細を確認してみると、識別率の問題と同様「その答えでなければならない」といったものではないところで大きな差異が生じることが判明した。

表6 Google Gemini を用いた再検討結果

	出題数	識別率<0(a)	B<>G(b)	a/出題数	b/出題数
1-1	10	4	1	40.00%	10.00%
1-2	10	3	10	30.00%	100.00%
1-3	7	1	7	14.29%	2.04%
Why	27	8	18	29.63%	66.67%
2-1	6	0	6	0.00%	100.00%
2-2	16	5	8	31.25%	50.00%
2-3	24	10	14	41.67%	58.33%
2-4	15	6	7	40.00%	2.67%
2-5	38	3	13	7.89%	34.21%
2-6	19	4	1	21.05%	5.26%
2-7	33	7	8	21.21%	24.24%
2-8	14	1	0	7.14%	0.51%
What	165	36	57	21.82%	34.55%
3-1	9	0	5	0.00%	55.56%
3-2	11	0	9	0.00%	81.82%
3-3	14	1	4	7.14%	28.57%
3-4	11	0	6	0.00%	54.55%
3-5	8	1	6	12.50%	75.00%
How	53	2	30	3.77%	56.60%
全体	245	46	105	18.78%	42.86%

凡例：B<>G Bing と Gemini が一致しない

### 5.3 別 AI を用いることによる効果

「分野 1-2」については、5-2 において識別率の例として挙げているが、「分野 2-1 What-社会におけるデータ」については、識別率と生成 AI による対比において全く逆の結果となった。この原因について、具体的に Bing が生成した問題と Gemini が回答した内容をいくつか対比して検討することで傾向を把握することにした (図 5)。

◆Bing による問題-回答
DXIによって競争環境を分析する際に必要なデータは何でしょうか? A: 自社の販売実績や顧客満足度などの内部データ B: 競合他社や新規参入者の販売実績や顧客満足度などの外部データ C: 市場規模や成長率、セグメント別需要などのマクロデータ D: データやデジタル技術に関する専門家やアナリストの意見や予測などのクオリティデータ E: 顧客やユーザーの行動や感情、動機などのマイクロデータ ANSWER: E
◆Gemini による回答
該当なし：競合企業の財務データ
◆Bing による問題-回答
データを扱う際には、データの所有権や利用権、個人情報や機密情報の保護など、様々な法的・倫理的な観点に配慮する必要があります。以下のうち、データを扱う際に遵守すべき法律や規則はどれでしょうか? A: 個人情報保護法 B: 不正競争防止法 C: 情報セキュリティマネジメントシステム (ISMS) 基準 D: すべて E: どれも違う ANSWER: D
◆Gemini による回答
A: 個人情報保護法
◆Bing による問題-回答

DXIによって競争環境を分析する際に重要な視点は何でしょうか?? A: 自社の強みや弱みを客観的に評価する視点 B: 顧客やユーザーのニーズや課題を深く理解する視点 C: 競合他社や新規参入者の戦略や行動を予測する視点 D: データやデジタル技術の最新動向を把握する視点 E: 自社の製品やサービスが顧客やユーザーにとってどのような役割を果たしているかを把握する視点 ANSWER: E
◆Gemini による回答
該当なし：グローバルな視点

図 5 Bing と Gemini の対比

Gemini に対するプロンプトとして「この問題を解いて [Bing の問題と選択肢を貼り付け] を投入したのだが、「選択肢に回答はありません-競合企業の財務データです」と表示される場合があったり、または選択肢から Bing の用意した回答と全く違う回答を選択するという場合もあったりした。

この結果を見ると、識別率の問題と同様に「競争環境を分析する際の必要なデータ」については Bing の用意した「選択肢 E」でも間違いではないが、Gemini の示した「競合企業の財務データ (Bing で最も近いのは競合他社という言葉含む選択肢 B)」でも間違いとはいえない。

また、「データを扱う際に配慮する必要があること」についても、Gemini は個人情報保護法のみ、と選択肢から判断しているが、Bing が用意した「すべて」には「個人情報保護法」も含まれるし、ISMS 基準も含まれるということで、これも「どちらもいえない」内容といえる。

同様に、競争環境を分析する際に重要な視点についても、Bing のいずれの選択肢も「なぜそれが間違っているのか」と問われると返答に窮するものであるし、Gemini の回答も「意見の一つとしてはありうる」ものである。

その他大きな乖離がある「分野 3-1, 3-2, 3-5」等についても起きていることは同様であり、各選択肢について「なぜそれが誤りなのか受験者に説明できる根拠が無い」ものを事前に把握するために、異なる AI を用いたクロスチェックの有効性は確認できたと考えている。

## 6. 今後の課題

### 6.1 スキルサーベイにおける生成 AI の活用方法

識別率の観点及び異なる生成 AI を用いるクロスチェックの観点のいずれもが「出題の意図があいまいな作問」について人間の介入が必要であることを示している。

当初の想定では「Bing を用いて生成した問題を中堅研修でサーベイとして回答させ、識別率の低い問題を排除する」ことで品質の高い生成 AI による問題 DB の構築ができると考えていたが、異なる生成 AI を用いたクロスチェックの結果、識別率を指標とするだけでは不十分で「たまたま識別率が正」になっている可能性もあることが判明した。

また、出題の意図があいまいになる原因として、作問プロンプトが「[16 分野]について 5 択問題を 10 問作成して」

という「分野」について幅広い解釈ができることも問題点であったと反省している。

これらの結果を活かして、16分野についてそれぞれ定められた更に細目である学習目標、例えば「分野 1-1 Why-社会の変化」については更にその学習項目例として「メガトレンド・社会課題とデジタルによる解決ーサステナビリティ：SDGs, 持続可能な開発」といったものが定められている。よって、プロンプトも「情報リテラシー標準 Why-社会の変化におけるメガトレンド・社会課題とデジタルによる解決ーサステナビリティ：SDGs, 持続可能な開発 の理解力を測定するための5択問題を3問作成して」とするような工夫が必要であることに気が付いた。

ちなみに上記プロンプトで実施してみたところ、Microsoft Copilot と Gemini の内容は3問とも一致していた。

これら2点の気づきを基に、プロンプトの解像度をより高めて、今回と同様に Microsoft Copilot による作問を Gemini で検証し、完全に一致した新たな問題 DB を作成し2024年7月～8月の中堅職員研修で検証することを予定している。この取り組みを積み重ねて DSS-L 領域においては日本初の「DSS-L を完全に網羅したスキルサーベイ問題 DB」の構築を完了させ、全国の自治体職員から受験可能な状態とすることで、受験した自治体職員が自らのスキルを把握できるようにするとともに、例えば団体規模ごとのスキルの偏在といったものがあるならば、それを「見える化」することで、総務省が実施している「全国均一」の自治体 DX 推進支援施策ではなく、自治体の実情に応じたまさに EBPM な DX 人材支援施策の立案に役立つものになればよいと考えている。

なお、2024年7月8日に受講が開始され、本論文執筆時において40名程度が受験済みであるが、その時点で Moodle の簡易レポートを用いて確認したところ、識別率がマイナスになっている問題は3%未満であり、極めて良好な状態に改善されていることが判明した。

このことから、プロンプトの改善と異なる AI を用いたクロスチェックを組み合わせる作問手法は改善に資すると推測できるが、研修開始日からすぐに取り組む職員は得点も高い傾向にあることから、全受験者の結果をもって最終的な有効性を確認し、新たに把握した問題点と共に、TCE にて発表したいと考えている。

## 6.2 スキルサーベイの「問い」を活用した教材作成

DX スキル標準で定められている学習目標の抽象度が高いことから、教材作成において「その学習目標を理解していると判断する問題」を明確にする、といったことにも今回生成した問題群が活用できると考えている。

例えば分野 1-1 に対応する問題として「第4次産業革命において、新たに重要視されるスキルは何ですか?」というものが生成されている。

これを活用して次は「第4次産業革命において、新たに重要視されるスキルについて5ページ程度の教材を作成して」というスクリプトを投入することで教材作成の効率化にも取り組み、抽象度の高い DX スキル標準を具体化する道筋を立てたいと考えている。

## 6.3 スキルサーベイデータを活用したコース設計

スキルサーベイを行うことで、多くの職員が苦手としている分野からコースを作成するというコースリリースの優先順位の設定は言うまでもないことだが、6.1 で触れたように「より詳細な学習目標」単位で問題生成をしたサーベイを行うことで、個別の分野の中でより丁寧に説明する必要があるものやその逆に軽く触れる程度にとどめておけばよいものといった濃淡をつけることで、「とっつきやすさ」を維持しつつより高い学習効果が得られる工夫に取り組みたい。

また、コースの受講データ～例えば動画の再生速度や飛ばし見をしている～といった受講態度とサーベイデータを組み合わせて分析することで、スキルサーベイで正答した内容に関する部分は飛ばしているが、誤り解答となった部分は何度も繰り返している、ということであれば誤り解答に関する説明に課題があるといった形で、スキルサーベイにおける各項目の正誤と受講態度の関連性について把握することによりデータに基づく教材改善に取り組みたい。

更に、スキルサーベイと動画コースの組み合わせにより、スキルサーベイの結果分野 1-1 の受講を促す、といった取り組みから更に踏み込んで、Microsoft AI Video Indexer といった動画解析 AI を援用することで、誤り解答となった用語の解説部分のみ頭出しする、といった工夫にも取り組んでいきたい。

## 6.4 地方自治体における学びの標準化の実現

本研究の出発点は、筆者の修士研究である「地方自治体職員の業務関連性に着目した DX 人材育成 e ラーニングの開発」であった。

当初は筆者が気になり、職場で話題になっている DX の話題(=Relevance)を基に e ラーニングコースを構築し、形成的評価により改善を重ねるといった取り組みを行っており、体系的な DX スキルについては明確な指標が持てない状況であった。

それに対して、指標となる DSS-L が策定されたことを活かして、本研究における「スキルサーベイ」と、そのサーベイ問題を活用した教材作成に継続し取り組んでいくことで、全16分野に関する研修コースを早急に完成させたい。

完成後は、スキルサーベイ結果で「弱点」とされた分野について受講を促すことで、これまで漠然と「DX 研修」として提供されていた内容を、知識が無い部分だけ効率的に学ぶことができる「個別最適化学習環境」を構築し、受講

後に再度サーベイを実施することで、カークパトリックの4段階モデルのレベル1をDSS-Lの全16領域について完全に確保できるシステムを完成させたい。

そしてこれによって地方自治体職員のDXスキルを早急に底上げすることに資するとともに、現在各自治体で試行錯誤している「DX人材育成」の取り組みを標準化することで、行政の効率化に貢献したいと考えている。

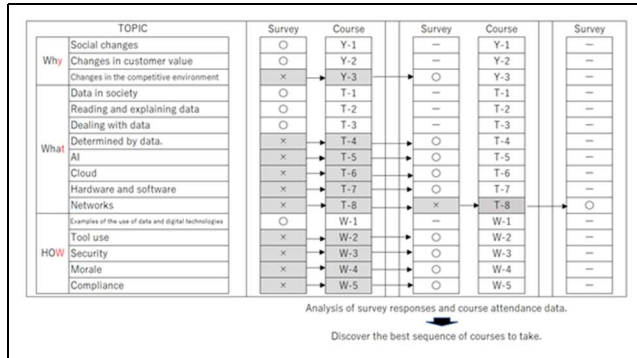


図6 個別最適化コースの全体イメージ

## 参考文献

- [1] ジョン・M. ケラー (著), 鈴木 克明 (監修, 翻訳) (2010) 学習意欲をデザインする: ARCS モデルによるインストラクショナルデザイン
- [2] 坂本昌宏: 地方自治体職員の業務関連性に着目したDX人材育成eラーニングの開発  
[https://idporta.gsis.jp/files/2022\\_sakamoto.pdf](https://idporta.gsis.jp/files/2022_sakamoto.pdf)
- [3] R. A. リーサー・J. V. デンプシー (編) 鈴木 克明・合田美子(監訳) (2013): インストラクショナルデザインとテクノロジー
- [4] 総務省: 自治体DXの推進  
[https://www.soumu.go.jp/denshijiti/index\\_00001.html](https://www.soumu.go.jp/denshijiti/index_00001.html)
- [5] IPA: DXリテラシー標準(DSS-L)概要  
[https://www.ipa.go.jp/jinzai/skill-standard/dss/about\\_dss-l.html](https://www.ipa.go.jp/jinzai/skill-standard/dss/about_dss-l.html)
- [6] 高橋信: IRT項目反応理論入門(2021) p47-48
- [7] 前田 悠翔, 田中 英武, 井垣 宏, 福安 直樹: 多様な問題パターンおよび難易度を考慮した言語系生成AIによるプログラミング演習問題自動生成手法の検討(2024)
- [8] 中鉢 直宏, 久野 靖, 角田 博保, 中山 泰一: 一般情報教育の教科書に基づいた生成AIを使用した情報プレースメントテストのための設問作成の試み(2024)
- [9] 布施 諒一, 林 浩一: 対話型生成AIを活用したTOEIC試験対策の一手法(2023)