

生成系 AI を用いた Moodle 選択問題自動生成に関する研究

北本 卓也

A Study on Automatic Generation of Moodle Multiple-Choice Questions Using Generative AI

KITAMOTO Takuya

(Received September 30, 2025)

キーワード：生成 AI、Moodle、選択問題自動生成

はじめに

大学教育における学習管理システム (LMS : Learning Management System) の普及は、教育実践の効率化と学習者支援の高度化を可能にしている。その中でも Moodle は、オープンソースで広く用いられる LMS として国際的に普及しており、多様な教育場面で利用されている。Moodle のクイズモジュールは柔軟性が高く、授業理解度の確認や formative assessment に有効であるが、問題作成の労力は依然として大きい。特に、多肢選択式問題 (Multiple-Choice Question ; MCQ) における誤答肢 (ディストラクタ) の設計は、教育的妥当性と難易度調整の観点から専門性を要する作業であり、教員の大きな負担となっている。

一方、近年急速に発展している大規模言語モデル (Large Language Models : LLM) は、自然言語処理分野に革新をもたらし、文章生成、要約、翻訳、さらには教育支援分野における活用が進んでいる。特に、教材に基づいた自動問題生成 (Automatic Question Generation : AQG) は、教育工学領域における有望な応用分野として注目されており、生成系 AI による実用的支援の可能性が広がっている。

本研究の目的は、生成系 AI を用いて Moodle に直接インポート可能な選択問題を自動生成する手法を提案し、その実用性と課題を明らかにすることである。具体的には、(1) Moodle XML フォーマットを出力形式としたプロンプト設計の方法論を構築し、(2) 教材資料を入力とする自動問題生成ワークフローを設計し、(3) 実証実験により生成結果の正確性・妥当性・効率性を検証する。本研究の成果は、問題作成に要する教員の負担軽減に寄与するとともに、教育現場における AI 活用の新たな枠組みを提示する意義を持つ。

本論文は以下の構成を取る。第 1 章で関連研究と技術的背景を整理し、第 2 章で提案手法を記述する。第 3 章で実証実験計画を示し、第 4 章で生成事例と初期観察を報告する。第 5 章では結論を述べ、研究の限界と今後の展望を論じる。

1. 関連研究と技術的背景

1-1 Moodle の問題形式と XML フォーマット

Moodle では、クイズ問題を XML 形式で一括インポートできる。このフォーマットは <quiz> 要素を基点とし、各問題は <question> タグ内に <name>、<questiontext>、<answer> 要素を含むことで定義される。従来研究では、教員が作成した問題を自動的に XML 化する支援ツールが開発されており、システムの基盤は整っている。しかし、問題文や選択肢そのものを自動生成する研究は限定的であり、AI 活用の余地が大きい。

1-2 生成系 AI と教育応用

近年の LLM (例 : GPT 系モデル、Claude、Gemini など) は、高度な自然言語理解・生成能力を持ち、教育支援に広範に応用されつつある。Zawacki-Richter et al. (2019) は、高等教育における AI 利用のレビュー

一を行い、教材生成や学習支援が主要な活用領域であることを示した。問題生成についても、資料に基づく自動問題生成の研究が増加しつつあり、本研究の着想を支える背景となっている。

1-3 先行手法の課題点

既存の AQG 研究では、誤答肢の自然さや資料内容との整合性に課題が残ることが報告されている。また、生成問題の XML 化において構文エラーやエンコーディング不整合が生じる技術的課題も指摘されている。したがって、生成過程における「前処理・プロンプト設計・後処理」の統合的設計が不可欠である。

2. 提案手法

本章では、生成系 AI を用いて Moodle に直接インポート可能な 4 択 10 問の選択問題 (Moodle XML) を自動生成するための設計と運用手順を述べる。手法の要は、(i) プロンプト設計による出力の強制整形、(ii) 教材コンテキスト指定による整合性確保、(iii) 選択肢生成ロジック、(iv) XML 出力制約の内在化、(v) 前処理と後処理の二段階品質保証、(vi) ワークフロー／ログ設計にある。

2-1 プロンプト設計戦略：役割付与・仕様列挙・否定指示

本研究の中心は、モデルに対して「望ましい出力形式を強制」しつつ、「教材整合性」と「教育的妥当性」を満たすよう誘導するプロンプトである。設計の基本方針は以下の通りである。

- 役割付与 (role assignment)：冒頭で生成者役割を「大学教授」と明示し、評価・出題の規範を持つ立場での文体・語彙選択を促す。
- 仕様列挙 (specification list)：問題数 = 10、選択肢数 = 4、正答 = 1、正答位置はランダム、重複禁止、タイトル・命名規則、XML の完全構造などを箇条書きで明記。
- 否定・禁止指示 (negative constraints)：「[cite start] [cite] 等のメタ表記を出力に含めない」「XML タグの欠落や入れ子崩れを起こさない」等のやってはならないことを明示。
- テンプレート提示 (format priming)：Moodle XML の完全テンプレートをプロンプト内に埋め込み、モデルが出力をテンプレートに上書き置換する形で生成するよう誘導。
- 一意性制約とランダムイズ：同一問題の禁止、正答位置の均等ランダム化（後述の後処理で分布検査）を指示。

設計上の要点

1. テンプレートを完全提示することで、モデルに「どのタグを必ず埋めるか」を明確化。
2. 命名規則 (タイトル-Number、Number は 01～) を明示し、後処理負担を軽減。
3. fraction=100/0 の二値制約を明記し、採点の一意性を確保。
4. 出力に不要メタ ([cite] 等) が入り込むのを明確に禁止。

2-2 コンテキスト指定

入力資料 (PDF・教科書・スライド) を明示的に指示することで、問題生成が資料に基づくことを保証する。

2-3 選択肢生成ロジック

正答は資料から抽出し、誤答肢は典型的誤解・類似概念・反対語などを基に生成する。さらに、正答位置はランダム配置とし、選択肢間の意味重複を回避する。

2-4 出力形式制約

XML タグの完全性、CDATA の使用、特殊文字のエスケープなどを明記し、Moodle に直接インポート可能な形式を担保する。

あなたは大学教授です。授業の理解度を確認するために授業内容に沿った moodle の選択問題を作りたいと思います。
下記の条件と添付した資料を元にして、moodle の選択問題を日本語で10問作ってください。

```
# タイトル
適当なタイトル
# 条件
- 選択肢の数は 4
- 選択肢の中に正しいものを1つだけ入れる。また正しい選択肢が何番目の項目であるかはランダムにする。
- 次の「XMLファイルのフォーマット」の形のXMLファイルを作る。ただし、「選択問題のNumber」(Numberは 01, 02, 03, ..., 09, 10, 11,
- Problem の所を日本語の問題文で置き換え、Choice1, Choice2, Choice3, Choice4 を日本語の選択肢で置き換える。
- A1, A2, A3, A4 はその選択肢が正解ならば 100 を、そうでないならば 0 を入れる。
- 同じ問題はつくらない。
- 「選択問題のNumberのフォーマット」のタイトル の所には上記の「タイトル」を入れる。
- [cite start] [cite] の情報は「選択問題のNumberのフォーマット」には入れない。
# XMLファイルのフォーマット
<?xml version="1.0" encoding="UTF-8"?>
<quiz>
選択問題の01
選択問題の02
選択問題の03
選択問題の04
選択問題の05
選択問題の06
選択問題の07
選択問題の08
選択問題の09
選択問題の10
</quiz>
# 選択問題のNumberのフォーマット
<question type="multichoice">
  <name>
    <text>タイトル-Number</text>
  </name>
<questiontext format="html">
  <text><![CDATA[<p> Problem </p>]]></text>
</questiontext>
<answer fraction="A1">
  <text>Choice1</text>
</answer>
<answer fraction="A2">
  <text>Choice2</text>
</answer>
<answer fraction="A3">
  <text>Choice3</text>
</answer>
<answer fraction="A4">
  <text>Choice4</text>
</answer>
</question>
```

図 1 システムへの指示文

2-5 入力資料の前処理

OCRによるテキスト化、余分な脚注やページ番号の除去、重要文抽出や要約を行い、入力テキストを簡潔かつ明瞭に整える。

2-6 出力後の後処理・検証

XMLタグの完全性、CDATAの使用、特殊文字のエスケープなどを明記し、Moodleに直接インポート可能な形式を担保する。

2-7 ワークフロー

最終的な流れは「資料準備 → 前処理 → 資料生成 → AI 実行 → 出力後処理 → Moodle インポート」となる。フィードバックループを設け、教員の修正を次回プロンプト設計に反映させる。

3. 実証実験計画

本章では、提案手法の有効性を検証するために実施予定の実証実験計画について述べる。対象科目や教材の選定理由、入力資料の性質、実験の手順、評価指標、想定されるリスクとその対策、ならびに実施計画の概要を体系的に記述する。

3-1 授業科目・教材選定理由

実験対象となる授業科目は、学部レベルの基礎科目（例：教育学部における情報教育関連科目や数学教育法関連科目）を予定している。これらの科目は、多人数を対象とし、知識の理解度確認を行う場面が多く、Moodle を用いた小テストや確認問題との親和性が高い。また、既存教材が PDF やスライド形式で整備されており、問題生成に利用可能な情報資源が豊富である点からも適している。

3-2 入力資料の特徴と分量

入力資料には、以下の種類を用いる予定である：

- 授業スライド（PowerPoint から PDF 化したもの）：章立てが明確であり、出題対象単位の分割が容易
- 教科書該当章の抜粋（約 20 ページ前後）：理論説明が中心で、定義・法則・事例など出題に適した素材を含む。
- 補助教材（配布プリント、図解資料）：理解を補う具体例や演習問題が含まれる。

総分量としては、数万字程度（A4 換算で 30～40 ページ程度）を対象とし、OCR・要約処理を経て文字化を行い、生成 AI で処理してまとめたものをプロンプトに与える。これにより、実験対象範囲として十分な多様性と量的基盤を確保する。

3-3 実験手順

実験の具体的手順は次の通りである：

1. 資料前処理：PDF からのテキスト抽出、OCR、生成 AI による要約・セクション分割を行い、出題範囲を設定する。
2. プロンプト実行：前処理済み資料を入力し、設定済みのプロンプトを用いて 10 問単位の Moodle XML 形式問題を生成する。生成は各資料について 2～3 回繰り返し、再現性と多様性を確保する。
3. 教員チェック・修正フェーズ：生成された問題について、授業担当教員が正答妥当性、誤答肢の適切さ、言語表現を確認し、必要に応じて修正する。この際、修正に要した時間を計測する。
4. Moodle へのインポートおよび確認：検証済み XML ファイルを Moodle にインポートし、問題文・選択肢の表示や採点挙動が正しく機能するかを確認する。

3-4 評価基準・指標

生成問題の評価は、以下の指標に基づいて行う：

- 正確性：問題文および正答が資料内容と一致しているか。
- 妥当性：誤答肢が自然で、学習者にとって紛らわしいが不適切ではない内容になっているか。
- 多様性・重複率：問題や選択肢に過剰な重複がないか、十分な多様性を有しているか。
- 言語表現品質：文法・用語・表現の流暢性と明瞭性。
- 実作業コスト削減効果：教員による修正・補正に要した時間を記録し、従来の問題作成と比較する。

これらを定量評価（正答率、修正時間、重複率）および定性評価（教員コメント）により多面的に分析する。

3-5 期待されるリスクと対策

実験に際して想定されるリスクと、その対策は以下の通りである：

- 的外れな問題生成 → 前処理段階で生成 AI による要約・セクション分割を徹底し、プロンプトの制約条件を強化する。
- XML 構文エラー → 生成 AI のプロンプトやパラメータを見直し、構文エラーの出にくい環境を整える。
- 誤答肢の偏りや不自然さ → 教員によるチェックを必須とし、誤答肢生成ルールの改善を継続的に行う。
- 出題の重複 → 再度生成系 AI を用いて重複を自動検出し、冗長性を回避する。

3-6 実験実施計画

実験は以下の計画に基づき実施する。

- スケジュール：年度後期授業（約 15 週）のうち、4 週分を対象として実施。
- サンプル数：各教材から 10 問×2 セット（合計約 20 問）を生成し、教員確認を経て最終データを作成。
- 実施場所：山口大学教育学部の Moodle 環境にて実施。
- 協力者：授業担当教員 1 名、補助的に学部学生数名が確認作業に参加予定。

4. 生成事例と初期観察

本章では、提案手法を用いて生成された問題事例を提示し、その初期的な観察結果を報告する。代表的な問題例を XML 出力形式とともに示し、良問例と問題点を対比しながら分析する。また、教員による修正内容や Moodle へのインポート過程で発生した事象について整理し、そこから得られる初期的な知見や仮説を提示する。

4-1 代表的な問題例

以下は、授業資料（例：数学に関する基礎資料）を入力として生成された問題例である。

問題文（例）：「空集合に関する基本的な性質として、正しいものはどれか」

選択肢：

1. 空集合は元として「0」を一つだけ持つ。
2. 空集合はただ一つではなく、議論の対象となる全体集合ごとに異なるものが存在する。
3. 空集合は、あらゆる集合の部分集合である。
4. 空集合を元に持つ集合は存在しない。

正答は選択肢 3 であり、XML 出力では fraction 属性を用いて正解ラベルが付与される。

4-2 良問例と問題点例の対比

初期生成結果には、以下のような特徴が観察された。

- 良問例：定義文や明確な概念に基づく設問は、正答が明確で誤答肢も自然に生成されやすい。上記例のように、文献やスライドの記述が直接問題化される場合は妥当性が高かった。
- 問題点例：資料中に曖昧な表現や複数の解釈を許す記述が含まれる場合、正答が一意に決まらない問題が生成されることがあった。また、誤答肢の一部に必ずしも誤りであると断定できないものが含まれることがあった。

4-3 教員チェック時の修正内容とパターン

教員によるレビュー段階では、次のような修正が多く観察された。

- 文言修正：語尾表現や用語統一のための調整（例：「…である」 → 「…であること」）。
- 誤情報補正：誤答肢に資料外の不適切情報が含まれていた場合、教育的に妥当な誤答に差し替え。
- 重複削除：複数の選択肢が意味的に同義である場合、統合や修正を実施。

これらの修正作業は、出題の正確性・妥当性を担保するうえで不可欠であった。

4-4 Moodle インポート時の動作確認とトラブル例

生成された XML を Moodle にインポートする際には、以下の問題が報告された。

- XML 構文エラー：タグ不一致や属性欠落による読み込み失敗。
- エンコーディング問題：日本語の特殊文字（例：全角記号）が原因でエラーを引き起こす場合があった。
- CDATA 範囲の不整合：<![CDATA[…]]> が閉じられていない出力により、問題文が破損して読み込まれるケース。

これらのトラブルは、後処理モジュールや構文検証ツールの導入により軽減可能であることが確認された。

4-5 初期観察に基づく仮説

初期的な分析から、以下の仮説が導かれる。

- 定義文や明確な記述に基づく問題は生成精度が高い → 教材中の定義・用語集部分を強調すれば、正確性の高い問題群を効率的に得られる可能性がある。
- 曖昧な記述は誤答肢の不適切性を引き起こしやすい → 教員による追加修正が必須となる領域である。
- 誤答肢の生成には資料外知識の混入リスクがある → プロンプトで「資料外知識を利用しない」制約を強める必要がある。

これらの観察結果は、今後の評価実験やプロンプト改良に向けた基礎的知見を提供する。

おわりに

本研究は、生成系 AI を活用して Moodle に直接インポート可能な選択問題（MCQ）を自動生成するための枠組みを提案し、プロンプト設計の方法論、前処理・後処理を含むワークフロー、実証実験計画、ならびに初期的な生成事例と観察結果を報告した。これにより、以下の成果と知見が得られた。

第一に、プロンプト設計の工夫により Moodle XML 構造を直接生成可能であることを確認した。具体的には、役割付与、出力制約の明記、テンプレート提示などを組み合わせることで、AI 出力を教育現場で即時利用可能な形式に整えることができた。この点は、教育工学的にも「AI を教材開発の一部工程に組み込む」という実践的価値を持つ。

第二に、教材中の定義や法則など明確な記述に基づく設問は精度が高く生成される一方で、曖昧な記述や複数解釈を許す文脈からは不適切な誤答肢が生じやすいことが明らかとなった。この観察は、AI による問題生成においては資料の種類や性質が出力の質に強く影響することを示唆している。

第三に、教員による最終チェックが依然として不可欠であることが示された。特に、誤答肢の自然さや教育的妥当性、資料外知識の混入の有無は AI 単独では保証しきれず、人的判断を伴う補正作業が必要であった。ただし、その修正は限定的であり、従来の問題作成プロセスに比べて大幅な省力化が可能である点も確認できた。

第四に、XML 構文上の問題やエンコーディング不整合といった技術的課題が観察された。これらは生成 AI を用いてダブルチェックを行うことによって大部分を解決可能であるが、教育現場での実運用に際しては安定性確保のためのシステムの工夫が欠かせない。

研究の限界：

本研究にはいくつかの限界が存在する。

1. 対象範囲の限定性：扱った教材や科目は限定的であり、他分野や高度な専門科目における適用可能性は未検証である。
2. 評価の依存性：成果の評価が主として教員レビューに依存しており、学習者の実際の解答データを用いた妥当性検証には至っていない。
3. モデル依存性：使用した生成系 AI の特性に結果が強く依存しており、モデル更新や他モデル利用時の再現性については担保されていない。
4. 問題生成の範囲：本研究は多肢選択問題に焦点を当てたが、記述式や組み合わせ問題など多様な形式には拡張していない。

今後の展望：

今後は以下の方向での発展が望まれる。

- 大規模・多様な実証実験：複数科目・学習者群を対象とした定量的評価を行い、生成問題の学習成果への影響を明らかにする。
- 教員-AI 協働型の環境設計：AI が自動生成した問題を教員が効率的に確認・修正できる「対話的問題生成支援システム」の構築が求められる。
- 適応的問題生成：学習者の回答ログや理解度データを解析し、個々の学習進度に応じて動的に問題を生成する仕組みの開発が有望である。
- 問題形式の拡張：多肢選択問題にとどまらず、記述式問題や応用的課題への展開を進め、生成系 AI の教育的応用範囲を広げる必要がある。

総括：

総じて、本研究は「生成系 AI を教育実践に統合するための一つの具体的枠組み」を提示したものであり、従来の教員負担を軽減しつつ教育の質を向上させる可能性を示した。今後の研究と改良を通じて、本手法が教育現場に定着すれば、AI が学習支援だけでなく「教材開発プロセスの協働的パートナー」として機能する未来が展望される。

参考文献

Sahoo, S., et al. (2024): A survey on prompt engineering for large language models: Methods and applications. *arXiv preprint*, arXiv:2402.07927.

引用文献

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019): Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16, (39). <https://doi.org/10.1186/s41239-019-0171-0>