



表示、クリック誘導などを組み込み、ICT に不慣れな教員でも反復的に運用できるように設計している。

## 2.2. 評価の持つ権力性

学校における評価は、学習の到達を測る手段であると同時に、「何を学ぶべきか」「どのように表現すべきか」を制度的に決定し、行為を方向づける装置でもある。評価が成績・進級・推薦などの配分と結びつくとき、それは単なる測定ではなく、学習者の時間配分、発言の仕方、文章の構成、さらには学びへの構えそのものにまで影響を及ぼす。この意味で評価は、学習者に外側から働く「圧力」（外部から働く力）であるだけでなく、学習者が自らを調整する「内側の規範」（自分で自分を縛り付ける力）としても作用する。

評価の権力性が問題化する局面は、とりわけ次の三点に整理できる。第一に「不可視(見えない)化」である。採点が採点者の経験や暗黙の重みづけに依存し、基準が十分に言語化されない場合、学習者は「なぜその点数なのか」を理解できず、改善の手がかりを失う。結果として評価は、学習支援よりも序列化となりやすい。第二に「規格(型にはめる)化」である。評価は、観察可能な(表面的な)成果(答案、発言、提出物)を通じて「望ましい書き方・振る舞い」を定型化し、「こう書けば点がもらえる」式の表現様式へ学習を収斂させる。第三に「内面(自分で自分を縛り付ける)化」である。学習者は評価者の期待を推測し、「減点されない書き方」「安全な結論」を選ぶようになり、試行錯誤や異論の提示よりも、評価に適合するパフォーマンスへと誘導される。

記述式評価は、知識再生に加えて、根拠の提示、論理の接続、要約、語彙選択など多層の判断を含むため、採点基準が暗黙知(基準が見えない、説明されてない状態)化しやすく、権力性が強まりやすい。さらに「主体性」など内面に近い概念を評価枠に取り込む場合、評価は学習者の行為だけでなく態度や人格へ接近し、従順さや同調が「良い主体性」として回収される危険を持つ。

## 2.3.『アジュール・カナ』作成の目的

本システムを作成した目的は、記述式評価を「自動化」することではなく、評価に内在する権力性を前提に、その作用を緩和する設計を与えることである。すなわち、採点を「確定判定」で終わらせるのではなく、採点の根拠と観点を開示し、修正と説明が可能な対話的プロセスへと組み替えることを目指す。

第一の目的は、不可視(見えない)化の緩和である。記述式採点では、どこを根拠として評価したのか、どの観点を重く見たのかが採点者の暗黙知に埋もれやすい。『アジュール・カナ』は、観点(例:知識・技能/思考・判断・表現等)を入力として明示し、それに対応するループリック案、減点理由候補、講評案を生成することで、「採点の理由」を言語化しやすい形に整える。これにより、評価を「結果だけ提示する手続き」から「理由を説明できる手続き」へ近づける。

第二の目的は、規格(型にはめる)化の緩和である。評価が「点が取れる型」を生み、学

習を表現の定型へ収斂させることは避けがたい。そこで本システムでは、正解の書き方を一つに限定するのではなく、「点になる要素」をいくつか用意して、様々な良い答え方を拾えるようにする。また、講評案を「良い点／改善点／次の一步」に分け、改善の方向を複線化することで、評価が一つの書き方へ固定する圧力を弱める。

第三の目的は、内面(自分で自分を縛り付ける)化の緩和である。学習者が「減点されない答え」を選び続けると、異論や試行錯誤が出にくくなる。『アジール・カナ』は、答案のどこをどの観点で見たのかを可視化し、改善点を具体化することで、学習者が「評価者の顔色を読む」よりも「根拠を確かめて書き直す」方向へ行為を転換できる条件を整える。特に、史資料に基づく論述では、根拠の提示(資料の数値・語句・時期への言及)を重視し、態度や印象で点が左右される設計を避ける。

以上をまとめると、『アジール・カナ』の目的は、評価の権力を、説明・再検討・対話の方向へできる限り振り向けることである。採点の観点と根拠が見える形にし、判断が修正できる余地を残すことで、学校から権力を可能な限り引きはがし、評価を学習支援の手続きへ近づけようとする実践である。

そのためにAIは「採点者」ではなく、観点・根拠・講評表現の候補を提示する支援役として位置づける。最終的な判断と説明責任は教員が担い、AIの提案を比較・修正できる設計にすることで、評価を「確定判定」から「説明と対話のプロセス」へ接続する。

### 3.運用例—歴史総合のワークフロー

#### 3.1.設計入力(字数上限・配点・観点の設定)



図 2 設計入力の UI

まず、教員は小問ごとに、①字数上限、②配点、③観点(例:知識・技能／思考・判断・表現等)を設定する。ここで重視するのは、「採点者の頭の中にある重みづけ」を外化することである。たとえば「資料に触れているか」「因果が書けているか」「用語が適切か」を観点または期待要素として明示すれば、採点の理由が説明可能になり、評価が印象や態度へ滑る危険を下げられる。

また、必要に応じて禁止事項(史実誤りに直結する語句、論点逸脱など)や、期待要素(必須語句、比較の有無、要因の個数など)も入力し、生成AIの出力が一般論に逃げることを抑える。

### 3.2. 問題投入（問題→OCR）

4. 以下の文章を読み、設問に答えよ。  
問1、左下の図は、19世紀半ばごろのイギリスの国王一家の写真であり、右下は、その当時理想化されていた女性像である。



i) 資料左と資料右の共通点から、資料左が描かれた理由を考え、資料左によってその作者は何を主張したいと判断できるか、あなたの考えを文章で表現しなさい。  
ii) 現代の視点から見て、この資料を見てあなたが考える疑問はなにか判断し、その疑問を「問い」として文章で表現しなさい。

図3 アップロードする問題の例

教員は、定期テストの問題（PDF／画像）をアプリに投入する。歴史総合では、写真・グラフ・史資料抜粋・地図などが混在するため、設問本文だけでなく、資料キャプション、注、設問番号の領域も含めてOCRを行う。

この段階の留意点は、OCR結果を「最終テキスト」とみなさず、教員の手で整形・確認することが可能である。OCRは入力の手軽であり、誤読があり得ることをシステム側も明示する。

### 3.3. 設問整形（OCR崩れの補正→要求の明確化）



図4 アップロードのUI

OCR結果は、段組の混線、改行崩れ、番号欠落などが起こりやすい。そこで『アジュール・カナ』では、対話型生成AIを用いて、行の結合、段落復元、番号補完を行い、設問文を「採点可能な形」に整える。

さらに、整形後の設問について、要求を短い文で明確化する（例：「資料Aから読み取れる変化を二点」「資料Bの語句を根拠に要因を説明」「比較を含める」等）。

この明確化は、採点の不可視性を緩和する第一歩であり、後のループリック設計の土台となる。

なお、整形後の設問は教員が必ず確認し、誤読が疑われる箇所（固有名詞、年代、数値など）があれば修正する。入力段階の確認を省略しないことが、誤りの連鎖を防ぐ条件である。

### 3.4. 模範解答生成（史資料との連関を必須にする）

整形済み設問と設計条件をもとに、対話型生成AIが模範解答案を生成する。歴史総合で重要なのは、模範解答が「それらしい一般論」ではなく、資料の情報（数値、語句、時期、主張）に触れているかである。

したがって、模範解答は次の条件を満たす形で生成させる。

- ① 字数上限内で、要求（理由・比較・因果・根拠）を満たす。
- ② 資料A・Bそれぞれに具体的に言及する（数値・語句・年代等）。
- ③ 推論が飛躍する場合は「要確認」として留保できる形にする。

模範解答は「正解の唯一形」ではなく、採点と講評の出発点となる叩き台として位置づける。

### 3.5. ループリック生成（期待要素・部分点・典型誤り）

次に生成AIが、各小問についてループリック案を生成する。ループリックは、満点条件を文章で述べるのではなく、採点者が確認できる「要素」に分解することがポイントであ

る。たとえば、変化を二点挙げているか、要因を資料Bの言い回しに即して説明しているか、因果が明示されているか、などである。さらに、部分点の条件と、典型的な誤り（論点ずれ、因果の欠落、資料への非言及）も併記する。

この工程により、採点の暗黙知が言語化され、学習者への説明が可能になる。評価の権力性（不可視化）を緩和する中核工程である。

### 3.6.答案投入(OCR/貼付)と小問分割

生徒答案をPDF/画像で投入し、OCRでテキスト化する。答案側でも崩れがあり得るため、必要に応じて手入力・貼付で補正できる導線を確認する。

その上で、答案を小問単位に分割し、字数不足・過多を検出する。記述式では字数が思考の痕跡を必ずしも示さないが、字数情報は「要約不足」「論点過多」の手がかりとして講評に活用できる。

### 3.7.仮採点と講評案(観点別に可視化し、留保を残す)

生成AIは、ループリックと模範解答を参照しながら、観点別の仮スコアと講評案を提示する。講評案は

- ・良い点（根拠の提示、資料参照、論理のつながり）
- ・改善点（欠けている要素、因果の弱さ、資料への非言及）
- ・次の一歩（どう直すか、何を追加するか）

に分けて出力させる。

重要なのは、AIが「断定」しないことである。史実の妥当性や資料解釈が絡む箇所は「要確認」として留保し、教員が照合すべきポイント（資料のどこを見ればよいか）を提示する。これにより、評価がAIの権威に吸収されることを避け、判断と説明の主体を教員側に保つ。

その後、その採点を生徒に根拠とともに示し、生徒がそこに「異議申し立て」を行い、その内容をAIにフィードバックする。

### 3.8.最終化(教員による修正・確定)と次回へのフィードバック

最後に教員が、生徒の「異議申し立て」と照合しながら仮採点・講評案を修正し、最終得点と所見を確定する。修正が入った場合は、どの要素をどう見直したかを簡単に記録し、次回のループリック改善へ回す。

この「修正の履歴化」は、採点の恣意性を減らし、教員間の基準共有や授業改善にも接続し得る。すなわち、『アジュール・カナ』は採点の省力化装置ではなく、評価を説明可能な対話過程へ移し替えるための実務的な枠組みである。

## 4.まとめ

本稿は、歴史総合における記述式評価が抱える「採点負荷」と「評価基準の不可視性（＝

権力性)」に対して、対話型生成 AI を用いた支援システム『アジール・カナ』がどのように介入し得るかを示した。結論として、『アジール・カナ』は採点を自動化して教員を置き換えるものではなく、①観点の言語化(ループリック化)、②根拠の提示と共有(講評案・確認ポイント)、③再検討可能性(複数案の比較と修正)を評価過程に組み込み、採点を「確定判定」から「説明と修正を含む対話的手続き」へ移し替えることを目的とする。

運用例として示したワークフロー(問題→OCR→整形→模範→ループリック→仮採点→最終化)により、教員は「点数を付ける作業」そのものよりも、①資料と照合した妥当性確認、②評価理由の説明、③次の学習へつなぐフィードバック設計に時間を振り向けやすくなる。また、採点の暗黙知が観点・期待要素として外化されることで、学習者にとっても「何が評価され、どう改善できるか」が見えやすくなり、評価が単なる序列化として経験される状況の緩和が期待できる。

さらに、本研究が持つ実践的含意として、対話型生成 AI を共同設計者として活用することで、専門的な開発経験の乏しい教員でも、要件整理と試作・改善を反復しながら、校務の具体的課題に即したツールを構築し得る点を示した。これは、生成 AI 活用を「授業での利用」に限らず、評価実務の改善という領域へ拡張する可能性を示唆する。

評価が不可避に持つ権力性を前提に、それを学習支援へ振り向けるための「設計」と「運用」を具体化することが、生成 AI 時代の記述式評価に求められる課題である。

<sup>1</sup> OCR : Optical Character Recognition (光学文字認識)。画像や PDF の中に写っている文字を、コピーできるテキストに変換する技術 のこと。

<sup>2</sup> User Interface (ユーザーインターフェース) の略で。人がアプリを操作するとき目にして触る部分のこと。たとえば、画面の配置、ボタンの種類、入力欄など、「使う人に見える・操作する部分」全体を意味する。