

マルチモーダル LLM を用いた記述メタデータ付与支援フローの設計 —大学所蔵史料を対象とした実験的検証—

鈴木翔太^{1,2} 加藤諭^{3,4} 片倉峻平³ 菅沼拓夫^{1,5}

概要：大学のデジタルアーカイブには、写真、文書、刊行物等の多様かつ膨大な史料が含まれるため、メタデータの整備には大きな負担が伴う。本研究では、OCR、物体検知、マルチモーダル大規模言語モデルを組み合わせ、統制語彙を参照しながら史料種別、形式・ジャンル、内容・主題、タイトル、説明文等の記述メタデータ候補を生成する付与支援フローを設計した。大学アーカイブズ所蔵史料をモデルとして、単一画像史料ならびに複数画像史料の史料全体 (Manifest) と各構成画像 (Canvas) を対象に付与実験を行い、生成結果の妥当性を大学史料館の専門家が評価した。本稿では、設計したフローの詳細と、実験から得られた有用性および課題を報告する。

キーワード：デジタルアーカイブ、記述メタデータ、マルチモーダル LLM、統制語彙

Design of a Descriptive Metadata Assignment Support Workflow Using a Multimodal LLM —An Experimental Study with University Archival Materials—

SHOTA SUZUKI^{†1,†2} SATOSHI KATO^{†3,†4}
SHUNPEI KATAKURA^{†3} TAKUO SUGANUMA^{†1,†5}

Abstract: University digital archives contain vast and heterogeneous materials, including photographs, documents, and publications, making metadata creation labor-intensive. This study designs a descriptive metadata assignment support workflow that combines OCR, object localization, and a multimodal large language model to generate candidate resource types, genre/forms, subjects, titles, and descriptions with reference to controlled vocabularies. Using materials held by Tohoku University as a model case, we applied the workflow to single-image resources and to both the Manifest-level resources and Canvas-level images of compound resources. An archival specialist evaluated the validity of the generated metadata candidates. This paper reports the workflow design, its utility, and the remaining challenges.

Keywords: digital archives, descriptive metadata, multimodal LLM, controlled vocabularies

1. はじめに

大学のデジタルアーカイブには、所蔵する文化・学術資源を社会の共有財産として公開するとともに、研究・教育によって得られた知見を社会へ還元する役割がある[1]。大学アーカイブズ所蔵史料は、写真、文書、刊行物、図面、物品写真等にわたり、言語、書写形態、年代も多様である[2][3]ため、これらの膨大な史料の検索性を向上させ、利活用を促進する上で、重要となるのがメタデータの整備である。

メタデータの作成には、史料画像に加え、収集経緯や出所等の事前情報を確認する必要がある、大きな人的負担を伴う[4][5]。また、所蔵史料は、単一画像で完結する単一画像史料と、複数画像から構成される複数画像史料に大別されるが、後者では、構成史料全体に対するメタデータのみが整備され、各構成画像の記述が十分でない場合があり、

個々の画像の検索・利用が制約されることが報告されている[5]。このため、史料の早期公開と検索性向上に向けて、メタデータ付与の効率化が求められる。

そこで本研究では、メタデータ付与の効率化を目的とし、単一画像史料と複数画像史料の双方に対応する記述メタデータ付与支援フローを設計する。具体的には、OCR、物体検知、マルチモーダル大規模言語モデル (MMLLM) を組み合わせたメタデータ付与支援フローを構築し、東北大学アーカイブズ所蔵史料の一部を対象に評価実験を行った。

2. 関連研究

デジタルアーカイブを対象としたメタデータ付与支援では、これまで OCR や画像認識等を活用した研究が行われてきた[6][7][8]。これらは AI による大規模処理と省力化の可能性を示した一方、出力は認識文字、物体ラベル、画像領域等の個別観測が中心であり、複数の観測結果を統合

1 東北大学大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University
2 東北大学データ戦略室
Data Strategy Office, Tohoku University
3 東北大学史料館
Tohoku University Archives

4 東北大学統合日本学センター
Center for Integrated Japanese Studies, Tohoku University
5 東北大学サイバーサイエンスセンター
Cyberscience Center, Tohoku University

して史料全体の文脈を記述メタデータとして構造化する点は限定的であった。

LLM の登場後は、タイトル・説明文等の既存属性からメタデータを補完する研究[9]や、OCR 結果を LLM で解釈して属性を抽出する研究[10]が行われている。また、画像・テキスト・表形式メタデータを統合する手法[11]、コンピュータビジョン、LLM 及び Knowledge Graph を組み合わせるフレームワーク[12]、MMLLM によりアーカイブ写真のタイトルや内容説明を生成する研究[13]も報告されている。これらのモデルの活用により、AI の役割は文字・ラベルの抽出から、史料の文脈を踏まえた属性候補の生成へ拡張した。

これらの研究は、AI によるメタデータ候補生成と複数技術の統合可能性を示す一方、新聞、スポーツ画像、ボードゲーム、絹織物、アーカイブ写真、初期刊本等、対象史料又は生成項目を限定したものが多く、抽出対象、候補語彙、正解データ及び評価基準を事前に定義しやすい。これに対し、大学アーカイブズ所蔵史料は膨大かつ多種多様であり、組織の多様性によって言語、書写形態、年代等にも幅がある[2][3][14]。

そこで本研究では、写真、文書、刊行物、図面、物品写真等が混在する大学アーカイブズ所蔵史料を対象に、OCR、物体検知、MMLLM を活用した記述メタデータ付与支援フローを設計した。具体的には、統制語彙に基づく複数の選択項目と自由記述を同時に生成し、単一画像史料及び Manifest/Canvas 階層を持つ複数画像史料を対象として、その有用性と課題を検証した。

3. 記述メタデータ設計

3.1 メタデータ付与対象

本研究では、ISAD(G)[15]及び Dublin Core[16]における記述要素を参照し、タイトル (title)、内容説明 (summary)、史料種別 (jps_type)、形式・ジャンル (material_form)、内容・主題 (content_topic) の記述メタデータを付与対象とした。識別子、権利情報、入手先等は、画像から推定するものではなく、所蔵機関の管理情報や典拠情報に基づき確定すべき項目であるため、本研究の対象からは除外した。

また、付与するメタデータ形式について、各史料の違いを表すタイトルと内容説明は自由記述とする一方、同義語、粒度差、項目間の概念混在を抑え、出力を比較・確認可能な単位へ揃えるために史料種別、形式・ジャンル、内容・主題については、次節に示す候補集合から複数語を選択するマルチラベル課題として定義した。

3.2 統制語彙の選定

候補集合の作成に先立ち、東北大学史料館が所蔵する史料に関する目録記述 89,975 件を対象に、分類的に用いられている非空のユニーク値を抽出した。その結果、写真系 13 種類、文書系 640 種類、計 653 種類の分類的記述が得られた。しかしながら、これらには、「集合写真」「新聞」のよ

うな形式、「学生運動」のような主題、「大学本部」や委員会名のような組織・史料群名が混在し、粒度と概念軸も一定ではなかったため、既存値をそのまま正解ラベルとはせず、史料種別、形式・ジャンル、内容・主題の三つの概念軸へ仕分け、それぞれ統制語彙を設定した。

統制語彙の設定について、大学は、教育、研究、管理運営、学生生活、地域連携等の多様な活動を担っており、その活動に伴って作成・収集される史料も、文書、刊行物、写真、図面、物品等、多種に及ぶ[2][3][14]ため、史料を大別する語彙、史料の形式を表す語彙、内容を表す語彙を、それぞれの役割に応じて組み合わせた。史料種別には、分野横断的なデジタル資源の集約基盤であるジャパンサーチ (JPS) の利活用スキーマ[17]を参照し、形式・ジャンルには、文書・刊行物を含む史料のジャンルの形式を扱う Library of Congress Genre/Form Terms (LCGFT) [18]および、封筒や科学機器等の視覚史料・物質文化を補完するために、Art & Architecture Thesaurus (AAT) [19]を併用した。内容・主題には、日本語史料との親和性を考慮し、国立国会図書館件名標目表 (NDLSH) [20]を採用した。

これらの参照語彙は膨大な語彙を含み、その全てを候補ラベルとして用いることは現実的ではないため、各語彙の定義及び階層を確認し、最終的に JPS 参照 15 語、LCGFT/AAT 22 語、NDLSH 15 語を選定した。各項目の参照語彙、候補数および選定語彙を表 1 に示す。なお、事前に設定した統制語彙だけでは表現が困難な場合に検索・確認上可能な候補を統制語彙外候補 (unmapped) として補助的に保持可能な設計とした。

表 1 統制語彙候補の構成

Table 1 Composition of controlled-vocabulary candidates

項目	参照語彙	候補数	選定語彙
史料種別	JPS	15	photo; docs; offdoc; periodical; poster; map; book; article; postcard; paint; architecture; craft; natural; dataset; other
形式・ジャンル	LCGFT/AAT	22	Photographs; Newspapers; Fliers (Ephemera); Tracts (Ephemera); Ephemera; Posters; Minutes (Records); Annual reports; Personal correspondence; Business correspondence; Maps; Architectural drawings; Conference materials; Programs (Publications); Newsletters; Periodicals; Yearbooks; Essays; group portraits; aerial photographs; envelopes; scientific instruments
内容・主題	NDLSH	15	学生; 学生運動; 建築物; 学校行事; 教育実習; 教育課程; 大学; 高等学校; 労使関係; 国際会議; 教育; PR; 災害; 航空写真; 写真

4. 提案フロー

本研究では、史料画像から直接取得できる情報に加え、史料の取得経緯や背景等、史料取得時に記録される情報（以下、事前情報）を利用して、記述メタデータ候補を生成する。画像由来情報として OCR 及び物体検知の結果を用い、これらを対象画像、事前情報及び統制語彙候補とともに MMLLM へ入力する。単一画像史料のメタデータ付与フローを図 1 に、複数画像史料の同フローを図 2 に示す。

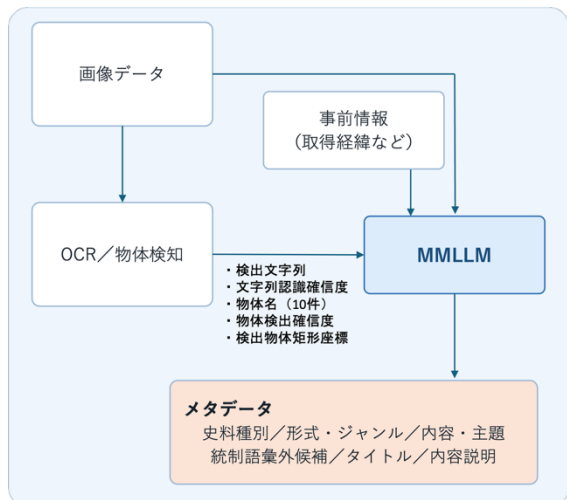


図 1 単一画像史料の記述メタデータ付与フロー

Figure 1 Descriptive metadata assignment workflow for single-image archival materials

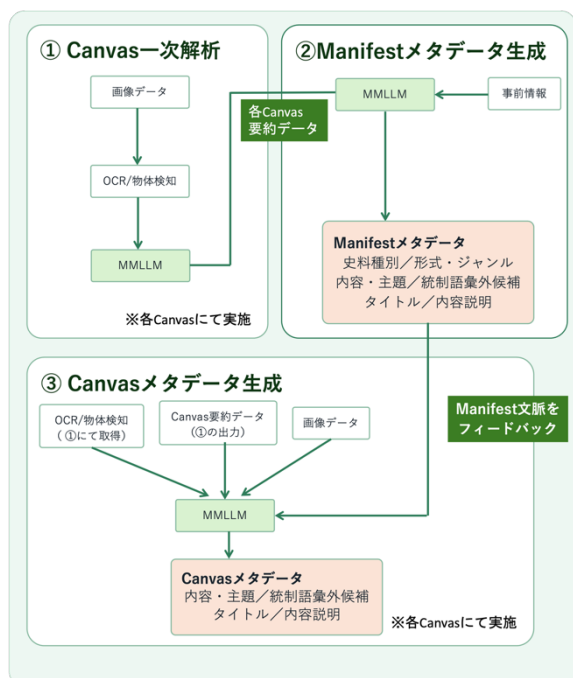


図 2 複数画像史料の記述メタデータ付与フロー

Figure 2 Descriptive metadata assignment workflow for multiple-image archival materials

4.1 単一画像史料

単一画像史料では、まず対象画像に OCR と物体検知を適用する。OCR からは、画像中から認識された文字列と、各単語の認識確信度を平均した値を取得する。物体検知からは、検出確信度上位 10 件について、検出された物体名、確信度及び画像内の位置を示す矩形座標（バウンディングボックス座標）を取得する。

次に、対象画像、OCR 及び物体検知の結果、事前情報を MMLLM へ入力し、史料種別、形式・ジャンル、内容・主題、統制語彙候補、タイトル及び内容説明の候補を生成する。なお、OCR 及び物体検知の結果は確定情報ではなく、確信度を伴う補助的な観測情報として扱い、プロンプトにて、確信度を参考に画像内容との整合性を確認し、信頼性が低い結果や意味的に不自然な結果を無理に採用しないよう指示する。このとき、史料種別は 1 件、形式・ジャンルは最大 2 件、内容・主題は最大 3 件を選択可能とし、統制語彙候補には件数上限を設けなかった。

4.2 複数画像史料

複数画像史料では、史料全体と各構成画像を区別する IIIF Presentation API 3.0[21]の Manifest/Canvas 構造を参照し、Manifest を史料全体、Canvas を構成表示面として、それぞれ以下のフローにて処理する。なお、複数画像史料では、Manifest の史料種別及び形式・ジャンルをそれぞれ 1 件とした。Manifest 及び Canvas の内容・主題には付与件数の上限を設けず、統制語彙候補はそれぞれ最大 3 件とした。

1. Canvas 一次解析

各 Canvas について、対象画像、OCR 結果及び物体検知結果を MMLLM へ入力し、画像単体から読み取れる内容を暫定的に構造化した Canvas 一次解析を生成する。事前情報は史料全体に対して付与された情報であり、個々の Canvas に直接対応するとは限らないため、この段階では入力しない。

2. Manifest メタデータ生成

各 Canvas の一次解析結果と事前情報を MMLLM へ入力し、史料全体を表す Manifest メタデータを生成する。この段階では、各 Canvas の画像を直接入力するのではなく、全 Canvas の一次解析結果を統合することで史料全体の内容を判断する。Manifest には、史料全体の史料種別、形式・ジャンル、内容・主題、統制語彙候補、タイトル及び内容説明を生成する。

3. Canvas メタデータ生成

各 Canvas について、対象画像、OCR 結果、物体検知結果、一次解析結果及び生成済みの Manifest メタデータを MMLLM へ入力し、Canvas メタデータを生成する。Manifest メタデータを入力することで、各 Canvas が史料全体の中でどのような位置づけにあるかを反映する。Canvas には、構成表示面ごとに異なる内容を持ち得る内容・主題、統制語彙候補、タイトル及び

内容説明を生成する。史料種別と形式・ジャンルは史料全体の種別及び形式を表す項目であり、各構成表示面へ重複して付与すると史料全体とページ単体の性質が混在するため、Manifest のみに付与する。

5. 実験方法

設計したメタデータ付与フローを評価するために評価実験を行った。実験には、東北大学総合知デジタルアーカイブ (ToUDA) で公開されている東北大学アーカイブズ所蔵史料のうち、東北大学史料館が所蔵する史料に関する目録記述と、それに対応する画像データを用いた。単一画像史料は、写真史料と文字主体史料を層化し、その構成比を考慮して 50 件を抽出した。複数画像史料は、5 史料 (計 105 Canvas) を対象とし、各史料の全 Canvas を処理して Manifest 及び Canvas のメタデータ候補を生成した。

生成されたメタデータ候補は主著者が確認し、既存メタデータとの差異や修正を要すると判断した事例を含むよう、単一画像史料 20 件、Manifest 5 件、Canvas 20 件の計 45 件を目的抽出した。単一画像史料には写真史料と文字主体史料の双方を、Canvas には 5 史料すべてを含めた。史料館所属の専門家 1 名が、対象画像又は構成画像と候補を確認し、ToUDA の既存メタデータを参考情報として、4=そのまま採用可能、3=軽微な修正で採用可能、2=大幅な修正が必要、1=採用困難の 4 段階で評価した。なお、候補が出力されなかった項目は評価対象外とし、統制語彙外候補は出力された 31 件のみを集計した。本実験に使用したモデル及び API を表 2 に示す。

表 2 実験に使用したモデル

Table 2 Models used in the experiment

処理	利用モデル / API
OCR	Cloud Vision DOCUMENT_TEXT_DETECTION
物体検知	Cloud Vision OBJECT_LOCALIZATION
候補生成	Vertex AI Gemini 3.5 Flash

6. 結果

6.1 専門家評価

評価結果を表 3 に示す。史料種別の平均は単一画像史料及び Manifest のいずれも 4.00 であった。内容・主題は単一画像史料 3.65、Manifest 4.00、Canvas 3.85、全体 3.78 であり、45 件中 42 件が 3 点以上であった。タイトルと内容説明は、Manifest で 3.80 と 4.00、Canvas で 3.70 と 3.80、単一画像史料で 3.15 と 3.10 であった。統制語彙外候補は 31 件で評価され、平均 3.58、25 件が 3 点以上であった。

単一画像史料では新聞 2 件及び公的通知文書 1 件、複数画像史料では Manifest 5 件が、それぞれ「人間による目録作成と遜色ない」と評価された。形式・ジャンルの平均は単一画像史料 3.55、Manifest 3.20 であり、2 点となった事例は 5 件であった。

表 3 評価結果

Table 3 Evaluation results

項目	単一画像 (n=20)	Manifest (n=5)	Canvas (n=20)	全体
史料種別	4.00	4.00	—	4.00
形式・ジャンル	3.55	3.20	—	3.48
内容・主題	3.65	4.00	3.85	3.78
統制語彙外候補	3.28	4.00	4.00	3.58
タイトル	3.15	3.80	3.70	3.47
内容説明	3.10	4.00	3.80	3.51

6.2 単一画像史料の定性的結果

単一画像史料の成功例を図 3 に示す。同左図では、史料種別に文書、形式・ジャンルに envelopes, タイトルに「東北大学金属材料研究所封筒 (和泉教授宛)」が適切に生成された。また、同右図では、史料種別に定期刊行物、形式・ジャンルに Newspapers, 内容・主題に学生運動、学生及び大学が選択され、発行日と号数を含むタイトル並びに主要記事の内容説明が適切に生成された。

一方で誤った内容を付与したものも見られた (図 4)。同左図では科学部寮の門を撮影した写真で表札「二高科学部寮」を「三高科学部寮」とし、同右図では「如春寮」を「和敬寮」と付与した。これらの名称はタイトルと内容説明の双方に用いられ、専門家は検索上の決定的な差異になり得ると評価した。

加えて、既存メタデータ (タイトル) と AI 生成タイトルとで乖離があるものも見られた (図 5)。同図の 4 件では、AI 生成タイトルは画像に写る人物、建物及び場面を記述したが、既存メタデータでは寒稽古、留学生ホームビジット、創立 50 周年記念式典及びボートレース応援団という具体的な行事名が含まれており、本実験にてこれらを再現することは不可能であった。

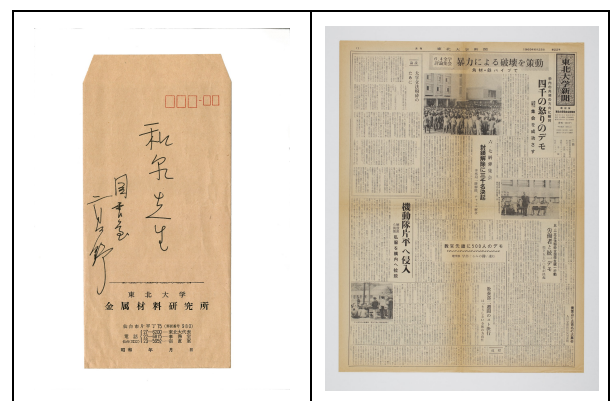


図 3 単一画像史料における記述メタデータ生成の成功例

Figure 3 Examples of successful descriptive metadata generation for single-image archival materials



図 4 画像中の固有名誤読例

Figure 4 Examples of proper-name misidentification in images



図 5 既存タイトルと相違があった例

Figure 5 Examples of discrepancies between existing and generated titles

6.3 複数画像史料の定性的結果

Manifest ではすべてについて、「人間による目録作成と遜色ない」との評価を得た。例えば、「教育に対する本学の基本的姿勢に関する申し入れ」では、全 9 Canvas の記述から、学生増募、教育条件、概算要求及び後半の統計史料を含む内容説明が生成された。

Canvas 一次解析と Manifest 文脈入力後の実出力を表 4 に示す。統計表では、一次解析時のタイトル「昭和 43 年度授業回数、聴講学生数分布調」に対し、Manifest 文脈入力後は「史料 2」が加わり、内容説明にも「教育に対する本学の

基本的姿勢に関する申し入れ」の添付史料であることが記述された。また、評議会規則改正案では、文書内容に加えて、東北大学第二改革委員会報告書の最終ページであることが記述された。

一方で Manifest 文脈による弊害についても確認された。複数枚史料の中で各セクションの区切りを表す空白ページについて、一次解析の段階では付与されなかった Manifest 由来の内容・主題が蛇足的に付与された。専門家評価においても、ページ自身に主題を裏付ける画像又は OCR の証拠がないと付記されていた。

表 4 Canvas 一次解析と Manifest 文脈入力後の生成結果
Table 4 Comparison of generation results from initial Canvas analysis and after incorporating Manifest context

対象	Canvas 一次解析	Manifest 文脈入力後
統計表	タイトル:「昭和 43 年度授業回数、聴講学生数分布調」 内容説明冒頭:「昭和 43 年度における授業回数と聴講学生数の分布を示した統計表 (史料 2) です。」	タイトル:「史料 2 昭和 43 年度授業回数、聴講学生数分布調」 内容説明冒頭:「昭和 44 年 (1969 年) 5 月に東北大学教養部教授会が作成した「教育に対する本学の基本的姿勢に関する申し入れ」の添付史料 2 です。」
評議会規則改正案	タイトル:「大学評議会規則の改定案および審議事項に関する文書」 内容説明冒頭:「大学の管理運営、評議会の権限、教員の採用・昇任基準、学生の定員やカリキュラム調整、および評議会規則の改定 (第 5 条、第 9 条、第 10 条等の変更案) について記された文書です。」	タイトル:「東北大学評議会規則改正案 (第 5 条・第 9 条・第 10 条等の改定および審議事項一覧)」 内容説明冒頭:「東北大学第二改革委員会による報告書の最終ページ (21-22 ページ) です。」

7. 考察

評価において、統制語彙から選択する項目 (史料種別、形式・ジャンル、内容・主題) では、史料種別は評価対象 25 件すべてが 4 点 (最高評価) であり、内容・主題は 45 件中 42 件 (93.3%) が 3 点以上であった。一方、形式・ジャンルは低い評価のものも散見された。

形式・ジャンルで低評価 (2 点) となった 5 件には、いずれも Ephemera が含まれていた。Ephemera はビラやチラシ等の一過性の印刷物・配布物を表す語であるが、文書一般の「その他」のような役割を担い、具体語が候補集合にな

い場合の退避先として広く用いられた。Ephemera を乱用することは、他の史料との境界を曖昧にし、史料の検索性を悪化させるため改善が必要であるが、統制語彙の適切な選択によって改善できる可能性が高い。本研究では、膨大な LCGFT/AAT 語彙から大学史料に関連する語彙を選別して候補集合を作成した。しかしながら、Ephemera が乱用された風刺画、大学組織の系統図及び委員会史料については、LCGFT/AAT の中に Caricatures, organizational charts, administrative records という適切な語彙が存在することが検証後に確認された[18][19]。本研究では、メタデータ付与の発散を防ぐために統制語彙を選別し、それらから複数の語を選択させるマルチラベルタスクとして MMLLM に実行させることを試みたが、膨大なシソーラスから候補語彙を適切に選定する手法の構築が必要であると考えられる。

統制語彙の拡張で対処可能なものがある一方で、委員会名、人名、施設名及び行事名など、史料の特徴を表す上で重要な固有名詞は、一般的な形式・ジャンル語彙で表すことが困難であるため、これらは別に管理する仕組みが必要である。本研究で採用した統制語彙外候補は、既存語彙の追加で解決できる概念と、ローカルに管理すべき固有名を見分けるための確認欄として利用可能であると考えられる。

また、メタデータ生成における入力情報である史料の OCR について、本研究では文字列のみならず平均単語 confidence も併せて MMLLM に入力したが、それが活用された例を図 6 に示す。

同図左側の大学新聞の例では文字情報が主であり、OCR から発行日、号数及び主要見出しが適切に検出され、また平均単語 confidence も 0.8008 と高く、その結果、メタデータ付与においてタイトルに OCR 情報が反映された。一方、同図右側の写真史料では、OCR にて「CPS08 利用下さい」という意味をなさないものが検出されており、ノイズとして扱うのが適切であると考えられる。本史料の平均単語 confidence は 0.4500 と低く算出されており、低い confidence とともに入力された OCR 文字列は、生成結果に反映されなかったものと考えられる。写真史料と文字情報が主である史料を同一システムにて付与する本研究においては、史料種別によって精度差が生じ得る OCR 等については、検出確度を参考情報として活用することが有用であると考えられる。

また、既存タイトルと本フローにて付与したタイトルとの相違(図 5)について、寒稽古やホームビジット等の行事名は実際の画像や事前情報に含まれていないため、タイトルとして付与できなかったものと考えられる。ただし、これは本研究の実験条件として、既存メタデータの備考及び解説を事前情報と定義したためであり、これらの行事名については別の事前情報が存在していた可能性が考えられる。そのため、MMLLM の誤認ではなく、今回の実験条件の制約が原因であると考えられる。

さらに、複数画像史料では、Manifest 文脈により、史料全体における各 Canvas の位置付けを補完することができた一方で、Canvas メタデータ再生成プロセスにおいて、空白ページに Manifest の内容・主題が過剰に継承されることが確認された。空白ページは、構成史料の構造を定める上で重要な役割を担うものの、それ自体に内容上の意味を持つことはないと考えられるため、これは空白ページを検出した際の付与ルールを別途設計することで改善可能であると考えられる。



史料	画像上の表記/OCR 出力	平均単語 confidence
大学新聞(a)	OCR:「東北大学新聞/1969 年 6 月 25 日/第 22 号/暴力による破壊を策動...」 タイトル根拠: image・ocr	0.8008
写真史料(b)	OCR:「CPS08 利用下さい」 生成タイトル:「涌谷町史料館を訪れた人々」 タイトル根拠: image・notes (OCR 文字列は未反映)	0.4500

図 6 OCR 出力の利用例 ((a) 大学新聞, (b) 写真史料)

Figure 6 Examples of the use of OCR output: (a) university newspaper and (b) photographic material

8. まとめと今後の課題

本研究では、JPS 参照候補、LCGFT/AAT、NDLSH への写像と自由記述を組み合わせた記述メタデータ付与支援フローを設計した。複数画像史料には IIIF Presentation API の Manifest/Canvas 構造を参照し、史料全体と構成表示面を区別して記述した。史料館専門家による 45 件の独立評価では、史料種別 25 件が全件 4 点、内容・主題の 3 点以上率が 93.3%であった。一方、形式・ジャンルでは初期候補集合の不足、自由記述では固有名の誤読と入力情報の不足が確認された。また、文字主体史料では OCR が記述生成の根拠となる一方、写真史料の OCR ノイズは生成結果に反映されなかった。複数画像史料では Manifest 文脈が Canvas の位置付けを補った一方、空白ページへの過剰継承も確認した。

今後は、専門家所見と統制語彙外候補の再照合を踏まえ、候補語彙、名称典拠への振り分け及び空白 Canvas への文脈

継承への対応を行う。さらに、生成した統制語彙候補、統制語彙外候補及び内容説明を基に検索用表現を別途設計し、ベクトル検索及びグラフデータベース等へ接続した上で、従来のキーワード検索と比較して検索性能と史料発見に与える効果を評価する。

謝辞 本研究は、JSPS 科研費 JP25H00379 (奨励研究「AIを活用した史料メタデータ作成および検索の実験的検証」)の助成を受けたものである。

参考文献

- [1] 半澤智絵：東北大学における大学デジタルアーカイブの構築，デジタルアーカイブ学会誌，Vol.8，No.2，pp.88-91 (2024). DOI: 10.24506/jsda.8.2_88
- [2] 片倉峻平，塚越柚季，半澤智絵，柳原幸子：東北大学総合知デジタルアーカイブ (ToUDA)：収録資料の特徴と AI 活用の展望，情報知識学会誌，Vol.35，No.2，pp.227-232 (2025). DOI: 10.2964/jsik_2025_020
- [3] 堀井洋，林正治，堀井美里，高田良宏，古畑徹：大学所蔵非文献資料を対象にしたリポジトリの構築，人文科学とコンピュータシンポジウム論文集，Vol.2011，No.8，pp.361-366 (2011). DOI: 10.24517/00028551
- [4] DeRidder, J. L., Presnell, A. A. and Walker, K. W.: Leveraging Encoded Archival Description for Access to Digital Content: A Cost and Usability Analysis, *The American Archivist*, Vol.75, No.1, pp.143-170 (2012). DOI: 10.17723/aarc.75.1.5641v61p422u0u90
- [5] Lapworth, E.: Assessing Large-Scale Digitization Using Web Analytics, *Digital Library Perspectives*, Vol.37, No.2, pp.133-150 (2021). DOI: 10.1108/DLP-09-2020-0095
- [6] 山本純子，濱田陽人，新名佐知子，中里正行：人工知能技術の活用によるメタデータ付与効率化の検証：画像認識とテキスト分類，デジタルアーカイブ学会誌，Vol.5，No.s1，pp.s5-s8 (2021). DOI: 10.24506/jsda.5.s1_s5
- [7] Lee, B. C. G., Mears, J., Jakeway, E., Ferriter, M., Adams, C., Yarasavage, N., Thomas, D., Zwaard, K. and Weld, D. S.: The Newspaper Navigator Dataset: Extracting and Analyzing Visual Content from 16 Million Historic Newspaper Pages in Chronicling America, *arXiv:2005.01583* (2020).
- [8] 青池亨：デジタルアーカイブからの情報抽出技術：画像からのテキスト・図表の抽出，デジタルアーカイブ学会誌，Vol.8，No.3，pp.115-119 (2024). DOI: 10.24506/jsda.8.3_115
- [9] 松澤有三，高橋陽一：大規模言語モデルを利用したメタデータの自動補完，デジタルアーカイブ学会誌，Vol.8，No.3，pp.111-114 (2024). DOI: 10.24506/jsda.8.3_111
- [10] 福田一史：LLM を用いた自動生成メタデータの精度評価：ボードゲームのパッケージ画像 OCR と固有表現抽出から，デジタルアーカイブ学会誌，Vol.9，No.s2，pp.s198-s201 (2025). DOI: 10.24506/jsda.9.s2_s198
- [11] Rei, L., Mladenic, D., Dorozynski, M., Rottensteiner, F., Schleider, T., Troncy, R., Lozano, J. S. and Salvatella, M. G.: Multimodal Metadata Assignment for Cultural Heritage Artifacts, *Multimedia Systems*, Vol.29, No.2, pp.847-869 (2023). DOI: 10.1007/s00530-022-01025-2
- [12] Ignatowicz, J., Kutt, K. and Nalepa, G. J.: Position Paper: Metadata Enrichment Model: Integrating Neural Networks and Semantic Knowledge Graphs for Cultural Heritage Applications, *arXiv:2505.23543* (2025).
- [13] Facklam, D., Sweeney, S. J., Majumdar, S. and Dmello Kamath, R. N.: Describing archival photographs using multimodal LLMs: a case study on evaluating vision-language model performance for creating descriptive metadata, *The Electronic Library*, advance online publication, pp.1-24 (2025). DOI: 10.1108/EL-06-2025-0270
- [14] Salse, M., Guallar-Delgado, J., Jornet-Benito, N., Mateo Bretos, M. P. and Silvestre-Canut, J. O.: GLAM Metadata in Museums and University Collections: A State-of-the-Art (Spain and Other European Countries), *Global Knowledge, Memory and Communication*, Vol.73, No.4/5, pp.477-495 (2024). DOI: 10.1108/GKMC-06-2022-0133
- [15] 国際公文書館会議 (ICA) : ISAD(G) : 国際標準アーカイブズ記述 第 2 版，国立公文書館訳 (2022) (オンライン)，入手先 [〈https://www.archives.go.jp/about/report/pdf/ISAD\(G\)2nd.pdf〉](https://www.archives.go.jp/about/report/pdf/ISAD(G)2nd.pdf) (参照 2026-07-17)。
- [16] DCMI Usage Board: Dublin Core Metadata Element Set, Version 1.1: Reference Description (online), available from [〈https://www.dublincore.org/specifications/dublin-core/dces/〉](https://www.dublincore.org/specifications/dublin-core/dces/) (accessed 2026-07-16).
- [17] 国立国会図書館：ジャパニサーチ開発者向け情報・利活用スキーマ概説 (オンライン)，入手先 [〈https://jpssearch.go.jp/static/developer/introduction/〉](https://jpssearch.go.jp/static/developer/introduction/) (参照 2026-07-16)。
- [18] Library of Congress: Library of Congress Genre/Form Terms (online), available from [〈https://id.loc.gov/authorities/genreForms.html〉](https://id.loc.gov/authorities/genreForms.html) (accessed 2026-07-16).
- [19] Getty Research Institute: Art & Architecture Thesaurus (online), available from [〈https://www.getty.edu/research/tools/vocabularies/aat/〉](https://www.getty.edu/research/tools/vocabularies/aat/) (accessed 2026-07-16).
- [20] 国立国会図書館：国立国会図書館件名標目表 (NDLSH) (オンライン)，入手先 [〈https://id.ndl.go.jp/auth/ndlsh/〉](https://id.ndl.go.jp/auth/ndlsh/) (参照 2026-07-16)。
- [21] IIIF Consortium: IIIF Presentation API 3.0 (online), available from [〈https://iiif.io/api/presentation/3.0/〉](https://iiif.io/api/presentation/3.0/) (accessed 2026-07-16).