

奨学金制度文書からの適格性判定フローの自動抽出

金谷 一輝[†] 宮森 恒[†]

[†] 京都産業大学情報理工学部情報理工学科 〒603-8555 京都府京都市北区上賀茂本山

E-mail: †{g2353351,miya}@cc.kyoto-su.ac.jp

あらまし 本研究では、奨学金制度文書から適格性判定フローを自動抽出における制度文書の適格性判定箇所を抽出する手法を提案する。各種支援制度における適格性判定は、現在も人手による文書解釈と判断に大きく依存しており、人手不足や判断の属人化といった課題がある。従来研究では、制度文書から条件要素を抽出する手法や、意思決定手順をルールベース表現へ変換する試みが行われてきたが、自然言語文書から完全な適格性判定フローを自動生成する手法は十分に検討されていない。本稿では、制度文書を対象として、受給資格要件を抽出し、表やフローチャートを構造化することに焦点を当てる。提案手法では、受給資格要件の抽出において VLM 単体を用いた場合と OCR を用いて文書のレイアウト情報を VLM に渡した場合を比較した結果、レイアウト情報を与えることで文書の抽出精度が向上することがわかった。

キーワード 制度文書, 適格性判定, 判断根拠, 意思決定支援, 大規模言語モデル, 行政手続き支援

1 はじめに

現在、我が国の行政サービスや社会保障制度の多くは、支援を必要とする個人自らが申し出ることによって初めて適用される申請主義に基づいている。しかし、これらの制度を規定する公的な募集要項やガイドラインは、法律用語などの専門的語彙や、複雑な条件分岐などを多用して記述されていることが多く、専門知識を持たない一般市民にとってその内容を正確に解読することは困難である場合も多い。その結果、本来は支援を受ける資格があるにもかかわらず、制度の複雑さゆえに自身の受給資格の有無を判断できずに申請を断念する事例や受給資格を満たす制度を自身で見つけることができないうえ、申請ができない事例などの申請漏れが課題となっている。

OpenFisca [1] などの Rules as Code (RaC) [4] の観点から制度の適格性判断を行うことのできるシステムは存在するものの、制度文書の適格性判定フローを手で構築するには、専門家による文書理解が必要であり、また作成に多大な時間を要する。加えて、受給要件などが変更された際にシステムの維持に課題が残る。そのため、適格性判定フローを自動構築できる手段が社会的に必要とされている。

本研究は、制度文書から適格性判定フローを自動構築するための基礎研究として位置づけられる。Nguyen らは [5] 手続き文書からフローチャートを自動生成するシステムを提案し、その構築過程が手順ステップの特定からフローチャートの組み立てまで 6 段階から構成されることを示している。本稿では、この過程の前提となる、制度文書からの受給資格要件の抽出に焦点を当てる。

制度文書は、フローチャートや表などの視覚的要素を含んでおり、視覚的に情報を処理することが必要である。従来、文書からの情報抽出には OCR (Optical Character Recognition) 技術が用いられてきた。近年、Vision-Language Model (VLM) が

画像とテキストを統合的に処理できる可能性を示している。しかし、日本語の制度文書における VLM 単体での文書処理と、レイアウト解析を経由した処理のどちらが効果的かは明らかでない。本稿では、VLM 単体による抽出と、レイアウト解析を経由した抽出を比較し、制度文書からの受給資格抽出におけるより効果的な手法を明らかにする。

本稿の構成は以下の通りである。第 2 節では関連研究について述べ、第 3 節では比較する 2 つの手法について詳述する。第 4 節では実験設定と結果を示し、第 5 節で本稿をまとめる。

2 関連研究

2.1 制度文書における情報処理

制度文書に含まれる受給資格要件には、条件分岐や場合分けなどの論理的構造があり、片山 [12] は、国民年金法の条文を述語論理により形式的に記述し、SMT ソルバー Z3Py を用いてその正しさを自動検証する手法を提案した。この研究は、被保険者の資格条件や年金の支給要件といった制度の規定が論理式として表現可能であることを示しており、制度文書に含まれる情報が構造的かつ形式的に扱えることを裏付けている。

また、法律文書を対象とした自然言語処理の取り組みとして、狩野ら [10] は国際コンペティション COLIEE を開催し、民法条文からの関連条文検索や含意関係認識といったタスクを設定している。参加システムには TF-IDF や BM25 などの情報検索手法に加え、BERT [2] などの深層学習手法が用いられており、法制度文書への NLP 技術の適用が進展していることを示している。

しかし、これらの研究はいずれもデジタル化済みのテキストデータを入力として前提としている。現実の制度文書は、行政機関から PDF 形式で配布されることが多く、受給資格の条件分岐をフローチャートで示したり、受給資格要件を表形式で整理したりするなど、視覚的なレイアウトによって情報を伝達す

る要素を含んでいる。このような視覚的要素に含まれる情報は、テキスト抽出のみでは分岐構造や要素間の関係性が失われるため、テキストベースの処理だけでは十分に捕捉できない。

本研究では、このような視覚的要素を含む制度文書を対象とし、文書画像からの受給資格要件の自動抽出を目的として、Vision-Language Model (VLM) および OCR を用いた手法の比較検討を行う。

2.2 Vision-Language Model を用いた文書理解

Shi ら [7] は GPT-4V [6] の OCR 能力を定量的に調査し、ラテン系以外の言語では認識精度が低下することを示している。また、緒方ら [11] は土木分野の日本語文書を対象にマルチモーダルモデルと OCR ツールの比較を行い、テキスト認識タスクにおいてマルチモーダルモデルが既存ツールに大きく劣ることを報告している。これらの知見は、日本語制度文書においても VLM 単体での処理に限界がある可能性を示唆しており、本研究で OCR を用いた場合との比較を行う動機となっている。また、文書のレイアウト情報とテキストを統合的に学習する LayoutLM [9] が提案されており、文書画像理解におけるレイアウト情報の有効性が示されている。しかし、日本語の制度文書への適用例は限定的であり、また VLM 単体では制度文書の文書構造の把握に限界がある可能性がある。

3 手 法

3.1 問題設定

本研究では、制度文書から受給資格要件を抽出するタスクを扱う。

入力: 制度文書の画像 $D = \{d_1, d_2, \dots, d_m\}$ (m はページ数)

制度文書は、大きく分けて、制度の説明、受給要件などの判断基準、申請手続きが記載されている。判断基準の記述形式は文書により異なり、文章のみで記述されている場合、フローチャートで記述されている場合、表を用いて記述されている場合などがある。本研究では、受給要件などの判断基準の抽出に焦点を当てる。

出力: 受給資格要件 $Y = (S, G)$

$S = \{s_1, s_2, \dots, s_m\}$ は文章および表形式で記述された受給資格要件の集合であり、 $G = (V, E)$ はフローチャートから抽出された構造である。 V は要件ノードの集合、 $E \subseteq V \times V$ は分岐・遷移関係を表すエッジの集合である。

出力は、文書内の該当箇所を抜き出したものであり、表やフローチャートなどはその構造を反映した形式で記述する。

3.2 手法 A: VLM 単体による抽出

手法 A では、制度文書の画像を VLM に直接入力し、受給資格要件を抽出する。VLM は画像内のテキストおよび視覚的要素を統合的に処理し、受給資格要件に該当する箇所を選択的に抽出する。

3.2.1 処理フロー

1. PDF 文書の各ページを PNG 画像に変換する
2. 各ページ画像とプロンプトを VLM に入力する

3. VLM が画像内から受給資格要件を抽出し、テキストとして出力する
4. 全ページの抽出結果を統合する

3.3 手法 B: レイアウト解析を利用した抽出

手法 B では、VLM への入力の前段階としてレイアウト解析を行い、文書の構造情報を明示的に付与した画像を VLM に入力する。これにより、VLM が文書のレイアウト構造を把握しやすくなることが期待される。

手法 A, B のシステム構造を図 2 に示す。

3.3.1 処理フロー

1. PDF 文書を yomitoku [3] に入力し、レイアウト解析を行う
2. yomitoku がレイアウト情報（見出し、段落、表など）を抽出し、種別をアノテーションした画像を生成する
3. アノテーションした画像とプロンプトを VLM に入力する
4. VLM が画像内から受給資格要件を抽出し、テキストとして出力する
5. 全ページの抽出結果を統合する

図 1 に、教育訓練給付金の制度文書に対する yomitoku のレイアウト解析結果を示す。各構造要素に対してバウンディングボックスと要素種別が付与されていることが確認できる。

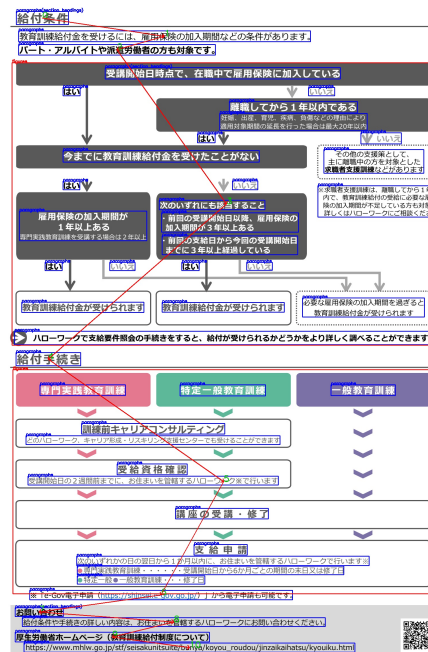


図 1 yomitoku によるレイアウト解析結果の例

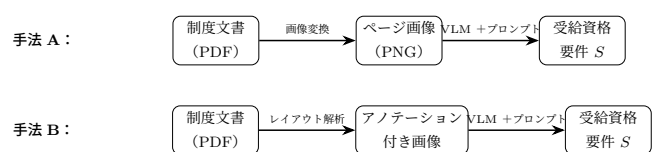


図 2 手法 A および手法 B の処理手順

4 実 験

4.1 実験設定

4.1.1 データセット

本研究では、厚生労働省、日本年金機構、全国健康保険協会および JASSO が公開する pdf 形式の制度文書 9 件（1-40 ページ、中央値 3 ページ）を用いた。内訳はフローチャートを含むものが 2 件、表形式を含むものが 6 件、いずれも含まないものが 2 件である。

4.1.2 実装環境

- VLM モデル: Qwen3-VL-30B [8]
- OCR ツール: yomitoku
- GPU: Quadro RTX 8000

4.2 評価指標

具体的には、以下の 2 点を評価対象として設定する。

- RQ1: 文書の該当箇所を正しく抽出できるか
- RQ2: フローチャートの構造が維持されるか

4.2.1 RQ1 の評価指標

手法 A (VLM 単体) と手法 B (レイアウト解析+VLM) について、文章および表形式を含む文書からの抽出を以下の指標で評価した。ここで、正解としたものは適切に文章を抽出できたものに絞り、要約などの意味的な一致は不正解とする。

- **適合率 (Precision):** 抽出された内容が受給資格要件として正しいか
- **再現率 (Recall):** 受給資格要件の記述が漏れなく抽出されているか
- **F1 尺度**
- **表形式維持率:** 表が含まれる文書において、表の構造が保持されているか

4.2.2 RQ2 の評価指標

手法 A (VLM 単体)、手法 B (レイアウト解析+VLM) について、フローチャートの構造維持を以下の指標で評価した。

- **フローチャートの構造維持:** ノードと分岐条件 (エッジ) が正しく抽出されているか
 - ノード抽出率: 正しく抽出できたノードの割合
 - エッジ抽出率: 正しく抽出できた遷移の割合

4.3 実験結果

4.3.1 RQ1: 文章・表形式の抽出性能

表形式を含む 6 件を含めた 9 件の制度文書を用いて、手法 A と手法 B の抽出性能を比較した結果を表 1 に示す。

表 1 RQ1: 文章・表形式の抽出結果

手法	Precision	Recall	F1	表形式維持率
手法 A	29.69%	45.80%	36.03%	60.00%
手法 B	35.60%	47.24%	40.60%	71.56%

実験の結果、手法 A (VLM 単体) は、適合率が 29.69%、再現率が 45.80%、F1 尺度が 36.03%、表形式維持率が 60.00%であり、

表 2 RQ2: フローチャート構造の抽出結果

手法	ノード抽出率	エッジ抽出率
手法 A (VLM 単体)	77.78%	46.67%
手法 B (レイアウト解析+VLM)	58.30%	66.67%

手法 B (レイアウト解析+VLM) は、適合率が 35.60%、再現率が 47.24%、F1 尺度が 40.60%、表形式維持率が 71.56%であった。

この結果から、全ての評価指標において手法 B の方が上回り、特に抽出精度と表形式維持率においてその差は顕著であることがわかった。

4.3.2 RQ2: フローチャート構造の抽出性能

フローチャートを含む 2 件の文書を用いて、2 つの手法の抽出性能を比較した結果を表 2 に示す。

実験の結果、手法 A (VLM 単体) は、ノード抽出率が 77.78%、エッジ抽出率が 46.67%であり、手法 B (レイアウト解析+VLM) は、ノード抽出率が 58.30%、エッジ抽出率が 66.67%、であった。

この結果からノード抽出率では手法 A の方が上回り、エッジ抽出率において手法 B の方が上回ることがわかった。

4.4 考 察

4.4.1 レイアウト情報の付与による抽出精度の向上

RQ1 の結果から、レイアウト情報を付与した手法 B が全ての評価指標において手法 A を上回った。両手法の出力を比較すると、手法 A には以下の 2 つの誤りの傾向が確認された。

第一に、受給資格要件に該当しない箇所の誤抽出である。例えば、ある文書において手法 A はマイナンバーの提出手続きに関する説明文を「受給資格」「支給対象者」「支給条件」といった見出しとともに抽出した。これらは申請手続きの説明であり、受給資格要件には該当しない。第二に、同一の文章が異なる見出しのもとで繰り返し出力される現象が確認された。これらの誤りにより、手法 A では抽出精度が低下したと考えられる。

一方、手法 B ではこれらの誤抽出が抑制された。これは、レイアウト情報の付与により、VLM が文書内の各領域の役割（見出し、段落、表、手続き説明など）を区別できるようになり、受給資格要件に該当する箇所を選択的に抽出できたためと考えられる。また、表形式維持率においても手法 B が上回った（60.00% → 71.56%）のは、レイアウトアノテーションにより表領域が明示されたことで、VLM が表の構造を保持したまま抽出できたためと考えられる。

また、RQ2 の結果から、VLM にレイアウト情報を与えることでエッジ抽出率が上がった。これは VLM がフローチャートの内容を読み取る能力の向上を裏付けていると考えられる。一方、ノード抽出率においては手法 A が上回った。この結果は、レイアウトアノテーションがフローチャートの領域認識には寄与する一方で、VLM がフローチャートの内部構造を正確に読み取ることができないことを示唆している。レイアウト情報によりフローチャートの存在は明示されるが、ノード間の接続関係やテキスト内容の正確な抽出には、さらなる工夫が必要であると考えられる。

5 まとめと今後の課題

本研究では、制度文書からの受給資格要件の抽出において、VLM 単体による手法とレイアウト解析による構造情報を付与した手法の比較を行った。9 件の制度文書を用いた実験の結果、レイアウト情報の付与により適合率・再現率・表形式維持率のいずれにおいても改善が確認された。特に表形式の維持において効果が顕著であった。一方、フローチャートの抽出ではエッジ抽出率が改善した反面、ノード抽出率では逆の傾向が見られ、レイアウト情報の効果は抽出対象の性質によって異なることが示された。今後の課題として、データセットの拡充、受給資格要件およびフローチャート構造の抽出精度の改善、フローチャート以外の判断基準形式への対応が挙げられる。

謝 辞

本研究の一部は科研費 23K11342 の助成を受けたものである。

文 献

- [1] Openfisca: Introduction.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.
- [3] Kotaro Kinoshita. Yomitoku: Document ai engine specialized in japanese document image analysis.
- [4] James Mohun and Alex Roberts. Cracking the code: Rule-making for humans and machines. *OECD Working Papers on Public Governance*, No. 42, pp. 0_1–109, 2020.
- [5] Nhat-Khiem Nguyen and Thanh-Tung Tran. Iuflowgen: An ai system for converting procedural texts into flowcharts. In *International Conference on Multi-disciplinary Trends in Artificial Intelligence*, pp. 222–234. Springer, 2025.
- [6] OpenAI. GPT-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.
- [7] Yongxin Shi, Dezhi Peng, Wenhui Liao, Zening Lin, Xinhong Chen, Chongyu Liu, Yuyi Zhang, and Lianwen Jin. Exploring ocr capabilities of gpt-4v (ision): A quantitative and in-depth evaluation. *arXiv preprint arXiv:2310.16809*, 2023.
- [8] Qwen Team. Qwen3 technical report, 2025.
- [9] Yiheng Xu, Minghao Li, Lei Cui, Shaohan Huang, Furu Wei, and Ming Zhou. Layoutlm: Pre-training of text and layout for document image understanding. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 1192–1200, 2020.
- [10] 狩野芳伸, 吉岡真治, 佐藤健ほか. Coliee: 法律文書の情報抽出および含意関係認識を行う国際コンペティション. 人工知能, Vol. 35, No. 3, pp. 377–384, 2020.
- [11] 緒方陸, 大久保順一, 藤井純一郎. 土木文書へのマルチモーダルモデル適用へ向けた調査. 人工知能学会全国大会論文集 第 38 回 (2024), pp. 3F1GS1001–3F1GS1001. 一般社団法人 人工知能学会, 2024.
- [12] 片山卓也. 国民年金法の述語論理による記述と検証 smt ソルバー z3py を用いたケーススタディ. コンピュータ ソフトウェア, Vol. 36, No. 3, pp. 3_33–3_46, 2019.