

RAG (Retrieval-Augmented Generation) 技術を応用した保育業務支援ツール Support tool for childcare operations using RAG (Retrieval-Augmented Generation) technology

岡田 怜士[†] 富島 悠介[‡] 丸山 桃子[†] 満永 拓邦[†]
Satoshi Okada Yusuke Tomishima Momoko Maruyama Takuho Mitsunaga

1 はじめに

近年、保育現場での業務過多が問題視されており、子ども家庭庁が ICT (情報通信技術) 化の推進や DX (デジタル・トランスフォーメーション) の実現に向けた事業に力を入れ始めるなどして対策を講じている[1]。保育現場の業務の中でも特に指導計画の立案と作成の業務に着目した。指導計画は厚生労働大臣告示である保育所保育指針によってその作成が規定されており、ある時期における保育のねらいや予想される子どもの活動、保育士が配慮すべき事項等が考慮される。加えて、対象が生後 3 歳未満の児童であれば月ごとに個別の計画を作成することが基本とされている[2,3]。東京都の調査によると、これらの保育の計画に関わる業務は現場の職員にとって負担が大きく、負荷が軽減できると保育の量、質の向上が図れるとした上で、ICT の活用による効率化の期待が高いものとなっている[4]。

他方で、近年は OpenAI 社の ChatGPT をはじめとした LLM (Large Language Models/大規模言語モデル) を用いたサービスが社会的に大きな影響を与えている。これらのいわゆる生成 AI (Generative Artificial Intelligence) はビジネスや教育現場など、あらゆる分野での活用が期待されている。また LLM に関する研究も盛んに行われており、回答精度を向上させる RAG (Retrieval-Augmented Generation/検索拡張生成) をはじめとした新たな応用的な技術も生み出されている。

これらの現状を踏まえ、本研究は LLM や RAG といった技術を用いて、保育現場の指導計画の立案と作成の業務を支援するツールを作成し、保育士の作業効率の向上及び業務負担の軽減に寄与することを目的とする。

2 関連研究・事例

保育現場における AI の導入に関する研究において、本村ら (2019) は保育現場の課題を示したうえで、施設への AI 導入例として、1.保育現場に適したアノテーションソフト、2.保育記録システム、3.危険予測システムの開発を提案した[5]。保育現場における指導計画の ICT 化に関する実証研究は多くないが、実際に企業が提供しているサービスやシステムは保育現場に普及している。株式会社コドモンが提供する ICT サービスでは、5000 を超える文例や過去の指導計画案、保育雑誌などを引用することが可能である[6]。また株式会社ルクミーが提供する ICT サービスでは、上記のような機能に加え、2023 年より ChatGPT を用いた文章生成の機能を追加した。この機能によって保育士は箇条書きの

項目から保護者向けの文章を自動で生成することができるようになった[7]。

また生成 AI の日本における文書作成業務への利用に関する研究として、原田 (2024) は外交実務における外交文書の英文及び和文テキストの要約を行い、効率性の改善を検証した[8]。また中罵ら (2024) は、教育現場における教員の授業設計業務にて、教育方法学の知見に基づいた文章を生成するためのプロンプトを提案した上で、生成 AI が科目の学習目標を明確化するという教員の作業の支援について、検証を通じてその可能性を考察した[9]。

さらに LLM に関する応用的な技術である RAG に関して、そもそも RAG は Lewis ら(2020)の論文によって初めて登場した概念で、パラメーターとして埋め込まれていない外部知識を用いながら文章を生成する技術である[10]。外部のデータを参照することで従来の LLM で課題とされていた生成文の幻覚 (Hallucinations) を低減することに繋がるだけでなく、Adobe 製品に特化した QA システムを作成した Suresh(2024)らのような分野特化型の生成 AI を普及させていくことにも繋がる[11]。

LLM と RAG を用いて保育現場の指導計画の立案と作成業務を支援するツールを作成するという本研究は、保育分野に特化した LLM による AI ツールを作成した点、また RAG により既存サービスと比較して実際の文書や記録に基づいた文章作成を行える点において新規性があると考えられる。

3 提案ツール

3.1 提案ツールの概要

前章までで述べた通り、本研究では LLM や RAG を用いて指導計画を生成するツールを作成する。なお、ここでの指導計画は生後 1 歳未満の児童 (以下、「乳児」) の月ごとの個別計画を対象とした。また指導計画の構成は「1.月のねらい、2.先月末の子供の姿、3.活動内容、4.保育士が配慮すべき事項、5.重大事例の検索」とする。本研究では指導計画の作成において、次の 2 つのタスクに分けて検証を行った。

タスク 1: 対象の乳児の日々の個別記録を参考に次の月の指導計画 (構成 1-4) を作成する。

タスク 2: 過去の保育現場での重大事故や事故に至らなかった (以下、「ヒヤリハット」) 事例を参考に次の月の指導計画 (構成 5) を作成する。

タスク 1 では LLM のみを用いて、タスク 2 では LLM と RAG を用いてツールを作成した。またそれぞれのツール作成においては、言語モデルは OpenAI 社が提供する GPT-3.5

[†] 東洋大学 Toyo University

[‡] 所属なし No Affiliation

Turbo、開発にあたってはオープンソースのフレームワークである LangChain を使用した。

3.2 タスク 1 の詳細

個別記録には対象の乳児に関わる日々の出来事や様子が記録されており、通常、保育士が毎日記入するものである。保育士が業務として月ごとの個別計画を作成する際は、対象となる児童の個別記録を参照し、その月の出来事や乳児の様子をまとめ、次の月の目標や具体的な活動を計画することが最も想定される。これを踏まえ、本研究ではタスク 1 における LLM へのプロンプトに個別記録を含ませることとした。

架空の乳児である C 児の 5 月の個別記録の一部を以下図 1 に示す。

5月1日（水）
登園時、保護者と離れる際に泣く姿があった。そして、ずり這いで部屋の中を移動し、興味を持った玩具に向かって移動をしていた。

5月2日（木）
睡眠の時間が一定になってきた。後追いが盛んになり、保護者、特に母親の後を追う姿が見られる。

5月3日（金）
「頂戴」と保育者が C 児の持っているものに対して言うと、渡す姿があった。また、指さしの意味が分かり、保育者の指さしをしたものの方向を見ることができるようになった。

・・・（中略）・・・

5月28日（火）
戸外に出ることを喜んでいて。外遊びではブルーシートの上に座り、外の空気に触れる心地よさを感じている姿があった。

5月29日（水）
新しい玩具に興味を示し、探索的な行動が活発になっている。睡眠のリズムも安定してきている様子がある。

5月30日（木）
音が鳴る玩具を気に入り、繰り返し使用し遊ぶ姿がある。時々、保育者の膝の上に行き、安心感を得る。

5月31日（金）
外遊びの際に、花や草を渡すと、握ったり、ちぎったりして遊ぶ姿があった。

図 1 架空の乳児である C 児の個別記録（一部省略、筆者作成）

なおここでの個別記録は日付と記録内容を要素としたデータベースに格納されており、入力時にリスト形式のテキストデータとしてプロンプトに含まれる。

またタスク 1 を実行するにあたって LLM を用いたツールの概要図を以下図 2 に示す。

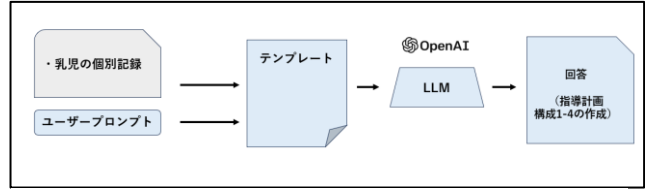


図 2 タスク 1 における作成ツールの概要図 ([12].Figure 1 を基に筆者作成)

今回は図 1 にある個別記録のデータと次の図 3 に示すテンプレートを用いて、LLM に 6 月の指導計画の作成を指示した。

```

template = """
あなたは保育士で乳児の指導計画を作成するスペシャリストです。

下記にする質問の構成は以下の通りです。
**質問の構成**
（乳児の名前）の個別記録を参照して、次の月の指導計画を作成してください。
===
（日にち）：（乳児の出来事）
#繰り返し
****

あなたは、質問にある個別記録を基に「月のねらい」、「先月末の子供の姿」、「活動内容」の3つを作成します。
月のねらい：個別記録に記載されている先月末の出来事をもとに、次の月の目標を立てる。簡条書きで4つ作成する。
先月末の子供の姿：個別記録に記載されている先月末の出来事をもとに、1カ月の出来事をまとめる。実際の出来事を引用しつつ100文字以上で作成する。
活動内容：個別記録に記載されている先月末の出来事や月のねらいをもとに、次の月の目標に対する具体的な活動目標を立てる。簡条書きで5つ以上作成する。
保育士が配慮すべき事項：上記の活動目標に対して、保育士が健康面や安全面で注意すべき事項をまとめる。簡条書きで3つ作成する。

**アウトプットのテンプレート**
1. 月のねらい：
2. 先月末の子供の姿：
3. 活動内容：
4. 保育士が配慮すべき事項：
****

**質問文**
{question}
****

"""
    
```

図 3 タスク 1 において LLM に入力したテンプレート

図 3 内のテンプレートの文章での {question} には、ユーザーの入力プロンプトと図 1 の個別記録のテキストデータが当

てはめられ、まとめて LLM への質問文となる。なお今回の入力プロンプトは「C 児の個別記録を参照して、次の月の指導計画を作成してください。」としている。

3.3 タスク 2 の詳細

保育の現場において、保育士が計画段階から実際に起こった事故や事故の危険があった事例を把握し予防の意識を持つことは非常に重要である。今回はタスク 2 を実行する上で、参考資料として株式会社日本経済研究所がとりまとめた重大事故防止対策に関わる調査研究の報告書[13]と、子ども家庭庁が公表する子どもの事故防止を目的としたハンドブック[14]の 2 つを採用した。これらの資料には重大事故やヒヤリハットの事例、また保育士が注意すべき点に関する情報が多く記載されている。

またタスク 2 を実行するにあたって LLM と RAG を用いたツールの概要図を以下図 4 に示す。

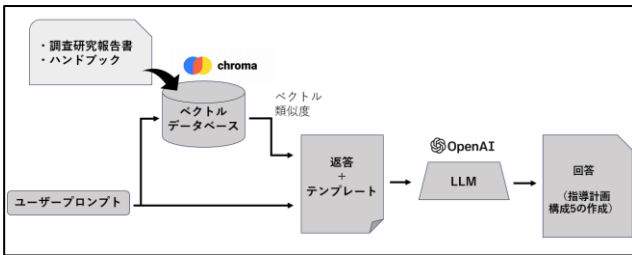


図 4 タスク 2 における作成ツールの概要図 ([12].Figure 1 を基に筆者作成)

またタスク 2 における出力までの手順は次の通りである。

手順 1. 2 つの参考資料を読み込み、テキスト部分のみを抽出して文章を 1 つにまとめる。

手順 2. 文章を 500 文字ごとに区切り、区切った文字列を OpenAI Embedding API を使用してベクトル化し、データベースに保存する。なお今回は保存先として ChromaDB を用いた。

手順 3. ユーザがプロンプトを入力する。

手順 4. プロンプトからベクトル類似度を用いてデータベース内の類似の文字列を検索する。

手順 5. 検索結果とテンプレートとプロンプトを基に LLM が文章を生成し、回答として出力する。

RAG に入力したテンプレートとプロンプトは以下図 5 の通りである。template 内の {context} にはデータベースから検索した類似の文字列が、また {question} にはユーザの入力プロンプトが当てはめられ、まとめて LLM への質問文となる。なお今回の入力プロンプトは「昼寝や午睡、睡眠であった事故やヒヤリハットについて」としている。

```

template = ""
あなたは乳児の指導を行う保育士です。
次のコンテキストと質問トピックに基づいて、該当するトピックにおける乳児の事故や危険のもととなり得るポイントを、事例を交えつつ教えて下さい。
コンテキスト: {context}
質問トピック: {question}

ただし、アウトプットの構成は以下の通りです。
===
5. 重大事例の検索:
(#質問トピック) に対しては、以下の様な危険と事例があります。
-事故や危険の事例
-事前の予防策と保育士が注意すべき点
#これを2つ出力する。
===
""
    
```

図 5 タスク 2 において LLM に入力したテンプレート

4 出力結果

各タスクの出力結果を以下に示す。

```

1. 月のねらい:
- はいはいの上達を促進し、安定した歩行を目指す。
- 他者への関心を高め、コミュニケーション能力を發展させる。
- 離乳食の食べ物の種類を増やし、食事の楽しさを感じる。
- 外遊びを通じて自然と触れ合い、感性を豊かにする。

2. 先月末の子供の姿:
C 児は 5 月末までにははいはいや伝い歩きが上達し、探索活動が盛んになりました。他者への関心も高まり、保育者や他の子どもたちとの関わりを楽しんでいます。離乳食にはまだ苦手意識がありますが、食事の楽しさを感じるようになってきました。

3. 活動内容:
- はいはいや伝い歩きを促進するための遊びを提供する。
- 他の子どもたちとの関わりを増やすためのグループ遊びを取り入れる。
- 離乳食の食材を多様化し、食べる楽しさを感じさせる。
- 戸外遊びを通じて自然と触れ合い、感性を刺激する。
- 音の出る玩具を活用し、聴覚を刺激する。

4. 保育士が配慮すべき事項:
- はいはいや伝い歩きの際に周囲の安全を確保する。
- 離乳食の際に食べ物の大きさや硬さに注意する。
- 戸外遊び時に日焼けやけがのリスクを考慮し、適切な対策を講じる。
    
```

図 6 タスク 1 の出力結果

5. 重大事例の検索：

昼寝や午睡、睡眠であった事故やヒヤリハットに対しては、以下の様な危険と事例があります。

- 事故や危険の事例：

1. 1歳児が昼寝中に少量の嘔吐をしている状況に気付かれなかった。
2. 2歳児が午睡中に熱性痙攣を起こした。

- 事前の予防策と保育士が注意すべき点：

1. 昼寝中や午睡中に子どもたちの呼吸や体調を定期的を確認することが重要です。
2. 熱性痙攣などの緊急事態に備え、保育士は救急救命講習を受講し、適切な対応ができるようしておく必要があります。

図 7：タスク 2 の出力結果

タスク 1 の出力結果 (図 6) を見ると、形式面ではテンプレートで指定した構成や文章の文字数、箇条書きの個数の通りに出力されていることがわかる。また内容面においても、先月末の出来事のまとめ (構成 2) として「はいはい」といった行動や他者との関わり、食事面など個別記録の内容に沿って広く言及できている。さらに、それに基づいた目標の策定 (構成 1) と具体的な活動内容と目標 (構成 3)、また保育士が注意すべき事項 (構成 4) が対応して出力されており、全体として整合性の取れた指導計画になっているといえる。

次にタスク 2 の出力結果 (図 7) を見ると、形式面ではテンプレートで指定した構成や箇条書きの個数の通りに出力されていることがわかる。内容面においても、図 6 における構成 4 の内容と図 7 の内容を比較すると、同じ配慮事項でもその具体性には大きな差がある。LLM 単体で出力された「4. 保育士が配慮すべき事項」の内容は全体のトピックには沿うものの、より一般的な内容に留まる。これに対して RAG を併せて用いた出力である「5. 重大事例の検索」の内容は、参考資料に掲載されている具体的な事例をもとに文章を作成しているため、保育士にとってより参考しやすい内容になるであろう。

またこれらのタスクは個別記録や RAG の参考データ、テンプレート等によって出力の形式や内容が変化するため、各園の方針や形式に沿った柔軟な利用も期待できる。

5 おわりに

本研究では保育現場における保育士の指導計画の立案、作成の業務に着目し、LLM や RAG を用いて乳児の月の指導計画を生成するツールを作成した。本研究で提示したツールは、日々の個別記録から次の月の達成目標等を生成する。さらに特定のトピックを入力することで関連する実際に起こった事故やヒヤリハットの事例とその対応策を生成する。これらの機能は保育士の作業効率の向上及び業務負担の軽減に寄与できると考える。

今後の展望として、本研究で作成したツールや出力結果が指導計画として適切であるかの評価が必要である。具体的には保育士による定性的な評価や、保育現場における実際の個別記録や指導計画を使用した文章の類似度の測定といった定量的な評価等が想定される。

本研究をはじめとした保育分野での LLM や RAG など生成 AI の活用に関する研究が、今後の保育 DX の実現に貢献していくことを強く期待したい。

参考文献

- [1] 子ども家庭庁 (2024) 「資料 6：保育現場での DX の推進について」.会議等.こども家庭審議会子ども・子育て支援等分科会.第 5 回 子ども・子育て支援等分科会.スライド 2,3.
https://www.cfa.go.jp/assets/contents/node/basic_page/field_ref_resources/481073ad-6d4f-4ddb-9f39-13370dbcef18/2eda89ec/20240219_councils_shingikai_kodomo_kosodate_YQvq3ix1_07.pdf
- [2] 厚生労働省 (2017) 「保育所保育指針」.第 1 章の 3.(1),(2).
https://www.mhlw.go.jp/web/t_doc?dataId=00010450&dataType=0&pageNo=1
- [3] 厚生労働省 (2018) 「保育所保育指針解説」 pp.48-50.
https://www.cfa.go.jp/assets/contents/node/basic_page/field_ref_resources/eb316dce-fa78-48b4-90cc-da85228387c2/f4758db1/20231013-policies-hoiku-shishin-h30-bunkatsu-1_24.pdf
- [4] 東京都福祉局(2020)「付属資料 II. 保育施設へのアンケート調査結果 (抜粋)」.子供家庭.保育サービス.「保育事業者の事務負担軽減等に関する調査・分析」報告書について.【保育事業者の事務負担軽減等に関する調査・分析 報告書】.スライド 18,19.
<https://www.fukushi.metro.tokyo.lg.jp/kodomo/hoiku/jimuhutan.files/03.pdf>
- [5] 本村陽一,村田知佐恵,大塚裕子,大森隆司,山田徹志(2019)「保育施設への AI 導入プロジェクトの課題と展望」.2019 年度人工知能学会全国大会 (第 33 回) .
- [6] コドモン.機能一覧.帳票作成 (指導計画・日誌等) .
<https://www.codmon.com/service/docs/>
- [7] ユニファ(2023)「保育 ICT ルクミー、おたよりと帳票に続き、連絡帳にも『ChatGPT』入力補助機能を追加搭載」.
<https://unifa-e.com/archives/002/202310/%E3%83%AB%E3%82%AF%E3%83%9F%E3%83%BC%E9%80%A3%E7%B5%A1%E5%B8%B3%E3%81%AB%E7%94%9F%E6%88%90%E7%B3%BB%E6%A9%9F%E8%83%BD%E6%90%AD%E8%BC%89%E3%80%82%E6%AF%8E%E6%97%A5%E3%81%AE%E9%80%A3%E7%B5%A1%E5%B8%B3%E4%BD%9C%E6%88%90%E3%82%92%E3%82%88%E3%82%8A%E5%8A%B9%E7%8E%87%E7%9A%84%E3%81%AB-final.pdf>
- [8] 原田武夫(2024)「大規模言語モデルの我が国外交における活用：外交文書の要約業務における実装に関する考察」.学習院女子大学紀要.26 号. pp.225-265.
- [9] 中嶋康二,南畑淳史(2024)「教員の授業設計を支援する学習目標明確化のための生成 AI プロンプトの検証」.教育総合研究叢書.17 号. pp.15-27.
- [10] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela(2020)*Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. arXiv:2005.11401.
- [11] Sanat Sharma, David Seunghyun Yoon, Franck Démoncourt, Dewang Sultania, Karishma Bagga, Mengjiao Zhang, Trung Bui, Varun Kotte(2024)*Retrieval Augmented Generation for Domain-specific Question Answering*. arXiv:2404.14760v2.
- [12] Karthik Suresh, Neeltje Kackar, Luke Schleck, Cristiano Fanelli(2024)*Towards a RAG based Summarization for Electron Ion Collider*. arXiv:2403.15729v1.
- [13] 日本経済研究所(2021)「教育・保育施設等における重大事故防止対策に係る調査研究 報告書」.令和 3 年度子ども・子育て支援調査研究事業.
<https://www.jeri.co.jp/parenting-r3/>
- [14] 子ども家庭庁「こどもの事故防止ハンドブックについて」.政策.こどもの不慮の事故を防ぐために.こどもの事故防止ハンドブックについて.
<https://www.cfa.go.jp/policies/child-safety-actions/handbook/>