

令和7年度 成果報告書

基本情報(公開)

事業名 プログラム名		戦略的イノベーション創造プログラム(SIP)第3期 統合型ヘルスケアシステムの構築
研究開発課題名		B-3 症例報告・病歴要約支援システム開発を通じた臨床現場支援
研究開発 担当者* 1	機関名	新医療リアルワールドデータ研究機構株式会社
	所属	データサイエンス室
	役職	室長
	氏名	岡田 昌史
実施期間*2		令和7年4月1日～令和8年3月31日

*1 委託研究開発契約書に定義

*2 年度の契約に基づき、本委託研究開発を行った期間又は中止までの期間

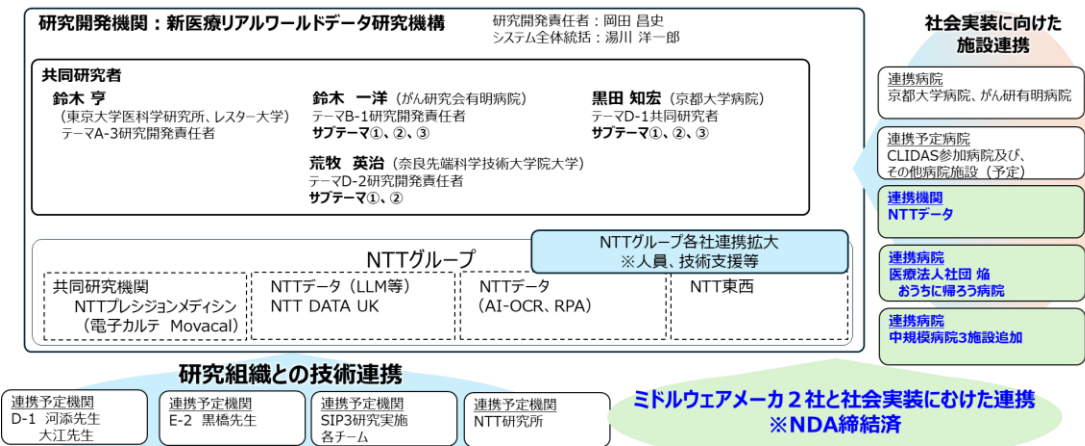
1. 研究開発テーマ概要(公開)

1.1 研究開発内容

医療現場の DX 推進に向け、退院時サマリー、診療情報提供書等の医療文書を対象に、独自アルゴリズムと汎用 LLM を組み合わせた半自動生成技術を開発することで、医療従事者の業務稼働削減を実現する。併せて生成過程で得られる構造化データをリアルワールドデータ(RWD)として蓄積し、創薬研究等に活用することで医療の発展に寄与する。

1.2 研究開発実施体制

新医療リアルワールドデータ研究機構を中心に、大学病院、電子カルテベンダ、ミドルウェア事業者、NTT グループ各社と連携し、研究開発・実証・社会実装を推進する。



1.3 研究推進のスケジュール

令和 7 年度は、実運用環境での継続的な実証を進め、次年度以降の商用化・展開に向けた準備を実施する。

	2023年度	2024年度	2025年度	2026年度	2027年度
癌領域	プロトタイプモデル検討・構築	検証、品質向上に向けたチューニング循環器へ拡大	複数施設へ展開	他研究成果を用いた品質向上	一部事業化展開
循環器	癌領域からスタート (単一施設)	実際の施設データでの検証 京大病院	他施設データでの検証 がん研有明病院	他施設展開	
テーマ①	Cyber Oncology システムに登録された構造化データでの出力研究	テーマ①にて生成した構造化データを用いた出力検証	性能評価 異なる施設データにおける品質検証	▲他研究成果を用いた品質向上	さらなる品質向上 他研究との連携 (D-2 E-2での成果物の取込み)
テーマ②	京都大学医学部附属病院にてプロトタイプ構築	自動生成された構造化データと手動入力したデータを比較検証	性能評価 異なる施設データにおける品質検証	▲他研究成果を用いた品質向上	さらなる品質向上 他研究との連携 (D-2 E-2での成果物の取込み)
テーマ③	導入に向けた基礎的調査の計画 ニーズ調査	テーマ①、テーマ②成果を組み合わせたシステム化 (既存サービスとの連携含む) 社会実装に向け商用化検討	▲2025年度末までに中核病院にてシステム構築	▲2025年度末までに症例報告・病歴要約支援システム一部事業化	▲2027年度末までに事業化
テーマ④	英国における先駆的取り組みの調査	英国における先駆的取り組みの調査 ※調査内容の拡大	テーマ①、テーマ②、テーマ③に関する精度向上に向けた支援及び、グローバル的な側面からのチェック、海外事例等の取込み支援	複数の電子カルテベンダとの連携検証	
			英国における臨床現場支援に向けたシステム開発動向の研究 ※英国における医療DX、電子カルテの根拠を超えた情報連携等に関する調査 ※テーマ②研究実績評価 (テーマ②で作成された構造化データの国際的な互換性の確認)		

2. 本年度の成果・進捗の概要(公開)

【技術】独自アルゴリズムにより薬剤処方・臨床検査等から治療変化点を抽出し、汎用 LLM と組み合

【経理様式 1 別添】【R7】

わせて文書生成の精度向上を図った。

【制度】医療現場における LLM 適正利用に向け、関連法令・ガイドラインを踏まえた検討を継続した。

【事業】大規模病院向け・中規模病院向けの二軸でプロトタイプ検証を進め、中規模病院向けは電子カルテへの実装・商用化準備を推進した。

【社会受容性】医師・医療従事者へのヒアリングおよびマーケティングリサーチを継続し、受容性向上の観点で改善点を整理した。

【人材】博士課程・修士課程学生を含む体制強化や、他大学研究室等との連携拡大を進めた。

<2025 年度上期取り組み>

テーマ①「医療現場支援ツール開発」

テーマ②にて生成した構造化データおよび、各種データから自動検出した治療における変異点抽出機能を用いた出力の評価結果を検証。電子カルテベンダ、医療文書管理ベンダと連携し、京都大学医学部付属病院の既存の医療文書作成環境から API を用いて連携できるプロトタイプを開発し、実際に先生方に利用してもらい評価を実施。また評価結果を元に、利用性向上をめざしチューニングを実施。

またテーマ④の検証で得られた情報等との比較検証を実施。

<アプリケーション画面遷移(2025 年 8 月時点版)1 大規模病院向け >

京都大学医学部付属病院にて、電子カルテに実装しているプロトタイプ版画面遷図は以下の通り。今回のシステムは医師がストレスなく利用できるように、これまで利用していた環境に自然に溶け込む形のシンプルな構成を実現しており、テスト利用した医師からのアンケートなどを基に精度向上をはかる。

また今回のプロトタイプモデルは、実際の病院での利用を考慮し、医療機関等で通常利用される 1U サーバに搭載可能な GPU で動くシステムを想定した環境で検証を実施している。

① 院内の文書作成システムから起動してテンプレートを選択する

文書生成

透過

×

対象とする退院日を指定して文章を生成してください。

退院日

入院歴から選択

2024/05/06

文書生成

{PR: 日付 / イベント名 / サマリ}

AIで生成する文章の定型文テンプレートを表示しています。
文章を生成してください。

文章はAIで自動生成したものです。内容が正しいか確認の上ご活用ください。

② 患者 ID・退院日に紐づくイベントリストを返し、文書作成ツールに提供する

文書生成

透過

×

対象とする退院日を指定して文章を生成してください。

退院日

入院歴から選択

2024/05/06

文書生成

- 2025/06/01 / 診察開始 / 外来患者の初診診察を開始

- 2025/06/02 / 血液検査 / 血液検査オーダーを実施

- 2025/06/10 / 診断確定 / 高血圧症の診断を確定

AIによって生成された文章を表示しています。
AI生成箇所をクリックすると選択できます。
選択すると、根拠と推定される診療録の記述の確認や、値の変更
を行えます。

文章はAIで自動生成したものです。内容が正しいか確認の上ご活用ください。

③ 行をクリックすると右側に根拠となる SOAP 情報が表示される。

文書生成

透過

×

対象とする退院日を指定して文章を生成してください。

退院日

入院歴から選択

2024/05/06

文書生成

- 2025/06/01 / 診察開始 / 外来患者の初診察を開始
 - 2025/06/02 / 血液検査 / 血液検査オーダーを実施
 - 2025/06/10 / 診断確定 / 高血圧症の診断を確定

AI生成結果に戻す

- 2025/06/02 / 血液検査 / 血液検査オーダーを実施

AI生成結果
 2025/06/02
 [S]
 現病歴：1年前から高血圧と診断。
 降圧剤を内服中
 [O]
 血圧：150/90 mmHg
 [A] 降圧目標値（140/90 mmHg未満）には未達。

文章はAIで自動生成したものです。内容が正しいか確認の上ご活用ください。

テーマ②「データ構造化」

1) 最新のオープンソース LLM の構造化性能評価結果

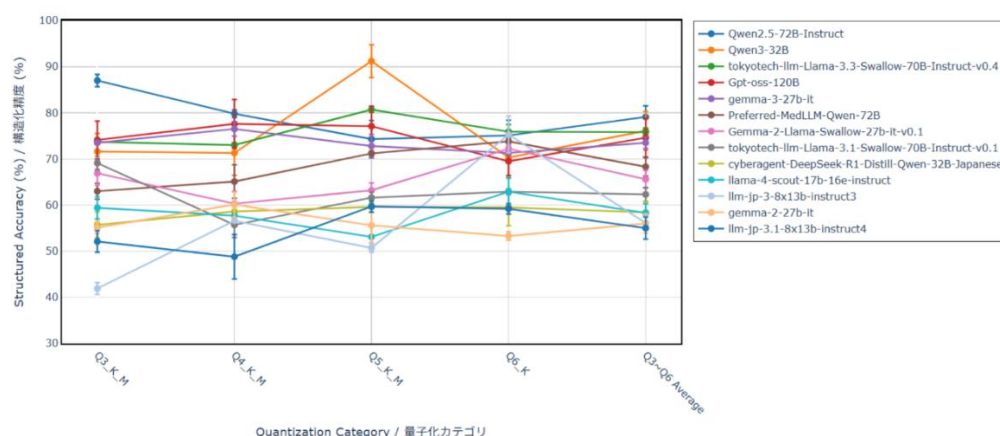
a) 構造化精度のモニタリング

ここでは、Hugging Face 上に公開される最新の LLM モデルを対象に、カルテに記載される経過記録などの非構造化テキストの読み取り精度を継続的に評価している。現時点での評価結果を下図に示す。図の横軸は、モデルの軽量化を目的に量子化するビット数を表しており、縦軸は構造化精度である。非構造化テキストとしてがん治療の実臨床を模したカルテの経過記録を新たに作成し、各 LLM モデルに対し構造化した臨床情報を抽出する指示を与えている。抽出する情報は昨年度の 3 倍以上の 160 項目程に拡張しており、その抽出精度を量子化ビット数毎に評価している。評価は各点において 6 回行い、その平均値と標準偏差をエラーバーで示している。現時点での評価結果は以下の通りである。

- ① 現在のモデルは、一昨年のモデルと比較して、構造化精度が全体的に約 20 ポイント向上しており、構造化精度は 50～90% 辺りに分布している。
- ② 一昨年のモデルでは、量子化ビット数が増えるほど構造化精度も向上する傾向が見られた。しかし最近のモデルでは、特定の量子化ビット数で精度が最大となる傾向があり、これは現行モデルの特徴といえる。
- ③ パラメータサイズ 32B モデルが 120B モデルよりも高い精度を示すケースが見られるようになっており、パラメータのサイズよりも「リーズニング(推論)能力」の強化が性能向上に大きく貢献していることが、本研究でも明らかになっている。
- ④ 医師国家試験問題や IgakuQA タスクは、LLM の医療領域における性能評価手法として広く用いられている。これらのタスクにおいて高い評価を得ているのが Preferred-MedLLM-Qwen-72B である。しかし、本研究において実施した実臨床のカルテ記載を読む試験では、カルテ内容の理解が不十分であることが明らかとなった。

- ⑤ Preferred-MedLLM-Qwen-72B は、Qwen2.5-72B-Instruct をベースに、日本語の医療コーパスを用いて継続的な事前学習を行ったモデルであることが報告されている。しかし本研究の試験では、継続事前学習を行っていない元の Qwen2.5-72B-Instruct の方が、カルテの構造化において高い精度を示している。すなわち日本語による継続事前学習が、むしろ構造化性能の低下を招いていることが明らかになった。臨床現場への LLM 導入にあたっては、医師国家試験や IgakuQA タスクだけでなく、本研究のような実臨床に即した評価試験が重要であることが示された。
- ⑥ tokyotech-llm-Llama-3.3-Swallow-70B-Instruct-v0.4 は、どの量子化ビット数においても安定して高い精度を示しており、特筆すべき点である。

LLM Model Structured Accuracy Comparison with Standard Deviation / LLMモデル別構造化精度の比較（標準偏差付き）



b) 構造化における精度劣化の要因

構造化精度の劣化要因に関する分析も実施しており、その結果を示したものが下図である。図の横軸には LLM のモデル名、縦軸にはカルテ記載項目のカテゴリを配置し、量子化ビット数はラジオボタンにより指定可能となっている。各カテゴリにおいて、LLM が取りこぼした臨床項目数を可視化しており、精度が劣化しているカテゴリは濃い赤色で表示される仕様である。

LLM に共通して見られる精度劣化の要因としては、検査結果、薬剤処方などの治療経過記録、有害事象、処置といった項目が挙げられる。評価は各点において 6 回実施し、「取りこぼした臨床項目数の平均値」を算出している。

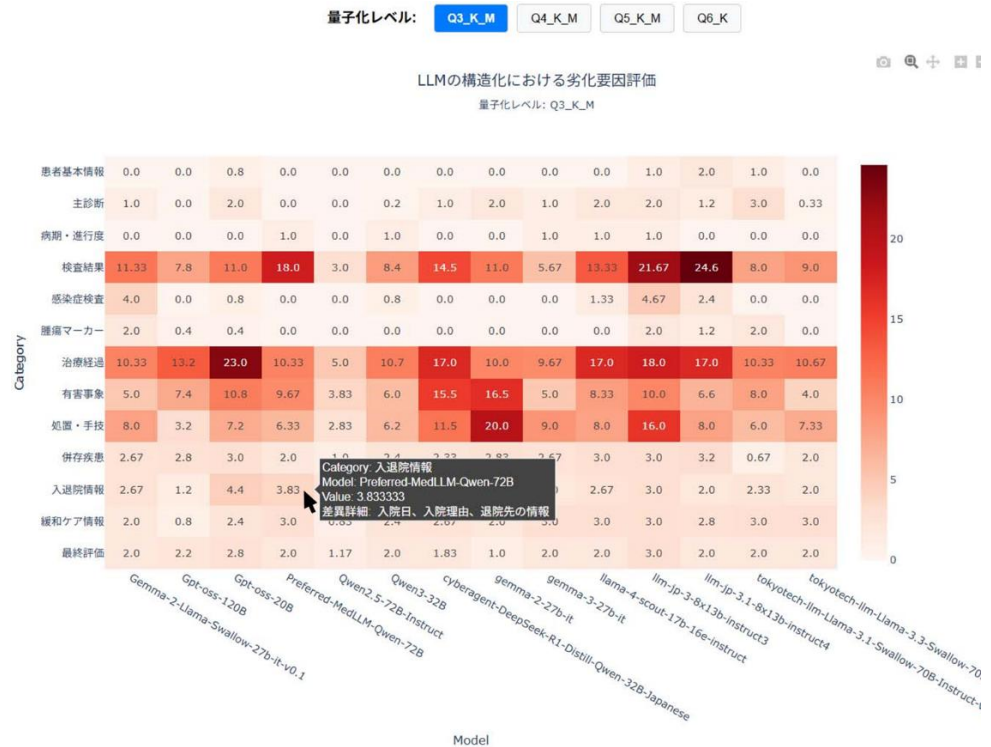
また、「取りこぼした臨床項目の平均値」にマウスカーソルを重ねることで、「具体的な取りこぼし内容」がポップアップ表示されるよう設計されており、図中の黒枠で示された部分がその表示例である。

以下に、本劣化要因分析によって明らかとなった点を列記する。

- ① 日本語による継続事前学習が構造化精度の劣化を招く要因として、薬物治療における治療ライン（一次、二次、三次治療など）の情報生成に失敗する傾向が確認された。この傾向は量子化ビット数には大きく依存しないことも判明している。
- ② カルテの経過記録には治療ラインが明記されていないため、使用されている薬剤が「抗がん剤」であるか「その他の治療薬」であるかを正確に判別する必要がある。加えて、薬物治療の前後関係を整理し、治療全体の構成と流れを把握した上で治療ラインを生成する能力が

求められる。この能力は構造化精度に大きく影響する要因であることが明らかとなっている。

- ③ 検査結果に関しては、感染症、腫瘍マーカー、バイオマーカー等の分類判定に失敗しているケースが多数確認されている。



2) カルテ情報の汎用的な構造化と診療経過の可視化

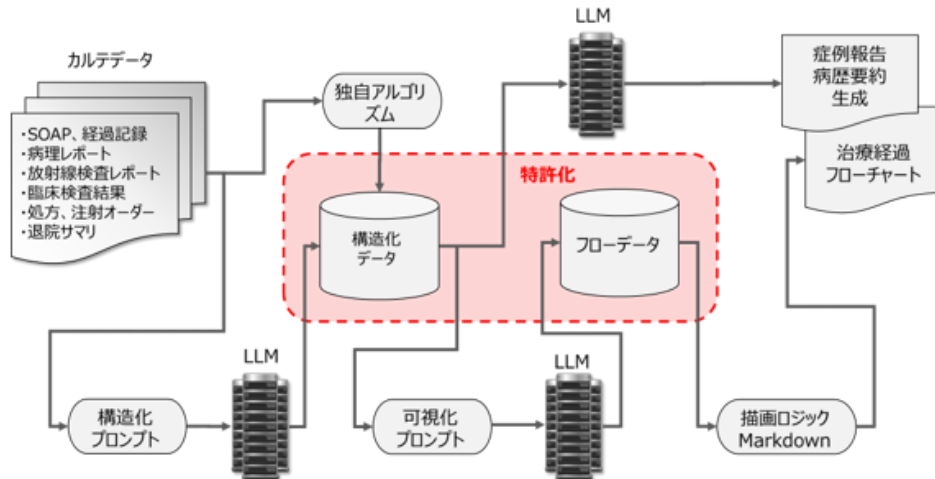
a) 把握すべき重要な臨床項目の抽出:

診療領域毎に把握すべき臨床項目があり、カルテに記載が無くとも部門システム等に情報が保管されている場合がある。そのため独自アルゴリズムやLLMに指示してカルテ記載の無い重要情報を抽出する仕組みを新たに構築している。データフローの概念を下図に示す。本研究におけるデータフローは以下の通りである。

- (ア) カルテデータには、SOAP や経過記録などの非構造化テキスト、および病理・放射線検査レポート、オーダー情報などが含まれている。これらを独自アルゴリズムによって構造化する流れがテーマ①であり、LLM と構造化プロンプトを用いて構造化する流れが本研究テーマ②である。
- (イ) 本研究テーマ②では、独自アルゴリズムによって得られた構造化データと、LLM によって抽出された構造化データを統合し、データベースに格納する。
- (ウ) 汎用的な統合データベースの構築技術の開発も、テーマ②の範疇に含まれる。
- (エ) 統合データベースの情報と LLM を活用して症例報告や病歴要約を生成する流れは、テーマ①の研究に位置づけられる。
- (オ) 本研究テーマ②では、さらに統合データベースの情報をもとに、LLM を用いて診療経過を可視化する技術の開発を進めている。

(カ) 可視化の手法としては、構造化データを LLM により時系列かつ治療カテゴリごとに整理し、フローチャート・データ(マークアップ言語:Mermaid 形式)へと変換する。

生成されたフローチャート・データは、描画ロジックを用いて PDF などに変換し、症例報告への添付資料、患者への説明資料、あるいは若手医師向けの研修資料として活用する。



b) 汎用的な構造化

(ア) LLM を用いた構造化技術とは

以下の文は、カルテに見られる経過記録の記載例である。

#. 胃癌(adeno:tub2,por,sig,HER(-))、腹膜播種、CEA:4.7、CA19-9:7.4、2019/8/5GF、2019/8/19CT で上記の診断。

この記載を JSON 形式で部分的に構造化したものが以下である。

```
"診断情報": {
  "主診断": {
    "診断名": "胃癌",
    "診断日": "2019-08-19",
    "組織型": ["adeno:tub2", "por", "sig"],
    "バイオマーカー": ["HER(-)"]
  },
  "病期・進行度": {
    "ステージ": "",
    "転移部位": ["腹膜播種"]
  },
  "腫瘍マーカー": {
    "CEA": "4.7",
    "CA19-9": "7.4"
  }
},
```


経過記録には、医師が重要と判断した情報が簡潔なメモとして記載されている。これらを臨床情報として構造化するためには、「診断名」「診断日」「組織型」など、データの意味を示すキーを新たに生成し、データを適切にマッピングする機能が LLM に求められる。LLM に対して単に構造化を指示するだけでは、モデルやバージョンによって構造やキーが千差万別となり、統一性のない出力が生成される。そのため、所定の構造に収めるための技術開発は不可欠である。

(イ) LLM を活用する汎用マスタの開発

一般にマスタデータと聞くと、RDB(リレーショナルデータベース)のマスタを想起することが多い。しかし、LLM を活用するためのマスタには、データの意味を示すキーとバリューのリストだけでなく、キーの位置を指定する「構造」も備えている必要がある。ここでは、LLM 用のマスタを JSON 構造に基づいて記述する。

具体的には、前述の JSON 形式の例において、バリューが空欄("")となっているものをマッピング構造とみなし、プロンプト内で LLM に対して指示を行う。また、JSON マスタへのマッピング手順もプロンプトに明記する。たとえば、がん薬物治療の経過記録には、治療ラインが明示されていないケースが多いため、治療の前後関係を整理し、治療全体の構成と流れを把握した上で治療ラインを生成し、それを JSON マスタにマッピングするよう指示する。

(ウ) LLM を活用する汎用マスタの構成とマッピング手法の知財化

カルテに記載されていない情報を LLM に抽出させる際、汎用的な構造を構造化プロンプトで指示する新たな仕組みを実用化した。この汎用的な構造は、データベース化後にさまざまな解析を行う上で非常に重要な要素となるため、LLM を活用する汎用マスタの構成とマッピング手法について特許を出願している。(特願 2025-139345)

c) 診療経過の可視化

(ア) 現行のクリニカルパスの限界と課題

クリニカルパスは、標準化された診療プロセスを提供することで医療の質と安全性を高める手法である。しかしながら、現行のパスは診療ガイドラインに基づいて一律に設計されており、患者の個別性(たとえば併存疾患、年齢、遺伝的背景、生活習慣など)を十分に反映できていない。その結果、現場ではパスからの逸脱が頻繁に発生し、医療者の判断に委ねられる場面が多くなっている。

(イ) LLM による個別化パス生成の可能性

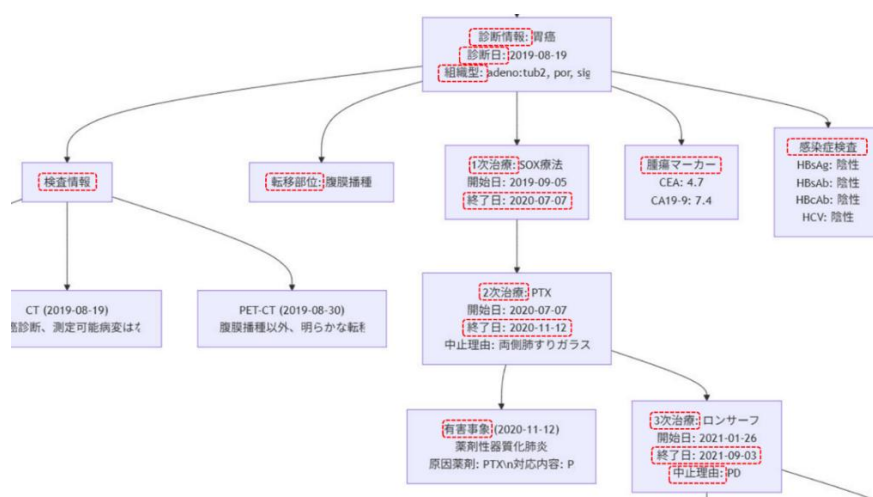
LLM は、非構造化テキストから意味を抽出し、文脈に応じた情報を動的に生成する能力を有している。カルテの記録には、患者の症状、治療反応、医師の判断、検査結果などが時系列で記録されており、これらを LLM が解析することで、個々の患者に最適化された診療パスを生成することが可能となる。このアプローチにより、以下のような利点が期待される：

- ・個別性の反映:患者ごとの特異性を考慮した診療プロセスの提示
- ・パス逸脱の抑制:標準パスとの乖離を減らし、医療の一貫性を確保
- ・意思決定支援:医師の判断を補完する形で、診療方針の選択肢を提示
- ・リアルタイム更新:治療経過に応じてパスを動的に修正・最適化

(ウ) 実現に向けた取り組み

汎用的なデータ構造を生成する仕組みを確立したことで、LLM を活用した診療経過の可視化手法も新たに構築することができた。その可視化の一例を下図に示している。図中の赤い破線枠はカルテには記載されていない情報を表しており、独自のロジックと LLM を用いてカルテから重要な構造化情報を抽出することで、診療経過をより分かりやすく表現できるようになっている。

診療経過の可視化は、患者への治療説明、重要な治療判断ポイントの情報共有、若手医師の教育等でも有用と考えている。



テーマ③「実臨床現場へのソリューション導入」

実際の社会実装に向けた、電子カルテベンダ、ミドルウェア提供事業者各社とのアライアンスへの取り組み。

<2025 年度下期取り組みの前倒し実施>

- ・テーマ①②成果を組み合わせたシステム化、既存サービスとの連携として、NTT プレジジョンメディシン社にて提供する「モバカルホスピタル」へ機能実装したプロトタイプモデルを開発
- ・実証テスト(京都大学病院、がん研有明病院および、「おうちにかえろう。病院」等でのテスト)と並行し、法制度・ガイドライン等の確認を行う
- ・本システムに関するマーケティングリサーチ、特に本システムに対する不満、拒否感等を中心にネガティブ情報等のリサーチも実施

モバカルホスピタルへ機能実装したプロトタイプ版に関しては、「おうちにかえろう。病院」において、実臨床で利用し先生方に評価を実施し、その結果を活かし精度向上に向けたチューニングを実施。

※2025 年 7 月時点では実運用には精度が十分でなかったが、チューニング等を実施し、8 月末

時点では実環境での運用が可能なレベルの精度となっている。

また今後は、入院期間が長期にわたるため電子カルテ記載内容が長文となる場合等において、精度を向上させるためのさらなるチューニングや、機能追加を実施していく。

- ・AI-OCR を用いた、紙、FAX で他医院から受領した紹介状の自動読み取り、電子カルテシステムへの取り組みに関しては、前年度作成した合成データ(5 種の紹介状を実際に FAX により送付し、そのデータをスキャナーでスキャンし読み取り、AI—OCR にて検証を実施)した結果、により検証を実施した結果、クラウド型 AI-OCR では 90%以上の精度で読み取りが可能なことを確認(利用した AI-OCR は市場シェア 1 位であり、NTT グループで取り扱いのある DX Suite を利用)。

この結果を受け、今年度上期において実際のおうちにかえろう。病院にて受け取った紹介状(50 患者分)を用いた事前検証を実施した。結果としては、特定の病院からの紹介状(急性期病院等)に関しては、フォーマットを読み込みチューニングすることで実用できる成果をだせるとが、診療所等多種多様なフォーマットで記載されたものは、フォーマット設定ができないため精度が低くなると想定される結果となった。

特に小規模な個人医院等では下イメージ図1のように、長文で記載され、検査結果や処置なども分内に記載されている事が多く、フォーマットを指定しないフリーフォーマット形式で読み取りを行う必要があり、AI 等認識においても費用が高額となってしまふと想定される。一方で急性期病院や、複数の医師を抱える医院等では、イメージ図2のようなフォーマットに記載されていることが多く、実運用が可能だと想定される結果となった。

今後は、これら事前検証を受け、第一ステップとして、事前にAI等で振り分ける(長文等で書かれたものを除外する)技術を組み合わせ、実環境での検証を実施し、その後LMM等異なる技術を組み合わせることを検討していく。

【イメージ図1】



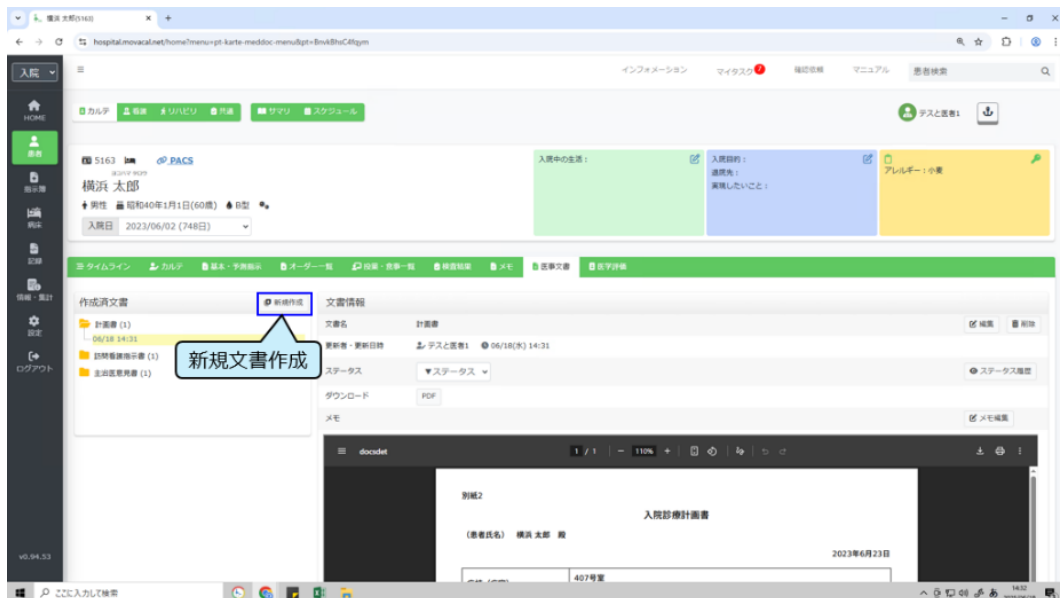
【イメージ図 2】

[illegible]

<アプリケーション画面遷移(2025年7月時点版)2 中規模病院向け >

「おうちにかえろう。病院」にて、電子カルテに実装しているプロトタイプ版画面遷移は以下の通り。
医師がストレスなく利用できるように、シンプルな構成を実現しており、実際に診療現場にて 2025 年 6 月から運用テストを実施中

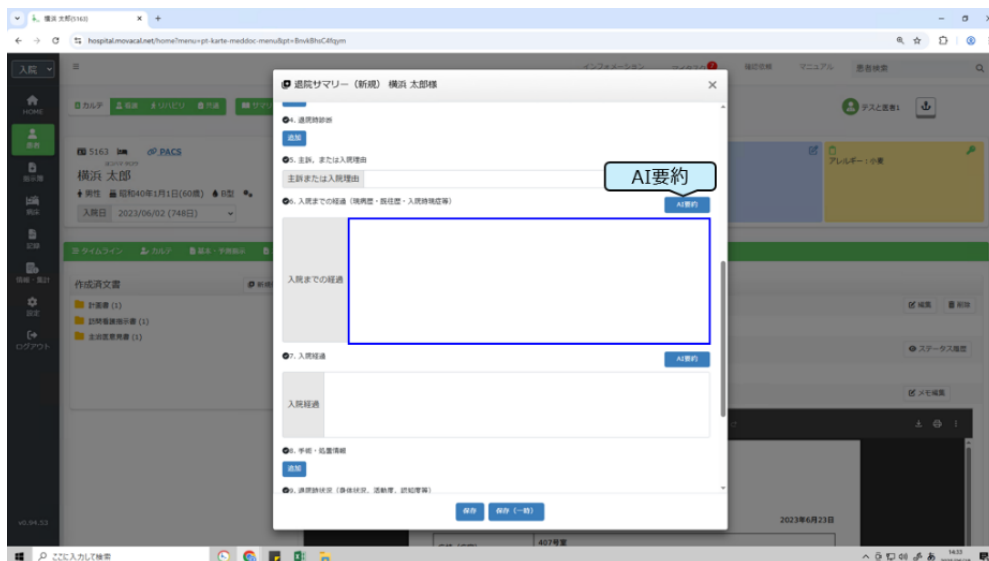
① 退院サマリー作成時に新規文書作成を押す



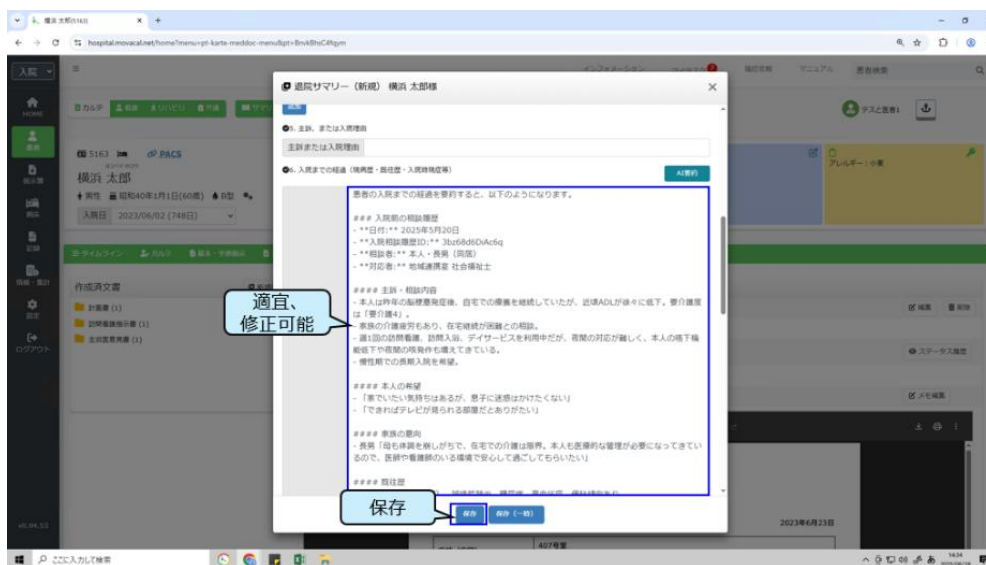
② 新規文書作成一覧から、退院サマリー作成を押す



③ AI 要約ボタンを押す



④ LLM にて自動で退院サマリーを生成、医師にて内容確認し保存を押す



⑤ AI 生成された文書を活用する(PDF 出力、印刷等)



テーマ④「英国等における先駆的取り組みの調査・連携」

- ・英国における開発のための研究計画書の作成を継続し、計画書の完成に向けた作業を実施。
- ・研究計画書の倫理委員会承認前に、合成データを用いたサマリー生成の検証に向けた合成カルテデータの準備。
- ・英国における調査実施内容については、別紙1(巻末)の通り

医師の労働負荷の状況

医療 DX の需要

代表的な技術紹介

これまで英国における先駆的取り組みについて調査を実施してきた結果、英国においては LLM を用いた医療文書の半自動構造化に必要な情報が、日本と比較すると電子カルテに十分に記載されており、LLM 導入などが進みやすいと考えていたが、実際の医療現場においては医師が不足しており、中央の大病院にはスキルの高い医師が十分に在籍するが、地方のかかりつけ医においては、

【経理様式 1 別添】【R7】

人数もスキルも不足している状態となっている。そのため地方のかかりつけ医等においては、逆紹介状等作成における、文書半自動生成技術など医療 DX が求められている状態である。今後研究を継続実施し国際的に活用できるソリューションを開発していき、これら状況改善に対するアプローチを検討していきたい。

<2025 年度下期取り組み>

テーマ①「医療現場支援ツール開発」

開発ツールの臨床評価完了と多施設展開準備を中心に、システム改良・評価・導入準備を集中的に実施しました。京都大学病院での医師参加評価を経て、がん研究会有明病院への展開調整と共同研究準備を進め、次年度以降の二拠点体制での開発基盤を整備した。

・主な活動と成果

ユーザ評価と総合評価取得:10 月以降、京都大学病院での医師参加ユーザレビューを反復実施。複数名の医師からフィードバックを収集し、11 月～12 月にかけて指摘事項を反映した改良を継続。2026 年 3 月に 6 名の医師チームによる最終実証評価試験を完了し、ツールの基本機能(退院サマリー等自動生成)の臨床有用性を確認した。

機能強化と追加開発:放射線検査レポート自動要約アルゴリズムの追加や入院経過表示アプリの調整、LLM ベースのテキスト生成モジュール精度向上を実施。外部モジュール(Gemini 等)の試用評価も行い、生成品質改善に取り組んだ。

多施設展開・連携準備:がん研有明病院向けのデータ連携コネクタ開発、院内要件に合わせたカスタマイズ設計、試験用 PC・ネットワーク機器のセットアップを進めた。院内説明による合意形成と共同研究契約締結準備を実施し、導入計画を整備した。

管理・報告:ステージゲート評価や学会発表準備(医療情報学会向け)を行い、次年度研究計画の再策定および最終報告書の取りまとめ準備を完了した。

テーマ②「データ構造化と可視化」

これまで構造化データは主に二次利用や臨床研究を目的としていたが、診療の質向上や医療従事者間でのコミュニケーション改善のためには、複雑な診療経過を直感的に理解できる形で提示することが重要である。特に、がん診療のような長期にわたる複雑な治療経過においては、時系列での治療内容や有害事象、治療効果の関係性を視覚的に把握できることが、次の治療方針決定や多職種連携において重要な役割を果たす。さらに、構造化精度が完全でない場合の目視修正を容易にする仕組みも臨床実装には不可欠である。下期では、これらの課題について詳細な検証を実施した。

a) LLM を用いた可視化方式

本研究では、LLM によって抽出された構造化データ(JSON 形式)を、診療フローチャートとして可視化する手法の開発に焦点を当てる。

- 1) 構造化データベースから 1 症例分のデータを取り出し、可視化プロンプトを用いて LLM でフローデータ:Mermaid 形式に変換する。
- 2) フローデータ(Mermaid)に変換されたデータをフローデータベースに格納する。
- 3) フローデータベースから 1 症例分のデータを取り出し描画ロジック(Markdown 描画)でフローチャート図に変換する。

4) 生成されたフローチャートの評価は以下の観点で行う:

- ・時系列の正確性: 診療イベントの時間的順序が正しく表現されているか
- ・関連性の明示: 治療と有害事象、検査と診断の関連が適切に配置されているか
- ・可読性: 医療従事者にとって理解しやすい表現となっているか
- ・情報の完全性: 構造化データの重要情報が漏れなく可視化されているか
- ・評価例として、図 11 に示すような治療と有害事象の関連性表現や、LLM が補完した暗黙的情報の明示化などを確認する。

5) Mermaid 形式への変換では、構造化の際に用いたものと同じ LLM:tokyotech-llm-Llama-3.3-Swallow-70B-Instruct-v0.4 を使用する。

b) 経過記録の可視化の検証と評価1

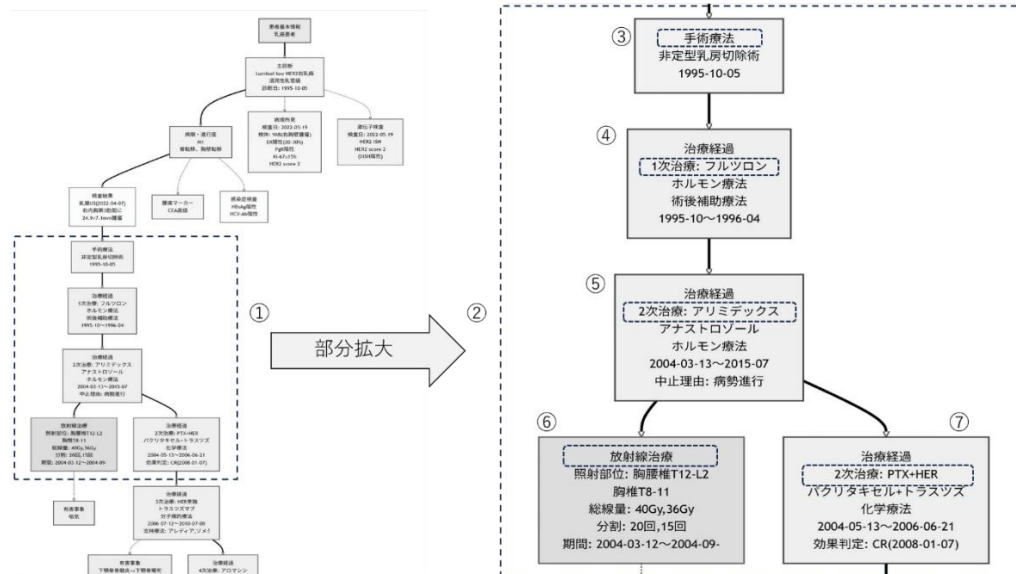
可視化プロンプトを用いて LLM によりフローデータ(Mermaid 形式)に変換した結果の一例を下図(Mermaid)に示す。生成された Mermaid 形式データを MarkdownPreview により描画したものを図(可視化)に示す。対象症例は 8 次治療まで実施された模擬症例であり、前半 4 次治療までを図の左側にフロー図として表示した。図中破線枠①は薬物治療開始時点を示し、その一部を拡大表示したものを破線枠②に示す。

構造化データから Mermaid 形式への変換成功率は 100%(n=20)であった。生成されたフローチャートについて、以下の 3 項目で評価を行った。

- 1)時系列の正確性: 20 症例中 12 症例(60%)で診療イベントの時間的順序が正しく表現された
- 2)関連性の明示: 治療と有害事象の関連が 19 症例(95%)、検査と診断の関連が 19 症例(95%)で適切に配置された
- 3)情報の完全性: 構造化データの主要項目(診断、治療、有害事象)が全症例で漏れなく可視化された。

<pre>graph TD A["患者基本情報
乳癌患者
診断日: 1995-10-05"]:::basic B["主診断
Luminal low HER2右乳癌
浸潤性乳管癌
骨転移、胸壁転移"]:::diagnosis C["病期・進行度
M1
転移部位: L1,Th9-11,L2
左肋骨、右胸壁"]:::stage D1["検査結果
乳腺US
2022-04-07"]:::test D2["右内胸第3肋間
24.9×7.1mm腫瘍
再発疑い"]:::test E1["病理所見
2022-05-19
VAB"]:::pathology E2["ER陽性(20-30%)
PgR陰性
Ki-67≤15%
HER2 score2"]:::pathology F1["遺伝子検査
HER2-ISH
2022-05-19"]:::genetic F2["HER2 score2
DISH陰性"]:::genetic G1["画像検査
MRI
2024-03-22"]:::imaging G2["Th11,L1,L2圧迫骨折
L2圧壊進行
Th11新規骨折"]:::imaging H["腫瘍マーカー
CEA高値"]:::marker I1["1次治療
フルツロン
1995-10~1996-04
術後補助療法"]:::treatment J1["手術療法
非定型乳房切除術
1995-10-05"]:::surgery K1["2次治療
アリミデックス
2004-03~2015-07
11年6ヶ月"]:::treatment L1["放射線治療
胸腰椎T12-L2
胸椎T8-11
2004-03~2004-09"]:::radiation L2["総線量40Gy,36Gy
20回,15回"]:::radiation L3["有害事象
嘔気"]:::ae M1["2次治療
PTX+HER
2004-05~2006-06
2年間"]:::treatment M2["効果判定
2008-01-07
CR"]:::evaluation</pre>	<pre>【 --- 以下省略 --- 】</pre>
--	-----------------------------

図(Mermaid)



図(可視化)

図(可視化)の破線枠②に示すように、LLM は診療録に明示的に記載されていない情報(治療カテゴリ、治療ライン情報等)を医学的知識に基づいて補完し、診療プロセスの明示化を実現している。時間軸は上から下に向かって表現されているが、フロー⑤と⑥については前後関係が逆転している。このように、治療イベントが時間的に近接している場合には、時系列の正確性が低下する傾向が認められた。また、④の 1 次治療は③の手術に付随する術後補助療法であり、⑤の 2 次治療(アリミデックス)は⑥の放射線療に対する放射線補助療法であることがわかる。さらに、⑦の 2 次治療(レジメン:PTX+HER)が本来のがん化学療法における 1 次治療に相当することが、本フロー図から明確に判読可能である。現行の LLM における構造化における誤りは、主として治療歴における複雑な治療ラインの認識領域に集中していることが本検証で確認された。

c) 経過記録の可視化の検証と評価2

経過記録のフローチャート化により治療行為の把握や治療効果、有害事象等を治療スタッフ間での確に共有できる可能性が示された。一方、フローチャートでは時間軸での表現が難しいと言う課題があるため、この課題を補完する手法が求められている。そのため同様に LLM を用いて経過記録をタイムラインに可視化する新たな仕組みを開発した。

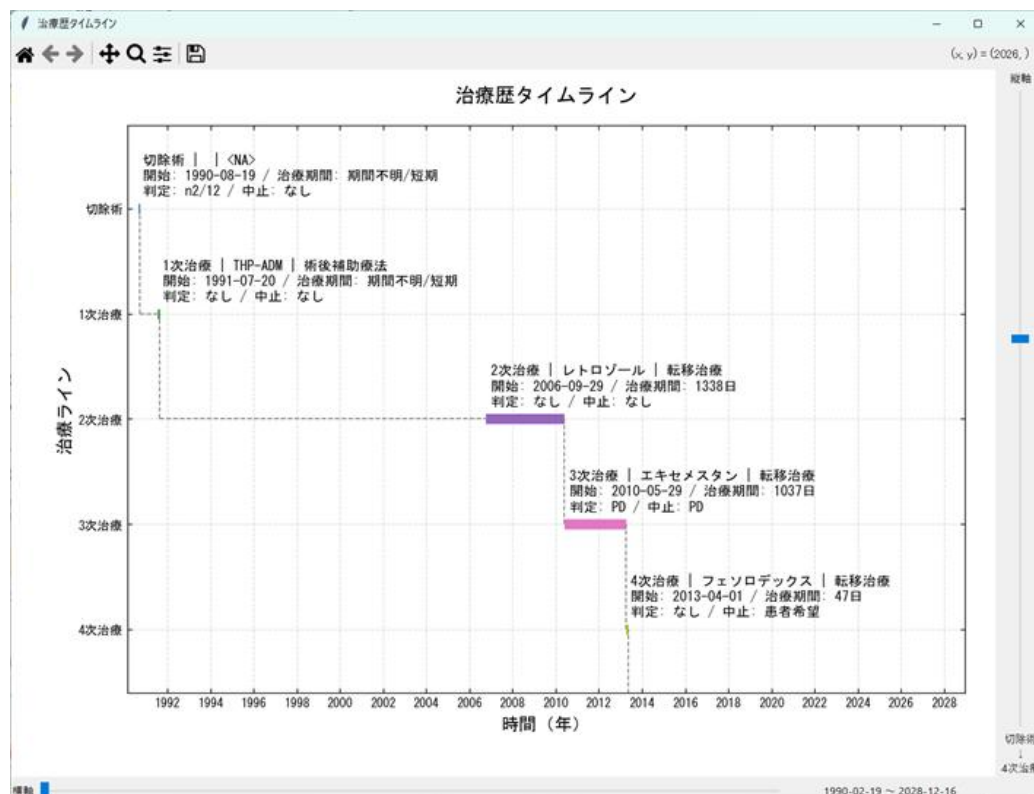
1) LLM を用いたタイムライン可視化方式

以下の手順で経過記録データをタイムライン・チャートに描画する。

- ・データベースのスキーマ(JSON 形式)指定による構造化データの抽出指示
- ・抽出データの時系列処理
- ・治療ライン、治療目的、レジメン名、開始日、治療日数=(終了日-開始日)、判定結果、中止理由項目の整理
- ・時系列データの表出力指示
- ・Python プログラムによるタイムライン・チャート描画

2) タイムラインに可視化の評価

上記の方式でデータベースから抽出したデータを LLM で時系列データに整理した後、Python プログラムで描画したものを図(タイムライン可視化)に示す。図より術後から再発に至る過程や各治療における奏功期間などを的確に把握することが可能であることが示された。



図(タイムライン可視化)

d) 経過記録可視化の考察

経過記録可視化の目的である「複雑な診療経過を直感的に理解できる形での提示」が達成されたことは重要な成果である。現行の LLM では構造化精度が 60~80%程度にとどまり、40~20%の誤りについては目視による修正が必要となる。機械的に JSON 構造化にマッピングされたデータを目視で修正することは実質的に困難であることから、本研究で提案する汎用的な可視化により直感的に誤りを確認できる仕組みは、構造化データベースの運用において不可欠な要素であると考えられる。さらに、LLM が暗黙的な診療プロセスを明示化できることは、癌診療のような長期にわたる複雑な治療経過において、次の治療方針決定や多職種連携の促進に寄与する可能性があり、臨床現場のニーズに直接的に応えるものと考えている。

テーマ③「実臨床現場へのソリューション導入」

下期において、実社会実装を見据えた電子カルテ関連事業者とのアライアンス構築を進め、文書生成・要約を中心とした医療 AI 機能の実用性検証と市場展開を推進した。

1. システム開発・実証

テーマ①②の成果を統合し、NTT プレジジョンメディシンの「モバカルホスピタル」へ機能実装した

プロトタイプを開発し、「おうちにかえろう。病院」で実証テストにおいて、実臨床評価を踏まえたチューニングにより、2025 年 8 月末には中規模病院向け「退院時サマリー半自動生成機能」実運用可能な精度に到達しており、10 月以降 実業務での利用中の状況(長期テストを実施し、3 月時点でバグやハルシネーションに関しても確認できたため、2026 年度商用化を実施していく)

令和 7 年度第4四半期において、中小規模病院向けに看護サマリー、紹介状の半自動生成を実施しており、これら機能も含め商用サービス化を実施していく

3 月末時点で紹介状半自動生成機能に関してはモバカルホスピタルだけでなく、モバカルネットにも実装してく予定であり、4 月以降実際のトライアル参加施設を募集していく。

※令和 8 年 3 月 モバカルホスピタルにおける退院時サマリー半自動生成機能の提供について報道発表実施

2. その他医療 DX に向けた実証(AI-OCR 等による紙の紹介状自動読取)

合成データ検証により、クラウド型 AI-OCR で 90%以上の読取精度を確認(DX Suite 使用)。実際の FAX データ(2 か月分)の検証を行い、90%程度の精度で読み取り可能なことを確認。今後は、異なる AI 技術(LMM 等)との組み合わせによる段階的検証を計画。

※LLM,LMM の発展により AI-OCR を凌駕するソリューションが2~3年で商用化されると考えられるため、既存技術・商材だけでなく、新たな技術に関しても検証を実施していく。

3. 市場展開

大規模・大学病院向け :ベンダ・ミドルウェア事業者とのアライアンスによるトライアル拡大。

中規模病院向け :NTT グループ連携により、電子カルテのオプション機能としてトライアル
社会実装に向けたパートナー開拓推進中
(SIP3 テーマ B2 プレジジョン 佐藤先生と連携スタート)

テーマ④「英国等における先駆的取り組みの調査・連携」

・合成データを用いたサマリー生成の検証に向けた合成カルテデータの準備と、LLM を用い電子カルテ情報をサマライズするデモシステムを構築。

・英国レスター大学における調査研究を推進すべく、大学との契約締結。

NHS 全体では、AI を用いた実証的なパイロット段階から、臨床現場への本格展開フェーズへと急速に移行している。特に、音声認識を活用したアンビエント AI による診療記録支援、臨床意思決定支援、業務自動化等は、慢性的な人員不足や医療従事者のバーンアウトといった課題への有効な対策として期待されている。

既存の研究プロトコルは、単一ツール・単一診療科に限定した評価設計であり、実運用・全体最適の観点では十分な示唆を提供できない状況であり、以下のような構造的課題が確認された。

- ・特定ユースケースに限定され、他部門・他ツールへの汎用性が乏しい
- ・技術的精度や作業時間短縮に重点が置かれ、職員体験・定着度・安全性の評価が不十分
- ・導入障壁や活用を促進する要因に対する定性的知見が限定的

これらを踏まえ、研究の Scope を AI 利活用全般を対象とする検証型スタディとし、本取組は研究倫理審査を伴う臨床研究ではなく、AI Service Evaluation(サービス評価)として位置づけることとし、以下の観点などから多面的な評価を実施する。

- ・定性調査 : 医師・看護師・事務職・デジタル／ガバナンス担当者への個別インタビュー・フォーカスグループインタビュー
- ・サーベイ調査 : ユーザビリティ、文書負荷、認知負荷、継続利用意向等の定量的把握
- ・利用・定着分析 : 利用頻度、利用中断等のログ分析
- ・安全・ガバナンス : セキュリティおよび関連する規制への対応状況

3. 成果物の公表(公開)

3.1 論文など(原著論文、学位論文、プロシーディングス、総説、解説、速報など)

【報道発表】

- ◇報道発表実績:2025 年 6 月 研究実施施設「おうちに帰ろう病院」との取り組みに関する報道発表
- ◇報道発表実績:2025 年 8 月 SIP3 特設ホームページの開設
- ◇報道発表予定:2025 年 3月 研究実施施設「おうちに帰ろう病院」との取り組みに成果及びサービスリリース発表

【論文】

- [1] 吉原 博幸, 岡崎 直観, 横田 理央, 加藤 康之, 鈴木 利明, 「電子カルテ由来臨床データの知識ベース化:ナラティブデータから LLM を用いて臨床データを構造化する」, 第 45 回医療情報学連合大会, 公募パネルディスカッション 2, 2-D-1, 2025(Proceedings).
- [2] 加藤 康之, 小田 哲, 石田 茂樹, 横田 理央, 吉原 博幸, 松本 繁巳, 武藤 学, 持田 剛, 岡田 昌史, LLM を用いた経過記録情報のユーザ開放型解析環境, 第 30 回日本医療情報学会春季学術大会, 2026.
- [3] 医療情報学会の原著論文(査読中)
加藤 康之、小田 哲、荒木 綾乃、持田 剛、岡田 昌史、湯川 洋一郎、石田 茂樹、横田 理央、岡崎 直観、吉原 博幸、松本 繁巳、武藤 学、
電子カルテの臨床研究活用を加速する:大規模言語モデルによる経過記録構造化手法の開発と評価.

論文数(総数)	(内国際誌)	(内国内誌)
3		

3.2 学会発表など((国内・国際)学会口頭発表・ポスター発表、招待講演で成果を公表したもの)

- ①吉原、岡崎、横田、加藤、鈴木「電子カルテ由来臨床データの知識ベース化 ーナラティブデータから LLM を用いて臨床データを構造化するー」
第 45 回医療情報学連合大会(第 26 回日本医療情報学会学術大会)、公募企画パネルディスカッション、2025 年 11 月 13 日

- ②2025 年 6 月日本在宅医療連合学会大会 ランチョンセミナーにおいて、本取り組みについて紹介
- ③2025 年 7 月第 79 回日本食道学会学術集会スポンサードセミナーにおいて本取り組み紹介

3.3 診療ガイドライン、省令、基準、日本薬局方、添付文書改訂、国の技術文書(通知)等への反映
該当無し

3.4 研修プログラム、カリキュラム、シラバス、教材、e-learning 等の公表
該当無し

3.5「国民との科学・技術対話」に対する取り組み
該当無し

3.6 その他
該当無し

以上