

# 専門職主導による治験関連文書構造化モデルの設計:生成AIを設計支援として用いたアプリ開発

林 和華子<sup>1,2)</sup>, 深瀬 恭子<sup>2,3)</sup>, 伊藤 陽一<sup>2,4)</sup>

1) 東北大学病院臨床研究推進センター, 2) 北海道大学 大学院医学院 社会医学講座 ヘルスデータサイエンス教室,  
3) 札幌医科大学 医療統計・データ管理学講座, 4) 北海道大学病院 医療・ヘルスサイエンス研究開発機構 プロモーションユニット データサイエンスセンター

※本発表は発表者個人の試行的な実践および見解に基づくものであり、所属機関のDX方針や生成AI活用方針を代表するものではありません  
本発表にあたり開示すべきCOIはありません。



北海道大学  
HOKKAIDO UNIVERSITY

## 背景・目的

### 【背景】

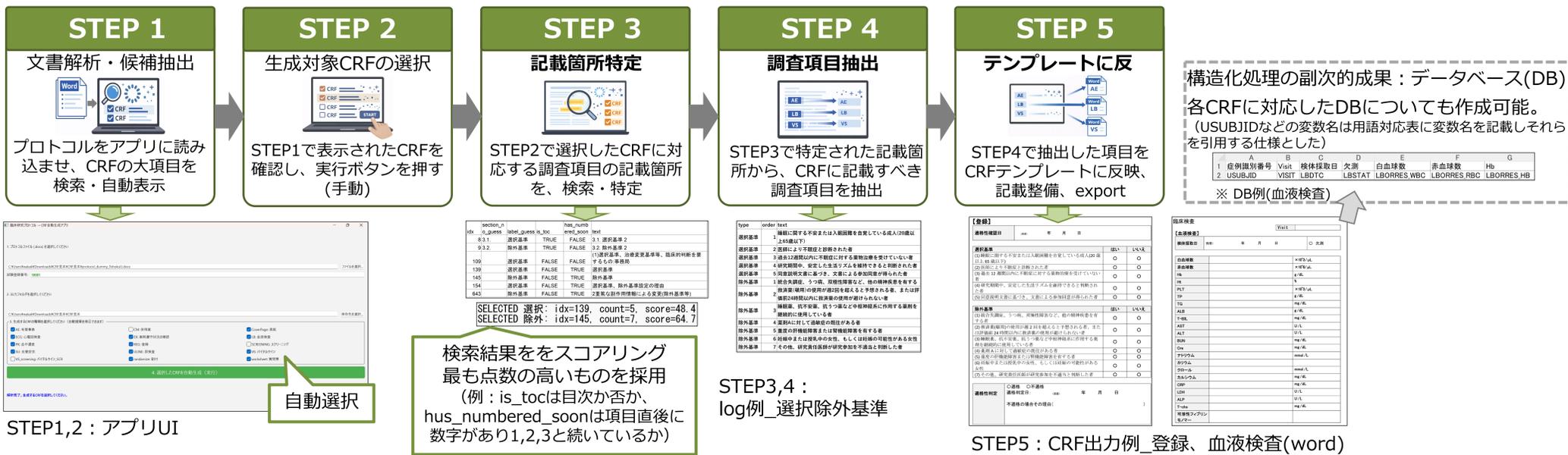
- 治験関連文書作成は専門家の暗黙知に依存しており、情報処理の構造化が十分に検討されていない点が課題である。
- 生成AIは非構造化情報の処理に有効な手段として注目されているが、機密性や倫理的配慮の観点から、治験・臨床研究分野の文書作成では生成AI等の活用が制限されている。
- 治験・臨床研究分野においては、生成AIへの依存を前提とせず、専門職が自ら業務工程を構造化し、自動化可能な処理単位として整理する方法論の検討が求められる。

### 【目的】

- 専門職が自ら業務工程を構造化し、生成AIを開発支援として用いながら文書作成業務の自動化モデルを設計できることを示す。

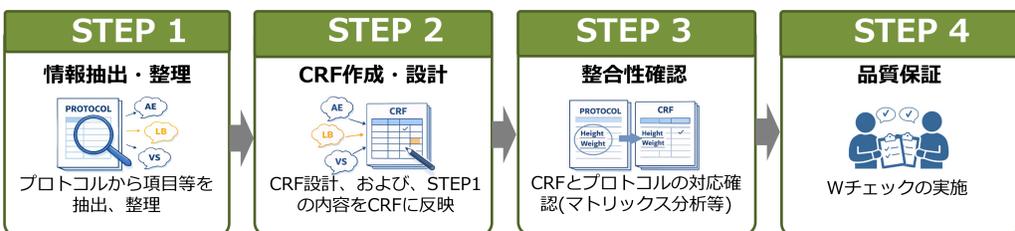
## アプリ動作の概要

1) CRF作成アプリの動作フロー : CRF作成工程を処理単位に分解し、各単位ごとにロジックとして実装した



## アプリ開発の概要

1) CRF作成工程の構造化と自動化対象の定義



★自動化対象: 構造化抽出・マッピング処理

### 2) システム構成

入力: プロトコル

処理: プロトコルから項目の検索、該当箇所の特定、項目の抽出、テンプレートへ反映

出力: CRF (Word形式)

実装環境: Python + PySide6 (GUI) ※アプリ実行時はAI・API・ローカルLLMを使用しない設計

3) 開発支援、コード作成における生成AI※の活用方法(※ChatGPT-5 有料版)

- CRF作成業務を処理が成立する最小単位に分解し、各工程の入出力仕様は筆者が定義。
- 定義した仕様に基づき、工程ごとに生成AIへコード作成を依頼した。
- 生成されたコードを実装・試行し、期待する挙動にならない場合は追加条件を提示して再依頼、または工程定義自体を見直すなど反復的調整を行った。
- 最終的な処理構造および工程分解の妥当性確認は筆者が行った。

## 結果・限界

### 【実装結果】

- 専門職が自ら業務工程を処理単位に分解し、アプリに外部API等を用いない状態で、プロトコルからCRFを生成する構造化モデルを設計・実装できることを示した。
- 本モデルにより、文書作成業務のうち一定の構造化可能な工程について自動化が可能であることが確認された。

### 【本研究の限界】

- 外部API等を使用しない設計としたため、高度な自然言語処理には限界があった。そのため、用語対応表等の補助定義を用いて処理の安定化を図った。
- また、本研究は業務工程の構造化可能性および実現可能性の検証を目的としたものであり、DXの運用定着や組織的変革そのものを評価したものではない。現場運用における包括的評価は今後の課題である。

## 開発対象

### 【開発対象】

- 治験関連文書作成業務に内在する処理構造を明示化し、自動作成を行うアプリケーションを設計する。
- 本研究では、プロトコルから症例報告書(以下CRF)を生成するデスクトップアプリケーション(以下アプリ)の開発を行う。

### 【プロトコルからCRF自動作成を選定した根拠】

- 筆者が専門職として文書作成業務の工程の分解及び、完成品を評価できるため。
- CRF作成は単なる転記作業ではなく、プロトコル記載内容の抽出・解釈・構造化を伴う専門的処理であり、外部API等を用いず、設計可能であると想定されたため。
- 本対象を通して、同様の処理構造を持つ他の治験関連文書への応用可能性を示す。

4) 外部APIを用いない設計におけるアプリ構成と外部資料の役割

- 外部APIを用いない環境での運用を前提とし、処理ロジックとは独立した外部定義資料を用いる構成とした。
- 副次的効果として、様式変更や用語追加が生じた場合でも、外部ファイルの修正のみで更新可能となる。

- CRFテンプレート (出力形式および項目構造を定義)
- 用語対応表 (プロトコル等における表記ゆれを正規化し、処理を安定化させる対応表)
- プロトコル (本アプリの構成要素ではなく、実行時に読み込む入力文書)

①CRFテンプレート 例\_登録、血液検査

②用語対応表(一部抜粋)

③今回使用したダミープロトコル(一部抜粋)

## 考察・今後の展望

### 【考察・今後の展望】

- 複数試験への再利用を前提とし、特定の試験構造に依存せず、プロトコル記載内容を探索・抽出する方式を採用した。これにより、試験ごとに処理構造を再設計することなく適用可能な構成とした。結果として、継続的な使用が可能となり、文書作成業務の効率化につながる可能性が示唆される。
- 本アプリの開発を通して、元となる文書(本研究ではプロトコル)から「該当項目の検索・特定」、「項目の抽出」、「テンプレートへの反映」といった一連の処理が構造化可能であることが示された。
- これらの工程は、専門職が日常業務において暗黙的に判断・整理してきたものであり、本研究で提示した構造化・自動化手法は、CRFに限らず総括報告書や各種研究文書など、他の医療文書作成業務にも応用可能であると考える。