

大規模言語モデルの医療現場での 活用と展望

東京大学 大学院医学系研究科
疾患生命工学センター 医工情報学部門
今井 健

自己紹介 (今井 健)



- 1999年 東京大学理学部数学科卒
- 2005年 東京大学大学院情報学環 学際情報学府
博士課程修了, 博士 (学際情報学)
- 2005年 東京大学医学部附属病院 企画情報運営部 特任助手
- 2008年 (兼任) 米国 Mayo Clinic 客員研究員
- 2014年 同 大学院医学系研究科 疾患生命工学センター
医工情報研究領域 講師
- 2016年 同 大学院医学系研究科 疾患生命工学センター
医工情報学部門 准教授
- 2016年 (兼任) 自治医科大学 データサイエンスセンター
特別特命教授

- 日本医療情報学会 理事, 学術委員長, 2026春季学術大会大会長
- WHO-FIC ITC Secretariat
- ISO/IEC JTC1/SC42 (AI) JWG3 (医用人工知能) Co-Editor, 国内主査
- ISO TC215 (医療情報) WG3 | SC1 | TF5 Expert

Agenda

- **大規模言語モデルの医療現場での活用と展望**

SIP3 生成AI での成果から、

①医療従事者の業務負担軽減、

②標準化・保健行政支援 の2つの観点で事例を紹介

- **医療従事者の業務負担軽減の観点から**

- SIP3生成AI で取り組んだこと
- 感染症発生届の下書き
- 医療DXの流れから「Patient Summary 用療養上のアドバイス」作成支援

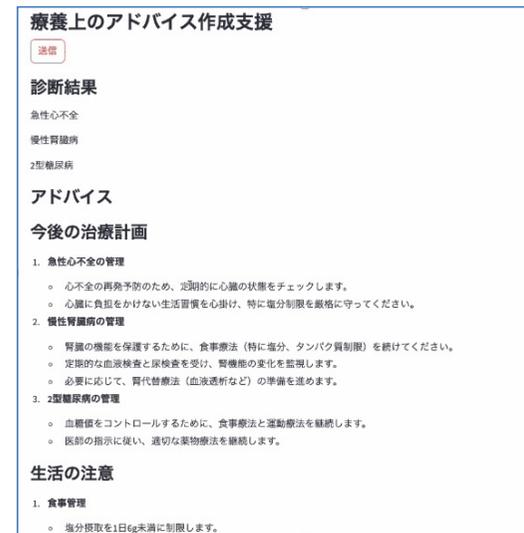
- **標準化・保健行政支援の観点から**

- 標準化の必要性と標準コード
- ICDとICF, ICD10からICD11へ
- 電子カルテ文書中の自由記載からのICD11/ICFコーディング
- 死亡票外因コーディングシステムへの活用

- **医療AI領域の国際標準化動向**

医療におけるLLM/LMMの活用例

- 診療記録・サマリー・各種報告書自動作成**
 - 経過記録、退院サマリー、看護引き継ぎ書の下書き
 - 画像から放射線読影報告書、病理レポートの下書き
- 質問応答/医学知識検索**
 - 診療ガイドライン、薬物相互作用の検索、臨床スコア評価
 - 病理画像に対する診断上の質問応答
- トリアージ・チャットボット**
 - 重症度選別、教育、遠隔モニタ
- 診療情報コーディング**
 - ICD/DPC/ICF等各種標準コードの自動決定
- 文献レビュー支援**
 - 論文スクリーニング
- 診療動画の解析**
 - 内視鏡/手術動画のリアルタイム字幕提示・手技認識



国内の生成AIモデルと課題

- **生成AIを用いたサービス開発**
 - 既に多くの国内企業が着手している
 - GPT, Gemini 等海外のモデルを使用し、サービス・アプリケーション開発に注力する手もある。
 - **手数料を今後ずっと払い続ける必要がある**
 - **機微な診療データを院外(海外企業)に出したくない**

- **国産のLLM/LMM の開発の必要性** (cf 各国のソブリンLLM開発の流れ)
 - **医療に特化したモデルをどのように開発するか？**
 - 2024年 内閣府 SIP (第3期)「統合型ヘルスケアシステムの構築における生成AIの活用」が開始
 - **最新の医療分野の知見**をどう適時組み込むか？
 - **質・安全性**をどのように担保するか？
 - **個人情報保護**の担保は絶対条件
 - **海外モデルと比較した独自の強み**を出せるか？ (cf. Med-PaLM2 by Google)
 - 海外モデルに翻訳の皮を被せる vs 国産モデル
 - 学習するためのデータ自体の質の高さ

- **医学医療分野での 生成AI等への“向き合い方” 議論の醸成**
 - 倫理的・法的側面からのさらなる議論の進展が必要
 - どんどんサービスは展開されていく

2024～内閣府 SIP での医療LLM/LMM 開発研究がスタート

テーマ1: 相澤先生(NII) チーム「」



- 国立情報学研究所が2000名を超えるコミュニティを中心に開発する LLM-JPシリーズをベース
- 日本最大の医療系コーパス442億トークンで追加事前学習、汎用ファインチューニングとユースケース特化型ファインチューニングを経て各種ソリューション化
- ユースケースによってはGPT-4を超える性能

ベースモデル	追加事前学習	ファインチューニング	学習済みモデル
LLM-jp-3-13b	汎用 Webク	医療系 コーパス	汎用インスト
	ロール	V2	ラクション
LLM-jp-172b	データ	79.6B	+ 医療系イン
	300B	79.6B	ストラクショ
LLM-jp-3-8x13b	Tokens	Tokens	ン
			SIP-jmed-llm-2-13b
			SIP-jmed-llm-2-172b
			SIP-jmed-llm-2-8x13b (MOEモデル)

①医療従事者の業務負担軽減の観点から

～サマリー・レポート下書きは
業務軽減の観点から非常に期待が高い～

これまでの国内開発事例 ～サマリ文書の下書き支援～

NEC、生成AIを搭載した電子カルテシステム「MegaOak/iS」の販売を開始

～医療文書の作成を支援し、医師の働き方改革に貢献～

2024年3月18日
日本電気株式会社

News Room >

経営戦略／業績／人事 >

サステナビリティ >

研究開発／新技術 >

サービス／ソリューション >

ハードウェア >

事例 >

イベント・セミナー >

発表年月で探す

年月を選択

キーワードで探す

キーワードを入力してください

NECは、「医師の働き方改革」への貢献に向けて、医療文書の作成を支援する生成AIを搭載した電子カルテシステム「MegaOak/iS(メガオーク アイエス)」の販売を、本年4月より開始します。生成AIを搭載した電子カルテシステムの販売は国内初(注1)となります。

今回新たに搭載する医療機関向け生成AI「MegaOak/iS AIメディカルアシスト」(注2)は、生成AIの1つである大規模言語モデル(Large Language Model : LLM)を用いています。第一弾として、電子カルテに記載の診療情報をもとに、診療情報提供書(紹介状)と退院サマリ(注3)に活用できる文章案を自動生成します。また今後、クラウドセキュア接続サービス(注4)を介して、「MegaOak/iS」以外の電子カルテシステムを導入している医療機関でも利用可能とする予定です。



例:退院サマリ

https://jpn.nec.com/press/202403/20240318_01.html より

患者サマリー用“療養上のアドバイス”作成支援

患者サマリー（療養計画書）イメージ

医療機関名	Aクリニック
医師氏名	厚生 太郎
主傷病名	
胃の悪性新生物<腫瘍>	▼
胃体部癌	
副傷病名	
潰瘍性大腸炎	▼
潰瘍性大腸炎性関節炎	
療養上の計画・アドバイス	
<ul style="list-style-type: none"> ・内服を継続しましょう。 ・1日〇分、〇〇程度の運動を行いましょう。 ・〇ヶ月ごとに血液検査を予定しています。 ・〇〇の福祉サービスの利用を検討しましょう。 ・〇〇の疾患について、診療所Aを受診してください。 	



電子カルテ情報共有サービス

(厚労省HPより)

「全国医療情報プラットフォーム」の仕組みの一つである、電子カルテ情報共有サービスは、全国の医療機関や薬局などで患者の電子カルテ情報を共有するための仕組みです。提供するサービスは次の4点です。

1. 診療情報提供書を電子で共有できるサービス。(退院時サマリーについては診療情報提供書に添付)
2. 各種健診結果を医療保険者及び全国の医療機関等や本人等が閲覧できるサービス。
3. 患者の6情報※を全国の医療機関等や本人等が閲覧できるサービス。
4. 患者サマリーを本人が閲覧できるサービス。

- ・国が進める“電子カルテ情報共有サービス”
- ・患者が自分の健康医療情報を確認できる時代に
- ・患者サマリー用の情報を作成する手間を削減
- ・SIP-JMedLLM 2 (MOE: 13B x 8) にて大幅に質向上

カルテテキストの一部と疾患説明文 (RAG用, 医師監修, 600疾患) の両方を与えて“療養上の計画・アドバイス”を生成

患者サマリー用“療養上のアドバイス”作成支援

例（急性心不全 + 慢性腎臓病 + 2型糖尿病患者）

指示文：

以下の患者カルテと疾患の説明文から、治療計画と内容、生活での注意についてこの患者に渡すアドバイスだけを箇条書きでわかりやすく短く書いてください。・・・

入力カルテ文章：

基本情報(入退院日以外)、主病名、病歴、入院後経過、退院時処方のセクションを利用

疾患説明文：

RAG 用に使用：右の例では「心不全、慢性腎臓病、2型糖尿病」の3つの解説文章を使用

療養上のアドバイス作成支援

テキストを入力してください

【基本情報】
 入院日：2019/10/24
 退院日：2019/10/31
 転帰：治癒
 フォローアップ：外来で

【主病名】
 急性心不全

確定診断名
 #1. 急性心不全
 #2. 慢性腎臓病

送信

**医師監修アドバイス
 +
 患者固有の内容も反映**

感染症発生届の下書き支援

自由入力病名を抽出

(基盤技術を利用)

疾患オントロジーに問い合わせ
該当疾患かチェック

該当疾患なら

感染症発生届下書き生成

別記様式 10

梅毒発生届

都道府県知事（保健所設置市長・特別区長） 殿

感染症の予防及び感染症の患者に対する医療に関する法律第12条第1項（同条第8項において準用する場合を含む。）の規定により、以下のとおり届け出る。

報告年月日 令和 年 月 日

医師の氏名 _____

従事する病院・診療所の名称 _____
上記病院・診療所の所在地(※) _____
電話番号(※) () - _____
(※病院・診療所に従事していない医師にあっては、その住所・電話番号を記載)

1 診断(検査)した者(死体)の種類
・患者(確定例) ・無症状病原体保有者 ・感染症死亡者の死体
2 性別
男 ・ 女
3 診断時の年齢(0歳は月齢)
歳 (月)
病 型
① 病型 1) 早期顕症梅毒(Ⅰ期Ⅰ、Ⅱ期) 2) 晩期顕症梅毒、 3) 先天梅毒、4) 無症候(無症状病原体保有者)
② HIV感染症合併の有無 1) 有 2) 無 3) 不明
4 症 状
・初期硬結(性器、肛門、口唇、口腔咽頭、その他) ・硬性下疳(性器、肛門、口唇、口腔咽頭、その他) ・鼠径リンパ節腫脹(無痛性) ・梅毒性バラ疹 ・丘疹性梅毒疹 ・扁平コンジローマ ・ゴム腫 ・心血管症状 ・神経症状 ・眼症状 ・骨軟骨炎 ・実質性角膜炎 ・感音性難聴 ・Hutchinson 歯 ・その他() ・なし
5 診 断 方 法
① 患者(確定例)の場合 ・病変からの病原体の検出(染色法、PCR検査) ・次の1)、2)の両方の抗体検査による血清抗体の検出 1) カルジオリピンを抗原とする検査 2) <i>T. pallidum</i> を抗原とする検査 ② 無症状病原体保有者の場合 ・次の1)、2)の両方の抗体検査による血清抗体の検出 1) カルジオリピンを抗原とする検査 (抗体価を記載、16倍相当以上が必要)

下書き対象

1 1 感染原因・感染経路・感染地域
① 感染原因・感染経路(確定・推定)
1 性的接触 (A. 性交 B. 経口) (ア. 同性間 イ. 異性間 ウ. 不明) (性風俗産業の従事歴(直近6か月以内)) 1) 有 2) 無 3) 不明 (性風俗産業の利用歴(直近6か月以内)) 1) 有 2) 無 3) 不明
2 静注薬物使用
3 母子感染(ア. 胎内・出産時 イ. 母乳)
4 輸血・血液製剤 (輸血・血液製剤の種類・使用年月・状況:)
5 その他()
6 不明
② 感染地域(確定・推定)
1 日本国内(都道府県 市区町村)
2 国外 (国名:)
詳細地域:)
3 不明

感染症発生届下書き支援

テキストを入力してください

【基本情報】
年齢: 73歳
性別: 男性
入院日: 2018/09/21
退院日: 2018/10/14
転帰: 軽快
フォローアップ: 外来で

【主病名】
梅毒感染症

確定診断名
#1. 梅毒感染症
#2. ネフローゼ症候群
#3. 急性腎障害

【病歴】
主訴: 下腿浮腫、全身倦怠感
現病歴: 2018年7月頃に名古屋の風俗店で性行為を行った。同年8月半ば頃から下腿浮腫が出現し、9月に腎臓の発赤・しこりが増悪してきたが、排尿時の痛みや灼熱感はなかった。9月18日に両側下腿浮腫を初めて自覚し、翌日より悪化した。9月20日に血清Alb低下、蛋白尿、肝障害を認め、

7日以内に提出

生成開始

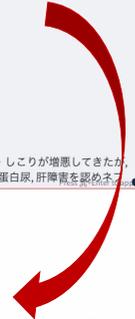
早期劇症梅毒Ⅱ型

**初期硬結、梅毒性バラ疹、
カルテ中の自由入力病名の梅毒疹、鼠径リンパ節腫脹**

**該当疾患である場合、発生届の必要項目と
性的感染(性交・異性間)**

カルテテキストの部分情報の両方をLLMに与え

“感染症発生届”の下書きを作成



②標準化・保健行政支援 の観点から

～診療情報の標準コード化は
医療情報標準化、AI開発の必須の基盤～

カルテ文中の診療情報から標準コードへの自動変換

自由入力病名・症状・所見・部位

- ICD10/11コード
- 診療報酬請求、各種統計、研究等で広く利用
- ICD11になって、より詳細な病態が記述可能に。コーディング負荷は極めて高い
- 標準病名マスター (ICD10対応) だけでは自由入力病名・フレーズ表現に対応できない

“右大腿骨転子部骨折”



NC723Z & XK9K (ICD11)
S7210 (ICD10)

生活機能情報

- ICFコード
- 病院外 (在宅介護・リハ) での予後 (患者状態) を把握

“端座位に介助を要する”



d4153.2+ (ICF)

アレルギー情報

- J-FAGYコード
- コード化されていれば薬剤アレルギー等のチェックを自動で行え、医療事故を未然に防止できる

“甲殻類アレルギーあり”



J9FC15000000 (J-FAGY)

ICD11時代のPost-Coordination

- (例) 右季肋部痛
 - 成句では登録なし
 - Post-Coordinationにてコードされることが前提
 - XK9K (右)
&
MD81.10 (上腹部痛)

さらに 季肋部 は部位の中で
“上腹部” の下位語 (XA3TD4)

ICD-11 (Foundation)

Search hypochondrium [Advanced Search]

Foundation URI : <http://id.who.int/>

Hypochondrium

Parent(s)

- Abdominal wall
- Upper abdomen

季肋部

解剖構造ツリー

▼ Extension Codes

- ▶ Severity Scale Value
- ▶ Temporality
- ▶ Aetiology
- ▶ Topology Scale Value

▼ Anatomy and topography

- ▶ Functional anatomy
- ▼ Surface topography
 - ▶ Head and neck
 - ▼ Trunk
 - ▶ Upper trunk
 - ▼ Lower trunk
 - ▼ Abdomen
 - ▼ Upper abdomen
 - Epigastrium
 - Hypochondrium**
 - Periumbilical region
 - Lateral lumbar region

Post-Coordination: XK9K & MD81.10 & XA3TD4 とする必要あり

コード化対象は単なる単語ではない

例) 生活機能評価 と ICF Coding

【診断名】 誤嚥性肺炎 #11 【併存疾患】 慢性閉塞性肺疾患 #11
 【病歴】 xx年y月z日 息切れを主訴に前医を受診。慢性閉塞性肺疾患の診断で入院し加療されたが、入院中に誤嚥性肺炎を繰り返し、徐々にADL低下を認めた。嚥下評価、呼吸リハビリおよびADLの改善を目的に、z+60日当院に転院となった。
 【社会的背景】 無職。両親と同居。病前はADL修正自立レベル。
 【現症・評価】 血圧 130/70mmHg、脈拍 68分、SpO₂ 96% (O₂ 5l投与下)。意識清明、心音整、心雑音なし。右下肺にcoarse crackleを聴取。腹部平坦かつ軟、圧痛なし、下肢浮腫なし。RSST 2回/30秒、喉頭挙上遅延。筋力上下肢それぞれMMT4レベル、握力右9kg、左6kg、筋緊張正常、感覚正常、基本動作：端座位保持は軽介助、起立・立位保持は中等度介助レベル。MMSE 22/30。ADL：FIM 運動項目 23/91、FIM認知項目 30/35。
 【検査】 胸部レントゲン写真：右下肺の透過性低下(+)。血液ガス(room air)：pH 7.45、pO₂ 66.2 mmHg、pCO₂ 45.1 mmHg、HCO₃ 30.5 mmol/L、BE 5.8mmol/L。嚥下造影検査：反射惹起遅延(+)。液体で少量の誤嚥あり、ペースト食で少量の残留あり。咳反射あるが微弱

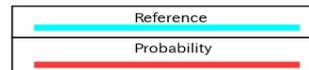
心身機能	活動と参加	環境因子
#11 呼吸機能低下(中等度) b440.3	#21 座位保持(軽度) d4153.1	#31 家族 e310.+2
#12 嚥下機能低下(中等度) b5105.2	#22 立位保持(中等度) d4154.2	#32 酸素投与 e1108.+2
#13 筋力低下(軽度) b730.1	#23 セルフケア(重度) d5.3	
#14 認知機能低下(軽度) b117.1		
	(潜在的な問題)	
	#24 仕事 d850	

ICD10/11/ICF/J-FAGYコード化システム

Step1) 対象情報の抽出 (F1 > 0.96)

【主病名】 薬剤性腎障害 【確定診断名】 #1. 薬剤性腎障害 #2. 肺小細胞癌 #3. 発熱性好中球減少症 【主訴】 右大腿の腫れ、体重減少 【現病歴】 2023年6月頃から1ヶ月で2kgの体重減少を認め、8月より右側の顔面が浮腫むようになった。8月11日に前医を受診し胸部X線で胸部異常陰影を指摘され、8月14日に当院を紹介受診した。胸部CTで右肺上葉に胸壁に直接浸潤する8.0mm大の腫瘍影と複数の縦隔リンパ節腫大を認めた。気管支鏡下肺生検を含めた全身精査により、小細胞肺癌 cT4N3M0 Stage III Cと診断した。1次治療としてアテゾリズマブ+カルボプラチン (CBDCA) +エトポシド (VP-16) を導入する方針とし、9月5日に入院となった。【既往歴】 高血圧症 アレルギー歴：なし 常用薬：アムロジピン10mg 【主な入院時現病】 身長 162 cm、体重 46 kg、血圧 128/66 mmHg、体温 36.8 °C、脈拍 88 回/分・整 SpO2：94% (室内気) 【意識】 清明 【聴診】 呼吸音 肺野に異常はない、呼吸音 右肺野全体で減弱している 【触診】 平坦・軟で圧痛はない、腸蠕動音異常はない 【四肢】 浮腫はない、チアノーゼはない 【神経学的所見】 異常はない 【入院時プロブレムリスト】 #1. 薬剤性腎障害 #2. 肺小細胞癌 #3. 発熱性好中球減少症 【入院時プロブレムリスト】 #1. 薬剤性腎障害 #2. 肺小細胞癌 #2-1. 上大静脈血栓症

- Encoderモデルベース (Decoderモデルを上回る)
- フレーズであっても高精度に抽出可能



Step2) 抽出用語／フレーズから標準コードへ変換 (Acc. > 0.9)

“右大腿骨転子部骨折”



NC723Z & XK9K (ICD11)
S7210 (ICD10)

“端座位に介助を要する”



d4153.2+ (ICF)

- Decoder モデルベース (SIP-Jmed-LLM2 13B)

一般カルテ文章からの情報抽出

Step1 (ICD)

テキストを入力してください

入院日：2019/06/28
退院日：2019/07/03
転帰：治癒
フォローアップ：外来で
【主病名】
急性腎障害

確定診断名

送信

LLMの多様な応用：保健医療行政支援

～ 人口動態調査死亡票の外因自動ICDコーディング～

様式第2号

人口動態調査死亡票

数字記入欄
0123456789 人口動態調査死亡票 平成 年 月 日 市区町村受付 統計法に基づく基幹統計調査

市区町村符号及び保健所符号 事件番号 平成 年 月 日 保健所受付 7 7 7 7

死亡票の種別
 1 自然死
 2 病死
 3 外因死
 4 不詳の死
 5 手術/解剖所見
 6 死因の種類
 7 追加事項

死亡した人の属性
 性別 男 女
 年齢 0歳未満 0歳以上1歳未満 1歳以上2歳未満 2歳以上3歳未満 3歳以上4歳未満 4歳以上5歳未満 5歳以上6歳未満 6歳以上7歳未満 7歳以上8歳未満 8歳以上9歳未満 9歳以上10歳未満 10歳以上14歳未満 15歳以上19歳未満 20歳以上24歳未満 25歳以上29歳未満 30歳以上34歳未満 35歳以上39歳未満 40歳以上44歳未満 45歳以上49歳未満 50歳以上54歳未満 55歳以上59歳未満 60歳以上64歳未満 65歳以上69歳未満 70歳以上74歳未満 75歳以上79歳未満 80歳以上84歳未満 85歳以上89歳未満 90歳以上94歳未満 95歳以上99歳未満 100歳以上

死亡した人の職業
 1 農林業・漁業・畜産業
 2 鉱業・採石業・建設業
 3 製造業・建設業
 4 商業・サービス業
 5 学業
 6 無職
 7 不明

死亡した人の居住地
 1 市区町村
 2 特別区
 3 特別区
 4 特別区
 5 特別区
 6 特別区
 7 特別区

死亡した人の死亡原因
 外傷性ショック T794
 胸部大動脈損傷、脳挫傷 S250
 全身打撲 T009
 糖尿病 E11

その他の付帯情報
 手術/解剖所見
 死因の種類
 追加事項

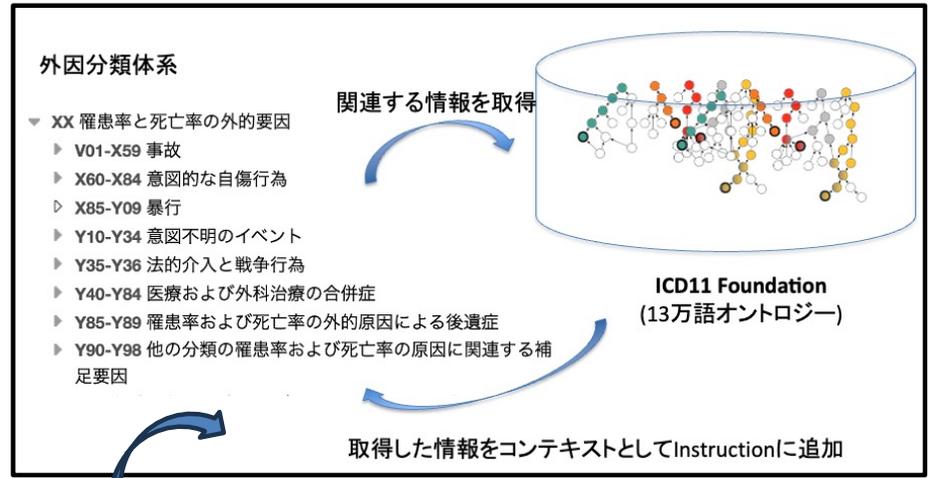
注：この様式は、平成30年1月1日から適用する。

▶ 基本的な情報

▶ 死亡の原因（傷病名欄）

▶ その他の付帯情報

- ・手術/解剖所見
- ・死因の種類：病死 / 自然死 / 外因死 / 不詳の死
- ・追加事項： 外因死の手段 / 状況
生後1年未満死の追加事項、その他付言すべきことから



“横断歩道歩行中に自動車と接触して受傷
救急室搬入時に心肺停止を確認”

“V031”

ここをAIで読み取り

生成AI + RAG

現在は職員が人手で
チェック（月4万件）

死亡外因ICDコード推定

死亡の原因

一酸化炭素中毒

I (イ)

I (ウ)

I (エ)

II

厚生労働省 人口動態統計支援

死因の種類

9 自殺

(人口動態・保健社会統計室にて試用中)

外因死の手段及び状況

ユニットバス内で練炭にて一酸化炭素を発生させた状況 (4) 欄推定、死亡したとき：令和X年X月X日頃、傷害が発生したとき

死亡外因送信

ICD11コード：PD02

- SIP-JMedLLM-2 を死亡票データ14万件でファインチューニング
- ICD10, ICD11 の双方でコーディング可能 (いずれも **96%の正解率**を達成)
- 厚生労働省内での人手作業の大幅な効率化
- 非常に高速な処理
- 外に情報を漏らさない運用が可能
- WHO会議でも日本の事例として大きな反響

難しいポイント

• 分類体系(ICD10/11/ICF)でのコード化戦略

– 逐次降下？



死亡外因コーディング

- 分類ならば、最上位から順に候補のリストから選択させて、降りていけば良い
- 排反ならばやりやすいが、実際は排反ではない。

– 用例ベース？



汎用診療情報コーディング

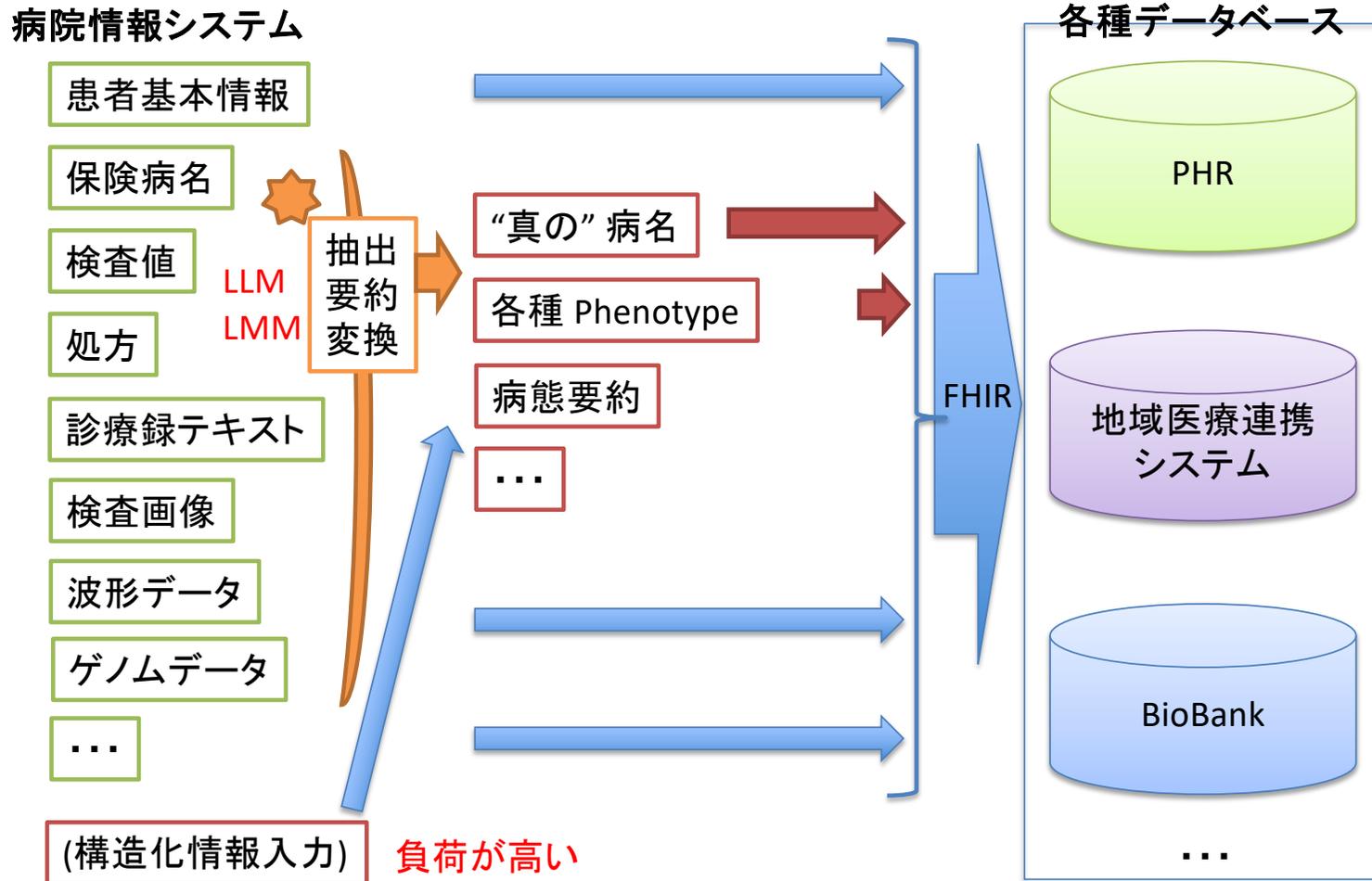
- 用例が少ない項目には到達できない

• バリエーションが多様

- コード化対象は単なる単語ではない。長いフレーズも。
- Post-Coordination
 - 複数のコードを組み合わせて複雑な病態を表現
 - 単純な構成原理ではない
- そもそも分類コードの数が多い
 - 約2万。Post-Coordinationを含めるともっと

診療情報とLLM/LMMが切り開く世界

～高品質・高精細 Deep Phenotype 情報の提供に向けて～



今後の活用展望

- **2～3年で急速に導入が進むと考えられる**
 - 業務軽減の需要は高い
 - 救急でのトリアージなど各領域での診療支援
 - エージェント型臨床助手へ
 - AI自動記載、コーディング、オーダーの下書きセットまで自動で
- **各種の課題**
 - ハルシネーション・抽出もれの対策
 - 医師による2重チェックはあるが・・・
 - プライバシー保護
 - バイアス
 - FTコスト
 - 規制・ガイドラインのさらなる充実
- **国産モデルのさらなる充実化**
 - 日本・EUのみならず、今後他国でもローカル言語・診療慣習に最適化したLLM/LMM開発が進む

ご清聴ありがとうございました



今井 健

東京大学大学院医学系研究科

疾患生命工学センター 医工情報学部門

imai@m.u-tokyo.ac.jp