

健康長寿産業の発展に向けたPHRデータ利活用について ー生成系AI・大規模言語モデルの可能性ー

矢野 裕一郎 MD, PhD, FAHA

順天堂大学 医学部 教授
AIインキュベーションファーム 副センター長

東京科学大学 医療イノベーション機構 特任教授

兼任所属



横浜市立大学 客員教授



香川大学
KAGAWAUNIVERSITY

客員教授



国立循環器病研究センター
National Cerebral and Cardiovascular Center
客員部長



Duke
UNIVERSITY

Adjunct Professor



国立大学法人
滋賀医科大学
SHIGA UNIVERSITY OF MEDICAL SCIENCE

客員教授



Stanford
MEDICINE

Center for Asian
Health Research
and Education

Global Faculty



RITSUMEIKAN
UNIVERSITY

立命館大学

客員教授



Adjunct Professor

データ駆動型ヘルスケア

1. 現状

病院内医療データの活用

2. 課題

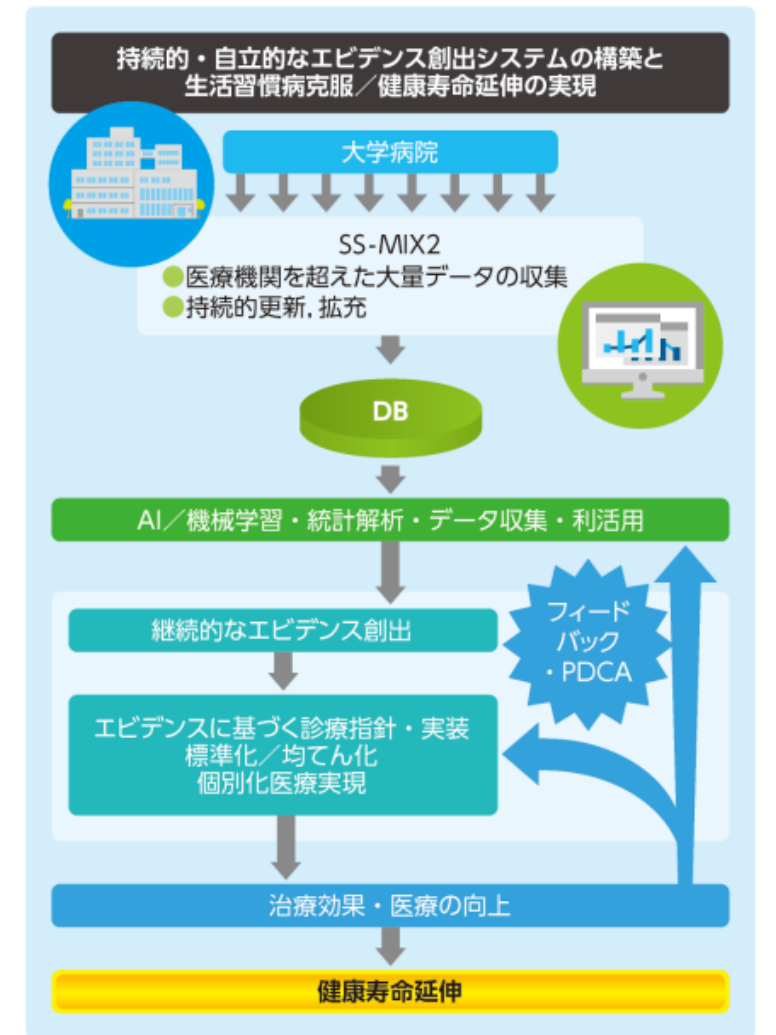
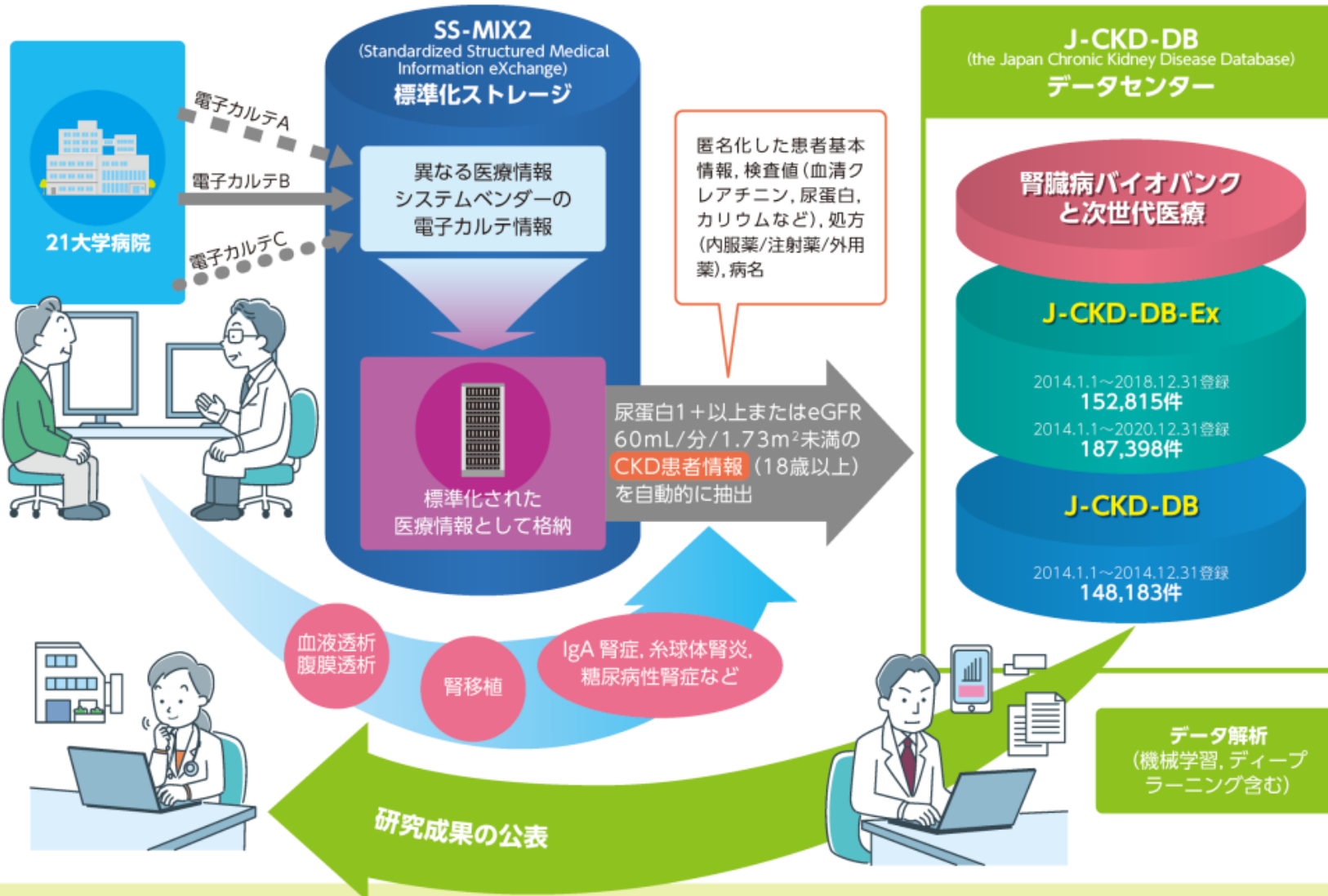
病院内医療データの限界・PHRの重要性

3. 未来

健康長寿産業の発展に向けたPHRデータ利活用
生成系AI・大規模言語モデルの可能性

J-CKD-Database

日本腎臓学会×日本医療情報学会×厚労省×AMEDが構築してきた
腎臓病総合レジストリー



J-CKD-DB

2014年

施設		登録件数
1	川崎医科大学	10,520
2	東京大学	22,008
3	九州大学	14,194
4	岡山大学	10,375
5	旭川医科大学	3,032
6	横浜市立大学	11,307
7	京都大学	15,915
8	新潟大学	8,759
9	金沢大学	6,911
10	名古屋大学	10,421
11	高知大学	6,734
12	筑波大学	8,725
13	和歌山県立医科大学	11,818
14	島根大学	2,665
15	香川大学	4,799
総登録件数		148,183

J-CKD-DB-Ex

2014年~2023年

施設		登録件数
1	川崎医科大学	57,320
2	東京大学	61,111
3	九州大学	63,085
4	岡山大学	52,274
5	旭川医科大学	11,169
6	横浜市立大学	51,679
7	埼玉医科大学	73,190
8	東京慈恵会医科大学	60,382
9	名古屋大学	29,066
10	新潟大学	抽出準備中
11	金沢大学	抽出準備中
12	京都大学	抽出準備中
13	滋賀医科大学	抽出準備中
14	奈良県立医科大学	抽出準備中
15	大阪公立大学	抽出準備中
16	大阪大学	抽出準備中
17	順天堂大学	抽出準備中
18	慶應義塾大学	抽出準備中
総登録件数		459,276

経時的なデータを有する慢性腎臓病患者
データを自動的に抽出

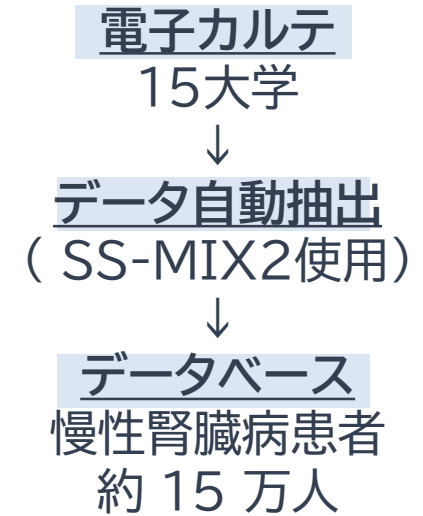
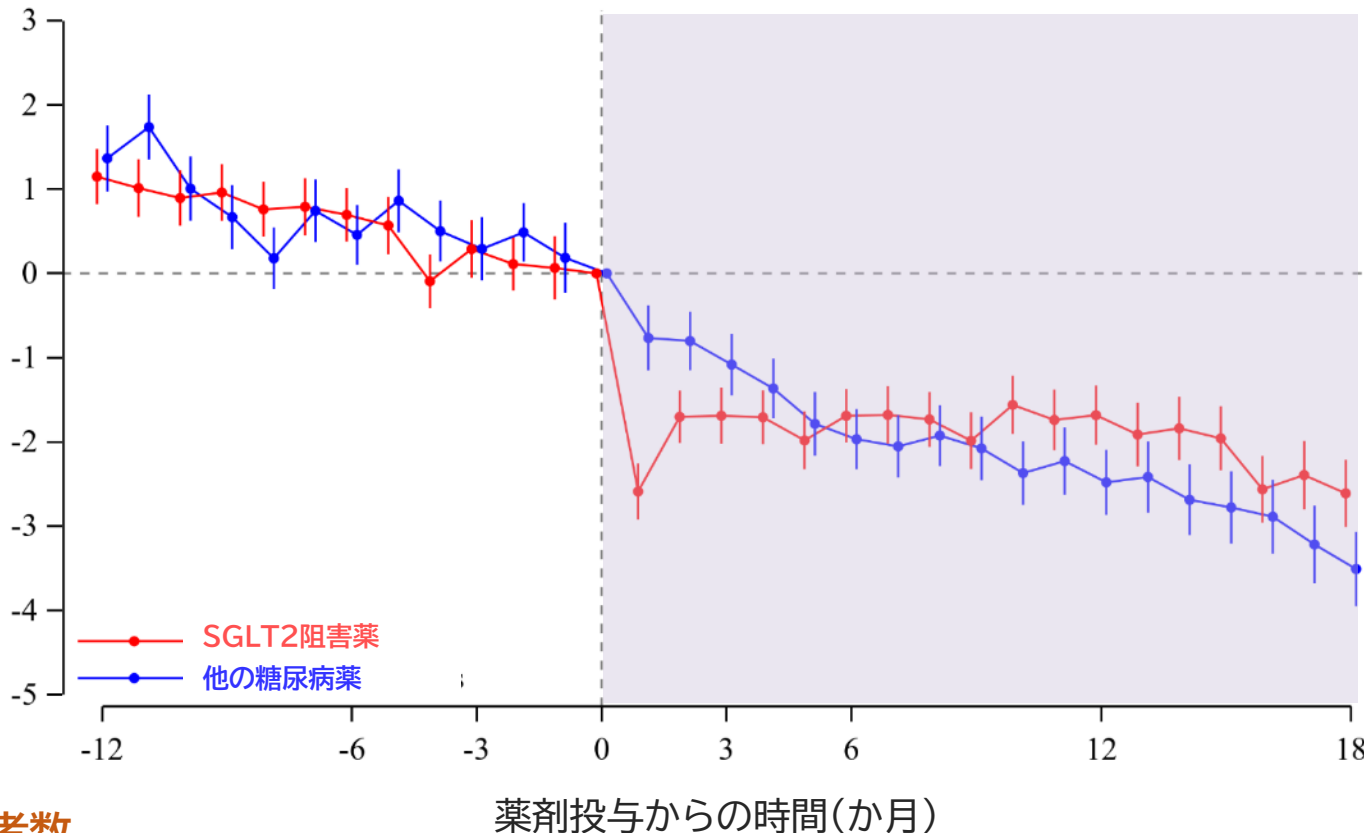
リアルワールドエビデンスの価値

RCTでは組み込まれない症例に対する効果の検証

SGLT2阻害薬 vs 他の糖尿病薬における腎臓への影響

薬剤投与

糸球体濾過量の平均変化値



データベース蛋白尿陽性
あるいは陰性例に対する
SGLT2阻害薬の効果の評価
(プロペンシティスコアマッチング)

リスク保有者数

SGLT2阻害薬	453	497	396	716	439	503	368	256
他の糖尿病薬	294	400	334	576	366	394	308	213

※SGLT2
:sodium-dependent glucose transporters 2

リアルワールドエビデンスの価値 ~経時的データの活用例~

研究例

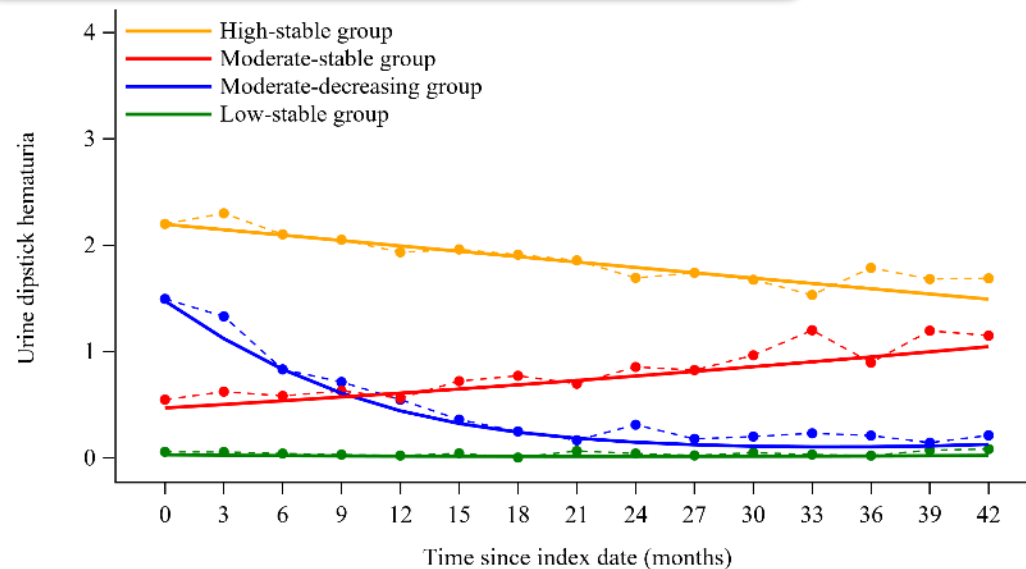
①

IgA腎症患者における血尿と蛋白尿の経時的変化

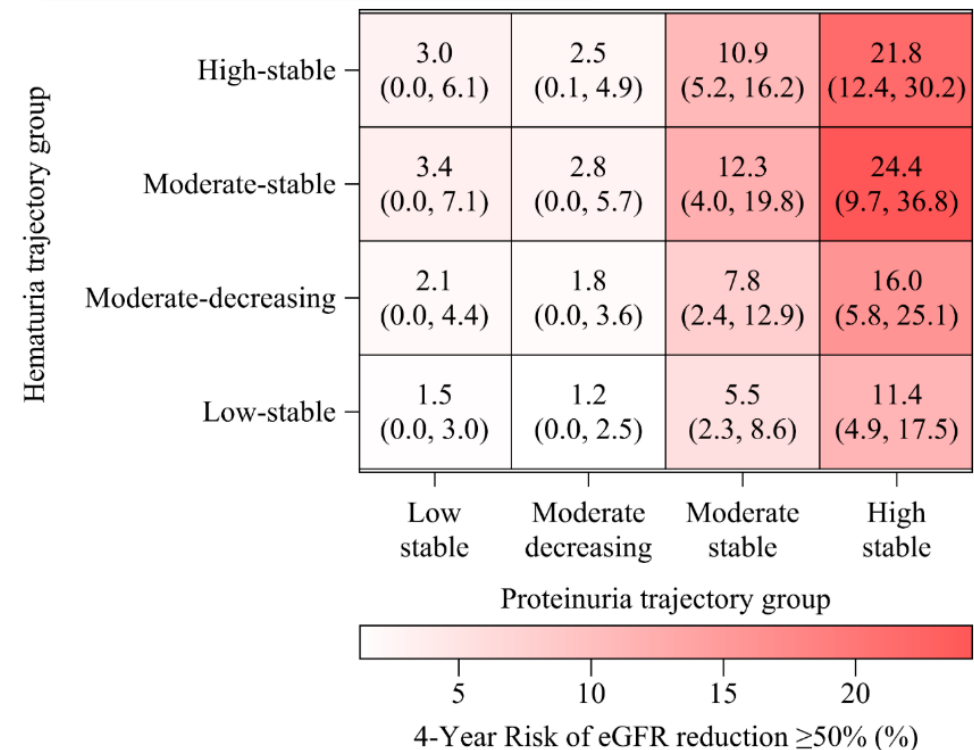
患者レジストリ (J-CKD-Database) 活用事例

データトラッキング (血尿等)

腎臓疾患リスク分析



	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42
High-stable group	248	171	155	118	108	89	94	87	87	84	72	68	72	63	58
Moderate-stable group	98	64	72	53	56	55	51	38	45	42	34	36	29	38	32
Moderate-decreasing group	136	95	101	93	85	64	64	56	43	56	50	37	47	36	43
Low-stable group	273	151	179	146	153	147	122	117	109	90	105	89	88	87	85



血尿と蛋白尿のトラッキングにより、高リスクのIgA腎症患者を同定が可能

データ駆動型ヘルスケア

1. 現状

病院内医療データの活用

2. 課題

病院内医療データの限界・PHRの重要性

3. 未来

健康長寿産業の発展に向けたPHRデータ利活用
生成系AI・大規模言語モデルの可能性

リアルワールドエビデンスの価値 ~経時的データの活用例~

研究例

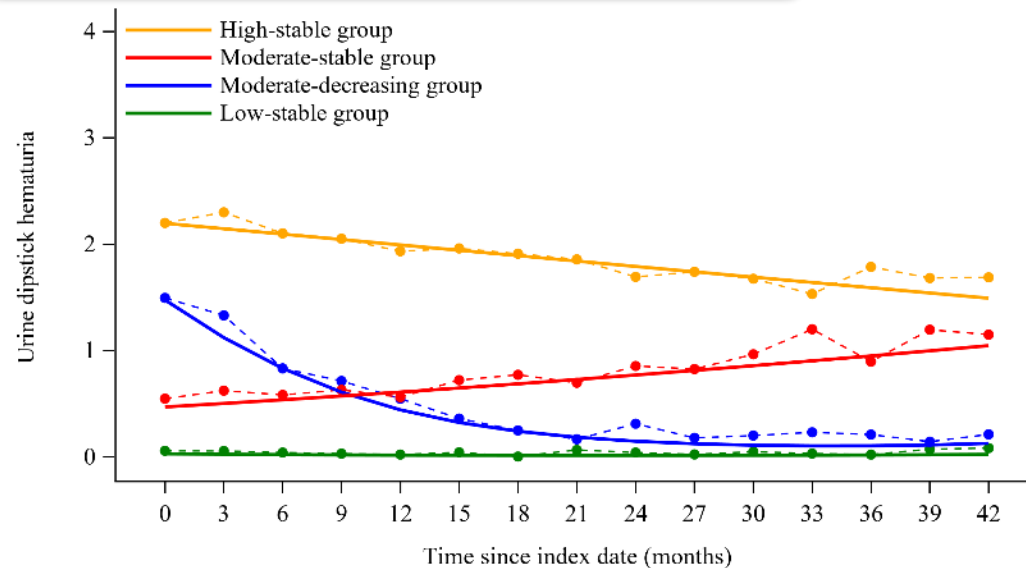
①

IgA腎症患者における血尿と蛋白尿の経時的変化

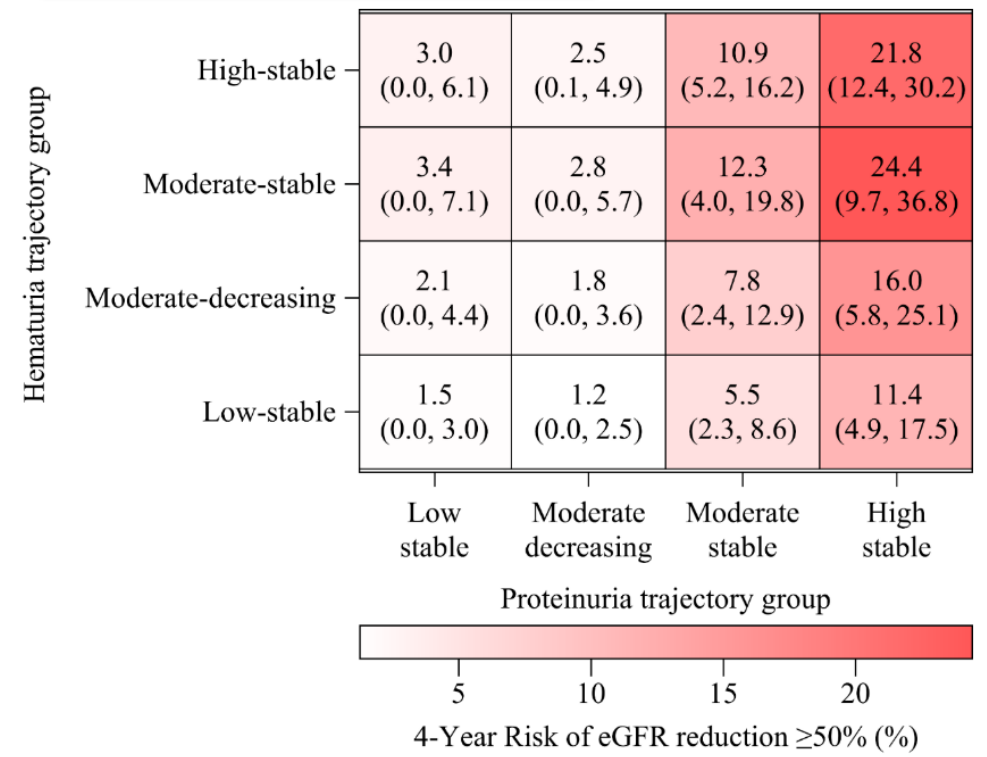
患者レジストリ (J-CKD-Database) 活用事例

データトラッキング (血尿等)

腎臓疾患リスク分析



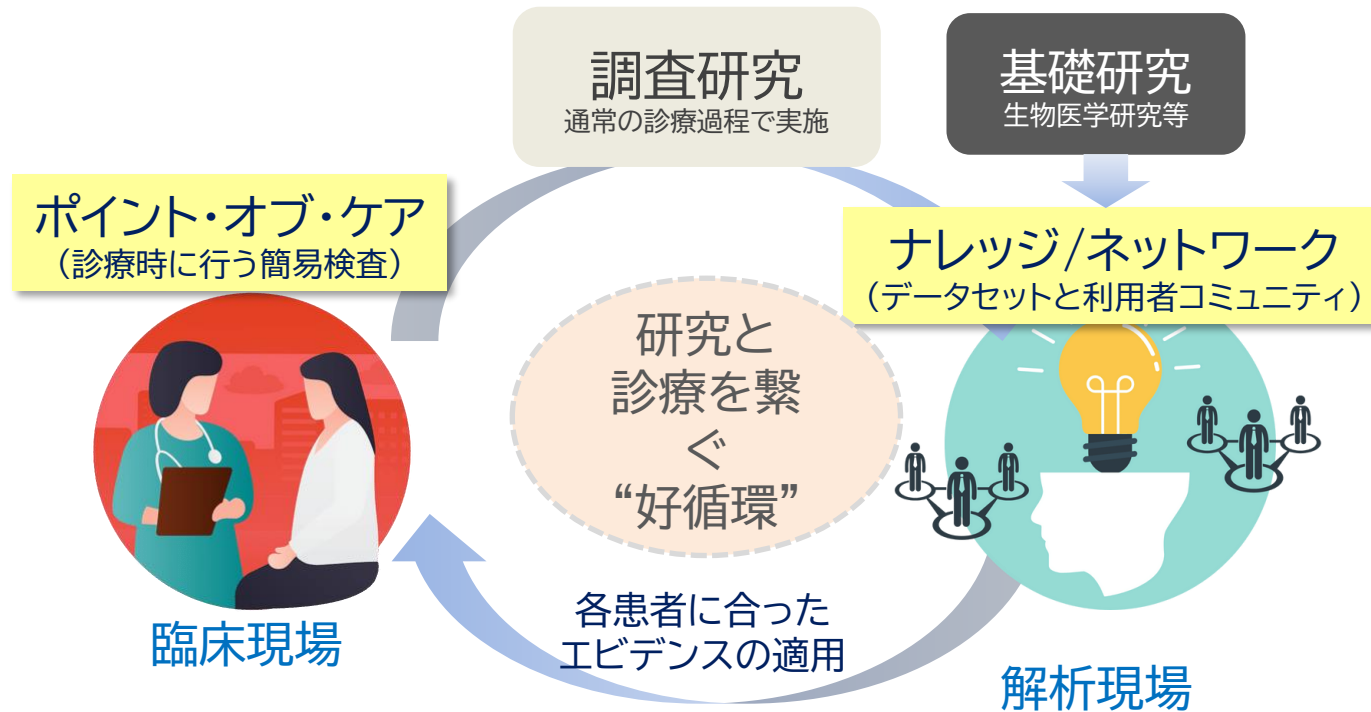
	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42
High-stable group	248	171	155	118	108	89	94	87	87	84	72	68	72	63	58
Moderate-stable group	98	64	72	53	56	55	51	38	45	42	34	36	29	38	32
Moderate-decreasing group	136	95	101	93	85	64	64	56	43	56	50	37	47	36	43
Low-stable group	273	151	179	146	153	147	122	117	109	90	105	89	88	87	85



血尿と蛋白尿のトラッキングにより、高リスクのIgA腎症患者を同定が可能

ヘルスケア領域のデータ連携 ~これからできること・すべきこと~

実社会の医療データの質を改善し、解析精度を向上するためには
“Learning Health System(LHS)*”の医療現場への実装が必要



【LHSの必要要素】

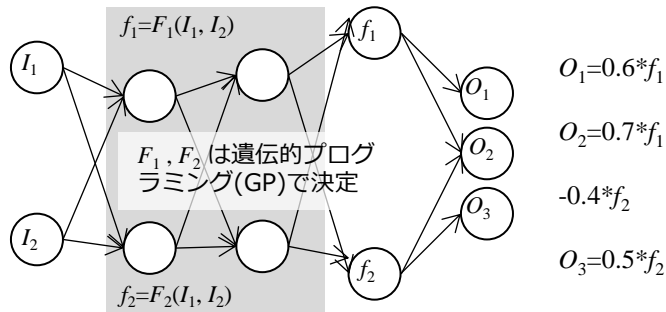
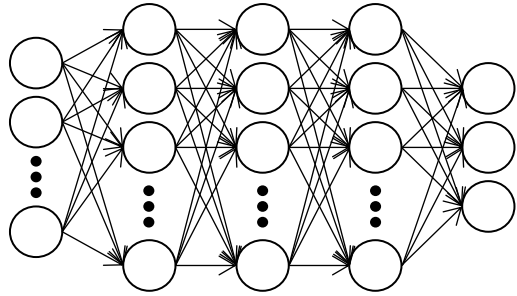
- ①データ収集システム: EHR・PHRシステム
- ②データ統合・インフラストラクチャ: データウェアハウス、データレイク、クラウドベースのストレージおよび処理プラットフォーム
- ③分析および意思決定支援ツール: AIや機械学習モデル、リアルタイムデータ分析とフィードバックシステム
- ④コミュニケーションと協調のためのツール
- ⑤ガバナンスとポリシーマネジメント: プライバシーとセキュリティのポリシー
- ⑥教育と訓練プログラム: 医療従事者および技術スタッフ向けの継続的な教育とトレーニング
- ⑦品質とパフォーマンスのモニタリングシステム:

2006年に米国医学研究所が公表したコンセプトに基づく
National Academies Press (US); 2007.ISBN-13: 978-0-309-10300-8ISBN-10: 0-309-10300-

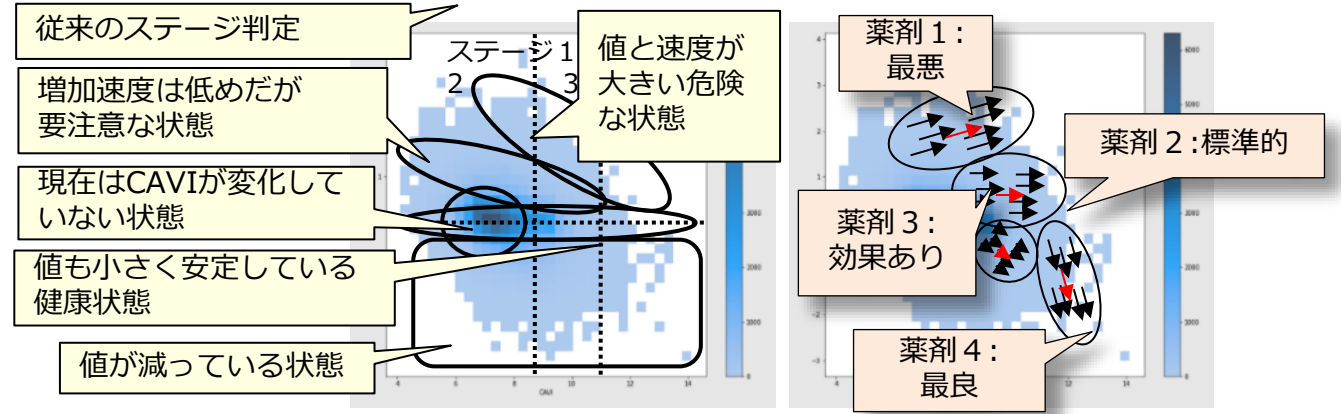
日々生成される情報(ナレッジ)が日々の診療に還元される“好循環”を生み出す

病院データだけでは、腎機能予測に限界がある

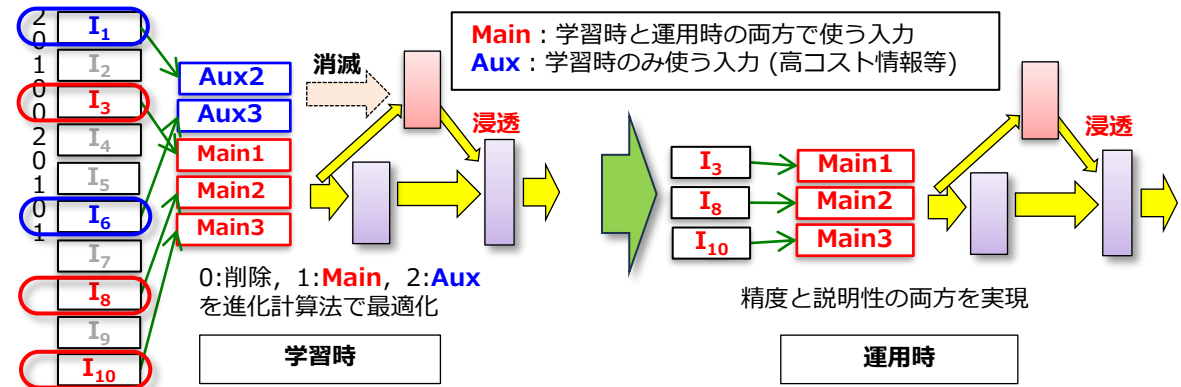
深層回路を線形回路化して
図や言葉で説明



CAVI(cardio-ankle vascular index)と dCAVI/dt の
2次元状態空間でデータを可視化

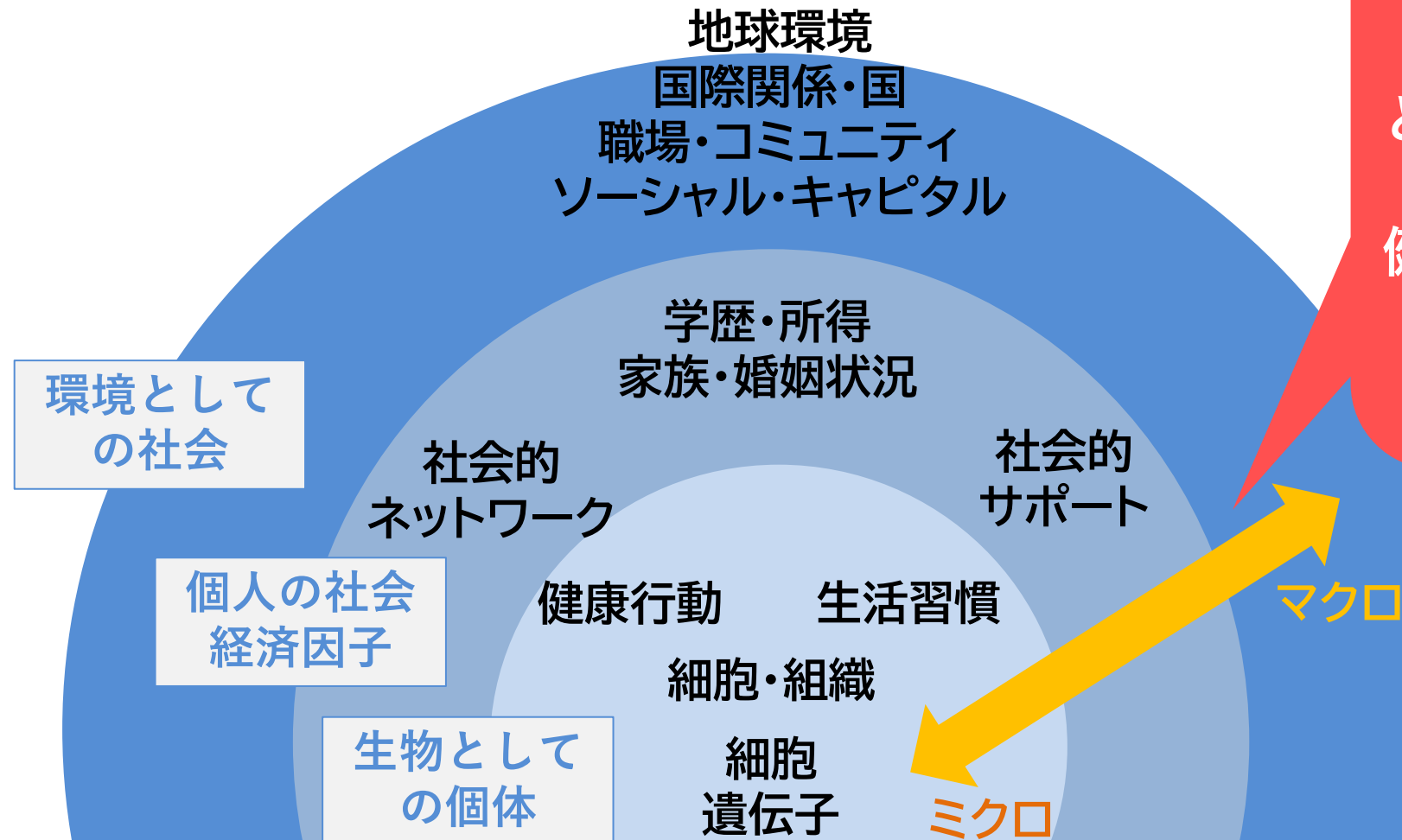


浸透学習法を用いて運用時の入力変数を究極まで削減



健康の社会的決定要因 (Social determinants of health)

健康の決定因子の階層構造

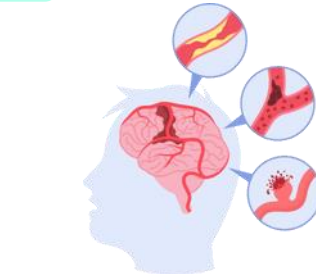
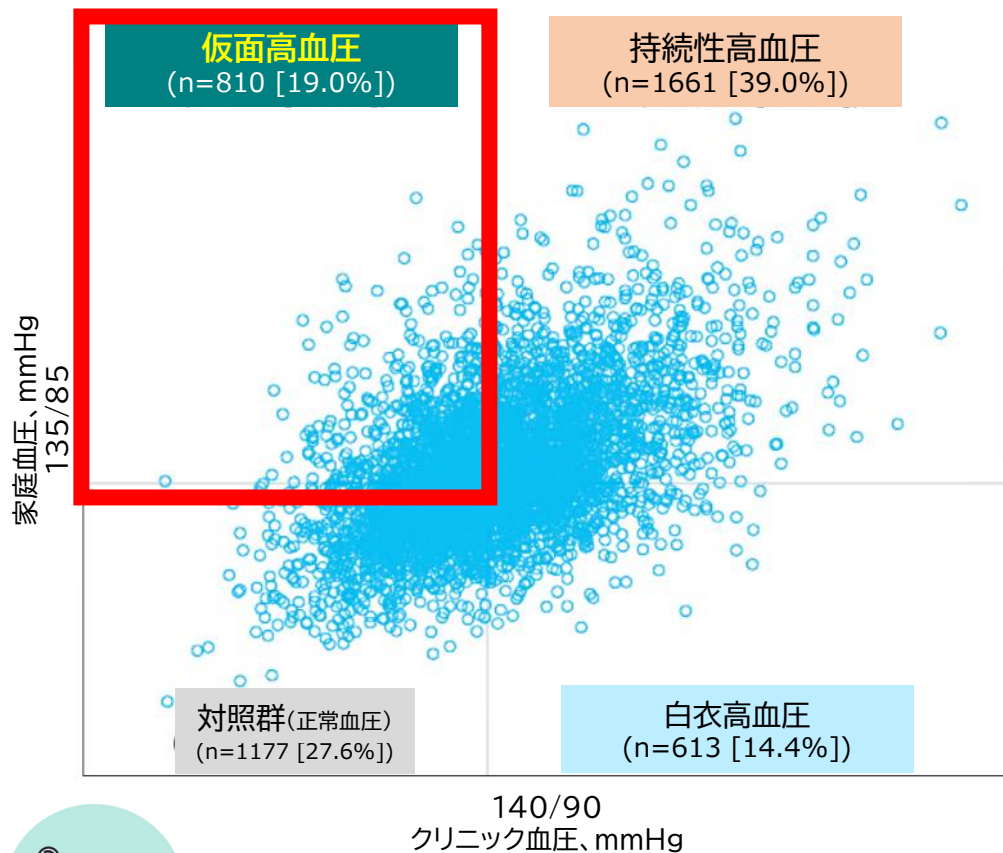


科学技術力を
駆使して、
どうモニタリング
(=可視化) し、
健康資産向上につ
なげるか？

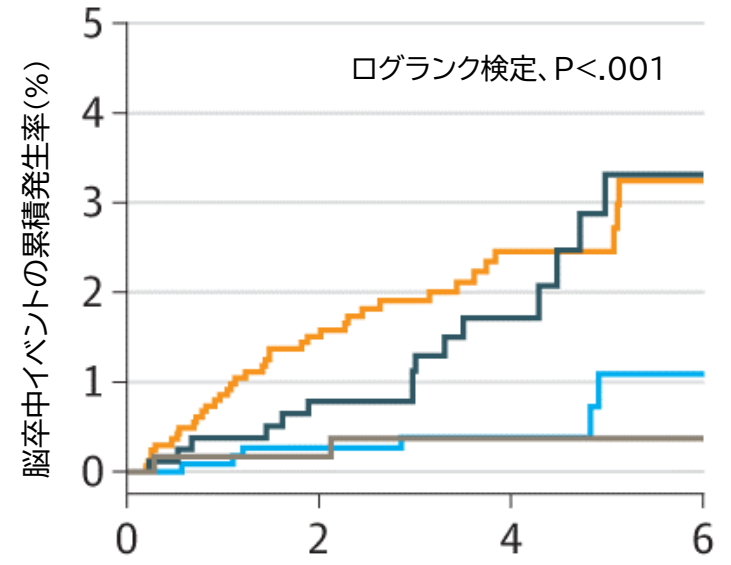
病院外で血圧が高いリスク

～仮面高血圧(隠れ高血圧)と心血管疾患に関する研究～

血圧と脳卒中リスクに関する疫学調査(日本)



診察外の血圧が高い人(仮面高血圧含む)は脳卒中リスクが高い



リスク人数(Number at Risk)

	1661	1373	720	313
持続性高血圧	1661	1373	720	313
仮面高血圧	810	681	394	187
白衣高血圧	613	516	255	99
対照群(正常血圧)	1177	1005	515	231

経過観察期間(年)

— 持続性高血圧 — 仮面高血圧 — 白衣高血圧 — 対照群(正常血圧)

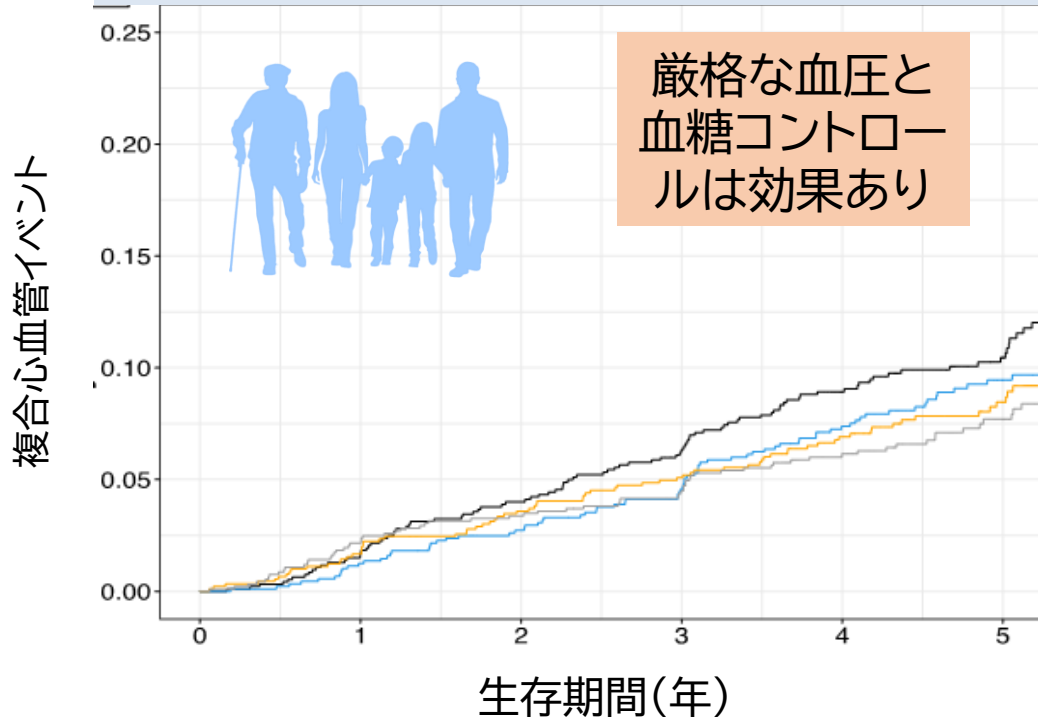
診察“外”での健康管理が重要

出所: Fujiwara T, Yano Y, et al. *JAMA Cardiol.* 2018

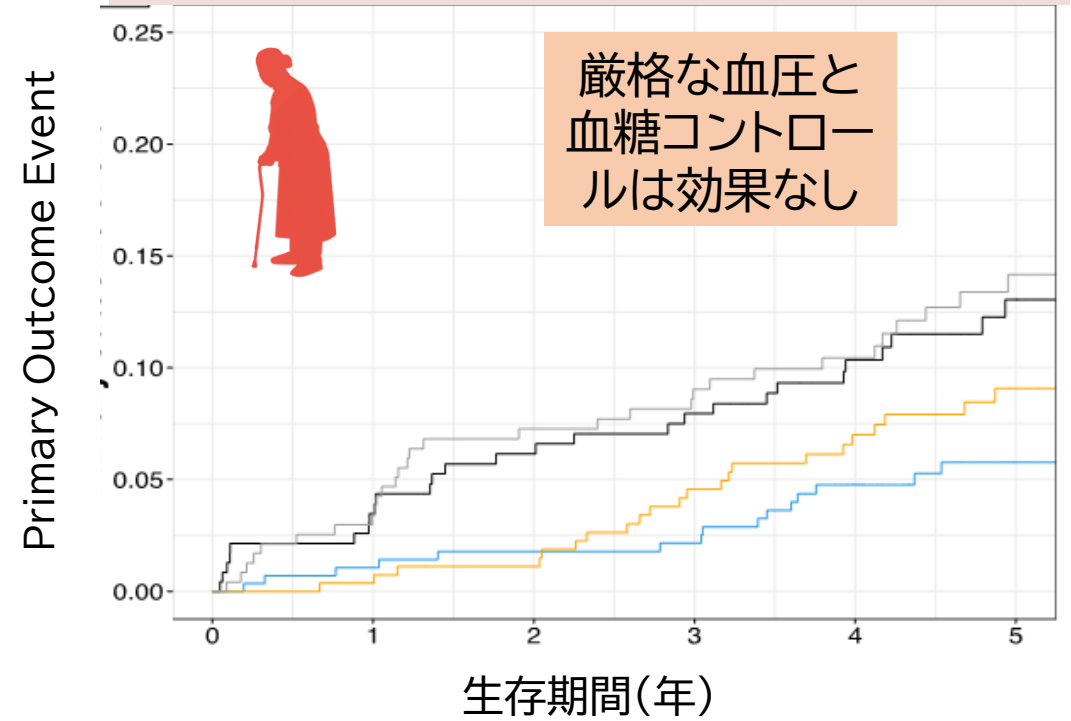
患者さんの社会的要因により、 治療の効果は異なる可能性がある

ACCORD-BP（厳格な血糖と血圧コントロールの効果を検証したRCT）

同居（家族等と居住）



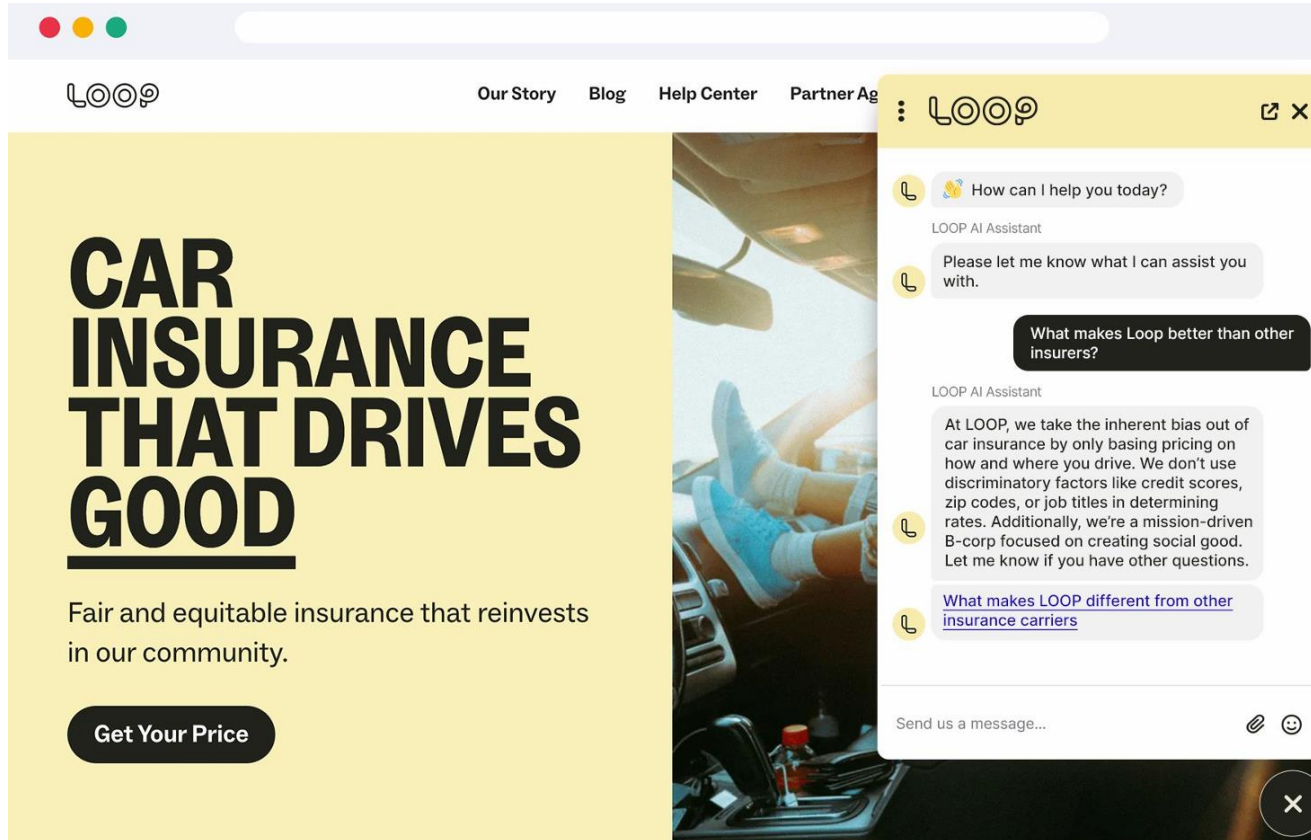
独居（一人で居住）



— 標準血糖コントロール+標準血圧コントロール
— 集中的な血糖コントロール(HbA1c<6.0%)+標準血圧コントロール

— 標準血糖コントロール+集中的な血圧コントロール
— 集中的な血糖コントロール+集中的な血圧コントロール(SBP<120mmHg)

健康の社会的決定要因 (Social determinants of health)の盲点



保険料の決め方

信用スコア、所得、職業 **特定のグループに対するバイアスが生じる**

AIを掲載したスマートフォンアプリを利用して、どこで(道路の種類、交通量、天候など)、どのように(速度、急ブレーキ、携帯電話通話など)運転しているか

保険料を請求するリスクを、高、中、低に分類

- 低い人には保険料の大幅値下げ
- 安全運転の動機付け
- データから運転に関するアドバイス(週一)→行動変容により保険料が変化することを可視化

顧客生涯価値を高める
(生成系AIにより増大)

Personal Records

バイアスを抑え、エクイティを向上させ、価値獲得

令和6年度診療報酬改定:医療DX推進体制整備加算が新設 →加速する、病院外におけるケア



診療情報 (EHR) と
患者情報 (PHR) の
連携

標準化規格

これまでの医療データ連携
(電子的診療情報交換推進事業 平成18年～)

これからの医療データ連携
(次世代医療基盤整備事業 平成26年～)

SS-MIX (SS-MIX2)
Standardized Structured Medical record Information eXchange

HL7 FHIR
Fast Healthcare Interoperability Resources

- 医療機関外とのネットワーク接続原則なし:オフライン
- データは、可搬メディア(CD/DVD)で、患者が運搬

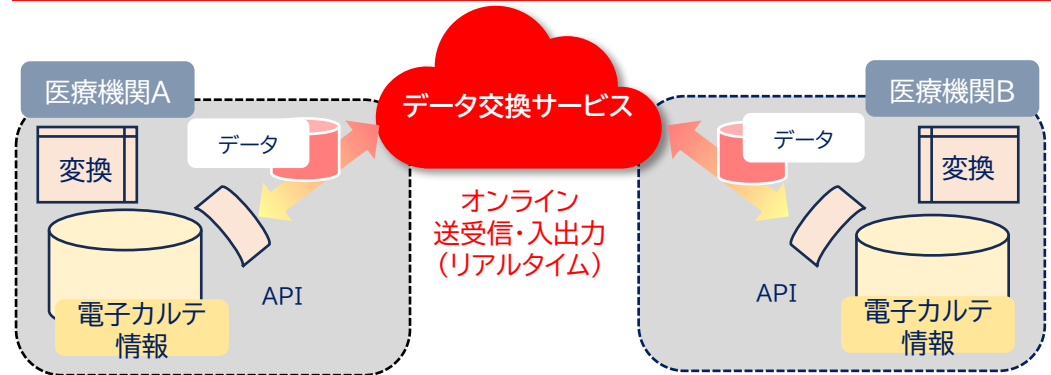
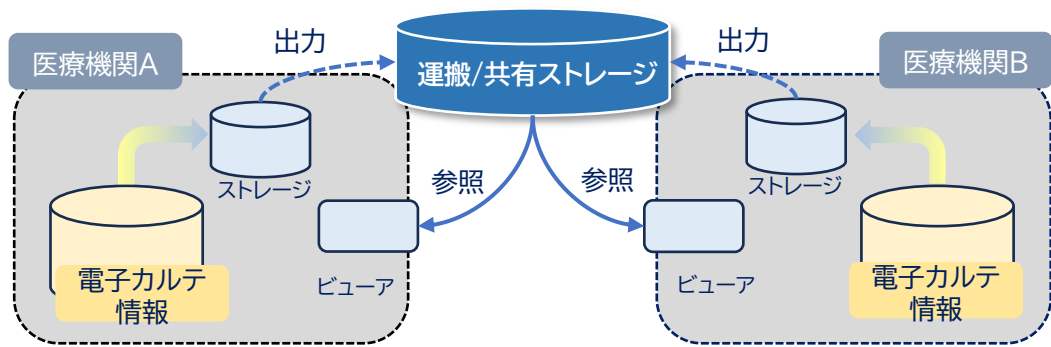
病院間での
情報共有

- 医療機関外とのネットワーク接続あり:オンライン
- データは、オンラインで、適時、送受信

- HL7 Ver.2.5形式のメッセージデータ
(医療分野独自フォーマット:データ区切り“|”)をファイルとし、フォルダ構成でストレージに格納し、可搬メディアに出力・運搬

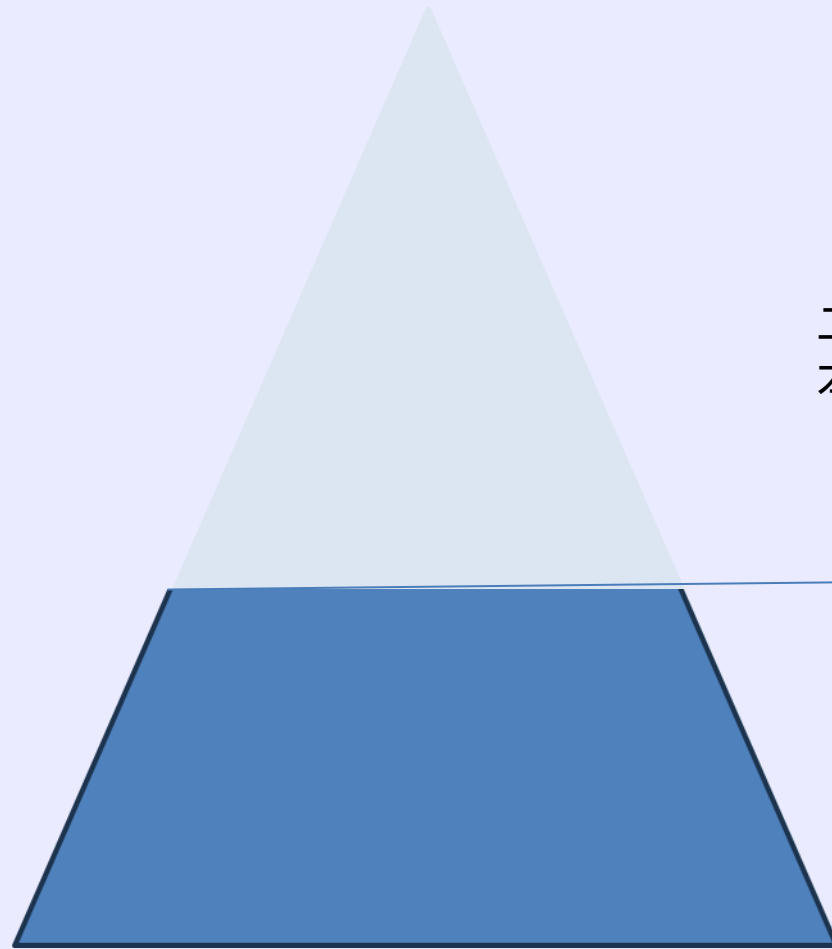
データ連携
の特徴

- Web(HTTPプロトコル)で一般的に用いられるデータ交換方式(REST通信+JSON形式)で、形式変換しやすいJSON/XMLフォーマットで送受信



引用元: 厚生労働省「健康・医療・介護情報利活用検討会 医療情報ネットワークの基盤に関するワーキンググループ資料(2022)」等を基に作成

データ駆動型研究の限界と、その先へ：**潜在意識**を科学する



ニーズ(顕在意識)
本人が自覚している思い

私は健康に気を遣っている
ので、健康的な飲み物
が買いたい

インサイト(潜在意識)
本人が自覚していない
隠れた心理

実は。。。
ジャンクフードを食べ
たいので、その背徳感？
から、健康的な茶を
買うように心がけて
いる

データ駆動型ヘルスケア

1. 現状

病院内医療データの活用

2. 課題

病院内医療データの限界・PHRの重要性

3. 未来

健康長寿産業の発展に向けたPHRデータ利活用
生成系AI・大規模言語モデルの可能性

日常生活に浸透してくる生成系AI

AI(人工知能)

機械学習(大量のデータをコンピューターが自動的に学習。
背後にあるルールやパターンを自動で認識)

ニューラルネットワーク(人間の脳の神経細胞を模倣した数学的モデル。DLの基盤技術)

ディープラーニング
(NNを多層に重ねて複雑なパターン認識や予測を行う)

**大規模言語
モデル**

Chat GPT
など

生成AIとは・・・

- 事前に学習した大量データからパターンやルールを識別
- 多くが、人間の脳の神経細胞(ニューロン)の構造を模したニューラルネットワークの一種である深層学習モデルで構築
- 最適な情報をつかってテキスト、出力画像、音声、デザインなどを新たに生成



身近な課題解決に活用できる
ソリューション

世界中の高血圧患者が日常生活で抱く疑問とは？ 回答の精度は（GPT-4）？

ChatGPTを用いた高血圧診療の“集合知”の臨床応用

Relevance of ChatGPT's Responses to Common Hypertension-Related Patient Inquiries

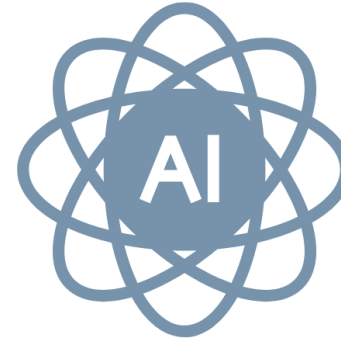
Yuichiro Yano¹, Akira Nishiyama², Yusuke Suzuki³, Satoshi Morimoto⁴, Takashi Morikawa⁵, Tomohito Gohda⁶, Hiroshi Kanegae, Naoki Nakashima⁷

Dialog-based artificial intelligence (AI) language models, powered by cutting-edge deep learning methodologies, emulate human-like interactions by adeptly responding to natural language queries. For patients, it serves as a readily accessible and convenient source of general insights on diseases, symptoms, treatments, and potential side effects. Incorporating dialog-based AI (eg, ChatGPT) into medicine may hold promise, but careful planning, ethical considerations, and continuous oversight are necessary to ensure patient safety and optimal health care outcomes.^{1,2}

Previous research has highlighted that ChatGPT provided largely appropriate responses to queries related to the prevention of cardiovascular diseases.³ However, these questions were formulated by experts in preventive cardiology and grounded in established guidelines. The evaluators assessing GPT's responses were cognizant of the fact that they originated from ChatGPT, introducing a possible bias. Moreover, the consistency in the quality of GPT's answers when posed in English compared with other languages (eg, Japanese) is yet to be conclusively determined.

responses were consistent across all 3 requests. From this set, we identified 21 pertinent questions for our investigation. English inquiry: we then made a similar request in English: "Please provide a list of frequently asked questions from patients with hypertension." ChatGPT-4 provided a list comprising 30 questions. Notably, 20 of these questions coincided with those from the Japanese list. Final selection: based on the above results, we included these 20 questions in our detailed study (Table). We posed each question to ChatGPT-4 in both Japanese and English and recorded the responses. For the evaluation phase, each response, both in Japanese and English, was assessed by 3 accredited hypertension or nephrology specialists (S.M., T.M., and T.G.). The reviewers were led to believe that a medical resident provided the answers, unaware that they were actually generated by ChatGPT-4. They were tasked with grading each response as either "appropriate" or "inappropriate" based on the answer's alignment with the reviewer's medical knowledge. Additionally, reviewers were asked to decide which version of the response (Japanese versus English) was superior in terms of accuracy, comprehensiveness, professionalism, and level of detail.

人工知能



VS

人間（研修医）



- ✓ **ChatGPT-4に高血圧に関する20の質問を提出（日本語、英語の2つのバージョンで作成）**
- ✓ **その回答を「研修医が答えた」としてマスキングして、高血圧専門家3名が評価**

ChatGPTでの回答の医学的に妥当

- 日本語の回答の85%は、研修医として医学的に妥当
- 多くの質問について英語の回答例の方が日本語よりも優れている、または同等

Open AI O1: 多段階の推論、思考の連鎖に焦点



例: Open AI

非営利法人OpenAI, Inc.とその子会社である営利法人OpenAI Global, LLCなどの法人から構成される人工知能(AI)の開発を行う米国企業より汎用人工知能が提供されている。大規模言語モデルのGPTファミリー、テキストから画像を生成するモデルのDALL-Eシリーズ、そしてテキストから動画を生成するモデルのSoraなどがある。

旧モデル

現在
2024.9~

Open AI o1 (Strawberry)



- 推論能力(Reasoning Capability)が高い

自らの誤りを学習し、推論を向上できる技術(STaR/スタンフォード大学)を使用してチューニングや自己学習を繰り返す

- 長期的な問題(Long-Horizon Tasks)に対応できる

短期的なタスクだけでなく長期的なタスクやデータ管理もおこなえ、複雑な意思決定や長期的なプロジェクト管理等多岐にわたる分野での応用が可能になる

- 高品質な合成データ(Synthetic Data)の生成ができる

データプライバシーの観点を気にせずにデータを大量生成して学習に扱うことができる

GPT-5はヘルスケアをどう変えるのか？



例: Open AI

非営利法人OpenAI, Inc.とその子会社である営利法人OpenAI Global, LLCなどの法人から構成される人工知能(AI)の開発を行う米国企業より汎用人工知能が提供されている。大規模言語モデルのGPTファミリー、テキストから画像を生成するモデルのDALL-Eシリーズ、そしてテキストから動画を生成するモデルのSoraなどがある。

将来
2024年末～

GPT-5 Orion

● 高い信頼性

事実とは異なる情報を生成してしまう問題(ハルシネーション)を改善して正確で信頼性の高い情報を提供できるようになる

● マルチモーダル処理

テキストだけでなく、画像・音声・動画などの複数のデータ形式を統合的に理解・処理できるようになる

● 高い自律性

人間が細かく指示を与えなくても、自ら目標を設定し、必要な情報を収集・分析しながら、自律的にタスクを遂行できるようになる

医療領域を革新

膨大なデータ分析を
高速かつ正確に行える

(論文、臨床試験データ
電子カルテ情報等)

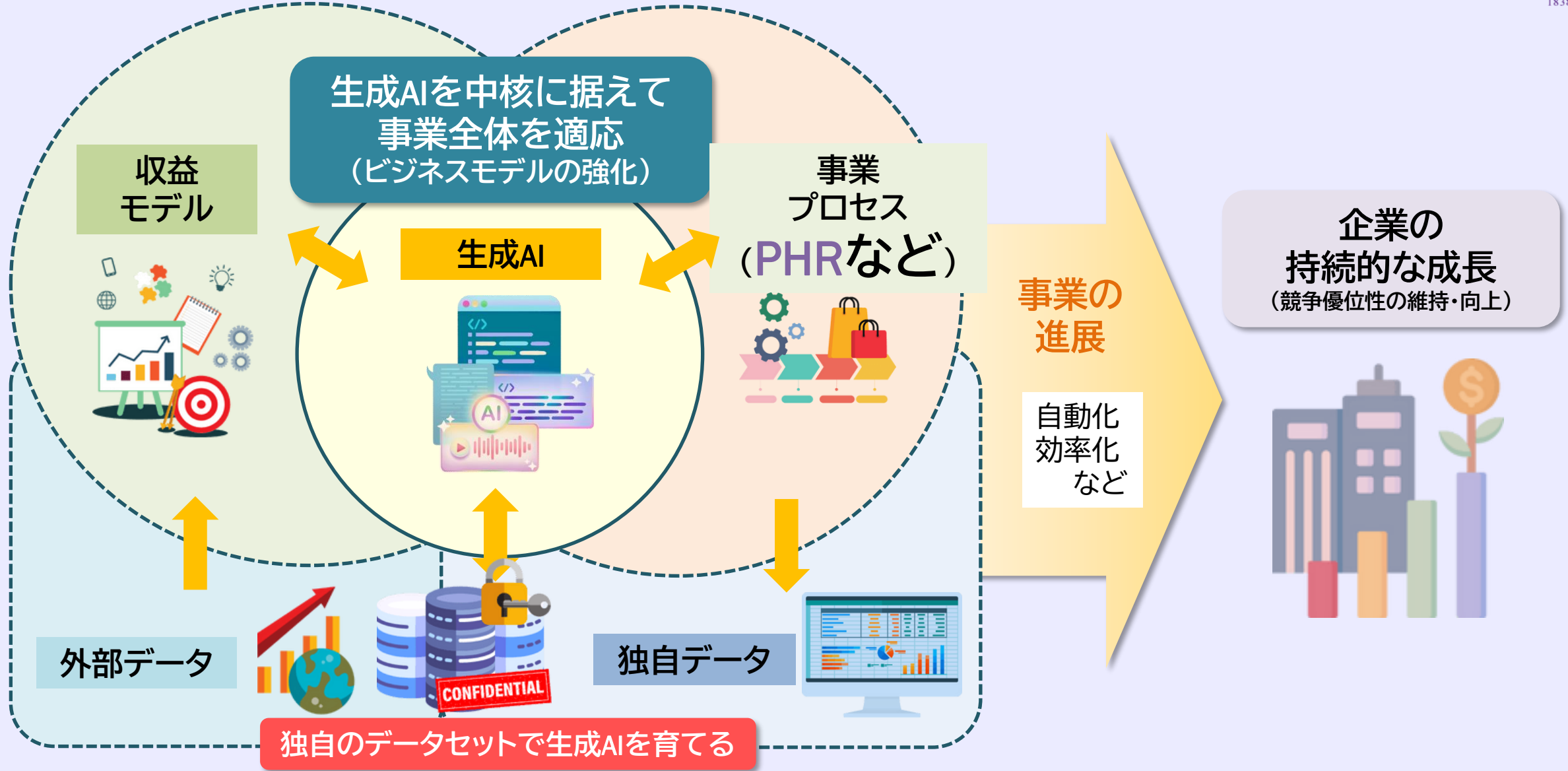


診療・創薬研究の
質・効率が向上する

診断・治療計画の個別化
新薬開発のスピードアップ等



生成AIを中核に据え、新たな価値獲得を目指す



Web検索は手動でなく自動的に(インタラクティブなAIエージェント)

— 個人のPHRデータに適したサービスを自動的に展開? —

例: ChatGPTのWeb検索機能を用いた飲食店検索



シカゴダウンタウンで、イタリアンの店をWEBで紹介してください。上司といくので、その辺かみしてください。

地元の飲食店について質問。
上司と一緒にという条件を追加

シカゴのダウンタウンで、上司との会食に適したイタリアンレストランをご紹介します。以下の店舗は、雰囲気、サービス、料理の質ともに高く評価されています。

GPT-4がウェブ検索を開始して
必要な情報を提示

- AI生成回答
- 情報ソース
- 詳細情報へのリンク



これらのレストランは、上司との会食にふさわしい環境とサービスを提供しています。訪問の際は、事前に予約を取ることをおすすめします。

強力かつ双方向的ツールとして機能
→ 現実世界で複雑な行動を(私の代わりに)実行できる、より強力にインタラクティブなAIエージェントの開発

革新的な意思決定・行動変容プロセス

-システムゼロ(System 0)の概念-

System 0

一旦、そして断続的にAIをはさむ



日常の認知タスク(テキストの校正や画像生成、情報検索など)はAIシステムに委ねる→従来の認知プロセスに新たな基盤を提供

生物学的でない人工的な知能の層として機能し、システム1とシステム2の入力を加工・強化する「前処理器」としての役割を果たす



生成AIによる認知拡張・補完

- 情報処理の強化
- 認知・分析の多角化
- 分散的知性の集約

System 1

ファスト・シンキング
(直感的かつ迅速な思考)



System 2

スロー・シンキング
(分析的かつ思慮深い思考)

AIと人間の相互作用が、現代の認知においてどのような基盤的役割を構築していくかが鍵

データ駆動型ヘルスケア

1. 現状

病院内医療データの活用

2. 課題

病院内医療データの限界・PHRの重要性

3. 未来

健康長寿産業の発展に向けたPHRデータ利活用
生成系AI・大規模言語モデルの可能性