

臨床研究における情報利活用

自治医科大学
永井良三

令和6年度予算案における 統合プロジェクトの概要

(目次)

1. 医薬品プロジェクト	325億円(AMED)+41億円(インハウス)
2. 医療機器・ヘルスケアプロジェクト	108億円
3. 再生・細胞医療・遺伝子治療プロジェクト	186億円
4. ゲノム・データ基盤プロジェクト	166億円(AMED)+16億円(インハウス)
5. 疾患基礎研究プロジェクト	150億円
6. シーズ開発・研究基盤プロジェクト	236億円
(参考)ワクチン開発・生産体制強化に関する事業	

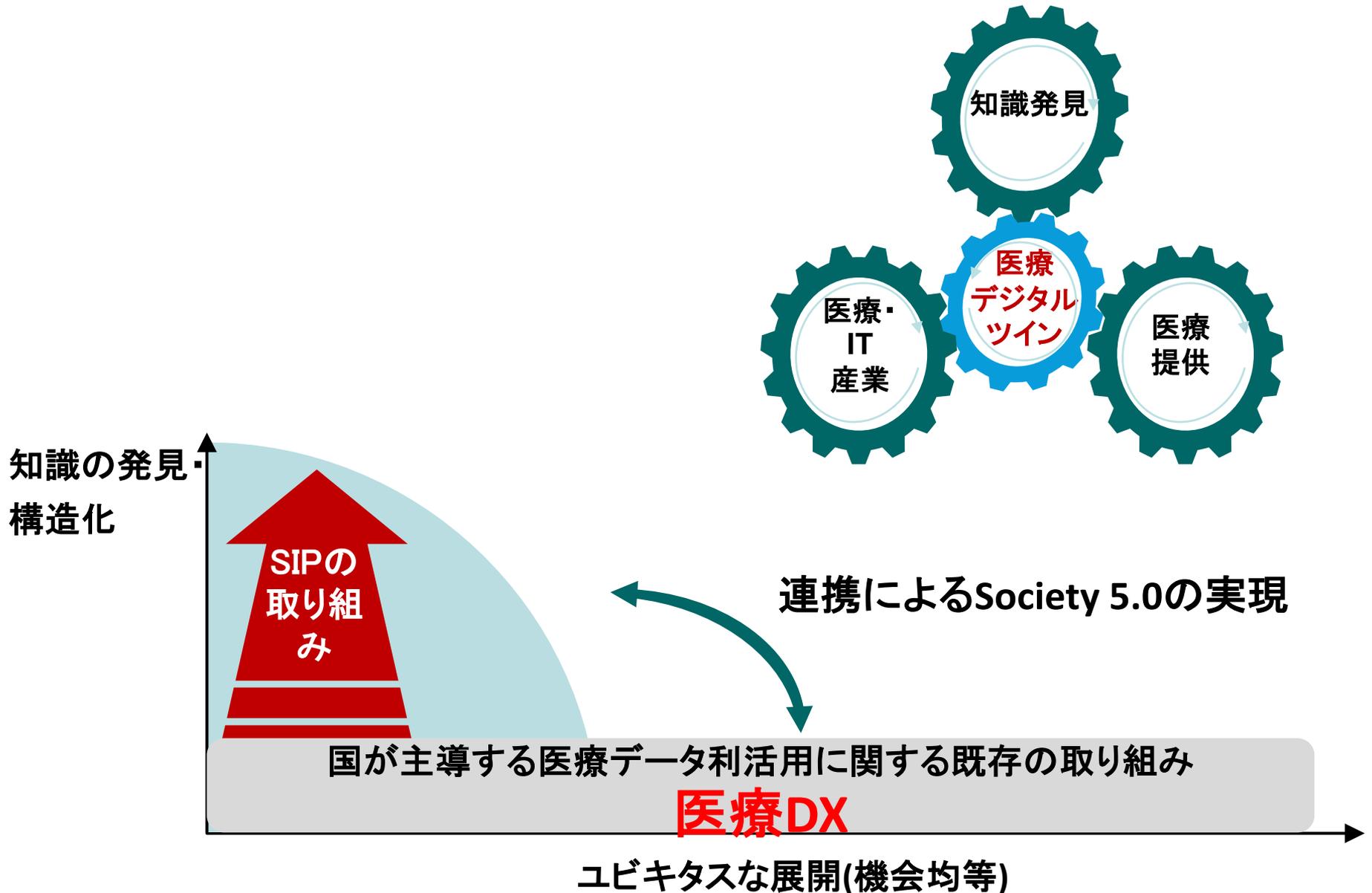
主に医療データを活用した研究

医療高度化に資するPHRデータ流通基盤構築事業
医工連携・人工知能実装研究事業 / 循環器疾患・
糖尿病等生活習慣病対策実用化研究事業 / 肝炎
等克服実用化研究事業 / 免疫アレルギー疾患実用
化研究事業 / 移植医療技術開発研究事業 / 障害
者対策総合研究開発事業(精神障害分野、身体・知
的・感覚器障害分野) / 女性の健康の包括的支援
実用化研究事業 / 「統合医療」に係る医療の質向
上・科学的根拠収集研究事業 / メディカルアーツ研
究事業

成育疾患克服等総合研究事業

予算額
27億円

第3期SIP: 統合型ヘルスケアシステムの構築



第三期SIP 統合型ヘルスケアシステムの構築 – 背景・概要

本SIPは、わが国の電子カルテEHRや個人電子カルテPHRを、HL7 FHIR準拠で標準化するための基盤構築をめざす。同時に、個人の健康医療情報を時系列で収集するシステムを構築する(ライフレコードデジタルツイン)。ライフレコードデジタルツインは、個人の健康管理システムとして社会実装するが、あわせて複数の基幹病院の臨床データを収集し、疾患(とくにがんと循環器疾患)の可視化と、個別化(層別化)医療等を可能とするためのシステムである。また、医療データの標準化システムおよび収集・分析システムの開発やビッグデータの二次利用により、医療の質の改善、患者の医療機関への受診支援、医師による鑑別診断、治療方針の決定、予後の予測、看護業務の支援、医薬品・医療機器等の評価などを容易に行えるシステムとする。さらに医薬品・医療機器・SaMD・医療ITなどの開発支援と医療政策支援も行う。これらのシステムを、自走性のあるシステムとして社会実装するとともに、本領域の若手人材を育成する。これにより、わが国の医療版Society 5.0を推進する。

基本的な考え方

- データの標準化、統合、収集
- 知識の発見と医療の質向上
- 受診・診療の効率化
- 医療行政への貢献

AIと著作権に関する考え方について

令和6年3月15日

文化審議会著作権分科会法制度小委員会

著作権法では、一定の場合には権利者の許諾を得ることなく著作物等を利用できる旨の、権利制限規定を設けている。

著作物に表現された思想又は感情の享受を目的としない利用(法第30条の4)や電子計算機による情報処理及びその結果の提供に付随する軽微利用等(法第47条の5)のほか、私的使用のための複製(法第30条)、引用(法第32条第1項)、学校その他の教育機関における複製等(法第35条)、営利を目的としない上演等(法第38条)といったものなどがあり…

- この文書(「本考え方」)は、生成 AI と著作権に関する考え方を整理し、周知すべく、文化審議会著作権分科会法制度小委員会において取りまとめられたものである。
- 本考え方は、その公表時点における、本小委員会としての一定の考え方を示すものであり、本考え方自体が法的な拘束力を有するものではなく、また現時点で存在する特定の生成 AI やこれに関する技術について、確定的な法的評価を行うものではないことに留意する必要がある。
- 今後も、著作権侵害等に関する判例・裁判例をはじめとした具体的な事例の蓄積、AI やこれに関連する技術の発展、諸外国における検討状況の進展等が予想されることから、引き続き情報の把握・収集に努め、必要に応じて本考え方の見直し等の必要な検討を行っていくことを予定している。

電子計算機による情報処理に対する規制緩和

製薬企業の要望

企業による医療データ活用

EU 創薬開発や安全性監視目的の活用は公衆衛生例外に含まれ使用可能

日本 学術研究目的での医療データ活用は可能（ただし研究機関を要する大学や組織に限定）

- 創薬開発や安全性監視については、オプトインではなく、オプトアウト対応を考慮すべき
- 公的データも匿名化ではなく、仮名化情報を使用可とすべき（追跡性、悉皆性）
- EUでは死亡者のデータはオプトアウト確認不要。
- 国内ではオプトアウトの確認ができないため、オプトアウトの意思確認が必要な場合には死亡者のデータを活用できない。将来的にはマイナポータルへの連絡を活用

IT事業者の要望

- データを取得した側の各サービス・基盤がお互いにデータ連携できるように規格の統一が必要

マイナポータル、PHR、オンライン診療、SaMD など

統合型ヘルスケアシステムの構築 全体構成

医学知識発見デジタルツイン

サブ課題A

研究開発支援・知識発見ソリューションの開発

- A-1 臨床情報プラットフォーム構築による知識発見 拠点形成
- A-2 臨床情報プラットフォームを活用したPHRによる 突然死防止・見守りサービス
- A-3 臨床情報プラットフォームと連携したPHRによる ライフレコードデジタルツイン開発

医療実践・患者支援デジタルツイン

サブ課題B

患者・医療機関支援ソリューションの開発

- B-1 がん診療についての統合的臨床データベースの社会実装
- B-2 電子問診票と個人健康情報（PHR）を用いた 受診支援・電子カルテ機能補助システムの開発
- 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
- B-4 看護師支援・医療の質向上（データに基づく看護師支援）：患者の生活のリアルタイム可視化によるインシデントリスク判定アルゴリズムの自動アップデートシステム及び自動服薬管理システムの開発
- B-5 医療機器・材料のトレーサビリティデータ収集・分析システムの構築を通じた医療機器開発・改良支援、医療資源最適化、病院経営最適化支援

地域医療デジタルツイン

サブ課題C

地方自治体・医療介護政策支援ソリューションの開発

- C-1 地方自治体の意思決定支援システム開発による、住民の医療資源アクセスと提供体制の最適配置・財源確保、地域共生社会のための安全ネットワークの実現
- C-2 患者の疾患状態及び施設間動態の可視化を可能とするシステム開発による地域医療構想の実現

サブ課題D

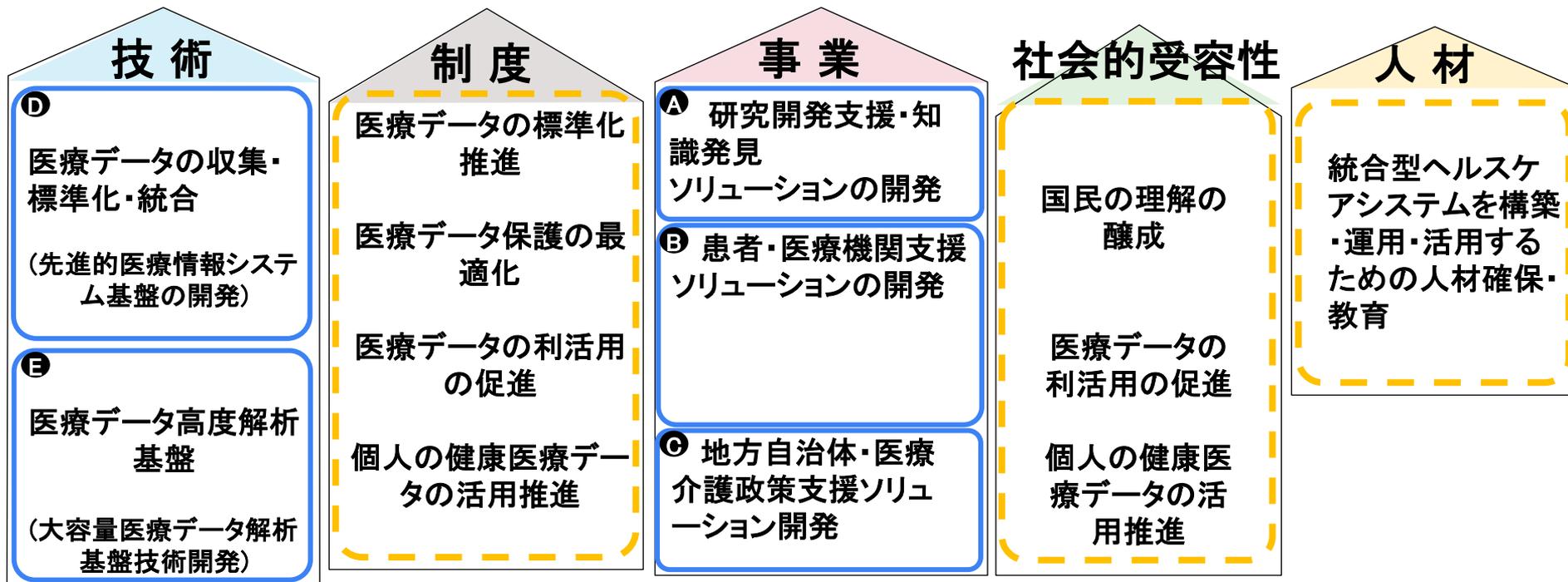
デジタルツインのための先進的医療情報システム基盤の開発

- D-1 医療機関・ベンダー・システムの垣根を超えた医療データ基盤構築による組織横断的な医療情報収集の実現
- D-2 統合型の医学概念・知識連結データベースの構築及び医療文書の自動分析基盤の整備
- D-3 僻地診療支援のためのクラウド型標準電子カルテサービスの研究開発

サブ課題E

大容量リアルタイム医療データ解析基盤技術の開発

- E-1 大容量医療データの高速処理・高効率管理・高次解析基盤技術の開発
- E-2 大規模医療文書・画像の高精度解析基盤技術の開発



生成AI時代の新たな課題

- 生成AI(大型言語モデル、大型マルチモードモデル)には大量のデータが必要
- AIの結果に対する検証システムにも大量データが必要
- AI開発に利用する医療情報の素性を明らかにする必要
(どこからどのように得た情報か、信頼性、同意、など)
- 医療情報の収集と活用の国際協調(入口規制から出口規制へ)

テーマ 1 医療LLM基盤の研究開発・実装

概要

- 世界的な生成AIの急速な進展により、医療に関する様々な業務をこれまでにはなかった方法で実現できる可能性がある。我が国においても国内の研究者等によりLLM開発が進展している。海外では大手テクノロジー企業が相次いで医療分野に乗り出しており、GPT-4などの汎用モデルの医療分野での利用に加え、Med-PaLM2など医療特化型のLLMも登場している。しかしながら、産業政策・競争政策・および国民に対するサービス品質の向上等の視点からは、**日本人の医療データを日本語で学習したLLMの開発に取り組む必要がある**。このため、**医療用の日本語版LLM基盤の構築を図る**。

テーマ 2 医療LMMの研究開発・実装

概要

- LLMについては、既に民間レベルでの開発が進んでいる。一方、画像等の言語以外の多様な情報を組み込んだ大規模マルチモーダルモデル (LMM) は「Med-PaLM Multimodal」などが登場しているが、未だ世界的に確立されていない。我が国の医療分野においてはLMMの学習に利用できる良質な医療データが存在していることから、**世界的な競争力の獲得も視野に入れて、医療用の日本語対応LMM基盤の構築を図る。**

テーマ 3 ソリューション/アプリケーション開発

概要

- 医療LLM/LMM基盤を構築し、社会実装につなげるためには効率的な構築手法の開発や構築されたLLM/LMMを活用したユースケースの具体化・実装が必要となる。そのため、**医療LLM/LMMの開発に資するソリューション開発、診療支援、保健行政支援、臨床研究支援など様々な業務の推進を図るLLM/LMMを活用したアプリケーション開発を行う。**

テーマ 4 医療データ基盤構築と運営手法の検討

概要

- 高精度な医療LLM/LMMを開発するためには、良質な医療データを継続的に確保し、それを研究者や開発者が利用できる仕組みが必要となる。一方、個人情報保護、倫理などデータ保護の面での条件が満たされた学習データの整備など、個々の研究開発事業でデータを確保するには様々な課題がある。そのため、**多様な医療データの収集・蓄積と医療LLM/LMMの研究開発主体による活用を可能とする医療データ基盤を構築する。**

本医療データ基盤の設計思想

本公募の医療データ基盤は、意味づけされた医療データを大量に確保し、使いやすい形に加工する

本医療データ基盤で想定するデータ蓄積形式

生成AIモデル開発では、テーブルデータではなく、大量のテキストデータと意味づけされたマルチモーダルデータの確保が重要

- **データの前処理**：文法上の誤りや重複画像の排除など、データ品質の担保が重要
- **データの蓄積**：異なる医療機関のデータを個別に蓄積可能
- **データの加工**：各データに対する意味合いの付与が重要（画像に対する疾患情報のアノテーション等）

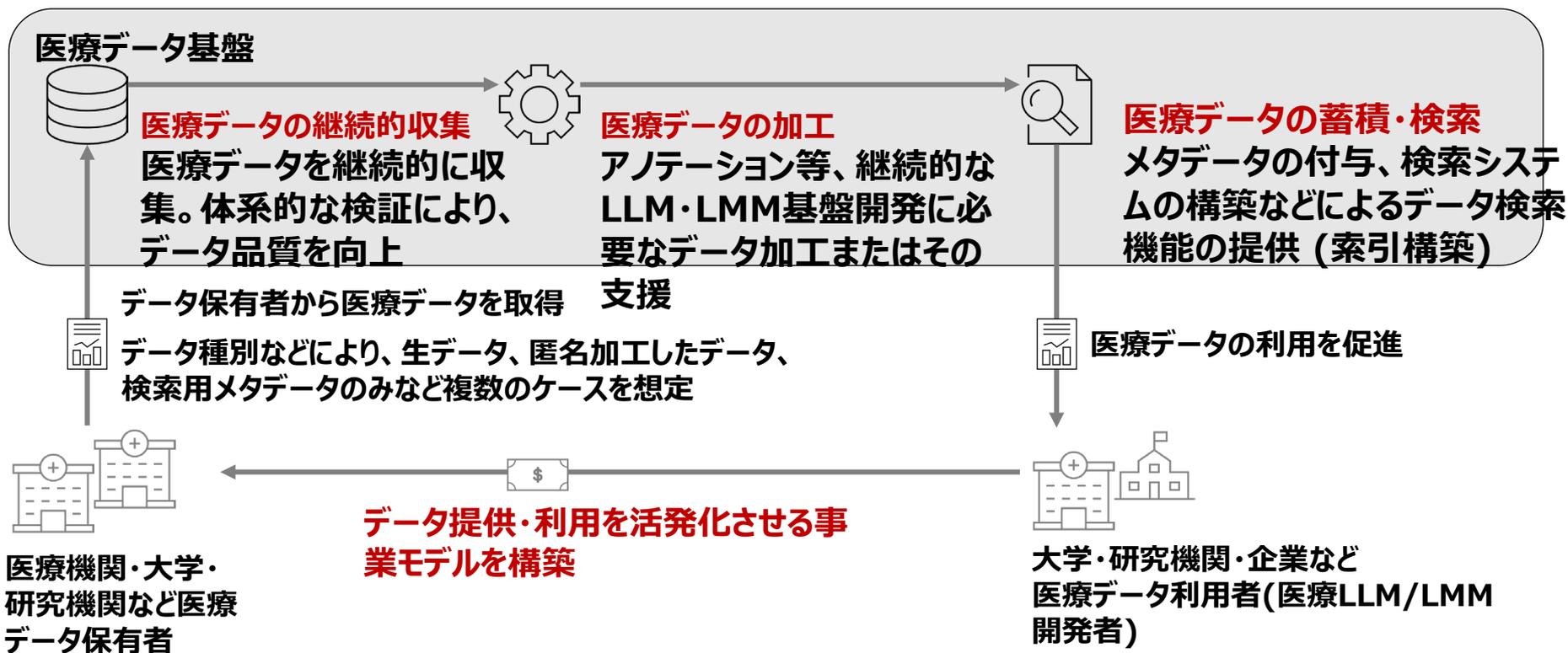
本医療データ基盤で想定する機能

LLM/LMM開発に使えるデータを継続入手し、多数の研究者が利活用できるよう、学習用の加工を実施・支援し、検索できる形で蓄積

- **大量に多様なデータを入手**：医療機関、他の医療データベース等からテキスト、画像・動画、テーブルデータなど幅広く収集
- **学習データに加工**：最新技術と業界知見に基づくテキスト整形や重複データ削除、アノテーションの実施・支援等
- **データ検索機能を提供**：メタデータ等を活用。医療データ基盤で常時保有しないものを含めデータカタログを提供
- **開発するLLM・LMMの管理**：さまざまなバージョンのモデルを管理

医療データ基盤のコンセプト概要

本医療データ基盤は、「医療データの継続的収集・加工・蓄積」を一貫して担う。さらに索引機能を構築し、生成AIモデル開発者にとって使いやすい仕組みとする



疾患基礎研究の振興

- 医療研究開発研究の基盤
- マウスからヒトへの体制づくり、基盤整備、意識改革が必要
ゲノム、細胞機能、症状、身体所見、検査所見、画像、治療への反応、予後、社会的インパクトを統合する時代
- 個別研究機関では達成できない規模の研究
- 研究はひらめき(セレンディピティ)と努力の積み重ね。交流、武者修行、組織の流動性、研究の精神の育成を推進する事業費
- 大学病院のシステム改革

大学設置基準(1956)で、大学病院は「学部の教育研究に必要な病院」として位置づけ。大学病院の機能が高度化したがる、その診療は、教育研究要員を振り向けざるをえないことが、研究力低下の最大の要因。