

千年カルテ2024



Life Data Initiative 代表理事



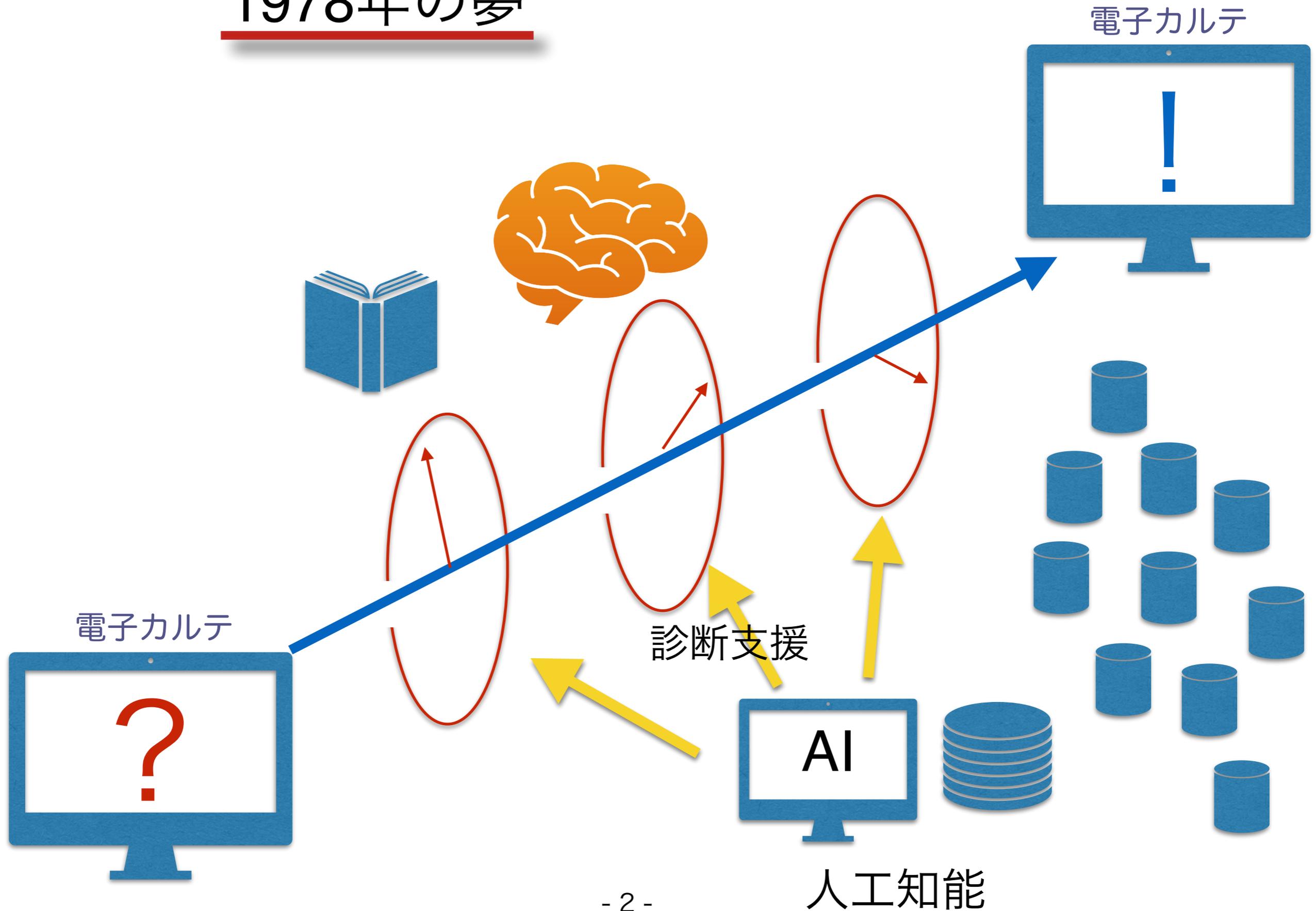
京都大学名誉教授



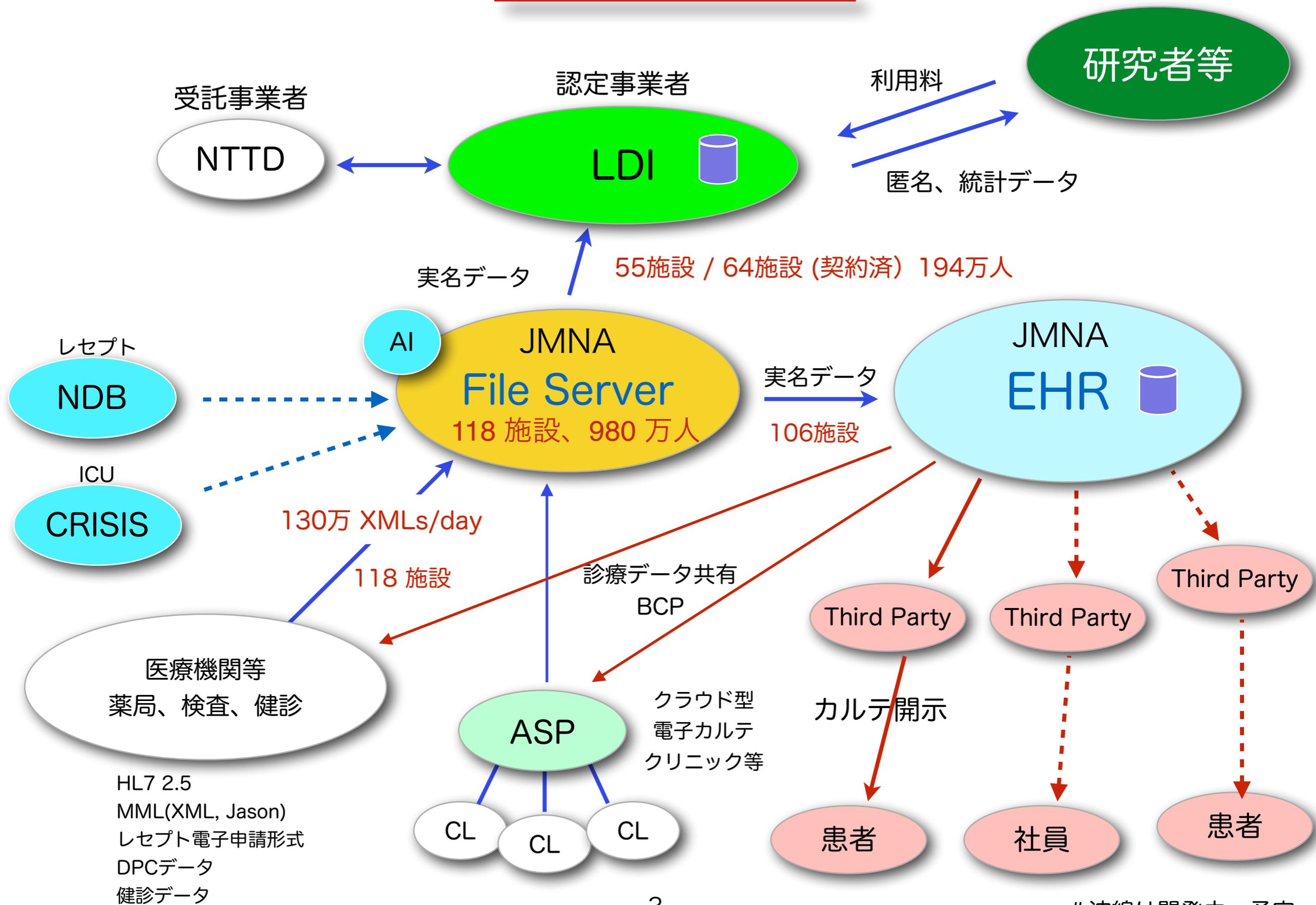
宮崎大学名誉教授

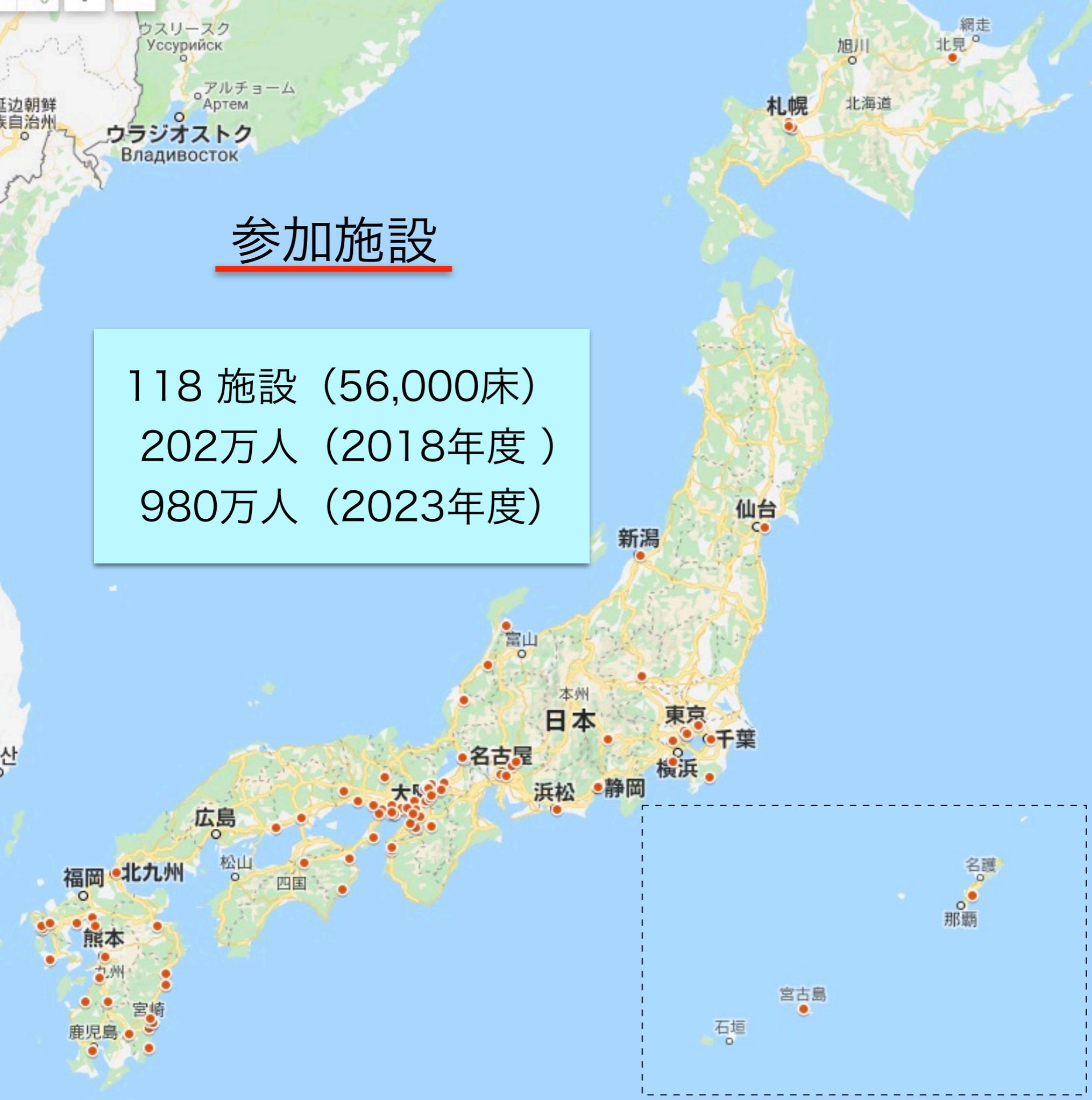
吉原博幸

1978年の夢



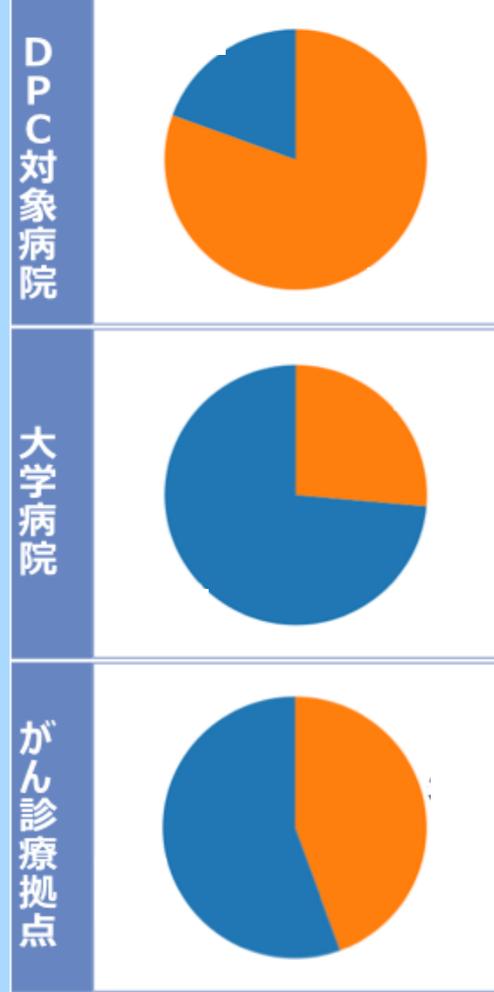
千年カルテ2024





参加施設

118 施設 (56,000床)
 202万人 (2018年度)
 980万人 (2023年度)



千年カルテのUserインターフェイス

DEMO

EHR3 on Cache (IRIS)

Brave ファイル 編集 表示 履歴 ブックマーク プロフィール タブ ウィンドウ ヘルプ

MML1次利用 gEHR 患者リスト

https://front.gehr.jp/csp/MML10/MML10WebEHRPtList.CSP

Home JMS GM FB 京都大学医の倫理委... CAL G drv EHR3アクセス WebMTG ナギフォルダ Active Improvisation by... LDI PM 千年カルテ

戻る

1人を表示します

選択子	施設	masterId	氏名	カナ氏名	fullname	性別	生年月日	年齢	死亡日	ZIP	住所	電話番号
選択	京都大学	06363448	吉原 博幸	ヨシハラ ヒロユキ		男性	1949(昭和24)-03-12	72歳		606-0803	京都府京都市左京区下鴨下川原町	(自宅) 075- (昼) 075-

Mac OS dock with various application icons including Finder, Music, Photos, Safari, and Microsoft Office.

EHR4 on Google Cloud

千年カルテ



患者検索

吉原

診療記録一覧

検査

処方

注射

報告

病名

経過記録

臨床サマリ

手術記録

検索開始日 1900-01-01 検索終了日 2099-12-31

検歴一覧

病名一覧

ユーザ管理

表示施設選択

京都大学医学部附属病院

施設患者番号

06363448

患者氏名

生年月日

検索

施設名	施設患者番号	患者氏名	患者カナ氏名	性別	生年月日
-----	--------	------	--------	----	------

No data available

Rows per page:

5

-

<

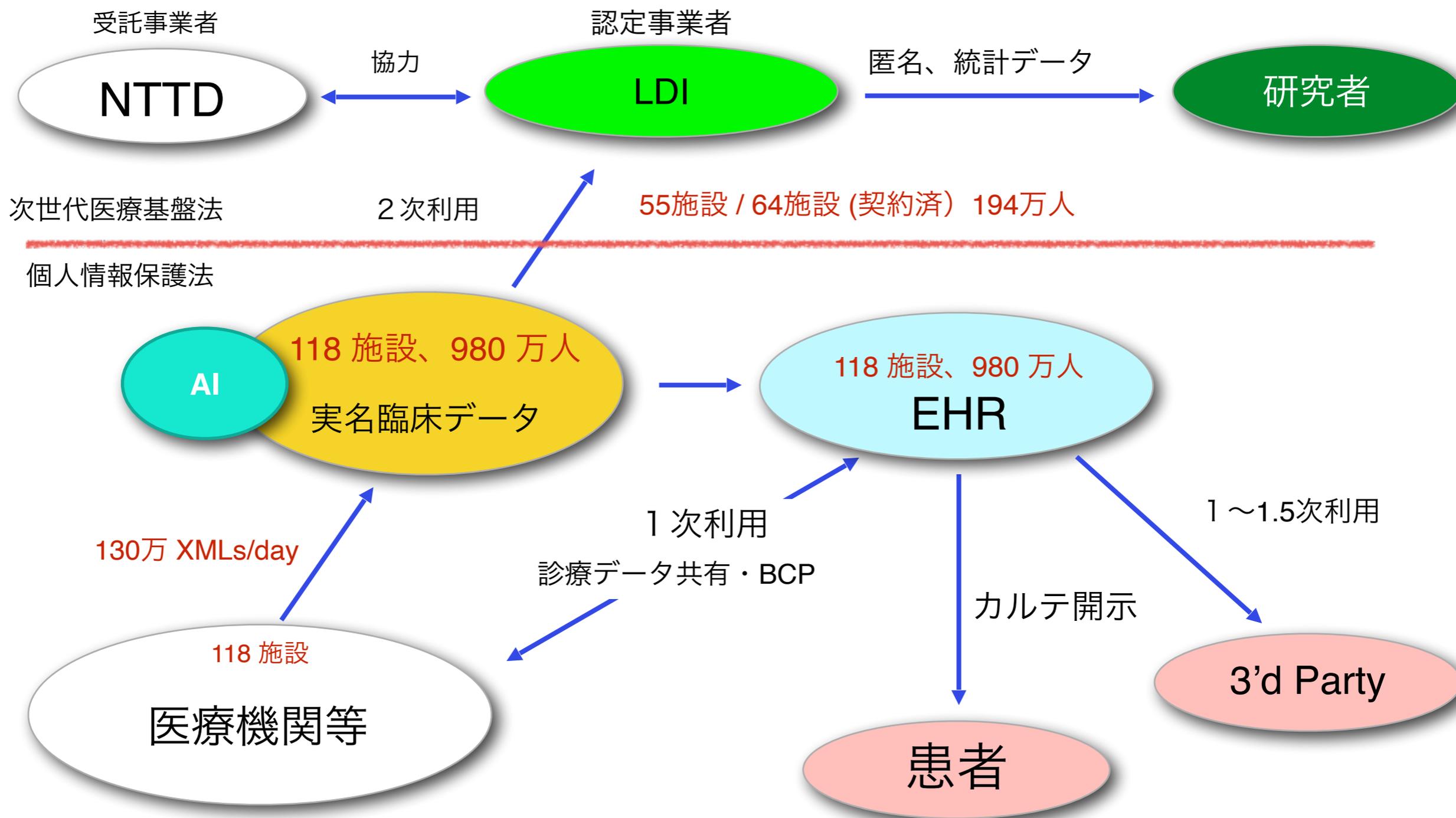
>

患者用アクセス

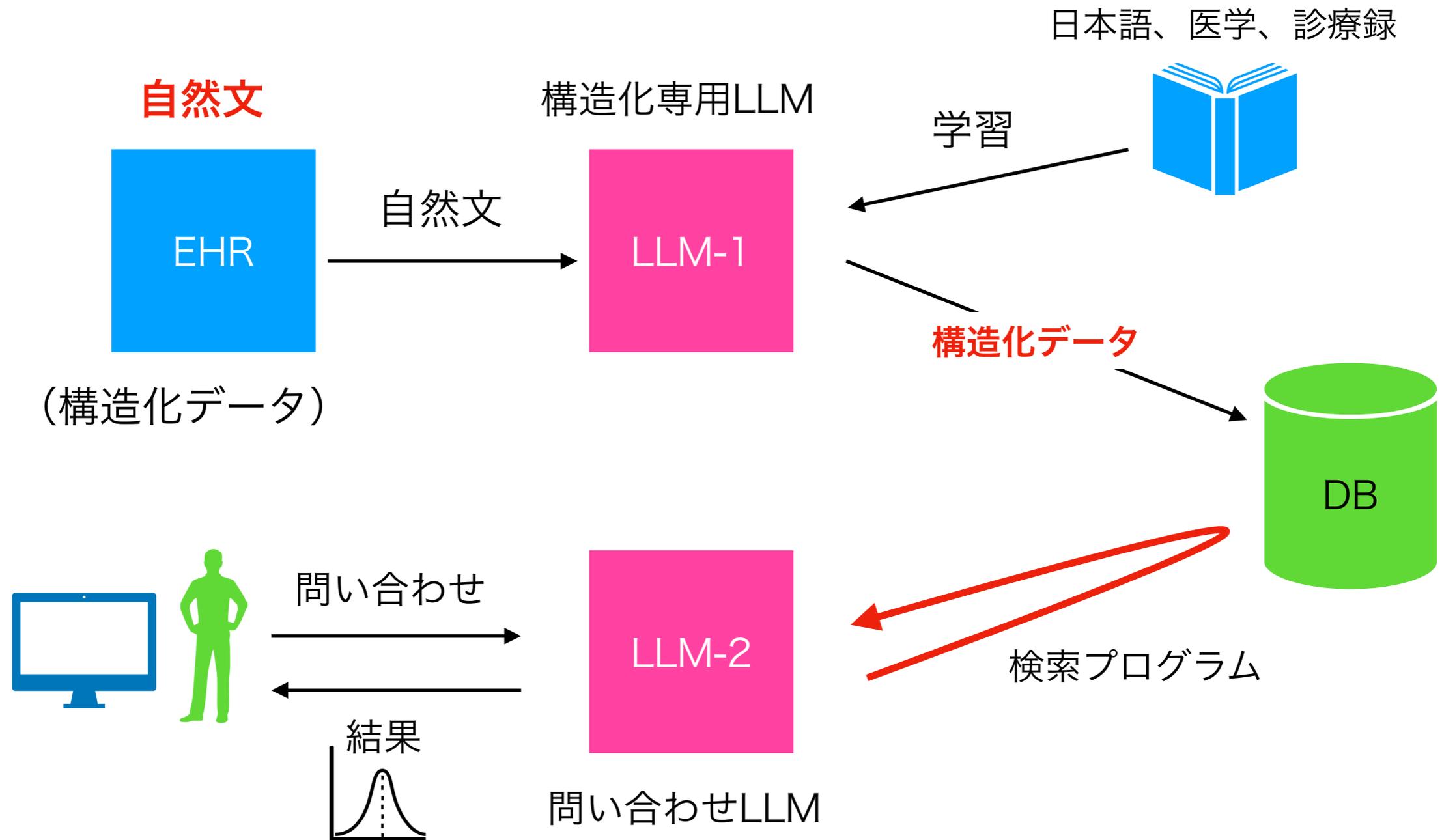
NOBORI App



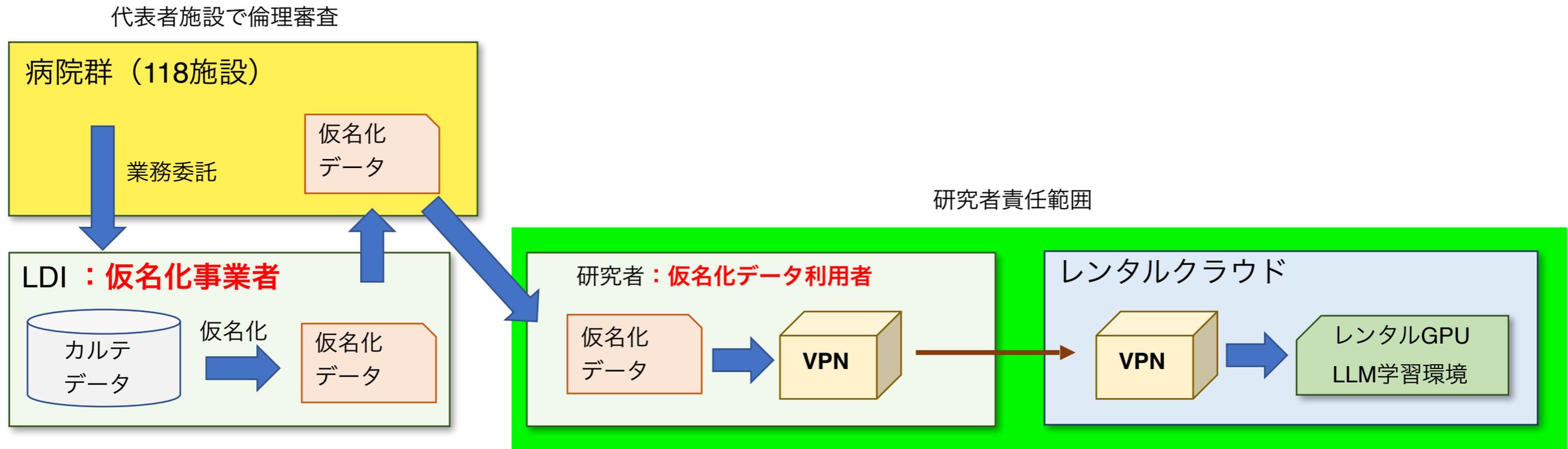
千年カルテ2024



EHRデータの知識化



個人情報保護法配下での医療データを用いたLLM強化学習



1) 個人情報上の整理

学習用の診療情報 (病院→LDI→京大等)

- ・ 研究に使用する学習用のデータは、LDIに保管されているものを使用
- ・ LDIで加工する際に、**カルテの自由記載から個人を特定する情報を完全に排除することは困難**
- ・ LLM学習用の診療情報は、次世代医療基盤法の匿名加工医療情報として扱うことはもちろんのこと、仮名加工医療情報として扱うことも難しい
- ・ 従って、**個人情報保護法上の学術研究の規定**を使う
病院から研究機関に個人情報である診療情報を提供
研究機関が学術研究の範囲で利用する
LDIによる診療情報加工により学習に不要な項目を削除
→漏えい等が発生した場合のリスクを低減するための措置

カルテを学習したLLM (LLM-1)

- ・ 学習用の診療情報を学習したLLMそのものは、学習用モデルとパラメータにすぎないので、個人情報には該当しない。
- ・ LLM-2も同様にそれ自体は個人情報ではない。

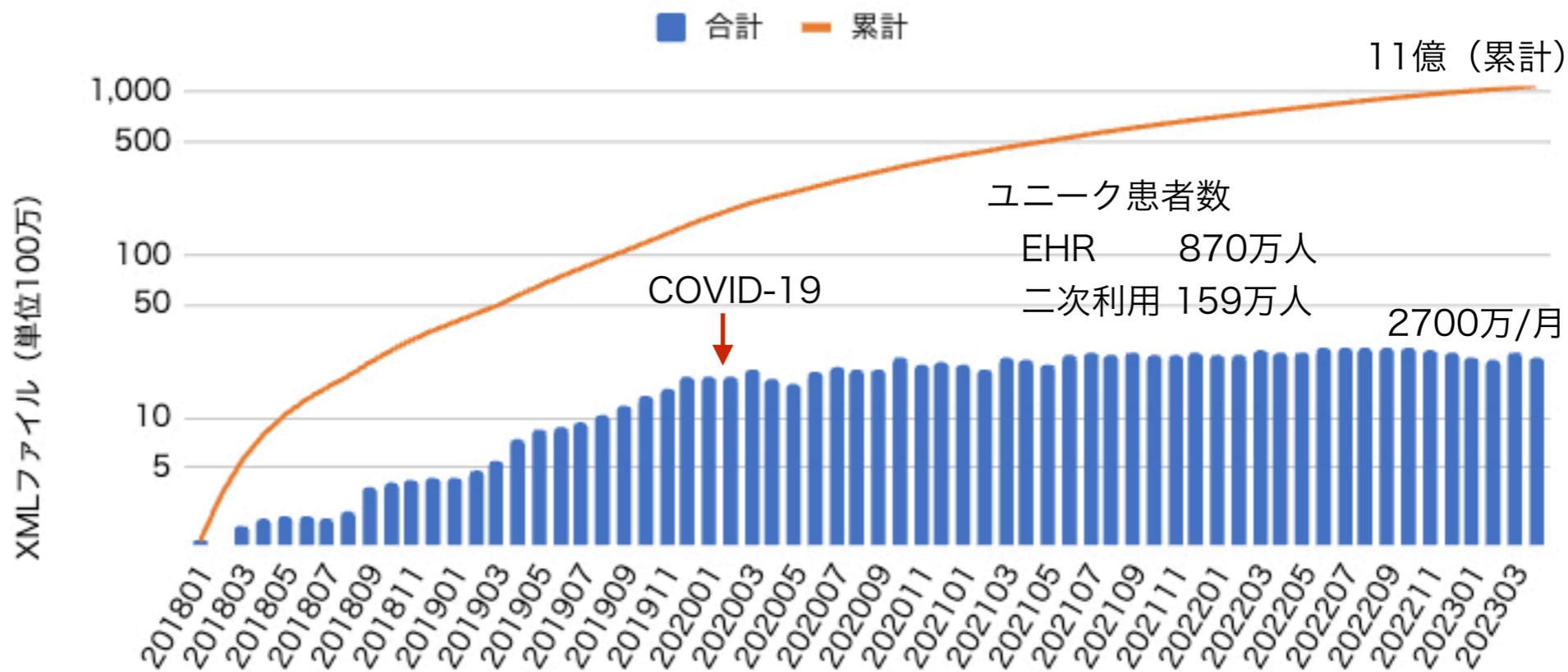
2) 当事者の契約関係

- 1 病院・LDI間の加工委託
- 2 病院・研究機関間の診療情報提供契約→検討着手
(病院と代表研究機関である京大)
- 3 研究機関間の共同研究契約
(厚労科研参加大学、会社)
- 4 研究機関と産総研の間の計算機資源の利用に関する契約

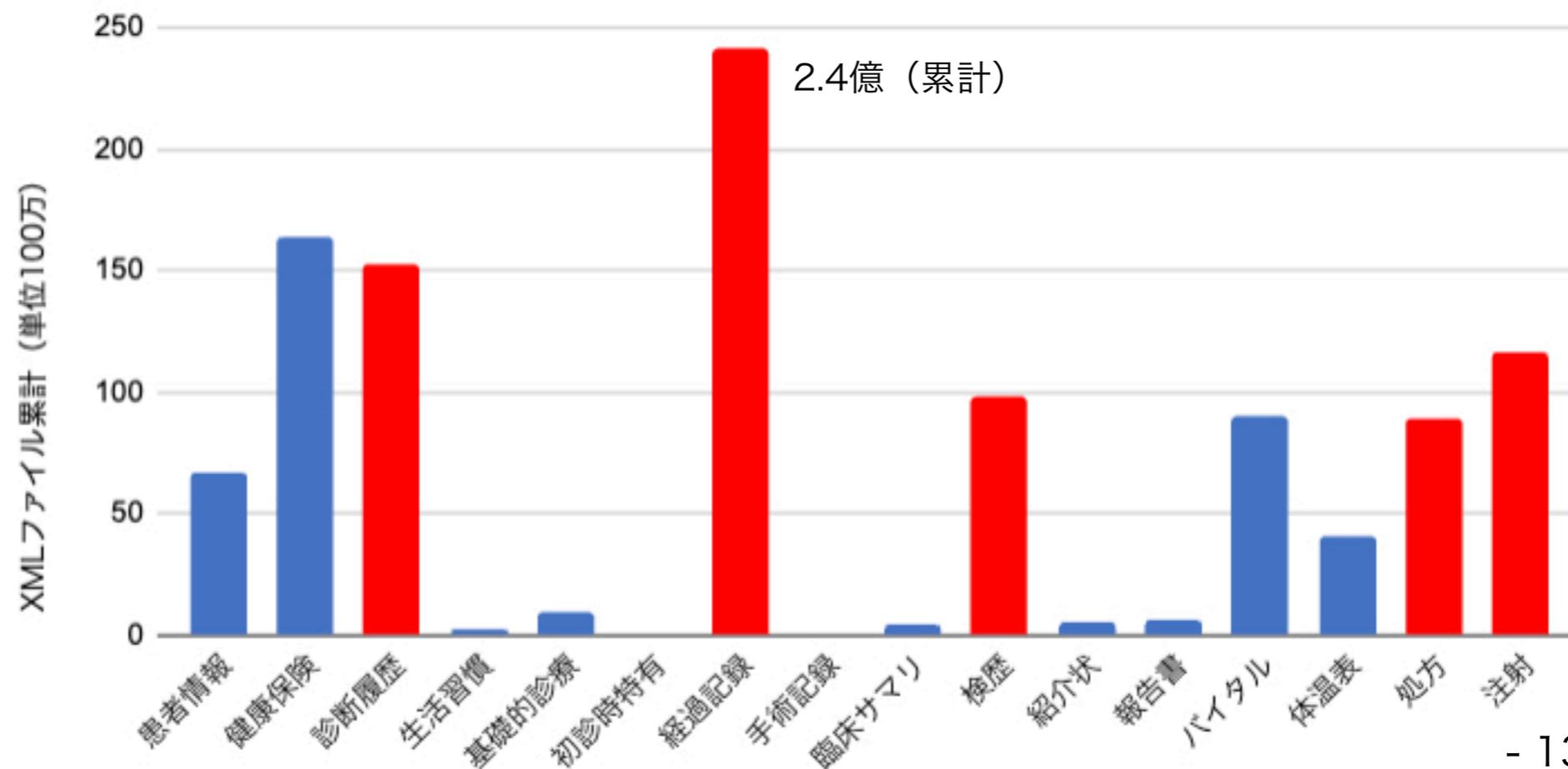
千年カルテへのXMLデータ出力状況

2018.01～2023.04

XML出力数（月次合計と累計）



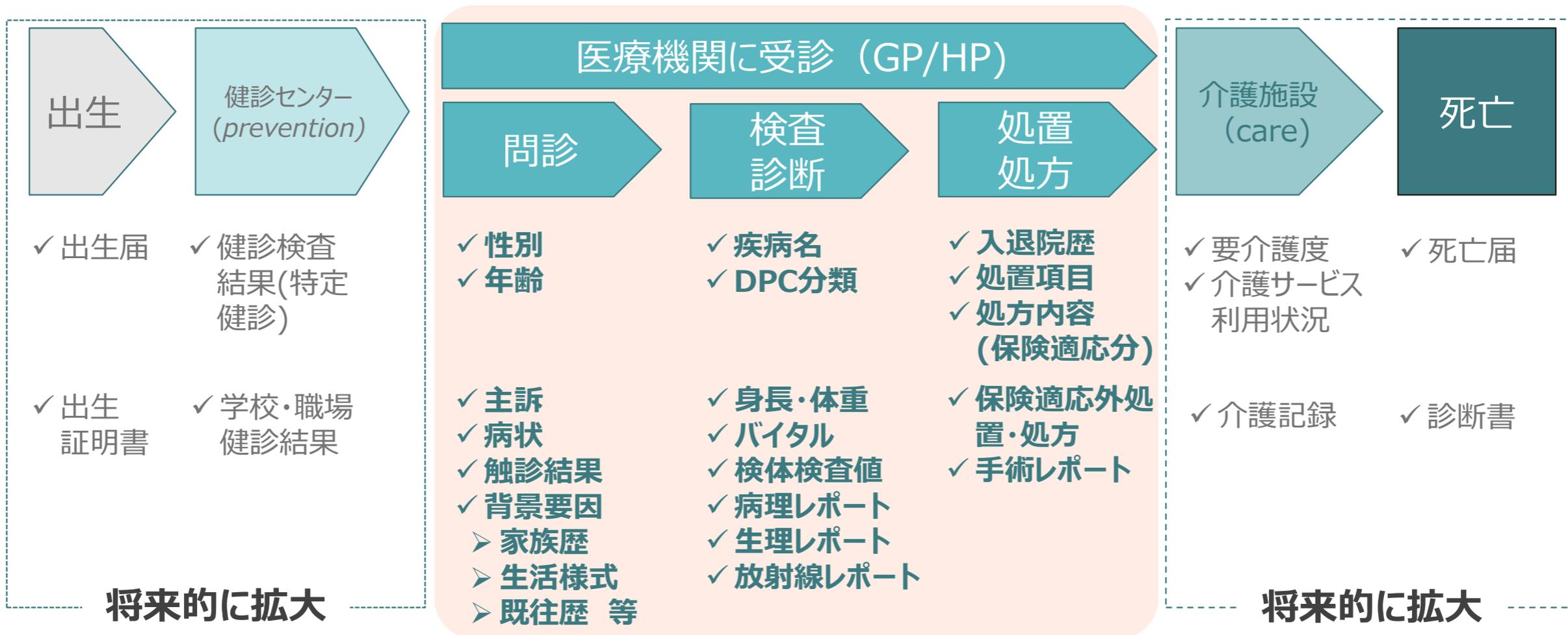
出力XMLファイル累計（種別）





千年カルテが目指す医療情報群

千年カルテは、患者さまの一生涯に渡る健康・医療情報の収集を目指し、高齢社会に向けた健康長寿社会の実現に向けて取り組んでまいります。



現在、政府も次世代医療基盤法の普及・拡大に向け自治体などを中心に情報の連携並びに市民サービスの向上に向け取り組みを始めています。



想定される活用シーン・方法

	分析的研究		記述的研究
アカデミア	<ul style="list-style-type: none"> 臨床試験の代替手段 <ul style="list-style-type: none"> エビデンス創出 適応拡大探索 外部対照群としての利用 評価基準開発 医療経済性評価 	<ul style="list-style-type: none"> クリニカルパス検討 フォーミュラリ検討 	<ul style="list-style-type: none"> 疾患／疾患治療実感調査 アンメットニーズ探索
製薬企業／医療機器等		<ul style="list-style-type: none"> 治験計画検討 製造販売後データベース調査 画像情報によるAI分析（準備中） 	

■ 二次利用実績（利用目的等審査委員会承認実績） 33件

（2023年12月時点）

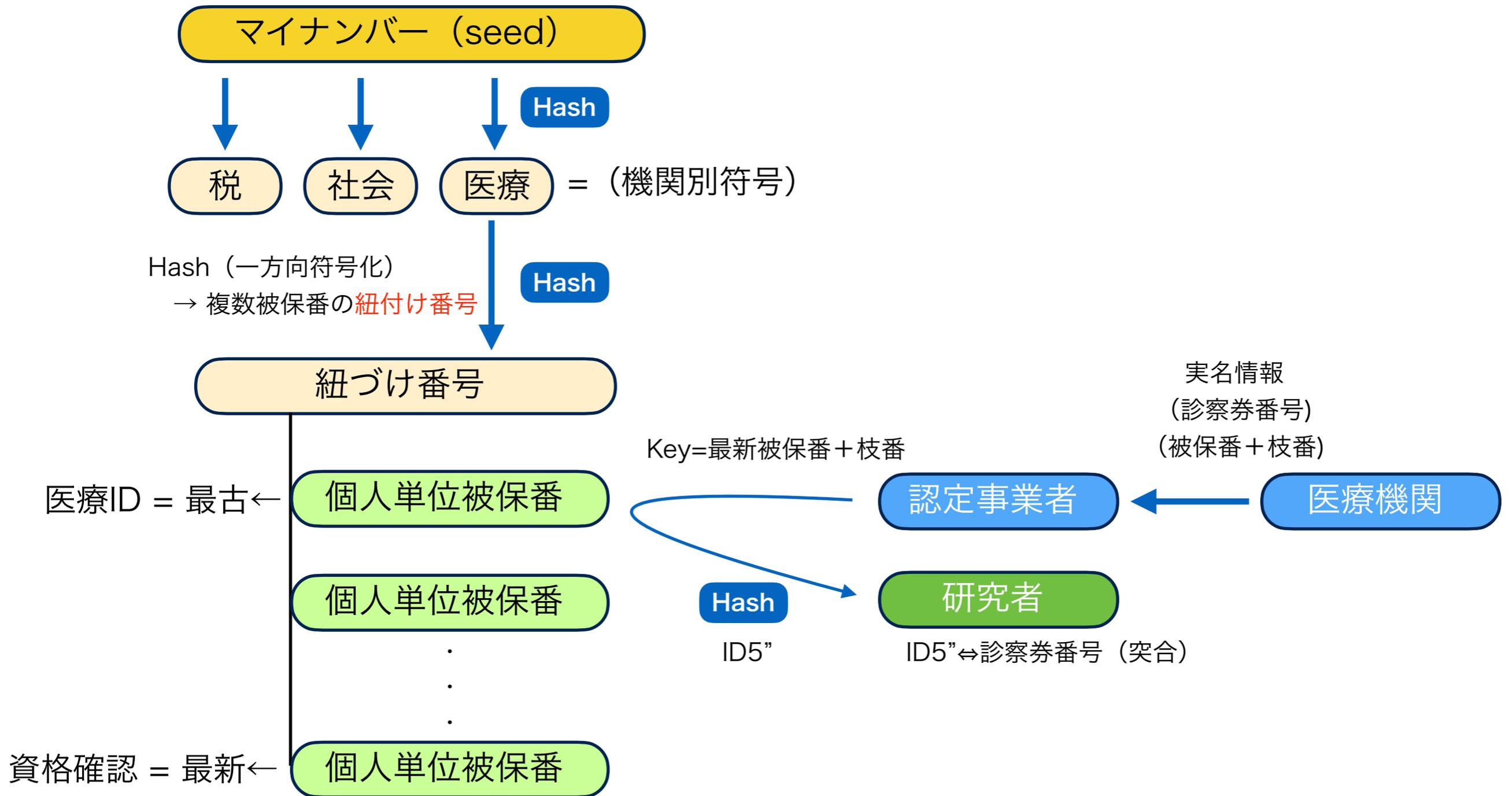
No.	承認日	課題名	活用データ項目			活用者区分
			電子カルテデータ	DPC調査データ	レセプトデータ	
1	2020年10月20日	乳癌のサブタイプ別、治療実態を探るための千年カルテデータのFeasibility	●	●	●	アカデミア
2	2020年10月20日	がん患者の臨床アウトカムにおけるEHRデータベースを用いた評価方法の後ろ向き研究	●	●	●	民間企業
3	2021年1月15日	自己免疫疾患領域における寛解指標のフィージビリティ確認	●	●	●	民間企業
4	2021年3月5日	検査値等を用いたウイルス性肝炎患者研究のフィージビリティスタディ	●	●	●	民間企業
5	2021年5月26日	検査項目の多施設突合手法開発を目的とした研究	●	-	-	アカデミア
6	2021年7月15日	非構造化データの評価方法確立を目的とした研究	●	-	-	民間企業
7	2021年7月15日	希少疾病領域における症状把握を目的としたフィージビリティ検証	●	●	●	アカデミア／民間企業
8	2021年7月15日	乳がんデータ項目に関するフィージビリティ調査	●	●	●	民間企業
9	2021年8月31日	匿名加工医療情報のAI研究への利活用可能性の検討	●	●	●	アカデミア
10	2021年9月28日	心不全データベース研究のためのFeasibility調査	●	●	●	民間企業
11	2021年10月26日	感染症に対するTreatment flow 及び 関連医療費の推計	●	●	●	民間企業
12	2021年10月26日	がん患者の臨床アウトカムにおけるEHRデータベースを用いた評価方法の後ろ向き研究—自然言語解析—	●	●	●	民間企業
13	2021年11月30日	肺がん・乳がん患者の治療実態把握及び病気の進展に関する因果探索	●	●	●	民間企業
14	2021年11月30日	電子カルテのテキストを活用したRECIST評価の辞書作成	●	●	●	アカデミア
15	2022年3月8日	電子カルテ情報を活用した、臨床試験の新規手法論開発	●	●	●	民間企業

■二次利用実績（利用目的等審査委員会承認実績） 33件

（2023年12月時点）

No.	承認日	課題名	活用データ項目			活用者区分
			電子カルテデータ	DPC調査データ	レセプトデータ	
16	2022年3月8日	電子カルテのテキスト情報を用いた癌患者の治療実態に関する検討	●	●	●	民間企業
17	2022年4月13日	希少疾病の罹患リスク予測モデル構築	●	●	●	民間企業
18	2022年5月18日	アウトカムバリデーションでのフィージビリティ研究	●	●	●	民間企業
19	2022年7月13日	千年カルテ二次利用データベースを利用した心不全患者の病態に対する因子探索	●	●	●	民間企業
20	2022年8月22日	電子カルテ情報を用いた有効性等に関する新規エビデンス創出の検討	●	●	●	民間企業
21	2023年3月13日	先天性代謝異常症患者の治療実態の把握	●	●	●	民間企業
22	2023年3月13日	消化管領域における治療実態調査	●	●	-	民間企業
23	2023年3月13日	感染症におけるTreatment flow及び関連医療費の推計	●	●	●	民間企業
24	2023年4月12日	がん患者の臨床アウトカムの薬剤群間比較におけるEHRデータベースを用いた評価方法の後ろ向き研究	●	●	●	民間企業
25	2023年4月12日	医学研究における匿名加工情報利用の最適化の検討	●	●	●	アカデミア
26	2023年5月10日	検査値の患者属性別統計	●	-	-	アカデミア
27	2023年8月24日	アウトカムバリデーションスタディの外挿可能性、代表性検討	-	●	-	民間企業
28	2023年9月27日	電子カルテを活用した病態進行を予測するAI開発	●	●	●	アカデミア/ 民間企業
29	2023年10月10日	血液がんにおける治療実態研究 （電子カルテ情報を活用した患者背景設定と臨床アウトカム評価）	●	●	●	民間企業
30	2023年10月10日	アレルギー免疫療法の投与継続理由及び中止理由の調査	●	●	●	民間企業
31	2023年10月10日	CKD・心不全関連疾患治療薬における治療継続期間毎のインサイト抽出	●	●	-	民間企業
32	2023年11月27日	リウマチ疾患、および抗リウマチ製剤による有害事象の検証	●	●	●	アカデミア
33	2023年11月27日	高齢心不全患者における診療ガイドラインに基づく標準的治療と再入院の関連	●	●	●	アカデミア

オンライン資格確認/国民医療ID 2023年



標準規格：文書別対応状況

2022.3現在

	HL7 Ver 2.x	HL7 Ver 3 CDA	HL7 FHIR	MML
患者情報	○	○	○	○
保険情報	○	x	2	○
病名	○	○	2	○
生活歴	x	△ (部分的)	1	○
アレルギー	○	○	1	○
経過記録	x	○	x	○
手術記録	x	x	x	○
臨床サマリー	x	○	○	○
検体検査	○	x	2	○
報告書*	x	○	○ (健診)	○
バイタルサイン	△	△ (部分的)	2	○
体温表	△	△ (部分的)	x	○
内服処方箋	○	x	○	○
注射	○	x	2	○
透析	x	x	x	○
処置	x	x	1	○

△：日本では未検証

数字は完成度 (x、1、2、○)

電子カルテベンダーからのデータ出力状況 (2022年5月)

ベンダー	レセプト	DPC	患者情報	病名	生活習慣	基礎的	初診時	経過	手術	サマリ	検歴	紹介	報告	バイタル	体温	処方	注射
A	○	○	○	○				○	○	○	○	○	○	○		○	○
B	○	○	○	○	○	○		○		○	○				○	○	○
C	○	○	○	○	○	○	○	○	○		○			○	○	○	○
D	○	○	○	○	○	○		○			○			○	○	○	○
E	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
F	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
G	○	○	○	○		○	○		○		○			○	○	○	○
H	○	○	○	○	○	○					○					○	○

データ形式はXML(MML)が主で、一部HL7v.2.5とCDA
 # 空白はデータ出力なし、またはPDF等でデータベース化困難

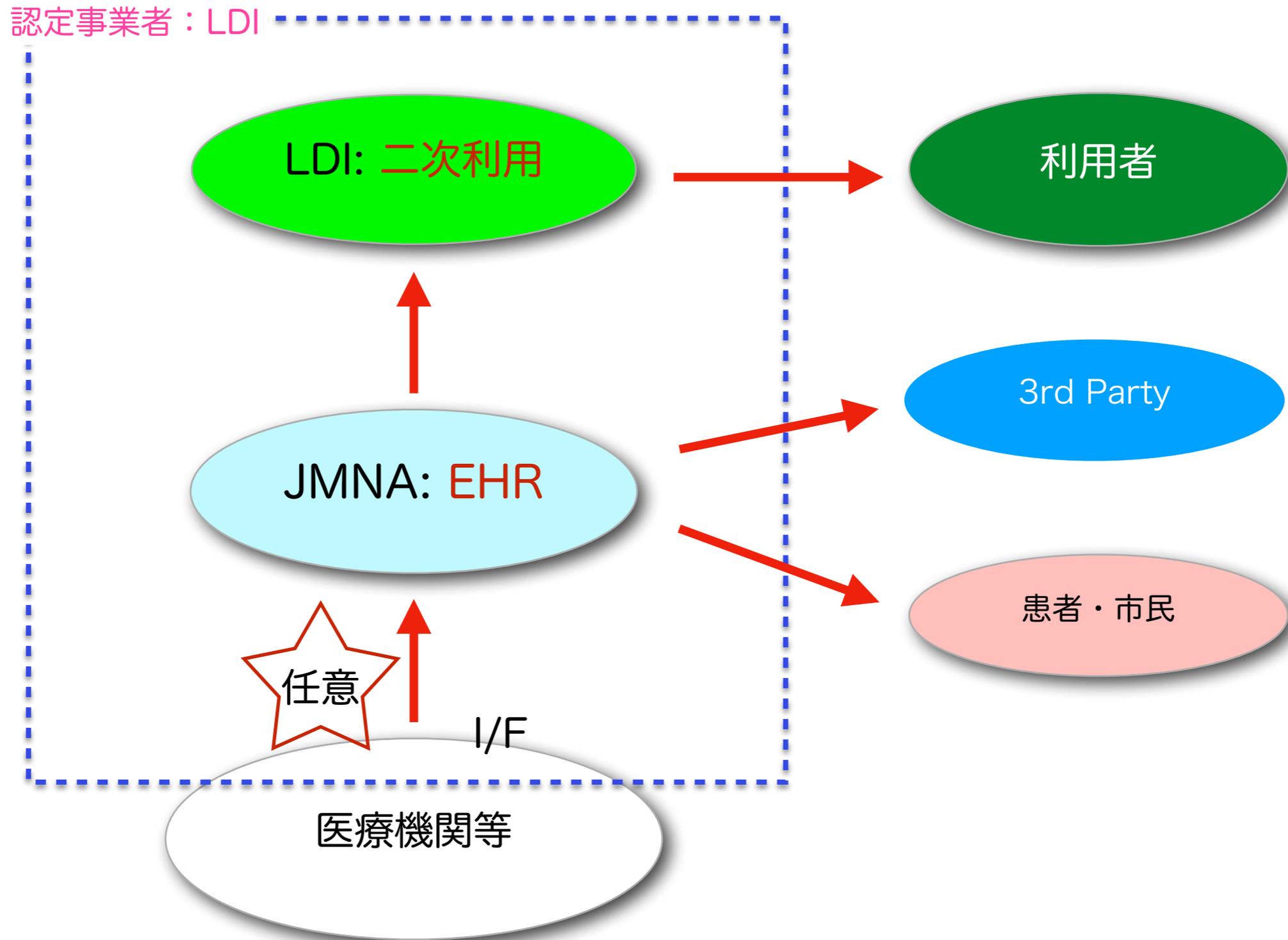
医療文書の**研究利用**状況（2023年3月）

研究者	レセプト	DPC	患者情報	病名	生活習慣	基礎的	初診時	経過	手術	サマリ	検歴	紹介	報告	バイタル	体温	処方	注射
1	○	○	○					○		○							
2		○	○	○	○			○	○	○					○	○	○
3																	
4		○	○	○				○		○	○						
5			○								○						
6		○	○							○							
7	○	○	○	○							○						
8		○	○	○				○	○	○			○				
9		○															
10		○	○	○				○		○	○		○		○		
11		○	○	○				○									
12		○	○	○	○			○	○	○					○	○	○
13	○	○	○	○							○				○	○	○
14		○	○	○						○							
15		○	○	○							○				○	○	○
16	○	○	○					○		○							
17	○	○	○	○							○				○		
18		○															
19		○	○	○							○				○		
20		○	○	○	○			○		○			○				

赤枠は研究での有用性を認めたデータ種別（アウトカムの記載等）

これまで

医療機関からのデータ取得も含めて
全て認定事業者が担当しているのが現状



認定事業におけるデータ収集

医療機関からのデータ取得と保存は国が行い、
認定事業者はサービスレイヤーに特化する。
これにより認定事業者の経営負担を軽減。
国レベルでのデータ収集は大幅に拡大。

国の運営

データ形式変換

国レベル
リポジトリ

義務化 データ抽出

医療機関等
電子カルテベンダー

認定事業者：LDI等

二次利用部門

匿名情報

研究者・企業

EHR部門

顕名情報

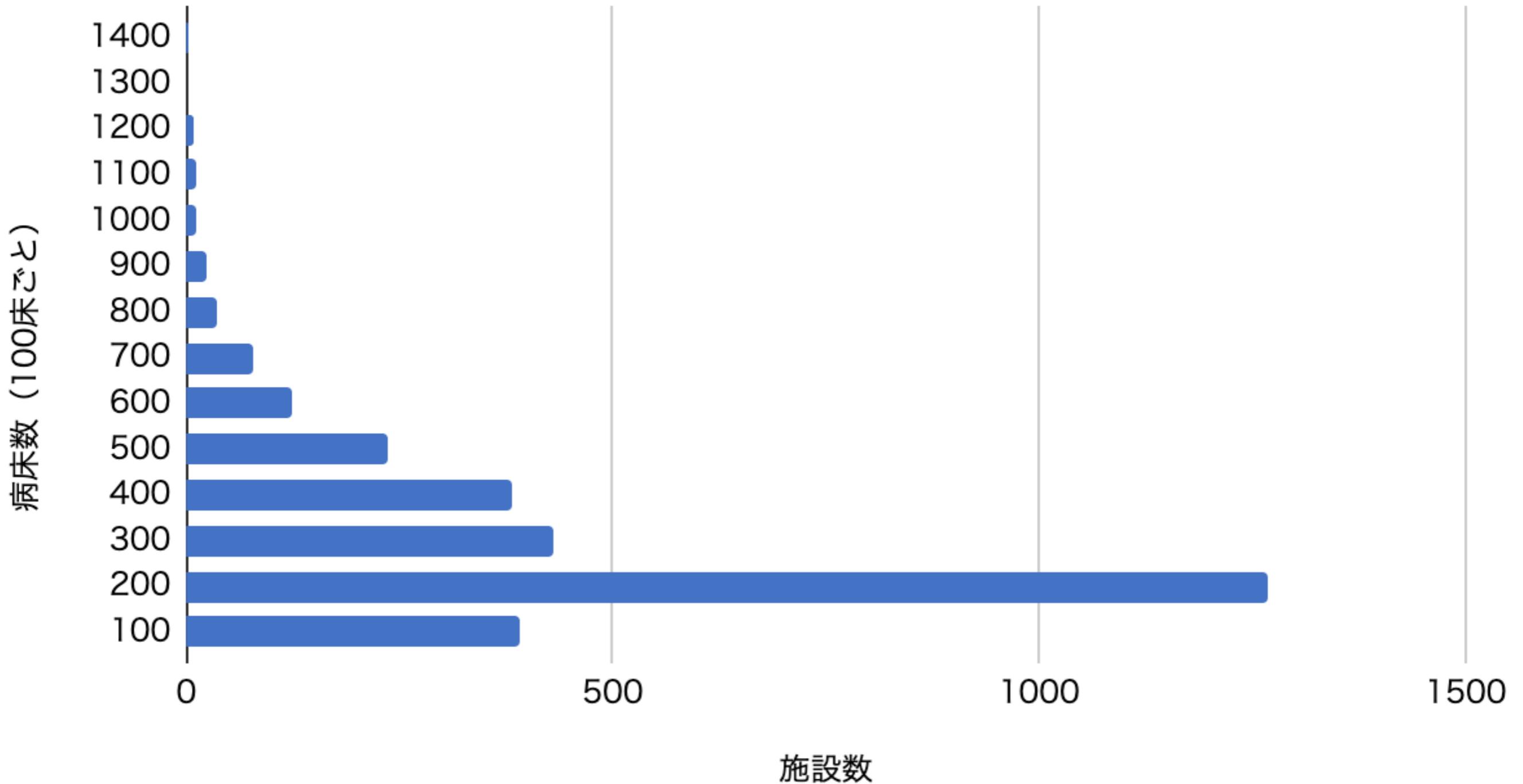
3rd Party

医療機関・患者

EHR事業者

医療機関・患者

病床数上位3000病院の施設数分布（100床ごと）



500床以上：約300病院

300床以上：約1000病院

200床以上：約2000病院



当面目指すは300床以上（1000病院）

医療情報収集＋データ保存管理（概算）

最初の10年：ベンダー10社、病院数1000

医療情報収集：151 億円/10年

データ保存管理：39.8 億円/10年

合計：190.8 億円/10年（19.1億円/年）

医療情報の有効利用：主な課題

医療情報は医療の発展のための国の共有財産

- データ提供の義務化
二次利用データ提供が現状では任意
法的強制力、インセンティブ（保健点数加算など）
- 運営コスト負担
認定事業者の負担が大き過ぎ、大規模データ収集は困難
国による：データ収集、オプトアウト受付代行
- 不明者、死亡者データの有効利用
現法では個人への通知が前提のため、二次利用不可能
- 行政・学校等のデータの利用
多省庁・自治体に権限分散
特に「死亡年月日」「死亡診断書」「学校健診」
法的強制力、インセンティブ（保健点数加算など）
- 準公的EHR管理機関の設立
モデルはカナダ Infoway（グラウンドデザイン、法案、標準化）