

# 医療デジタルツインの発展に資するデジタル医療 データバンク構想

## 研究開発とSociety 5.0との橋渡しプログラム (BRIDGE)

研究開発等計画書  
(令和5年度様式)

令和5年6月  
厚生労働省

○実施する重点課題に○を記載（複数選択可）

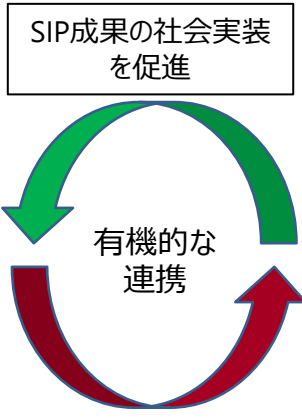
業務プロセス転換・政策転換に向けた取組	次期SIP/FSより抽出された取組	SIP成果の社会実装に向けた取組	スタートアップの事業創出に向けた取組	若手人材の育成に向けた取組	研究者や研究活動が不足解消の取組	国際標準戦略の促進に向けた取組
	○					—

○関連するSIP課題に○を記載（主となるもの）

持続可能なフードチェーン	ヘルスケア	包括的コミュニティ	学び方・働き方	海洋安全保障	スマートエネルギー	サーキュラーエコノミー	防災ネットワーク	インフラマネジメント	モビリティプラットフォーム	人協調型ロボティクス	バーチャルエコノミー	先進的量子技術基盤	マテリアルの事業化・育成エコ
	○												

# 資料 1 「医療デジタルツインの発展に資するデジタル医療データバンク構想」の全体像（位置づけ）

**【内閣府・SIP第3期】統合型ヘルスケアシステムの構築**  
 \*医療デジタルツインの構築  
 \*医療デジタルツインで開発した病院及び個人電子カルテの規格が、中核病院の電子カルテとPHRの標準化  
 \*HL7 FHIRに準拠した電子カルテシステムの導入



**【本施策(BRIDGE)】医療デジタルツインの発展に資するデジタル医療データバンク構想**

デジタル化/構造化された医療データバンク\*  
 [PRISM事業で構築された基盤を活用]

\*基盤となる診療情報統合管理システムは参考資料1を参照

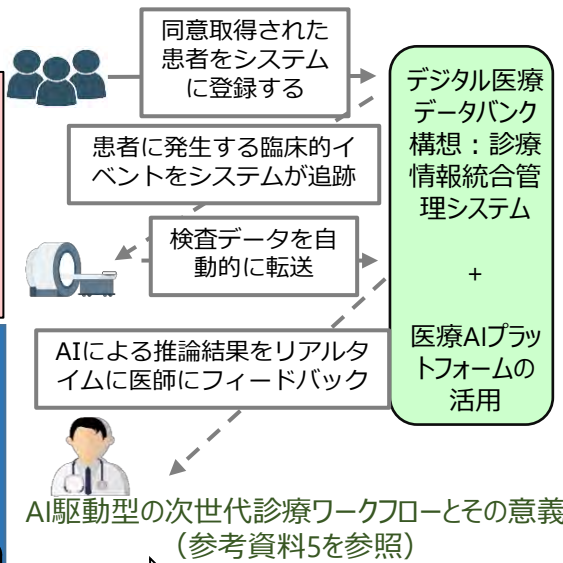
- 診療情報
- 薬剤情報
- レセプト情報
- ゲノムデータ
- 放射線画像データ
- 病理画像データ
- 内視鏡画像データ
- 皮膚画像データ
- ...

**内閣府・次世代医療基盤法認定事業者等**

**厚生労働省・政策科学総合研究事業 [22AD0201]**  
 保険医療分野におけるデジタルデータのAI研究開発等への利活用に係る倫理的・法的・社会的課題の抽出及び対応策の提言のための研究

匿名加工医療情報・仮名加工医療情報作成に関する連携 (参考資料7を参照)

医療デジタルデータの匿名加工・仮名加工手法に関する連携 (参考資料8を参照)



- AI駆動型の次世代診療ワークフローとその意義 (参考資料5を参照)
- 応用例 → **AI駆動型の次世代診療ワークフローの実現及び医療DX人材の育成プログラム**
- 産学連携 → **創薬への応用 [参考資料6]** (AIを積極的に活用し企業との連携を強化)
- パートナー企業出資による事業化
- 産学連携 → **医療機器開発** (AI SaMDの開発を中心に企業との連携を強化)
- 基礎研究 → **学術研究** (世界トップレベルの成果発表を目指す基盤)

- 診療負担の軽減**  
病変の検出や経時画像の計測を自動で行う診療支援AIの開発
- 医療安全の向上**  
高リスク症例を医師が診断する前に検出し、ランク付けするAIの開発
- 研究活動の支援**  
医用画像を起点とした診療情報構造化やAIモデル構築を簡便化する環境の提供
- 医療DX人材の育成**  
診療データの収集、統合、運用管理等の各専門分野を担う医療DX人材の育成

SIP/PDの提案・意見

## 資料2 「医療デジタルツインの発展に資するデジタル医療データバンク構想」の概要

### 【背景・現状・課題】

令和4年6月3日に閣議決定された統合イノベーション戦略 2022において、第5期科学技術基本計画で提案された「Society 5.0」が我が国が目指すべき社会であり、引き続きその実現に注力することが明文化されている。Society 5.0の目指すところは、人工知能（AI）・ビッグデータ・IoT・ロボットなどの新たな技術を取り入れてイノベーションを創出し、一人一人のニーズに合わせる形で社会的課題を解決する新たな社会を構築することである。一方、第6期科学技術・イノベーション基本計画の中でも指摘されているように、新型コロナウイルス感染症に対する蔓延防止の過程で、Society 5.0実現において重要なデジタルトランスフォーメーション（DX）に関する我が国の遅れが顕在化され、その対応は喫緊の課題となっている。以上のことからSociety 5.0実現に向けDXを強力に推進する体制を構築することが急務である。

### 【施策内容】

これまで我が国においても複数のバイオバンクが創設され医学の進歩に貢献してきたが、従来のバイオバンクは血清・手術検体など生物学的材料が中心であり、同一患者の診療情報・ゲノム情報・医用画像情報（放射線画像・内視鏡画像など）・薬剤情報などが紐づいたデジタルデータのバンクは未だ存在しない。そこで、本施策では本邦初のデジタル医療データバンクを構築し、AI駆動型の次世代診療ワークフローの実現、創薬や医療機器開発への応用を目指す。特に現時点において次世代医療基盤法下における匿名化加工医療情報に該当する医用画像は存在しないため、厚生労働省・政策科学総合研究事業（22AD0201）とも連携しながら医用画像に関する匿名加工・仮名加工に関する技術的な方法論を確立させ、本施策において積極的に活用する。また、国立がん研究センター/理化学研究所/産業技術総合研究所などの機関は、これまでPRISM事業において本施策の基盤となる統合データベースの構築などの成果を既に出しているため（参考資料2-4を参照）、それらの基盤を本施策においても積極的に活用していく予定である。

### 【研究開発等の目標】

- ・本邦初のデジタル医療データバンクを構築する。
- ・創薬や医療機器開発に応用するうえでのパートナー企業から出資を募り事業化を推進する。
- ・SIP第3期課題「統合型ヘルスケアシステムの構築」と有機的に連携し、HL7 FHIRに準拠したシステムの導入など研究目標や医療データベースの仕様を合わせるとともに、SIP課題の早期社会実装を支援する。
- ・AI駆動型の次世代診療ワークフローの実現及び医療DX人材の育成を目指す（レジデント、大学院生及び博士研究員など若手人材が主な対象で毎年度5名以上若手人材の参加を必須とする）。

### 【社会実装の目標】

- ・デジタル医療データバンクを事業化させる（上記パートナー企業からの出資とデータ使用料が主な収益）。
- ・AI駆動型の次世代診療ワークフローの基盤を構築する。
- ・本施策で構築されたデジタル医療データバンクを6ナショナルセンターや大学病院などに展開していく。

### 【対象施策の出口戦略】

- ・SIP第3期課題「統合型ヘルスケアシステムの構築」への導出を予定している。  
→BRIDGEにおいて作成するデータベースに含まれるデータそのものは、国立がん研究センターに蓄積されたデータが中心となるため、初めはSIPとは独立したデータベースとなるが、共有のための諸手続きを迅速に進め、BRIDGEの成果としてのデータベースはSIPのデータベースに統合していく予定である。  
→ AI駆動型の次世代診療ワークフローの実現に向けた取り組みの成果は、SIPのサブ課題B「患者・医療機関支援ソリューションの開発」に導出する予定である。
- ・対象施策で蓄積された成果や知見を次世代医療基盤法認定事業者や各省庁で実施する施策に反映させていく。

### ○統合イノベーション戦略や各種戦略等との整合性

統合イノベーション戦略 2022において「Society 5.0の実現」が我が国が目指すべき社会であることが明文化されているが、AI等を含む先端技術を活用したデジタルツインの構築等によるデジタル社会の形成は、Society 5.0実現のための根幹をなすものである。本施策はSIP第3期課題「統合型ヘルスケアシステムの構築」と有機的に連携しながら医療デジタルツインの構築に大きく貢献することが期待される。

### ○重点課題要件との整合性

本施策で構築するデジタル医療データバンクは、SIP第3期課題「統合型ヘルスケアシステムの構築」が目指す医療デジタルツインの構築に貢献し、またその社会実装を促進すると考えられる（重点課題の「次期SIP/FSより抽出された取組」及び「SIP成果の社会実装に向けた取組」に該当）。また、医療DXの人材育成は重要課題「若手人材の育成に向けた取組」に該当する。

### ○SIP型マネジメント体制の構築

SIP第3期課題「統合型ヘルスケアシステムの構築」のPD（SIPPD）がSIPにおける推進委員会での意見を踏まえつつ、本施策の提案、助言及び支援を行う。

### ○民間研究開発投資誘発効果、財政支出の効率化

本施策は本邦初のデジタル医療データバンクを構築することで、同一患者の診療情報・ゲノム情報・医用画像情報（放射線画像・内視鏡画像など）・薬剤情報などが紐づいたデジタルデータを一気通貫に解析が可能となるため、創薬や医療機器開発における大変重要なリソースとなると考えられる。製薬企業や医療機器メーカーから出資を募るとともに（約10億円程度を想定）、営利企業のデータ利用を有料化（データ使用料の徴収：約20億円/年程度を想定）を予定している。

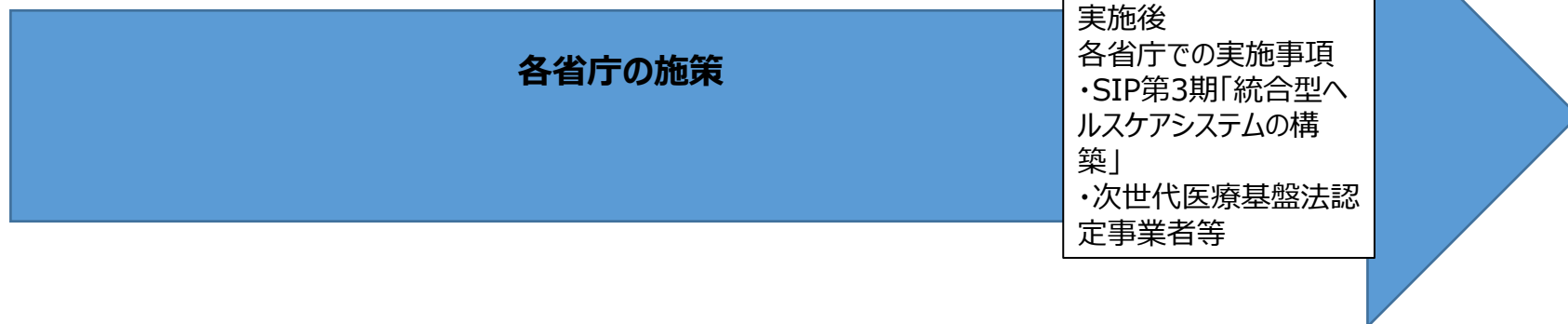
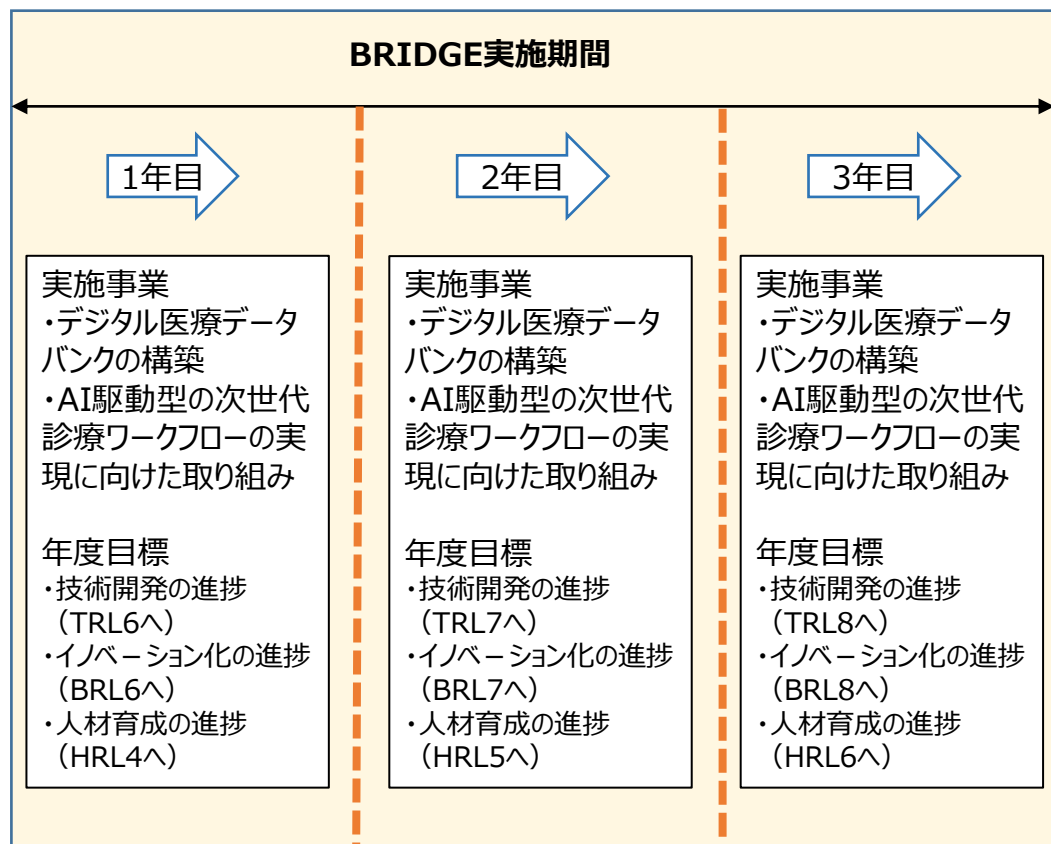
### ○民間からの貢献額（マッチングファンド）

デジタル医療データバンクの構築には、病院情報システムや診療情報統合管理システムの構築などに精通している企業の参画が必須である。具体的には、新規購入費／保有品利用料として1億円、人件費（20人 x 1500万円）として3億円の計約4億円程度を想定している。

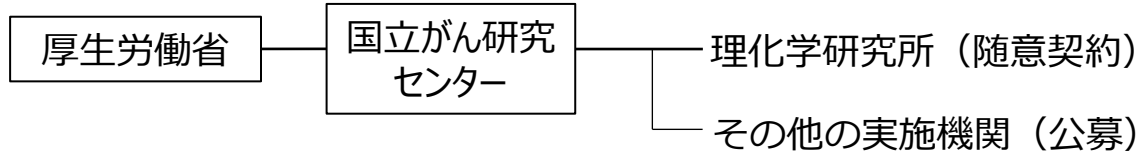
### ○想定するユーザー

- ・国内/海外の製薬企業
- ・国内/海外の医療機器メーカー
- ・学術機関及び医療機関
- ・行政機関

# 資料4 イノベーション化に向けた工程表



## 実施体制



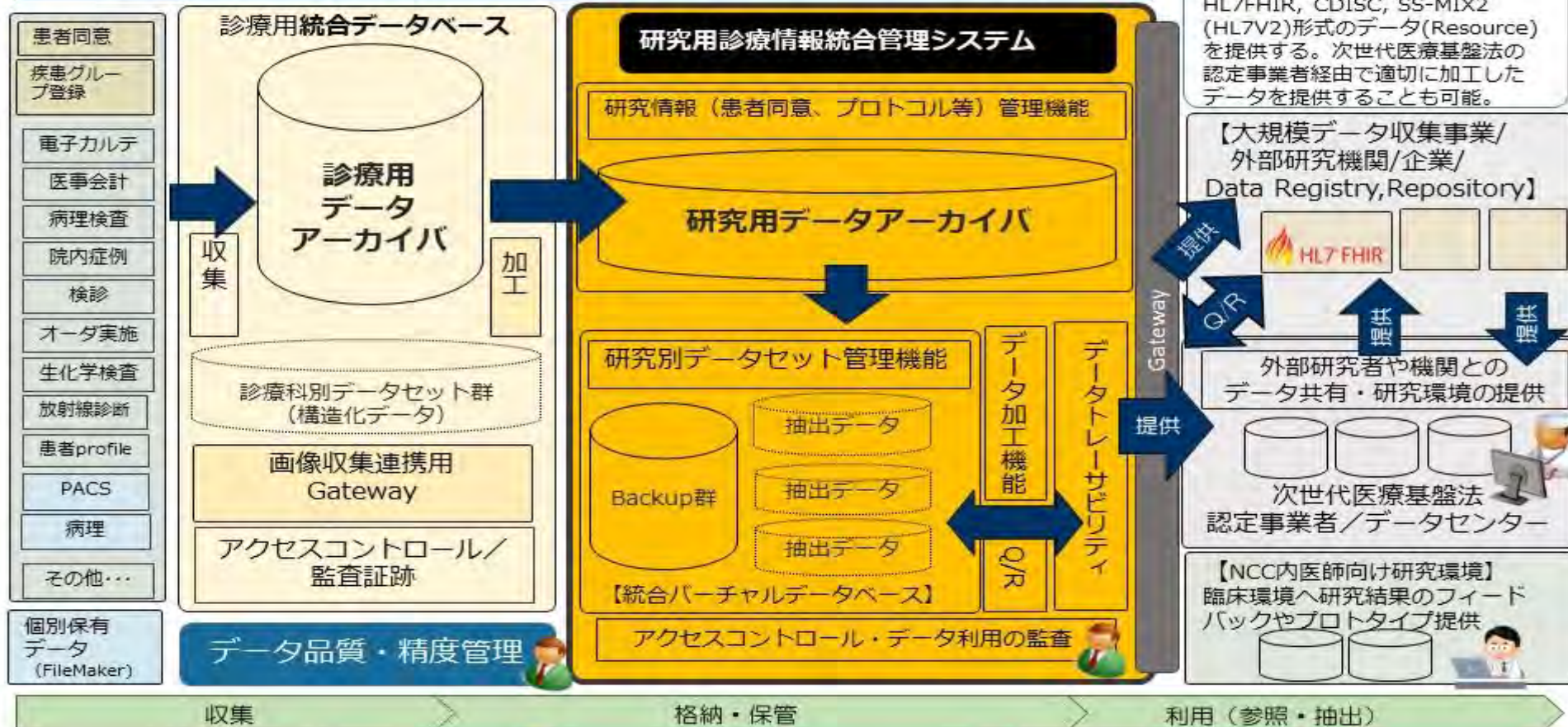
SIP第3期課題「統合型ヘルスケアシステムの構築」のSIPPDが、SIPにおける推進委員会での意見を踏まえつつ、本施策の提案、助言及び支援を行う。また、デジタル医療データバンクの構築は、全国のがん診療連携拠点病院等422施設への診療支援体制を構築している国立がん研究センターを中心に実施する。本施策の基盤となるPRISM事業の成果及び知見を継承できる施設として、理化学研究所は世界最高レベルのスパコン「富岳」及びAI用スパコン「RAIDEN」を有し、本施策で計画しているAI解析環境を準備できる唯一の機関であるため、公募によらず随意契約で実施する。さらに、本施策を発展させることを目的に、本施策の目標に合致したデータやリソースを保持した適切な機関と連携していくことも予定しており、それらの機関は公募で実施する。

○施策全体の目標 ……本施策では本邦初のデジタル医療データバンクを構築し、AI駆動型の次世代診療ワークフローの実現、医療DX人材の育成、産学連携に基づく創薬や医療機器開発へ応用するとともに、質の高い学術研究を行うことを目標とする。特に次世代医療基盤法下における匿名加工医療情報に該当する医用画像が未だ存在しない状況を鑑み、厚生労働省・政策科学総合研究事業とも連携しながら、匿名加工・仮名加工における技術的な方法論も確立させながら推進する。本施策の成果は事業化することを想定しており、施策終了後はSIP第3期課題「統合型ヘルスケアシステムの構築」への導出を予定している。

テーマ等（※個別に目標を設定している場合）	当年度目標	目標の達成状況（年度末報告）
①デジタル医療データバンクの構築	<p><u>R5年度</u></p> <ul style="list-style-type: none"> <li>・技術開発の進捗：医療AI研究開発を志向した医療データプラットフォームの構築と実装（TRL6へ）</li> <li>・イノベーション化の進捗：デジタル医療データバンクの創薬・医療機器開発への応用[1件以上のPOCの取得]（BRL6へ）</li> <li>・人材育成の進捗：レジデント、大学院生及び博士研究員など若手人材の参加を必須とし育成する（5名以上）[若手人材はテーマ①及び②双方に参加]（HRL4へ）</li> </ul>	(-)
②AI駆動型の次世代診療ワークフローの実現に向けた取り組み	<p><u>R5年度</u></p> <ul style="list-style-type: none"> <li>・技術開発の進捗：診断または治療をサポートするAIの開発（TRL6へ）</li> <li>・イノベーション化の進捗：開発した医療AIの社会実装を目指した取り組み[薬事承認を目指したPMDAとの面談・実証実験の推進など（1件以上）]（BRL6へ）</li> <li>・人材育成の進捗：レジデント、大学院生及び博士研究員など若手人材の参加を必須とし育成する（5名以上）[若手人材はテーマ①及び②双方に参加]（HRL4へ）</li> </ul>	(-)

# 参考資料 1 本施策の基盤となる診療情報統合管理システムの概要

## 研究用診療情報統合管理システム概要 (GOAL)



統合データベースに登録されたデータは共有され、必要な場面で引用、確認、検索、抽出、データ加工の上提供可能となる

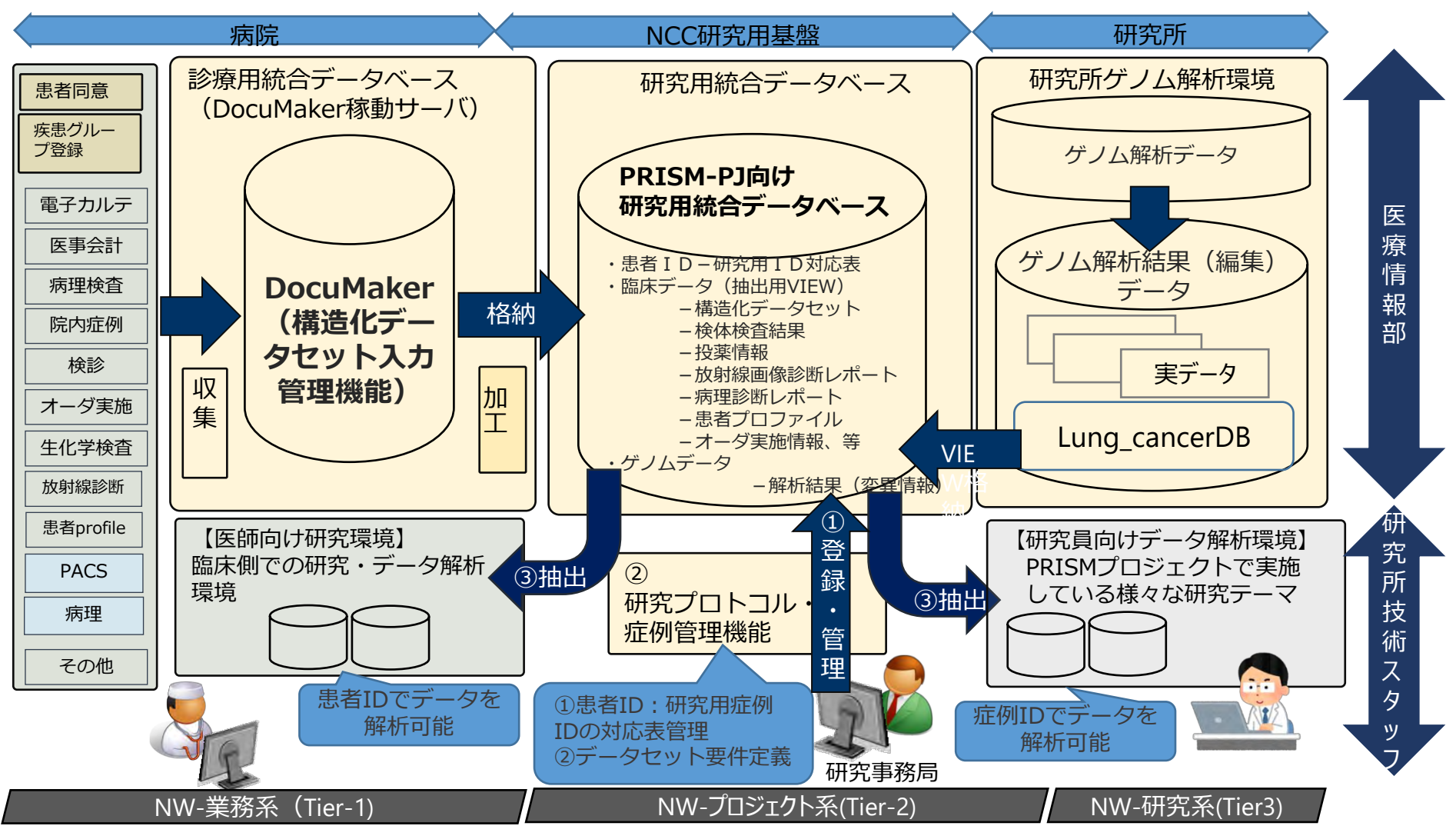
中央病院内に存在する様々なシステムを情報の発生源とし、テキスト、画像データを収集し管理する。診療科ごとに維持しているFileMakerデータと同等の構造化データ入力・管理機能を提供する。研究に必要な患者の同意や研究対象症例の登録機能を備える。

センター全体の研究を実施するための診療情報を統合的に管理する仕組みを維持するためのシステム。データ格納・保管機能の他に、研究テーマごとの厳密なアクセスコントロール、抽出したデータのトレーサビリティ、監査機能やデータ加工機能を保有する。

データ利用者に応じたデータ抽出、共有環境の他に、多施設共同研究や計算センターへのデータ転送などを視野に入れたデータ管理機能を実現する。患者を想定したデータ利用状況監査機能を提供する。



# 参考資料2 これまで構築してきた統合データベースシステムの概要 (PRISM成果)



## 【運用フロー概要】

医療情報部：データセットそのものを設計、構築、設定を担当する。Documaker画面の拡張が必要な場合はファインデックスに依頼し対応（作業量により若干の費用発生）

研究所技術スタッフ：

- ①患者ID-症例IDの対応表管理。
- ②研究テーマごとに必要なデータセットのとりまとめ（要件定義）。
- ③-1一括データ抽出の場合は直接作業、③-2オンデマンドの抽出が必要な場合は抽出画面の要件定義、操作サポートをご担当いただく。

データ移送／匿名化

アノテーション

学習モデルの設定

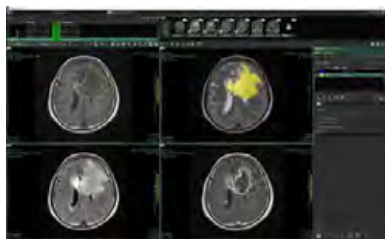
学習

AI活用

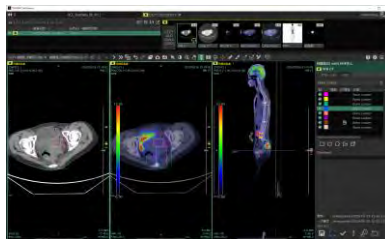
プロジェクト管理



アノテーション

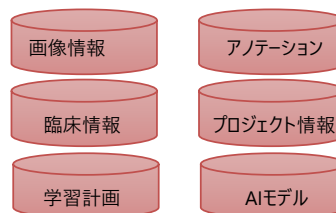


SYNAPSE SAI viewer<sup>\*1</sup>,  
SYNAPSE VINCENT<sup>\*2</sup>で培った  
画像編集技術を利用可能



複数シリーズを使った高度な  
アノテーションも可能

学習管理



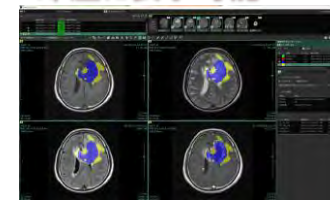
AI開発に必要な情報を  
一元管理



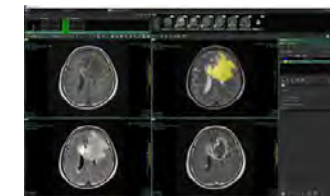
- ・正解データの進捗管理
- ・AI学習用のPCスケジュール管理
- ・学習の自動実行
- ・学習モデルの管理

AI実行

学習した結果を  
即座に表示可能



学習済モデルを用いて、  
更なる追加学習データを  
効率的に作成

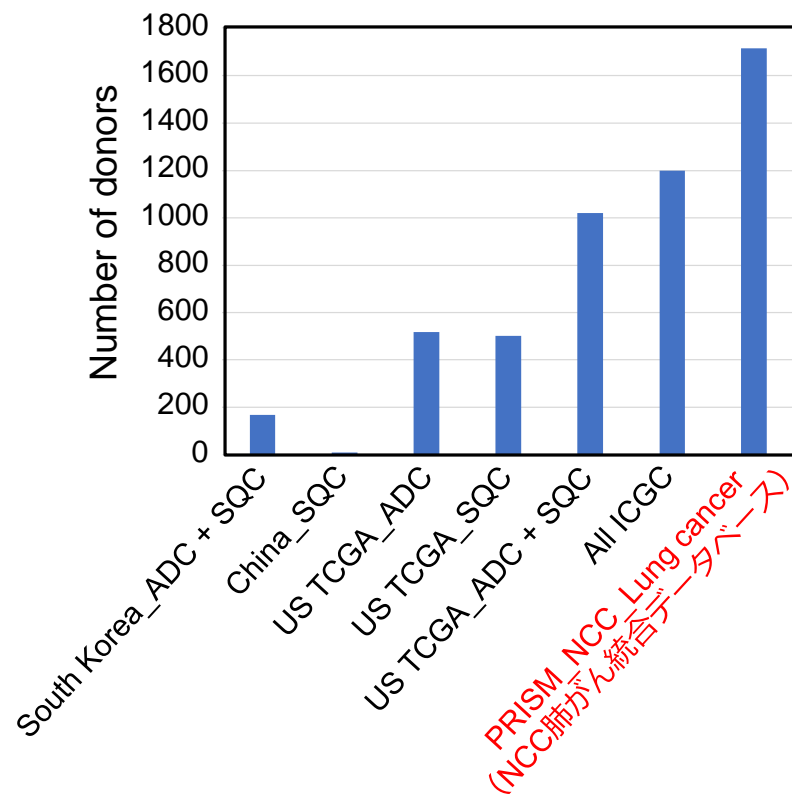


AI開発を All in One でサポートするプラットフォームを実現します！

\*1 販売名：画像診断ワークステーション用プログラム FS-V686型  
認証番号：231ABBZX00028000

\*2 販売名：販売名：富士画像診断ワークステーション FN-7941型  
認証番号：22000BZX00238000

## 肺がんデータベースにおける症例数の国際的な比較



ICGC Data Portal lung cancer (<https://dcc.icgc.org/q?q=lung%20cancer>)  
に2022年5月2日にアクセス

## 肺がん統合データベースの内訳

	TCGA		PRISM_NCC
	ADC	ADC + SQC	
臨床情報	585	1,089	1,714
オミックスデータ			
全エクソーム解析	582	1,084	1,556
RNA-seq解析	515	1,016	1,682
DNAメチル化解析	579	1,082	403*
全ゲノム解析	99	-	413*
ChIP-seq (H3K27Ac)解析	-	-	222*

\*全ゲノム解析・ChIP-seq・DNAメチル化解析に関しては、pan-negative肺がん症例に集中して解析を行った。

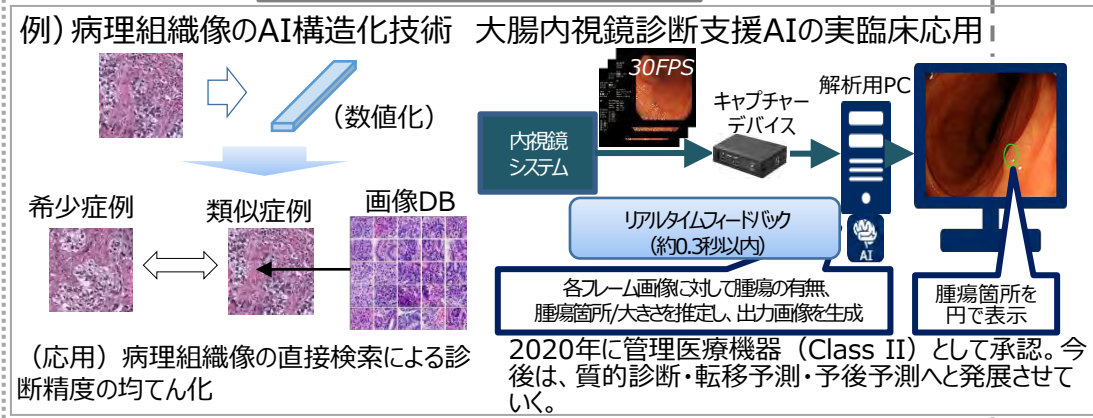
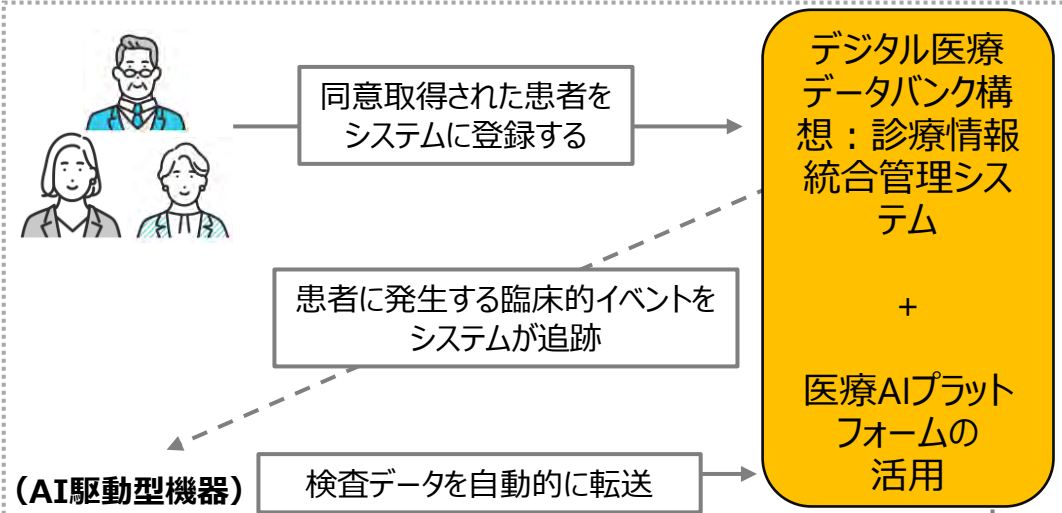
**AMEDの全ゲノムプロジェクトと連携しており、データは統合していくことを予定している。**

【PRISMと全ゲノムプロジェクトのデータを合わせた肺がん症例数 (参考資料)】

\*肺がん症例数 (臨床情報) : 約2200症例

\*全ゲノム解析データ : 約900症例

\*RNA-seqデータ : 約2100症例



新型コロナウイルス感染症の拡大下において医療システムの課題として認識された、柔軟で強靱な医療提供体制の構築、デジタル化・オンライン化を実現する。創薬研究、デジタル化・リモート化やAI・ロボットの活用を促進する事で、「イノベーション国家」としてのプレゼンスを世界に示していく。

AIによる推論結果をリアルタイムに医師にフィードバック



### 診療負担の軽減

病変の検出、候補診断名の提示や経時画像の計測を自動で行う診療支援AIの開発

→ 読影や確定診断に必要とする時間の短縮による診療スループットの向上、医療従事者の負担軽減、定量的な病変評価による治療の質の向上

### 医療安全の向上

高リスク症例を医師が診断する前に検出し、ランク付けするAIの開発

→ 高リスク症例の早期発見による患者予後の向上、見落とし防止による医療安全向上

### 研究活動の支援

医用画像を起点とした診療情報構造化やAIモデル構築を簡便化する環境の提供

→ 多くのモダリティの臨床情報の定量的統合、医療AI研究活動の支援

### 医療DX人材の育成

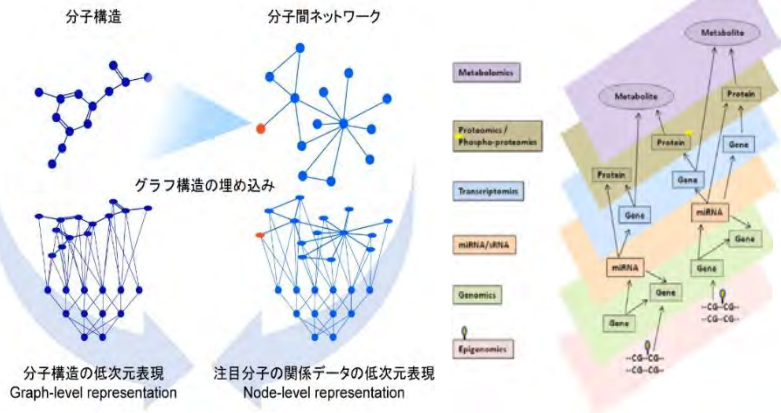
診療データの収集、統合、運用管理等の各専門分野（データ品質管理、システム構築、セキュリティ、法令順守等）を担う医療DX人材の育成。資格化を見越した教育システムの構築

→ 高精度なデータ利活用環境の整備

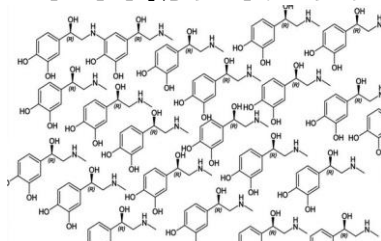
## 化合物-マルチオミクス情報の相関抽出

本施策で構築した薬剤投与に伴う *in vitro* マルチオミクス・プロファイルの解析データに基づいて、化合物の分子構造とオミクス・プロファイルの変化量の対応推定を行う。

## 網羅的分子情報



## 化合物ライブラリ



### 深層表現学習を用いた化合物のSAR展開

深層学習を用いた化合物構造のバーチャルな最適化アルゴリズムにより最終的に導出する化学構造を決定する。

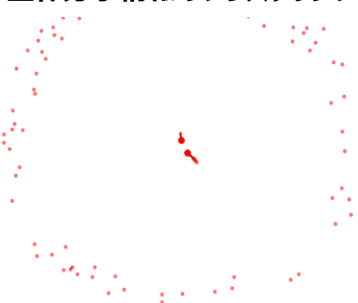
## タンパク質立体構造



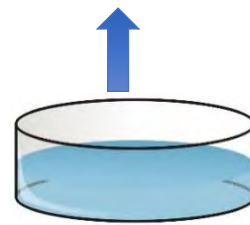
### 標的候補たんぱく質に対する化合物の結合性予測

RIKEN「富岳」を用いたタンパク質-化合物の結合予測あるいは深層学習を用いた推定によりリード化合物を決定。

### 臨床的表現型を教師とした生体分子情報のクラスタリング



### Bayesian Networkや Knowledge-based Network (RIKEN SIGN-BN, KEGG, etc...)



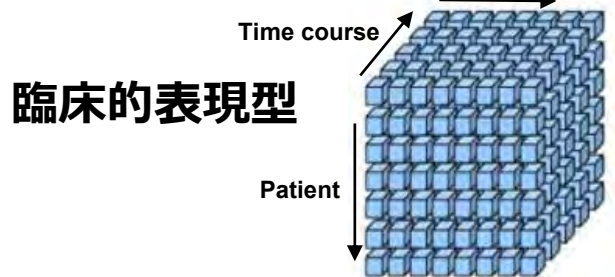
### *in silico*により予測された因子に対する実験系による検証

ネットワーク上のハブやボトルネックとして現れる因子に対して、実験系による抗腫瘍効果の検証を行い、標的候補タンパク質を決定。

### マルチオミクス・プロファイルの差別的情報に注目した薬剤探索

疾患の罹患や増悪時、治療反応群と不応群間の差異として現れたオミクス・プロファイルの変化と逆相関するものが候補化合物と期待される。

### 臨床的表現型のクラスタ Clinical descriptors



深層学習によりマルチモーダルな情報を教師データとしてオミクスデータをクラスタリング

### 臨床的表現型を教師データとする分子情報のクラスタリングを起点とした薬剤探索

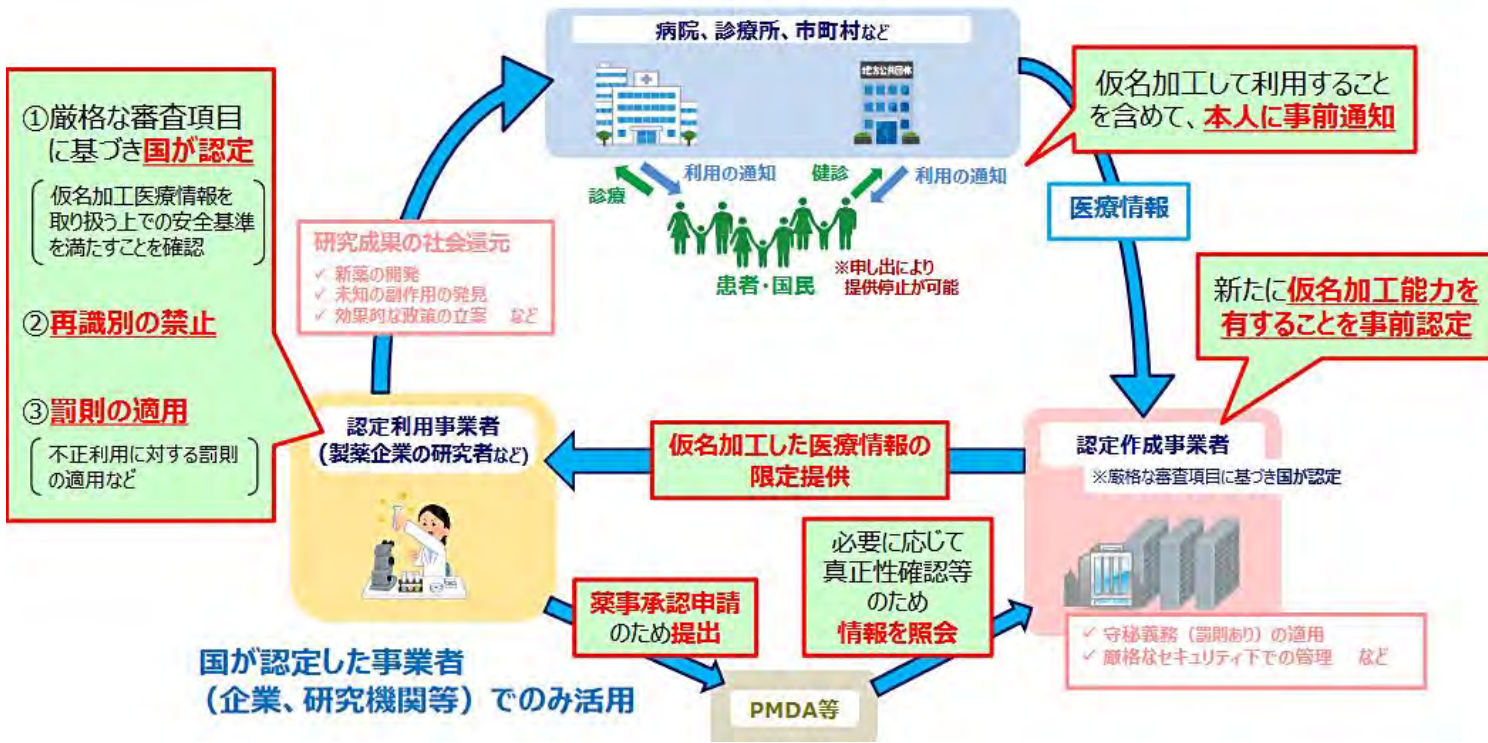
多様な臨床的表現型のクラスタを教師データとしてクラスタリングを行い、ベイズ推定で求めた因果ネットワークや知識ベース・ネットワークにマッピングする。

医療データバンクのデータに基づき構築

# 参考資料 7 次世代医療基盤法との連携：改正案が閣議決定

- 医療情報の研究ニーズ、社会的便益の観点から、新たに「仮名加工医療情報」の作成・提供を可能とする
- その際、個人情報の保護の観点から、仮名加工医療情報の提供は国が認定した利活用者に限定

2023年3月3日（金）に次世第医療基盤法の改正案が閣議決定された。



現行法の匿名加工医療情報に加えて、新たに「仮名加工医療情報 ※」（仮称）を創設することとし、以下のように法令上の措置を検討。

→ 仮名加工医療情報を作成・提供する事業者を国が認定する仕組みを新たに設ける。

→ 上記認定事業者から、安全管理等の基準に基づき国が認定した利活用者に限り、仮名加工医療情報を提供可能とする仕組みを設ける。

→ 薬事承認申請のため、PMDA等に対し、利活用者からの仮名加工医療情報の提供、認定事業者からの元データ提出を可能とする。

※ 仮名加工医療情報：他の情報と照合しない限り、個人を特定できないよう加工した情報。個人情報から氏名やID等の削除が必要だが、特異な値等の削除等は不要。

2023年3月9日にアクセス

出典：首相官邸 第7回 次世代医療基盤法検討ワーキンググループ  
 (https://www.kantei.go.jp/jp/singi/kenkouiryuu/data\_rikatsuyou/jisedai\_iryokiban\_wg/dai7/siryuu1.pdf)

本施策においては次世代医療基盤法下における匿名加工医療情報及び仮名加工医療情報（改正案で新設）も適切に活用していく。

### 研究課題名

保健医療分野におけるデジタルデータのAI 研究開発等への利活用に係る倫理的・法的・社会的課題の抽出及び対応策の提言のための研究（22AD0201）

### 求められる成果

- ① 保健医療分野におけるデジタルデータ（病理画像、CT・MRI 画像、手術動画、ゲノムデータ等）のAI 研究開発等への利活用に係るELSI の抽出、国際的な動向も踏まえた対応策の提言
  - ② 保健医療分野におけるデジタルデータのAI 研究開発等への利活用に係るガイドライン案（デジタルデータの加工手法、加工基準を含む）の策定、及び研究者等が参考にできる事例集等の作成
  - ③ 保健医療分野におけるデジタルデータのAI 研究開発等への利活用に係る国内外のELSI の議論の動向の調査・分析の結果
  - ④ ゲノムデータの個人識別性に該当する範囲について、科学的な観点、海外の動向を踏まえた総合的な解釈に関する提言
  - ⑤ ゲノムデータの持つ個人識別性に関する国内外の議論の動向の調査・分析の結果
- ※ 上記の成果物は、個人情報保護委員会事務局等の関係機関への横展開が可能

出典：令和4年度 厚生労働科学研究費補助金公募要項（三次）（<https://www.mhlw.go.jp/content/10600000/000965669.pdf>）

本施策を推進する上で医療データの匿名加工・仮名加工は重要な過程であるため、厚生労働省・政策科学総合研究事業 [22AD0201]と連携し、利便性が高くかつ安全性が高い匿名加工・仮名加工手法のガイドライン作成に協力しながら本施策にもその手法を取り入れていく。