



医療等分野における情報の保護と利活用に関する実態調査 報告書

2025年3月28日

目次

略語一覧	3	4. 医療等情報の二次利用に係る取組	94
用語一覧、報告書の記載	9	4.1 実運用も考慮した具体的な運用・システムの整理	95
1. 会議の運営	10	4.2 情報提供依頼の実施概要	108
1.1 技術作業班の会議運営	11	4.2.1 情報提供依頼書	109
1.2 技術作業班の概要および開催実績	15	4.2.2 要件定義書	132
1.3 技術作業班での主な意見	18	4.3 次年度に向けた論点項目の洗い出し	224
1.4 技術作業班委員の個別ヒアリング	21	5. 医療等情報の二次利用推進に向けた海外におけるデータ活用基盤及び技術動向調査について	239
2. 諸外国の医療情報の取扱いについての把握・分析	35	5.1 データ活用基盤等調査事業	243
2.1 諸外国の調査	35	5.2 技術動向調査事業	355
2.1.1 EHDS	36	6. 総括	378
2.1.2 国際標準コード	64		
2.1.3 OMOP	70		
2.1.4 オプトアウト	76		
2.1.5 医療情報等の海外移転について	81		
3. 医療等情報を収集するデータベース（レジストリ）調査について	85		
3.1 公的データベースの現状調査	86		
3.2 国内におけるビジテイング解析環境の比較	91		

略語一覧 (1/6)

略語	スペル	説明	対象国
ABPI	Association of the British Pharmaceutical Industry	英国製薬産業協会	英国
AHRQ	Agency for Healthcare Research and Quality	医療研究品質庁	米国
ANS	Agence du Numérique en Santé	デジタルヘルス庁	フランス
ANSM	Agence Nationale de Sécurité du Médicament et des Produits de Santé	医薬品及び健康製品の安全のための国家機関	フランス
API	Application Programming Interface	アプリケーション・プログラミング・インターフェース	共通
APMS	Adult Psychiatric Morbidity Survey	成人精神疾患調査	英国
APR-DRG	All Patient Refined Diagnosis IAGNOSIS RELATED GROUPS	全患者の改良診断IAGNOSIS関連群	共通
ARIA	Active Postmarket Risk Identification and Analysis	積極的な市場後リスクの特定と分析	米国
ASPE	Assistant Secretary of Health and Human Services for Planning and Evaluation	計画・評価担当保健社会福祉次官補	米国
ATIH	Technical agency for information on hospitalization	病院情報技術庁	フランス
BA	Business Associate	Covered Entitiesの業務をサポートする事業者の総称	米国
BAA	Business Associates Agreement	ビジネスアソシエイツ契約	米国
BEST	Biologics Effectiveness and Safety	生物製剤の有効性と安全性	米国
CAG	Confidentiality Advisory Group	守秘義務アドバイザリーグループ	英国
CBER	Center for Biologics Evaluation and Research	生物製剤評価研究センター	米国
C-CDA	Consolidated Clinical Document Architecture	統合診療記録文書構造	米国
CCW	Chronic Conditions Data Warehouse	慢性疾患データウェアハウス	米国
CDA	Clinical Document Architecture	診療記録文書構造	共通
CDC	Centers for Disease Control and Prevention	疾病管理予防センター	米国
CDER	Center for Drug Evaluation and Research	医薬品評価研究センター	米国
CDS	Commissioning Data Sets	データセットのコミッションング	英国
CCG	Clinical Commissioning Group	クリニカル・コミッションング・グループ、国内各地域の医療政策・予算権限を持つ組織	英国
CE	Covered Entities	医療提供者、保険者等医療関連サービス事業者の総称	米国
CESREES	Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé	健康分野の研究・研究及び評価のための倫理及び科学委員会	フランス
CFR	Code of Federal Regulations	連邦規則	米国
CMDC	Medical Causes of Death	国立保健医学研究の管轄するデータベース	フランス
CMP	Civil Money Penalty	民事制裁金	米国
CMS	Centers for Medicare & Medicaid Services	メディケア&メディケイドサービスセンター、主要な医療プログラムを管理する機関	米国
CNAM	Caisse Nationale de l'Assurance Maladie	国民健康保険基金	フランス
CNIL	Commission Nationale de l'Informatique et des Libertés	情報処理及び自由に関する国家委員会	フランス

略語一覧 (2/6)

略語	スペル	説明	対象国
CNRIPH	Commission National des Recherches Impliquant la Personne Humaine	人対象研究全仏委員会	フランス
CNSA	Caisse nationale de solidarité pour l'autonomie	全国自律連帯金庫	フランス
CPP	Comité de Protection des Personnes	人保護委員会	フランス
CPRD	Clinical Practice Research Datalink	臨床実践研究データリンク	英国
CSP	Code de la Santé Publique	公衆衛生法典	フランス
DB	Database	データベース	共通
DCMS	Department for Science, Innovation and Technology	デジタル・文化・メディア・スポーツ省	英国
DICOM	Digital Imaging and Communications in Medicine	医用画像および関連情報の国際標準	共通
DOJ	Department of Justice	アメリカ合衆国司法省	米国
DP	Data Partner	データパートナー	米国
DPA	Data Protection Act	データ保護法	EU
DPC	Danbury Psychiatric Consultants	ダンバリー精神科コンサルタント	米国
DPC	Diagnosis Procedure Combination	匿名診療等関連情報	日本
DPO	Data Protection Officer	データ保護責任者	EU
DUA	Data Use Agreement	データ利用契約	米国
DQF	Data Quality Framework	データクオリティフレームワーク	欧州
DSA	Data Sharing Agreements	データ共有契約	英国
DSP	Digital Staff Passport	デジタルスタッフパスポート	日本
DX	Digital Transformation	デジタルトランスフォーメーション	共通
EC	European Commission	欧州委員会	欧州
ECDS	Emergency Care Data Set	救急医療データセット	英国
EDPS	European Data Protection Supervisor	欧州データ保護監察機関	EU
EEA	European Economic Area	欧州経済領域	EU
EGFR	Epidermal Growth Factor Receptor	ヒト上皮増殖因子受容体	共通
EHDS	European Health Data Space	欧州健康データスペース	EU
HER	Electronic Health Record	電子健康記録	共通
EMR	Electronic Medical Record	電子カルテ	共通
ENS	Mon Espace Santé	マイヘルスペース	フランス
EU	European Union	欧州連合	欧州
FDA	Food and Drug Administration	食品医薬品局	米国
FDAAA	Food and Drug Administration Amendments Act	改正食品医薬品局法	米国
FHIR	Fast Healthcare Interoperability Resources	手早く設計し導入できる保健医療分野の相互運用性リソース	共通

略語一覧 (3/6)

略語	スペル	説明	対象国
FIH	First-in-Human	ヒト初回投与試験、ヒトに初めて薬を投与する段階の治験	共通
FY	Fiscal Year	会計年度	共通
GAfREC	Governance Arrangements for Research Ethics Committees	研究倫理委員会ガバナンス協定	英国
GCP	Good Clinical Practice	医薬品の臨床試験の実施の基準	英国
GDPR	General Data Protection Regulation	EU一般データ保護規則	EU
GP	General Practitioner	総合診療医、かかりつけ医	共通
GSK	Glaxo Smith Kline	グラクソスミスクライン、製薬企業	英国
GTAC	Gene Therapy Advisory Committee	遺伝子治療諮問委員会	英国
HAD	Hospitalization-Associated Disability	入院関連機能障害	共通
HARP	HRA Assessment and Review Portal	HRA評価レビューポータル	英国
HCUP	Healthcare Cost and Utilization Project	ヘルスケア コスト アンド ユーティリゼーション プロジェクト	米国
HDAB	European Health Data Access Body	EU各国が指定した、健康データの二次利用に関する業務を行う機関	EU
HDH	Health Data Hub	ヘルスデータハブ、国内の医療データを集約したデータベース	フランス
HDR UK	Health Data Research UK	ヘルスデータリサーチ、国立医療データサイエンス研究所	英国
HES	Hospital Episode Statistics	外来や入院等の病院データセット	英国
HHS	United States Department of Health and Human Services	米国合衆国保健福祉省	米国
HIC	Healthcare Intelligence Cloud	医療・介護データ等解析基盤	日本
HIE	Health Information Exchange	診療情報交換、必要な診療情報の共有に焦点をあてた概念	米国
HIN	Health Information Network	健康情報ネットワーク、州間データ交換に重点を置く全国的なネットワーク	米国
HIPAA	Health Insurance Portability and Accountability Act	経済的及び臨床的健全性のための医療情報技術に関する法律	米国
HITECH	Health Information Technology for Economic and Clinical Health Act	医療保険の相互運用性と説明責任に関する法律	米国
HL7	Health Level 7	ヘルスレベル7、医療情報のうち文字情報に関する規格	共通
HRA	Health Research Authority	医療研究機構	英国
HCSRN	Health Care Systems Research Network	医療システム研究ネットワーク	米国
IAS	Individual Access Services	個別アクセスサービス	米国
ICO	Information Commissioner's Office	英国個人情報保護監督機関	英国
ICD-9-CM/ICD-10-CM	International Classification of Diseases 9th Revision, Clinical Modification	医療行為の分類	米国
IDE	Integrated Development Environment	統合開発環境	共通
IDR	Integrated Data Repository	統合データリポジトリ	米国
IHE	Integrating the Healthcare Enterprise	医療情報システムの相互接続性を推進する国際的なプロジェクト	共通

略語一覧 (4/6)

略語	スペル	説明	対象国
ICNARC	Intensive Care National Audit & Research Centre	全国集中治療監査・研究センター	英国
INSERM	L'Institut national de la santé et de la recherche médicale	国立保健医学研究所	フランス
IOM	Institute of Medicine	科学アカデミー医学研究所	米国
IRAS	Integrated Research Application System	REC・CAGの倫理審査申請プラットフォーム	英国
IRB	Institutional Review Board	倫理審査委員会	米国
ITK	Interoperability Toolkit	NHSがシステム・データの標準化を図るために公開している仕様	英国
JLAC	Japan Laboratory Code	臨床検査項目分類コード	日本
KAKuG	Kranken- und Kuranstaltengesetz des Bundes	連邦病院および健康法	オーストリア
KDAP	KPRB Data and Analytic Platform	KPRBデータ分析プラットフォーム	米国
KID	Kids' Inpatient Database	小児入院患者データベース	米国
KP	Kaiser Permanente	カイザーパーマネンテ、アメリカ合衆国の健康維持機構	米国
KPRB	Kaiser Permanente Research Bank	カイザーパーマネンテリサーチバンク、バイオバンク	米国
LDS	Limited Data Set	リミテッドデータセット	米国
LHB	Linkage Honest Broker	プライバシー保護レコードリンケージを実行するリンケージオーネストブローカー	米国
MACRA	Medicare Access and CHIP Reauthorization Act	メディケアアクセスおよびCHIP再承認法	米国
MDPH	Maison Départementale pour les Personnes Handicapées	障害者のためのデータベース	フランス
MHLDDS	Mental Health and Learning Disabilities Data Set	メンタルヘルスおよび学習障害データセット	英国
MHMDS	Mental Health Minimum Data Set	メンタルヘルスミニマムデータセット	英国
MHRA	Medicines and Healthcare products Regulatory Agency	医薬品医療製品規制庁	英国
MHSDS	Mental Health Services Data Set	メンタルヘルスサービスデータセット	英国
MIPS	Merit-based Incentive Payment System	成果報酬制度	米国
MoDREC	Ministry of Defence Research Ethics Committee	防衛省研究倫理審査委員会	英国
MPS	Master Patient Service	マスターパシエントサービス	英国
MRC	Medical Research Council	医学研究会議	英国
MSL	Multi-Study Licence	マルチスタディライセンス、許可された研究を無制限に実施できるアクセスオプション	英国
NCATS	National Center for Advancing Translational Sciences	国立トランスレーショナル科学高度化センター	米国
NCI	National Cancer Institute	国立がん研究所	米国
NCPeH	National Contact Point for eHealth	e-ヘルスの国内連絡窓口機関	EU
NDB	National Database	レセプト情報・特定健診等情報 データベース	日本
NDC	National Drug Code	全米医薬品コード	米国
NDOO	National Data Opt-Out	ナショナルデータオプトアウト	英国

略語一覧 (5/6)

略語	スペル	説明	対象国
NEDS	Nationwide Emergency Department Sample	全国救急局サンプル	米国
NHS	National Health Service	国民保健サービス	英国
NHIN	Nationwide Health Information Network	全米医療情報ネットワーク	米国
NIH	National Institutes of Health	国立衛生研究所	米国
NIHR	National Institute for Health and Care Research	国立健康研究所	英国
NIS	Network and Information Systems	ネットワークと情報システム	EU
NREAP	National Research and Ethics Advisors’ Panel	英国研究倫理審査諮問委員会	英国
NSCLC	Non-Small Cell Lung Cancer	非小細胞肺癌	共通
N3C	National COVID Cohort Collaborative	全米COVIDコホート共同研究	米国
OCR	Office of Civil Rights	人権保護局	米国
ODS	Organization Data Services	固有の識別コード (ODS コード)および付随する参照データを発行および管理するサービス	英国
OMOP-CDM	Observational Medical Outcomes Partnership Common Data Model	分析用標準化を実現するオープン規格	共通
ONC	The Office of the National Coordinator for Health Information Technology	医療IT全米調整官室	米国
ONS	Office for National Statistics	国家統計局	英国
OPCRD	Optimum Patient Care Research Database	最適患者ケア研究データベース	英国
PACS	Picture Archiving and Communication Systems	医療用画像管理システム	共通
PCD	Patient Confidential Data	患者機密データ	英国
PHI	Personal Health Information	個人医療情報	米国
PHR	Personal Health Record	個人の健康・医療・介護に関する情報	共通
PII	Personally Identifiable Information	個人を特定できる情報	共通
PMDA	Pharmaceuticals and Medical Devices Agency	独立行政法人医薬品医療機器総合機構	日本
PMSI	French National Administrative Database	退院サマリー	フランス
PQRS	Physician Quality Reporting System	医療品質の報告を円滑に行うシステム	米国
QHIN	Qualified Health Information Network	認定医療情報ネットワーク	米国
QI	Quality Improvement	品質向上	共通
RCE	Recognized Coordinating Entity	認定調整機関	米国
REC	Research Ethics Committees	研究倫理審査委員会	英国
RIF	Research Identifiable Files	識別可能ファイルの調査	米国
RIPH	Research project Involving Human Subjects	対人の研究	フランス
RSNA	Radiological Society of North America	北米放射線学会	米国

略語一覧 (6/6)

略語	スペル	説明	対象国
RWD	Real World data	リアルワールドデータ	共通
RWE	Real World Evidence	リアルワールドエビデンス	共通
SCDM	Sentinel Common Data Model	Sentinel共通データモデル	共通
SDE	State Designated Entity	非政府状態指定機関	米国
SDE	Secure Data Environment	安全なデータおよび調査分析プラットフォーム	英国
SDRC	State Data Resource Center	州データリソースセンター	米国
SGSS	Second Generation Surveillance System	COVID-19 第 2 世代監視システム	英国
SID	State Inpatient Databases	州入院患者データベース	米国
SNDS	National Health Data System	フランス国立医療データベース	フランス
SNIRAM	French Health Insurance System Database	健康保険データベース	フランス
SOC	Sentinel Operations Center	センチネルオペレーションセンター	米国
SOP	Standard Operating Procedures	標準作業手順書	共通
SPE	Secure Processing Environment	安全な処理環境	EU
SPF	Santé publique France	公衆衛生局	フランス
SSA	Social Security Administration	社会保障局	米国
SUS	Secondary Uses Services	NHSが提供する医療データの二次利用サービス	英国
TEFCA	Trusted Exchange Framework and Common Agreement	信頼された交換フレームワークと共通契約書	米国
TEHDAS	Towards the European Health Data Space	欧州の健康データ空間に向けた共同行動	欧州
THIN	The Health Improvement Network	健康向上ネットワーク	フランス
THL	Finnish Institute for Health and Welfare	フィンランド国立健康福祉研究所	フィンランド
TRE	Trusted Research Environment	トラステッドリサーチエンパイロメント、Visiting解析環境	英国
VBPM	Value Based Payment Modifier	バリューベースペイメントモディファアー	米国
VDI	Virtual Desktop Infrastructure	仮想デスクトップ	共通
VDW	Virtual Data Warehouse	仮想データウェアハウス	米国
VRDC	Virtual Research Data Center	仮想リサーチデータ環境	米国
WG	Working Group	ワーキンググループ	共通

用語一覧、報告書の記載

用語一覧

本資料では、次の用語を下表の定義にて使用する

NO	用語	定義
1	(医療情報の)一次利用	患者本人の医療的利益に資する目的での利用
2	(医療情報の)二次利用	その他の目的での利用
3	オプトイン	事前の通知・公表を行ったうえで、本人から同意を取得する方式
4	オプトアウト	事前の通知・公表を行ったうえで、本人から異議が無ければ同意があったものとみなす方式
5	データ主体	当該個人情報に帰属する本人(患者)のことを示す
6	スポンサー	臨床試験の開始、管理及び資金調達の設置に責任を負う個人、企業、機関又は組織のこと
7	監督機関	情報保護分野を監督する職権を有する機関
8	広範な同意	同意取得の時点では目的を完全に特定できない場合、一定の分野や範囲内の目的に情報を利用することに対して同意の意思を示すこと
9	仮名加工情報	個人情報を加工して他の情報と照合しない限り特定の個人を識別できないようにした情報
10	匿名加工情報	特定の個人を識別することができないように個人情報を加工し、当該個人情報を復元できないようにした情報
11	リビングウィル	人生の最終段階における医療やケアについて事前に本人の意思を表明または共有する方法で、書面による生前の意思表示
12	情報処理	情報に対して行われる、データ処理を含む操作の体系的実施。データ通信、オフィスオートメーションなどの操作
13	データ処理	データに対して行われる操作の体系的実施
14	情報システム	情報処理システムとこれに関連する人的資源、技術的資源、財的資源などの組織上の資源とからなり、情報を提供し配布するもの
15	情報処理システム	データ処理システム及び装置であって情報処理を行うもの。事務機器、通信装置などを含む
16	データ保護	故意もしくは偶然によって権限をもちに行われるデータの開示、変更又は破壊に対する適切な管理的、技術的又は物理的な防御手段の実現
17	情報管理	情報処理システムにおいて、情報の取得、分析、保存、検索及び配布を制御する機能
18	データ管理	データ処理システムにおいて、データに対するアクセス、データの記憶の実行・監視及び入出力操作の制御をする機能
19	データバンク	利用者が参照できるように構成された特定の事柄に関するデータの集合
20	データベース	複数の適用業務分野を支援するデータの集まりであって、データの特性とそれに対応する実体との間の関係を記述した概念的な構造に従って編成されたもの
21	相互運用性	それぞれの機能単位に固有な特性に関する知識を利用者がほとんど又は全く必要とせず、各機能単位が互いに通信し、プログラムを実行し又はデータを転送する能力
22	端末	システム又は通信ネットワーク内の機能単位であって、データを入れたり取り出したりするもの

報告書の記載

※No.14以降、日本産業規格(JIS)の情報処理に関する用語に準拠

- 本報告書に記載されている情報は、調査時点のものであり、公開情報等を基礎としております。これら入手した情報自体の妥当性・正確性については、責任を負いません
- 本報告書のヒアリング調査に記載されている内容は、ヒアリング対象者及びヒアリング対象機関からの回答を基礎としております
- 本報告書における分析手法は、多様なものがありうる中での1つを採用したに過ぎず、その達成可能性に関して、いかなる保証を与えるものではありません
- 本報告書が本来の目的以外に利用されたり、第三者がこれに依拠したりしたとしてもその責任を負いません

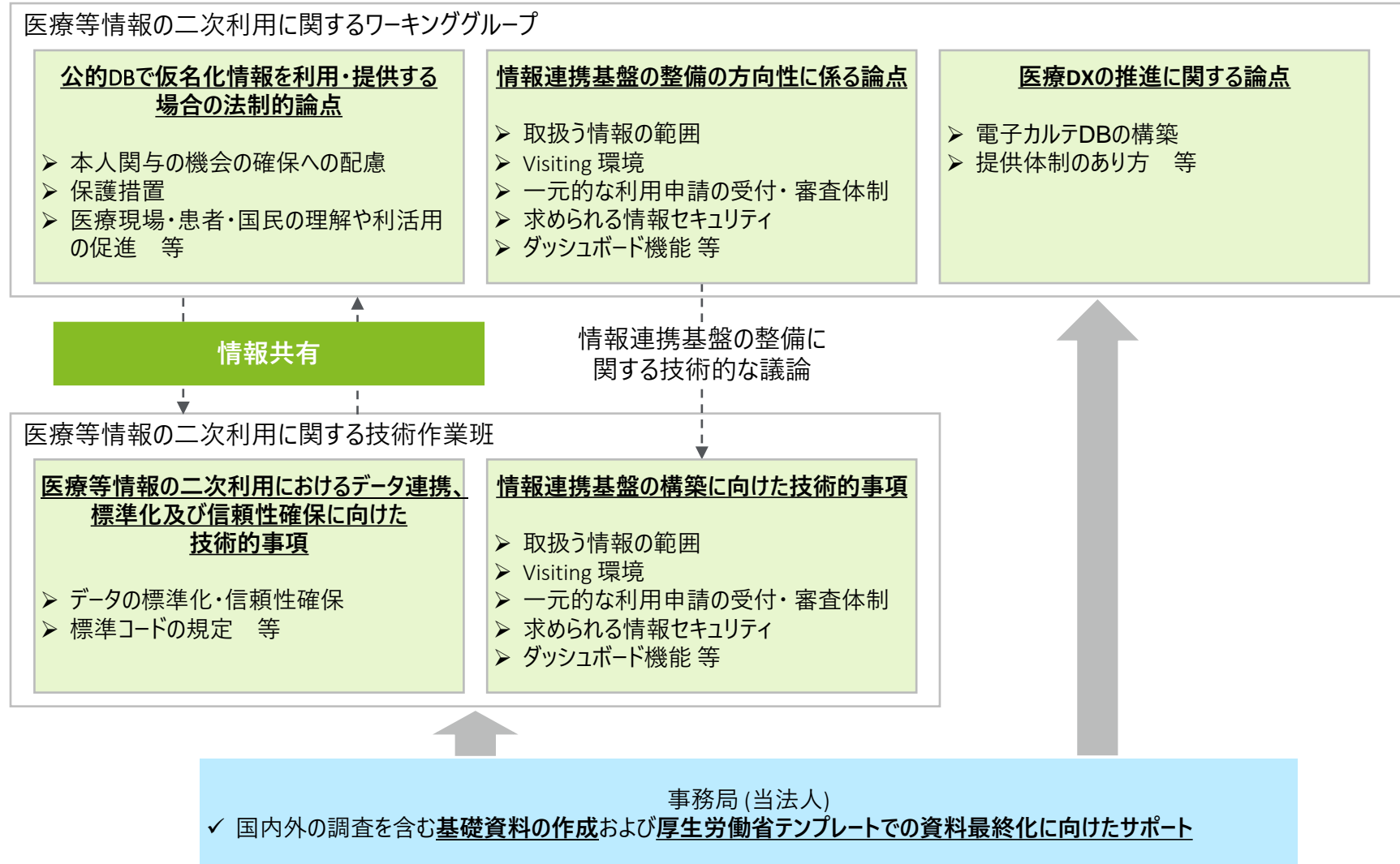
1. 会議の運営

1.1 技術作業班の会議運営

WGおよび技術作業班の円滑な会議運営に向け、海外事例調査および基礎資料の作成を実施した

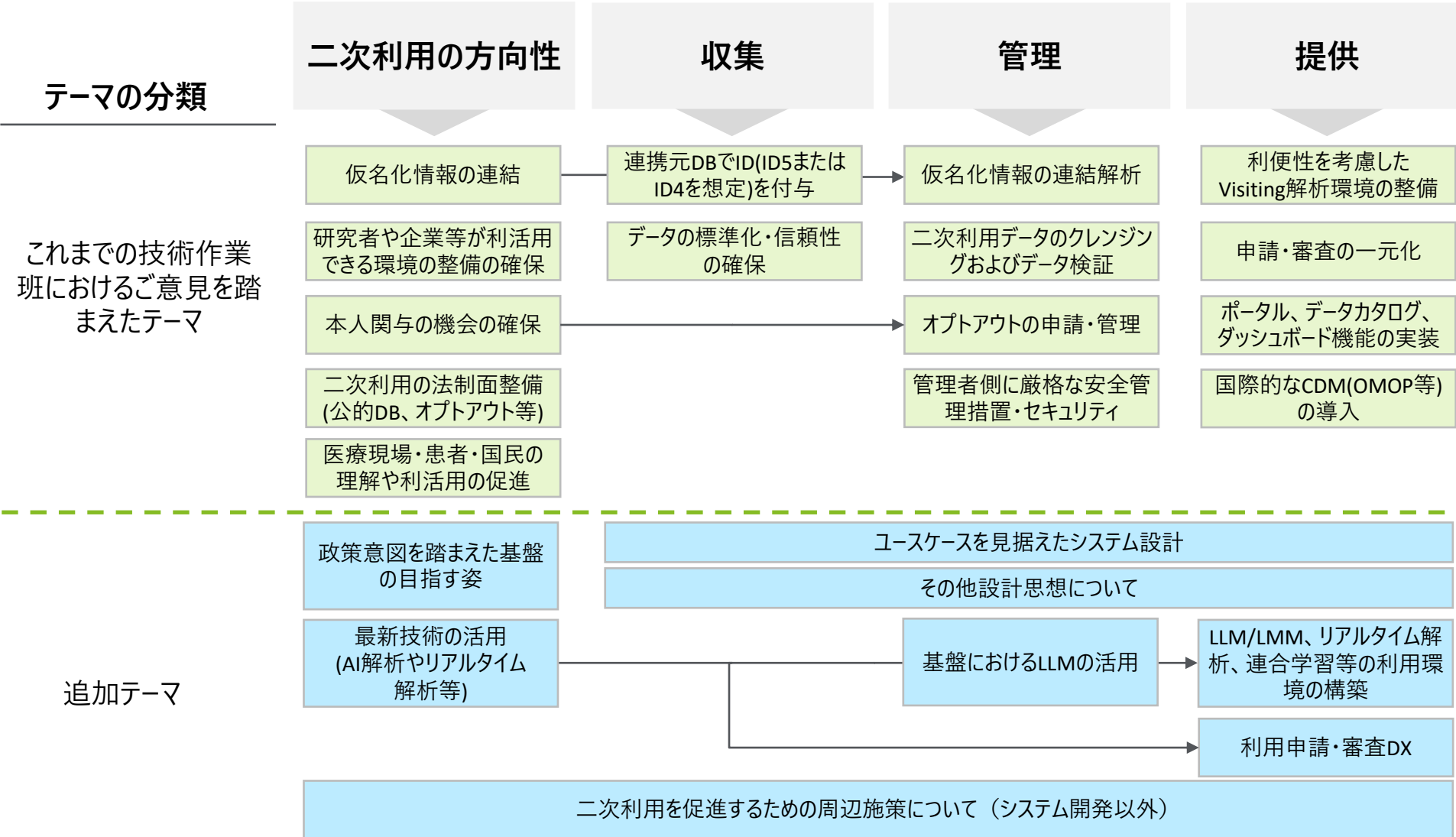
WGおよび技術作業班の関係性

会議における論点



これまでの技術作業班の議論を踏まえたテーマは以下の通りであり、令和7年度の会議運営においては、政策意図を踏まえた基盤の目指す姿や最新技術の活用といったテーマを中心に議論した

技術作業班のテーマ



1回あたりの会議の開催に向けた準備スケジュールと役割は概ね以下の通り実施した

WGおよび技術作業班の準備スケジュール

技術作業班1回あたりの流れ

	開催準備	構成員への事前レク	WG/技術作業班の開催	記録の作成および構成員への確認
日程	4週間前 ※日程は構成員がより参加できるように事前に先の日程を調整	1週間前	前日～当日	会議後3日～5日以内
当法人の役割	<ul style="list-style-type: none">✓ 関係者(構成員、貴室等)との日程調整✓ 会議室の確保✓ 会場レイアウトの準備・手配✓ 開催に関する資料の送付<ul style="list-style-type: none">・ 開催案内の送付・ 出欠票の送付・ 謝金等の書類送付✓ 会議の資料作成	<ul style="list-style-type: none">✓ 事前レクチャーの日程調整✓ オンライン会議等の準備	<ul style="list-style-type: none">✓ 会議資料一式の送付✓ 会議URLの再通知✓ 当日の運営支援<ul style="list-style-type: none">・ 会場準備・ ウェブ会議等の操作✓ 議事録作成	<ul style="list-style-type: none">✓ 関係者への議事録の送付✓ 議事録(案)の最終化<ul style="list-style-type: none">・ 構成員への確認・ 関係参加者への確認✓ 意見の整理

1.2 技術作業班の概要および開催実績

技術作業班の構成員は、厚生労働省により医療情報の関係者、個人情報法制の学識経験者等から選定された

委員構成

厚生労働省により、構成員は下記のメンバーが選定された

氏名	所属
足立 昌聰	一般社団法人医療データベース協会 ディレクター 情報処理安全確保支援士・弁護士・弁理士
小笠原 克彦	一般社団法人日本医療情報学会代表理事
岡田 美保子	一般社団法人医療データ活用基盤整備機構理事長
岡村 浩史	大阪公立大学医学研究科臨床検査・医療情報医学講師
清水 央子	東京大学情報基盤センター客員研究員
田辺 里美	情報処理推進機構デジタル改革推進部主任研究員
山口 光峰	独立行政法人医薬品医療機器総合機構医療情報科学部長

敬称略・50 音順

ワーキンググループの日程確定後、以下の議題に基づき計4回の検討会運営を支援した

開催実績

	第3回	第4回	個別ヒアリング①	個別ヒアリング②
日時	2024年5月20日(月) 15:00～17:00	2024年6月17日(木) 13:00～15:00	2024年9月	2024年10～11月
場所	厚生労働省会議室	厚生労働省会議室	オンライン開催	オンライン開催
参加者 (構成員)	計7名（全員参加） 足立昌聰、小笠原克彦、岡田美保子、岡村浩史、清水央子、田辺里美、山口光峰	計7名（全員参加） 足立昌聰、小笠原克彦、岡田美保子、岡村浩史、清水央子、田辺里美、山口光峰	計7名（全員参加） 足立昌聰、小笠原克彦、岡田美保子、岡村浩史、清水央子、田辺里美、山口光峰	計7名（全員参加） 足立昌聰、小笠原克彦、岡田美保子、岡村浩史、清水央子、田辺里美、山口光峰
会議形式	オンライン開催	オンライン開催	オンライン開催	オンライン開催
議題	1. 電子カルテ情報の二次利用及び情報連携基盤の構築等について 2. その他	1. 電子カルテ情報の二次利用及び情報連携基盤の構築等について 2. その他	1. 政策意図を踏まえた基盤の目指す姿について 2. 最新技術の活用について 3. ユースケースについて	1. 情報連携基盤の目指す姿 2. 政策目的を達成する上でのユースケース及びシステムの考え方

1.3 技術作業班での主な意見

医療等情報の二次利用にあたり、医療機関における各マスターの管理手順やデータの標準化の重要性に関する意見があがった

技術作業班で頂いた主なご意見

小項目	構成員等からの主な発言内容
①：医療等情報の二次利用におけるデータ標準化、信頼性確保	<ul style="list-style-type: none">✓ 各医療機関に対し、<u>各マスターの管理手順のようなものを示すことができれば信頼性確保に有用</u>なのではないか。各医療機関においては、必ずしも知見がない担当者がデータ管理をしている状況があるため、ガイドのようなものがあると参考になる✓ 6情報はOMOPとの相性もよいと考えられる。国際共同研究もしやすくなると思うので、二次利用を進めていくことが重要✓ <u>データを利活用するためには標準化が非常に重要</u>ということはこれまでずっと指摘されている。なかなか進んでこなかった背景があるが、<u>今後進めるためには、イニシアチブをとり、要になる組織が必要</u>なのではないか。社会保険診療報酬支払基金が医療DXの推進の主体となるのであれば、医療等情報のルールの発信などを担っていくということが考えられるのではないか✓ <u>マスターの整備は重要であるのと同時に、時系列も含めてコードを整えることが必要</u>。例えば、薬剤のYJコードでも、販売会社が変わればコードが変わる。経時的な分析を行う際に、コードが変わっていてできないことがあるので、時系列管理も念頭に整備していくことが必要✓ 民間DBにおいては、マスタの整備やデータの読み替えに、多数の工程をかけて行っている。効率的な処理を行っていくということには賛成。また、データの欠けは、情報連携基盤上では補えないので、<u>入力時点、医療機関において標準化されたものが入力されていることが大事</u>。データの上流で、医療機関に入力していただくことと、情報連携基盤側でマスターの整備などを行うこと、しっかり分けて議論をすることが大事✓ データの品質に関して、<u>レギュトリーの目的に耐えうる品質を担保していくためには、MID-NETの取組が参考</u>になる。この作業班でも議論を深めていくべき✓ 薬剤でも検査でも、コードが様々ある。それぞれ何のコードが使われるのかということを念頭に議論することが重要。ICDについては、ICD10からICD11に変わっていく議論もあると思うが、ICD10とICD11は仕様が大きく違い、ICD11を使うことで滞る面もあると思うので、注意が必要。JLAC10についても、JLAC11に変えるのはなかなか難しい面があり、関係者と議論していくことが重要。公的DBとして品質保証する部分は二次利用の部分だと思うので、<u>範囲をよく理解して議論することが必要</u>ではないか✓ データの標準化については、ある程度ルールを決めて、<u>入力側が強制力やメリットを感じるような形でグリップを行っていくことが必要</u>ではないか

情報連携基盤に具備する機能や既存のHICとの関係性、セキュリティの在り方について 様々な意見があがった

技術作業班で頂いた主なご意見

小項目	構成員等からの主な発言内容
②：情報連携基盤に関する検討事項	<div>✓ HICを拡張するのか、新たな情報連携基盤を作るのかという部分で違う部分はあるかと思うが、いずれにしても安全管理措置の部分はHICに寄せればよいのではないかと。一方、利用者側に求める要件については、利便性を考慮して、利活用者が使いやすいように、何が使いにくいのかをHICを参考にしながら、簡略化できるようにしていくのがよいのではないかと。</div> <div>✓ セキュリティの点では、ガバナンスを標準化する観点だと、HICのセキュリティの監視の部分は踏襲して共通的に使っていくというのがあり得るのではないかと。新たに基盤を作るということであれば、管理も二重三重になり、セキュリティホールになり得る。クラウドであれば、比較的安全にゾーンを区切れるので、それぞれのゾーンで目的にあった使い方をしていくことが考えられる。Visiting環境については、許されたユーザーだけアクセスできるということであれば、ガバナンスも担保できるし、利便性もあがると思う。</div> <div>✓ セキュリティの要件を考えるに当たっては、リスクベースで議論することが必要。Visiting環境でどのような情報が扱われるのかというのが議論の前提になると思われるので、議論の順序については配慮が必要。</div> <div>✓ セキュリティは安全に倒すと、いくらでもコストがかかるし利便性も下がる。絶対に守るべきものと、ある程度流通してもいいものとで格付けをつけて、セキュリティ対策を行う方向性もある。また、ゼロトラストという概念があるが、専用線を使うと高価になるため、ネットワーク上はゼロトラストということで信頼せず、利活用者の身分を証明していただくことによるセキュリティ対策を講じていくべきではないか。</div> <div>✓ 技術的なセキュリティだけでは難しい部分もある。利活用者の倫理感を醸成するため、一定の講習を受けていただくような対策も重要。</div>

1.4 技術作業班委員の個別ヒアリング

ヒアリング資料

ヒアリング目的および項目は以下の通り。これまでの技術作業班での意見を踏まえ、情報連携基盤の目指す姿や運用面・システム面に関する意見収集を実施した

ヒアリング概要

- これまでの技術作業班での意見を踏まえ、情報連携基盤の運用やシステムに関する検討を進めてきた。これに基づき、情報連携基盤の目指す姿やユースケース、技術面に関する意見を収集することを目的として、委員へのヒアリングを実施した
- このヒアリングを通じて、現状のシステムの課題や改善点を整理にし、今後の運用方針や技術的な方向性を具体化していくための意見を収集した

【情報連携基盤の目指す姿】

- 情報連携基盤の目指す姿として、現場や研究機関等のニーズや期待に沿った目標設定を念頭に考えているが、目指す姿の方向性に修正が必要な点・その理由や、具体的な改善点について
- 政策目標を踏まえた、データ利用者、データ提供者の視点の妥当性

【政策目的を踏まえたユースケース】

- 医療等情報の二次利用の政策目標を達成するにあたり、重要と思われるユースケースはどのようなものか
- ユースケースの更なる精緻化に向けて、具体的な改善点について

【情報連携基盤でデータ提供のあり方としてスコープとするべき技術】

- 情報連携基盤の目的、ユースケースを踏まえると、最新技術（例: LLM/LMM、リアルタイム分析、連合学習等）の活用も考慮されると思われるため、情報連携基盤で対応すべき技術について
- それら技術の有効性や導入に際しての課題、また他国の事例など

【情報連携基盤でのデータ管理等におけるLLMの活用、また、利用申請・審査DX】

- 昨今の技術の進展を踏まえると、情報連携基盤内にLLMを取り入れ（スクラッチ開発等）、情報の加工や利活用に生かしていくという方向性も考えられるが、いかがか
- 利用申請・審査にDXの視点を取り入れる（あり方の検討、デジタル化、AI等導入）について

【情報連携基盤における設計思想】

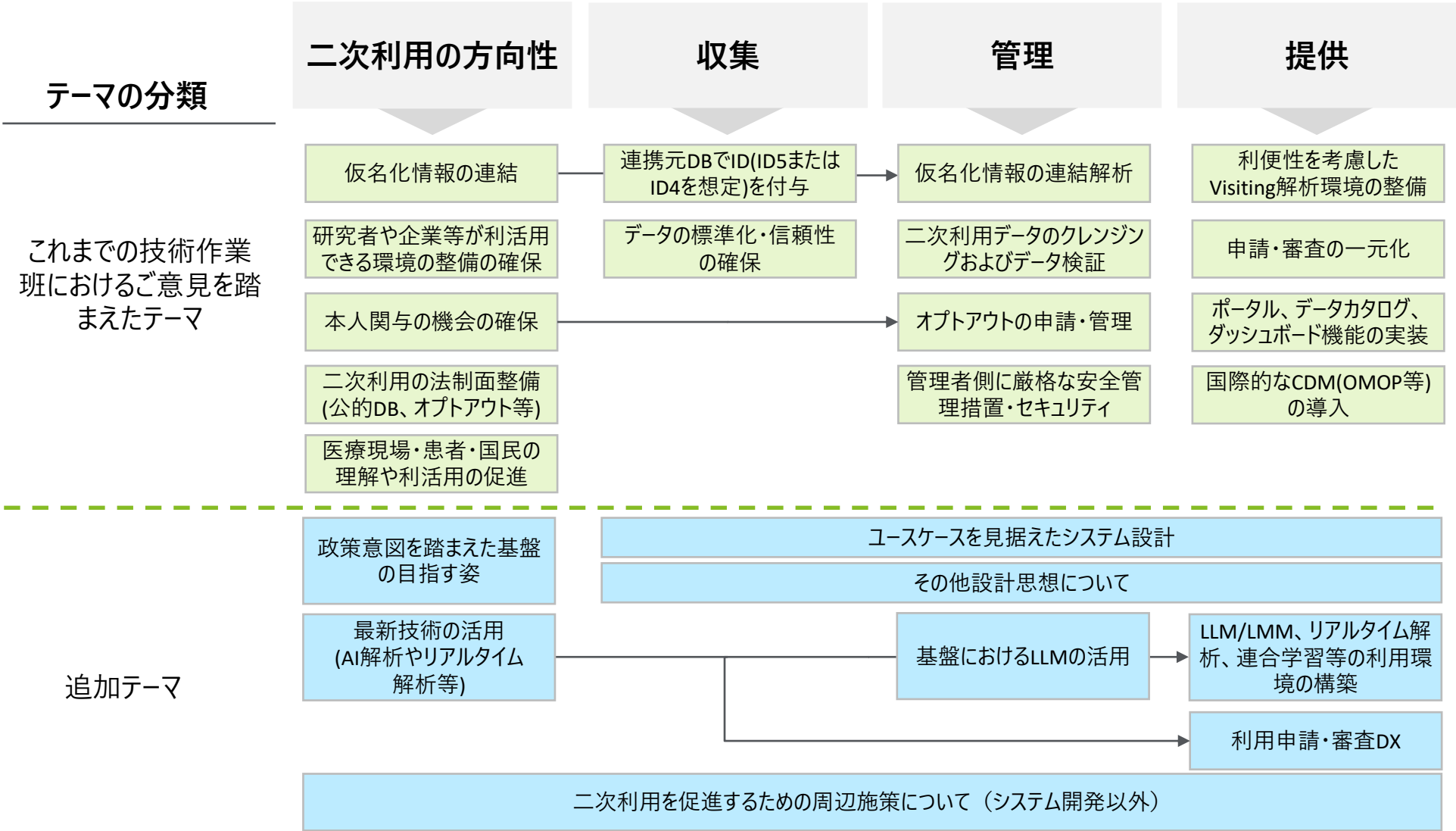
- 情報連携基盤における設計思想として、「ユーザー中心設計」「セキュリティバイデザイン」「ゼロトラストアーキテクチャ」などを想定
これらの設計思想について、他に取り入れるべき設計方針や特に重要視すべき点があるか

【二次利用を促進するための周辺施策について（システム開発以外）】

- 例示した内容やそれ以外に実施すべき項目等

これまでの二次利用の技術作業班の議論を踏まえたテーマは以下の通りであり、本日は主に追加テーマについてご意見をいただきたいです。本日は、これまでの議論で出てこなかった論点も含めて、長期的な目線で情報連携基盤の理想像について考えたく、忌憚ないご意見をお願いいたします

情報連携基盤で踏まえるべきテーマ



情報連携基盤の目指す姿

背景・目的

- 医療等情報の二次利用を通じて、「国民の健康増進、より質の高い医療・ケア、医療の技術革新（医学研究、医薬品開発等）、医療資源の最適配分、社会保障制度の持続性確保（医療費の適正化等）、次の感染症危機への対応力の強化などにつなげていく（規制改革実施計画（令和5年6月16日閣議決定））」必要がある。
- 一方、医療等情報の活用に当たっては、医療現場の理解を得つつ、本人の権利利益を保護し、国民の患者の理解を得ながら、適切に二次利用を推進する必要がある
- 日本では、仮名化情報の利用の制約や、データベースの整備や利用申請手続きに係る負荷が多く、リアルワールドデータの研究利用が進みにくい状況にある。これらの課題を解決し、医療等情報を適切に活用するための環境整備が求められている

目指す姿

データ利用者

以下のデータ利用者の医療等情報の二次利用を通じて、政策目標の実現を目指す

政策	医学研究	医療	開発およびイノベーション	健康管理/教育
政府・自治体 <ul style="list-style-type: none">EBPM による政策立案と検証・改善感染予防や流行時の危機管理対策の強化医療費や予算分配の最適化による財政持続可能性の検討	学術・研究機関等 <ul style="list-style-type: none">臨床データの解析による疾病リスク因子、疾患予後予測の解明長期データ連携による治療・薬剤の影響評価国際共同研究の推進AIアルゴリズム研究の推進	医療機関 <ul style="list-style-type: none">診療や治療方針の向上のための教育としてデータ利用研修医や医学生の臨床教育やケーススタディ	民間企業 （製薬・医療機器会社等） <ul style="list-style-type: none">臨床情報の解析による市販後調査等、医薬品・医療機器の安全性評価研究・イノベーション推進による産業育成	患者・教育機関 <ul style="list-style-type: none">医療・健康に関する統計データ参照自身の健康管理や病気の予防を目的とした教育コンテンツ統計学習への活用

情報連携基盤の構築
公的DB等の情報の利用や連結解析が可能な環境を整備



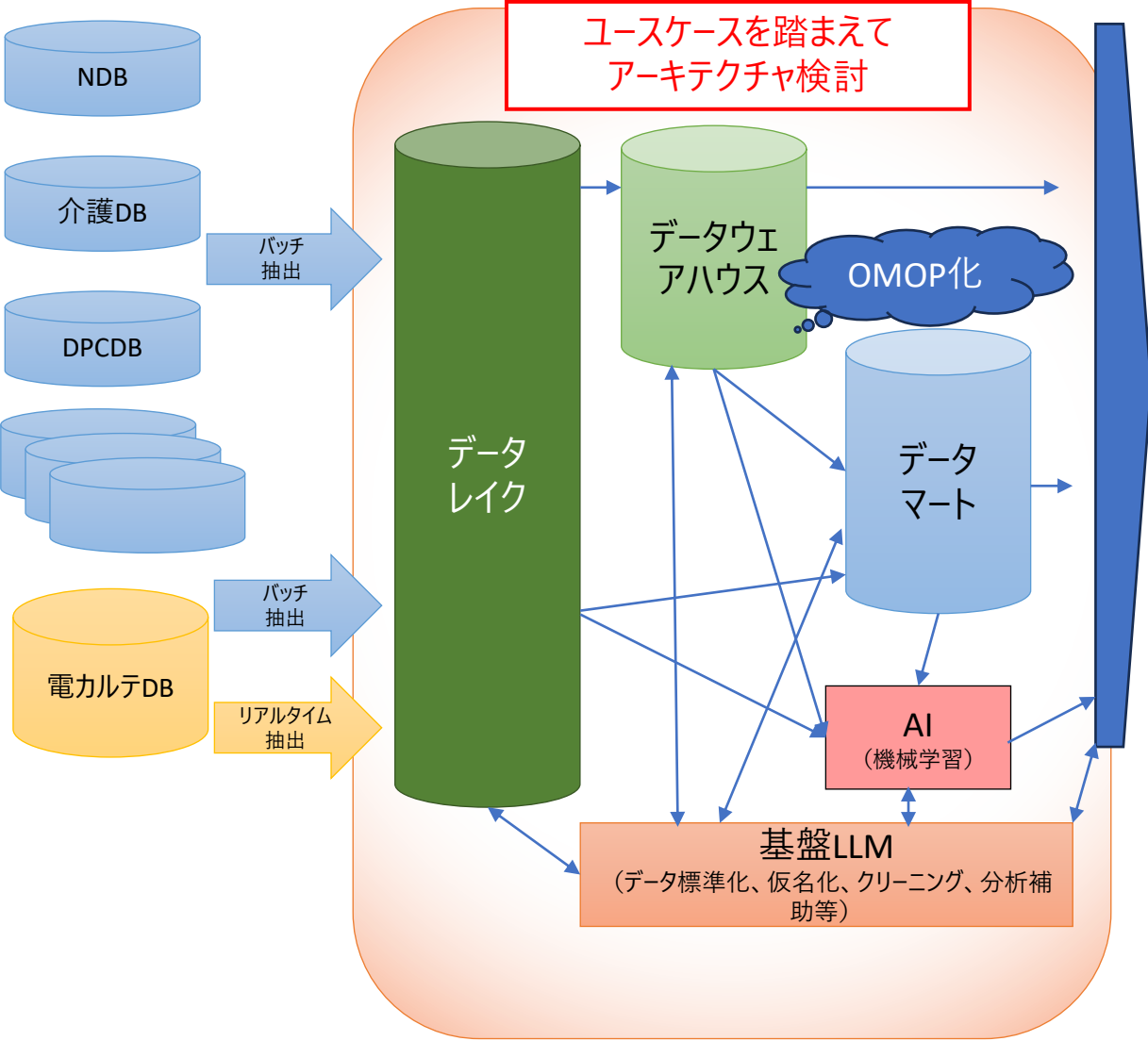
医療機関等・患者

- 安全なデータの活用
- 理解増進
- データ提供に対するインセンティブの付与(研究の成果フィードバック等)
- 権利利益の保護

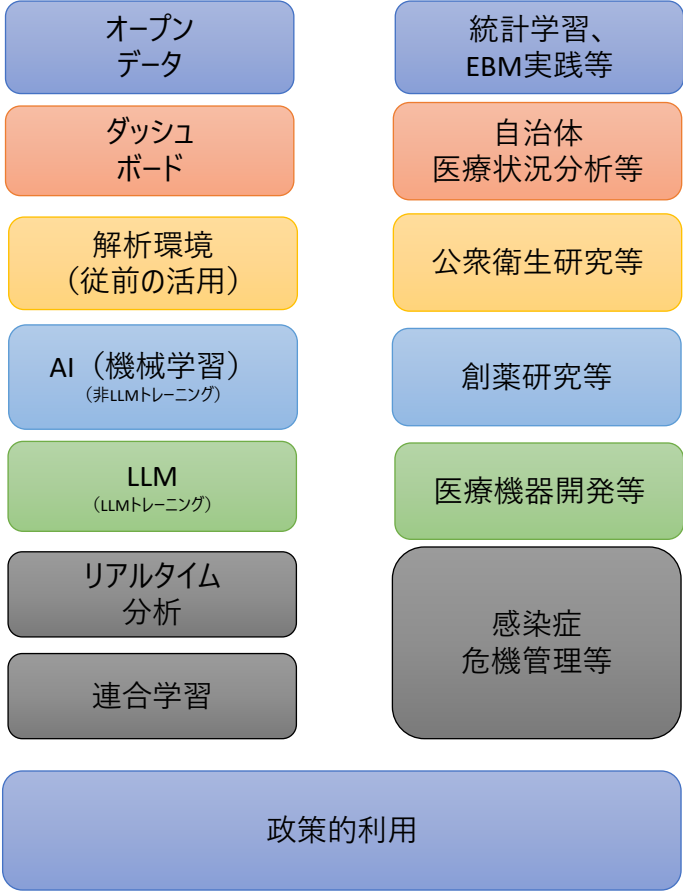
円滑なデータ提供等により、医療等情報の二次利用が可能な環境を構築する

データ提供者

ユースケースを見据えたシステム設計（イメージ）、最新技術の活用(AI解析やリアルタイム解析等)



《ユースケース例》



参考：各公的DBの保有データ

請求系データベース				臨床データ系データベース
NDB	介護DB	DPCDB	障害福祉DB	電子カルテ情報DB (仮称)
<ul style="list-style-type: none"> 傷病名（レセプト病名） 投薬 健診結果 等 	<ul style="list-style-type: none"> 介護サービスの種類 要介護認定区分 ADL情報等 	<ul style="list-style-type: none"> 傷病名・病態等 施設情報 等 	<ul style="list-style-type: none"> 障害の種類 障害の程度 障害支援区分認定データ 障害福祉レセプト 等 	<p>【3文書】</p> <ul style="list-style-type: none"> 診療情報提供書 退院サマリー 健康診断結果報告書 <p>【6情報】</p> <ul style="list-style-type: none"> アレルギー 感染症 薬剤禁忌 検査 処方 傷病名

疾患系データベース				
全国がん登録DB	難病DB	小慢DB	感染症DB	予防接種DB
<ul style="list-style-type: none"> がんの罹患 診療内容 転帰 等 	<ul style="list-style-type: none"> 告示病名 生活状況 各種検査値 等 	<ul style="list-style-type: none"> 告示病名 発症年齢 各種検査値 等 	<ul style="list-style-type: none"> 感染症の名称・症状 診断方法 初診年月日・診断年月日 発病推定年月日等 	<ul style="list-style-type: none"> ワクチン情報 接種場所 副反応の症状 等

情報連携基盤における設計思想等

目標設定

設計思想

設計原則

- 公的DB等を一元的かつ安全に利用・解析できる「情報連携基盤」を構築する
- データ収集にあたっては、データの連結性を高め、データの正確性を確保し、定期的な更新頻度を維持することが求められる
- 二次利用しやすいデータベースを構築するため、収集したデータのクレンジングを行う等の品質管理が求められる

アーキテクチャパターン

- 将来的な変更や拡張を容易にし、必要に応じて個別のサービスをスケールアップできる柔軟性が求められる

開発・運用組織と管理

- 開発・運用を見据えた体制等

セキュリティ

- 情報システムの利用環境、攻撃者や手法の変化に対応するため、ゼロトラストセキュリティモデルの考え方も取り入れたアーキテクチャで設計・構築する
- 常時のシステム監視・ログの分析により不審な挙動を早期発見し、未然防止や被害の局所化を可能とする
- データ利用者の利便性も考慮する

- ① - ユーザー中心設計 (User-Centered Design)
- ② - デザインシンキング (Design Thinking)
- ③ - データ駆動設計 (Data-Driven Design)
- ④ - ドメイン駆動設計 (Domain-Driven Design: DDD)

- ⑤ マイクロサービスアーキテクチャ (Microservices Architecture)
- ⑥ イベント駆動アーキテクチャ (Event-Driven Architecture: EDA)
- ⑦ 疎結合アーキテクチャ (Loose Coupling Architecture)
- ⑧ サービスの独立性と疎結合 (Service Independence and Loose Coupling)
- ⑨ APIファースト設計 (API-First Design)

- ⑩ 継続的インテグレーションとデリバリー (CI/CD)
- ⑪ インフラストラクチャ・アズ・コード (Infrastructure as Code, IaC)
- ⑫ クラウドセンターオブエクセレンス (Cloud Center of Excellence, CcoE)
- ⑬ エンタープライズアーキテクチャ (Enterprise Architecture)

- ⑭ セキュリティバイデザイン
- ⑮ ゼロトラストアーキテクチャ

二次利用を促進するための周辺施策について（システム開発以外）

1. 教育プログラムの開発と普及

医療従事者や研究者、企業関係者や行政関係者等を対象に、データ二次利用に関する教育プログラムを開発・提供し、データ活用の理解とスキルを向上させる

2. データ利活用者のチェンジマネジメントの強化

医療等情報の二次利用を促進するために、データを活用する組織内での変革を円滑に進めるためのリーダーシップ育成やコミュニケーション戦略を強化し、デジタルトランスフォーメーション（DX）を推進する

3. 組織体制の整備と強化

医療等情報を政府として効果的に活用するための組織体制について検討する。一例としては、政府、医療機関、企業、学術機関が協働するコンソーシアムを設立し、データ二次利用を促進する体制を検討する

4. 社会的理解と信頼の醸成

広報活動や市民参加型イベント等を通じて、データ二次利用の意義や安全性を周知し、社会的信頼を高める

情報連携基盤を通じた政策目標の実現として、大規模データ解析や創薬の研究・開発など医療等情報のさらなる活用に期待が寄せられた

ヒアリング結果概要（1/5）

二次利用を通じた政策目標の実現

- ✓ 都道府県の医療データ分析事業を進めており、レセプトデータベースを構築する研究を進めている。これにより高齢者の認知症有病率や分布、多剤併用の状況を明らかにすることができると考える。データベースを構築するにあたっては先ず、行政や行政と連携するコンサル会社・大学が活用しやすいように整えるべき
- ✓ 請求系のデータには非常に多くの情報が含まれており、もっと広く活用すべき（例：病院・在宅医療のデータのリンク）
- ✓ 大規模なDBの存在が医療DXをさせ、カルテ入力支援等のAI技術に活用される可能性がある。また、個人情報保護の観点から日本の医療情報が国際的なAI開発に活用されにくい状況にあると思う
- ✓ 製薬会社は利益を追求しながらも良い製品を開発する責任があるが、利益追求が悪く見られる偏見がある。データは開発目的だけでなく、さまざまな研究に活用すべき。また、製薬会社は質の高いデータを必要としており、AIやゲノムの前に、データの質の向上が優先されるべき

情報連携基盤の目指す姿

- ✓ データ利用者・提供者の観点について、データが充実すればより価値は高まっていくと考えるが、国としてNDBを集約していくにあたり、自治体ベースのEHRは不要になっていくのか、それとも住み分けがされていくのか、役割分担や場合によっては連携についても検討すべきである
- ✓ NDB等を連携するにあたり、スマートフォンのデータやPHRのデータを紐づけて使いたいという要望も出てくるのではないかと
- ✓ 情報連携基盤で収集したものを次世代医療基盤法の枠組みを活用して連携できるようになる等、次世代医療基盤法との接点も検討すべきである
- ✓ ユーザーのリテラシー水準が違うため、自治体や医療機関などのユーザーは情報連携基盤を直接扱うことが難しいことが想定される。統計解析やクエリを直接操作できる「プロユーザー向けのツール」を提供するの、BIツールのような定型的な解析を希望する「リテラシー水準が低めのユーザー向け」にするのか、どちらに焦点を当てるかが重要である
- ✓ 医療データは患者主体の視点から扱うべきであり、患者の立場を基にデータの利用方法を整理し進めるべきであるとする
- ✓ 国民に対し、データ利用の利点を理解してもらうだけでなく、医療データの扱いについての理解を深めてもらう必要がある。そのため、医療データの利活用が患者にもたらす利益を明確に示すべきである

想定ユースケースにおいては、行政、アカデミア、企業などで様々な利用例があげられた

ヒアリング結果概要（2/5）

想定されるユースケース

【政策・行政】

- ✓ 感染症モニタリングへの活用、薬剤耐性菌の地理的な分布などの可視化などの公衆衛生の向上への活用
- ✓ 自治体職員も簡単にアクセスできるオープンデータやダッシュボードの整備や、コンサル会社の介入などが有益と考える

【アカデミア】

- ✓ OMOP（コモンデータモデル）を用いた国際共同研究の促進への活用
- ✓ 研究と政策を連携させ、データを政策に落とし込むことが重要。この際、誰が司令塔となるかが重要。医療経済研究機構や国立保健医療科学院が中心となり、大学と連携する方法が考えられる
- ✓ 従来は基礎研究などに基づいて仮説を構築していたが、今後はデータドリブンによって仮説を立てる取組がますます活性化していくことが想定される。仮説やシーズを自動生成し、再現性、因果推論、メカニズムの探求に結び付けるようなケースが想定できる。仮説を生むところまでは技術の発展で可能になると考えるが、メカニズムの探求は動物実験に委ねるしかないとする
- ✓ 仮想人体生成モデルへの活用として、臨床検査のデータに欠損値が多い場合、AIが使いづらいという課題がある。そこで、患者の性別や年齢などの基礎情報を入力し、それをもとに欠損値を自動補完する仮想人体生成モデルというものがある。ビッグデータがあればそういったものも、より正確な数値が出せるようになり他の研究にも使えるようになるのではないだろうか
- ✓ 臨床意思決定支援システム(CDSS)への活用として、リアルタイムでのAPI連携が可能であれば、情報連携基盤のデータをソースに診療領域別にCDSSを作成し、何かあれば臨床医にアラートが行くような使い方も考えられる

【企業】

- ✓ 臨床研究、市販後調査、メディカルアフェアーズなどが主なユースケースとして想定
- ✓ 特に、新薬の探索的研究では、データベースを活用したシミュレーションが可能である。また、適応拡大、業務効率化、有害事象（AE）の管理、市販後調査の改善にも活用が期待できる
- ✓ AI機器開発について連携する公的データベースには臨床データや検査値が不足しており、AI向けの活用には限界がある。また、生活習慣に関連するデータも不足しているため、10個の公的データベース以外のデータとの連携ができるかどうかが課題である
- ✓ 診断補助系のAI（SaMD等）は増えているが、直接診断を行うタイプのAIには、さらに詳細なデータが必要であり、特に画像や波形データの活用が求められている

情報連携基盤に組み込むべき技術として、LLMやLMMなどのAI活用の重要性が伺えた

ヒアリング結果概要（3/5）

情報連携基盤でデータ提供のあり方としてスコープとするべき技術

- ✓ **NDBのデータは文字や数値が中心だが、画像データも取り込むことで国際的な競争力を高める必要がある。**また、日本の医療情報は詳細であり、この強みを活かせば高精度なデータレイクやデータマートを構築できると思う
- ✓ **LLM、LMMは基盤での技術活用と、研究ユースケースでの活用の両方が重要**である
- ✓ **他の基盤との連携**が考えられる。例えば診療領域別の目的でつくられたAIを連携させていくようなこともできればと思う（参考企業：Sakana AI、生成AI（人工知能）スタートアップ企業）
- ✓ **NDBには悉皆性があるものの、同一施設が一貫して認識されない問題があり、マスターデータ管理が不十分。**ICD-10からICD-11、将来的なICD-12への移行に伴い、**時系列管理とデータの正確な紐付けが重要。**また、経営者変更の情報が反映されないケースもあるため、**データベースの有効活用にはマスター管理の強化が求められている**
- ✓ AI開発のための**計算資源を情報連携基盤の内外にどう配置するかにより、機密性を保つための工数が変わる**ため、慎重な検討が必要である
 - 計算資源が基盤外に配置される場合、理論上は連合学習や他のデータベースと組み合わせた解析が可能であるが、それはプロユーザー向けの機能であり、**標準機能として提供するには過剰である**と考えられる
- ✓ AI開発ではローデータに戻って検証を行う必要があるため、匿名化や仮名化されたデータでも元の個別データにアクセスできないと開発が進まないという課題がある。そのため、**自由度の高いサンドボックス環境でローデータを参照できる開発環境を整備する必要があり、解析環境とは区別して考えるべきである**
- ✓ 生成AIについて米国の省庁では、申請プロセスや規制業務において、生成AIを利用することで申請書の評価や電子カルテの要約などで業務効率化が実現されている。今後、AI技術のさらなる発展が見込まれる中で、**業務支援分野においても生成AIを積極的に活用すべきである**と考えられる
- ✓ **JLACの検索機能を強化し、病院が提供する情報をダウンロードできる仕組みを構築**すべきである。
- ✓ 諸外国では、**検査内容に対応する標準コードを付与して返す機能の導入**は技術的に可能であり、特に検査や薬剤に関する標準化が重要であるため、これを導入する価値がある
- ✓ **標準コードのサーバーを取り入れる**ことは、効率的な運用のために重要であるとする

データ利用時の審査については、DX化を進め、審査側、利用者側の負担軽減するよう業務効率化への意見があげられた

ヒアリング結果概要（4/5）

情報連携基盤でのデータ管理等におけるLLMの活用、また、利用申請・審査DX

- ✓ DXは不可欠だが、国民にそのメリットをしっかりと説明する必要がある。特にマイナンバーカード導入に関しては色々騒がれているためフィジビリティや研究を進めながら丁寧に説明することが求められる
- ✓ 今後、僻地医療におけるAIアシストが不可欠になると思う。特に医療従事者が減少する可能性に備え積極的に取り組むべき
- ✓ 現在DVD等で情報が送られているが、将来的にはデジタルの利点を活かしてリアルタイムに情報にアクセスできる環境が必要だと感じている。アクセス権限に応じて利用範囲を区別し、現場の医師が診療時に過去の症例を参考にできるようなシステムが理想的。これにより、最新の医療情報へのアクセスが可能になり、診療の質向上に繋がる。データの利用範囲に自由度を持たせ、より開放的で利用しやすい環境を整える必要がある
- ✓ 機能を解説できるLLM、プロンプトで研究したい内容を入れると提案してくれるようなLLMがあればよい
- ✓ 審査のオンライン化については「なぜやらないのか」という姿勢であり、現在のPDF変換作業をやめてオンライン化を進めるべき
- ✓ 重要なのはプロセスの短縮で、申請からデータ提供までに1年以上かかるのは問題。14日での提供は難しくても、数カ月でデータを提供できるデジタル化が必要だと考えている
- ✓ 申請・審査の一元化はしたほうがよく、政策決定に時間がかかってしまうようなことを避けるため公的利用のケースは分けて設計したほうがよいと考えるが、その際機関指定ではなく業務指定とできないだろうか。法律で委託されている業務の範囲などは簡略化可能という立て付けがよい
- ✓ 全体感はよいと思うが、構築コストが膨らむことが懸念される。構築している間にも技術は発展するので、何を盛り込むべきかを慎重に判断すべき。AIをどうするかが一番の争点かと思うが、例えば医薬品の安全対策など政策に取り入れられるということにすれば国費が使いやすいのではないかと
- ✓ クリーニングについてはキー情報を明確にする必要があるためまずそのルール作成が必要と考える
- ✓ 情報連携基盤の構築に際し、特別な法制度に基づいてユニークな情報を入力させる必要があると考えられる。入力ボリュームが大きくなり、途中で入力結果を保存できなければ、作業負担を軽減するにはcsv形式などで自動的にデータを取り込む仕組みが効果的である可能性が高い。ユーザー側からは、形式審査の自動化よりも入力項目の削減が求められる可能性がある。ただし、審査の透明性は非常に重要であり、さらに不服審査に対応できるエビデンスの確保も不可欠である。
- ✓ 生成AIを利用することで申請書の評価や電子カルテの要約などで業務効率化を図り、業務支援分野において生成AIを積極的に活用すべきであると考えられる

二次利用にあたっては、二次利用に関するメリット等を周知していくことが重要な方針であるといった意見があげられた

ヒアリング結果概要（5/5）

二次利用を促進するための周辺施策について（システム開発以外）

- ✓ マイナンバーカード普及に伴い、医療情報への関心が高まっており、メリットを広めるためにイベントや研究が必要
- ✓ 医師会を通じて待合室にポスターを掲示することが効果的である。「皆さんの診療が新しい医療技術に活用されています」「皆さんのデータが貢献しています」といったメッセージを伝えることが、データ活用に対する理解を深めるための第一歩ではないか
- ✓ 教育プログラムの開発やデータのユースケースに関しては、既に使える人は活用しており、「こうやって使える」というアプローチでは不十分である。データ活用の具体的なメリットを示すことが重要で、反対意見に妥協して使えないデータにしてしまうのは避けるべき。成功例や有益な事例を積極的に示し、良い成果をアピールすることで、反対派の理解を得る方が効果的だと考える
- ✓ オプトアウトを導入しない場合、公衆衛生上のメリットを強調し、その意義を十分に説明することが必要である。一方、オプトアウトを導入する場合、手続きコストやオプトアウトの機会を保証することがKPIとなり、方針に大きな影響を与える可能性がある。そのため、オプトアウトを導入しない方針を早期に明示することが望ましい
- ✓ 日本の医療情報標準化は十分に普及しておらず、具体的な目標や方向性が不明確である。標準の認定は進んでいるが、現場での普及が遅れている。二次利用に関わる標準化政策を進めるためには、国家としての一貫した枠組みと明確な方針が必要である

2. 諸外国の医療情報の取扱いについての把握・分析

2.1 諸外国の調査

2.1.1 EHDS

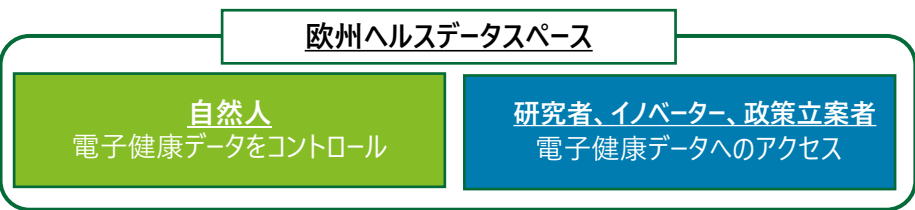
EHDS規則は自身の電子健康データへのアクセスや権利の行使、各加盟国の不均一なGDPR実施や解釈、規格の違い等に対応する欧州特有のデータ共有の枠組みである

欧州ヘルスデータスペース(EHDS)の概要

■ EHDS提案の背景

- GDPRで自身のデータ(健康データを含む)に対する自然人の権利が保護されているにもかかわらず、国内および国境を越えた電子健康データへのアクセスや送信等、自身の電子健康データに対する権利の行使が困難である
- 加盟国ごとの不均一なGDPR実施や解釈が電子健康データの二次利用の障壁になっている
- 規格の違いによる限定的な相互運用性により、デジタルヘルス分野において別のEU加盟国への参入を妨げている
- COVID-19パンデミックで健康上の緊急事態対応における電子健康データの重要性がより一層示された

■ EHDSの目的



- EU域内の自然人に対して、自身の電子健康データのコントロールを保証する
- EUおよび加盟国のガバナンス機構と安全な処理環境による法的枠組みを構築する
- 規則を調和させ、デジタルヘルス製品とサービスの真の単一市場に貢献し、医療システムの効率性を高める

EHDS規則案	
正式名称	Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space
公表日	2022年5月3日
関係組織	欧州委員会、欧州保健デジタル庁
概要	<u>健康に特化した欧州の共通データスペース</u> で、電子健康データへのアクセスと共有に関する健康固有の課題に対処するための提案(説明覚書より)
関連法規制	GDPR、データガバナンス法案、データ法案、NIS指令、EUサイバーセキュリティ法、EUサイバーレジリエンス法案

EHDS導入の準備状況、スケジュール

■ 一次利用のインフラ(MyHealth@EU)

➢ EHDS規則案以前からの努力義務のため、既に一部で開始されており、2025年までに25か国が段階的に参加を予定している

➢ 2023年2月時点で一次利用が可能な国は11か国(エストニア、オランダ、クロアチア、スペイン、チェコ、フィンランド、フランス、ポーランド、ポルトガル、マルタ、ルクセンブルク)であるが、利用可能なサービス状況は異なる

■ 二次利用のインフラ(HealthData@EU)

➢ 2022年10月から2年間の予定でパイロットプロジェクトを実施中

• EU4Healthプログラムにより各国の健康データインフラ(HDH、Findata等)や欧州医薬品庁、欧州疾病予防管理センター等、計16組織でコンソーシアムを組成

• 役割は以下の2点

- ✓ データソースプラットフォーム(Node)のネットワークを開発、展開
- ✓ 二次利用のインフラをEU全域に展開するための実現性、関心、能力を評価

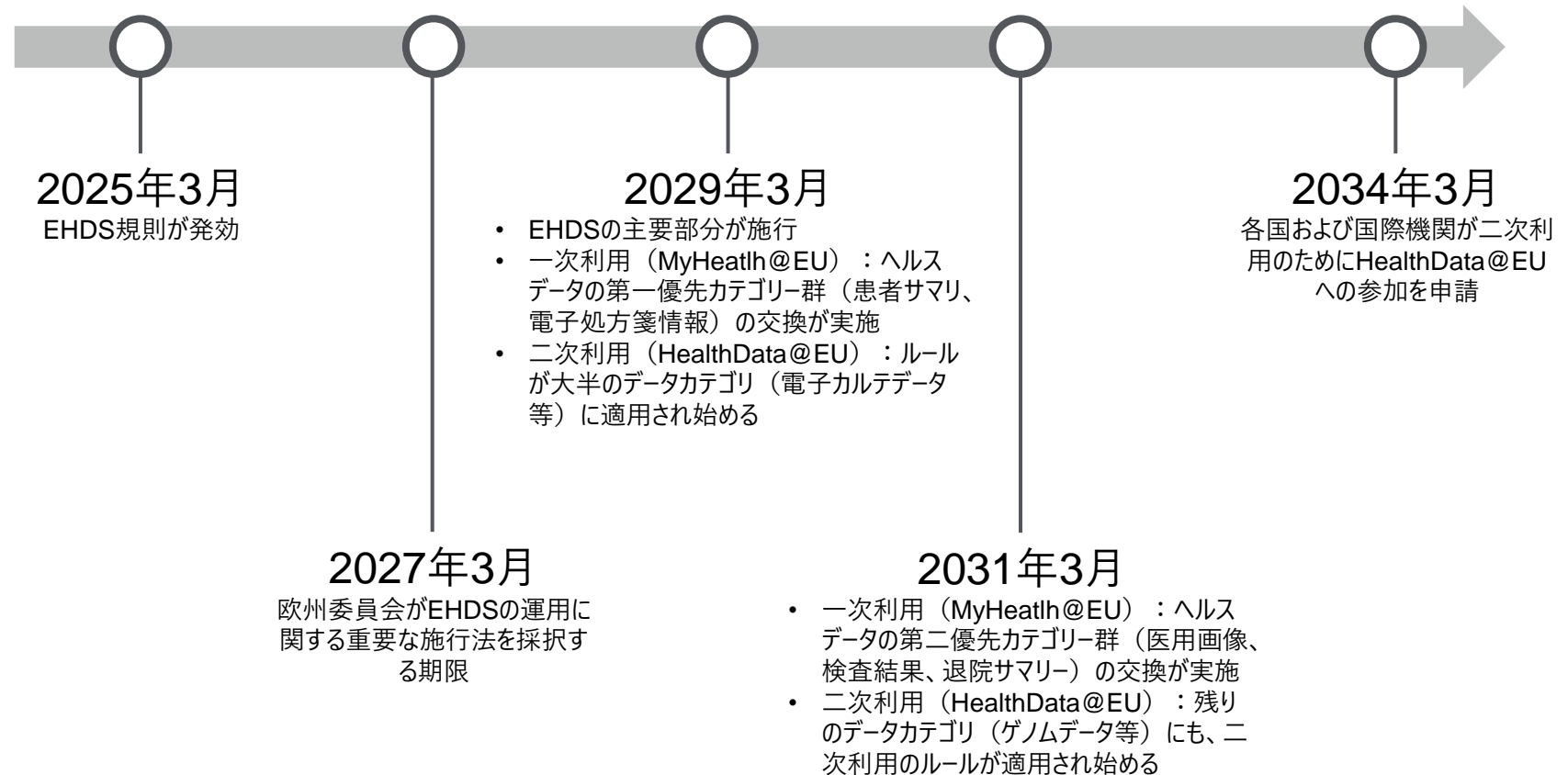
2024年4月、欧州議会・欧州評議会によるEHDS修正案の採択に関する合意がなされ、2025年3月に公布された

EHDS規則の最新動向（1/2）

2022年05月03日	欧州委員会 (European Commission) によりEHDS規則案が公表される 規則案はEU内のデータ保護基準を順守しつつ、国境を越えた共有を含め、欧州市民が自身の医療データを管理できるようにすること (一次利用)や、研究、技術革新、公共政策目的の医療情報の利活用 (二次利用) の促進を目的としている
2023年06月14日	EHDS規則案に係る初めての議論が雇用・社会政策・保健消費者問題評議会によって行われる 総じて規則案は歓迎されるが、 <u>医療情報の保護を確保するためのセキュリティ体制について再検討する必要性</u> が提起される
2023年02月10日	欧州議会 (European Parliament) がEU GDPRの関連規定との整合性を確立した修正案を発表 主な追加項目は以下の通り： <ul style="list-style-type: none">・ 二次利用のためのオプトアウトに係わる規定・ 医療情報の不正利用時の金銭的罰則等の法的措置の設定による説明責任の強化・ 医療情報をEU域内に保存する義務に係わる規定・ EHDS Boardによるさらなる包括的なガバナンス規定
2023年02月16日	EHDS修正案の評価をENVI (環境・公衆衛生・食品委員会) 及びLIBE (市民自由・司法・内務委員会) が担うことが決定 欧州議会での審議に向け、欧州議会規則第58条(Parliament Rules of Procedure)に基づきENVI及びLIBEにEHDS修正案の内容を確認・評価することが決定される
2023年11月28日	ENVI及びLIBEによるEHDS修正案の見解文章 (Draft Position) がMEPs(欧州議会の代表者)で仮採択 ENVI及びLIBEによるEHDS修正案の見解文章 (Draft Position) が欧州議会の代表者で95票賛成、18票反対、10票棄権の結果で仮採択される。ENVI 及びLIBE の見解文章 (Draft Position) の主要項目は以下の通り： <ul style="list-style-type: none">・ 二次利用にあたり、特定の機微情報を利用する場合は明示的な同意を必須とし、それ以外の場合はオプトアウトの仕組みを設ける・ 市民に、健康データアクセス機関に異議を申し立てる権利を与える。その場合、非営利団体が代理で申し立てることも可能とする・ 二次利用が禁止されるケースの拡充 (例: 労働市場や金融市場での利用)
2023年12月6日	欧州議会(European Parliament)がEHDS修正案を採択 欧州議会は <u>EHDS修正案を516票賛成、95票反対、20票棄権の結果で採択した</u> 。今後欧州評議会(European Council)で規則案に関する最終的な議論が行われる予定
2024年4月24日	欧州議会と欧州評議会によるEHDS修正案の採択の合意
2025年3月5日	2025年3月5日、EU連合官報に正式に掲載。2025年3月26日に発効し、適用に向けた移行フェーズが始まる

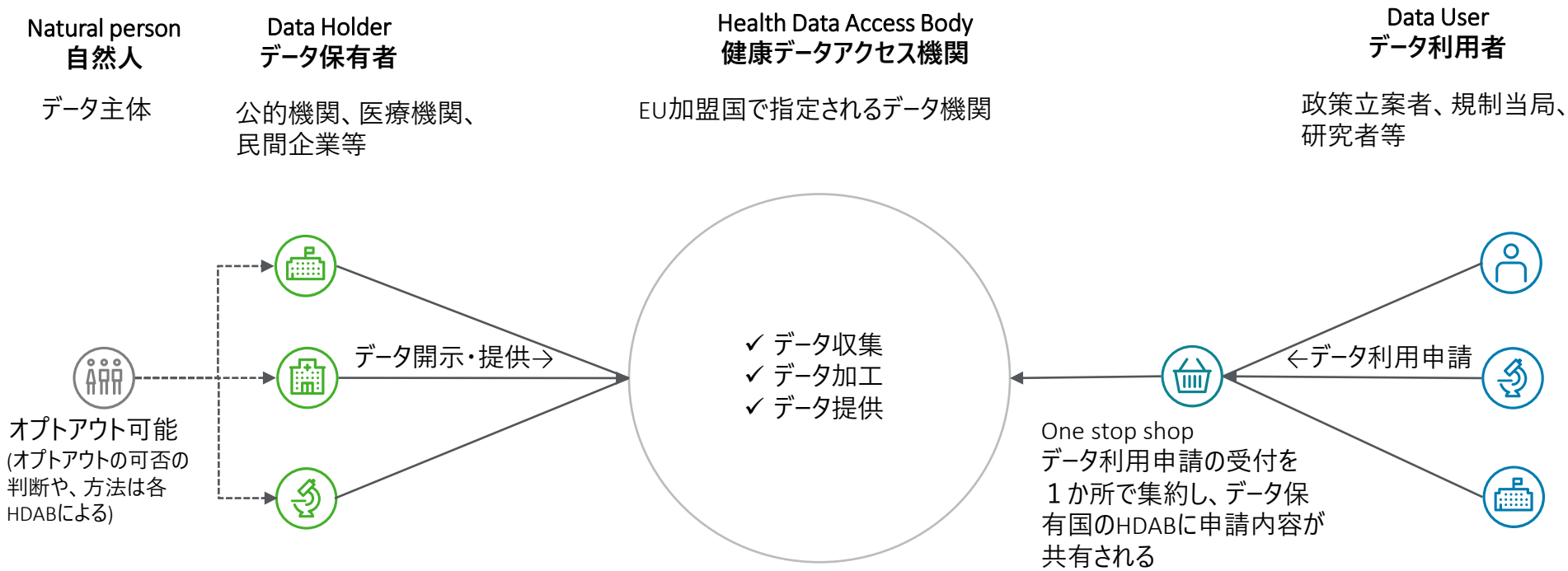
EHDS規則は2025年から段階的に調整され、二次利用においては、2029-2033にかけてデータカテゴリへのルール適用、各国での二次利用の参加申請は2034年に予定されている

EHDS規則の主なマイルストーン（2 / 2）



2025年3月5日にEHDS規則が公表されたが、今後、implementation actを制定して具体的な施策を検討する

EHDS規則のイメージ



GDPRの法的根拠	データ保有者	Health Data Access Body	データ利用者
GDPR第6条1項	(c) 法的義務: 各国のHDABへのデータ提供義務 (EHDS第41条1項)	(e) 公共の利益	(e) 公共の利益 (f) 正当な利益
GDPR第9条2項	(h) 医療関係、(i) 公衆衛生、(j) 研究活動		

出所: EHDS規則前文37 <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197>

データカテゴリにウェルネス・アプリケーション、死亡レジストリ、臨床研究・臨床調査等からのデータが追加された

第33条 二次利用電子データの最小区分(1/2)

「2024年4月24日 EHDS規則修正案」に基づく内容

条文番号	概要
1	<p>健康データ保有者は、本章の規定に従って、以下のカテゴリの電子データを二次利用できるようにしなければならない</p> <ul style="list-style-type: none">a. <u>EHRからの電子健康データ</u>b. <u>社会的・経済的、環境的、行動的な健康決定要因を含む、健康に影響を与える要因に関するデータ</u>ba. <u>医療ニーズ、医療に割り当てられる資源、医療の提供及び医療へのアクセス、医療支出及び資金調達に関する集計データ</u>c. 人の健康に影響する<u>病原体データ</u>d. <u>調剤・請求・償還データ等のヘルスケア関連管理データ</u>e. ヒトの遺伝子、<u>エピゲノム及びゲノムデータ</u>ea. <u>プロテオミクス、トランスクリプトミクス、メタボロミクス、リポドミクス及びその他のオミックスデータ等のヒトの分子データ</u>f. 医療機器を通じて<u>自動的に生成される個人の電子健康データ</u>fa. <u>ウェルネス・アプリケーションからのデータ</u>g. 自然人の治療に関与する医療従事者の<u>職業的地位、専門性、所属機関に関するデータ</u>h. <u>人口ベースの健康データレジストリ（公衆衛生レジストリ）</u>i. 医療レジストリと<u>死亡レジストリからのデータ</u>j. <u>規則(EU)536/2014、規則(SOHO)、規則(EU)2017/745、規則(EU)2017/746の対象となる臨床試験、臨床研究、臨床調査からのデータ</u>k. 医療機器からの<u>その他の健康データ</u>ka. <u>医薬品及び医療機器のレジストリデータ</u>l. 健康に関する研究コホート、アンケート及び調査から得られたデータ(最初の結果発表後)m. バイオバンク及び<u>関連データベースからの健康データ</u>

二次利用電子データについて、加盟国はデータのカテゴリの追加、データ処理及び仕様に関する規則を定めること、さらに追加的な保護措置を国内レベルで導入することができる

第33条 二次利用電子データの最小区分(2/2)

「2024年4月24日 EHDS規則修正案」に基づく内容

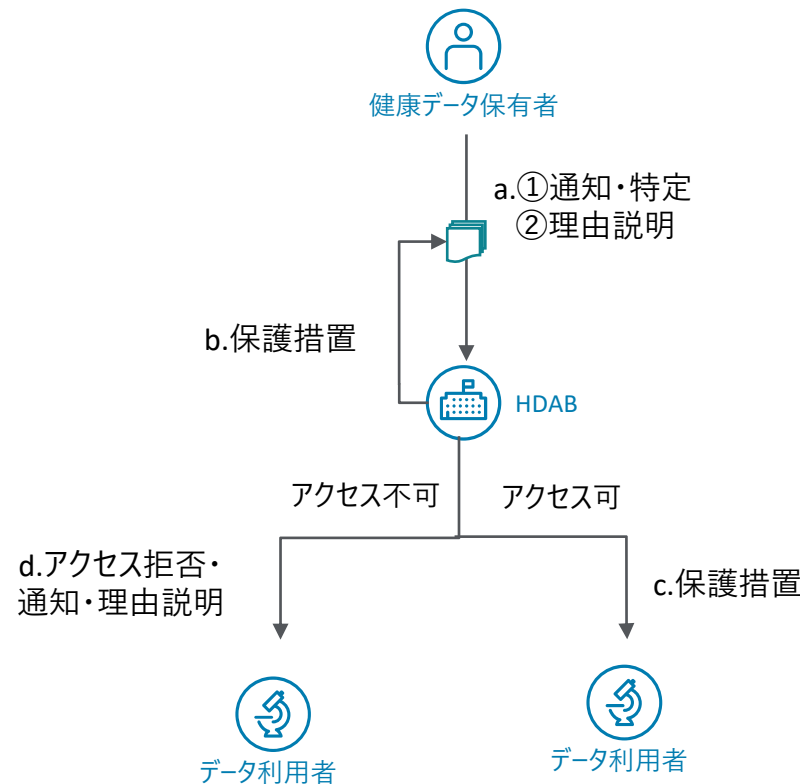
条文番号	概要
6	公的機関が緊急事態においてデータを入手する場合、HDABの支援を受けながらデータ処理または他のデータとの結合のための技術支援を提供できる
8	<u>加盟国は国内法によって電子健康データの追加カテゴリを二次利用に規定することができる</u>
8a	<u>加盟国は、訂正、注釈及びエンリッチ等電子健康データの処理及び使用に関する規則を定めることができる</u>
5	<u>加盟国は、データの機密性と価値を保護することを目的としたより厳格及び追加的な保護措置を国内レベルで導入することができる。それらの保護措置等を遅延なく欧州委員会に通知しなければならない</u>

知的財産権、企業秘密等の規定に保護される電子健康データは規定に従って二次利用のために利用可能とされることが追加された

第33a条 知的財産権と企業秘密 (新設)

「2024年4月24日 EHDS規則修正案」に基づく内容

1. 知的財産権、企業秘密等に保護される電子健康データは、二次利用のために利用可能とされるものとする。その場合、以下が適用されるものとする
 - a. 健康データ保有者は自己が保有するデータセットの説明を医療情報アクセス機関に通知する際に、遅くともHDABからの要請があった際に以下の情報を提供しなければならない
 - ① 該当電子健康データを、HDABに通知し、特定する
 - ② データセットのどの部分に該当するかを示し、そのデータがなぜ特別な保護を必要とするのかを説明する
 - b. HDABは、該当電子健康データの保護に特定の適切かつ比例的な措置をすべて講じなければならない
 - c. データアクセス許可証を発行する際、HDABは、特定の電子健康データへのアクセスを法的、組織的、技術的措置に条件付けることができる。このような措置は健康データ保有者と健康データ利用者との間の契約上の取決めを含めることができる
 - d. データのアクセスにデータ保護権を侵害するような重大なリスクを伴う場合、HDABはその点において健康データ利用者へのアクセスを拒否することができる。HDABは、この拒否を健康データ利用者に通知し、アクセスを提供できない理由を説明するものとする。健康データ保有者及び健康データ利用者は、第38b条に従って苦情を申し立てる権利を有する



HDABのタスクにデータ収集・処理・提供及び情報公開について内容が追加された

第37条 HDABのタスク

「2024年4月24日 EHDS規則修正案」に基づく内容

分類	条文番号	概要
データ収集	第37条(1)(a),第37条(1)(j)	<ul style="list-style-type: none">■ 健康データ保有者からの要求された必要なデータの<u>受領</u>■ 品質保証
データ処理	第37条(1)(a),第37条(1)(d),第37条(1)(f)	<ul style="list-style-type: none">■ データ処理<ul style="list-style-type: none">✓ データ保有者から収集したデータの結合、<u>編集、仮名化または匿名化等の加工</u>■ 知的財産権や企業秘密等データ保護の措置
データ提供	第37条(1)(a),第37条(1)(m)	<ul style="list-style-type: none">■ <u>データアクセスの許可・提供</u>■ <u>健康データ利用者及び健康データ保有者が本規則に定める要件を遵守していることを監視・監督する</u>■ 安全な処理環境の提供
情報公開	第37条(1)(l),第37条(1)(q),	<ul style="list-style-type: none">■ 以下の情報を公開する<ul style="list-style-type: none">✓ 電子健康データソースと性質及び利用可能にするための条件を含んだデータセットカタログ✓ 健康データの申請及び要求✓ <u>健康データの申請に対する許可または却下 (発行30営業日以内)</u>✓ <u>コンプライアンス違反に対する措置</u>✓ 健康データ利用者から通知された結果✓ <u>第35e条(自然人に対するHDABの義務)を遵守するための情報システム</u>✓ <u>第三国または国際的な組織の国内連絡ポイントとの連絡に関する情報</u>
その他	第37条(1)(k),第37条(1)(n),第37条(1)(o),第37条(1)(r)	<ul style="list-style-type: none">■ データ収集・処理・提供を管理する管理システムの維持■ HealthData@EUを通じて、他の加盟国との二次利用の促進■ 欧州委員会への二次利用に関する助言■ 自然人に対する義務 (第35e条参照)

出所:https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ43/AG/2024/04-09/1299790EN.pdf; <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197>

新設された第35条fでは、全ての加盟国に二次利用に対するオプトアウトメカニズムの構築を義務付けているが、各加盟国は国内法で正当な例外を設けることが可能である

第35条f – 二次利用に対するオプトアウトの権利 (新設)

「2024年4月24日 EHDS規則修正案」に基づく内容

分類	条文番号	概要
オプトアウトの権利	第35条f (1)	<ul style="list-style-type: none">■ 自然人は、理由を述べることなくいつでも自己に関する医療情報の二次利用に対しオプトアウトする権利を有する■ オプトアウトの行使に関する取消しも可能とする
加盟国の義務	第35条f (2),(2a),(7)	<ul style="list-style-type: none">■ 加盟国は、アクセスしやすく理解しやすいオプトアウトメカニズムを提供し、自然人が医療情報の二次利用に関する意思表示できるようにする必要がある■ 個人がオプトアウトした場合、その個人に関連するオプトアウト後に収集した識別可能な新規データは二次利用に使用してはならない。なお、オプトアウト前のデータは影響を受けない■ データ主体の識別が不要な場合、オプトアウトの対応のために識別情報を収集する必要はない
例外条件	第35条f (3)-(6)	<ul style="list-style-type: none">■ 加盟国は、オプトアウトに対する正当な例外を国内法で設けることができる。その際、以下の全ての条件が満たされる必要がある<ul style="list-style-type: none">・ 公衆衛生分野に係る業務を遂行する公共機関、EU機関、又は団体等による申請であり、利用目的が第34条第1項(a)-(c)に該当するもの、又は公共の利益につながる科学研究であること・ データが他の手段でタイムリーに入手にできない場合・ データの申請者が第46条(1)(h)や第47条(2)(be) に基づく正当性を提示した場合■ 例外は基本的権利と自由の本質を尊重し、公共利益を達成するものでなければならない■ 例外は再識別の禁止等の要件を遵守し、国内法には個人の権利保護に関する規定を含む必要がある■ 加盟国は採用した国内法を委員会に通知し、改正がある場合も同様に通知する必要がある

(参考資料)

第35条f – 二次利用に対するオプトアウトの権利 関連条文

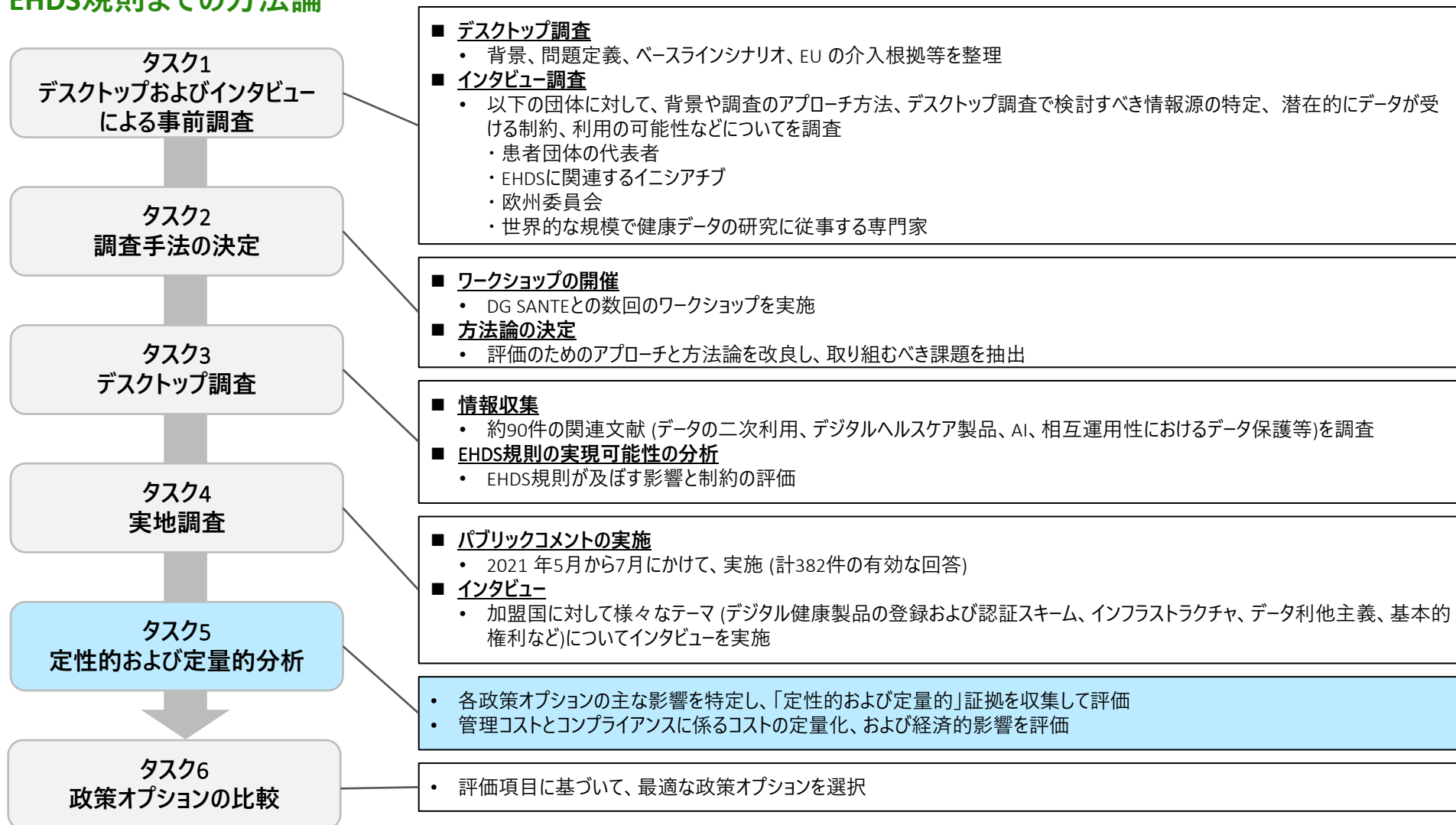
「2024年4月24日 EHDS規則修正案」に基づく内容

条文番号	概要
第34条(1)(a)-(c)	<ul style="list-style-type: none">■ データ申請者が以下の目的の遂行のためにデータを必要とする場合に限り健康データの二次利用を許可する<ul style="list-style-type: none">• 公衆衛生や労働衛生に係る公共利益（例: 国際的な健康脅威に対する保護・管理措置、医療機器や医薬品の安全性確保等）• 公的機関やEU機関等による政策立案や特定の規則・基準の策定等• 国際、EU、又は国家レベルの統計
第46条(1)(h)	<ul style="list-style-type: none">■ データ申請者が第35条(f)に基づくオプトアウトの例外申請を希望する場合は採択された国内法による根拠を必要とする
第47条(2)(be)	<ul style="list-style-type: none">■ データ申請者は申請時以下の情報を提供する必要がある：<ul style="list-style-type: none">• データの処理が GDPR 第6条(1) (Lawfulness of processing) または EU規則 2018/1725第5条(1) (Lawfulness of processing) 及び第10条(2) (Processing of special categories of personal data) にどのように準拠するかに関する説明

EHDS 二次利用の影響評価

欧州委員会は、以下のタスク1～6の手順でEHDS規則の取組を整理した

EHDS規則までの方法論



欧州委員会では、検討の結果、EHDSは「B/B+」オプションを選択した

ヘルスケアデータ活用に関する影響評価 – 政策オプション

政策オプションA: EUレベルでの調整の強化	政策オプションB / B+: 中強度の規制的介入	政策オプションC / C+: 高強度の規制的介入
<p>一次利用</p> <ul style="list-style-type: none"> ➤ 国家デジタルヘルス当局のネットワーク化の義務化 ➤ 電子健康データの相互運用性に関するガイドラインの策定 ➤ EHRシステム、デジタルヘルス製品、モバイルウェルネスアプリケーションの相互運用性の確保 	<ul style="list-style-type: none"> ➤ 権利と執行のための国家デジタルヘルス当局の指定 ➤ MyHealth@EUの義務化 ➤ EHRシステムと医療機器に対する最低要件の義務化 ➤ ウェルネスアプリの推奨仕様と認定およびラベリング ➤ EUレベルで確率された電子健康データに対する患者の権利 ➤ (Bのみ) 健康データを提供するEHRシステムと医療機器に対する自己申告型品質ラベルの義務付け ➤ (B+のみ) 健康データを提供するEHRシステム及び医療機器に対する第三者認証の義務付け 	<p>(オプションBに以下を追加)</p> <ul style="list-style-type: none"> ➤ EHRシステム、健康データを提供する医療機器、ウェルネスアプリの第三者認証の義務化 ➤ (Cのみ) 既存のEUの機関等(既存の機関の権限を拡大)を通じたEU内の調整 ➤ (C+のみ) 新たに設立されるEuropean Digital Health Bodyを通じてのEUの調整
<p>二次利用</p> <ul style="list-style-type: none"> ➤ 国家レベルでのデータアクセス付与のための所轄官庁の自主的協力 ➤ MyHealth@EUの二次利用への拡張 ➤ 健康データの二次利用に関するガイドラインやAI開発・検証に関する行動規範の策定 	<ul style="list-style-type: none"> ➤ 各国Health Data Access Bodyの指定 <ul style="list-style-type: none"> ・ EUが保有するデータへの各機関からのアクセス ・ 安全な処理環境でのデータ処理 ・ データ利他主義(data altruism) <ul style="list-style-type: none"> ※公益に資する目的のために、個人や組織がデータを提供するという概念 ・ 二次利用のデータアクセス、AI、データ品質に関するEUレベルでの専門家グループの組成 ➤ 二次利用のための新たなEUの分散型インフラへの参加義務化 ➤ 自己申告によるデータ品質ラベルの義務化やデータの Kategorisierung ➤ データ利他主義の監督 ➤ 二次利用のための欧州共通の法的目的、AIの開発や検証、標準化に関するHealth Data Access Bodyによる支援 	<p>(オプションBに以下を追加)</p> <ul style="list-style-type: none"> ➤ 二次利用のための新たなEUの中央集権型インフラへの参加義務付け ➤ Health Data Access Bodyによるデータ品質に関する最低要件を満たす第三者認証 ➤ EHDS上のすべての健康データは、セマンティック相互運用性要件 (Semantic Interoperability) を使用して構造化されること ➤ (Cのみ) 既存のEUの機関等(既存の機関の権限を拡大)を通じたEU内の調整 ➤ (C+のみ) 新たに設立されるEuropean Digital Health Bodyを通じてのEUの調整

政策オプションの評価は、以下の軸に基づいて評価を実施している

タスク5 定性的および定量的分析「評価項目」

EHDSにおける健康データの一次利用および二次利用のための政策オプションは、以下の項目で評価された

評価項目	内容
有効性 - 政策目標への貢献 (Effectiveness-Contribution to the policy objectives)	<ul style="list-style-type: none">■ 個人の健康データへのデジタルアクセスとコントロールの向上による個人のエンパワーメント■ デジタルヘルスおよび製品の単一市場の育成によるデータ経済の解放■ 研究、イノベーション、政策立案および規制活動への電子健康データの二次利用による効果
社会的な影響 (Social impacts)	<ul style="list-style-type: none">■ 一次利用: デジタルヘルスケアがもたらす健康の公平性への影響■ 二次利用: 特定目的のために細分化されたデータへのアクセスおよび高度なレベルのヘルスデータの共有
基本的な権利と自由への影響 (Impacts on fundamental rights and freedoms)	<ul style="list-style-type: none">■ 個人の基本的な権利と自由への影響
環境への影響 (Environmental impacts)	<ul style="list-style-type: none">■ 医療資源の利用効率化による環境への配慮
効率性 - 便益とコストの比較 (Efficiency-comparison of benefits and costs)	<ul style="list-style-type: none">■ 投資とコンプライアンス遵守に係るコスト: デジタルインフラ等の投資コストおよびガイドラインや要件等を含む専門家グループのコスト■ コスト削減と便益: 電子健康データがもたらす経済的価値や医療費の削減および時間■ 競争力と市場: 競争力の向上や市場に与える影響
一貫性 (Coherence)	<ul style="list-style-type: none">■ 内部的な一貫性: 他のEUの法的枠組み、政策、イニシアティブとの一貫性■ 外部的な一貫性: 国際的な法的、政策的枠組みとの一貫性■ 加盟国間での政治的な実現可能性

政策オプションBの評価は、提案するガバナンスシステムにより、データ共有の障壁を取り除き、研究やデジタルヘルス製品の開発が大きく促進すると評価されている

タスク5 定性的および定量的分析「政策オプションBの評価」

評価項目		評価結果の概要
有効性 - 政策目標への貢献 (Effectiveness-Contribution to the policy objectives)	定性	<ul style="list-style-type: none">✓ 各国のHDABの指定と専門家グループによるEUの協力は、研究目的でのデータ共有の障壁を取り除き、各国の研究チーム間の既存の競争に役立ち、EULEVELでの一貫した枠組みの構築を支援できる✓ 分散型インフラの構築は、データへのアクセスを容易にするため、デジタルヘルス製品の開発をより大きく促進する
社会的な影響 (Social impacts)	定性	<ul style="list-style-type: none">✓ 新たな医学的発見、より優れた予測能力、より優れた予防措置、大規模な健康リスクへの適応・最適化・対応能力の向上をもたらす可能性がある✓ ただし、適切な保護措置がとられない限り、マイナスの影響を及ぼす可能性もある。特に、セキュリティ違反や不正アクセスがあった場合、そのようなデータに悪影響を及ぼす可能性がある
基本的な権利と自由への影響 (Impacts on fundamental rights and freedoms)	定性	<ul style="list-style-type: none">✓ 国が指定したHDABが責任を負うことにより、個人データとプライバシーの保護にプラスの効果をもたらすことが予想される✓ ただし、ウェルネスアプリ、遠隔医療等からの個人データは、デジタルヘルス産業やデジタル産業（クラウド、アプリケーションプラットフォーム等）を含む第三者に提供されるため、適切なセーフガードなしでは個人データとプライバシー保護にマイナスの影響を与える可能性がある
環境への影響 (Environmental impacts)	定性	<ul style="list-style-type: none">✓ 政策決定（例：政策措置の改良に使用されるエネルギーと紙）および研究（例：追加データが利用可能になるとプロセスを複製しなければならないことによるデジタル汚染）に関連する様々なプロセスに必要な資源を削減することができる✓ ただし、研究・イノベーションを支援するための大規模なデジタルインフラの構築、大量のデータトラフィックとストレージ、デジタル機器の製造は、デジタル汚染につながる可能性がある
効率性 - 便益とコストの比較 (Efficiency-comparison of benefits and costs)	定性	<ul style="list-style-type: none">✓ 研究、新しい製品やサービスの開発、そしてよりエビデンスに基づいた政策立案や規制活動を支援することが期待される
	定量	<ul style="list-style-type: none">✓ データ技術による医療における革新的な製品やサービスの開発による医療データのさらなる再利用による患者、医療提供者、イノベーターの価値の向上から、今後10年間で係るインフラ費用（4〜7億€）に対して、54億€の大きな経済的なベネフィットを得られると推定される✓ デジタルヘルスとEHRシステムの製品は、年間20%〜30%の成長、ウェルネスアプリケーション市場は15%〜20%の成長が見込まれる
一貫性 (Coherence)	定性	<ul style="list-style-type: none">✓ EHDS規則案の目的に関して、加盟国間で幅広いコンセンサスが得られているようである✓ しかし、加盟国は自国の健康データへのアクセス要件に関する自治を放棄することはないと思われる。従って、研究者が欧州の中央機関にデータ利用を申請し、各加盟国が研究利用の許可を与える仕組みも考えられる✓ 各国のHDABをゲートウェイとして機能するだけでなく、ガイドラインを作成しEDPS 及び EDPB と緊密に連携して、匿名化・仮名化などの特定サービス提供者としての役割を果たすことが考えられる

EHDSでは、二次利用に係る費用3.5～7億€に対して、ベネフィットを54億€程度と想定している

EHDSの枠組み内のデータの二次利用によるコストとベネフィット

■ オプションBでは、二次利用 によって10年間で合計54 億ユーロの収益または削減効果があると試算されている (期間: 10年間)

コスト		ベネフィット	
概要	金額	概要	金額
欧州委員会と加盟国による専門家のガバナンスネットワークの構築 ※HDABフレームワークの一致性確保や二次利用の管理のため	1,300万€ (約22億623万円)	データ利用者が健康データへのアクセスや再利用を効率化することによる費用削減効果	34億€ (約5770億1400万円)
加盟国によるHDABの設置と機能強化 ※既存組織のユニットとしてHDABを設置する場合は4~15 FTE(Fulltime Equivalent: フルタイム当量)、HDABを独立法人として新規設置する場合は15~50 FTEと推定	3,900万-1億5,700万€ (約66億1869万円-266億4447万円)	政策立案者と規制当局のための情報の透明性による費用削減効果	8億€ (約1357億6800万円)
欧州委員会と加盟国によるHDAB欧州ネットワークのインフラ配備と運用 ※加盟国によるHDABのインフラ配備、各国のHDABをEU-wideインフラに接続、サポートサービス等の費用を含む	1億7,600万-2億8,700万€ (約298億6896万-486億677万円)	健康データのさらなる再利用により、患者や医療提供者が得られる価値	12億€ (約2036億5200万円)
データ利用者がEHDSの健康データにアクセスと利用できるようにするための支援 ※データの相互運用性を保証するため	9,700万-2億400万€ (約164億6187万-346億2084万円)	計	54億€(約9164億3400万円)
データ保有者が行うデータ品質ラベルのため ※品質認証コスト、更新料、ラベルコスト等を含む	2,500万€-8,100万€ (約42億4275万-137億4651万円)		
計	3億5,000万-7億4,200万€(約593億9850万-1259億2482万円)		

出所： https://health.ec.europa.eu/publications/impact-assessment-european-health-data-space_en

推奨される政策オプションBを導入することで、データ利用者は効率的に健康データにアクセスし再利用できるようになり、34億円の費用を削減できる

EHDSの枠組み内のデータの二次利用によるベネフィット(1/2)

ベネフィット	詳細	試算に使われる要素と仮定			
<div>データ利用者が健康データへのアクセスや再利用を効率化することによる費用削減効果</div> <div>(34億€、約5770億1400万円)</div>	<div><div><div>● 個々の同意に30分</div><div>● データアクセス機関に支払う手数料</div><div>● かつ複数の国にまたがってデータを収集する等と仮定</div></div><div>現在の細分化されたデータアクセス方式の場合、10年間でデータ利用者が関連機関に支払うコストが27億€に達する可能性がある」と指摘</div></div>	データアクセス機関（DAB）のコスト	内部コスト(1リクエスト当たりに係る管理コスト等)	<div>• 2,100~2,600€/1リクエストあたり</div>	
		データ利用者のコスト	データ収集に係る費用	<div>• リクエスト量 * 同意取得時間 ※国ごとに100件のリクエストを想定 ※1件のリクエスト当たりに係る同意取得時間を30分と想定</div>	
			DABへの手数料	<div>• リクエスト基本料金(1,000€) + 処理コスト(115€/h) • ※処理コスト = 処理時間 * 1時間当たりの価格 • ※1,000€/1リクエストあたり、115€/h（平均処理時間は10~20h)</div>	
		データ管理者のコスト		測定不可	
		➤ 試算にあたっての補足情報			
		リクエスト数	<div>• 国ごとに241件から600件</div> <div>• 10年間で年間5%～10%増加する ※フィンランドとデンマーク: 400-500件/年 フランス: 1600件/年の参考情報あり</div>		
		DABの1リクエスト当たりのリクエスト処理の作業量	<div>• 2.5人日 ※本来3人日から導入する場合0.5人日削減</div>		
		データ利用者における1リクエストを準備するための作業量	<div>• 15～20人日</div>		
		※数値と割合の根拠について記載なし			

政策立案者と規制当局のための情報の透明性が高まることに加え、健康データのさらなる再利用が可能になることで、8億と12億ユーロのベネフィットをもたらすことができる

EHDSの枠組み内のデータの二次利用によるベネフィット(2/2)

ベネフィット	詳細	試算に使われる要素と仮定						
<p>政策立案者と規制当局のための情報の透明性による費用削減効果</p> <p>(8億€、約1357億6800万円)</p>	<ul style="list-style-type: none">10年間で約16億€の利益をもたらす。そのうち、8億€がEHDS導入によるRWE(real-world evidence) の活用強化による利益と推定フェーズⅢの医薬品開発コストを10%削減できるがん、糖尿病、循環器、呼吸器、神経系の薬剤費を節約できる等と仮定 <p>※1点目の残り8億のベネフィットは、欧州医薬品庁(EMA)が主導するリアルワールドデータのデータ分析およびネットワーク(DARWIN: Data Analysis and Real-World Interrogation Network)を通じて、規制当局の健康データの利用を支援することによって創出されることを想定</p>	<p>➢ 医薬品開発コストの削減額(1億8400万€) = フェーズⅢ 医薬品コスト(1,840万€)×医薬品数(年間100品)×10%</p> <table><tr><td>各フェーズの医薬品臨床試験コスト</td><td>フェイズⅠ: 290万€ (約4億9215万円) フェーズⅡ: 740万€ (約12億5585万円) フェーズⅢ: 1,840万€(約31億2264万円)</td></tr><tr><td>医薬品数</td><td>年間100品</td></tr></table> <p>➢ 薬剤費の削減額(5000万€) = 薬剤費×5%</p> <table><tr><td>薬剤費削減</td><td>EU中規模の国では5%、年間5000万€(約84億8550万円) ※新薬の価格上昇に伴い、削減額が年々と増加すると推定</td></tr></table>	各フェーズの医薬品臨床試験コスト	フェイズⅠ: 290万€ (約4億9215万円) フェーズⅡ: 740万€ (約12億5585万円) フェーズⅢ: 1,840万€(約31億2264万円)	医薬品数	年間100品	薬剤費削減	EU中規模の国では5%、年間5000万€(約84億8550万円) ※新薬の価格上昇に伴い、削減額が年々と増加すると推定
各フェーズの医薬品臨床試験コスト	フェイズⅠ: 290万€ (約4億9215万円) フェーズⅡ: 740万€ (約12億5585万円) フェーズⅢ: 1,840万€(約31億2264万円)							
医薬品数	年間100品							
薬剤費削減	EU中規模の国では5%、年間5000万€(約84億8550万円) ※新薬の価格上昇に伴い、削減額が年々と増加すると推定							
<p>健康データのさらなる再利用により、患者や医療提供者が得られる価値</p> <p>(12億€、約2036億5200万円)</p>	<p>AIベースのシステムなど、データ集約型技術による健康分野の革新的な製品やサービスの開発を通じて、健康データのさらなる再利用を推進することによって、患者、医療提供者等の関係者により高い価値を提供できる</p>	<p>※推奨オプションBでは価値が1.0%上がると推定される</p> <table><tr><td>全体データの価値</td><td>3,060億€~5,330億€(2020~2028年) ※コンサルによって算出</td></tr><tr><td>EUのGDPに占める健康支出の割合</td><td>8.3% ※コンサルによって算出</td></tr><tr><td>EUのGDPに占める健康分野の研究開発の割合</td><td>0.03% ※WHOにより算出</td></tr></table>	全体データの価値	3,060億€~5,330億€(2020~2028年) ※コンサルによって算出	EUのGDPに占める健康支出の割合	8.3% ※コンサルによって算出	EUのGDPに占める健康分野の研究開発の割合	0.03% ※WHOにより算出
全体データの価値	3,060億€~5,330億€(2020~2028年) ※コンサルによって算出							
EUのGDPに占める健康支出の割合	8.3% ※コンサルによって算出							
EUのGDPに占める健康分野の研究開発の割合	0.03% ※WHOにより算出							

Healthdata@EUパイロットプロジェクト

EHDS規則の実施に向けたコンソーシアム活用（TEHDAS）およびパイロットプロジェクトを実施し、ユースケース実証や二次利用のガイドラインや仕様等の作成を進めている

EHDS規則における取組

	2022 (R4)	2023 (R5)	2024 (R6)	2025 (R7)	2026 (R8)	2027 (R9)	2028～ (R10～)
EHDS規則案	★ 規則案の提出			★ 規則案の合意			
TEHDAS1	● 2021/2～2023/07 ● ヘルスデータの二次利用に関する欧州の原則を策定						
Healthdata@EU パイロットプロジェクト	● 2022/10～2024/9 ● 「ITインフラの開発・展開」、「ユースケース実証」の実施						
TEHDAS2			● 2024/5～2026/12 ● HDAB(Health Data Access Bodies)の設立に向けてガイドラインと技術仕様を作成				

※TEHDAS（Towards the European Health Data Space）：国民の健康、公衆衛生、研究、イノベーションのために、国境を越えたヘルスデータの使用を改善するため活動する26のEU加盟国等から組成されたコンソーシアム

出所： <https://ehds2pilot.eu/project/>、 <https://tehdas.eu/tehdas1/>、 <https://tehdas.eu/>

欧州ではEHDS規則案における二次利用（Healthdata@EU）の推進に向けて2022年からパイロットプロジェクトを実施した

Healthdata@EUパイロットプロジェクト（1/3）－ 概要と実績

概要

【担当機関】

• 欧州委員会

【目的】

Healthdata@EU に係る以下の実証

① ITインフラの開発・展開

メタデータ検索ポータル、データアクセスリクエストポータル、その他の中央サービスを含む、データプロジェクトの二次利用のためのヨーロッパ初のネットワークを構築し、プロジェクト内ではデータカタログをクエリし、データアプリケーションフォームを利用し、複数のデータにアクセス要求できる

② ユースケースの実行および検証

研究、イノベーション、政策立案、規制活動のために、さまざまな参加国間でデータを共有するためのデータガバナンス、データ品質、データインフラストラクチャに関する標準を開発、調査、分析

【参加規模】

• 欧州内17のパートナー機関など

【取組内容】

• 9つのWP（ワークパッケージ）を設け、プロジェクト全体のタスクと技術的タスクに分類し実施

実績

実績	発表日	WP	概要
<u>概念実証 (POC)</u>	2023/4/4	WP5	POCでは、インフラストラクチャのセットアップ、データ検出とデータ許可のワークフローの実装など、ソリューションを構築するための多くのトピックに対処 ➢ <u>データの検出</u> 研究者が関連するデータセットを発見できるようデータカタログの公開 ➢ <u>データ許可</u> 研究者がデータセットにアクセスするための要求に対する検証
参加ノードの法的枠組みの中で、ヘルスデータへのアクセスとデータ使用のための共通の要素の特定	2023/7/25	WP7	プロジェクトに関与する HDAB、研究インフラストラクチャ、EU 機関の共通要素と主な相違点（データ範囲、データアクセス申請、料金、データ提供の形式等） を特定した
欧州連合(EU)域内での電子ヘルスデータの国境を越えた利用を促進するための共通データ申請フォーム	2024/4/15	WP7	ノード間のバリエーションに対応するため、EHDS規則案とGDPRの要件を満たした、 ヘルスデータ二次利用のデータアクセス申請方法とフォーム を作成した
利用可能なメタデータカタログと使用中のメタデータ標準のランドスケープ分析	2024/4/16	WP6	ヘルスデータの二次利用を促進するために、 メタデータカタログ(データセットをリスト化するカタログ) を作成した

※2024年11月時点

9つのワークパッケージ(WP)はプロジェクト全体のタスクと技術的タスクに分かれている

Healthdata@EUパイロットプロジェクト（2/3）－ワークパッケージ一覧

分類	WP	WP名	リーダー	役割
全体のタスク	WP1	プロジェクト調整	Health Data Hub	プロジェクト全体の調整、プロジェクト管理、財務管理、リスクと倫理管理を提供及び監視する
	WP2	普及とコミュニケーション	欧州公衆衛生協会	コンソーシアム外の関係者への普及とコミュニケーション活動を監督する
	WP3	プロジェクトの評価	Health Data Hub	プロジェクトが計画どおりに実施され、目標を達成しているかどうかを確認するために実施されるアクションを監督する
	WP4	EHDS ノードになるための推奨事項	デンマーク保健データ局 (Danish Health Data Authority)	IT、法律、メタデータ、データ標準、経済的原則を網羅している将来のEHDSに関する推奨事項とテスト済みパイロットインフラストラクチャのスケールアップを行う
技術的タスク	WP5	ITインフラストラクチャ	Health Data Hub、Health Data Lab	ITインフラストラクチャの実装でノード間の情報交換を行い、eDelivery(データ交換に使用されるソフト)プロトコルで、ノード間のデータ処理と検出サービス及びホスティングソリューションを提供する
	WP6	メタデータ標準	ベルギー国立公衆衛生研究所(Sciensano)	データコレクションの表示と記述に使用される記述メタデータテンプレートを標準化し、Health DCAT-AP(メタデータレコード仕様アプリケーションプロファイル)拡張機能に基づいた共通の記述メタデータモデルを標準化する
	WP7	規制及び法令遵守	バイオバンキング及びバイオ分子資源研究インフラストラクチャ- 欧州研究インフラストラクチャコンソーシアム(BBMRI-ERIC)	EHDSのデータアクセスフォームを構築し、5つのユースケースにおけるデータの法的・倫理的な使用とセキュリティ要件、GDPRのコンプライアンスに関するガイドラインを提供する
	WP8	データの相互運用性、品質、保護	バイオバンキング及びバイオ分子資源研究インフラストラクチャ- 欧州研究インフラストラクチャコンソーシアム(BBMRI-ERIC)	データ相互運用性の研究、データの共通データモデルや用語の標準化、ノード間のデータ交換から研究者へのデータ配信までのデータの相互運用性と品質管理を扱う
	WP9	ユースケース管理	Health Data Hubとデンマーク保健データ局 (Danish Health Data Authority)	5つのユースケースの実装を調整し、共通データモデルに基づいてHealthDCAT(メタデータレコード仕様)拡張機能とメタデータカタログを埋めるプロジェクト管理を行う

出所 : <https://ehds2pilot.eu/work-packages-2/>

幅広い研究及び政策トピックを網羅する5つのユースケースが選定され、それに基づき必要なデータ標準、法的要件、コスト及び経済モデルに関する推奨事項を作成する

Healthdata@EUパイロットプロジェクト（3/3）－ ユースケース概要

- データを利活用する可能性を示すために、**幅広い研究及び政策トピックを網羅する5つのユースケース**を実行する
- ユースケースに基づき、ネットワークの拡張に必要な**データ標準、法的要件、コスト及び経済モデルに関する欧州委員会への推奨事項を作成する**

No	ユースケース名	目的	リーダー	参加国・機関
1	感染症サーベイランス(AMR: 薬剤耐性)	薬剤耐性に着目した感染症サーベイランスへのEHDS活用の実現可能性を実証	欧州疫病予防管理センター (European Centre for Disease Prevention and Control)	フィンランド、ベルギー、ノルウェー、クロアチア
2	心血管代謝性疾患	欧州諸国における心血管代謝性疾患のケア経路の比較と予測モデルの構築、人工知能の活用	Health Data Hub	不明
3	COVID-19患者の血栓症	COVID-19患者の血栓症リスクの理解を深める	欧州医薬品庁(European Medicines Agency)	デンマーク、フィンランド、クロアチア、フランス、ダーウィンEU(R)ネットワーク(EUにおける医療データからリアルワールドデータを生成するネットワーク)
4	Covid-19の検査、ワクチン接種、入院	COVID-19の検査、ワクチン接種、入院を、一般の人々と健康への悪影響やその他の悪影響のリスクが高い集団のサブグループの間で比べる	ベルギー国立公衆衛生研究所(Sciensano)	ベルギー、デンマーク、クロアチア、フィンランド、ハンガリー、フランス
5	大腸がん	臨床データとゲノムデータを動員して連鎖させ、大腸がんの理解を深める	Elixir(欧州ライフサイエンスインフラストラクチャー)	ベルギー、ハンガリー、デンマーク

WP5～7におけるデータ接続ネットワーク、法的枠組の共通点や相違点、データ申請フォーム、メタデータ等についての5つの実績が公表されており、3つの実績が公表予定になっている

実績概要

■ 公表済み

実績	発表日	WP	概要
概念実証 (POC) 通信	2023/4/4	WP5	フランス、デンマーク、フィンランドのノードと欧州中央サービスを接続するネットワークを提供及び展開し、 <u>欧州データセットカタログの公開、単一のエントリポイント</u> の提供を実現する
参加ノードの法的枠組みの中で、健康データへのアクセスとデータ使用のための共通の要素の特定	2023/7/25	WP7	プロジェクトに関与するHDAB、研究インフラストラクチャ、EU 機関の <u>共通要素と主な相違点 (データ範囲、データアクセス申請、料金、データ提供の形式等)</u> を特定した
欧州連合(EU)域内での電子医療データの国境を越えた利用を促進するための共通データ申請フォーム	2024/4/15	WP7	ノード間のバリエーションに対応するため、EHDS規則案とGDPRの要件を満たした、 <u>ヘルスデータ二次利用のデータアクセス申請方法とフォーム</u> を作成した
利用可能なメタデータカタログと使用中のメタデータ標準のランドスケープ分析	2024/4/16	WP6	ヘルスデータの二次利用を促進するために、 <u>メタデータカタログ(データセットをリスト化するカタログ)</u> を作成した
既存のメタデータテンプレートからHealthDCAT-APへの移行ワーキンググループ議事録	2024/4/16	WP6	HealthDCAT-APの設計を目的とし、 <u>インフラ内のメタデータカタログの相互運用性を確保するための重要なプロパティ</u> について議論した

■ 公表予定

- ✓ HealthDCAT-APの拡張
- ✓ データの相互運用性、クエリ、交換に関する標準の推奨事項
- ✓ 品質管理、保証、出所の推奨事項

パイロットでは計17組織がパートナーとして参画している

パートナー一覧

所属国	組織名	役割	設立年	属性	概要
オランダ	欧州医薬品庁(European Medicines Agency)	パートナー	1995	EU機関	医薬品の科学的評価、監督、安全性監視機関
クロアチア	クロアチア公衆衛生研究所 (Croatian institute of public health)	パートナー	1893	政府機関	クロアチア共和国に関連するすべての健康統計データの収集を担当する機関
スペイン	国際分類協会連合(IACS)	パートナー	2002	研究機関	スペインアラゴン州の生物医学研究管理機関
	スペイン保健省(MISAN)	パートナー	1978	政府機関	政府の医療政策や計画及び市民の健康保護の権利を確保するための国家総局からの支援機関
スウェーデン	欧州疫病予防管理センター (European Centre for Disease Prevention and Control)	パートナー	2005	EU機関	感染症に対する欧州の防御を強化することを目的としたEUの機関
デンマーク	デンマーク保健データ局 (Danish Health Data Authority)	WP4リーダー	2015	公的研究機関	保健省の一部で患者の治療、開発、計画、研究のための包括的で有効な健康データを確保するために、国民健康登録簿を開発および実行する機関
ドイツ	ヘルスデータラボ (Health Data Lab)	WP5リーダー	2019	公的研究機関	ドイツのデータ二次利用のためのデータ分析機関。ドイツ全国法定健康保険基金連合会のデータの品質保証等も行っている
ノルウェー	ノルウェー電子保健局 (The Norwegian Directorate of eHealth)	WPリーダー	2016	政府機関	eヘルスの運営と調整担当機関

出所 :<https://ehds2pilot.eu/partners/>

パイロットでは計17組織がパートナーとして参画している

パートナー一覧

所属国	組織名	役割	設立年	属性	概要
フランス	Health Data Hub	全プロジェクトのコーディネーター、WP1・WP3・WP5・WP9リーダー	2019	公益団体	フランスのヘルスデータ二次利用機関
ベルギー	ベルギー国立公衆衛生研究所 (Sciensano)	WP6リーダー	2018	公的研究機関	ベルギーの連邦公衆衛生大臣と連邦農業大臣の権限の下で運営されている連邦研究機関
ハンガリー	国立病院総局 (ORSZÁGOS KÓRHÁZI FŐIGAZGATÓSÁG – OKFŐ)	パートナー	2020	政府機関	医療提供とサービスの調整、開発、品質保証担当機関
フィンランド	フィンランドの社会・健康データ許可局(Findata)	パートナー	2019	政府機関	社会医療データの二次利用と個人情報保護支援機関
複数加盟国	バイオバンキング及び生体分子資源研究基盤-欧州研究基盤コンソーシアム(BBMRI-ERIC)	WP7、WP8リーダー	2019	EU機関	バイオバンクと生体分子資源の分散型研究基盤
複数加盟国	欧州公衆衛生協会(European Public Health Association)	WP2リーダー	1992	非政府組織	公衆衛生の政策立案強化組織
複数加盟国	オルファネット(Orphanet)	パートナー	1997	国際非営利団体	健康情報システムやRDレジストリで使用される希少疾患の標準的な命名法(ORPHAコード)の作成ネットワーク組織
複数加盟国	EBRAINS	パートナー	2019	国際非営利団体	最先端の脳の神経学関連研究のプラットフォームにアクセス可能なオープンインフラ
複数加盟国	Elixir	パートナー	2013	政府機関	ライフサイエンスデータの一次利用、二次利用支援組織

出所 : <https://ehds2pilot.eu/partners/>

EUでは2024年5月にTEHDAS2プロジェクトを立ち上げ、各国のHDABの設立に向けてガイドラインと技術仕様を作成し、EHDSの実施を支援する

TEHDAS2の概要

名称	Towards the European Health Data Space Joint Action(TEHDAS2)
資金提供	EU4Healthプログラム※
コーディネーター	Sitra(フィンランド独立記念基金)
期間	2024年5月～2026年12月
基礎	✓ TEHDAS1(2021.02～2023.07) ✓ HealthData@EUパイロットプロジェクト
関係者	✓ 非加盟国を含む欧州29カ国の60以上の組織により協力 ✓ 進行中のプロジェクトとイニシアチブと連携
ゴール	✓ 国境を越えたヘルスデータの円滑な利用を可能にする ✓ ヘルスデータの二次利用を活発化させることで、ヘルス分野における欧州の研究とイノベーションの競争力を高める
主な取り組み	✓ 各国の HDAB(Health Data Access Bodies)の設立に向けてガイドラインと技術仕様を作成し、EHDS(European Health Data Space)の実施を確保するための支援を提供する

※EU4Healthプログラムとは、EUにおける最大規模の公衆衛生イニシアティブで、2021-2027年の期間で53億ユーロの予算が割り当てられ、EU加盟国、医療機関、関連組織に対して財源支援を提供している

ワーク パッケージ	<ul style="list-style-type: none">① 管理及び調整(Management and coordination)② コミュニケーション(Communication)③ 評価(Evaluation)④ 協力モデル(Collaboration models)⑤ データ・ディスカバリ(Data discovery)⑥ データへのアクセス(Access to data)⑦ 安全かつセキュアな処理(Safe and secure processing)⑧ 市民サービス(Serving citizen)
--------------	--

2.1.2 国際標準コード

主な国際標準コードおよびコモンデータモデルは以下の通り

国際標準コード

分類	コード(略語)	管理団体	説明	利用状況
病名	ICD-10、ICD-11	WHO	WHOが管理する国際疾病分類	<ul style="list-style-type: none">ICD-11は132カ国で導入中(2024年5月視点)
医療行為	SNOMED CT	SNOMED International	状態、診断、症状、徴候、手順、観察、および一部の臨床検査、薬物、およびデバイスも表すなど、多領域にわたる臨床概念を表す臨床用語集	<ul style="list-style-type: none">80カ国以上で利用(2025年2月HP情報)
医薬品	ATC	WHO	WHOが統括管理する国際的な統計を取るために考案された薬品の分類	<ul style="list-style-type: none">不明 ※ただし、フランスやフィンランド等多くの国で採用
検査	LOINC	Regenstrief Institute, Inc.	臨床検査、バイタルサインなどの臨床測定、観察、文書を明確に識別する一連のユニバーサルコードと構造化された名称	<ul style="list-style-type: none">196カ国
	DICOM	NEMA	医用画像の国際規格	<ul style="list-style-type: none">情報なし(各国で利用)

CDM

分類	名称(略称)	管理団体	説明	利用状況
医療アウトカム 共通データモデル	OMOP-CDM	OHDSI	正式名称Observational Medical Outcomes Partnership Common Data Modelであり、国際的標準データモデルとして医療ビッグデータ分析を推進するために整備されたデータモデル	<ul style="list-style-type: none">49カ国534データベースマッピングされた患者記録9億5,600万件 ※2024/4時点

国際疾病分類であるICD-10、11および医療画像等の国際標準であるDICOMは世界各国で導入が進められている

ICD10、11

名称	International Classification of Diseases
管理団体	WHO(World Health Organization)
概要	<ul style="list-style-type: none">ICD(国際疾病分類)は世界保健機関(WHO)によって、1994年から正式運用されたICD-10、その後継としてICD-11は2022年から正式に運用が開始された最新の国際疾病分類システムとし、異なる国や地域、異なる時期に収集された死亡率や罹患率データの体系的な記録、分析、解釈、比較を可能としている
利用状況	<ul style="list-style-type: none">ICD-11は132カ国で導入中(2024年5月視点) →72カ国が翻訳作業を含む導入中 →50カ国が導入パイロットを実施/拡大

DICOM

名称	IDigital Imaging and Communications in Medicine
管理団体	NEMA(National Electrical Manufacturers Association：米国の医療機器等の業界団体)
概要	<ul style="list-style-type: none">DICOMは医療用画像および関連情報の国際標準であり、臨床使用に必要なデータと品質で交換できる医療用画像の形式を定義しているほぼすべての放射線科、心臓病学の画像診断、放射線治療機器(X線、CT、MRI、超音波など)に実装されている
利用状況	<ul style="list-style-type: none">情報なし(ただし、各国で利用されている)

SNOMED CTは世界でもっとも包括的で多言語の医療用語集であり、80カ国以上で導入が進んでいる

SNOMED CT

名称	SNOMED CT
管理団体	SNOMED International ※英国の非営利団体、SNOMED CTを所有し、管理、開発を実施
概要	<ul style="list-style-type: none">• SNOMED CTは世界で最も包括的で多言語の臨床医療用語集であり、包括的かつ科学的に検証された臨床コンテンツを備えたリソース• 電子ヘルスデータにおける臨床コンテンツの一貫した表現を可能にし、他の国際標準にマッピングされている• SNOMED CTは単なる診断のコード化システムではなく、症状や兆候、臨床所見なども幅広くカバーしている• SNOMED CTを導入することで、ヘルスデータが共通化され、人日の新たな健康問題の早期発見、集団の健康状態の監視、変化する臨床実践への対応を促進できる
利用状況	<ul style="list-style-type: none">• 80カ国以上で利用 (2025年2月ホームページ掲載情報)

SNOMED CTの臨床範囲

- ✓ SNOMED CTには、症状や臨床所見、さらに何万もの手術、治療、診断手技が含まれている
- ✓ 観察可能な情報(例えば心拍数)が含まれるほか、身体構造、生物、物質、医薬品、物理的物体、物理的力、検体、その他健康記録の中やその周辺に記録される必要のある多くの種類の情報を表す概念も含まれる
- ✓ 包括的なスコープにより、同じ医療記録システムで複数の互換性のないコード体系や重複するコード体系をサポートする必要性が低くなり、これはSNOMED CTが、地域および国際的なレベルで、医療機関間で一貫したコミュニケーション、検索、および処理を行うための共通用語となり得ることを意味している

出所 : <https://www.snomed.org/what-is-snomed-ct>

ATCは、WHOが統括管理する国際的な統計を取るために考案された薬品の分類である

ATC

- ✓ ATC分類システムは以下の14の主要な解剖学的または薬理学的グループに分類されている
- ✓ 医薬品は、各医薬品に1つの ATC コードのみを使用するという基本原則に基づいて、主な治療用途に応じて分類される (投与経路と場合によっては強度によって定義される)

名称	Anatomical Therapeutic Chemical Classification
管理団体	WHO(World Health Organization)
概要	<ul style="list-style-type: none">• WHOが統括管理する国際的な統計を取るために考案された薬品の分類• 主に薬効、作用部位、化学的な特徴に基づいて5段階までのレベルで薬品进行分类する• 主な用途として、医薬品使用研究や副作用モニタリングなどの疫学研究における利用がWHOにより推奨されている
利用状況	<ul style="list-style-type: none">• 不明 ※ただし、フランスやフィンランドなど多くの国で採用されている

LOINCは医療における測定値、観察結果、文書を識別するための共通言語であり、196か国にわたる約24万の登録者により利用されている

LOINC	
名称	LOINC (Logical Observation Identifiers Names and Codes)
管理団体	Regenstrief Institute, Inc. ※医療の質向上、患者の健康増進、医療コストの削減を目的とした研究と技術開発に取り組んでいる米国の非営利の医療研究機関
概要	<ul style="list-style-type: none">LOINCは、<u>医療における測定値、観察結果、文書を識別するための共通言語</u>（識別子、名称、コードのセット）である臨床結果の交換や集約を可能にし、医療提供、成果管理、研究に活用できるよう、普遍的なコードと構造化された名称を提供することで、測定や観察が可能なものを明確に識別できるようにしている<u>臨床検査、バイタルサインや身体計測値</u>などの臨床的な測定、標準化された調査票など、多岐にわたる測定項目の豊富なカタログを提供する。パネル、フォーム、文書などの項目をまとめたコードも含まれている
利用状況	<ul style="list-style-type: none"><u>196か国</u>にわたる約24万の登録者より利用されているLOINCのコンテンツおよびドキュメントは、各国の要件に対応するためにフランス語、中国語、ドイツ語、イタリア語、スペイン語等20種の言語に翻訳されている
資金提供	
<ul style="list-style-type: none">National Library of Medicine(米国国立医学図書館)Assistant Secretary for Technology Policy / Office of the National Coordinator for Health IT(技術政策担当次官補室 / 国家医療IT調整室)Centers for Medicare & Medicaid Services(CMS)Regenstrief FoundationJohns Hopkins UniversityLOINCコミュニティの寄付者	

出所: <https://loinc.org/atlas/>;

2.1.3 OMOP

OMOP-CDMは2008 年からOHDSIにより推進された国際的標準データモデルで、現在約9億5,600万件の患者記録がマッピングされている

OMOP-CDM概要

正式名称	Observational Medical Outcomes Partnership Common Data Model 観察的医療アウトカムパートナーシップ共通データモデル
利用開始年	2008年
推進団体	OHDSI(共通データ形式を使った医療ビッグデータ分析を推進するオープンサイエンスコミュニティ)
目的	国際的標準データモデルとして医療ビッグデータを統一し、今後の分析に準備する
導入実績 (2024.4まで)	49か国534データベース マッピングされた患者記録：9億5,600万件

OMOP-CDM特徴

① 国際的に広く導入されている

OMOP-CDMは、49か国の534のDBに広く採用されている。さらに、OHDSIコミュニティによる積極的なサポートとグローバルなネットワークを通じて知識とリソースが共有され、国際的なコラボレーションが促進されている

② 解析を支援する分析ツールが多く提供されている

OHDSIにより提供されているツールを通して、データの品質管理、データマッピング及びデータ分析までの各段階での作業を効率化し、迅速かつ正確な結果が得られる

③ リアル・ワールド・データ（RWD）分析に特化している

OMOPはリアル・ワールド・データ（RWD）分析のために設計されたデータ・モデルであり、電子医療記録、管理請求、レジストリ、病院システム、ゲノミクス、バイオバンク等、様々な患者レベルの幅広いデータの標準化に対応している

EUにおいては、EHDSに関連する欧州医薬品の調査等のためのネットワークDARWIN EU® 及び欧州健康データ・エビデンスネットワークがOMOPを導入している

EHDSにおけるOMOPの導入

■ DARWIN EU®

- DARWIN EU®（Data Analysis and Real World Interrogation Network in the European Union：EUデータ分析及びリアルワールド調査ネットワーク）は医薬品の安全性及び有効性の調査等を目的に欧州医薬品庁(EMA)と欧州医薬品規制ネットワークによって設立されたネットワークである
- DARWIN EU®は欧州医薬品規制ネットワークをEHDSと繋げており、EHDSの開発や、健康データの二次利用に関する欧州の原則 (TEHDAS) の実現にも貢献している
- DARWIN EU® は、その業務の一環として、OMOP 共通データモデルと OHDSI ツールを使用している

■ EHDEN

- EHDEN（European Health Data & Evidence Network：欧州健康データ・エビデンスネットワーク）はOMOPに標準化された大規模なDBの連合ネットワークを構築することで、ヨーロッパにおける健康データの発見と分析を支援するネットワークである
- EHDENは2022年10月20日にEHDSに関する欧州委員会の提案を支持するグループに加入した
- EHDENアカデミーでは無料のOMOPの教育リソースを公開している

出所： <https://www.ohdsi.org/wp-content/uploads/2023/11/OHDSI-Book2023.pdf>, <https://darwin-eu.org/index.php/about/ehds>, <https://www.ehden.eu/ehden-joins-stakeholders-consensus-response-to-the-proposed-european-health-data-space/>

OHDSIは元データの品質管理からデータマッピング、データ分析のツール群を提供している

OHDSIにより提供されているツール

OHDSIでは、医療データに熟知していない医療関係者であっても一定の分析ができるように充実した分析ツール群を提供している

分類	ツール名	概要	主な機能
品質管理	ACHILLES	データ品質のチェックを可視化するR言語パッケージ	データ品質の自動チェック、概要レポート生成
	Data Quality Dashboard	データ品質の自動チェックとレポート生成	データ品質の自動チェック、品質問題の特定、詳細なレポート生成
データマッピング	ATENA	標準化されたボキャブラリーの検索とダウンロード	用語の検索、用語のダウンロード
	WhiteRabbit	データベースのETLを準備する	データベースのスキャン、データ詳細情報のレポート生成
	Rabbit-In-a-Hat	スキャンされた情報の読み取りと表示	テーブルとフィールド間のマッピング、ETLプロセスのドキュメント生成
	Usagi	コードマッピングを手動で作成するプロセスを支援するインターフェイスを提供する	用語マッピングの提案、マッピングの手動調整、マッピング結果のエクスポート
データ分析	ATLAS	Webベースのインターフェイスで分析の設計と実行の簡易化を図る	コホート作成、大規模データ分析、語彙検索、患者集団に関する特性の評価、患者レベルの予測等
	HADES	大規模分析のためのR言語ライブラリ	特性評価、集団レベルのインが効果推定、患者レベルの予測等

既存の医療に関するCDMはI2b2、OMOP、Sentinel、PCORnetである。そのうちOMOPの利用国及び論文数が最も多い

CDM比較

CDM 名称	利用開 始年	主な 利用国	概要	参考値(論文数※)		国際共同研究例
				Google Scholar	Pub Med	
I2b2	2004年	米国、 ヨーロッパ	ゲノミクス時代に医療記録と臨床研究データを統合するために、必要なツールを研究者に提供するために開発された柔軟に適応できるデータモデルである	574本	495本	米国、ヨーロッパ、南米、シンガポールの生物学と臨床現場を統合する情報科学による大規模な COVID-19 研究ネットワークのサポート
OMOP	2008年	49か国	リアルワールドデータ（RWD）を標準化し、医療ビッグデータ分析を目的としたデータモデルである	8,220本	371本	<ul style="list-style-type: none">米国、韓国、日本、ドイツにおいて、OMOPを用いて490万人の高血圧患者における複数の降圧薬の臨床転帰に対する効果と安全性を包括的に比較東アジア諸国と米国における新しい抗糖尿病薬の普及:国際比較研究米国、台湾、香港、日本、韓国における抗精神病薬使用の発生率、有病率および処方パターン
Sentinel	2008年	米国	米国食品医薬品局（FDA）が推進する医薬品、ワクチン、医療機器の安全性の評価を迅速に実行するために開発されたデータモデルである	589本	81本	台湾における医療製品安全性監視システムの構築:米国 Sentinelシステム共通データモデル構造の台湾国民健康保険研究データベースへの適応
PCORnet	2014年	米国	米国の患者中心の比較臨床効果研究を効率的に実施するために、臨床情報等を標準化する共通データモデルである	751本	191本	－

※2024年9月17日キーワード検索により集計

研究のユースケースは、主に人口レベルの推定、特性評価、患者レベルの予測に分けられる

OHDSIにおけるリアルワールドデータを用いた研究

ユースケース分類	研究分類	研究例	概要
人口レベルの推定	臨床応用	フェブキシスタット(FebuxostatCVRisk)の心血管イベントと死亡のリスク	痛風患者におけるフェブキシスタットの有害な心血管イベントと死亡リスクをアロプリノールとの比較
		COVID-19治療のための既存薬の評価	既存のレトロスペクティブデータ(後ろ向き研究)を使用して、COVID-19治療の薬剤の安全性を評価し、潜在的な有効性の推定
	メソッドリサーチ (手法の有効性の評価)	交絡因子除去ツールの実証的評価	観察医療データから治療効果を偏りなく推定する機械学習アルゴリズムである医療交絡因子除去ツールの実証的評価
		ワクチンの有効性のための方法の使用の評価	これまでの因果効果推定法の評価に関する研究と同様に、観察データを用いてワクチンの有効性を推定する方法を体系的評価
特性評価	臨床応用	がん患者におけるCOVID-19のベースライン特性	がん患者におけるCOVID-19の特徴付け
		非介入的管理を受けている前立腺癌患者の長期転帰	さまざまな疾患リスクプロファイルで非治療目的の治療で管理された前立腺がん患者の長期的な転帰の評価
	メソッドリサーチ	フェデレーテッドネットワーク研究モデルを使用したメンタルヘルスクア提供での健康公平性の評価	請求データとEHRデータを既存の調査データとともに利用したフェデレーテッドネットワーク調査研究
		COVID-19ワクチンの有害事象、発生率の補正	4つの方法を使用して、分散データベースネットワーク全体で表現型エラーに対してCOVID-19ワクチンの有害事象の背景発生率の修正と評価
患者レベルの予測	臨床応用	COVID-19による死亡の予測	COVID-19による入院後30日以内の死亡を予測するモデルの検証
		肺がん予後	患者ががんに罹患していない来院後3年以内の新たな肺がんのリスクを予測するために開発された患者レベルの予測モデル
	メソッドリサーチ	パンデミック予測	COVID-19予測モデルのパフォーマンスと、ワクチン接種などのイベントによってパフォーマンスが時間の経過とともにどのように影響を受けるかの調査
		ディープラーニングの比較	低リスク集団における肺がんの3年リスクの予測、認知症の10年リスクの予測、うつ病の1年リスクの双極性障害の予測に対する異なるディープラーニングモデルの比較

2.1.4 オプトアウト

新設された第35条fでは、全ての加盟国に二次利用に対するオプトアウトメカニズムの構築を義務付けているが、各加盟国は国内法で正当な例外を設けることが可能である

EHDS第35条f – 二次利用に対するオプトアウトの権利 (新設)

分類	条文番号	概要
オプトアウトの権利	第35条f (1)	<ul style="list-style-type: none"> ■ 自然人は、理由を述べることなくいつでも自己に関する医療情報の二次利用に対しオプトアウトする権利を有する ■ オプトアウトの行使に関する取消しも可能とする
加盟国の義務	第35条f (2),(2a),(7)	<ul style="list-style-type: none"> ■ 加盟国は、アクセスしやすく理解しやすいオプトアウトメカニズムを提供し、自然人が医療情報の二次利用に関する意思表示できるようにする必要がある ■ 個人がオプトアウトした場合、その個人に関連するオプトアウト後に収集した識別可能な新規データは二次利用に使用してはならない。なお、オプトアウト前のデータは影響を受けない ■ データ主体の識別が不要な場合、オプトアウトの対応のために識別情報を収集する必要はない
例外条件	第35条f (3)-(6)	<ul style="list-style-type: none"> ■ 加盟国は、オプトアウトに対する正当な例外を国内法で設けることができる。その際、以下の全ての条件が満たされる必要がある <ul style="list-style-type: none"> ・ 公衆衛生分野に係る業務を遂行する公共機関、EU機関、又は団体等による申請であり、利用目的が第34条第1項(a)-(c)に該当するもの、又は公共の利益につながる科学研究であること ・ データが他の手段でタイムリーに入手にできない場合 ・ データの申請者が第46条(1)(h)や第47条(2)(be) に基づく正当性を提示した場合 ■ 例外は基本的権利と自由の本質を尊重し、公共利益を達成するものでなければならない ■ 例外は再識別の禁止等の要件を遵守し、国内法には個人の権利保護に関する規定を含む必要がある ■ 加盟国は採用した国内法を委員会に通知し、改正がある場合も同様に通知する必要がある

(参考資料)

EHDS第35条f – 二次利用に対するオプトアウトの権利 関連条文

条文番号	概要
第34条(1)(a)-(c)	<ul style="list-style-type: none">■ データ申請者が以下の目的の遂行のためにデータを必要とする場合に限り健康データの二次利用を許可する<ul style="list-style-type: none">• 公衆衛生や労働衛生に係る公共利益（例: 国際的な健康脅威に対する保護・管理措置、医療機器や医薬品の安全性確保等）• 公的機関やEU機関等による政策立案や特定の規則・基準の策定等• 国際、EU、又は国家レベルの統計
第46条(1)(h)	<ul style="list-style-type: none">■ データ申請者が第35条(f)に基づくオプトアウトの例外申請を希望する場合は採択された国内法による根拠を必要とする
第47条(2)(be)	<ul style="list-style-type: none">■ データ申請者は申請時以下の情報を提供する必要がある：<ul style="list-style-type: none">• データの処理が GDPR 第6条(1) (Lawfulness of processing) または EU規則 2018/1725第5条(1) (Lawfulness of processing) 及び第10条(2) (Processing of special categories of personal data) にどのように準拠するかに関する説明

英国のNational data opt-outが適用されないケースとして、「明示的な同意」、「公衆衛生上のリスク」、「優先する公共の利益」、「法律による要求」を挙げている

National data opt-out operational policy guidance

National data opt-out operational policy guidance

【全53ページ】

1. 目次
 - －目的
2. National data opt-outについて
 - －法的枠組と根拠
 - －他のオプトアウトとの相互作用
 - －National data opt-outのポリシー
 - －死亡患者の取り扱い
3. オプトアウトの設定
 - －対象
 - －設定するチャネル
4. 適用される組織
 - －医療および社会保障制度
 - －医療および介護(民間)
 - －データ管理者および処理者
5. National data opt-outはいつ適用されるか
6. National data opt-outが適用されないケース
7. 特定の組織または目的に対する考慮事項
 - －国境を越えたデータフロー 等
8. National data opt-outの適用
 - －適用のタイミング
 - －データの使用 等
9. National data opt-outの分析
 - －National data opt-outの公開 等
10. National data opt-outの順守
 - －データセキュリティ
 - －National data opt-outへの準拠 等

法的枠組と根拠

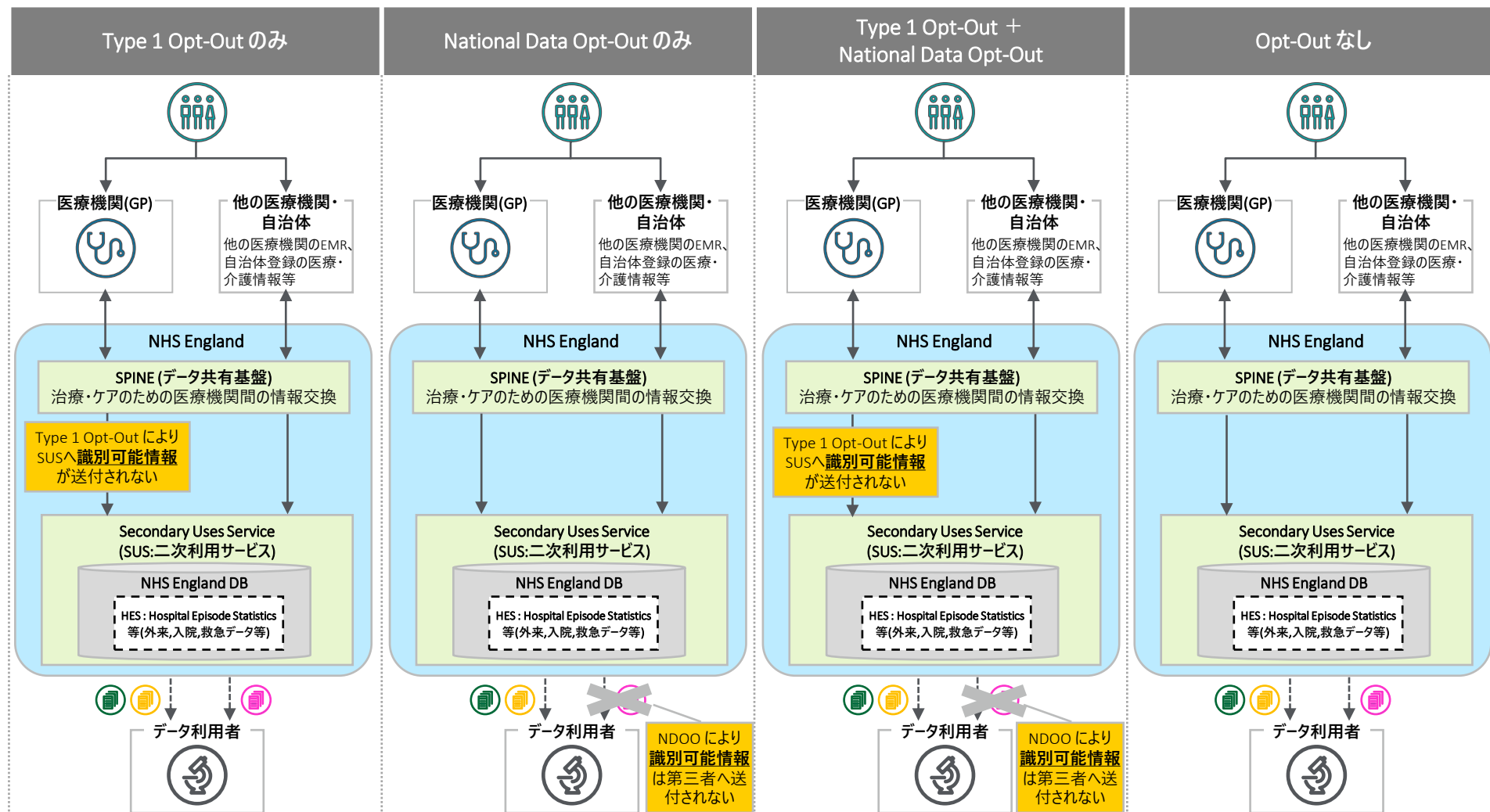
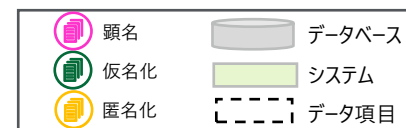
- ✓ 法的根拠
データ保護法 (DPA 2018 および GDPR) では、データ管理者は、個人データのすべての処理が公正、合法、透明性などの原則に沿って次のことを保証する必要がある
 - ・GDPR6条の要件
 - ・GDPR9条(特別なデータカテゴリの取扱い)の要件
 - ・Common Law の要件(例えば、NHS法第251条のデータの使用に関する同意および法定承認について等)

National data opt-outが適用されないケース

- ✓ 同意
特定の目的のために患者から明示的な同意が得られた場合には適用されない
- ✓ 感染症と公衆衛生へのリスク
感染症や公衆衛生に対するリスクの監視と制御に必要な患者の機密情報の開示には適用されない
以下の目的で患者情報を合法的に処理することができる
 - ・感染症や公衆衛生に対するその他のリスクの診断
 - ・そのような病気とリスクの傾向の認識、蔓延の制御・防止
 - ・予防接種プログラムの実施、有効性、安全性
 - ・ワクチンや薬品の副作用
- ✓ 優先する公共の利益
データを開示することの公益性が、秘密を保持することの公益性に勝る場合
 - ・General Medical Council(GMC)のガイドラインに基づいた銃創や刃物による負傷の報告
- ✓ 法律または裁判所命令により要求される情報

英国のオプトアウトが適用されるのは本人識別可能な情報のみであり、ゆえに匿名化情報及び仮名化情報は対象にならない

NHS Englandにおけるオプトアウトの仕組み



出所: <https://digital.nhs.uk/services/national-data-opt-out/understanding-the-national-data-opt-out>, <https://digital.nhs.uk/services/national-data-opt-out/operational-policy-guidance-document>, <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/about-the-gpdr-programme>

2.1.5 医療情報等の海外移転について

米国、EUを中心として5つの海外ゲノムバンクに対して調査を行った。一般市民を対象にしたバンクや特定の疾患を対象にしたバンクなど設置背景は各ゲノムバンクによって異なる

2024年7月時点

各国のゲノムバンク概要

		米国	英国		フィンランド	フランス
主要ゲノムバンク		All of Us	UK Biobank	Genomics England	THL Biobank	CNRGH
設置主体		NIH (国立衛生研究所)	DHSC他 (保健・社会福祉省)	DHSC (保健・社会福祉省)	THL (国立保健福祉研究所)	CEA (フランス原子力庁)
設立年		2015	2006	2013	2014	2008
対象者		一般市民	40-69歳の一般市民	希少疾患患者 がん患者	一般市民	自己免疫疾患、がん、感染症、希少疾患等の罹患者
収集物	サンプル	・血液、尿、唾液	・血液、尿、唾液	・血液、唾液、組織等	・血液、尿、唾液、DNA、組織等	・血液、細胞、体液、組織等
	データ	・EHRデータ ・身体データ ・生活習慣アンケート ・フィットビットデータ等	・身体データ ・臓器等画像データ ・生活習慣アンケート等	・EHRデータ ・NHS England に格納されたデータ (HES、画像、死亡登録データ等) ・生活習慣アンケート等	・検査データ ・身体データ ・生活習慣アンケート等	非公開
サンプル / データ数		824,000 (参加者) 582,000 (サンプル)	500,000 (参加者) ※2006年-2010年の間にリクルート	118,448以上 (ゲノムデータ)	230,000以上 (サンプル、データ)	12,000程度 (ゲノムデータ)
利用制度		・利用機関・利用者登録が必要 ・行動規範・倫理研修の受講必須 ※国内外問わず (2023年-) ※患者向け説明文では民間企業の利用可能性について言及	・利用者および利用機関登録が必要 ※国内外問わず ※民間企業の制限なし	・利用者および利用機関登録が必要 ※国内外問わず ※民間企業の制限なし	・利用者および利用機関登録が必要 ※現段階新規申請は受付けていない ※国内外問わず ※民間企業の制限なし	・CNRGH内部、もしくは共同研究者に限られる
データ保管形態		研究用データは以下の処理を行う： ・識別可能情報を削除し、仮名化 ・EHRデータをOMOP-CDMに変換し、クレンジング	・識別可能情報は別で保存し厳重にアクセス管理を施している ・データはiCD10に変換	・識別可能情報を除き、仮名化 ・識別可能情報は別のストレージで保管し厳重に管理	・識別可能情報を除き、研究IDを付与した上で保管。 ・識別可能情報は別ストレージで保管	非公開
データ提供形態		データの内容や加工程度で以下の3種類に分類しアクセスを管理 ・Public tier dataset (一般公開) ・Registered tier dataset (研究者用) ・Controlled tier dataset (研究者用)	・研究者に提供するサンプル・データには識別可能情報が含まれない ・データの内容や加工程度でTier分けを行いアクセス管理	・Rare Disease Tiering process を採用しデータアクセスを管理	・識別子が取り除かれた状態で研究者に提供 ・データのTier分けに関する情報はないが、プロジェクトベースでDBが構成されている	非公開
第三者提供	サンプル	提供の可能性について言及 ※Materials transfer agreement 必須 ※海外移転に関する記載なし ※個人情報には含まない	承認を得た研究者は利用可能 ※Sample release assay を満たしていることが条件 ※海外移転に関する記載なし	承認を得た研究者は利用可能。ただし、別途同意がない限り、英国外へ移転されることはない	HP上で海外への移転の可能性について言及	機密保持契約を結んだ共同研究者との共有に言及 ※サンプルやデータ種別の明記なし
	データ	国内外問わずクラウドポータルで利用可能 (ダウンロードに関する記載なし) ※統計データは一般公開	国内外問わず承認を得た研究者は Research Analysis Platform 上で利用 (アクセス) 可	国内外問わずリモートデスクトップ環境下で利用可※情報のコピーは不可 ※匿名 (anonymized) の分析結果はエアロック機能を介しダウンロード可	仮名化された状態でデータを転送 ※同意者が選択可 ※データ種別の明記なし	－

消費者向けのDNA検査を提供する23andMe社と、ゲノムデータを解析してAIにより個別化医療を提案するTempus社は、ゲノム関連データを収集し、データベースを構築している

2024年7月時点

ゲノムデータを収集している民間企業

		米国	米国
企業名		23andMe, Inc.	Tempus
設立年		2006年	2015年
概要		<ul style="list-style-type: none">個人ゲノム解析およびバイオテクノロジーを行う企業である<ul style="list-style-type: none">遺伝子検査キットによって消費者の祖先情報、遺伝的健康リスクおよび薬理遺伝学的特性等に関するレポートを提供し、個別化医療やライフスタイルの提案を行う蓄積された顧客データは、23andMeにおいて遺伝性疾患の研究者により調査分析されることができる	<ul style="list-style-type: none">AIとバイオテクノロジーを活用して精密医療を提供する企業である<ul style="list-style-type: none">がん等の疾患に対する治療法の最適化を支援するために、遺伝子データや臨床データを解析し、AIにより個別化医療の提案を行う研究者はTempus+プログラムを通じてデータ分析プラットフォームLens及びデータライブラリーを利用し、非識別化された患者記録にアクセスすることができる
データ収集	収集物	唾液サンプル	血液サンプル、唾液サンプル、腫瘍サンプル等
	データ収集方法	米国にあるラボLabcorpによって唾液サンプルを解析	2,000以上の医療機関との連携、臨床試験の実施等によりデータを収集
	データ数	<ul style="list-style-type: none">研究参加者300万人表現型データ10億データポイント 等	<ul style="list-style-type: none">50ペタバイトのマルチモーダルデータ700万件の非識別化された研究記録100万件の画像データ100万件の臨床データとゲノム情報がリンクされた記録 等
データ利用	データ種類	<u>非識別化された(De-Identified)遺伝情報、表現型データ※1、バイオーム等のデータ</u>	<u>非識別化された(De-Identified)研究記録、画像データ、全転写産物プロファイル※2を含む記録、分子データ、表現型データ等</u>
	利用申請	非営利団体では <u>業務記述書(Statement of Work/SOW)とデータ移転契約(Data Transfer Agreement/DTA)等</u> を提出して申請する	Tempus+コミュニティに加入する必要がある
海外移転	同意書	グローバルに事業を展開しているため、 <u>顧客の個人情報</u> は <u>居住国以外の国に転送・保管・処理される可能性がある</u> と言及	検査の実施、医療提供者への連絡、事業運営管理等の目的遂行には、米国にあるTempusの施設を含む、 <u>顧客の個人データの海外へ転送および保管する可能性がある</u> と言及
	実績	230件の研究に利用 →ホームページ上では全例の公表はないが、研究パートナーの例として、ノッティンガム大学(英)、クイーンズランド大学(オーストラリア)、オーフス大学(デンマーク)等の学術機関がある	全例の公表はないが、Tempus+にてデータを利用している研究パートナーの例として、GlaxoSmithKline (英)、AstraZeneca(英)等の製薬企業がある

※1 表現型データ：生物の複合的で観察可能な特徴や形質を表す情報
※2 全転写産物プロファイル：特定の生物試料から得られた包括的なRNAシーケンシングデータを含むデータエントリー

ゲノムデータの海外移転に関する法規制

項目		米国		英国(EU GDPRも同様)
法令・法案		バイオセキュア法案 (draft of the BIOSECURE Act)		HIPAA
設立経緯		2023年12月と2024年1月に米連邦議会上院・下院に提出された 2024年5月に下院委員会で修正案が承認された 2024年7月時点では上院での審議待ち中		1996年に制定
目的		中国、北朝鮮、ロシア、イラン・イスラム共和国等の外国企業が米国のバイオテクノロジー分野において国家安全保障に対する脅威をもたらすことを防ぎ、 <u>特定の外国企業との契約を禁止することで米国国民の遺伝情報の不正使用を防止する</u>		医療関係事業者に遵守が義務付ける「Privacy Rule」及び「Security Rule」、当該医療情報に関する安全管理措置やその基準も定めている
ゲノムデータ 海外移転に 関する規制	対象となる データの定義	マルチオミクスデータ(Multiomic data)が規制の対象になる マルチオミクスデータとは、ゲノム(Genomic)データ、エピゲノム(Epigenomic)データ、トランスクリプトーム(Transcriptomic)、プロテオーム(Proteomic)、メタボロミクス(Metabolomic)のデータが含まれているデータタイプのことを指している		個人データの処理、英国国外への移転を行うための要件を定める 処理または移転を行う者が遵守すべき義務を定める
	規制の内容	懸念されるバイオテクノロジー企業(Biotechnology Companies of Concern)が製造また提供するバイオテクノロジー機器・サービスを利用し、 <u>契約を締結・延長・更新することが禁止される</u> 以下に基づく、米国の国家安全保障に危険をもたらす企業では懸念されるバイオテクノロジー企業と定義される ✓ バイオテクノロジー機器またはサービスを通じて得られたマルチオミクスデータを敵対国政府に提供する ✓ または明示的かつ十分な説明を受けた上での同意なしに、バイオテクノロジー機器またはサービスを通じて人間のマルチオミクスデータを入手する		遺伝データの海外移転については、一般的な個人データと同様の規制が設けられている。 ※なお、UK GDPRにおける遺伝データの定義は以下の通り 遺伝データ(genetic data) → 自然人の遺伝または獲得した遺伝的特性に関する個人データで、その自然人の生理機能または健康状態に関する固有の情報を提供し、特に当該自然人の生物学的サンプルの分析から生じるもの
	適用除外	商業的または公開されている人間のマルチオミクスデータの取得、使用、配布する場合では、禁止事項が適用されない		個人データの海外移転について、以下の制限がある → 十分性認定を受けた国への移転： → 自由に移転可能(EEA加盟国※1、日本、カナダ 等) ・ 十分性認定を受けていない国への移転： → 原則として移転禁止 → ただし、以下の例外に該当する場合は可能 ・ 適切な保護措置が講じられている場合 (データ輸出者と輸入者がデータ保護に関する一定の要件を満たした契約を締結している場合等 が該当) ・ 法令上の例外に該当する場合 ・ (本人同意がある場合、公共の利益に関わる必要性がある場合等 が該当) ※1： EEA(European Economic Area)： EUに加盟することなくEU市場に参加することができる枠組み。EU加盟国27国に加え、3か国が参加している
				—

出所： https://www.jetro.go.jp/ext_images/_Reports/01/b0226c404f93f434/20220001rev1.pdf、<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/special-category-data/what-is-special-category-data/#scd3>、<https://www.congress.gov/bill/118th-congress/house-bill/7085/text>、<https://statics.teams.cdn.office.net/evergreen-assets/safelinks/1/atp-safelinks.html>

3. 医療等情報を収集するデータベース (レジストリ) 調査について

3.1 国内におけるビジティング解析環境の比較

国内におけるビジティング解析環境の比較(1/2)

項目		MID-NET	次世代医療基盤法		医療・介護データ等解析基盤(HIC)	
			(参考)Ⅰ型認定	Ⅱ型認定		
解析環境の名称		MID-NET接続環境	-	-	医療・介護データ等解析基盤	
解析環境の運営主体		独立行政法人医薬品医療機器総合機構(PMDA)	-	認定仮名加工医療情報作成事業者(以下、「認定作成事業者」)	厚生労働省	
解析環境の概要		利活用者のオフィスから専用端末を用いてMID-NETシステムにおける利活用者作業領域にリモートアクセスし同領域内で解析等を実施	同法による認定(Ⅰ型認定)を受けた利用者が利用者自ら整備した環境で仮名加工医療情報を保存・利用	同法による認定(Ⅱ型認定)を受けた利用者が認定作成事業者が管理するビジティング環境で仮名加工医療情報を利用	利用者がクラウド上の解析基盤を利用してNDB等のデータ解析等を実施	
利用条件等	利用に係るガイドライン		医療分野の研究開発に資するための匿名加工医療情報及び仮名加工医療情報に関する法律についてのガイドライン		医療・介護データ等解析基盤(HIC)の利用に関するガイドライン	
	審査		PMDAが有識者会議に意見を聴取したうえで、審査基準に則り審査		認定事業者が設置する審査委員会で審査	各DBの専門委員会で審査
	使用端末	貸与端末	○	-	○	-
		利用者端末	-	○	(認定事業者から利用許可された端末)	
	解析作業場所		利活用者のオフィスで部外者が入室できない施錠管理した領域（建物、フロア、部屋等）	施設設備への立ち入り・機器の持ち込みを制限した管理区域(施錠・監視カメラ常時監視 等)	-	施錠し、立ち入れるものを制限した区画(国内に限られ、事前に申し出が必要)
	ログイン認証		二要素認証(顔認証・ユーザーID/パスワード(PMDAが発行))	二要素認証	認定作成事業者より発行された認証情報による認証	二要素認証
	パスワードルール		8桁以上かつ英字(大)・英字(小)・数字・記号から3種以上混在(定期変更なし)	-	-	13桁以上の英数字、記号(定期変更なし)
	ソフトウェア等の追加導入		-	(認定作成事業者との取り決めによる)		追加導入不可
	利用できるソフトウェア等		SAS製品(あらかじめ作成したSASプログラムを利用するためには、ファイル持込申請が必要)			R/Rstudio、SPSS、Stata、Python
	セキュリティ対策(例)		<ul style="list-style-type: none">覗き見防止対策:端末Webカメラによる顔認証盗撮(画面撮影・録画)防止対策:端末のWebカメラによる盗撮防止対策不正アクセス対策:外部記録媒体利用時のウイルスチェック	<ul style="list-style-type: none">覗き見防止対策:のぞき見対策のシートの貼付盗撮(画面撮影・録画)防止対策：画面撮影の禁止不正アクセス対策:ウイルス対策ソフトパターンファイル更新	(認定作成事業者との事前の取り決めに基づき定められたポリシーに則り、適切に端末を利用しなければならない)	<ul style="list-style-type: none">覗き見防止対策:間仕切りの設置、スクリーンセーバーの設定盗撮(画面撮影・録画)防止対策:利用画面のスクリーンショット取得等の禁止不正アクセス対策:ウイルス対策ソフトのパターンファイル更新、公衆LAN接続の禁止、外部記録媒体利用時のウイルスチェック
	教育・研修		MID-NETを利活用する者は、安全管理措置、システム操作等に関するPMDAの研修受講が必要	利用者は、データの取扱者に対して、安全管理措置等に関する研修実施が必要		利用者はデータの取扱者、医療・介護データ等を取り扱う上で必要な教育・訓練の実施が必要

出所：独立行政法人医薬品医療機器総合機構「MID-NETの利活用を検討又は利活用中の方向け情報」:<https://www.pmda.go.jp/safety/mid-net/0002.html>
内閣府「次世代医療基盤法ガイドライン」:<https://www8.cao.go.jp/iryou/hourei/pdf/guideline.pdf>
厚生労働省保険局「医療・介護データ等解析基盤（HIC）の利用に関するガイドライン」:<https://www.mhlw.go.jp/content/12400000/001174849.pdf>
© 2025. For information, contact Deloitte Tohmatsu Group.

国内におけるビジティング解析環境の比較(2/2)

名称	MID-NET	次世代医療基盤法	医療・介護データ等解析基盤(HIC)																																				
利用料金	<div>■ 製造販売後調査： 利用1件あたり42,123,000円</div> <div>■ 製造販売後調査以外の調査： 利用1件あたり10,820,000～21,061,500円</div> <div>PMDAにデータを提供している協力医療機関向けにはより低額な料金が設定されている</div>	<div>■ 利用1件あたり約20,792,727円 (2023年度の推計値)</div> <div>一般社団法人ライフデータイニシアティブ(LDI)の2023年度の公表資料より、以下の計算式で算定した推計値、 (年間収益)÷(年間で、同法人の審査委員会において利活用が承認された件数)</div>	<div>■ トライアルデータセットの解析環境(3-6か月間)： 利用1件あたり約600,000～1,200,000円(目安)</div> <div>■ 特別抽出等の解析環境(6か月間)： 利用1件あたり約1,500,000～2,800,000円(目安)</div> <div>トライアルデータセット及び特別抽出の内容は以下の通り。なお、令和6年度のHIC機能拡充に伴い手数料設定が見直される見込み</div> <div><ul style="list-style-type: none">トライアルデータセット：NDBの各年1,4,7,10月の医科・DPC・調剤レセプトから外来1%,入院10%程度を抽出し機微な項目を除いたプリセットのデータセット特別抽出：利用者の申し出に応じてNDBから抽出・提供するデータ</div>																																				
利用実績	<table><tr><td></td><td>行政</td><td>企業</td><td>その他企業・アカデミア</td></tr><tr><td>2023年度</td><td>15件</td><td>4件</td><td>1件</td></tr><tr><td>2022年度</td><td>30件</td><td>3件</td><td>-</td></tr><tr><td>2021年度</td><td>30件</td><td>3件</td><td>-</td></tr></table> <div>MID-NETの2023年度利用実績は、2024年1月末までの件数である</div>		行政	企業	その他企業・アカデミア	2023年度	15件	4件	1件	2022年度	30件	3件	-	2021年度	30件	3件	-	<table><tr><td></td><td>-</td></tr><tr><td>2023年度</td><td>12件</td></tr><tr><td>2022年度</td><td>7件</td></tr><tr><td>2021年度</td><td>-</td></tr></table>		-	2023年度	12件	2022年度	7件	2021年度	-	<table><tr><td></td><td>本省利用</td><td>第三者提供</td></tr><tr><td>2023年度</td><td>-</td><td>-</td></tr><tr><td>2022年度</td><td>255件</td><td>72件</td></tr><tr><td>2021年度</td><td>218件</td><td>74件</td></tr></table> <div>HICの利用実績として掲載している件数は、いずれもNDBの抽出件数である</div>		本省利用	第三者提供	2023年度	-	-	2022年度	255件	72件	2021年度	218件	74件
	行政	企業	その他企業・アカデミア																																				
2023年度	15件	4件	1件																																				
2022年度	30件	3件	-																																				
2021年度	30件	3件	-																																				
	-																																						
2023年度	12件																																						
2022年度	7件																																						
2021年度	-																																						
	本省利用	第三者提供																																					
2023年度	-	-																																					
2022年度	255件	72件																																					
2021年度	218件	74件																																					
備考	MID-NETにおける利用料金、利用実績は、MID-NET接続環境(リモート接続)を利用する場合に限らず、PMDAのオンサイトセンターを利用する場合も含む情報である。	次世代医療基盤法における利用料金(推計値)及び利用実績は、一般社団法人ライフデータイニシアティブ(LDI)の情報である。また、改正次世代医療基盤法の施行前の情報についているため、匿名加工医療情報に係る利用料金(推計値)、利用実績である。	令和6年度のHIC機能拡充に伴い、HICガイドラインの内容が見直される予定である。																																				

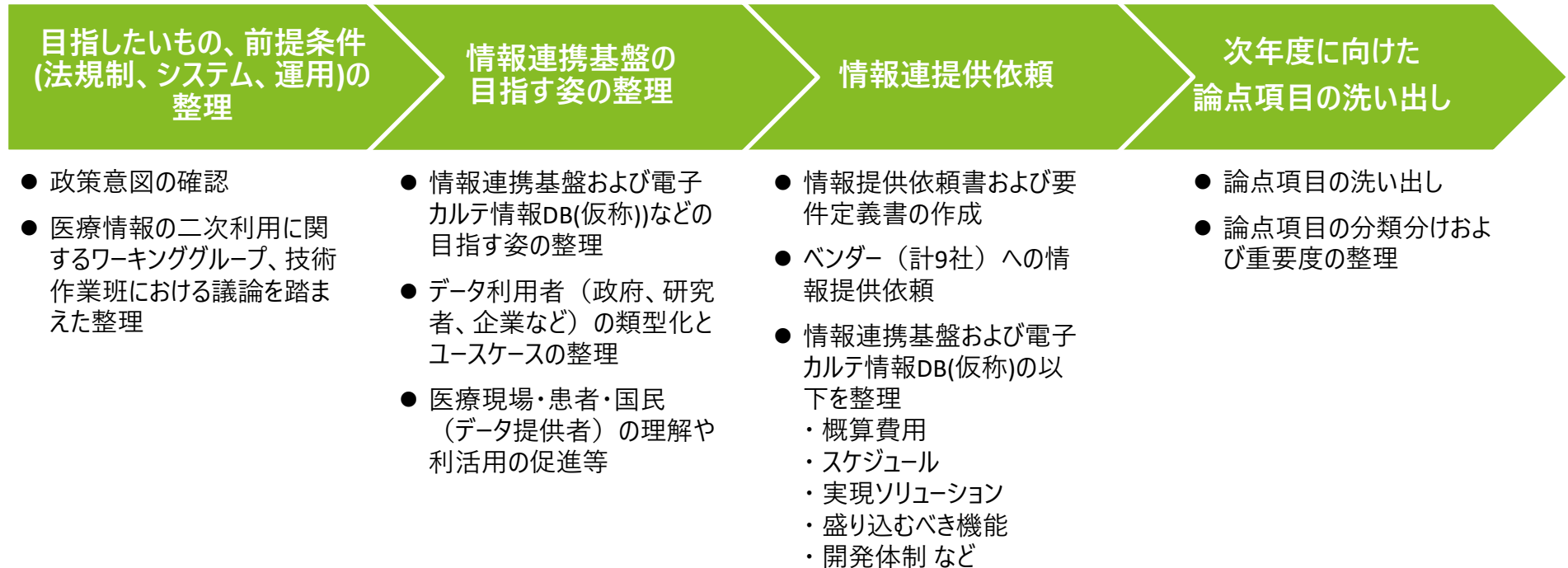
出所：
MID-NET® シンポジウム 2024「～新ステージへの幕開け～MID-NET®の利活用促進に向けた取り組み」：<https://www.mhlw.go.jp/stf/houdou/0000174868.html>
厚生労働省「匿名医療保険等関連情報データベース（NDB）の第三者提供よくあるご質問（FAQ）」：<https://www.mhlw.go.jp/content/12400000/001253089.pdf>
厚生労働省第165回社会保障審議会医療保険部会 資料2「NDBデータの利活用の更なる促進について」：<https://www.mhlw.go.jp/content/12401000/001114695.pdf>
一般社団法人ライフデータイニシアティブ「事業報告書」、「収支決算書」：<https://www.ldi.or.jp/plan>

4. 医療等情報の二次利用に係る取組

4.1 実運用も考慮した具体的な運用・システムの整理

二次利用に係る具体的な運用・システムの整理は以下の工程で整理した

具体的な運用・システムの検討の進め方(案)



**目指したいもの、前提条件(法規制、システム、
運用)の整理**

情報連携基盤の前提条件

(医療情報の二次利用に関するワーキンググループ、技術作業班における議論を踏まえた整理)

全体 / フェーズ	あるべき姿	前提条件
全体アーキテクチャ	<ul style="list-style-type: none"> ■ 医療等情報の二次利用を推進するため、「電子カルテ情報共有サービス」で共有される電子カルテ情報と厚生労働大臣が保有する医療・介護関係のDBや一定の要件を満たした民間DB等を一元的かつ安全に利用・解析できる「情報連携基盤」を構築する ■ 現行の公的DBでは匿名化情報の利活用のみが定められており、研究利用への期待が大きい仮名化情報を利用できるようにする ■ 本人関与の機会の確保への配慮として、仮名化情報を提供するに当たり、保護措置等を講ずることで本人の権利利益を適切に保護する ■ 情報連携基盤のセキュリティについて、不正なアクセスや高度なセキュリティ攻撃からデータ漏えい等を防ぐための安全管理措置など、必要なセキュリティ要件を整理する必要がある ■ データ利用者の利便性の観点で、利用申請の受付窓口、審査体制は原則一元化し、審査の手順や内容を統一化する 	<ul style="list-style-type: none"> ✓ データソースは電子カルテ情報共有サービスおよび公的DB等とする ✓ <u>医療情報を一元的かつ安全に利用・解析できる情報連携基盤を構築する</u> ✓ 匿名化情報では研究等が困難であり正当な理由がある場合に限り、<u>仮名化情報の連結解析を可能</u>とし、個人の特定リスクも考慮して適切に審査を実施する ✓ 本人からの利用停止の求めに対応できるよう<u>二次利用に対するオプトアウトを可能</u>とする ✓ <u>管理者側に厳格な安全管理措置</u>を設け、具体的な要件(利活用者の認証、ログの保存・監視・活用によるデータトレーサビリティの確保等)を整理する必要がある。また、<u>データ利用者に対して利便性も考慮して解析環境等の整備</u>を行う ✓ 情報連携基盤のデータ利用に係る<u>申請および審査は一括</u>で行う ※医学系倫理指針の要件を満たすものとし、各研究機関における倫理審査委員会の審査は必ずしも求めない
データ収集	<ul style="list-style-type: none"> ■ 電子カルテ情報共有サービスおよび公的DB等からのデータ収集にあたっては、データの連結性を高め、データの正確性を確保し、定期的な更新頻度を維持することが求められる 	<ul style="list-style-type: none"> ✓ <u>連携元DBでID(ID5またはID4を想定)を付与</u>した上で、元の識別子(被保険者番号、氏名等)は削除し情報連携基盤に連携する
データ管理	<ul style="list-style-type: none"> ■ 二次利用しやすいデータベースを構築するため、収集したデータのクレンジングを行う等の品質管理を行う 	<ul style="list-style-type: none"> ✓ <u>二次利用データのクレンジングおよびデータ検証</u>を行い、二次利用データの品質を確保する
データ抽出	<ul style="list-style-type: none"> ■ 様々なデータをデータ利用者の要望に応じて迅速に抽出できる ■ 利用可能なデータを可視化するためのメタデータ(データカタログ等)が整備されている 	<ul style="list-style-type: none"> ✓ データ利用者の研究への利便性を担保するため、<u>データは事前に連結し主な利用ニーズに合わせたデータセットを提供</u>する
データ活用	<ul style="list-style-type: none"> ■ 仮名化情報はVisiting解析環境での利用を基本とし、データ利用者の利便性も考慮して解析環境等の整備を行う ■ データ利用者のデータ利用を支援するために、利用者ポータルを整備し、利用可能なデータを一覧化するデータカタログ、オープンソースのデータを簡易に集計・分析するダッシュボード機能を設ける 	<ul style="list-style-type: none"> ✓ データ利用者は<u>Visiting解析環境（オンライン上の解析環境）での利用を基本</u>としてデータの解析を行う ✓ <u>ポータル機能、データカタログ、ダッシュボード機能</u>を設ける

情報連携基盤サービスでは下記のユーザーによる利用を想定している

情報連携基盤サービスの利用ユーザーの種別とユースケース例

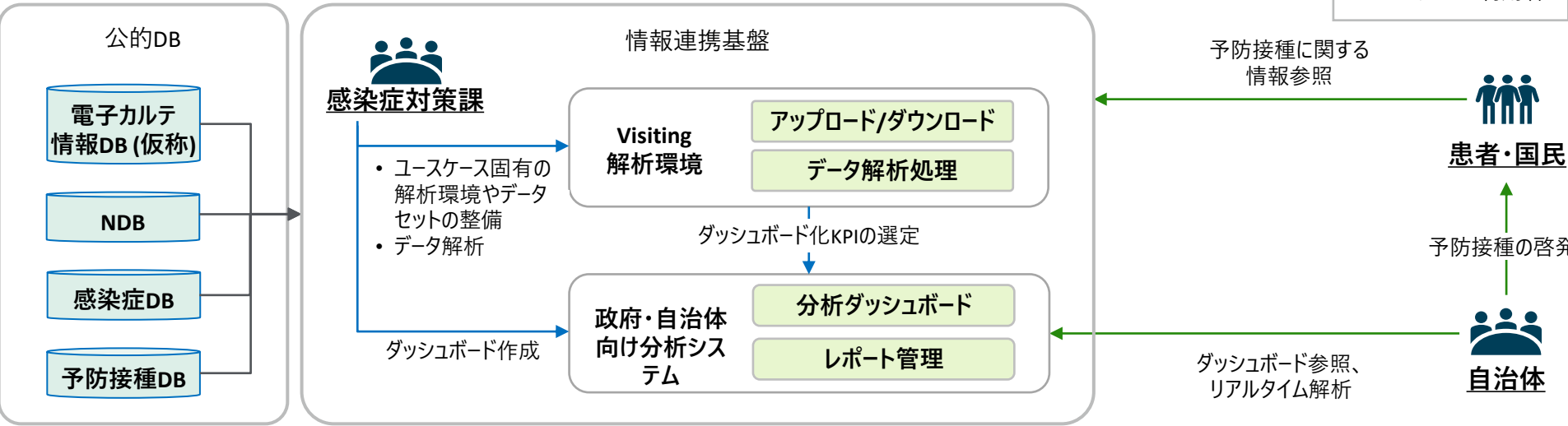
利用ユーザー		ユースケース例	
政府の政策担当者	EBPM による政策立案と検証・改善	生活習慣病予防対策推進	✓ 生活習慣病予防KPIの可視化 ✓ 生活習慣病発症リスクの予測
		重症患者診療体制整備事業	✓ 重症患者診療体制整備事業への活用
		予防接種データ分析	✓ 予防接種データの可視化および施策への活用
		疾病別(がん、難病等)医療費分析	✓ 疾病に関する医療費等の分析
		医療計画策定	✓ 5疾病6事業分析
		医薬品の安全対策	✓ 医薬品の安全性サーベイランス
		希少疾患	✓ 希少疾病患者の統計調査等
		介護データ分析	✓ 地域包括支援センター等におけるデータ共有 ✓ 高額介護給付費の削減
		感染予防や流行時の危機管理対策の強化	✓ 新型コロナ薬剤処方に係る公費支援の在り方の検討 ✓ 感染症罹患状況の可視化 ✓ 感染症のリアルタイムデータ分析（平時・有事）
		医療費適正化	✓ 医療費適正化計画（都道府県データブック） ✓ 地域別の医療費分布や傾向の可視化、医療費の適正化に向けた施策評価
自治体の政策担当者	地域医療の最適化	地域医療計画	✓ 再編統合等の医療資源の最適化
		行政施策の評価	✓ 感染症等のワクチン接種者数や感染者数の分析
学術・研究機関の研究者	研究・論文の増加	臨床研究	✓ 患者レベルの予測(発症リスク、薬剤の副作用、治療効果、再入院リスク等) ✓ 疾病のコホート分析 (治療法の効果、予防策の有効性、生存率分析等) ✓ LLMを活用した新規医療や診断補助アプリケーションの開発
		手法評価	✓ 研究方法論やデータ解析手法の開発、評価、改善
	国際共同研究の促進	各国の国際比較	✓ 連合学習を活用した医療・健康データの国際共同研究 ✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等
民間企業 (製薬)	治験、臨床試験、市販後調査等	医薬品の有効性・安全性評価 臨床情報の解析による市販後調査等、	✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析 ✓ 臨床情報の解析による患者ニーズの特定 ✓ 費用対効果の評価、既存薬剤との比較、分配の適正評価等
民間企業 (医療機器)	治験、臨床試験、市販後調査等 AI機器開発の推進	医療機器の有効性・安全性評価 臨床情報の解析による市販後調査等、 AI搭載機器の機械学習	✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等 ✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析 ✓ AI搭載機器やそのための機械学習の研修開発
民間企業 (再生医療等製品)	治験、臨床試験、市販後調査等	再生医療等製品の有効性・安全性評価 臨床情報の解析による市販後調査等、	✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等 ✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析
医療機関	臨床教育、診断補助	過去症例の参照や診断補助等	✓ 過去症例にアクセスし、現実に即したケーススタディや教育 ✓ LLMを用いた診断補助など
患者・国民	医療・健康への理解向上	自身の健康増進や予防への理解促進	✓ 医療・健康等の情報参照等

各地域の予防接種DB及び電カル情報DB、NDB、感染症DBとの統合でデータを分析し、ダッシュボードで可視化された情報に参照することで予防接種計画の策定等に役立つ

ユースケース例：予防接種の可視化

ユーザー像	<ul style="list-style-type: none">厚生労働省健康・生活衛生局 感染症対策部 予防接種課	利用するDBとデータ項目	<ul style="list-style-type: none">電子カルテ情報DB<ul style="list-style-type: none">✓ 6情報：傷病名、アレルギー情報、感染症情報、薬剤禁忌情報、検査情報、処方情報NDB<ul style="list-style-type: none">✓ 傷病名（レセプト病名）、投薬、健診結果等感染症DB<ul style="list-style-type: none">✓ 感染症の名称・症状、診断方法、初診年月日・診断年月日、発病推定年月日等予防接種DB<ul style="list-style-type: none">✓ ワクチン情報、接種場所、副反応の症状等
背景	<ul style="list-style-type: none">予防接種に関して国民向けの情報提供ツールとしてホームページが中心的な役割を果たしているが、情報提供のあり方が一元化されていない現段階では感染症DBにコロナの情報しかなく、今後様々な感染症の情報を追加する予定である		
目的	各地域の予防接種DBにあるデータを統合し、電子カルテ情報DB、NDB、感染症DBと結合することによって、地域別の予防接種計画の策定等に活用する		

<データフロー>



各地域の予防接種DB及び電カル情報DB、NDB、感染症DBとの統合でデータを分析し、ダッシュボードで可視化された情報に参照することで予防接種計画の策定等に役立つ

ユースケース例：予防接種の可視化

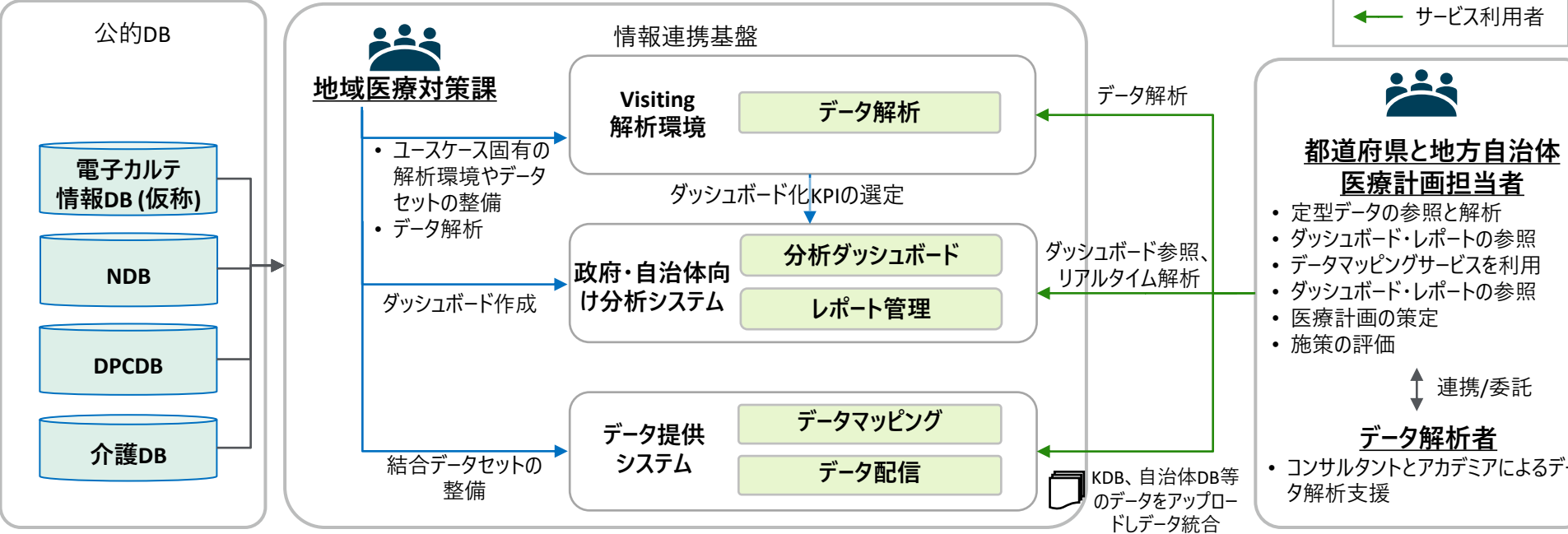
ユースケース	概要	利用DB				機能			アウトプット
		NDB	感染症DB	予防接種DB	電カルDB	ダッシュボード	Visiting解析環境	リアルタイム解析	
予防接種の接種意思の分析	地域や年齢層、性別などの属性別に、市民の予防接種に対する意思を分析し、接種を躊躇する要因を特定することで、接種促進のための効果的な施策を立案する			○	○	○	○		<ul style="list-style-type: none"> 接種意思分析レポート 接種促進施策計画書
予防接種の実施状況の把握	地域ごとの予防接種の実施状況をリアルタイムで把握し、接種率の低い地域や年齢層を特定する。これにより、効果的な介入策を講じて接種率を向上させる計画を立案する			○	○	○	○	○	<ul style="list-style-type: none"> 実施状況把握レポート 地域別接種状況マップ 接種計画改善提案書
予防接種のモニタリング及び効果の評価	予防接種後の健康状態を継続的にモニタリングし、接種の効果を評価する。接種後の感染症発生率、重症化率や副作用状況を分析し、予防接種の有効性を定量化する	○	○	○	○	○	○	○	<ul style="list-style-type: none"> 効果評価レポート モニタリングダッシュボード 接種効果評価ガイドライン
予防接種の推奨接種頻度	各種予防接種の最適な接種頻度を検討するため、年齢や健康状態に応じた接種スケジュールをデータに基づいて策定する	○	○	○	○	○	○	○	<ul style="list-style-type: none"> 推奨接種頻度レポート 接種頻度ガイドライン 最適接種スケジュール
予防接種の地域別最適化	地域ごとの感染症発生状況と予防接種率を分析し、地域特性に応じた最適な予防接種計画を策定する	○	○	○	○	○	○	○	<ul style="list-style-type: none"> 地域別予防接種最適化レポート 地域別接種率マップ 予防接種計画書

公的DBの患者情報および人口動態データ等进行分析することで、地域ごとの医療需要やニーズを把握・予測し、地域医療資源の最適化を実現することができる

ユースケース例：地域医療資源の最適化

ユーザー像	<ul style="list-style-type: none">厚生労働省医政局・地域医療対策課都道府県と地方自治体の医療計画担当者	利用するDBとデータ項目	<ul style="list-style-type: none">電子カルテ情報DB<ul style="list-style-type: none">✓ 3文書6情報NDB:<ul style="list-style-type: none">✓ 診療行為の内容、頻度、医療機関別の患者数、請求点数等DPCDB<ul style="list-style-type: none">✓ 入院情報、退院情報、患者情報、診断情報等介護DB<ul style="list-style-type: none">✓ 要介護認定情報、介護レセプト情報等
背景	<ul style="list-style-type: none">地域ごとの医療リソースの偏在が進んでおり、医療サービスの均等化が課題になる高齢化や人口減少により、医療ニーズの変化が生じ、資源の最適配分が求められる		
目的	地域ごと医療資源分布の偏在及び医療ニーズを可視化し、医療資源配分及び医療費を適正にする		

<データフロー>



公的DBの患者情報および人口動態データ等进行分析することで、地域ごとの医療需要やニーズを把握・予測し、地域医療資源の最適化を実現することができる

ユースケース例：地域医療資源の最適化

区分	ユースケース	概要	利用DB				機能		アウトプット
			NDB	感染症DB	DPCDB	電カルDB	ダッシュボード	Visiting解析環境	
現状把握	医療資源の地域分布状況及び格差・偏在分析	地域別医療資源マップを作成し、各地域における医療機関、医療従事者、医療機器などの資源分布を可視化し、医療資源の偏在を特定する	○		○		○		<ul style="list-style-type: none">医療資源分布レポート医療資源配分計画
	患者の流入・流出状況分析	患者の地域間移動や受診傾向を分析し、地域ごとの医療需要の充足度を評価できる患者流入・流出分析システムを構築し、必要な医療サービスの提供や地域医療の充実に役立てる情報と提供する	○		○		○	○	<ul style="list-style-type: none">患者流入・流出レポート地域間の需要ギャップ分析地域医療サービス改善提案
	地域別の疾病発生率と医療ニーズの分析	地域ごとの疾病発生率や患者属性（年齢、性別など）を分析し、地域別疾病ニーズを特定し、地域ごとに必要とされる医療サービスの分布を明確にする	○	○	○	○	○	○	<ul style="list-style-type: none">地域別の疾病分析レポート地域別の医療ニーズレポート医療サービス提供方針提案
改善対応	地域医療機関の役割分担と連携強化による医療提供体制の最適化	地域ごとに異なる医療ニーズや資源の分布に応じて、各医療機関の役割を明確にし、地域間での連携を強化することで、効率的な医療提供体制による資源偏在の緩和を実現する			○	○	○	○	<ul style="list-style-type: none">医療機関ごとの役割分担リスト医療資源調整計画
	地域の医療サービス供給やニーズに合わせた病院の再編統合	地域の医療ニーズや供給の状況に基づき、病院の再編や統合を実施し、効率的かつ持続可能な医療提供体制を構築する。各病院の診療報酬や医療収益も考慮し、財政面でも安定した運営が可能な医療リソース配分を実現する	○		○			○	<ul style="list-style-type: none">地域病院再編統合計画診療科・病床数の再配置計画診療報酬改善による財務シミュレーション
	地域別医療費の適正化	地域別に診療報酬や医療費の傾向を分析し、不均衡のある領域を特定する。不要な診療や検査を削減するための基礎データを提供し、地域別医療費の適正化の実現する	○	○		○	○	○	<ul style="list-style-type: none">地域別医療費分析レポート医療費の削減提案書
将来予測	地域別疾病発生予測と医療資源配分計画	各地域の疾病発生率や人口動態を基に将来の医療需要を予測し、病床や医療機器、人員など必要な医療資源の配分計画を策定する	○	○	○	○		○	<ul style="list-style-type: none">地域別疾病発生予測レポート医療資源配備計画将来医療ニーズに基づくリソース最適化計画

電子カルテ情報や公的DB等の大規模データを活用した学習により、医療用LLMを構築し、臨床での診断補助を目的としたアプリケーション開発が期待される

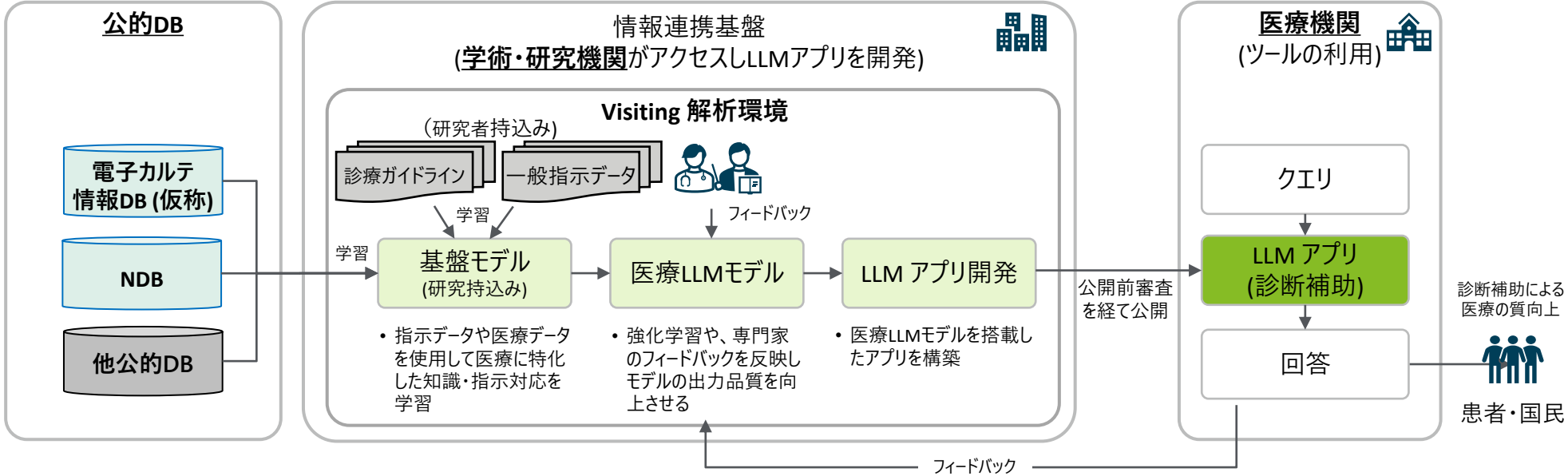
ユースケース例：LLMを活用した電子カルテ情報のリアルタイム分析による診断補助アプリケーション開発

ユーザー像	学術・研究機関
利用目的	診断補助
背景	<ul style="list-style-type: none">診断の効率化と質の向上により、医療全体の質向上が期待される地域医療における人材不足の対策として、AIアシストの導入が有効と考えられる2021年以降、LLM関連の医学研究が急増しており、米国、インド、中国などが積極的に取り組んでいる

利用するDBとデータ項目(案)	<ul style="list-style-type: none">電子カルテ情報DB<ul style="list-style-type: none">✓ 3文書：診療情報提供書、退院時サマリー、健康診断結果報告書✓ 6情報：傷病名、アレルギー情報、感染症情報、薬剤禁忌情報、検査情報、処方情報NDB<ul style="list-style-type: none">✓ 診断アウトカムの時系列情報その他対象とする疾患領域に該当するDB<ul style="list-style-type: none">✓ 診断情報上記の他、診療ガイドラインや一般指示データなど
-----------------	--

※臨床に影響するため、データの品質やバイアスの管理、有効性の検証等が重要課題となる

<データフロー>



情報連携基盤の目指す姿の整理

情報連携基盤の目指す姿

背景・目的

- 医療等情報の二次利用を通じて、「国民の健康増進、より質の高い医療・ケア、医療の技術革新（医学研究、医薬品開発等）、医療資源の最適配分、社会保障制度の持続性確保（医療費の適正化等）、次の感染症危機への対応力の強化などにつなげていく（規制改革実施計画（令和5年6月16日閣議決定））」必要がある。
- 一方、医療等情報の活用に当たっては、医療現場の理解を得つつ、本人の権利利益を保護し、国民の患者の理解を得ながら、適切に二次利用を推進する必要がある
- 日本では、仮名化情報の利用の制約や、データベースの整備や利用申請手続きに係る負荷が多く、リアルワールドデータの研究利用が進みにくい状況にある。これらの課題を解決し、医療等情報を適切に活用するための環境整備が求められている

目指す姿

データ利用者

以下のデータ利用者の医療等情報の二次利用を通じて、政策目標の実現を目指す

政策	医学研究	医療	開発およびイノベーション	健康管理/教育
政府・自治体 <ul style="list-style-type: none">EBPM による政策立案と検証・改善感染予防や流行時の危機管理対策の強化医療費や予算分配の最適化による財政持続可能性の検討医薬品の安全対策	学術・研究機関等 <ul style="list-style-type: none">臨床データの解析による疾病リスク因子、疾患予後予測の解明長期データ連携による治療・薬剤の影響評価国際共同研究の推進AIアルゴリズム研究の推進	医療機関 <ul style="list-style-type: none">診療や治療方針の向上のための教育としてデータ利用研修医や医学生の臨床教育やケーススタディ	民間企業 （製薬・医療機器会社等） <ul style="list-style-type: none">臨床情報の解析による市販後調査等、医薬品・医療機器の安全性評価研究・イノベーション推進による産業育成	患者・教育機関 <ul style="list-style-type: none">医療・健康に関する統計データ参照自身の健康管理や病気の予防を目的とした教育コンテンツ統計学習への活用

情報連携基盤の構築
公的DB等の情報の利用や連結解析が可能な環境を整備



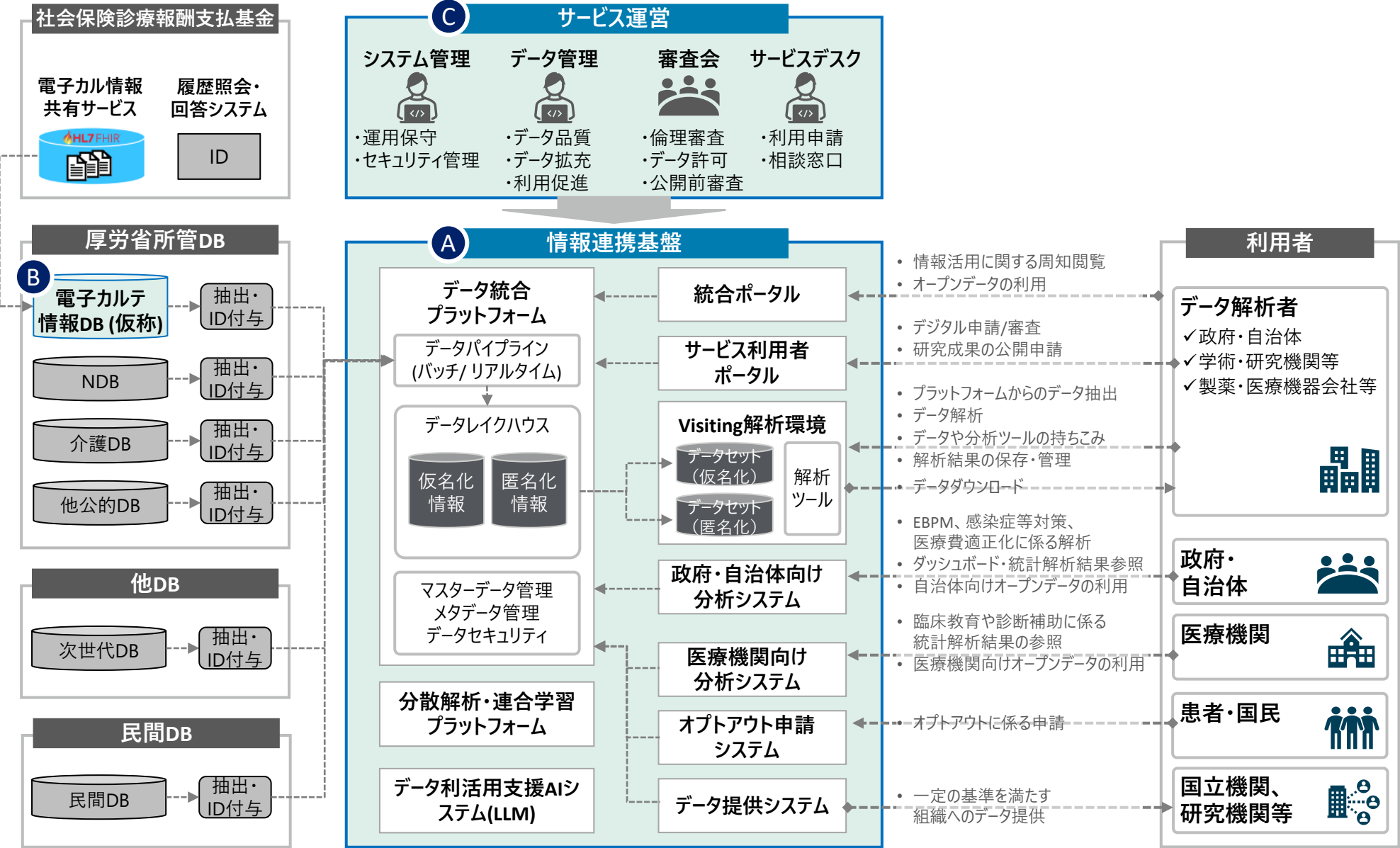
医療機関等・患者

- 安全なデータの活用
- 理解増進
- データ提供に対するインセンティブの付与(研究の成果フィードバック等)
- 権利利益の保護

円滑なデータ提供等により、医療等情報の二次利用が可能な環境を構築する

データ提供者

本事業の全体像は下記の通り。A. 情報連携基盤とB. 電子カルテ情報DB（仮称）の2つのシステムに加え、C. 情報連携基盤のサービス運営体制を構築する



4.2 情報提供依頼の実施概要

4.2.1 情報提供依頼書

事業の概要

情報連携基盤の目指す姿

背景・目的

- 医療等情報の二次利用を通じて、「国民の健康増進、より質の高い医療・ケア、医療の技術革新（医学研究、医薬品開発等）、医療資源の最適配分、社会保障制度の持続性確保（医療費の適正化等）、次の感染症危機への対応力の強化などにつなげていく（規制改革実施計画（令和5年6月16日閣議決定））」必要がある。
- 一方、医療等情報の活用にあたっては、医療現場の理解を得つつ、本人の権利利益を保護し、国民の患者の理解を得ながら、適切に二次利用を推進する必要がある
- 日本では、仮名化情報の利用の制約や、データベースの整備や利用申請手続きに係る負荷が多く、リアルワールドデータの研究利用が進みにくい状況にある。これらの課題を解決し、医療等情報を適切に活用するための環境整備が求められている

目指す姿

データ利用者

以下のデータ利用者の医療等情報の二次利用を通じて、政策目標の実現を目指す

政策	医学研究	医療	開発およびイノベーション	健康管理/教育
政府・自治体 <ul style="list-style-type: none"> EBPM による政策立案と検証・改善 感染予防や流行時の危機管理対策の強化 医療費や予算分配の最適化による財政持続可能性の検討 医薬品の安全対策 	学術・研究機関等 <ul style="list-style-type: none"> 臨床データの解析による疾病リスク因子、疾患予後予測の解明 長期データ連携による治療・薬剤の影響評価 国際共同研究の推進 AIアルゴリズム研究の推進 	医療機関 <ul style="list-style-type: none"> 診療や治療方針の向上のための教育としてデータ利用 研修医や医学生の臨床教育やケーススタディ 	民間企業 (製薬・医療機器会社等) <ul style="list-style-type: none"> 臨床情報の解析による市販後調査等、医薬品・医療機器の安全性評価 研究・イノベーション推進による産業育成 	患者・教育機関 <ul style="list-style-type: none"> 医療・健康に関する統計データ参照 自身の健康管理や病気の予防を目的とした教育コンテンツ 統計学習への活用

情報連携基盤の構築
公的DB等の情報の利用や連結解析が可能な環境を整備



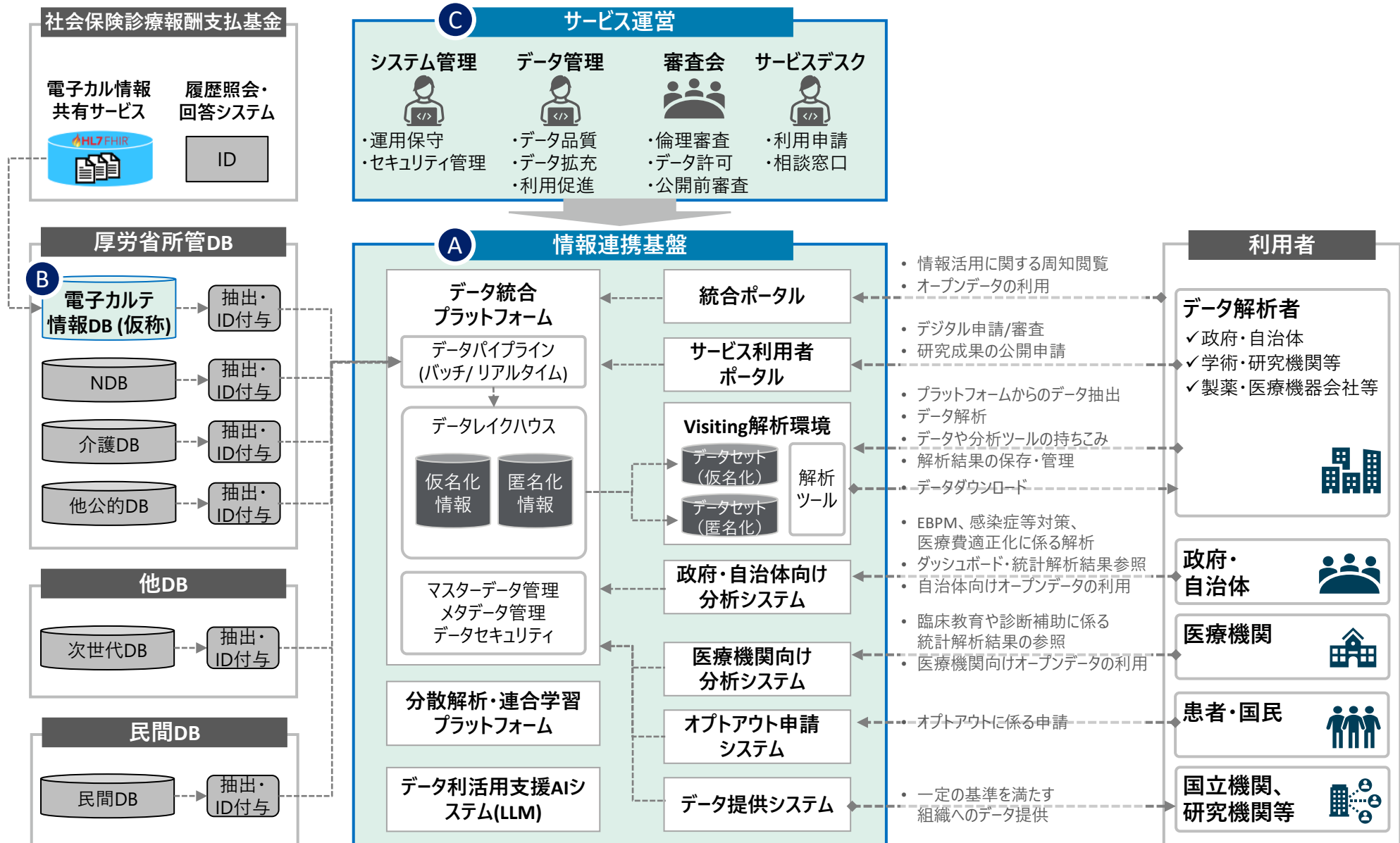
医療機関等・患者

- 安全なデータの活用
- 理解増進
- データ提供に対するインセンティブの付与(研究の成果フィードバック等)
- 権利利益の保護

円滑なデータ提供等により、医療等情報の二次利用が可能な環境を構築する

データ提供者

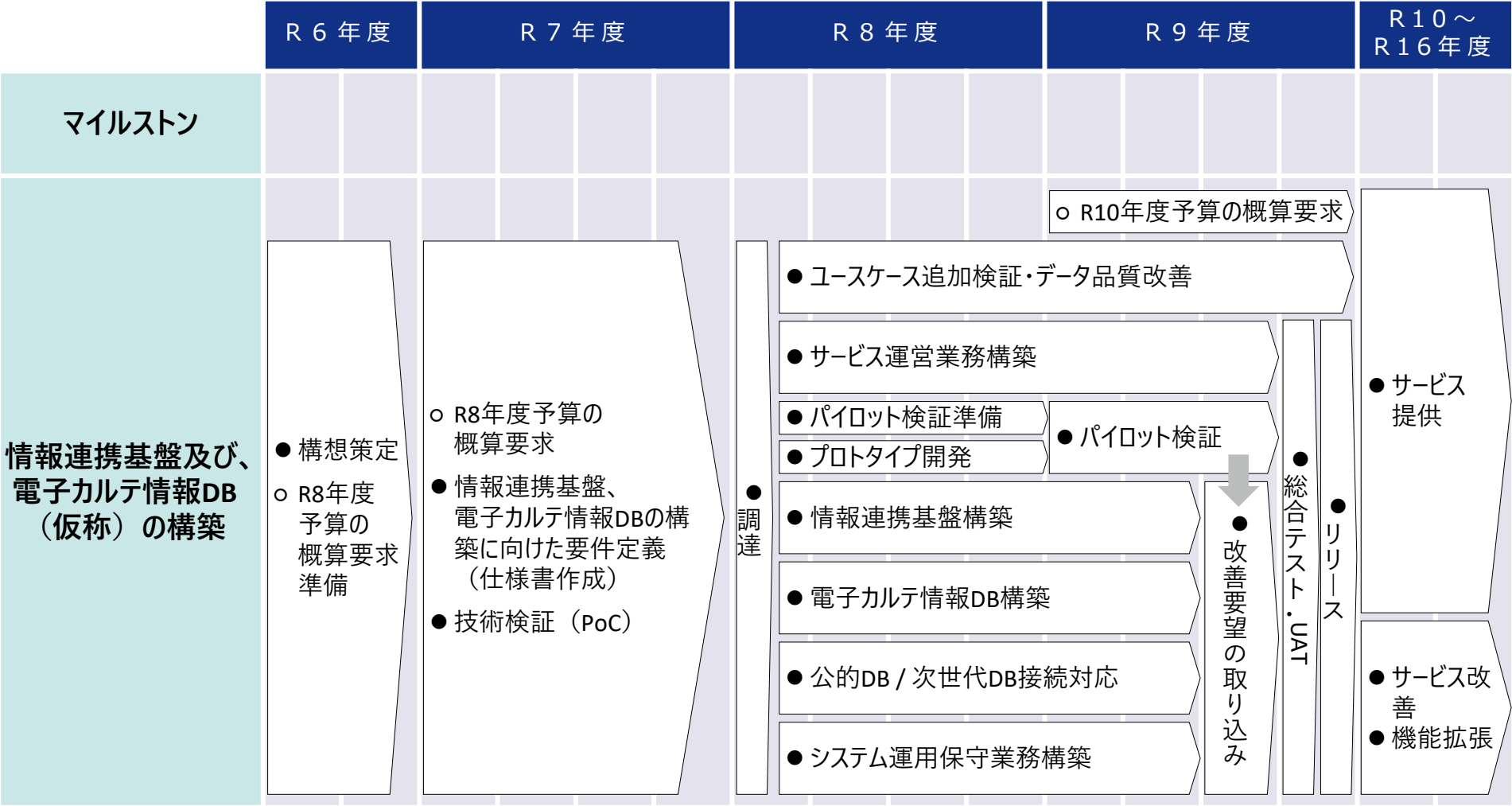
本事業の全体像は下記の通り。A. 情報連携基盤とB. 電子カルテ情報DB（仮称）の2つのシステムに加え、C. 情報連携基盤のサービス運営体制を構築する



R7年度に要件定義と調達仕様書の作成、技術検証を実施し、R8年～R9年で情報連携基盤と電子カルテ情報DB（仮称）を構築し、R10からサービス提供を開始する

今後の年間スケジュール

○：翌年度に向けて対応する事項
●：当該年度の実施事項



R10年からのサービス提供は段階的に行う方針とし、それに合わせ機能も段階的に拡張する (R6年～R7年で、サブシステム内の機能レベルで開発スコープを詳細に定義する)

サービス展開方針と開発スコープ方針

* 政府・自治体、アカデミアの利用を想定

	R8年度	R9年度	R10年度	R11年度
サービス展開方針	(開発中)	一部利用者*へ パイロット公開 <ul style="list-style-type: none">一部の利用者にサービスをパイロット公開し運営体制の習熟を図る分析ユースケースの実装と効果検証を行いデータ利活用のスタートを切る	全利用者へ公開 <ul style="list-style-type: none">全利用者への公開に必要な運営体制の実現ユースケースの想起を誘うような取り組み事例公開	更なるサービス強化 <ul style="list-style-type: none">サービス利用ユーザ種別の拡大（医療機関等）利用者のニーズに合わせたサービス機能の拡充サービス運営を省力化するシステムの実現
開発スコープ	✓ プロトタイプを作成し、R10年度の全利用者に向けたパイロットテストを実施 開発スコープ案) <ul style="list-style-type: none">統合データプラットフォームVisiting解析環境政府・自治体向け分析システムデータ提供システム統合、サービス利用者、オプトアウト申請ポータルデータ利活用支援AIシステムID管理システム、契約管理システム、サービス管理システム公的DB、統合ネットワーク	✓ 全利用者への公開に必要な機能の拡充 開発スコープ案) <ul style="list-style-type: none">医療機関向け分析システム統合データプラットフォームへのデータソース追加（次世代DB等）	✓ 利用者のニーズや、サービスの拡張計画に合わせた継続的な開発 開発スコープ案) <ul style="list-style-type: none">統合データプラットフォームへのデータソース追加（民間・学会レジストリ等）分散解析・連合学習プラットフォーム	

情報提供依頼事項

情報提供依頼事項は下記の通り

情報提供依頼事項サマリ

情報提供依頼事項		
1	回答方法	
<ul style="list-style-type: none">● 後述する情報提供書作成要領に沿って、情報提供書を作成すること● 別紙 1 見積回答書へ、工数・費用等必要情報を記載すること		
2	回答スケジュール	
<ul style="list-style-type: none">● 回答期限: R7年〇月〇日 17時迄		
3	スコープ	対象システムとシステム要件
<ul style="list-style-type: none">● 対象システム:<ul style="list-style-type: none">・ 情報連携基盤、及び電子カルテ情報データベース（仮称）● 対象システムの要件:<ul style="list-style-type: none">・ 別紙 2 を参照のこと		
4		対象作業
<ul style="list-style-type: none">● 下記を想定<ul style="list-style-type: none">・ 対象システムの設計・構築・テスト・リリースに関する作業費用・ 対象システムの運用保守業務構築費用・ R10年～R16年の7年間分の運用保守費用		

情報提供書には下記項目を含めること

1. 見積回答方法（情報提供依頼書回答作成要領）

<p>情報提供書 目次案</p>	<ol style="list-style-type: none"> 1. プロジェクト概要 2. 開発スコープ 3. プロジェクト推進の成功要因プロジェクトアプローチ 4. 成果物 5. スケジュール 6. 体制
<p>回答依頼 事項</p>	<ol style="list-style-type: none"> 1. 情報連携基盤、電子カルテ情報DB（仮称）に盛り込むべき追加の機能提案 2. 情報連携基盤、電子カルテ情報DB（仮称）の実現ソリューションとソリューション提案理由 3. 各公的DBとの接続を見据えた際の必要なタスクや想定されるリスク・課題 ※公的DBの運営・保守業務を行っている場合、ご回答ください 4. 継続的なサービス改善を実現するための高い拡張性を実現するアーキテクチャに関する提案 5. 技術ロックインを防止するためのアーキテクチャに関する提案 6. R8～R9に必要なプロジェクトタスクに対する提案（追加タスクの有無等） 7. 妥当と思われる各タスクの期間・スケジュール

別紙 2. 見積回答書の指定様式に沿って、各システム構成要素の工程毎の総人月・平均単価・期間を記載すること。人件費以外にサービス費用やライセンス費用も記載すること

1. 見積回答方法（別紙 1. 見積回答書作成要領）

【参考見積回答について】

- 本開発を進めて頂くうえでコストとしてかかる、開発費やサービス利用料・ライセンス料等を見積回答書へご記載ください
※「見積対象作業」、「4.情報連携基盤構築」、「5.電子カルテ情報DB構築」、「サービス利用料ライセンス料」の各シートにご記載ください
- 薄緑色に網掛けがしてある部分に工数や数量、数量単位、単価をご記入ください
- 御社側からの新規ソリューションなどの提案も可能な範囲で含めてください（その場合は、依頼事項との明確に区分できるよう、備考欄にその旨を記載する）
- その他の不明点などについては、随時ご連絡ください

■ 見積対象作業					
・薄緑色に網掛けがしてある部分に工数や単価をご記入ください					
Phase	構成要素	工数(人月)	単価(円)	金額(円)	小計(円) *
1. プロジェクト管理・推進					0 *
	開発計画			0	*
	開発管理			0	*
2. サービス構築・アプリ開発支援					0 *
	サービス構築・アプリ開発支援（問い合わせ対応・仕様情報提供）			0	*
3. プロトタイプ開発					0 *
	設計			0	*
	開発			0	*
	テスト			0	*
	ランニング費用			0	*
4. 情報連携基盤構築					0 *
	設計			0	*
	開発			0	*
	テスト			0	*
	ランニング費用（7年分のシステム運用保守費）			0	*
5. 電子カルテ情報DB構築					0 *
	設計			0	*
	開発			0	*
	テスト			0	*
	ランニング費用（7年分のシステム運用保守費）			0	*
6. 公的DB接続対応					0 *
	調査			0	*
	連携方式設計			0	*
	設計			0	*
	構築			0	*
	テスト			0	*

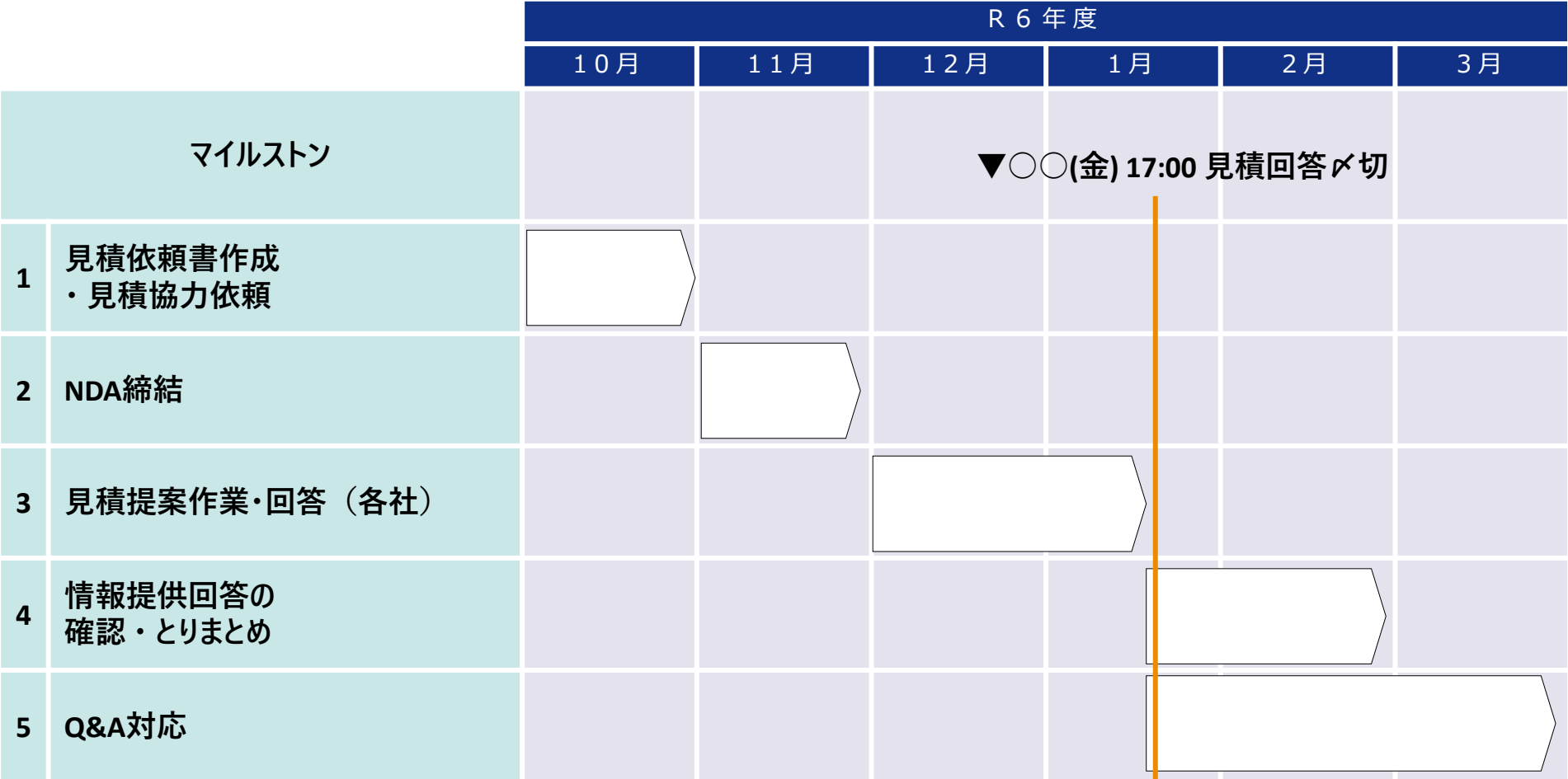
4. 情報連携基盤構築参考見積

・薄緑色に網掛けがしてある部分に工数や単価をご記入ください

No	機能名	回答（選択式）	イニシャルコスト（単位：円）			
			人件費			イニシャルコスト計
			設計	開発	テスト	
例	-	C：個別開発により対応	¥100,000,000	¥100,000,000	¥100,000,000	¥300,000,000
①	データ統合プラットフォーム					¥0
②	分散解析・連合学習プラットフォーム					¥0
③	統合ポータル					¥0
④	オプトアウト申請ポータル					¥0
⑤	サービス利用者ポータル					¥0
⑥	Visiting解析環境					¥0
⑦	政府・自治体向け分析システム					¥0
⑧	医療機関向け分析システム					¥0
⑨	データ提供システム					¥0
⑩	データ活用支援AIシステム					¥0
⑪	統合ネットワーク					¥0
⑫	ID管理システム					¥0
⑬	契約管理システム					¥0
⑭	サービス管理システム					¥0
	合計		¥0	¥0	¥0	¥0

見積にご協力いただける各社においては、〇〇までに見積提案回答をいただきたい
 また回答提出後の質疑応答についてもご協力いただきたい

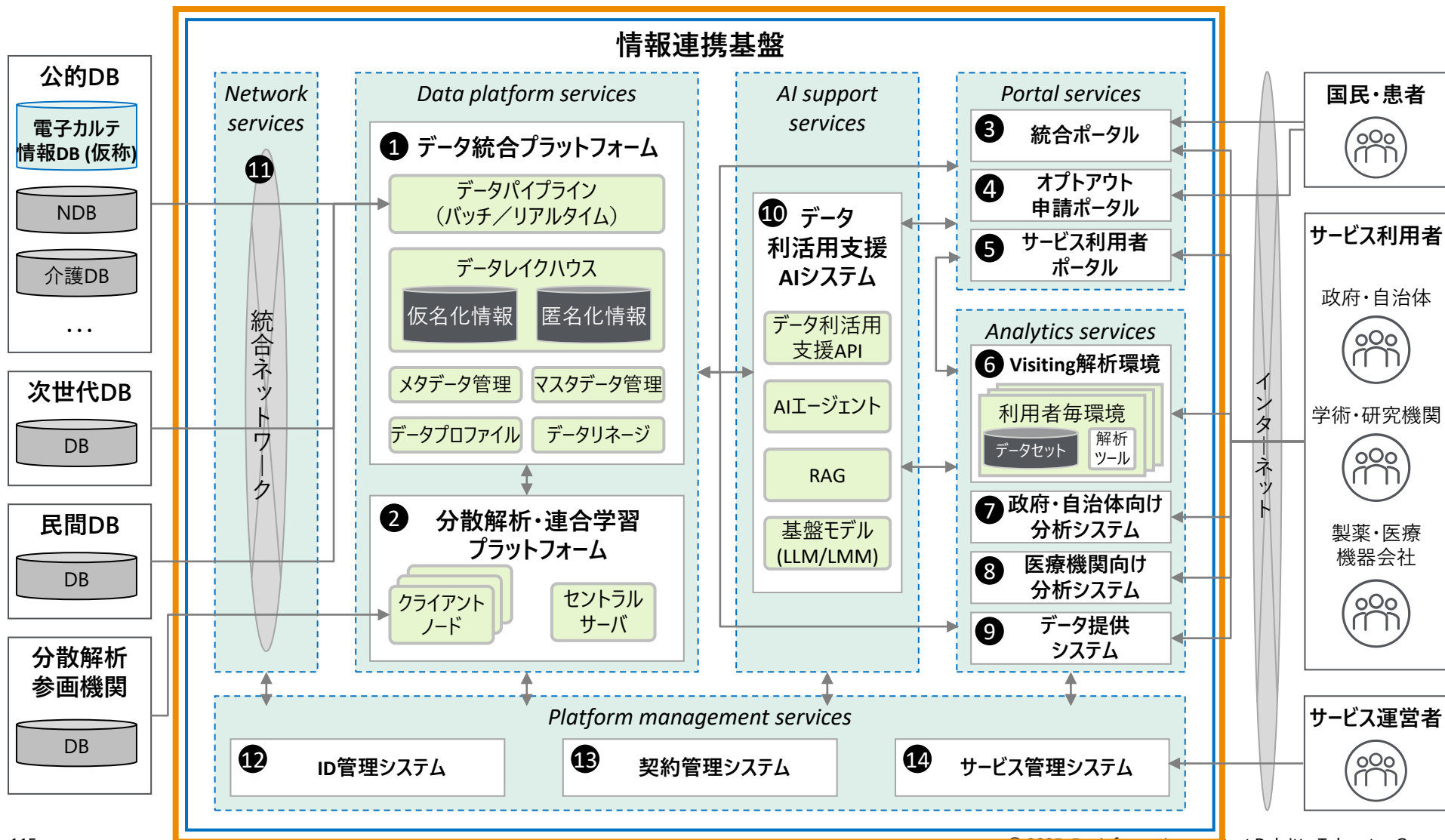
2.見積回答スケジュール



情報連携基盤に対する情報提供依頼対象は14要素となる。適宜「別紙2」を参照すること

3.構築対象システムの構成 — A. 情報連携基盤

情報提供依頼スコープ



情報連携基盤は、構成要素が多岐に渡るため下記の情報提供を依頼したい

3.情報提供依頼事項

対象	情報提供依頼事項
<ul style="list-style-type: none"> ① データ統合プラットフォーム ⑤ サービス利用者ポータル ⑥ Visiting解析環境 ⑦ 政府・自治体向け分析システム ⑨ データ提供システム ⑩ データ利活用支援AIシステム 	<ul style="list-style-type: none"> ● 情報提供依頼事項 <ul style="list-style-type: none"> ✓ 貴社が開発する場合の採用ソリューションと採用理由 ✓ 追加で盛り込むべき要求事項とその理由 ✓ 継続的な改善を実現する拡張性の高いアーキテクチャに関する提案 ✓ 技術ロックインを防止するためのアーキテクチャに関する提案 ● 見積 <ul style="list-style-type: none"> ✓ 要求事項を元にした、概算工数の積み上げによる見積り ※ 不足情報はQA等で適宜補足予定
<ul style="list-style-type: none"> ③ 統合ポータル ④ オプトアウト申請システム ⑧ 医療機関向け分析システム ⑪ 統合ネットワーク ⑫ ID管理システム ⑬ 契約管理システム ⑭ サービス管理システム 	<ul style="list-style-type: none"> ● 情報提供依頼事項 <ul style="list-style-type: none"> ✓ 類似システムで採用されたソリューション情報の提示 ● 見積 <ul style="list-style-type: none"> ✓ 類似システムの実績等を元にした、参考金額の提示
<ul style="list-style-type: none"> ② 分散解析・連合学習プラットフォーム 	<ul style="list-style-type: none"> ● R10以降開発想定のため、追加での情報提供をご依頼予定

電子カルテ情報DB（仮称）に対する情報提供依頼対象は4要素となる。適宜「別紙2」を参照すること

3.見積スコープ（対象システムとシステム要件）－ 電子カルテ情報DB（仮称）

見積スコープ

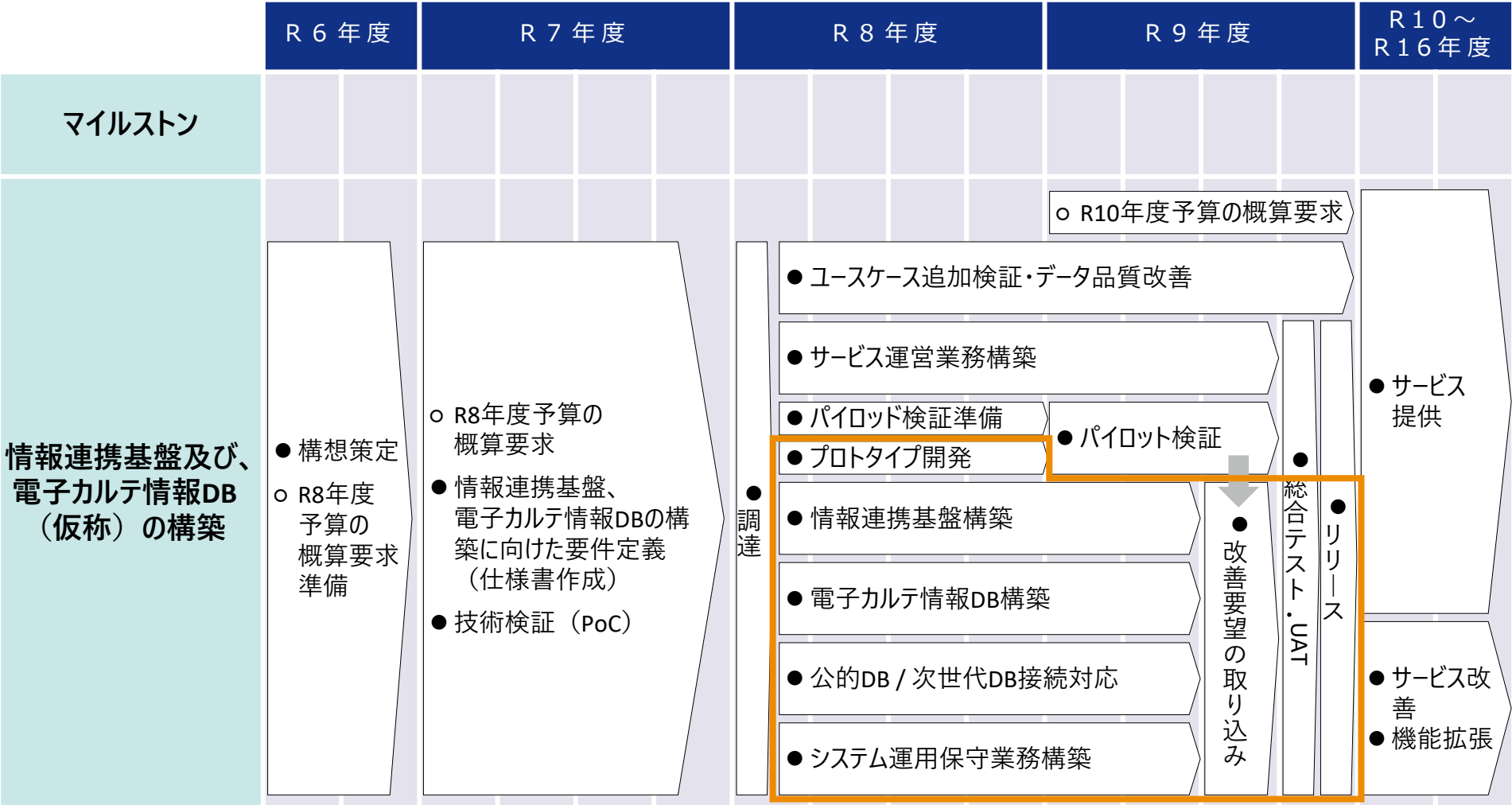


見積スコープは、情報連携基盤及び電子カルテ情報DBの構築～リリース迄の作業に加え、システム運用保守業務の構築と、R10～R16年度の7年間のシステム運用費用が対象

3.見積スコープ（見積対象作業）

○：翌年度に向けて対応する事項
●：当該年度の実施事項

見積依頼スコープ



R8～R9年度に実施を想定する作業の内、見積スコープに含めるべき作業は下記の通り
各作業の実施要件・成果物は、後述する「4. 作業実施内容に関する事項」を参照のこと

3. 見積スコープ（見積対象作業 — R8年度～R9年度作業概要）

見積依頼スコープ

		R 8 年度				R 9 年度			
		1 Q	2 Q	3 Q	4 Q	1 Q	2 Q	3 Q	4 Q
見積スコープ外の 関連アクティビティ		調 達	データソース側接続対応						
			サービス運営業務構築						
			パイロット計画・準備			パイロット実施			トレーニング
1	開発管理	開発計画	開発管理						
2	サービス構築・アプリ開発支援		サービス構築・アプリ開発支援（問い合わせ対応・仕様情報提供）						
3	プロトタイプ開発	要件定義	プロトタイプ開発			パイロット実施支援			
4	情報連携基盤構築	要件定義	基本設計	詳細設計・構築・テスト		結合テスト		改善要望 の取り込み	
5	電子カルテ情報DB構築	要件定義	基本設計	詳細設計・構築・テスト		結合テスト			
6	公的DB接続対応	調査	連携方式設計	詳細設計・構築・テスト		結合テスト			
7	システム運用業務構築	運用要件定義	運用業務設計・構築			結合テスト			
8	総合テスト							計画・準備	総合テスト
9	データ移行	移行計画・設計	移行手順設計 （移行ツール開発）			移行リハ	移行		差分移行
10	システムリリース								リリース計画

本見積依頼書に関する、見積要件（作業実施内容）は下記の通り

4. 作業実施内容に関する事項（1 / 5）

作業実施内容		
1.	開発管理	開発計画作成 <ul style="list-style-type: none"> ● プロジェクト開始後1ヶ月以内に開発計画を策定し、プロジェクト責任者・プロジェクト管理者と合意すること ● 開発計画は、「デジタル・ガバメント推進標準ガイドライン」を遵守して作成し、最低限下記を含めること
		開発管理 <ul style="list-style-type: none"> ● 開発計画で合意した管理ルールに沿って、進捗管理・課題管理・リスク管理し、プロジェクト責任者・プロジェクト管理者に報告を行うこと ● 成果物について、品質管理計画で定めた方法に沿って品質管理し評価すること ● プロジェクトの変更管理／構成管理について、管理ルールに沿って対応すること ● 要件定義・設計が進み、開発ボリュームの精緻な見積りが可能になった段階で、ファンクションポイントやCOCOMO2等の科学的な見積り手法を用いて工数見積もりを行い、精緻な計画策定と予実管理を行うこと
2.	サービス構築・アプリ開発支援 <ul style="list-style-type: none"> ● 情報連携基盤と電子カルテ情報DB（仮称）の開発と並走する、サービス構築プロジェクトやアプリケーション開発プロジェクトの円滑な推進のために、開発中のシステム仕様等の情報提供や、関連プロジェクトからの各種問い合わせ・リクエスト対応業務を行うこと ※ システム開発チームが多岐に分かれることが想定され、円滑なプロジェクト間のコミュニケーションのために窓口となる要員を配置いただくことを想定 	

本見積依頼書に関する、見積要件（作業実施内容）は下記の通り

4. 作業実施内容に関する事項（2 / 5）

			作業実施内容
3.	プロトタイプ開発	要件定義	<ul style="list-style-type: none"> ● 調達仕様書に記載された要求事項について、提案内容を踏まえての実現方式やシステム制約等、実現スコープをプロジェクト内で詳細に合意すること ● 要件定義結果については、関連プロジェクトのレビューを受け内容を合意すること
		プロトタイプ開発	<ul style="list-style-type: none"> ● 情報連携基盤の最初の試作段階として、基本的な機能を実現するためのシステムを作成する
		パイロット実施支援	<ul style="list-style-type: none"> ● 作成したシステムを限定的な環境で実際に運用し、運用状況を確認・評価するフェーズを支援する
4.	情報連携基盤構築	要件定義	<ul style="list-style-type: none"> ● 調達仕様書に記載された要求事項について、提案内容を踏まえての実現方式やシステム制約等、実現スコープをプロジェクト内で詳細に合意すること ● 要件定義結果については、関連プロジェクトのレビューを受け内容を合意すること
		基本設計	<ul style="list-style-type: none"> ● 要件定義内容を元に基本設計を実施する ● 基本設計の終盤にシステム構成要素間の整合を確認する期間を設けること ● ユーザーインターフェース設計内容等、関連プロジェクトに影響する部分についてはレビューを受け内容を合意すること ● <u>OMOP等、標準的なデータモデルの設計</u>
		詳細設計・構築・テスト	<ul style="list-style-type: none"> ● 基本設計の内容を元に詳細設計・構築（開発）・単体テスト・内部連結テストを実施すること
		結合テスト	<ul style="list-style-type: none"> ● 関連するシステムを含めた、システム全体の結合テストを実施すること
5.	電子カルテ情報DB（仮称）構築		※ 同上

本見積依頼書に関する、見積要件（作業実施内容）は下記の通り

4. 作業実施内容に関する事項（3 / 5）

作業実施内容		
6.	公的DB接続 対応	調査
		連携方式設計
		詳細設計・構築・ テスト
		結合テスト

本見積依頼書に関する、見積要件（作業実施内容）は下記の通り

4. 作業実施内容に関する事項（4 / 5）

作業実施内容		
7.	システム運用 業務構築	運用要件定義
		運用業務設計・構築
		運用トレーニング
8.	総合テスト	計画・準備
		総合テスト実施

本見積依頼書に関する、見積要件（作業実施内容）は下記の通り

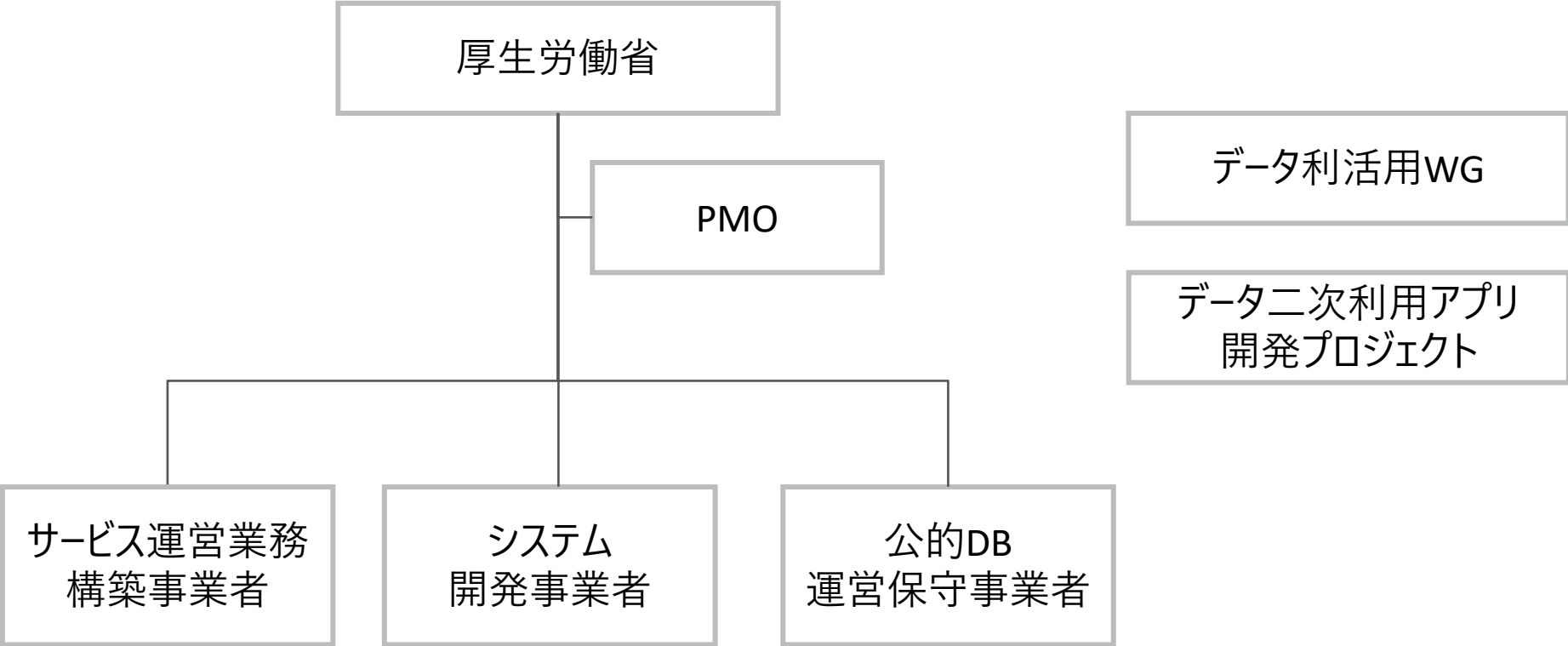
4. 作業実施内容に関する事項（5 / 5）

			作業実施内容
9.	データ移行	移行計画・設計	<ul style="list-style-type: none"> R10から情報連携基盤でデータ分析が可能なように、統合データプラットフォームに各公的DBのデータを移行する 移行計画と移行設計を実施して、各公的DBシステム管理者側と合意する
		移行手順設計	<ul style="list-style-type: none"> 移行計画・移行設計に基づいて、詳細な移行手順を設計する 必要に応じて移行を支援するツールの開発も実施する
		移行リハ	<ul style="list-style-type: none"> 移行手順設計に基づいて実際に移行作業のリハーサルを実施する
		データ移行	<ul style="list-style-type: none"> 移行リハーサルの結果を踏まえ、データ移行作業を実施する
10.	システムリリース	リリース計画	<ul style="list-style-type: none"> サービス開始に向けたシステムリリースの計画を策定すること システムリリース計画については、サービス運営構築プロジェクトと擦り合わせる
		リリース	<ul style="list-style-type: none"> システムリリース計画に沿って、リリース作業を実施すること

想定開発体制

情報連携基盤と電子カルテ情報データベース（仮称）の開発体制は下記を想定している

想定開発体制



4.2.2 要件定義書

目次

1. 業務要件

2. 機能要件

3. 非機能要件

Appendix-1 サービス利用フロー詳細

Appendix-2 ユースケース例

1. 業務要件

1.1. サービス利用ユーザーの種別と目的及び利用イメージ

1.2. サービス運営体制・運用業務一覧






1. 業務要件

1.1. サービス利用ユーザーの種別と目的及び利用イメージ

1.2. サービス運営体制・運用業務一覧

情報連携基盤サービスでは下記のユーザーによる利用を想定している

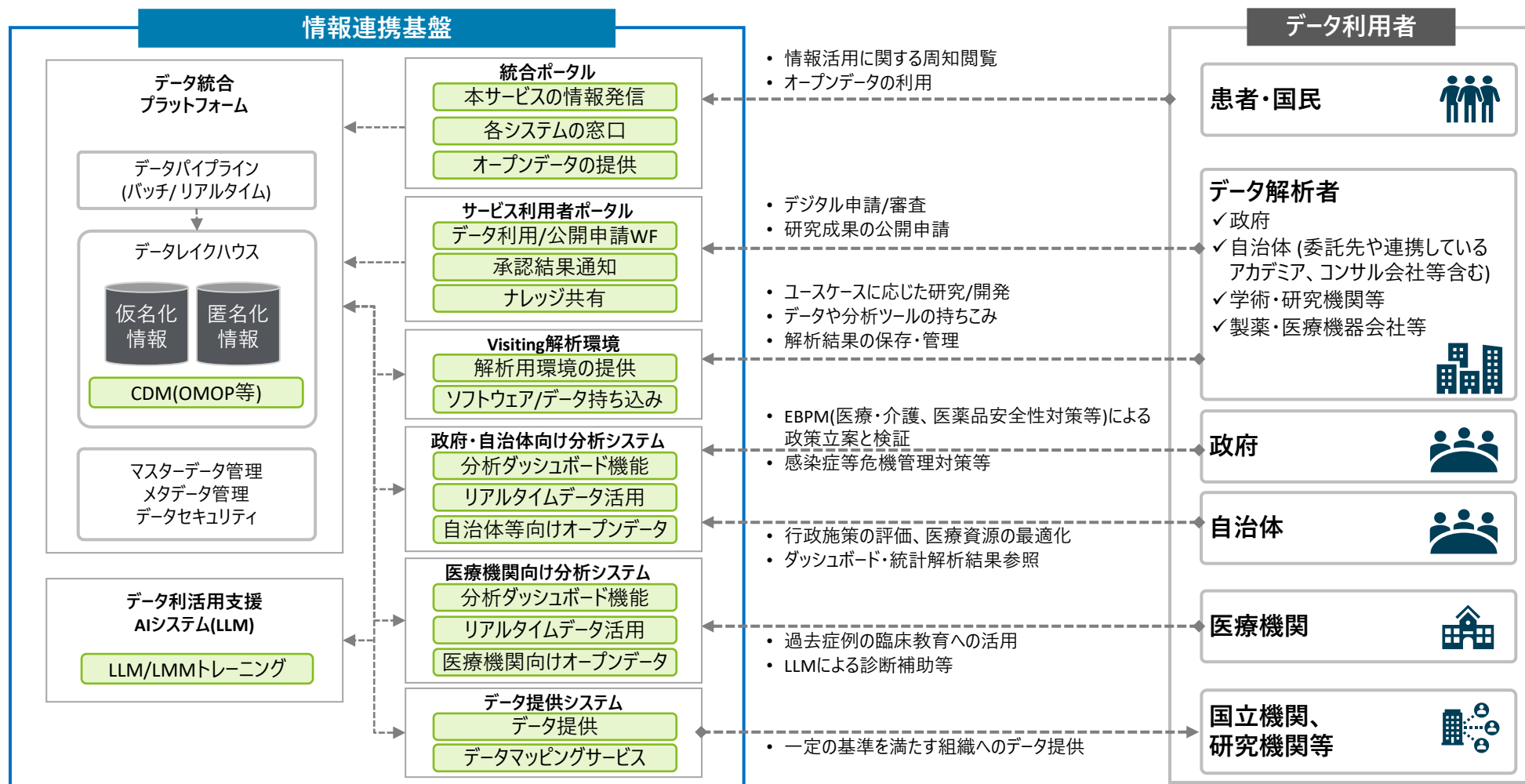
情報連携基盤サービスの利用ユーザーの種別と目的

	利用ユーザー種別	利用目的
サービス 利用者	1  政府・自治体の政策担当者	<ul style="list-style-type: none">✓ EBPM による政策立案と検証・改善✓ 感染予防や流行時の危機管理対策の強化✓ 医療費や予算分配の最適化による財政持続可能性の検討✓ 医薬品の安全対策
	2  学術・研究機関の研究者	<ul style="list-style-type: none">✓ 臨床データの解析による疾病リスク因子、疾患予後予測の解明✓ 長期データ連携による治療・薬剤の影響評価✓ 国際共同研究の推進✓ AIアルゴリズム研究の推進
	3  医療従事者	<ul style="list-style-type: none">✓ 診療や治療方針の向上のための教育としてデータ利用✓ 研修医や医学生の臨床教育やケーススタディ
	4  民間企業 (製薬・医療機器会社等)	<ul style="list-style-type: none">✓ 医薬品の有効性・安全性評価、臨床情報の解析による市販後調査等✓ 医療機器の有効性・安全性評価、臨床情報の解析による市販後調査等、AI搭載機器の機械学習✓ 再生医療等製品の有効性・安全性評価、臨床情報の解析による市販後調査等
	5  患者・教育機関	<ul style="list-style-type: none">✓ 医療・健康に関する統計データ参照✓ 自身の健康管理や病気の予防を目的とした教育コンテンツ✓ 統計学習への活用

情報連携基盤サービスでは、一元のかつ安全なデータ利用・解析環境の提供ならびに各サブシステムを通じて、多岐にわたるサービスを提供する

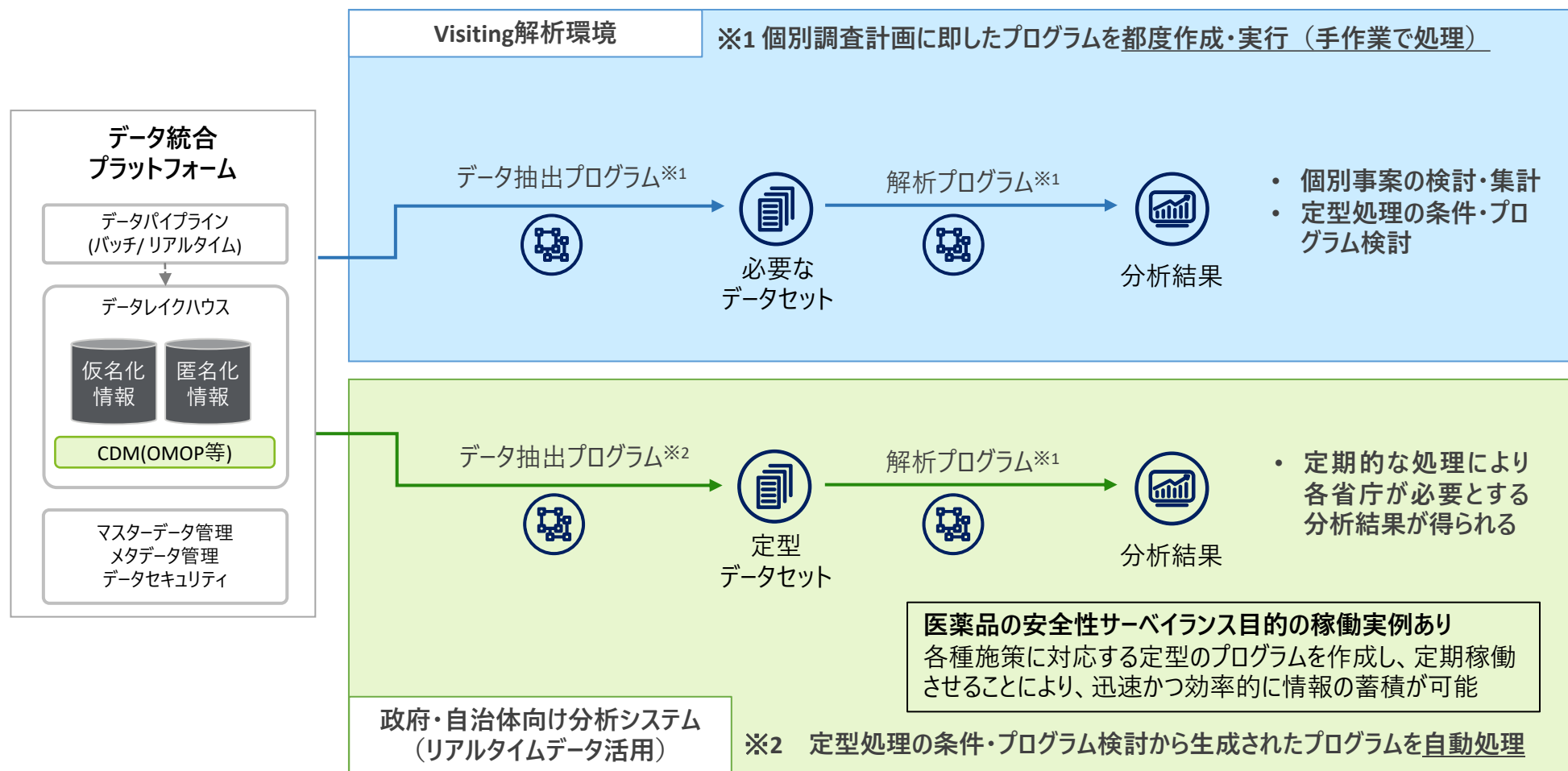
情報連携基盤サービスとユースケースの考え方

- 情報連携基盤では公的DB等にリモートアクセスできるVisiting解析環境の提供を通じて、データの一元的かつ安全な活用を促進する
- また、政府・自治体や医療機関向けシステム等の提供により、地域医療、行政サービス、社会保障分野において多岐にわたるサービスを提供する



政府・自治体の担当者を例とした場合のVisiting解析環境と政府・自治体向け分析システムのユースケースは下記の通り（※詳細なユースケースはAppendix-2を参照のこと）

ユースケース①（例：政府・自治体の政策担当者による分析）

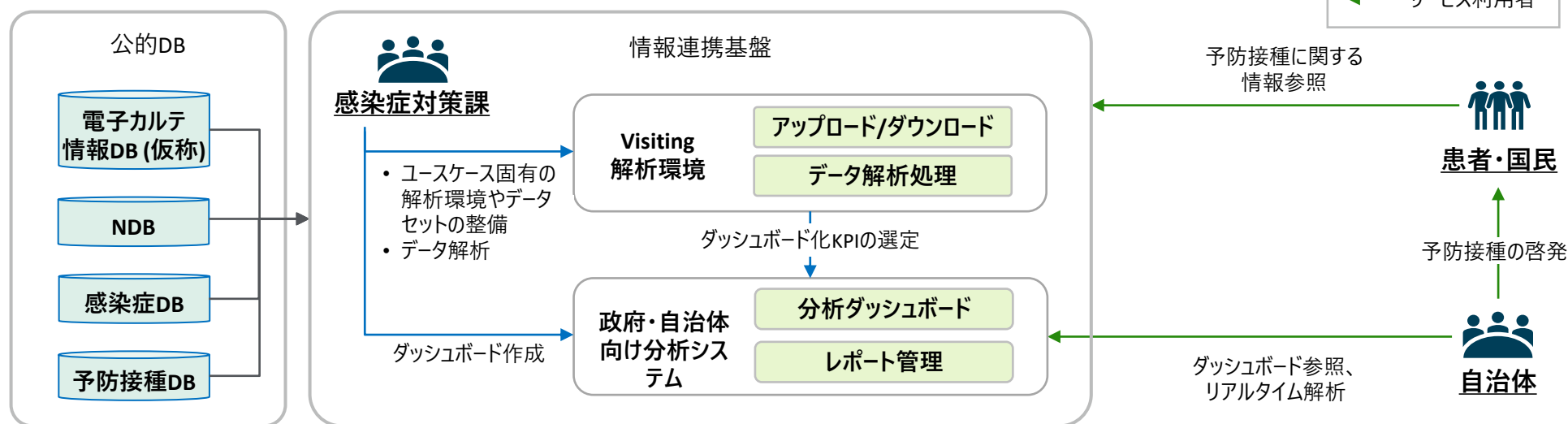


各地域の予防接種DB及び電カル情報DB、NDB、感染症DBとの統合でデータを分析し、ダッシュボードで可視化された情報に参照することで予防接種計画の策定等に役立つ

ユースケース②－１（例：政府・自治体の政策担当者による予防接種の可視化）

ユーザー像	<ul style="list-style-type: none"> 厚生労働省健康・生活衛生局 感染症対策部 予防接種課 	利用するDBとデータ項目 <ul style="list-style-type: none"> 電子カルテ情報DB <ul style="list-style-type: none"> ✓ 6情報：傷病名、アレルギー情報、感染症情報、薬剤禁忌情報、検査情報、処方情報 NDB <ul style="list-style-type: none"> ✓ 傷病名（レセプト病名）、投薬、健診結果等 感染症DB <ul style="list-style-type: none"> ✓ 感染症の名称・症状、診断方法、初診年月日・診断年月日、発病推定年月日等 予防接種DB <ul style="list-style-type: none"> ✓ ワクチン情報、接種場所、副反応の症状等
背景	<ul style="list-style-type: none"> 予防接種に関して国民向けの情報提供ツールとしてホームページが中心的な役割を果たしているが、情報提供のあり方が一元化されていない 現段階では感染症DBにコロナの情報しかなく、今後様々な感染症の情報を追加する予定である 	
目的	各地域の予防接種DBにあるデータを統合し、電子カルテ情報DB、NDB、感染症DBと結合することによって、地域別の予防接種計画の策定等に活用する	

<データフロー>



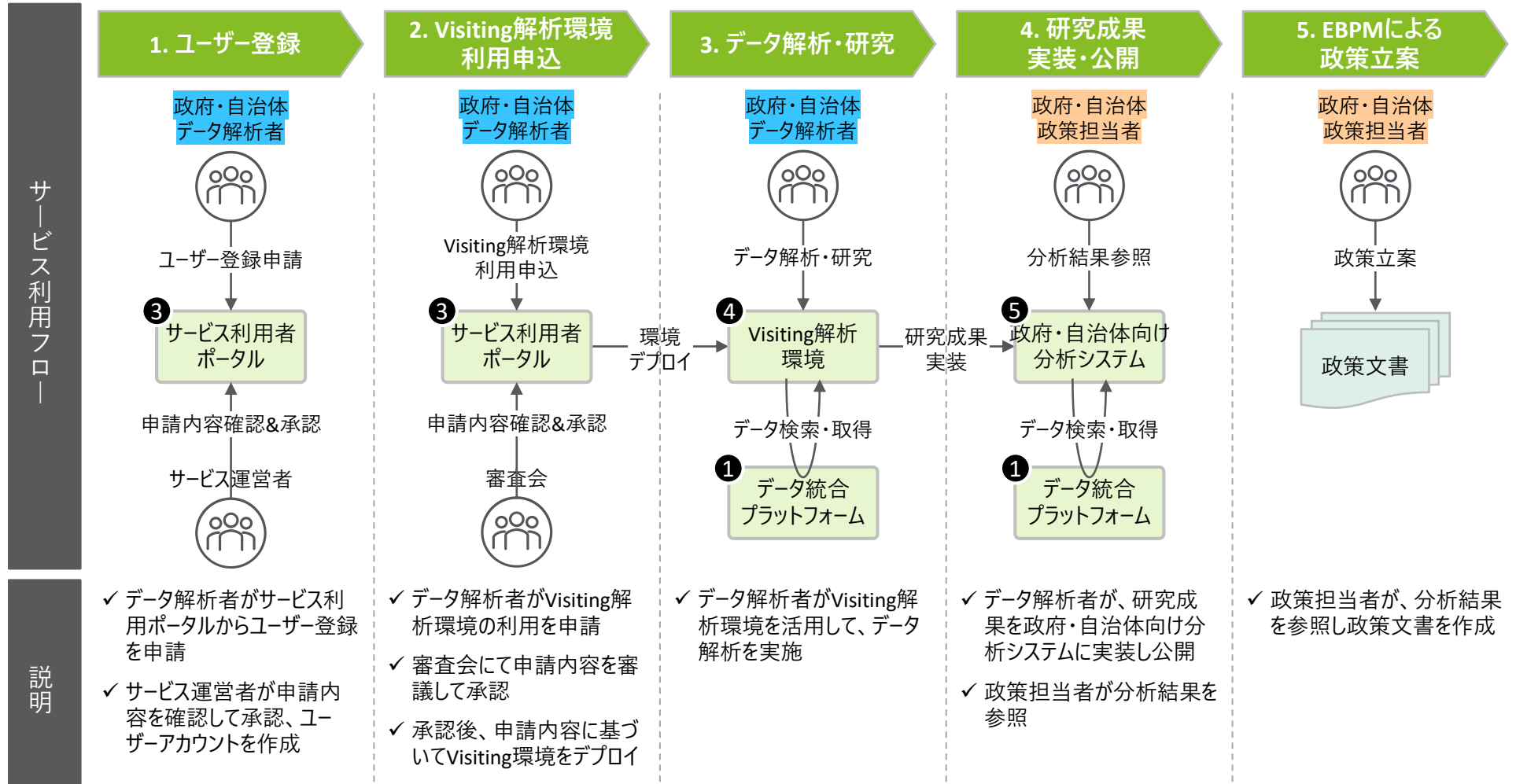
各地域の予防接種DB及び電カル情報DB、NDB、感染症DBとの統合でデータを分析し、ダッシュボードで可視化された情報に参照することで予防接種計画の策定等に役立つ

ユースケース②－２（例：政府・自治体の政策担当者による予防接種の可視化）

ユースケース	概要	利用DB				機能			アウトプット
		NDB	感染症DB	予防接種DB	電カルDB	ダッシュボード	Visiting解析環境	リアルタイム解析	
予防接種の接種意思の分析	地域や年齢層、性別などの属性別に、市民の予防接種に対する意思を分析し、接種を躊躇する要因を特定することで、接種促進のための効果的な施策を立案する			○	○	○	○		<ul style="list-style-type: none"> 接種意思分析レポート 接種促進施策計画書
予防接種の実施状況の把握	地域ごとの予防接種の実施状況をリアルタイムで把握し、接種率の低い地域や年齢層を特定する。これにより、効果的な介入策を講じて接種率を向上させる計画を立案する			○	○	○	○	○	<ul style="list-style-type: none"> 実施状況把握レポート 地域別接種状況マップ 接種計画改善提案書
予防接種のモニタリング及び効果の評価	予防接種後の健康状態を継続的にモニタリングし、接種の効果を評価する。接種後の感染症発生率、重症化率や副作用状況を分析し、予防接種の有効性を定量化する	○	○	○	○	○	○	○	<ul style="list-style-type: none"> 効果評価レポート モニタリングダッシュボード 接種効果評価ガイドライン
予防接種の推奨接種頻度	各種予防接種の最適な接種頻度を検討するため、年齢や健康状態に応じた接種スケジュールをデータに基づいて策定する	○	○	○	○	○	○	○	<ul style="list-style-type: none"> 推奨接種頻度レポート 接種頻度ガイドライン 最適接種スケジュール
予防接種の地域別最適化	地域ごとの感染症発生状況と予防接種率を分析し、地域特性に応じた最適な予防接種計画を策定する	○	○	○	○	○	○	○	<ul style="list-style-type: none"> 地域別予防接種最適化レポート 地域別接種率マップ 予防接種計画書

政府・自治体の担当者を例とした場合のサービス利用フローの概観は下記の通り （※詳細なサービス利用フローはAppendix-1を参照のこと）

サービス利用フロー（例：政府・自治体の政策担当者）



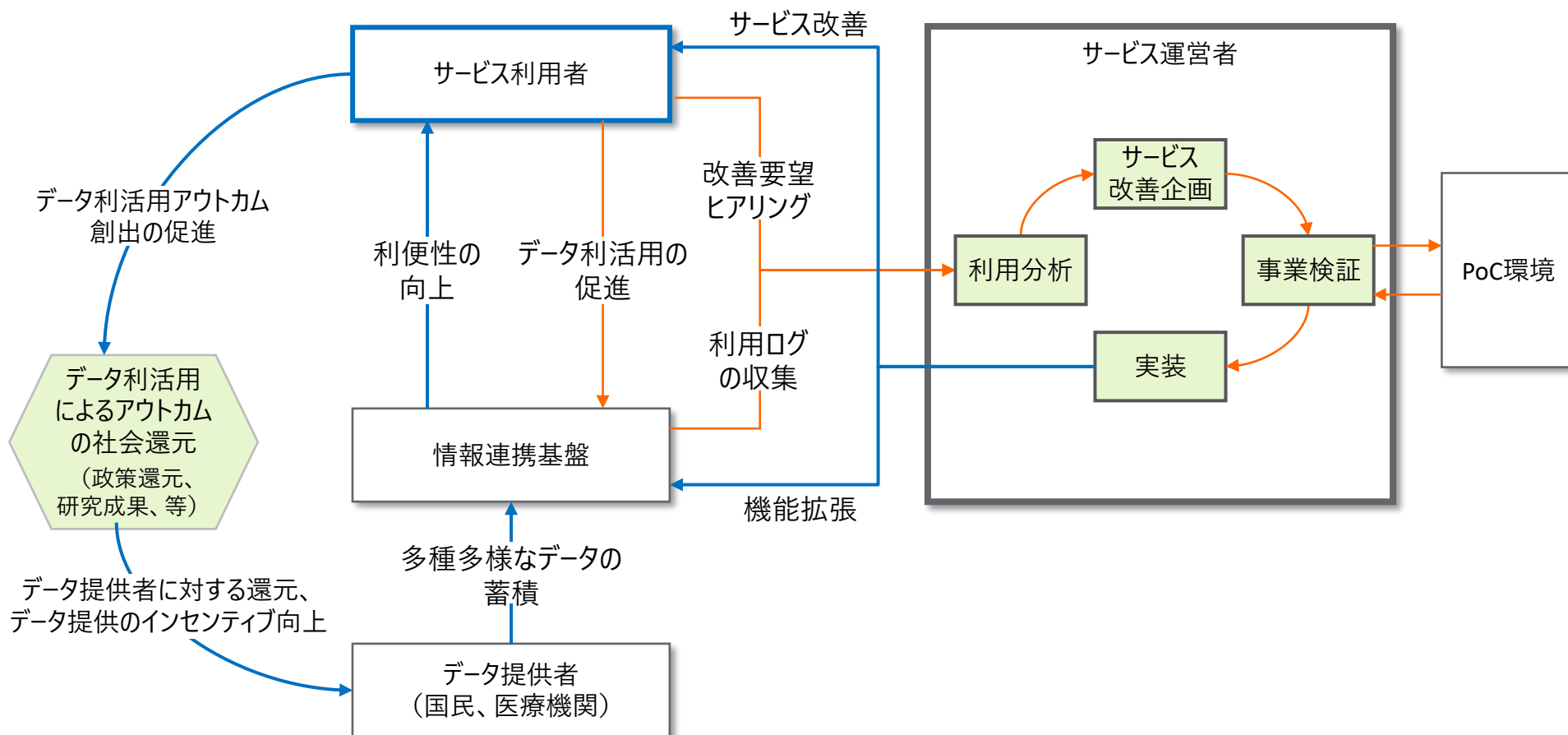
1. 業務要件

1.1. サービス利用ユーザーの種別と目的及び利用イメージ

1.2. サービス運営体制・運営業務一覧

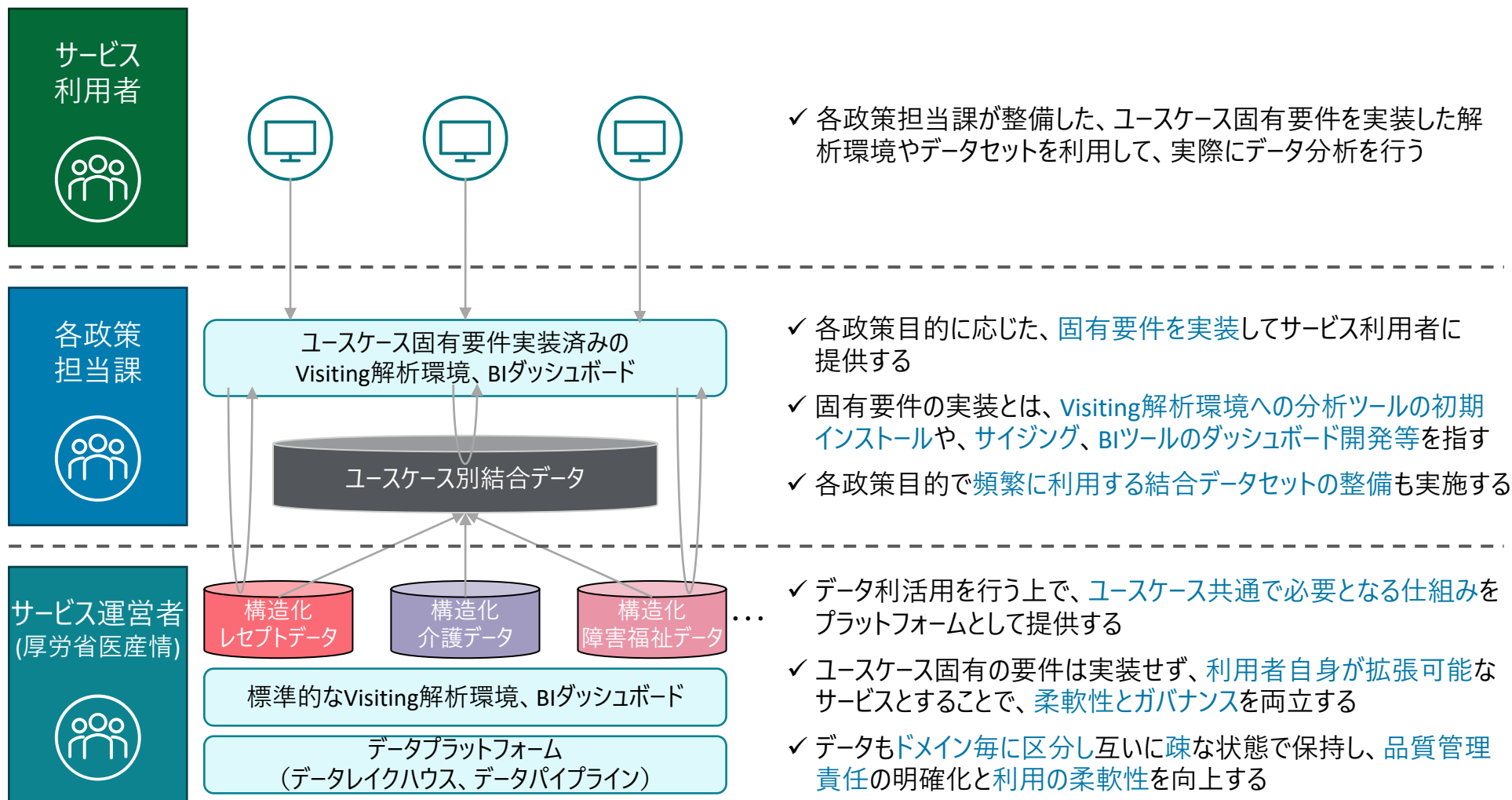
本事業では、サービス利用者の便益を継続的に向上することを主眼に運営を行い、データ活用によるアウトカムの社会還元を促進し、持続的な医療等情報活用社会を構築する

業務の運営方針①－利用者の便益向上を主眼に置いた事業運営モデル



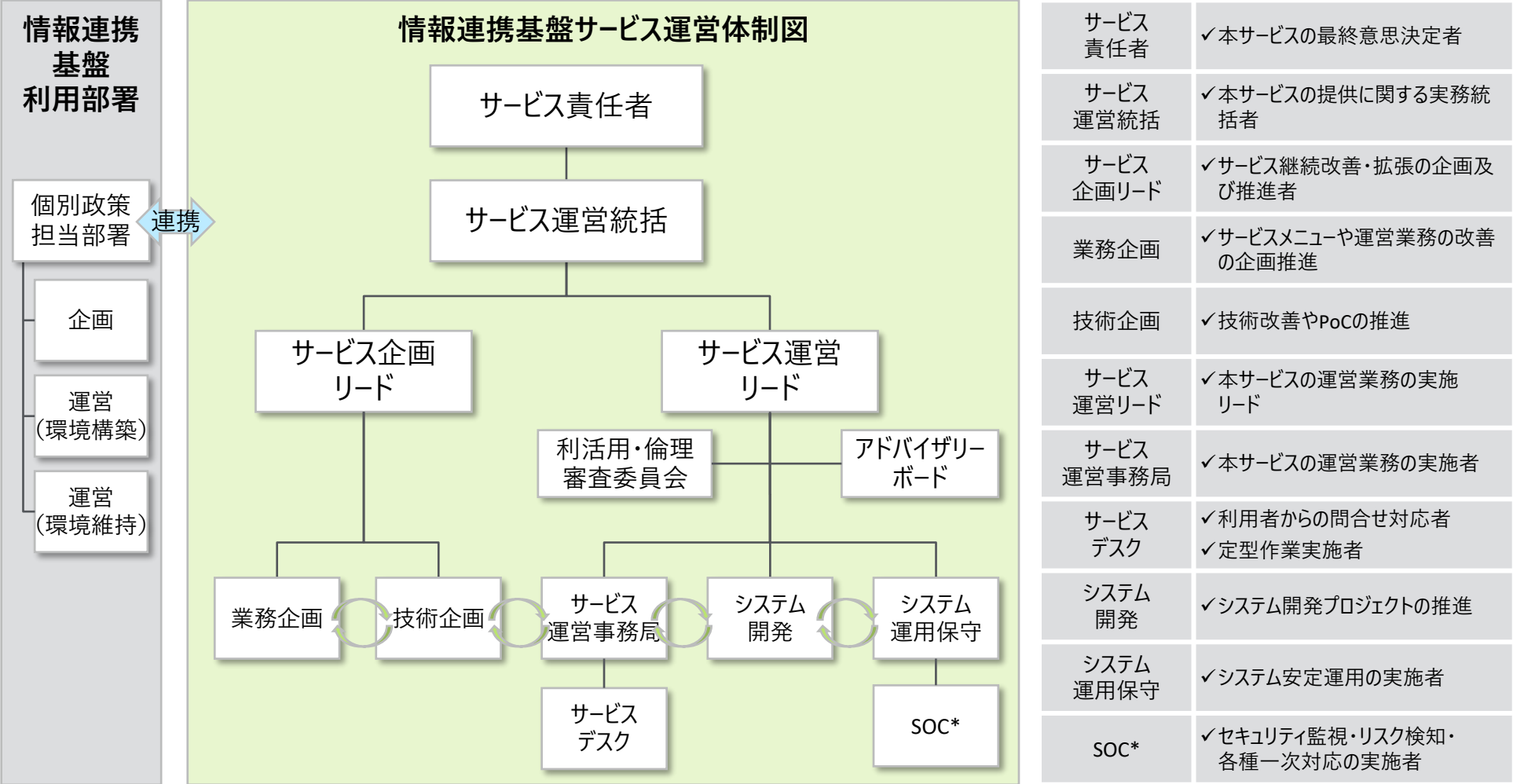
本事業では、利用者と各政策担当課、運営者の協働により、データ利活用におけるガバナンスの確保と、柔軟性&利便性の向上を両立させるスキームを構築する

業務の運営方針②－利用者と運営者の協働によるガバナンスと柔軟性の両立



本サービス事業の運営体制は下記11種類の役割で構成する想定である
企画・開発・運営・運用のチームがシームレスに連携し、迅速な改善を可能な体制を構築する

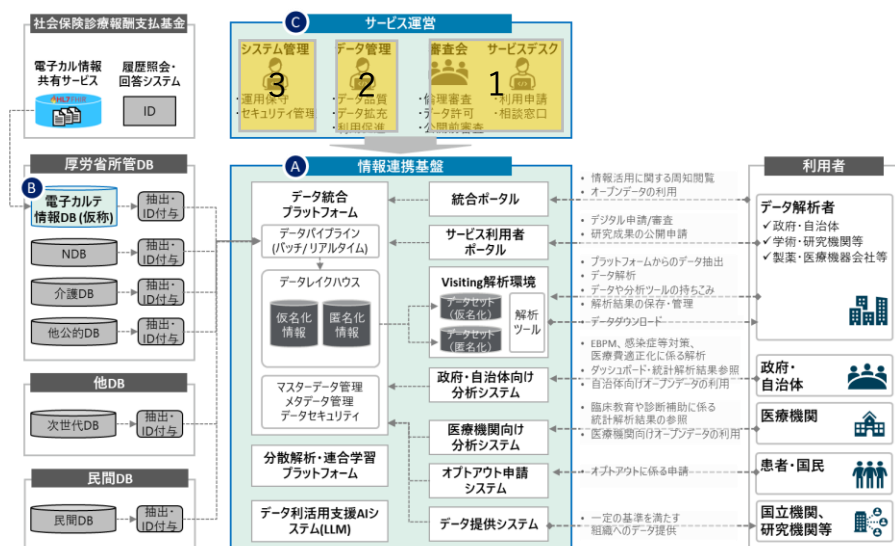
想定する本サービス事業の運営体制



* SOC: Security Operation Center

情報連携基盤サービスの運営に必要な業務を下記の3カテゴリに分けて考える

情報連携基盤サービスにおける業務の構造化と各業務の概要



業務カテゴリ

実施業務

1. サービス提供

データ利用者に対してサービス提供やナレッジの展開

- 【1-1】相談受付業務
- 【1-2】利用者ポータル対応業務(申請・審査等)
- 【1-3】データ提供サービス業務(定型データ等)
- 【1-4】データ提供サービス業務(特別抽出、データ突合等)
- 【1-5】患者ポータル対応業務(オプトアウト対応を含む)
- 【1-6】データ活用ガイドライン(教育コンテンツ等)策定
- 【1-7】データ・ツール等の持込/持出に係る対応業務
- 【1-8】継続的サービス改善

2. データ管理

データの品質管理やアクセス管理、セキュリティ等

- 【2-1】データ設計・構造管理
- 【2-2】バリデーションチェック・クレンジング業務
- 【2-3】マスターおよび標準用語作成・管理
- 【2-4】データカタログ作成
- 【2-5】セキュリティ・アクセス管理
- 【2-6】ルール違反に対する罰則等の対応

3. システム管理

サービス提供、データマネジメントで必要となる仕組みの提供

- 【3-1】情報連携基盤の運営管理
- 【3-2】各種サービス/ツールの運営管理
- 【3-3】継続的システム(LLM等)改善

前述を踏まえて、情報連携基盤サービスの運営に必要な業務は以下の通り

情報連携基盤サービスの運営業務一覧

業務カテゴリ	業務プロセス名	業務説明
1. サービス提供	1-1. 相談受付業務	・ 情報連携基盤の利用に関する問い合わせやデータ利用に係る相談の受付等をサポートする（質問対応、使用方法の説明、技術的な問題解決などを含む）
	1-2. 利用者ポータル対応業務(申請・審査等)	・ データ利用者の申請や審査の管理を行う。申請内容の確認、審査結果の通知、申請状況の管理など、効率的なサポートを実施する
	1-3. データ提供サービス業務(定型データ等)	・ 利用許可の下りた案件に対して、定型データを抽出・加工し、適切な形式で提供する
	1-4. データ提供サービス業務(特別抽出、データ突合等)	・ 特別抽出やデータ利用者からのデータ突合申請等などの特別な対応
	1-5. 患者ポータル対応業務	・ 患者や国民がアクセスできるポータルを運営する ・ オプトアウトに係る情報および申請窓口 ・ データ二次利用の意義や安全性等の適切な情報提供を行う
	1-6. データ活用ガイドライン策定	・ データを適切かつ有効に利用するための教育コンテンツやガイドラインを策定する
	1-7. データ・ツール等の持込/持出に係る対応業務	・ データ利用者からの要望に応じて、データやツール等の持込/持出に係るチェックや審査を行う
	1-8. 継続的サービス改善	・ データ提供サービスにおける継続的なサービス改善を図る
2. データ管理	2-1. データ設計・構造管理	・ 情報連携基盤におけるデータの構造設計および管理を行う（データモデルの作成や CDM(OMOP等)対応などを含む）
	2-2. バリデーションチェック・クレンジング業務	・ データの品質を確保するために、データの整合性や完全性を検証し、誤りや重複データの修正・削除やデータのコーディングなどを行う
	2-3. マスターおよび標準用語作成・管理	・ マスターおよび標準用語を管理し、検索・照会を可能にする
	2-4. データカタログ作成	・ データレイクに格納されているデータから提供可能なデータセットの概要やメタデータを記載したデータカタログを作成する
	2-5. セキュリティ・アクセス管理	・ ユーザ認証、Visiting 解析環境等の監視など安全管理措置を実施する ・ 情報連携基盤へのデータアクセス権限を管理
	2-6. ルール違反に対する罰則等の対応	・ 情報連携基盤の運営にあたり、データ利用者等のルール違反に対する対応を行う
3. システム管理	3-1. 情報連携基盤の運営管理	・ 情報連携基盤全体の運営管理を行う（システム運用、障害対応、システムパフォーマンスの最適化など）
	3-2. 各種サービス/ツールの運営管理	・ 情報連携基盤で利用される各種ツールやサービス（データ提供ツール、ポータルサイトなど）の運営管理を行う
	3-3. 継続的システム（LLM等）改善	・ データ利用者の利用状況や要望に応じて、LLM等のシステムの継続的な改善を図る

2. 機能要件

- 2.1. 機能に関する事項（情報連携基盤）
- 2.2. 機能に関する事項（電子カルテ情報DB（仮称））
- 2.3. データに関する事項
- 2.4. インターフェースに関する事項

2. 機能要件

2.1. 機能に関する事項（情報連携基盤）

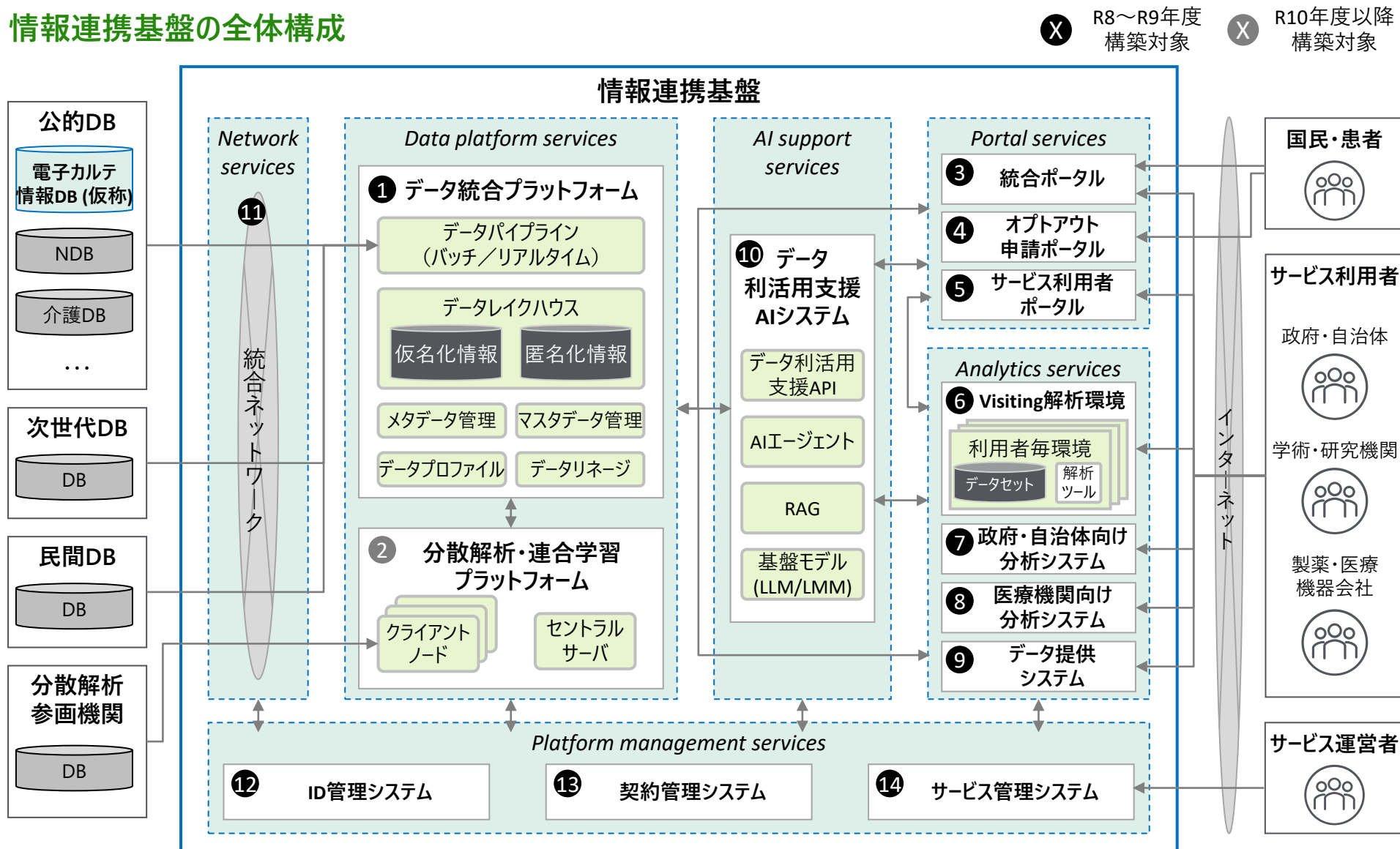
2.2. 機能に関する事項（電子カルテ情報DB）

2.3. データに関する事項

2.4. インターフェースに関する事項

情報連携基盤は下記14の要素で構成される。②の分散解析・連合学習プラットフォームはR10年度以降の開発を想定している

情報連携基盤の全体構成



情報連携基盤の各システムは以下の機能概要を想定している

情報連携基盤を構成する各システムの概要（1 / 2）

No	システム名称	機能概要
①	データ統合プラットフォーム	<ul style="list-style-type: none"> データ統合プラットフォームは、各データソースから仮名化データと匿名化データをリアルタイム／バッチで収集し、データレイクハウスに蓄積する データレイクハウスに蓄積されたデータは、適切なアクセス制御機構を経てデータ利用システムからアクセス可能となる また、データマネジメントに必要な各種機能を具備し、API経由でサービス運営者とサービス利用者からアクセス可能な構成を取る
②	分散解析・連合学習プラットフォーム	<ul style="list-style-type: none"> 分散解析及び連合学習は、データを中央に集約することなく、複数機関で保持するデータを活用して分析や機械学習モデルの訓練を行う手法 参画機関は、データを中央に集約する必要が無いため、個人情報提供に関する同意取得への対応や、自機関のデータ資産の流出対策等が不要となり、複数機関横断での研究や分析、機械学習が行いやすくなるというメリットがある 一方で、解析の精度を担保するためには、参画機関のデータを標準化するためがあるため、CDM（Common Data Model）の導入が必要となる システムとしては、クライアントノードと言う、参画機関毎に分離されたプライベートなDBと計算リソースをクラウド上に構築する機能も提供する
③	統合ポータル	<ul style="list-style-type: none"> 本システムは、サービス利用者、サービス利用を検討者、及び国民全般に広く本サービスに関する情報を発信するためのシステムである 情報発信内容としては、本サービスの説明、データ利活用方針、サービス内容、メンテナンス情報、サービス利用状況、データ種別や件数、セキュリティ対策状況、今後の計画と多岐に渡る また、各システムの窓口システムとしてのWEBリンク機能や、オープンデータの提供機能も提供する
④	オプトアウト申請ポータル	<ul style="list-style-type: none"> 本システムは、患者・国民等が自身のデータ提供に対して、不同意である旨を入力および申請することができる機能を提供する 申請の正当性を担保するために申請者の本人確認（認証）が必要となるが、国民を一意に認証できる仕組みをどのように実現するか、今後検討が必要となる（図ではマイナポータルとの連携のイメージで記載） 申請内容をサービス運営者が確認した後、①データ統合プラットフォームでオプトアウト対象データを識別できるように、オプトアウト情報の連携を行う機能を具備する
⑤	サービス利用者ポータル	<ul style="list-style-type: none"> 本システムは、利用者からの各種リクエスト窓口の一元化とワークフローのデジタル化による迅速なサービス提供を実現するシステムである 本システム上の各種承認結果をVisiting解析環境や、データ提供システムに連携し、データ利活用におけるガバナンスを担保する また、データ利活用の促進のために、研究成果の公開や、データ分析に関するナレッジ提供、データカタログ等の機能を具備する
⑥	Visiting解析環境	<ul style="list-style-type: none"> Visiting解析環境は、サービス利用者に対してプライベートな解析用の環境を提供するサービスである データ解析者はDaaS（Desktop as a Service）により提供される仮想デスクトップにアクセスし、解析作業を実施する またサービス運営者が研究プロジェクト毎に多数作られるVisiting解析環境を効率良く管理するための機能を提供する
⑦	政府・自治体向け分析システム	<ul style="list-style-type: none"> 政府・自治体の各政策担当がEBPMや政策立案・評価に必要なデータ分析結果をダッシュボードとして作成し、参照するシステムである ダッシュボードのみならず、EBPMや政策立案・評価に必要なデータ分析レポートや研究論文、調査報告書等を管理する機能を有する また、コンシェルジュ機能を具備し、AIによる必要情報の収集や文章生成の支援を行い業務効率化を実現する

情報連携基盤の各システムは以下の機能概要を想定している

情報連携基盤を構成する各システムの概要（2 / 2）

No	システム名称	機能概要
⑧	医療機関向け分析システム	<ul style="list-style-type: none"> サービス利用者である医療機関等が過去の症例データにアクセスし、疾患動向や類似症例の傾向を把握することで診断を支援する 臨床や医療統計、疫学管理を通じて現実に即したケーススタディや教育を行い、医療従事者の知識向上に寄与するシステムである 基本的に具備する機能は、「⑤ 政府・自治体向け分析システム」と同様である
⑨	データ提供システム	<ul style="list-style-type: none"> 本システムは、国立研究機関等の事前に承諾が得られた機関／法人が有するシステム環境にデータを提供する仕組みである 今後、具体的な連携方式等が議論される予定だが、現時点ではAPIによるリクエストドリブン配信方式と定期配信方式を想定している また、サービス利用者ポータルと連携し、事前承認されたデータのみが提供されるよう管理する仕組みを具備する
⑩	データ利活用支援AIシステム	<ul style="list-style-type: none"> サービス利用者ポータルにおけるコンシェルジュ機能や、サービス運営者の運営業務の効率化を実現するなど、LLMを活用した支援機能を提供する仕組みである 情報連携基盤の様々なシーンでLLMを活用した業務効率化／ユーザビリティ向上を実現するため、独立したシステムとして設計としている
⑪	統合ネットワーク	<ul style="list-style-type: none"> 情報連携基盤は、公的DBに加え、次世代DB、民間DBとの接続を想定しており、各接続先はそれぞれ別のクラウドサービスやオンプレミスのデータセンターで稼働している。従って、各データソースとセキュアにシームレスに接続するための連携用ネットワークを構築する必要がある 情報連携基盤を構成するシステムも、様々なクラウド上に実装される可能性があり、クラウド間を接続する閉域ネットワークが必要となる また、Internetからのユーザーアクセス通信や、Internetへのアウトバウンド通信をセキュアに行うための仕組みも必要となる
⑫	ID管理システム	<ul style="list-style-type: none"> 情報連携基盤が提供する各サービス／システムにログインするユーザーのIDとパスワード、ユーザー情報、グループを一元管理するサービス シングルサインオン機能を提供し、サービス利用ユーザ、運営者はID管理サービスに認証を行えば各システムの個別ログインが不要となる 基本的にはIDaaS（クラウド型ID管理サービス）での実装を想定している
⑬	契約管理システム	<ul style="list-style-type: none"> サービス利用者の利用しているサービスに対して利用料課金を行うことを目指し、サービス利用者の契約情報を管理する機能を具備する サービス運営者が、契約管理に必要なサービスメニューや課金プランを作成して管理する機能を具備する 本システムで作成したサービスメニューや課金プランをサービス利用者ポータルに連携し、ユーザーがメニューより利用申請が行えるようにする
⑭	サービス管理システム	<ul style="list-style-type: none"> サービス管理システムは、本サービスを安定的にかつ効率良く運営するための機能を提供する サービス運営者はインターネット経由でサービス管理用DaaSを用いて、各種サービス管理のための機能にアクセスする 同様にセキュリティパッチ等の資材持込や、本番環境ログの持出等のために、サービス管理用のAirlockも具備する

利用者便益の継続的な向上のために、情報連携基盤は高い拡張性を備えることを重視し、迅速かつ低コストで機能拡充が可能なアーキテクチャを採用する

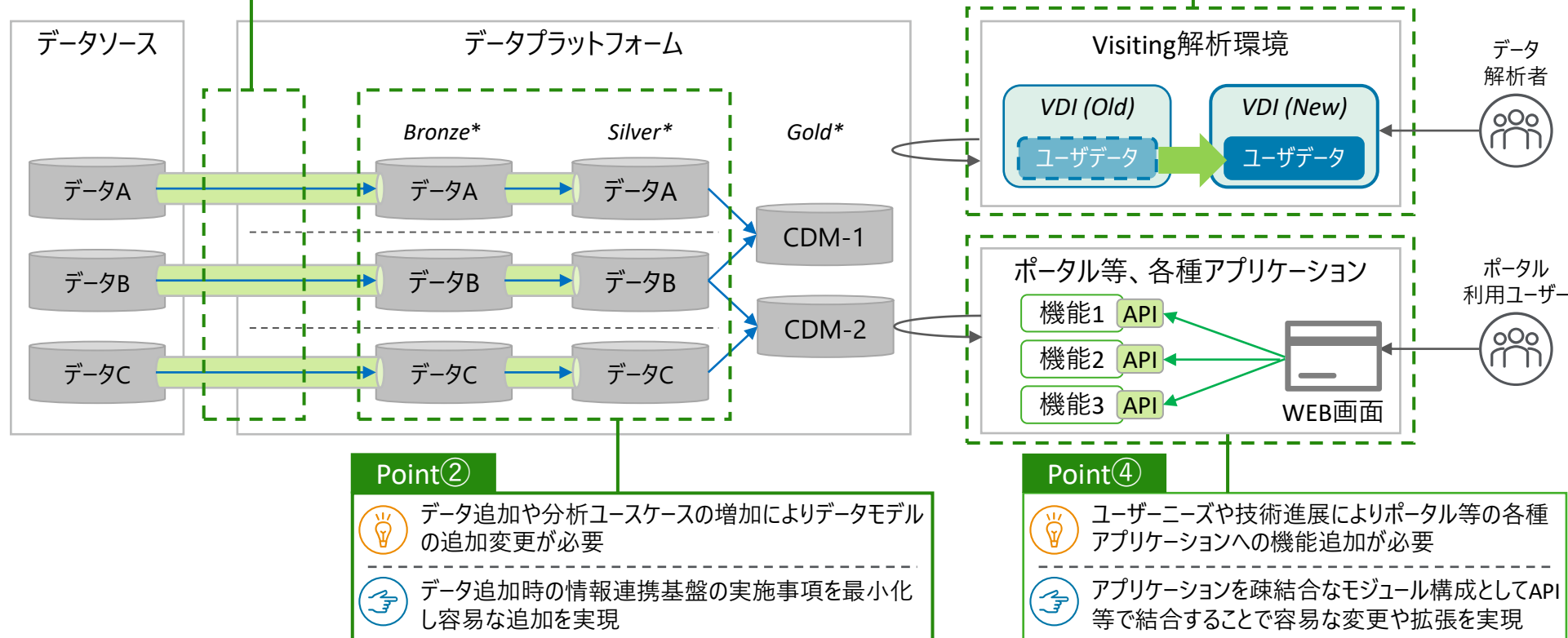
システムの実現方針① – 高い拡張性（拡張容易性）を実現するアーキテクチャ

Point①

- 💡 データ追加によるデータパイプラインの追加が必要
- 👉 パイプラインの処理をパターン化して、部品化を進め開発生産性を向上

Point③

- 💡 ユーザーニーズや技術進展によりVisiting解析環境で利用可能なアプリケーションの拡充が必要
- 👉 Visiting解析環境をImmutableインフラストラクチャとして運用することで容易な拡張を実現



Point②

- 💡 データ追加や分析ユースケースの増加によりデータモデルの追加変更が必要
- 👉 データ追加時の情報連携基盤の実施事項を最小化し容易な追加を実現

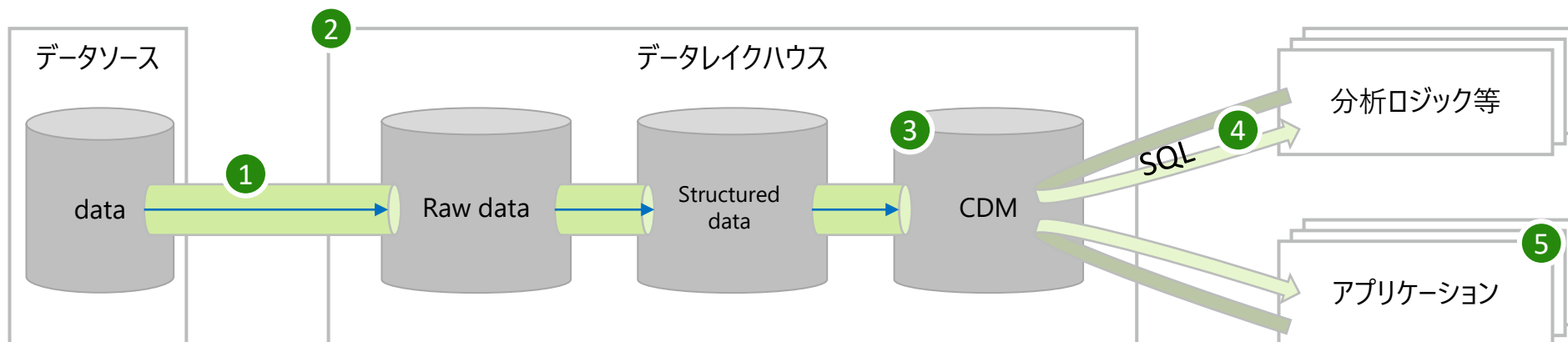
Point④

- 💡 ユーザーニーズや技術進展によりポータル等の各種アプリケーションへの機能追加が必要
- 👉 アプリケーションを疎結合なモジュール構成としてAPI等で結合することで容易な変更や拡張を実現

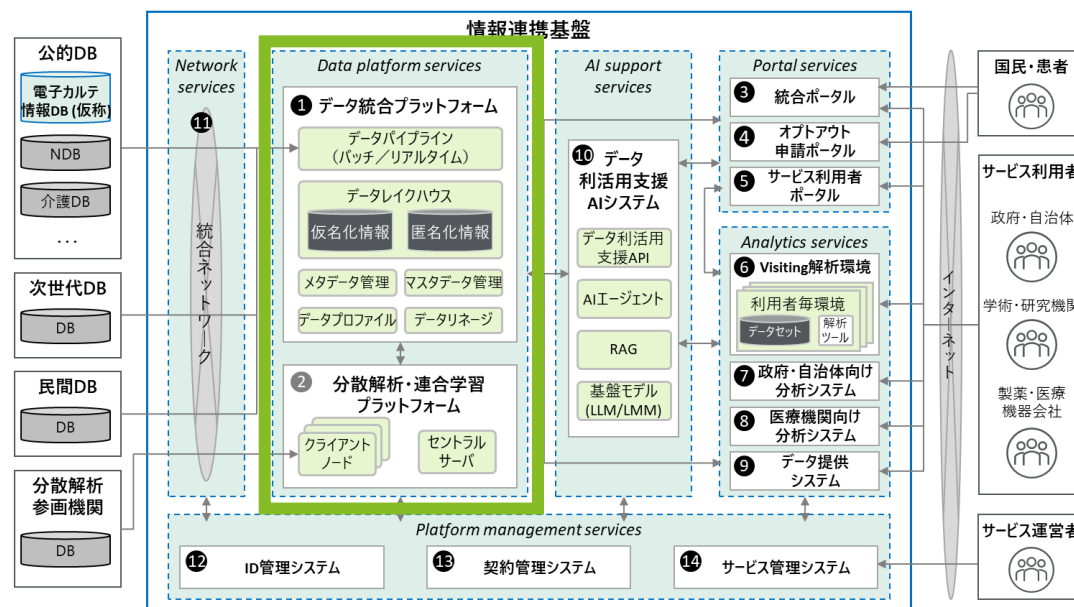
技術ロックイン防止の観点から、「移行性」を考慮した技術・製品選定を行う

システムの実現方針②－技術ロックイン防止を考慮した技術・製品選定

- ① データパイプラインに実装されるデータ処理ナレッジが固有の製品にロックインされないように、広く浸透しているプログラミング言語で実装可能な製品を採用する
- ② データレイクハウスに蓄積される大量データが、特定の製品にロックインされないように、データ移行課金を踏まえたクラウドサービス選定や、**Open Table Format**等のデータ移行性に配慮した製品選定を行う
- ③ データ分析やアプリケーションが直接利用するデータについては、可能な限り特定製品に固有ではないデータモデルとし、データレイクハウス製品変更時のアプリケーション影響を最小化し、移行性を向上する
- ④ データ分析やアプリケーションがデータアクセスする際は、極力標準規格に沿ったSQLを利用し、データレイクハウス製品変更時のアプリケーション影響を最小化し、移行性を向上する
- ⑤ アプリケーションは疎結合なモジュール構造を取り、API等のオープンなIF仕様で連携することで機能追加や変更容易性を保ち、技術ロックインを排除する。SaaSを採用する場合も、**SaaS上に過度な作り込みを行わず**移行性を考慮した実装を行う

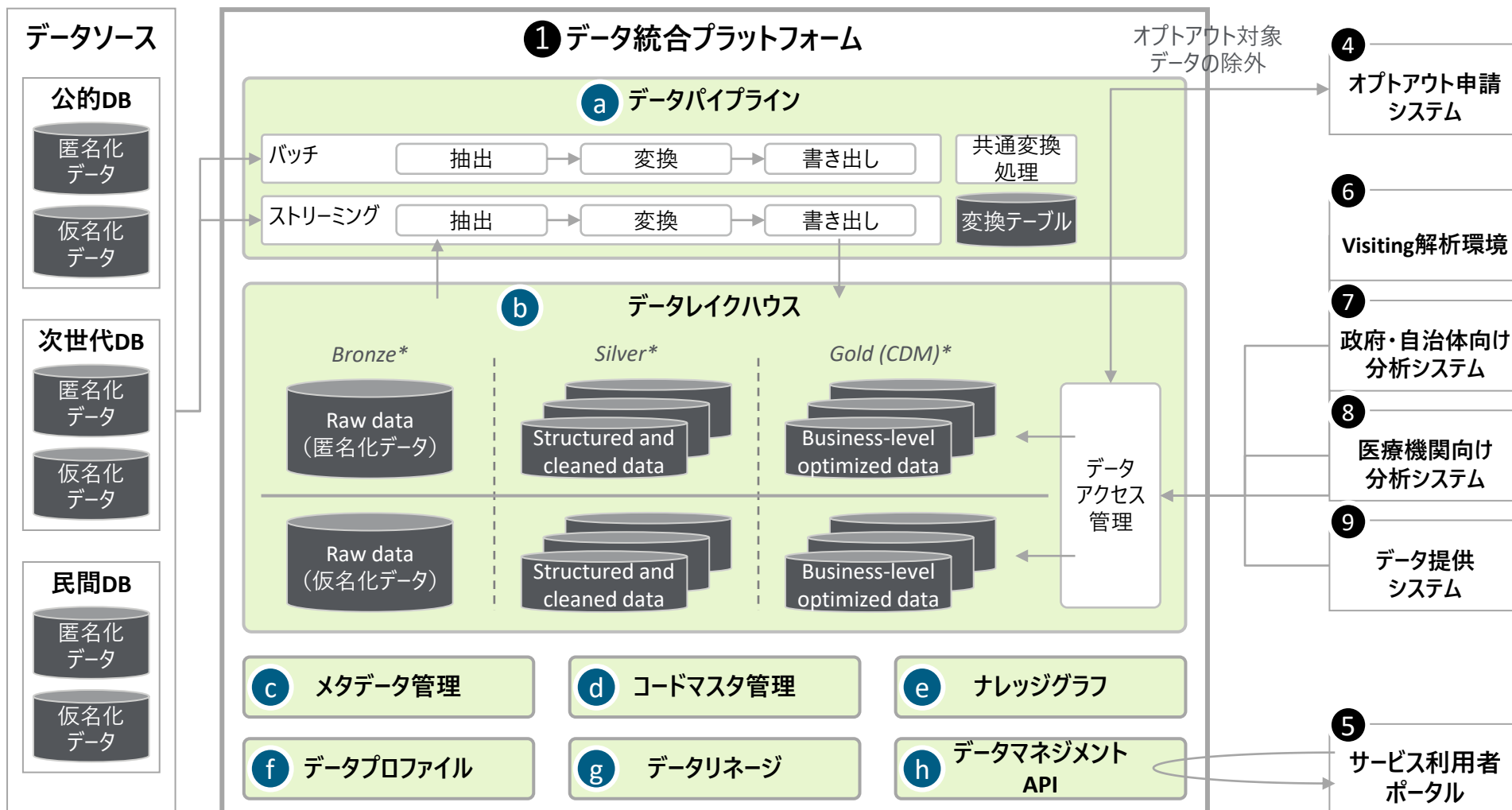


Data platform services



① データ統合プラットフォーム — 概要構成

- データ統合プラットフォームは、各データソースから仮名化データと匿名化データをリアルタイム／バッチで収集し、データレイクハウスに蓄積する
- データレイクハウスに蓄積されたデータは、適切なアクセス制御機構を経てデータ利用システムからアクセス可能となる
- また、データマネジメントに必要な各種機能を具備し、API経由でサービス運営者とサービス利用者からアクセス可能な構成を取る



* Medallion Architecture (<https://dataengineering.wiki/Concepts/Medallion+Architecture>)

① データ統合プラットフォーム — 機能概要説明 (1/2)

機能概要

a. データパイプライン

- データソースからデータを収集し、変換し、データレイクハウスに格納する機能
- 定期的にバッチ処理でデータを収集するパイプラインと、リアルタイムに（短い時間間隔で）データを収集するリアルタイムパイプラインの2種類を具備して、それぞれに必要な機能を具備する
- データ変換処理については、変換処理・バリデーション処理・クレンジング処理が効率的に行えるなど、ユースケースに関する変換処理をより効率的に行える仕組みや機能を有することが望ましい
- またパイプライン処理を部品化し、再利用可能にすることでパイプライン開発の生産性を向上可能なことも必要である

b. データレイクハウス

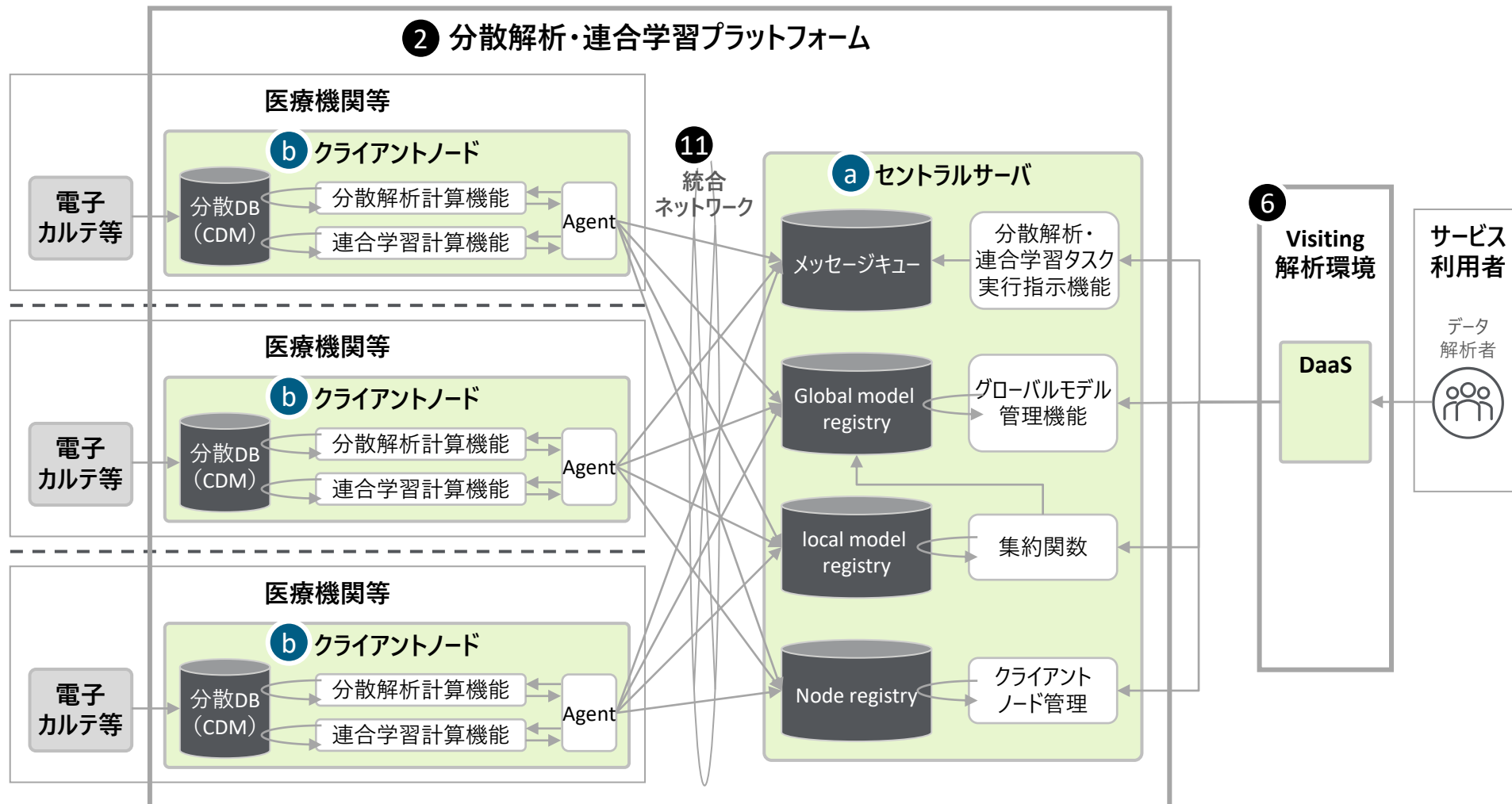
- データレイク（DL）とデータウェアハウス（DWH）それぞれの強みを活かし、データを統合的に扱うことを可能とするアーキテクチャ
- 統合的とは、DL→DWHのデータ移動を行わずとも、DLとDWHに対して透過的にSQL検索が行えることや、構造化データと非構造化データを分断せずに利用できる点等の特徴を指す。またデータの二重持ちを削減してストレージコストを低減させることも目指す
- 2024年時点でソリューションとしては、データレイクとDWHの連携を強めたものや、データレイクハウスとして完全に統合されたソリューションが出ているが、概念に捕らわれずにデータ利活用面・性能面・管理面・コスト面で最適なものを選択する必要がある
- データ構造としては、メダリオンアーキテクチャ（<https://dataengineering.wiki/Concepts/Medallion+Architecture>）を採用して、データをBronze、Silver、Goldにレベル分けして保持することで、データの利活用を柔軟にする
 - Bronze:**
 - ✓ データソースの構造を維持し、最低限の変換を行ったデータを保持する。ローデータとしてSilverデータを再生成可能とする
 - Silver:**
 - ✓ Bronzeデータをクレンジング・構造化・バリデーション・不要情報削除・テーブル統合を行い汎用的に利用できるデータを提供
 - ✓ ドメインを跨いだデータ統合は行わず、データを疎な状態に保つことで汎用性を維持する
 - Gold:**
 - ✓ ドメインを跨いだデータ統合も行い、分析ユースケースに最適化されたデータを提供する
 - ✓ OMOP等のCDMへの対応はGoldレベルと位置付ける
- データアクセスコントロール機能を有し、事前承認を得たデータのみアクセス可能なように制御する
- また、オプトアウト申請ポータルと連携し、オプトアウト対象データを除外する機能を具備する

① データ統合プラットフォーム — 機能概要説明 (2/2)

	機能概要
c. メタデータ管理	<ul style="list-style-type: none"> データを管理するためのデータであるメタデータを管理する機能。ビジネスメタデータ、テクニカルメタデータ、オペレーショナルメタデータ等を管理する想定 管理機能としては、メタデータの項目追加・変更・削除に加え、変更履歴の保存や差分表示機能を具備する ヘルスケア業界に関わらず、メタデータの管理項目に関する共通仕様が定まっていないため、今後検討が必要
d. コードマスタ管理	<ul style="list-style-type: none"> データの一貫性や結合を実現するために必要な、コードマスタを管理する機能 管理機能としては、管理対象マスタの追加・変更・削除に加え、変更履歴の保存や差分表示機能を具備する コードマスタとして何を管理するべきか、本プラットフォームとしてどのコード体系を採用するか今後検討が必要
e. ナレッジグラフ	<ul style="list-style-type: none"> オントロジー（データに対する前提知識）をグラフ型構造で保持する機能 データの臨床上／ビジネス上の“意味”を管理する用語辞書としての役割と、用語同士の関係性を管理することが可能となる データ分析者がデータを検索・解析する際や、生成AIによるデータ利活用支援で用いることを想定している
f. データプロフィール	<ul style="list-style-type: none"> データ内容や品質に関する傾向を分析・評価・可視化する機能。データ利用者がクレンジングを検討する際の情報として利用する 各項目値の欠損率、値のバラつき、値種別、数値フィールドの最大値／最小値・偏差、データ間の関連性を分析する AIや機械学習による、PIIの自動判定等の機能も具備する
g. データリネージ	<ul style="list-style-type: none"> データの出自や来歴（どの様な変換を経て来たのか）に関する情報を管理して可視化する機能 利用者が、入手したデータの正当性の把握や、データ変換不備についてデータパイプラインの変換過程の中から原因を調査可能にする 基本的には本プラットフォームに格納後から追跡を開始する想定だが、データソース側との協議し格納前の情報の追跡可否についても今後検討することを想定
h. データマネジメントAPI	<ul style="list-style-type: none"> 「c. メタデータ管理」、「d. マスタデータ管理」、「e. ナレッジグラフ」、「f. データプロフィール」、「g. データリネージ」等のデータマネジメントに関する情報を参照・登録・更新・削除可能にするAPI 例えばメタデータは、「⑤ サービス利用者ポータル」のデータカタログなど、他システムでの利用も想定しておりAPIで連携することが必須 他のデータマネジメントに関する情報もポータル等の情報発信や、申請業務等で利用することが想定される

② 分散解析・連合学習プラットフォーム — 概要

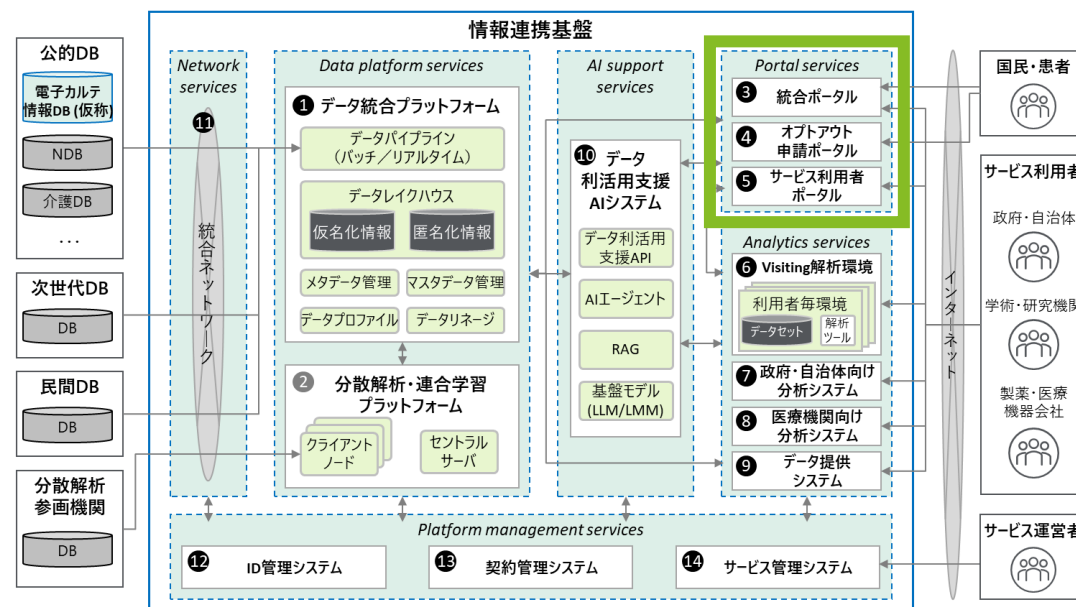
- 分散解析及び連合学習は、データを中央に集約することなく、複数機関で保持するデータを活用して分析や機械学習モデルの訓練を行う手法である
- 参画機関は、データを中央に集約する必要が無いため、個人情報提供に関する同意取得への対応や、自機関のデータ資産の流出対策等が不要となり、複数機関横断での研究や分析、機械学習が行いやすくなるというメリットがある
- 一方で、解析の精度を担保するためには、参画機関のデータを標準化するため、CDM（Common Data Model）の導入が必要となる
- システムとしては、クライアントノードと言う、参画機関毎に分離されたプライベートなDBと計算リソースをクラウド上に構築する機能も提供する



② 分散解析・連合学習プラットフォーム — 機能概要説明

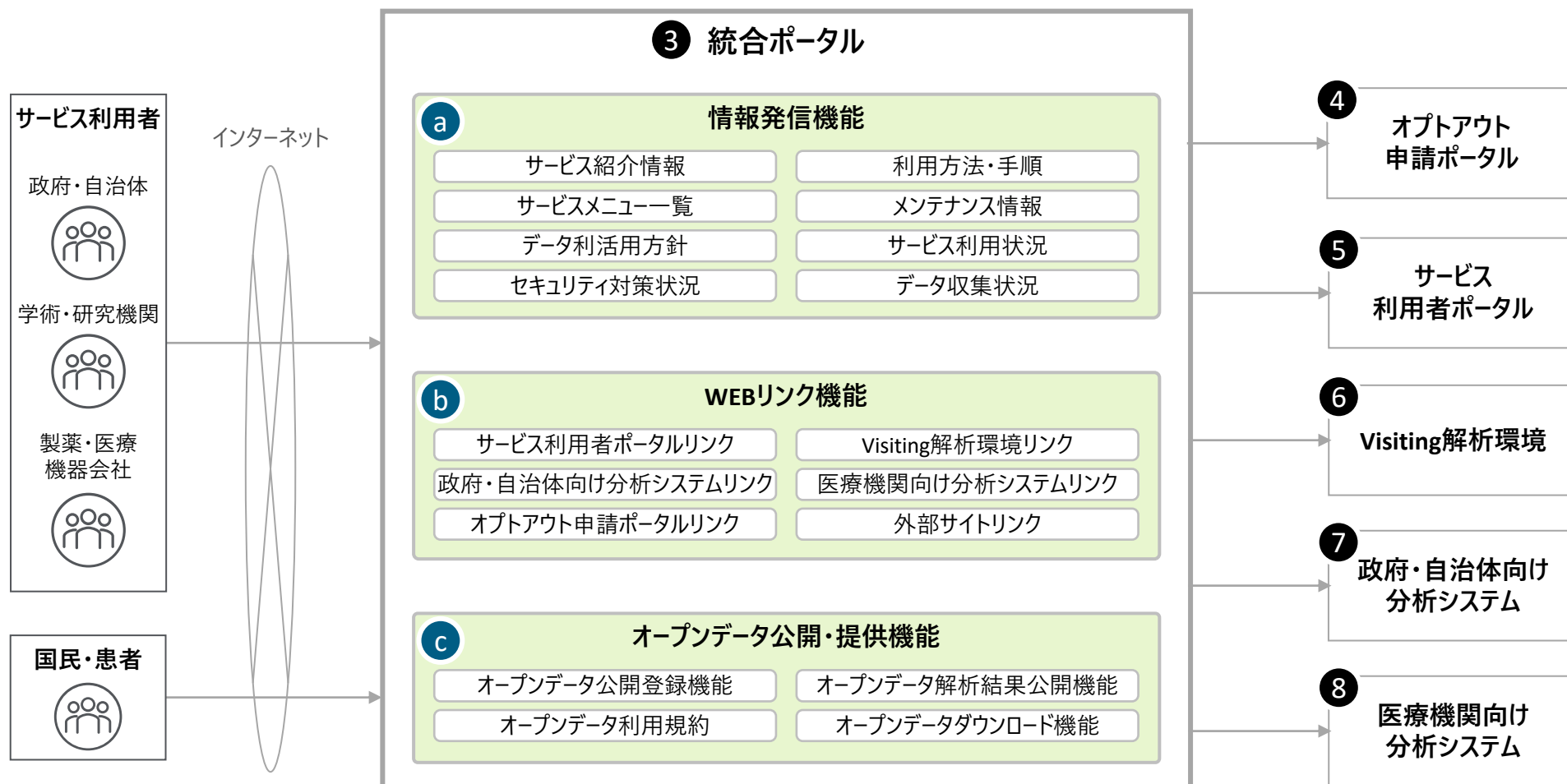
機能概要	
a. セントラルサーバ	<ul style="list-style-type: none">分散解析・連合学習の制御を行うサーバ分散解析・連合学習のタスクを各クライアントノード向けのメッセージとしてキューに格納する機能連合学習前のグローバルモデルを格納して、各クライアントノードに連携する機能を具備する各クライアントで学習済みのモデル（ローカルモデル）を受け取って集約し学習済みのグローバルモデルを生成する機能を具備する各クライアントノードの一覧の管理、ステータスの管理を行う機能を具備する
b. クライアントノード	<ul style="list-style-type: none">分散解析・連合学習プラットフォームに参画する各機関向けのプライベートなデータベースと計算リソースを稼働させる環境各参画機関は自身でクライアントノード環境を構築することも可能であるが、情報連携基盤サービスとして各参画機関向けに専用環境を構築して提供することも想定するクライアントノードには、Agent機能を具備し、セントラルサーバからのタスク実行指示の取得や、グローバルモデルのダウンロード、ローカルモデルのアップロードに加え、自ノードの稼働状況をモニタリングしてセントラルサーバに通知する機能を有する

Portal services



③ 統合ポータル — 概要構成

- 本システムは、サービス利用者、サービス利用を検討者、及び国民全般に広く本サービスに関する情報を発信するためのシステムである
- 情報発信内容としては、本サービスの説明、データ利活用方針、サービス内容、メンテナンス情報、サービス利用状況、データ種別や件数、セキュリティ対策状況、今後の計画と多岐に渡る
- また、各システムの窓口システムの役割としてのWEBリンク機能や、オープンデータの提供機能も提供する

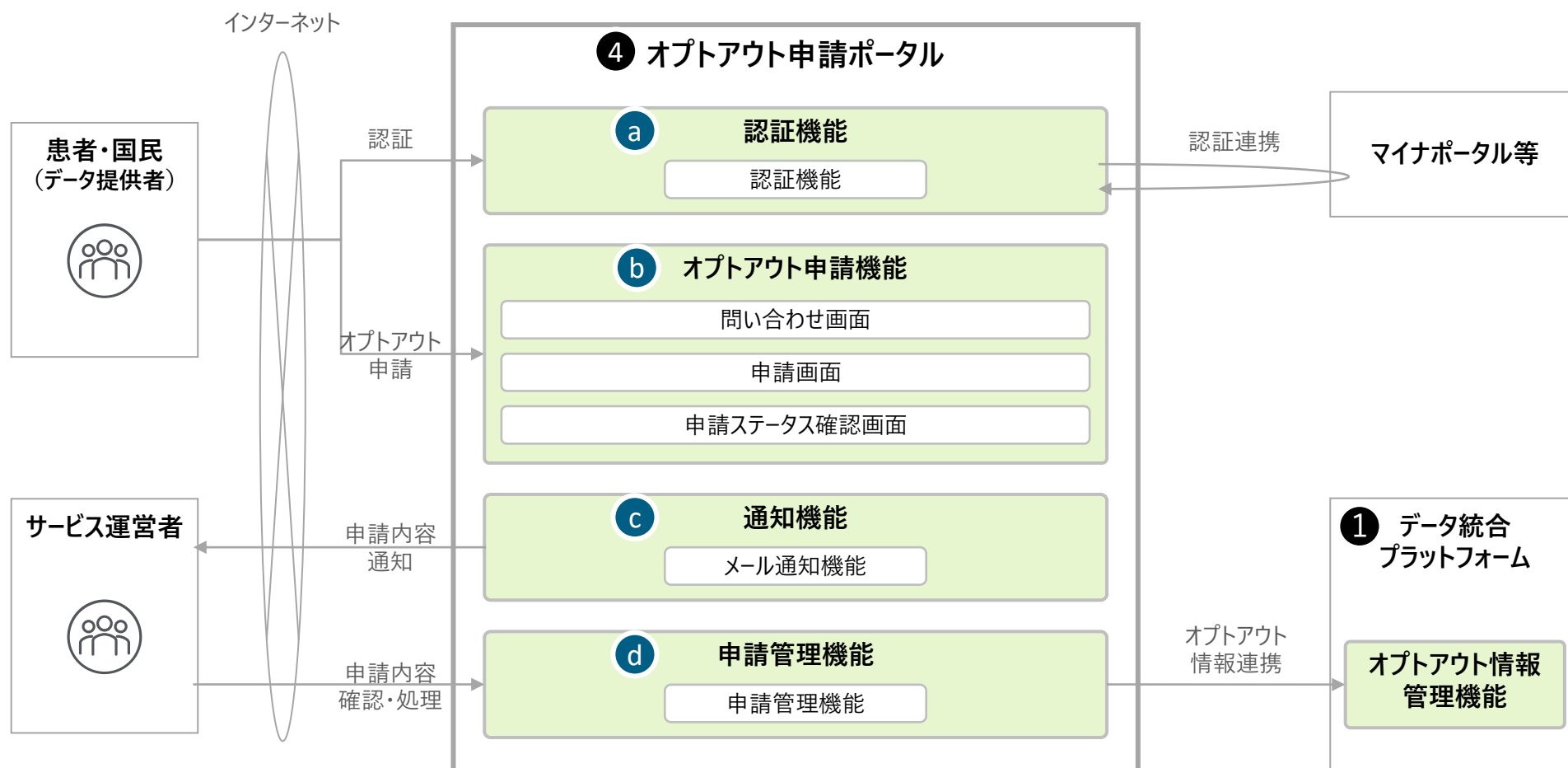


③ 統合ポータル — 機能概要説明

	機能概要
a. 情報発信機能	<ul style="list-style-type: none"> 本サービスに関する情報を集約し、サービス利用者のみならず広く国民全般に情報を発信するための機能 具体的なコンテンツは今後検討を重ねる必要があるが、一例として下記を想定している <ol style="list-style-type: none"> サービス紹介情報 … 本サービスの構築背景・目的、サービス概要等に関する説明情報を発信する サービスメニュー一覧 … 本サービスが提供しているサービスメニュー一覧と概要説明情報を発信する サービス利用方法・手順 … サービス利用見込みユーザ向けに、利用方法・利用手順に関する情報を発信する データ利活用方針 … 仮名化データに関する取り扱いなど、個人情報保護やデータ利活用方針を発信する サービス利用状況 … 本サービスの利用プロジェクト数や利用者数を示し、サービスの有用性を発信する セキュリティ対策状況 … データ統合管理する仕組みとして必要なセキュリティ対策が行われていることを発信する メンテナンス情報 … 本サービスに関するメンテナンスやサービス利用不可時間に関する情報を発信する データ収集状況 … 本サービス上に蓄積されているデータ種別・データ量・件数等に関する情報を発信する
b. WEBリンク機能	<ul style="list-style-type: none"> 本サービスが提供する各システムへのWEBリンクを集約して、サービス利用者をナビゲートする機能 また、外部の関連サイトへのWEBリンクも掲載する
c. オープンデータ公開・提供機能	<ul style="list-style-type: none"> オープンデータを公開し提供する機能 本サービスの利用ユーザアカウントを持たない者でもデータのダウンロードが可能となる オープンデータに関する利用規約や、オープンデータの解析結果（例）を掲載し、利用促進を図る

④ オプトアウト申請ポータル — 概要構成

- 本システムは、患者・国民等が自身のデータ提供に対して、不同意である旨を入力および申請することができる機能を提供する
- 申請の正当性を担保するために申請者の本人確認（認証）が必要となるが、国民を一意に認証できる仕組みをどのように実現するか、今後検討が必要となる（図ではマイナポータルとの連携のイメージで記載）
- 申請内容をサービス運営者が確認した後、①データ統合プラットフォームでオプトアウト対象データを識別できるように、オプトアウト情報の連携を行う機能を具備する

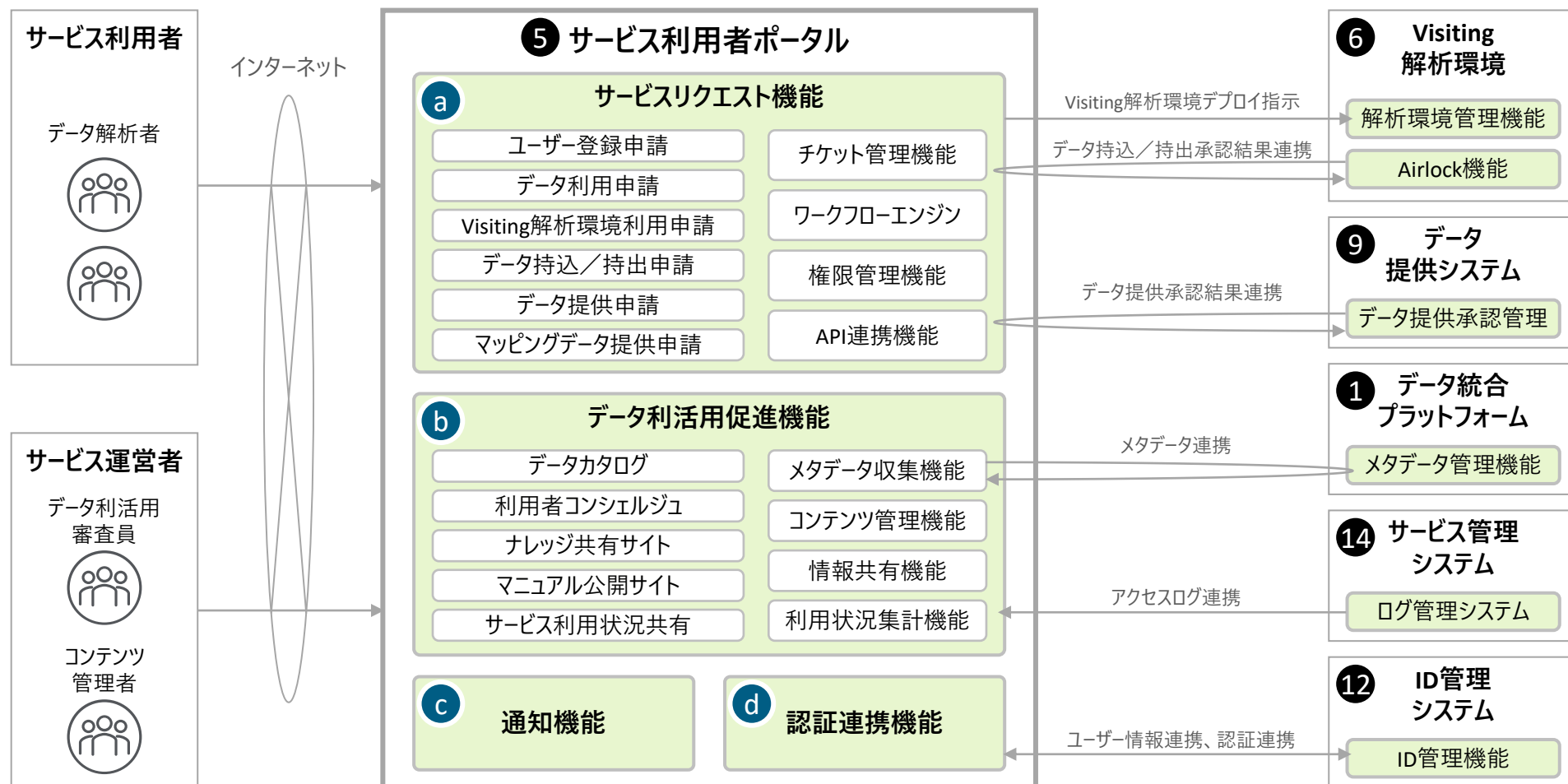


④ オプトアウト申請ポータル — 機能概要説明

	機能概要
a. 認証機能	<ul style="list-style-type: none">申請の正当性を担保するために、申請者本人を認証する機能国民を一意に認証できる仕組みをどのように実現するか、今後検討が必要となる未成年等、代行申請に係る認証・認可に関しても議論が必要
b. オプトアウト申請機能	<ul style="list-style-type: none">国民が自身のデータ利活用について、不同意である旨を入力および申請することができる機能即オプトアウトにならない場合をも想定し、広く問い合わせ窓口も具備する自身の申請内容に対する、サービス運営者側での対応状況やステータスを確認できる機能を具備する
c. 通知機能	<ul style="list-style-type: none">申請があった旨をサービス運営者にメール通知する機能
d. 申請管理機能	<ul style="list-style-type: none">サービス運営者が、オプトアウト申請について、内容の検索・参照・ステータス管理・対応内容の記録を行うことが可能な機能サービス運営者にて確認と所定の対応を実施後、「④ データ統合プラットフォーム」にてオプトアウト対象データを識別可能にするために、オプトアウト情報を連携する機能を具備する

⑤ サービス利用者ポータル — 概要構成

- 本システムは、利用者からの各種リクエスト窓口の一元化とワークフローのデジタル化による迅速なサービス提供を実現するシステムである
- 本システム上の各種承認結果をVisiting解析環境や、データ提供システムに連携し、データ利活用におけるガバナンスを担保する
- また、データ利活用の促進のために、研究成果の公開や、データ分析に関するナレッジ提供、データカタログ等の機能を具備する



⑤ サービス利用者ポータル — 機能概要説明

機能概要

a. サービスリクエスト機能

- サービス利用者からの各種リクエストへの対応業務の円滑な推進と、データ利活用に関するガバナンスを確実に効かせるための機能
- ワークフローの自動化／半自動化、承認証跡を含む対応情報の確実かつ効率的な記録と保全を実現する
- 各承認結果は、「⑥ Visiting解析環境」の環境デプロイ作業や、「⑨ データ提供システム」の提供先確認処理等、他のシステムで利用することが想定されるため、APIによるワークフロー情報の連携機能を具備することが必須となる
- サービス利用者からのリクエスト種別は現時点で下記 6 種を想定している
 - ユーザ登録申請 … サービス利用開始前にユーザ登録申請を行う機能。登録後、ユーザIDが払い出される
 - データ利用申請 … データカタログの利用したいデータを申請する機能。承認後データへのアクセス権が付与される
 - Visiting解析環境利用申請 … ⑥ Visiting解析環境の利用申請機能。目的・用途等の他、必要な環境スペックも申請する
 - データ持込／持出申請 … ⑥ Visiting解析環境へのデータ持込／持出を申請する機能。Airlock機能と連動する
 - データ提供申請 … 蓄積されたデータの外部提供（連携）を申請する機能。⑨データ提供システムと連動する
 - マッピングデータ提供申請 … サービス利用者の保持データに本サービスで保持するデータを付加して提供を依頼する機能
 - 研究成果公開申請 … 研究成果の公開を申請する機能
 - 問合せ申請 … 本サービス全般に関する問い合わせを行う機能

b. データ利活用促進機能

- サービス利用者のデータ利活用を促進するための情報提供を行う機能
- 具体的には、下記 5 つの機能提供を想定している
 - データカタログ … データ統合プラットフォームに蓄積しているデータのメタデータを検索・表示する機能
 - 利用者コンシェルジュ … Chat GPTによる利用者のサポート機能。目的に応じた利用すべきデータの推奨等、各種支援を行う
 - ナレッジ共有サイト … データ利活用に関するナレッジ共有機能。利用者自身がコンテンツを作成して公開することが可能
 - マニュアル公開サイト … サービス運営者が作成する公式のマニュアルを掲載する機能
 - サービス利用状況共有 … 利用者数やアクセス数、良く利用されるデータ名等、サービス利用状況を集計して可視化する機能

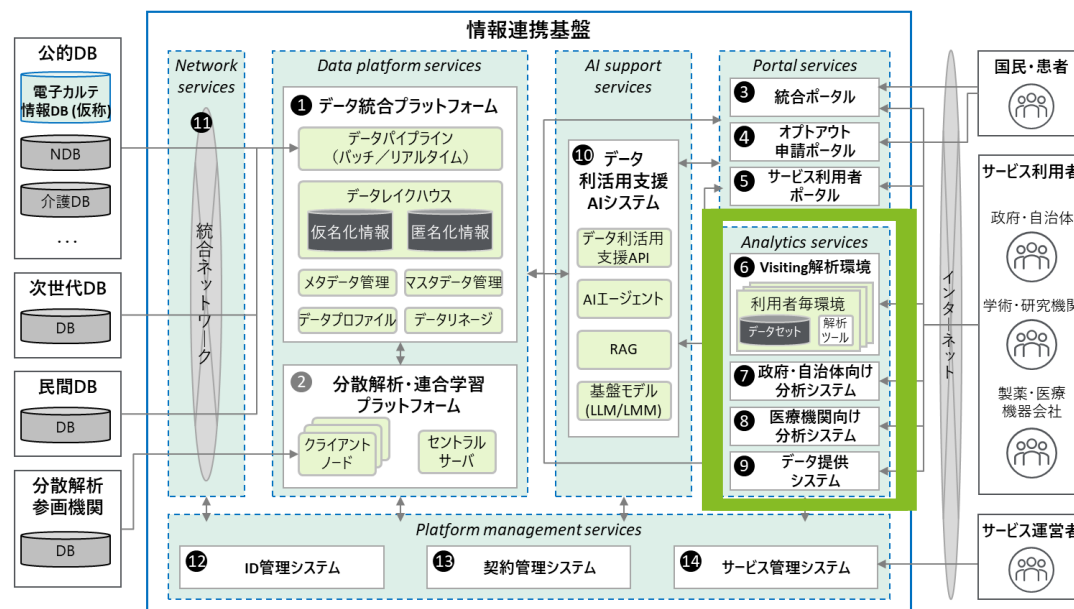
c. 通知機能

- サービス利用者向けに各種通知（メール通知）を行うための機能
- 申請ワークフローで対応者や同報者に通知することや、ナレッジ共有サイト等のコンテンツ更新時に利用者に通知することを想定
- 通知先のメールアドレスは、ID管理システムで管理している連絡先情報から同期する想定
- 通知内容の作成・管理機能、送信タイミング事前設定機能、メールテンプレート機能等、通知業務を支援する機能を具備する想定

d. 認証連携機能

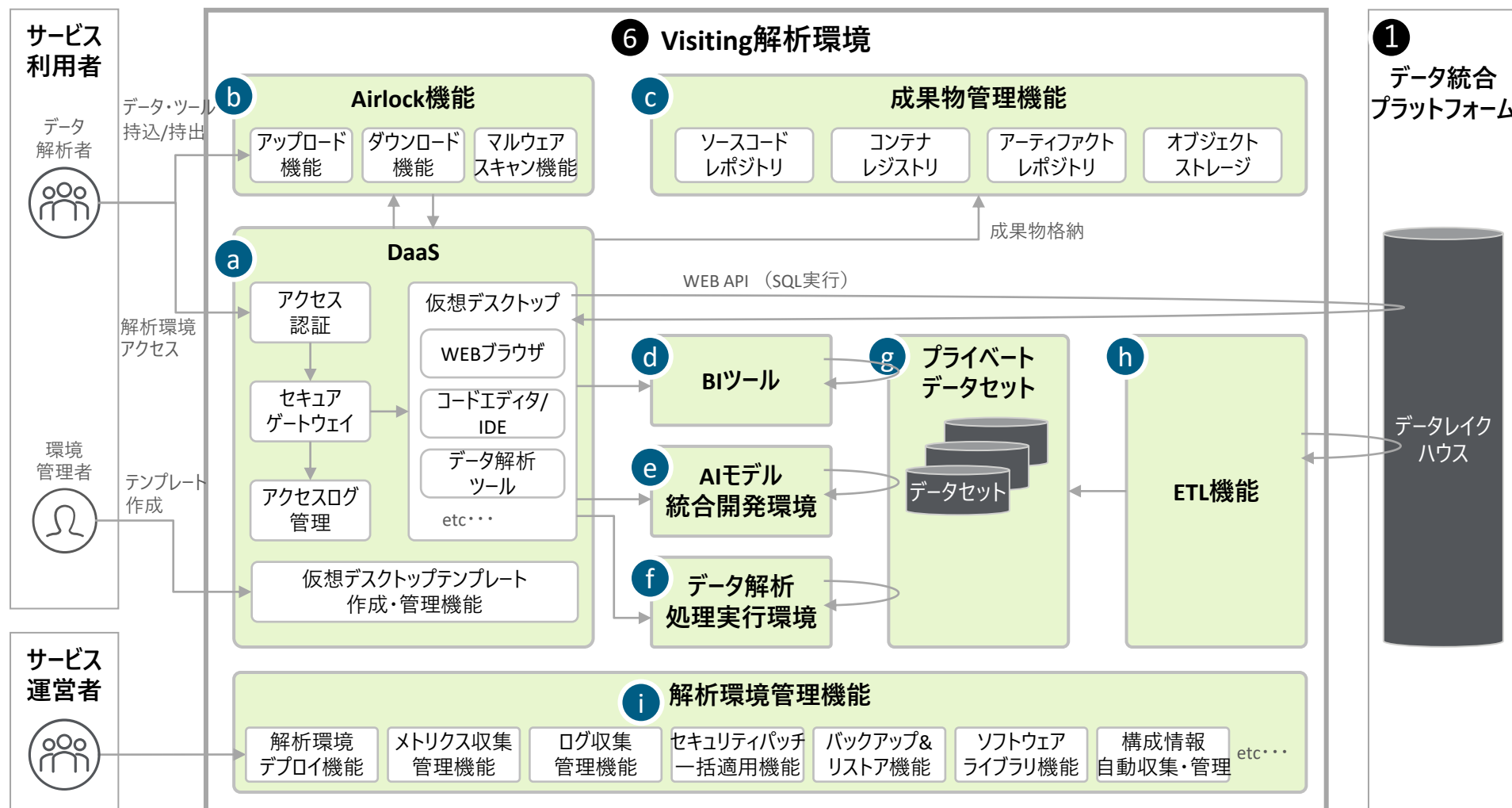
- サービス利用者ポータルにアクセスしてきた、サービス利用者を認証する機能
- 認証処理は、「⑫ ID管理システム」で行うことを想定しており、「⑫ ID管理システム」と認証情報の連携を行うことを想定している（シングルサインオン）

Analytics services



⑥ Visiting解析環境 — 概要構成

- Visiting解析環境は、サービス利用者に対してプライベートな解析用の環境を提供するサービスである
- データ解析者はDaaS（Desktop as a Service）により提供される仮想デスクトップにアクセスし、解析作業を実施する
- またサービス運営者が研究プロジェクト毎に多数作られるVisiting解析環境を効率良く管理するための機能を提供する



⑥ Visiting解析環境 — 機能概要説明 (1/2)

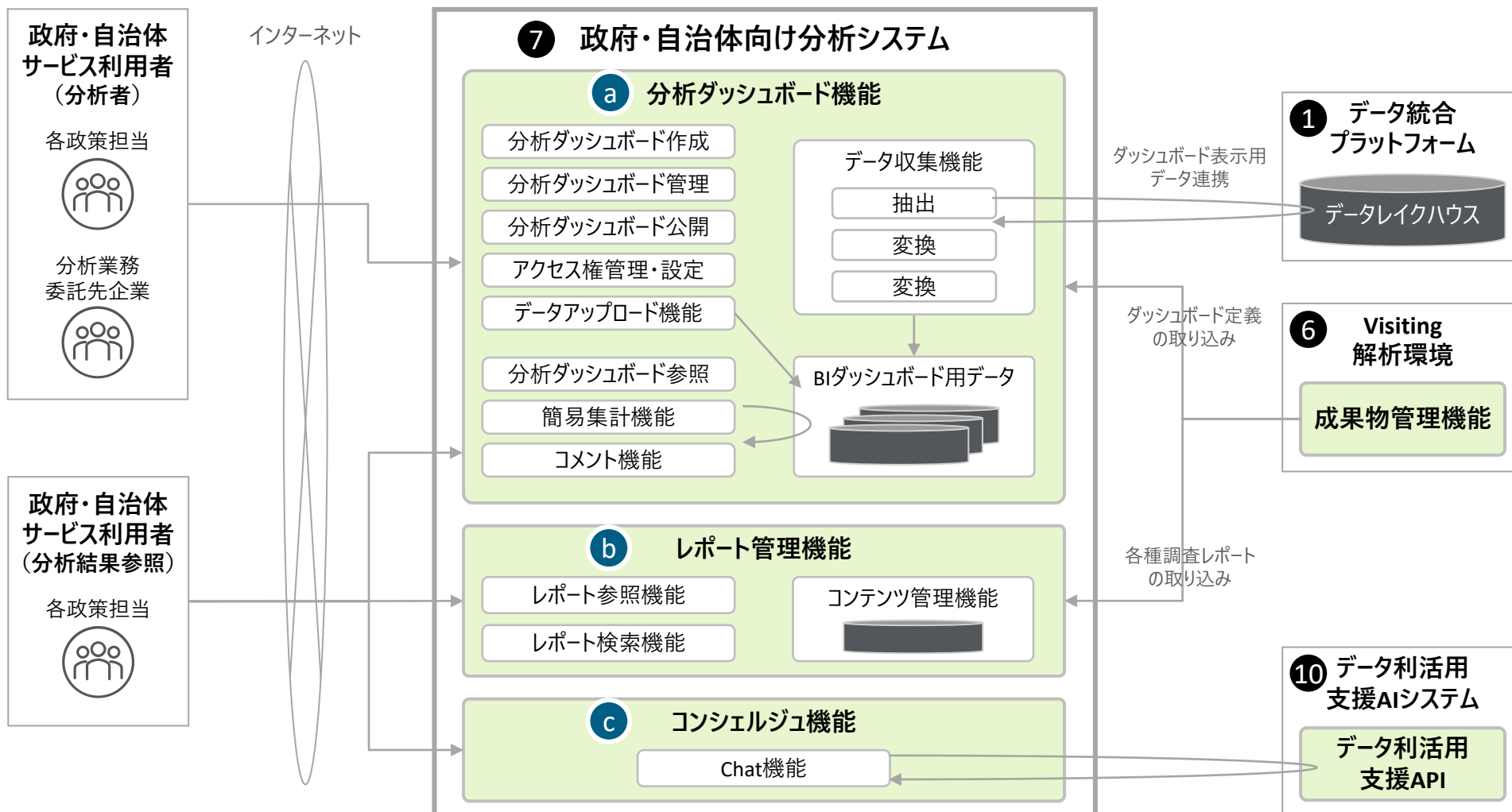
	機能概要
a. DaaS	<ul style="list-style-type: none"> データ分析者がVisiting解析環境を利用する際に必ず経由する仮想デスクトップ環境 プリインストールされたWEBブラウザやコードエディタ、データ解析ツールの他、外部から持ち込んだツールをインストールすることが可能 サービス利用者側の環境管理者が、自プロジェクトで利用するツールをインストールした仮想デスクトップイメージをテンプレートとして保存しておき、テンプレートから仮想デスクトップを立ち上げることが可能（サービス運営者の承認は必要）
b. Airlock機能	<ul style="list-style-type: none"> 仮想デスクトップへのデータやツールの持込、及び仮想デスクトップからのデータや成果物の持出を実現する機能 持込／持出対象のデータをアップロード後、「⑤ サービス利用者ポータル」で承認済みのデータがダウンロード可能となる 外部からデータを持ち込む際にマルウェアスキャンを行う機能を具備し、スキャン済みのデータのみが持ち込めるように制御を行う
c. 成果物管理機能	<ul style="list-style-type: none"> Visiting解析環境で作成した成果物（ソースコード、コンテナ、ビルド済みモジュール、データファイル、ドキュメント、等）を管理する機能 成果物の版管理と仮想デスクトップをimmutableに運用するために成果物を仮想デスクトップ外部で管理する設計としている CICDパイプラインと統合可能のように、基本的にWEB APIを備えることを必須とする
d. BIツール	<ul style="list-style-type: none"> Visiting解析環境内で、利用プロジェクト毎にプライベートに利用できるBIツール 「g. プライベートデータセット」のデータを可視化することに加え、データレイクハウスのデータを直接可視化することも可能とする 単純な可視化に加え、トラジェクトリー解析等の機械学習を活用したデータ解析が可能なツールを採用する
e. AIモデル統合開発環境	<ul style="list-style-type: none"> WebベースのAIモデルの統合開発環境（IDE） トレーニングデータのプリパレーションや、AIモデル構築・トレーニング・評価・開発したモデルのパッケージングまでを一貫して行える パッケージングしたモデルは、「b. Airlock機能」で外部に持ち出して利用することが可能
f. データ解析処理実行環境	<ul style="list-style-type: none"> 解析用のプログラムを実行する環境。DaaSでは事足りない大規模データ処理や長時間に及ぶ解析処理を実行する環境を想定 実行に必要なライブラリや、リソースをサービス利用者が設定可能にする 単一の仮想サーバ上での実行に加え、Apache Spark等の大規模データ分散処理の実行も可能とする
g. プライベートデータセット	<ul style="list-style-type: none"> データレイクハウスから抽出したデータや、外部から持ち込んだ解析用データを格納するプライベート利用可能なデータ格納領域 データセットの格納先にはRDBの他、KVSやドキュメントDB、ビッグデータ用のカラムナードB、オブジェクトストア（非構造データ用）等、扱うデータに合わせて自由に選択が可能。もちろん複数選択も可

⑥ Visiting解析環境 — 機能概要説明 (2/2)

	機能概要
h. ETL機能	<ul style="list-style-type: none">データレイクハウスからデータを抽出し加工・変換をして「g. プライベートデータセット」にロードするための機能データ解析者が容易に扱えるようなWEBベースの開発エディタや、ノーコード／ローコードによるETL処理開発機能も検討するETL処理の定義はXML等のテキストデータとしてコードレポジトリで管理可能とする
i. 解析環境管理機能	<ul style="list-style-type: none">サービス管理者が、サービス利用者に提供するVisiting解析環境をデプロイ・モニタリング・管理するための機能複数環境を一括して管理可能にするためにメトリクスやログを統合管理し、セキュリティパッチ等の一括配布・適用を可能とする環境障害時にバックアップから復旧可能なように定期バックアップも取得するDaaS環境等にインストールが可能なソフトウェアライブラリ機能を提供し、利用者が個別ソフトウェアを持ち込む頻度を減らす構成情報の自動収集・管理機能を具備し、不正なソフトウェアのインストール検知や障害時の復旧対応迅速化を目指す

⑦ 政府・自治体向け分析システム — 概要構成

- 政府・自治体の各政策担当がEBPMや政策立案・評価に必要なデータ分析結果をダッシュボードとして作成し、参照するシステムである
- ダッシュボードのみならず、EBPMや政策立案・評価に必要なデータ分析レポートや研究論文、調査報告書等を管理する機能を有する
- また、コンシェルジュ機能を具備し、AIによる必要情報の収集や文章生成の支援を行い業務効率化を実現する

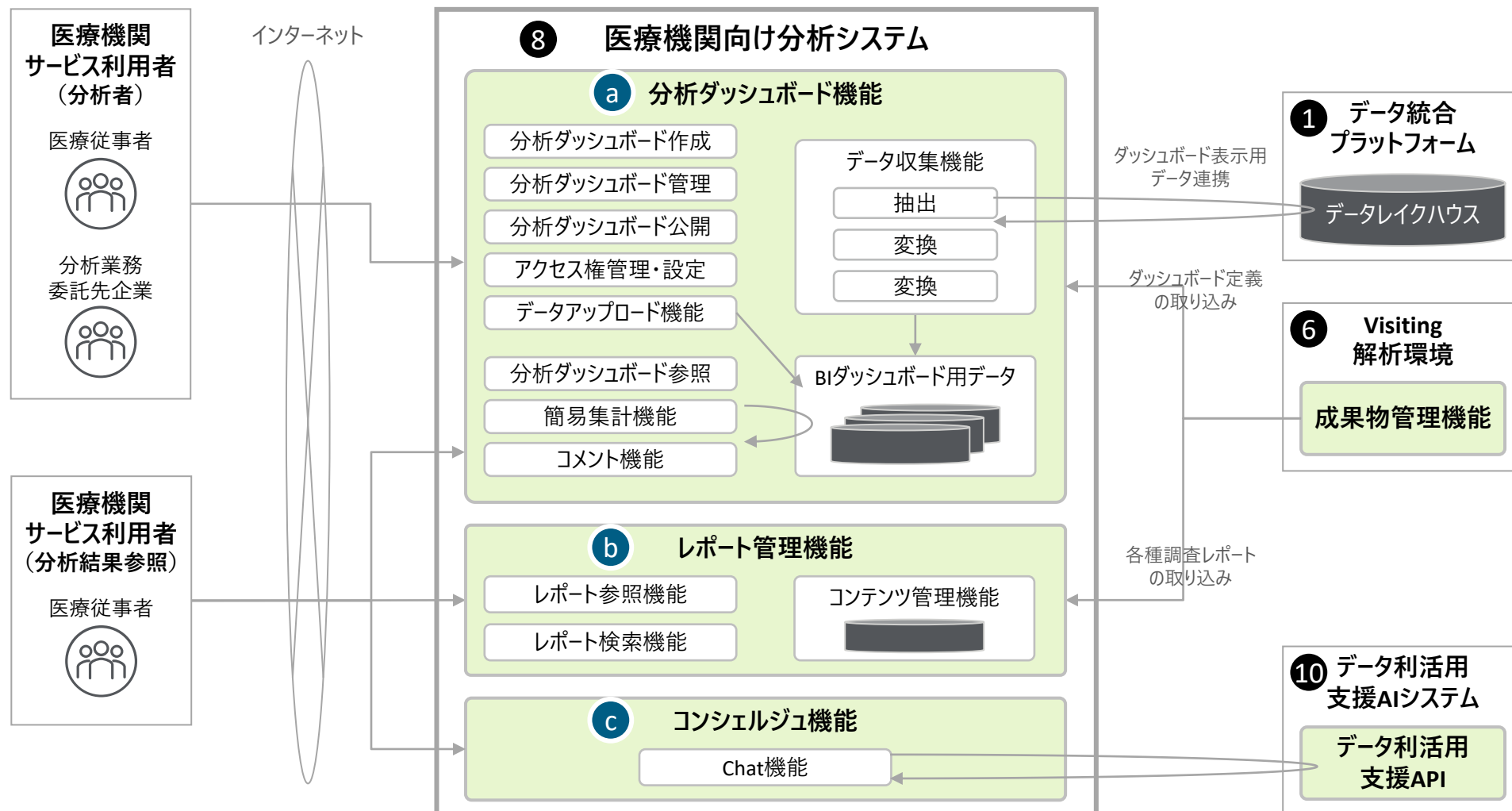


⑦ 政府・自治体向け分析システム — 機能概要説明

機能概要	
a. 分析ダッシュボード機能	<p>ダッシュボード作成・管理機能</p> <ul style="list-style-type: none">「⑥ Visiting解析環境」の分析成果を元に、政府・自治体の政策担当者がEBPMや政策立案・評価に必要なデータ分析結果をダッシュボードとして作成し公開する機能作成したダッシュボードの版管理・公開制御・アクセス制御を可能とするデータレイクハウスに存在しないデータをダッシュボードで可視化するユースケースを考慮して、データアップロード機能を具備するダッシュボードに必要なデータをデータレイクハウスから収集・変換・蓄積する機能を有する <p>ダッシュボード参照機能</p> <ul style="list-style-type: none">作成されたダッシュボードを参照する機能。ダッシュボードへのコメント書き込み機能や、並び替えや集計軸の変更等の簡易的な集計機能を具備する
b. レポート管理機能	<ul style="list-style-type: none">「⑥ Visiting解析環境」等で作成した分析レポート・論文・調査結果報告書を管理し掲載する機能政府・自治体の政策担当者がEBPMや政策立案・評価に必要な情報を集約して業務を効率的に行うための支援機能全文検索機能を具備し必要な情報を横断検索可能にする
c. コンシェルジュ機能	<ul style="list-style-type: none">政府・自治体の政策担当者がEBPMや政策立案・評価する際に必要な情報の収集や、収集した情報のサマリ等、生成AIによる業務支援を行う機能「⑩ データ利活用支援AIシステム」と連携して本機能を実現する

⑧ 医療機関向け分析システム — 概要構成

- サービス利用者である医療機関等が過去の症例データにアクセスし、疾患動向や類似症例の傾向を把握することで診断を支援する
- 臨床や医療統計、疫学管理を通じて現実に即したケーススタディや教育を行い、医療従事者の知識向上に寄与するシステムである
- 基本的に具備する機能は、「⑦ 政府・自治体向け分析システム」と同様である

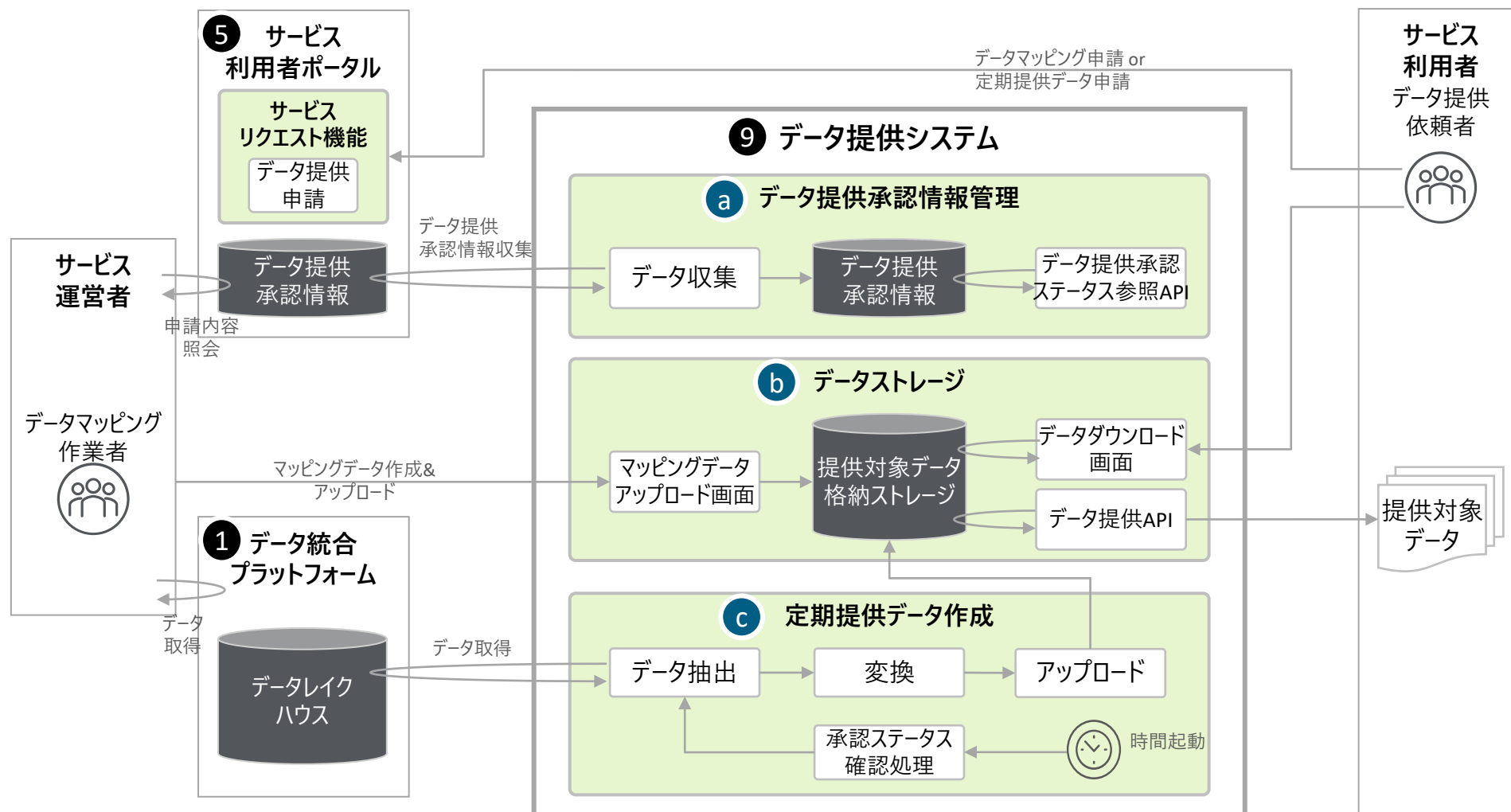


⑧ 医療機関向け分析システム — 機能概要説明

機能概要	
a. 分析ダッシュボード機能	<p>ダッシュボード作成・管理機能</p> <ul style="list-style-type: none">「⑥ Visiting解析環境」の分析成果を元に、医療機関等に勤務する医療従事者が疾患動向や類似症例の傾向を把握するデータ分析ダッシュボードを作成し公開する機能作成したダッシュボードの版管理・公開制御・アクセス制御を可能とするデータレイクハウスに存在しないデータをダッシュボードで可視化するユースケースを考慮して、データアップロード機能を具備するダッシュボードに必要なデータをデータレイクハウスから収集・変換・蓄積する機能を有する <p>ダッシュボード参照機能</p> <ul style="list-style-type: none">作成されたダッシュボードを参照する機能。ダッシュボードへのコメント書き込み機能や、並び替えや集計軸の変更等の簡易的な集計機能を具備する
b. レポート管理機能	<ul style="list-style-type: none">「⑥ Visiting解析環境」等で作成した分析レポート・論文・調査結果報告書を管理し掲載する機能医療従事者が疾患動向や類似症例の傾向を把握するために必要な情報を集約して業務を効率的に行うための支援機能全文検索機能を具備し必要な情報を横断検索可能にする
c. コンシェルジュ機能	<ul style="list-style-type: none">医療従事者が疾患動向や類似症例の傾向を把握する際に必要な情報の収集や、収集した情報のサマリ等、生成AIによる業務支援を行う機能「⑩ データ利活用支援AIシステム」と連携して本機能を実現する

⑨ データ提供システム — 概要構成

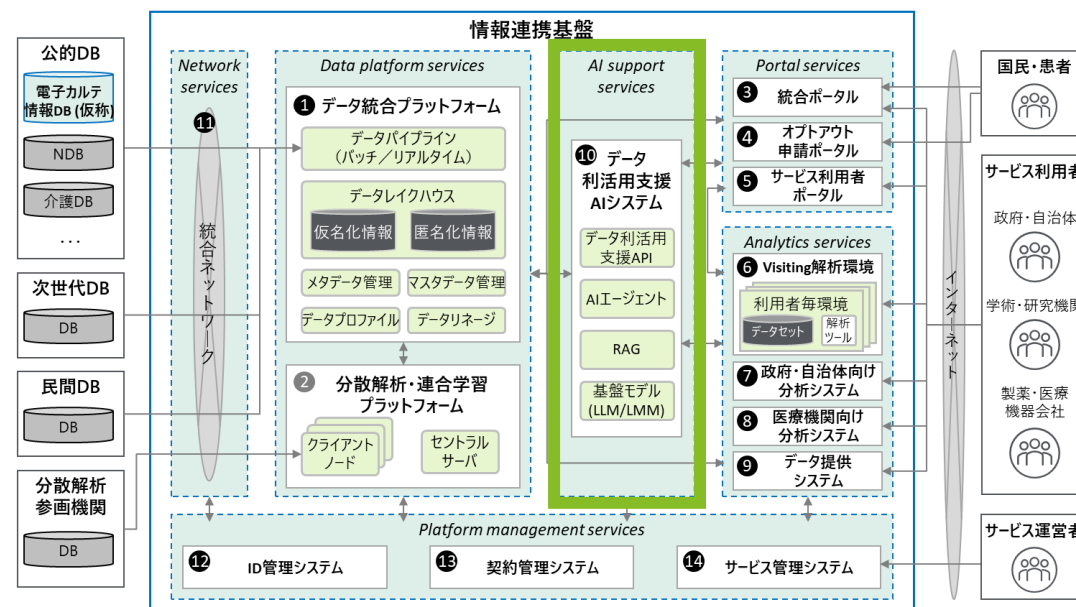
- 本システムは、国立研究機関等の事前に承諾が得られた機関／法人が有するシステム環境にデータを提供する仕組みである
- サービス利用者ポータルと連携し、事前承認されたデータのみが提供されるよう管理する仕組みを具備する
- また、サービス運営者がサービス利用者からの申請に基づきデータを付加して提供するデータマッピングサービスも提供する



⑨ データ提供システム — 機能概要説明

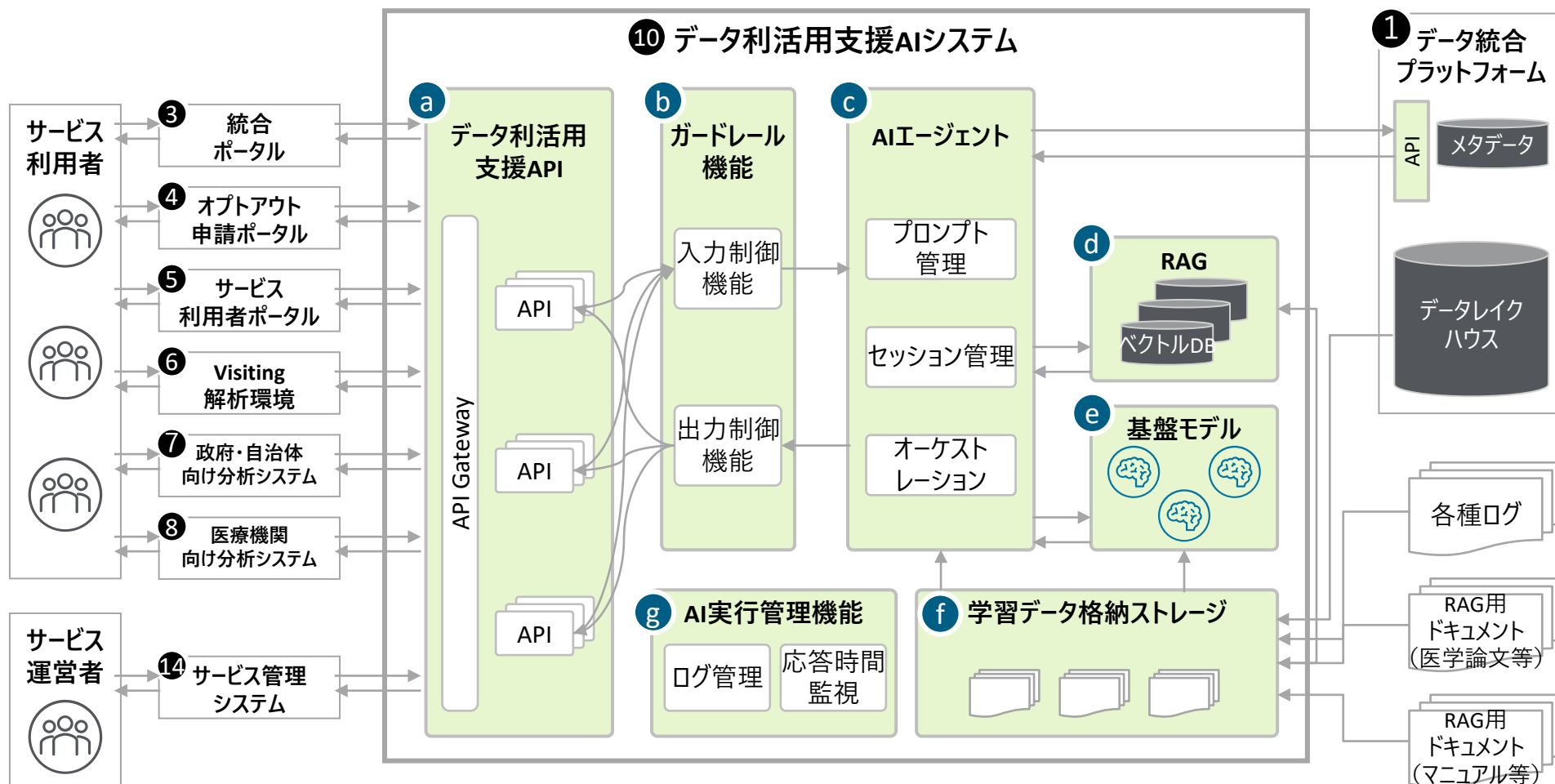
	機能概要
a. データ提供承認情報管理	<ul style="list-style-type: none">「⑤ サービス利用者ポータル」から定期的にデータ提供に関する承認証跡データを取得して管理する機能システム間の依存度を下げ、メンテナンス性を向上するため、承認証跡データを自システム内に保持する設計としているb. API連携や、c. 定期ファイル送信時に、提供対象データと提供先の組み合わせについて、事前に承認が行われていることをチェックするAPIを提供する
b. データストレージ	<ul style="list-style-type: none">提供するデータを格納する機能サービス利用者はデータダウンロード画面からデータを取得するか、データ提供API経由でデータを取得可能となるサービス運営者が、データマッピング申請内容に応じて作成したマッピングデータを格納する画面も有する
c. 定期提供データ作成	<ul style="list-style-type: none">事前定義されたスケジュールにて、定期的にサービス利用者に対して提供するデータを作成する機能データ提供の度に、提供するデータ内容と宛先が事前承認内容と相違ないことをチェックする処理を行い、相違ない場合にのみデータ提供を可能にする作成したデータは、b. データストレージに格納して、サービス利用者がデータをダウンロード可能なようにする

AI support services



⑩ データ利活用支援AIシステム — 概要構成

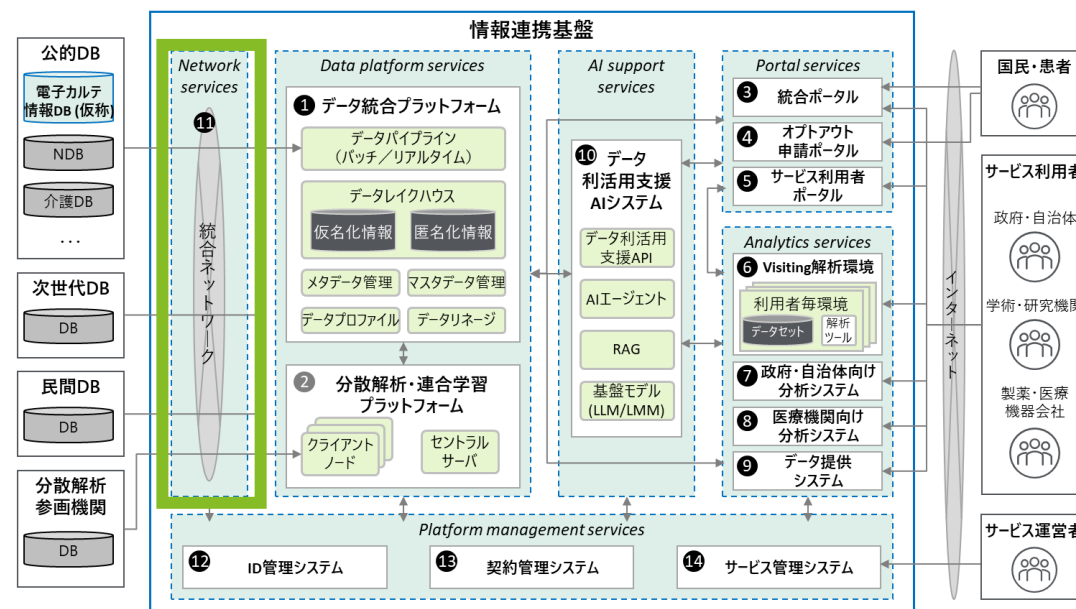
- サービス利用者ポータルにおけるコンシェルジュ機能や、サービス運営者の運営業務の効率化を実現するなど、LLMを活用した支援機能を提供する仕組みである
- 情報連携基盤の様々なシーンでLLMを活用した業務効率化／ユーザビリティ向上を実現するため、独立したシステムとして設計としている



⑩ データ利活用支援AIシステム — 機能概要説明

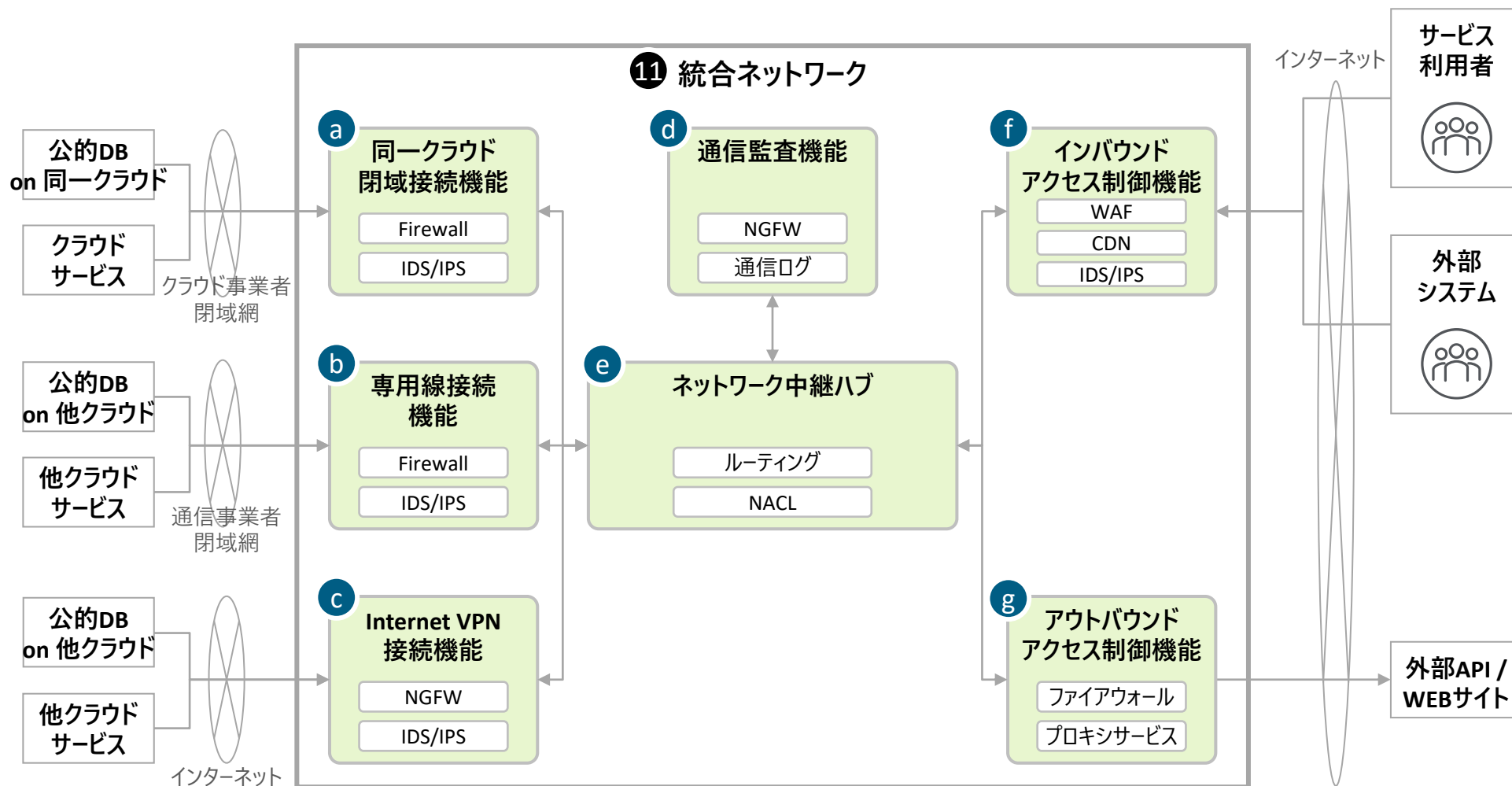
	機能概要
a. データ利活用支援API	<ul style="list-style-type: none"> LLMを活用したデータ利活用支援機能を提供するAPI サービス利用者ポータルや、統合データプラットフォームなど、複数のシステムから利用されることを想定しAPIとして提供している APIクライアントからの入力バリデーションや、基盤モデルにリクエストを連携する処理を実装する
b. ガードレール機能	<ul style="list-style-type: none"> 事前にポリシーを定義して、ユーザーから基盤モデルへの入力や、基盤モデルからの出力の内容をチェックして、不適切な回答やワードをブロックする機能 個人情報のブロックや有害コンテンツに対する制御を行う
c. AIEージェント	<ul style="list-style-type: none"> ユーザーが求める結果を達成するために必要なタスクを、複数のタスクに分割してタスクの実行計画を策定するAI 各タスクごとに適切なAPIを実行し、全てのタスクの結果を踏まえユーザが求める結果を返す ユーザーと生成AIの対話で解にたどり着くチャット形式や、タスクを必ず決まった順序で実行するワークフローと使い分けを行う
d. RAG	<ul style="list-style-type: none"> ユーザーからの問合せに対して、関連ドキュメントを検索して情報を取得し、取得情報をプロンプトに渡してLLMで結果を整形してユーザーに返す機能。基盤モデルが学習していない情報についても回答することが可能 関連ドキュメントをベクトル化して、RAG用の検索データベースに格納する機能を具備する
e. 基盤モデル	<ul style="list-style-type: none"> LLM / LMM等、ユースケースに沿った最適なモデルを実行することができる機能 複数の基盤モデルを比較・評価する機能を具備する
f. 学習データ格納ストレージ	<ul style="list-style-type: none"> 基盤モデルのファインチューニング用の学習データや、RAGで使用する関連ドキュメントを格納するストレージ
g. AI実行管理機能	<ul style="list-style-type: none"> プロンプト入力内容や、基盤モデルの回答結果内容をログ出力する機能。AIの精度改善やユーザーからのニーズを分析に活用する ユーザーからのリクエスト数に対して十分なスループット、ターンアラウンドタイムをログ出力する機能。性能改善やキャパシティプランニングに活用する

Network services



11 統合ネットワーク 概要

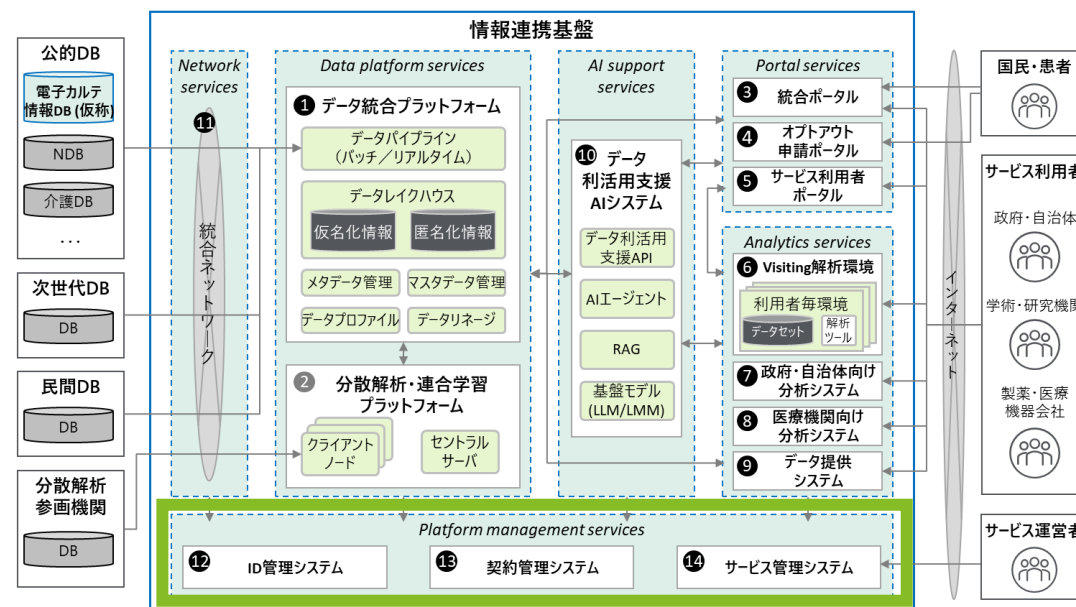
- 情報連携基盤は、公的DBに加え、次世代DB、民間DBとの接続を想定しており、各接続先はそれぞれ別のクラウドサービスやオンプレミスのデータセンターで稼働している。従って、各データソースとセキュアにシームレスに接続するための連携用ネットワークを構築する必要がある
- 情報連携基盤を構成するシステムも、様々なクラウド上に実装される可能性があり、クラウド間を接続する閉域ネットワークが必要となる
- また、Internetからのユーザーアクセス通信や、Internetへのアウトバウンド通信をセキュアに行うための仕組みも必要となる



11 統合ネットワーク ― 機能概要説明

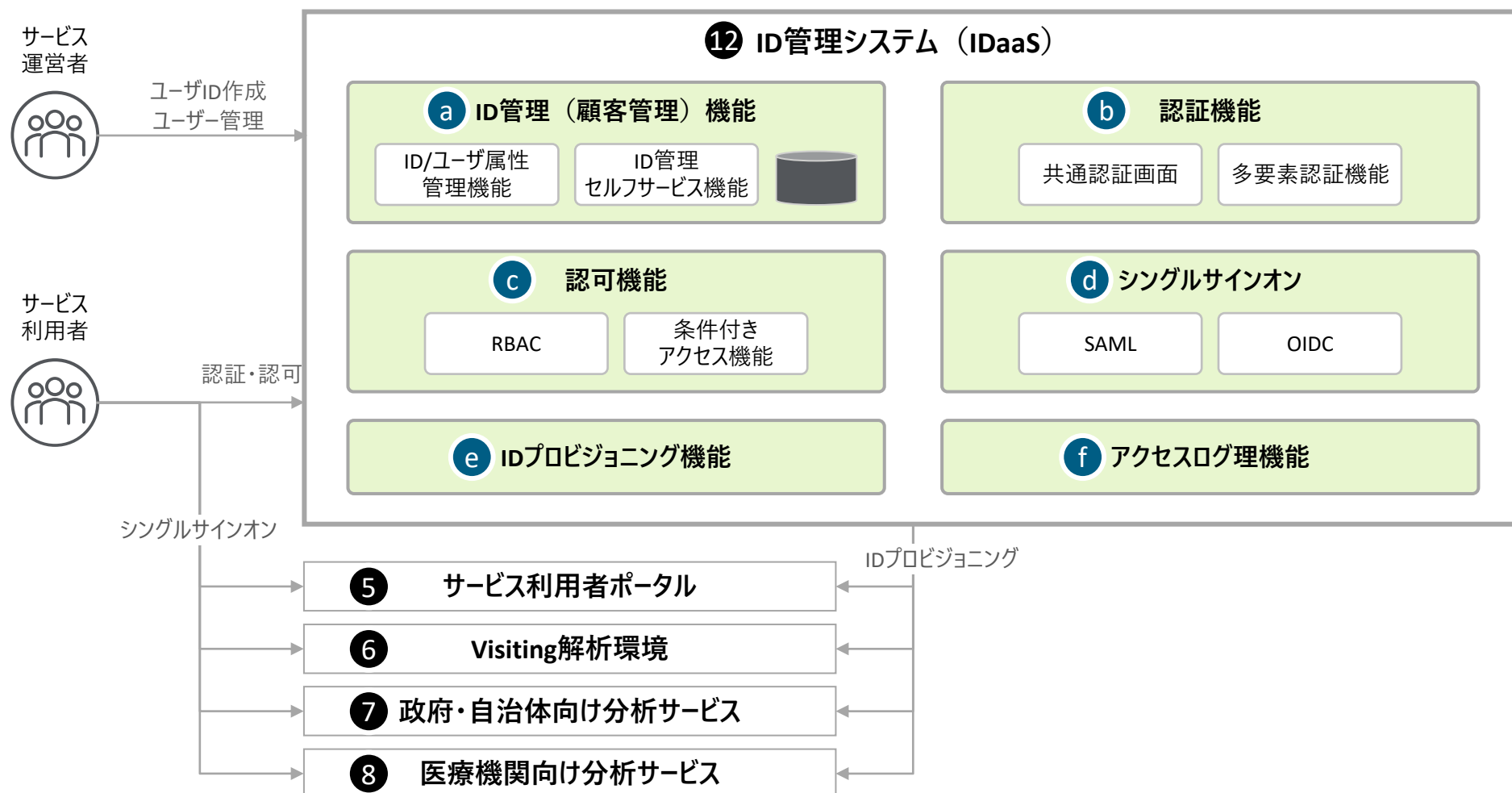
	機能概要
a. 同一クラウド閉域接続機能	<ul style="list-style-type: none"> 同一クラウドのネットワーク間を閉域接続する機能。VPC Peering等の機能を想定 接続先に不正な通信が行かないように、FirewallやIDS/IPS機能を具備する（最低限はFirewallのみ）
b. 専用線接続機能	<ul style="list-style-type: none"> 通信事業者が提供する、オンプレミスネットワークや、他クラウドサービスとの専用線接続を実現する機能 接続先に不正な通信が行かないように、FirewallやIDS/IPS機能を具備する（最低限はFirewallのみ）
c. Internet VPN接続機能	<ul style="list-style-type: none"> 各データソースや他クラウドサービスとInternet VPNによる閉域接続を実現する機能 接続先に不正な通信が行かないように、FirewallやIDS/IPS機能を具備する（最低限はFirewallのみ）
d. 通信監査機能	<ul style="list-style-type: none"> ネットワークセグメント間を流れる通信を監査して、不正な通信をブロックしてネットワークセキュリティ強化するための機能 Next gen Firewall機能の具備等を想定 通信ログを記録して、「⑭ サービス管理システム」に連携してセキュリティ分析が行えるようにする
e. ネットワーク中継ハブ	<ul style="list-style-type: none"> 同一クラウド内の複数の仮想ネットワークセグメント、専用線接続セグメント、VPN接続セグメント、通信制御機能を収容するセグメント、インバウンドアクセス制御機能を収容するセグメント、アウトバウンドアクセス制御機能を収容するセグメント等、複数のネットワークセグメント間の通信を中継するハブ機能 ルーティング機能や、ネットワークアクセスコントロール機能を具備する
f. インバウンドアクセス制御機能	<ul style="list-style-type: none"> Internetから情報連携基盤へのユーザアクセスや外部システムアクセスを行う際に、本機能を経由することで適切な制御を行う 制御機能としては、WAFや、DDOS対策用のIPS/IDS、CDN等を想定している
g. アウトバウンドアクセス制御機能	<ul style="list-style-type: none"> 情報連携基盤からInternet経由で外部APIや外部WEBサイトにアクセスする際に、本機能を経由することで適切な制御を行う 制御機能としては、URLフィルタリング（ホワイトリスト）、ソースNAT、通信先IPアドレス制御等を想定している

Platform management services



12 ID管理システム（IDaaS） — 概要構成

- 情報連携基盤が提供する各サービス／システムにログインするユーザーのIDとパスワード、ユーザー情報、グループを一元管理するサービス
- シングルサインオン機能を提供し、サービス利用ユーザ、運営者はID管理サービスに認証を行えば各システムの個別ログインが不要となる
- 基本的にはIDaaS（クラウド型ID管理サービス）での実装を想定している

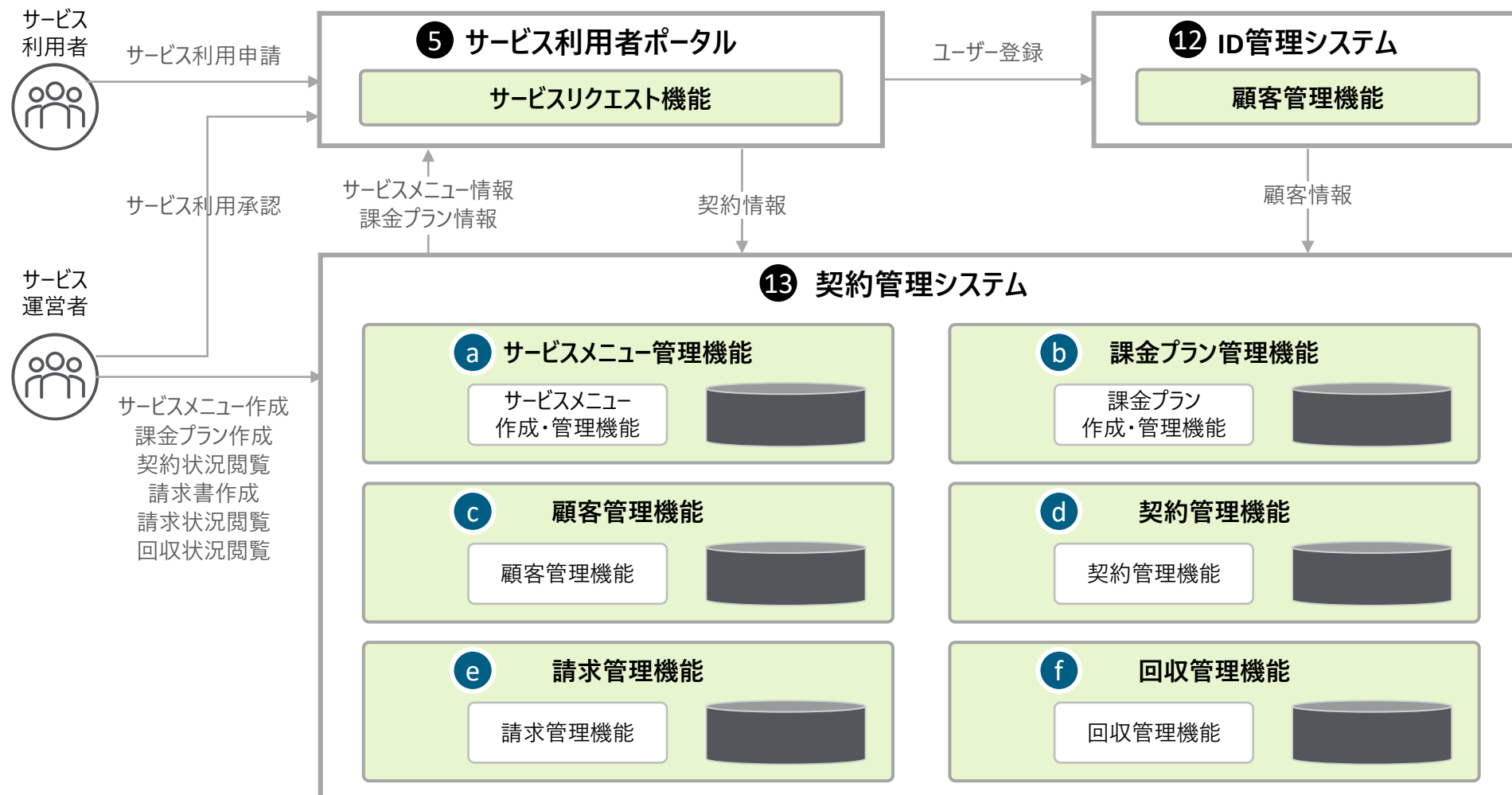


12 ID管理システム（IDaaS） — 機能概要説明

	機能概要
a. ID管理機能	<ul style="list-style-type: none"> サービス利用者及びサービス運営者のID、ユーザー属性情報を一元的に管理する機能 IDの新規登録は、サービス運営者が審査を行った後にサービス運営者が作成を行う IDに紐づくユーザーのパスワードや属性情報の変更を、ユーザ自身で行わせる「管理セルフサービス機能」を具備する
b. 認証機能	<ul style="list-style-type: none"> ユーザー認証を行う機能 セキュリティ強化のために、「多要素認証」機能を必須で具備する
c. 認可機能	<ul style="list-style-type: none"> RBAC（Role base access Control）機能により職責による権限管理が行える機能を具備する アクセス元のIPアドレスや、利用デバイス、アクセス先の組み合わせによりアクセス制御を行う「条件付きアクセス」機能を具備する
d. シングルサインオン機能	<ul style="list-style-type: none"> 1度の認証で、本サービスを構成する全てのシステムに自動でログインすることを可能とする機能 システム毎に個別にIDとパスワードを入力する手間、覚える手間を無くしてユーザビリティを向上する SAMLやOIDC等の、クラウドサービスとのシングルサインオンが可能な方式をサポートすることを想定
e. IDプロビジョニング機能	<ul style="list-style-type: none"> ID管理サービス上で作成・変更・削除したIDやIDに紐づくユーザー属性情報を、本サービスを構成する各システムに同期する機能 本機能により、各システム側での個別のID作成・管理が不要となりサービス運営コストの削減と、IDの権限削除漏れによるセキュリティリスクを低減することを目指す
f. アクセスログ管理機能	<ul style="list-style-type: none"> ユーザーのログイン成功・失敗や、パスワード変更、ユーザ属性情報変更等、関連するログを収集・管理する機能 本機能で集めたログは、「14 サービス管理システム」に連携して、SOC等が必要なセキュリティ分析を行えるようにする

⑬ 契約管理システム — 概要構成

- サービス利用者の利用しているサービスに対して利用料課金を行うことを目指し、サービス利用者の契約情報を管理する機能を具備する
- サービス運営者が、契約管理に必要なサービスメニューや課金プランを作成して管理する機能を具備する
- 本システムで作成したサービスメニューや課金プランをサービス利用者ポータルに連携し、ユーザーがメニューより利用申請が行えるようにする

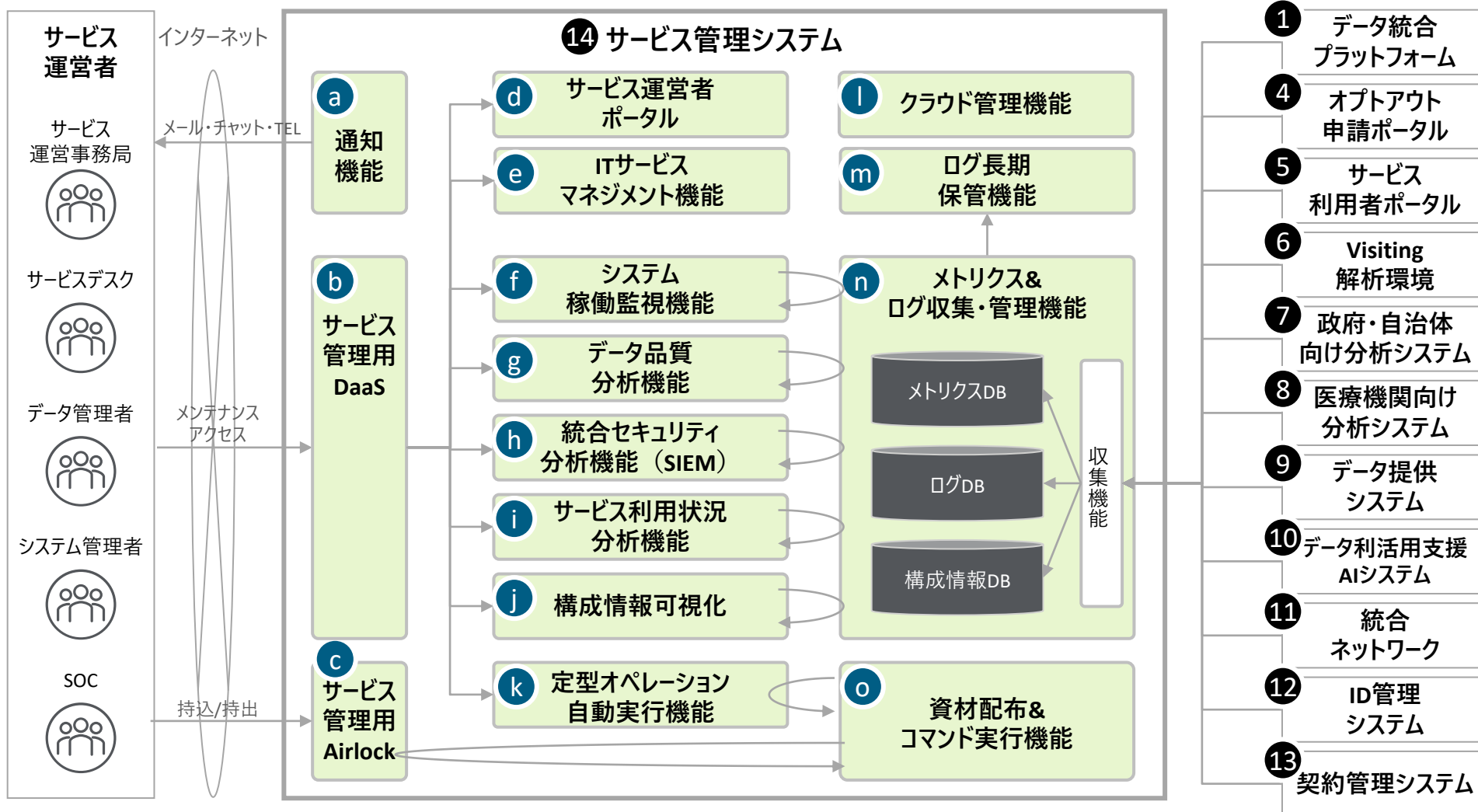


13 契約管理システム－機能概要説明

	機能概要
a. サービスメニュー管理機能	<ul style="list-style-type: none">サービス利用ユーザーが選択するサービスメニュー・サービスの説明・サービスレベル等を作成・管理する機能作成したサービスメニューを「⑤サービス利用者ポータル」に連携して、サービス利用ユーザーがメニューから利用したいサービスを申請できるようなユーザー体験を実現する
b. 課金プラン管理機能	<ul style="list-style-type: none">サービスメニューに対して課金プランを設定・管理することができる機能定額課金・固定従量課金・段階変動従量課金などに対応することを想定
c. 顧客管理機能	<ul style="list-style-type: none">サービス利用ユーザー（契約者）の情報を管理する機能「⑫ ID管理システム」と連携して必要なユーザー属性情報を受け取る
d. 契約管理機能	<ul style="list-style-type: none">サービス利用ユーザーが契約したサービスメニュー・課金プラン・期間等の情報を登録・管理することができる機能過去の契約履歴を保持して参照することも可能とする
e. 請求管理機能	<ul style="list-style-type: none">契約情報に対して請求書を作成して請求情報を管理する機能請求書のテンプレートの作成機能も具備する
f. 回収管理機能	<ul style="list-style-type: none">金融機関等の収納情報に基づいて、各請求に対する消込が行える機能

14 サービス管理システム — 概要構成

- サービス管理システムは、本サービスを安定的にかつ効率良く運営するための機能を提供する
- サービス運営者はインターネット経由でサービス管理用DaaSを用いて、各種サービス管理のための機能にアクセスする
- 同様にセキュリティパッチ等の資材持込や、本番環境ログの持出等のために、サービス管理用のAirlockも具備する



14 サービス管理システム — 機能概要説明（1/2）

	機能概要
a. 通知機能	<ul style="list-style-type: none"> サービス障害の検知時や、不正ログインの検知時等、サービス運営者による対応が求められるインシデントが発生した場合に、通知を行う機能 通知方法として、メール通知、チャット投稿、電話通知を行う機能を具備する
b. サービス管理用DaaS	<ul style="list-style-type: none"> サービス運営者はリモートによる対応を行う想定とし、安全に本番環境アクセスが可能なサービス運営者用のDaaSを具備する サービス運営に関連するすべてのシステムは、サービス管理用DaaSからのみアクセス可能のようにネットワークレベルで制限する
c. サービス管理用Airlock	<ul style="list-style-type: none"> サービス管理のために必要なデータやモジュールの持ち込み（例：セキュリティパッチ等のモジュール）や、データの持ち出し（例：ベンダーに送付するログ等）を安全に行うための機能 サービス管理者の承認を得た資材のみが、持込／持出可能のように制御する
d. サービス運営者ポータル	<ul style="list-style-type: none"> サービス運営者向けの周知情報や、各人の対応タスク、各システムへのリンク等、サービス運営者が効率的に運営業務を行うための情報集約を行うポータル機能
e. ITサービスマネジメント機能	<ul style="list-style-type: none"> インシデント管理、問題管理、変更管理、リリース管理等、ITサービスマネジメントプロセスを適切に管理・記録し、運営者同士で必要情報を連携するための機能 チケット管理ツールを想定
f. システム稼働監視機能	<ul style="list-style-type: none"> 本サービスを構成するシステムの稼働状況やリソース使用状況の監視を行う機能 過去の稼働状況やリソース使用状況の時系列グラフを提供し、ある断面での状況以外に過去から傾向の把握や、将来予測が行える機能を具備する
g. データ品質分析機能	<ul style="list-style-type: none"> 収集したデータの品質状況（データ件数、欠損率、変換エラー率、等）を可視化してデータ品質を分析する機能 データパイプラインの実行状況や所要時間の可視化も可能とする
h. 統合セキュリティ分析機能（SIEM）	<ul style="list-style-type: none"> ネットワークログ、アクセスログ、API実行ログ、構成変更イベント等、各種ログをリアルタイムで相関分析して、セキュリティインシデントを自動的に発見してアラート発出する機能 SOCが本機能を利用してセキュリティインシデントの検知・対応・復旧までを行う

14 サービス管理システム — 機能概要説明（2/2）

	機能概要
i. サービス利用状況分析機能	<ul style="list-style-type: none"> ポータルやVisiting解析環境へのユーザアクセス数や、データプラットフォームへのSQL実行回数、データ取得量、スループット等、サービス利用状況を分析する機能 分析結果を踏まえ、サービス改善やキャパシティプランニングに活用する
j. 構成情報可視化&差分検知機能	<ul style="list-style-type: none"> 構成情報の収集管理と、構成変更履歴の管理、事前定義したルールに基づく構成変更時のアラート通知機能を具備する サービス運営者や外部からの侵入者による不正な構成変更によるセキュリティリスクを防止することを目的としている また、本サービスを利用しているプログラムに脆弱性が見つかった場合に、対応範囲を迅速に特定する際にも利用する
k. 定型オペレーション自動実行機能	<ul style="list-style-type: none"> サービス運営時で実施が必要な定型オペレーションを自動化する機能 オペレーション手順をランブックとして保存しておき、オペレーションの効率化・自動化を可能とする スケジュール登録したオペレーション手順（ランブック）を自動実行する機能の他、アドホックに実行することも可能とする
l. クラウド管理機能	<ul style="list-style-type: none"> クラウドのガードレール機能の設定による故意または予期せぬ変更によるセキュリティホール発生の抑止を実現する 利用クラウドへのオペレーションコンソールを提供する クラウドの利用料の分析機能を提供する
m. ログ長期保管機能	<ul style="list-style-type: none"> 監査証跡として1年以上の長期保管が求められるログを収集して保管する機能 クラウドストレージ利用料の削減のため、一定の保存期間が経過した後に自動的に長期アーカイブに移動する機能や、所定の保管期間が経過後に自動でログを削除する機能を具備する
m. メトリクス&ログ収集・管理機能	<ul style="list-style-type: none"> 本サービスを構成する全てのシステムからメトリクス情報やログ情報、構成情報をリアルタイムに収集してデータベースに蓄積する機能 蓄積データは、「f. システム稼働監視」や「h. 統合セキュリティ分析」等、サービスを安定運用するための各種解析と監視に活用される 収集したログの内、証跡として長期保管が必要なものは、「m. ログ長期保管機能」に連携する
o. 資材配布&コマンド実行機能	<ul style="list-style-type: none"> 本システムで稼働するサーバに、セキュリティパッチ等のモジュール配布やリモートからの一括コマンドを実行可能にする機能 サーバ台数が増大した際のメンテナンス作業を効率化するための機能 PaaS/SaaSの活用により本システムで稼働するサーバ台数が少ない場合は、費用対効果を踏まえ実装しないことも想定している

2. 機能要件

2.1. 機能に関する事項（情報連携基盤）

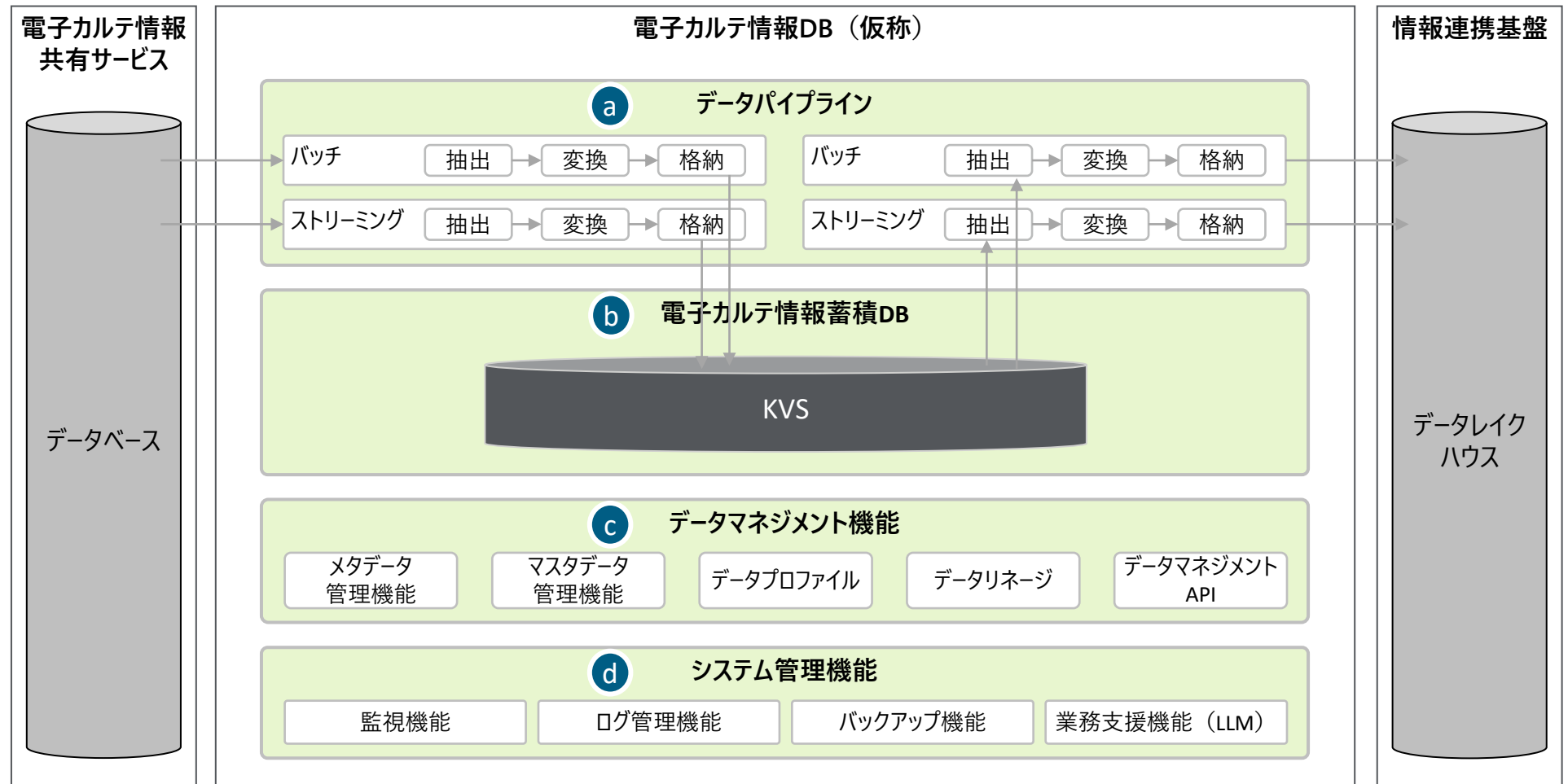
2.2. 機能に関する事項（電子カルテ情報DB）

2.3. データに関する事項

2.4. インターフェースに関する事項

電子カルテ情報データベース（仮称）のシステム構成は下記の通り バッチまたはリアルタイムで電子カルテ情報共有サービスのデータを収集・蓄積・提供する

構築対象システムの構成 — ②電子カルテ情報データベース（仮称）



2.2. 電子カルテ情報データベース（仮称） — 機能概要説明

	機能概要
a. データパイプライン	<ul style="list-style-type: none">「電子カルテ情報共有サービス」に蓄積された3文書6情報を収集・変換し、「b. 電子カルテ情報蓄積DB」に格納する機能及び、「b. 電子カルテ情報蓄積DB」に格納されたデータを情報連携基盤のデータレイクハウスに格納する機能を具備する変換処理では、ID5による仮名化処理等を行う想定データをリアルタイムに収集・連携するストリーミングパイプラインと、定期的にバッチ処理を行うバッチパイプラインの2種類を具備する
b. 電子カルテ情報蓄積DB	<ul style="list-style-type: none">データパイプラインが収集した、「電子カルテ情報共有サービス」に蓄積された3文書6情報を長期蓄積する機能3文書6情報はFHIR準拠したJSON形式のデータ構造となっているため、データベースとしてはKVS（Key Value Store）を想定している
c. データマネジメント機能	<ul style="list-style-type: none">データマネジメントに必要な、メタデータ管理、マスタデータ管理、データプロフィール（データ品質分析・可視化）、データリネージの機能を提供する機能データマネジメントに係る機能をAPI経由で操作（参照・作成・変更・削除）可能にする
d. システム管理機能	<ul style="list-style-type: none">本システムを運用するための監視機能、ログ管理機能、バックアップ機能を提供する業務効率化のためのLLMの活用を検討する

2. 機能要件

2.1. 機能に関する事項（情報連携基盤）

2.2. 機能に関する事項（電子カルテ情報DB）

2.3. データに関する事項

2.4. インターフェースに関する事項

情報連携基盤がデータを収集する各公的DBの一覧と、各公的DBが保有するデータは下記の通り

各公的DBの保有データ

請求系データベース				臨床データ系データベース	検診系データベース
NDB	介護DB	DPCDB	障害福祉DB	電子カルテ情報DB (仮称)	自治体検診医療機関データベース
<ul style="list-style-type: none"> 傷病名（レセプト病名） 投薬 健診結果 等 	<ul style="list-style-type: none"> 介護サービスの種類 要介護認定区分 ADL情報等 	<ul style="list-style-type: none"> 傷病名・病態等 施設情報 等 	<ul style="list-style-type: none"> 障害の種類 障害の程度 障害支援区分認定データ 障害福祉レセプト 等 	【3文書】 <ul style="list-style-type: none"> 診療情報提供書 退院サマリー 健康診断結果報告書 【6情報】 <ul style="list-style-type: none"> アレルギー 感染症 薬剤禁忌 検査 処方 傷病名 	<ul style="list-style-type: none"> がん検診 骨粗鬆症検診 歯周疾患検診 肝炎ウイルス検診等

疾患系データベース				
全国がん登録DB	難病DB	小慢DB	感染症DB	予防接種DB
<ul style="list-style-type: none"> がんの罹患 診療内容 転帰 等 	<ul style="list-style-type: none"> 告示病名 生活状況 各種検査値 等 	<ul style="list-style-type: none"> 告示病名 発症年齢 各種検査値 等 	<ul style="list-style-type: none"> 感染症の名称・症状 診断方法 初診年月日・診断年月日 発病推定年月日等 	<ul style="list-style-type: none"> ワクチン情報 接種場所 副反応の症状 等

各公的DBの運用開始時期、データ収集に関する仕様、データ保管仕様は下記の通り

各公的DBの仕様（概要）

項目		NDB	介護DB	DPCDB	iDB	難病・小慢DB	全国がん登録DB	
運用開始時期		2009年	2013年	2017年	2024年	2019年	2016年	
データ 収集	収集データ内容	レセプト、特定健診情報、死亡情報	介護レセプト、要介護認定情報、LIFE情報	DPCデータ	発生届情報等	臨床調査個人票、医療意見書	がんの罹患等の情報、死亡者情報票	
	収集元・収集方法	医療機関・保険者の情報を審査支払機関で匿名化したうえで収集	市町村・事業所等の情報を匿名化したうえで収集	医療機関から調査事業受託者を經由して収集(医療機関から頭名で受領したデータを調査事業受託者が匿名化の上DBに格納)	都道府県等から報告された頭名情報を格納	・医療機関から医師が臨個票、意見書登録 ・患者が自治体に認定申請し、自治体から認定結果・同意情報を登録	医療機関、都道府県からがん登録システムを通じてオンラインで情報収集(一部CD,USBメモリを使用して収集する場合もあり)	
	収集頻度	1か月に1回 レセプト:1回/月 特定健診情報:1回/年 死亡票：1回/年※2024年～収集予定	1か月に1回	3か月に1回	現在格納されているCOVID-19については、感染症法上の位置づけが変更されたため情報の更新はない	随時	随時 (都道府県による入力)	
	クレンジング	審査支払機関経由で受領したデータをそのまま収集	受領したデータをそのまま収集	様式間の矛盾確認、異常値確認等を実施し複数回に渡り再提出・確認を実施	受領したデータをそのまま格納	受領したデータをそのまま収集	ロジカルチェック等のクレンジングを実施し、誤りは医療機関に確認を実施	
データ 保管	容量	総容量	72TB	10TB	240TB	72 GB程度	1.1TB	13.3TB
		年間あたり収集容量	4.5TB	0.08GB	27TB	保有データは2020年2月3日～2023年5月7日分のみ、今後新たな収集は予定していない	0.9TB	0.07TB
	形式		CSV,Parquet等	DB形式(PostgreSQL) ZIP形式(定型データセット)	CSV形式	Parquet形式	臨個票・意見書：PDF形式、RDB形式 (JSON型など)	DB形式
	保管期間		規定なし（2008年以降の蓄積データを保管）	定めなし	定めなし	定めなし	臨個票・意見書等は5年保存及び5年経過後破棄、本省の扱うデータは保存期限を設けない	診断日から100年

2. 機能要件

2.1. 機能に関する事項（情報連携基盤）

2.2. 機能に関する事項（電子カルテ情報DB）

2.3. データに関する事項

2.4. インターフェースに関する事項

見積もり対象のパイプライン本数は350本とする。算出ロジックは下記の通り

開発対象のデータパイプライン本数（見積り前提値）

	見積前提数	見積前提
A) データソースの種類	<ul style="list-style-type: none"> 70種類 	<ul style="list-style-type: none"> NDBを7種類（外来レセ、入院レセ、DPCRレセ、調剤レセ、歯科レセ、特定健診、特定保健指導を想定、各レセプトの提供形式（IR,REなどレコード識別情報）までは含めない） NDBを基準に1データベースあたり7種類×10=70種類
B) データを取り込むバッチパイプラインの本数（データソース → Bronzeレベル）	<ul style="list-style-type: none"> 70本 	<ul style="list-style-type: none"> データソースの構造をそのままデータレイクに保持する前提のため、データ取り込みパイプラインはデータソースの種類と同数とする
C) データを正規化するバッチパイプラインの本数（Bronze → Silverレベル）	<ul style="list-style-type: none"> 70本 	<ul style="list-style-type: none"> 正規化処理もデータソースの構造をそのままに、主に項目値のバリデーションや変換を行う処理の前提とし、パイプラインはデータソースの種類と同数とする
D) データをビジネス観点で最適化するバッチパイプラインの本数（Silver → Gold）	<ul style="list-style-type: none"> 70本 	<ul style="list-style-type: none"> 現時点でGoldのデータ構造は想定されていない 見積もり前提としては、Silverのデータ種類と同数のGoldデータを作成するものとし、パイプラインの本数をデータソースの種類と同数とする
E) リアルタイムパイプラインの本数	<ul style="list-style-type: none"> 9本 	<ul style="list-style-type: none"> 3文書6情報（9種類）を全てリアルタイムで処理する前提
F) バッファ（開発対象パイプラインの増加見込み数）	<ul style="list-style-type: none"> 150% 	<ul style="list-style-type: none"> 現時点で見積り前提数が粗いため、多めにバッファ数を見込む
パイプライン開発対象合計本数	<ul style="list-style-type: none"> 350本 	<ul style="list-style-type: none"> $(B + C + D + E) \times F = 328.5 \div 350$

3. 非機能要件

情報提供を依頼するにあたり必要となる、情報連携基盤及び電子カルテDB（仮称）の非機能要件について、IPAの非機能要求グレードを元に前提事項を定義する

No.	非機能	概要	本情報提供依頼における非機能用要求の前提事項
1	可用性	システムがどれだけの時間、正常に稼働するかを示す指標	<ul style="list-style-type: none"> 運用スケジュール：運用時間、計画停止有無 業務継続性:対象業務範囲、サービス切替時間、業務継続要求度 目標復旧水準:RPO（復旧地点）、RTO（復旧時間）、RLO（復旧レベル）、大規模災害時再開目標 稼働率 可用性確認
2	性能・拡張性	システムが要求されるパフォーマンスを満たし、将来的な拡張に対応できるかを示す指標	<ul style="list-style-type: none"> 通常時業務量:ユーザ数、同時アクセス数、データ量、オンラインリクエスト件数、バッチ処理件数 業務増大度:通常時業務量の各項目増大率 保管期間 レスポンス（通常時/ピーク時）:オンラインレスポンス順守率、バッチレスポンス順守率 リソース拡張性（利用率/拡張性）:CPU、メモリ
3	運用・保守性	システムの運用や保守がどれだけ効率的に行えるかを示す指標	<ul style="list-style-type: none"> 運用時間:通常、特定日 バックアップ:外部データの利用可否、利用範囲、自動化範囲、取得間隔、保存期間 運用監視:監視情報、間隔 保守運用:計画停止の有無、保守作業自動化範囲 運用環境:環境有無（開発/試験用）、マニュアル準備レベル、リモートオペレーション、外部システム接続有無 サポート体制:保守契約（ハードウェア/ソフトウェア）範囲、ライフサイクル期間 その他:内部統制対応有無、サービスデスク設置有無
4	移行性	既存システムから新システムへの移行がどれだけスムーズに行えるかを示す	<ul style="list-style-type: none"> 移行スケジュール:システム移行期間、システム停止可能日時、並行稼働有無 システム展開方式:拠点展開ステップ数、業務展開ステップ数 移行対象:設備・機器の移行内容、移行データ量、移行データ形式
5	セキュリティ	システムのセキュリティ対策がどれだけ強固であるかを示す	<ul style="list-style-type: none"> 前提・制約条件、セキュリティリスク分析/診断:社内規定・ルール、リスク分析範囲、ネットワーク/Web診断有無 アクセス・利用制限:認証機能、利用制限 データ暗号化:伝送データ暗号化有無、蓄積データ暗号化有無 不正監視:ログ取得、ログ保管期間、不正監視対象 各種対策:ネットワーク対策、マルウェア対策、Web対策
6	システム環境・エコロジー	システム特性とシステムが環境に与える影響やエネルギー効率を示します	<ul style="list-style-type: none"> 制約条件:構築時制約条件、運用時制約条件 システム特性:ユーザ数、クライアント数、拠点数、展開地域、製品指定有無 適合企画:製品安全規格取得有無、環境保護規格取得の有無 機材設置環境:耐震/免震、設置スペース

【出典】 [システム構築の上流工程強化（非機能要求グレード）紹介ページ | アーカイブ | IPA 独立行政法人 情報処理推進機構](#) 非機能要求グレード2018本体（日本語版）

情報連携基盤・電子カルテ情報DB（仮称）で目指す非機能要求レベルは下記の通り データの2次利用基盤として適切な目標レベルとしつつ、一部システムは災害対策も行う

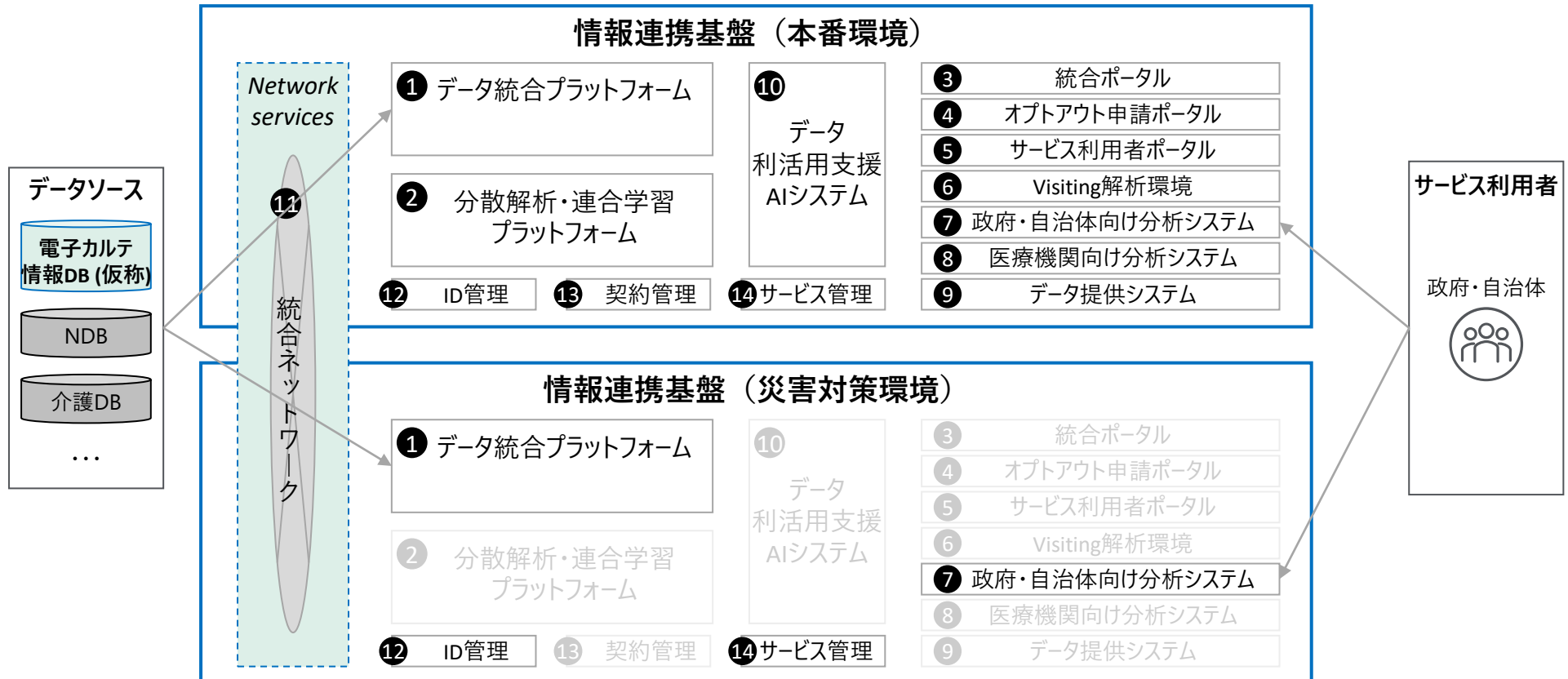
情報連携基盤・電子カルテ情報DB（仮称）で目指す非機能要求レベル

No.	非機能	主要メトリクス	非機能要件レベル
1	可用性	業務継続性・目標復旧水準	<ul style="list-style-type: none"> フェイルオーバー（ホットスタンバイ） 冗長化：地理的分散
		稼働率	<ul style="list-style-type: none"> 99.9%（年間8-9時間程度の停止）
		災害対策	<ul style="list-style-type: none"> 災害時に利用を想定する政府・自治体向け分析システム: 災害対策環境を具備し1日以内に切り替えを行う 上記以外のシステム: データ損失が無いようにバックアップを転送するが、災対環境は具備しない
2	性能・拡張性	レスポンス	<ul style="list-style-type: none"> オンライン応答時間：1秒～3以内。サービス利用者のユーザビリティを考慮して適切な目標時間を設定する スループット：1,000リクエスト/秒
		リソース拡張性	<ul style="list-style-type: none"> 必要に応じて自動スケーリング必要
3	運用・保守性	運用時間	<ul style="list-style-type: none"> 24時間365日、計画停止有
		バックアップ、運用監視、保守運用	<ul style="list-style-type: none"> 対象範囲：システム全体 自動化：完全自動化 監視間隔：秒～分単位監視 バックアップ間隔：日次
		運用環境	<ul style="list-style-type: none"> マニュアル準備レベル：詳細マニュアルおよびトレーニング資料
4	移行性	システム展開方式	<ul style="list-style-type: none"> 業務展開ステップ数：段階移行
5	セキュリティ	アクセス・利用制限	<ul style="list-style-type: none"> 多要素認証を必須とする
		データ暗号化	<ul style="list-style-type: none"> 伝送/蓄積データ共に暗号化有
		不正監視	<ul style="list-style-type: none"> リアルタイム監視、監査ログ取得
		SOC体制の具備	<ul style="list-style-type: none"> 具備する
6	システム環境・エコロジー	—	—

「政府・自治体向け分析システム」は、感染症の発生傾向の分析ダッシュボード等、災害時等の広域障害発生時においても利用継続が求められるため、災害対策環境を具備する

災害対策環境の考え方

- 災害対策環境には、政府・自治体向け分析システムに加え、データのリアルタイム更新に必要なデータ統合プラットフォーム、ユーザ認証機能を有するID管理システム、サービス運用のためのサービス管理システムを具備し、必要最低限の構成を取る
- 災対環境は平時は停止しておき、必要なデータを同期しておく。本番環境障害時に災対環境を立ち上げ利用を開始する（コールドスタンバイ方式）
- 災害対策環境の実現のためには、データソースとなる公的DBや電子カルテ情報DB等の対策も必要のため、今後の検討課題とする



情報連携基盤運用に要する非機能要件（重要項目抜粋）の考え方（1／6）

1. 可用性

	分類	項目	要件	考え方・背景
1	運用 スケジュール	運用時間	<ul style="list-style-type: none"> 24時間無停止 	<ul style="list-style-type: none"> 感染予防や危機管理対策への活用など、有事の際への利用を想定し、常時アクセス可能とすべきであるため
2		計画停止の有無	<ul style="list-style-type: none"> 計画停止有り 利用者の少ない時間帯での停止を想定 	<ul style="list-style-type: none"> 24/365無停止を必須とする性質のシステムではないため 但し、終日停止は1.運用時間に倣い想定しない
3	業務継続性	対象業務範囲 (稼働率の算出範囲)	<ul style="list-style-type: none"> システム毎に定義する 	<ul style="list-style-type: none"> システムごとに要求レベルが異なるため領域ごとに定義
4		サービス切替時間	<ul style="list-style-type: none"> バッチ連携処理:24時間未満 リアルタイム処理:60分未満 	<ul style="list-style-type: none"> バッチ連携は最頻でも1日に1回連携する想定のため、24時間以内に切り替われば業務影響無し リアルタイムは現時点では仮値として60分としている
5		業務継続の要求度	<ul style="list-style-type: none"> 単一障害時は業務停止を許容せず、処理を継続させる 	<ul style="list-style-type: none"> コスト最適な範囲で冗長化構成をとる サービス切替時間に合わせた運用想定
6	目標 復旧水準	RPO (復旧地点)	<ul style="list-style-type: none"> データプラットフォームのデータは日次でバックアップを取り1日前に戻せること Visiting解析環境の成果物はリアルタイムでレプリケーションし、データ損失を許容しない 	<ul style="list-style-type: none"> データプラットフォームの損失データはデータソースから再取得 Visiting解析環境の成果物はユーザビリティを優先しデータ損失（研究成果の損失）を許容しない
7		RTO (復旧時間)	<ul style="list-style-type: none"> 24時間以内 	<ul style="list-style-type: none"> データ二次利用のためミッションクリティカルではないと仮定し、1.運用時間に倣い、24時間以内の復旧を想定
8		RLO (復旧レベル)	<ul style="list-style-type: none"> システム毎に定義する 	<ul style="list-style-type: none"> 連携用ネットワークは、障害発生以降の未連携データを復旧後に連携する想定
9		大規模災害時 再開目標	<ul style="list-style-type: none"> 政府・自治体分析システムとその稼働に必要なシステムは1日以内。他は一週間以内 	<ul style="list-style-type: none"> 政府・自治体向け分析システムは感染症動向分析等、有事の際の利用も想定しているため
10	稼働率	稼働率	<ul style="list-style-type: none"> 99.9% 	<ul style="list-style-type: none"> 1年間で8-9時間程度の停止を許容する前提

情報連携基盤運用に要する非機能要件（重要項目抜粋）の考え方（2 / 6）

2. 性能・拡張性

	分類	項目	要件	考え方・背景
1	通常時 業務量	ユーザ数	<ul style="list-style-type: none"> 不特定多数のユーザが利用 	<ul style="list-style-type: none"> R10開始当初は最大100プロジェクト×5人＝500人が利用する想定とする
2		データ量	<ul style="list-style-type: none"> 主要なデータ量のみが明確である 	<ul style="list-style-type: none"> 稼働当初は現在の336TB＋120TB（年間あたりデータ収集容量（32.5TB/年）×稼働までの4年で蓄積）で約450TBと算定 年間あたりデータ収集容量（32.5TB/年）×20年（以降、データ圧縮し保管）で650TB超と算定
3	業務増大度	ユーザ数	<ul style="list-style-type: none"> 不特定多数のユーザが利用 	<ul style="list-style-type: none"> R17（リリースから7年後）に1,000プロジェクト×5人＝5,000人が利用する想定とする R10の500人から毎年増して行く想定
4		データ量	<ul style="list-style-type: none"> 主要なデータ量のみが明確である 	<ul style="list-style-type: none"> 連携元データ種の追加等により、3年後目安に20年（以降、データ圧縮し保管で650TB超）×1.2倍を見込む
5	保管期間	保管期間	<ul style="list-style-type: none"> 20年（以降、データ圧縮し保管） 	<ul style="list-style-type: none"> 疾病の要因分析では10年以上前のデータを利用する場合があること、また海外の情報連携基盤では20年と設定しているが、蓄積データの廃棄は行わないため

情報連携基盤運用に要する非機能要件（重要項目抜粋）の考え方（3／6）

3. 運用・保守性

	分類	項目	要件	考え方・背景
1	運用時間	運用時間	<ul style="list-style-type: none"> 24時間無停止（1.可用性 1.運用時間同様） 	<ul style="list-style-type: none"> 感染予防や危機管理対策への活用など、有事の際への利用を想定し、常時アクセス可能とすべきであるため
2	バックアップ	外部データの利用可否	<ul style="list-style-type: none"> 外部データは利用できない 	<ul style="list-style-type: none"> 情報連携基盤の連携元データは存在するものの、外部データの加工データ含めた情報が主となるため
3		利用範囲	<ul style="list-style-type: none"> データの長期保存（アーカイブ） 	—
4		取得間隔	<ul style="list-style-type: none"> 日次で取得（データプラットフォーム部分） レプリケーション（Visiting解析環境） 	<ul style="list-style-type: none"> 可用性要件に合わせて設定
5	運用監視	監視情報	<ul style="list-style-type: none"> リソース監視を行う 	<ul style="list-style-type: none"> システム特性より、エラー監視のみならず、障害未然防止含む監視を行い、安定運用する必要があるため
6		間隔	<ul style="list-style-type: none"> リアルタイム監視（分間隔） 	<ul style="list-style-type: none"> システム特性より、自動リアルタイム監視必要 コスト最適化を踏まえ、分間隔としている
7	運用環境	マニュアル準備レベル	<ul style="list-style-type: none"> システムの通常運用と保守運用のマニュアルを提供する 	<ul style="list-style-type: none"> 将来にわたる運用のため、整備が必要
8	その他	サービスデスク設置有無	<ul style="list-style-type: none"> 新規設置予定 	<ul style="list-style-type: none"> 流用できる体制が無い想定のため

情報連携基盤運用に要する非機能要件（重要項目抜粋）の考え方（4 / 6）

4. 移行性

	分類	項目	要件	考え方・背景
1	移行 スケジュール	システム移行期間	• 2年未満	• 全体スケジュール（案）に基づく
2		システム停止可能日時	• 利用の少ない時間帯（夜間など）	• 可用性に倣う
3		並行稼働有無	• 無し	• 新規システムのため
4	システム 展開方式	拠点展開ステップ数	• 単一拠点のため規定無し	• 情報連携基盤というシステム特性上、単一拠点となる想定
5		業務展開ステップ数	• 全業務一斉切り替え	• リリース計画のフィジビリティ、想定リスクの観点で適切な展開計画をご提案頂きたい

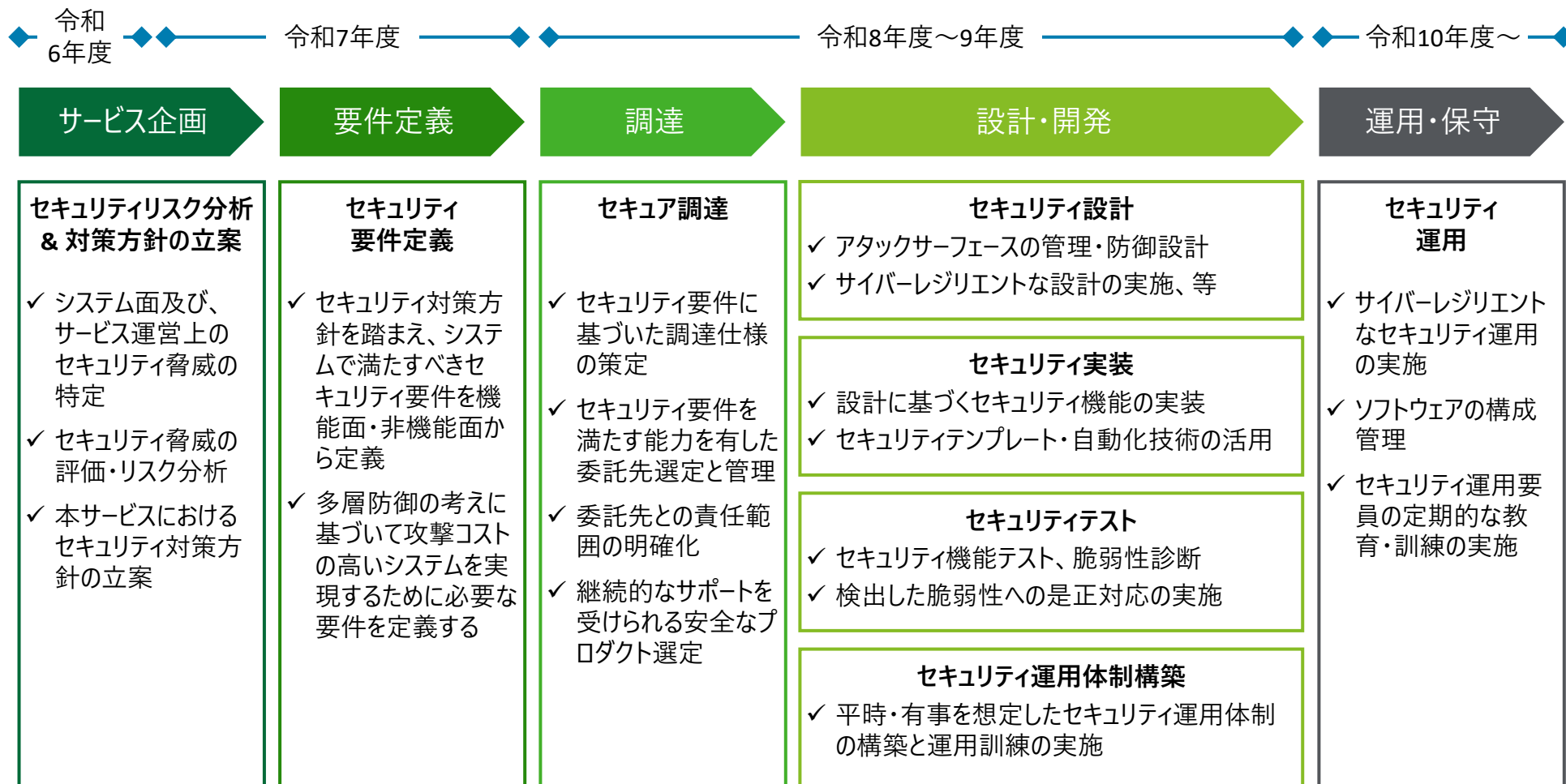
情報連携基盤運用に要する非機能要件（重要項目抜粋）の考え方（5 / 6）

5. セキュリティ

	分類	項目	要件	考え方・背景
1	前提・制約条件	社内規定・ルール	<ul style="list-style-type: none"> 有り 	<ul style="list-style-type: none"> 順守すべき規程、ルール、法令、ガイドラインは別紙一覧参照 SOC(Security Operation Center) の監視が必要 Security by Designに即した工程計画・実行
2	アクセス・利用制限	認証機能	<ul style="list-style-type: none"> 複数回の認証 	<ul style="list-style-type: none"> 強固な認証方式(デジタル署名、多要素認証 等)を必要とする
		ゼロトラストセキュリティ	<ul style="list-style-type: none"> ゼロトラストに基づく認証認可機能 	<ul style="list-style-type: none"> 境界防御に加えよりセキュリティ強度を高めるため
3	データ暗号化	伝送データ暗号化有無	<ul style="list-style-type: none"> 重要情報を暗号化 	<ul style="list-style-type: none"> 管理する情報の性質上、重要情報は暗号化を前提
4		蓄積データ暗号化有無	<ul style="list-style-type: none"> 重要情報を暗号化 	<ul style="list-style-type: none"> 管理する情報の性質上、重要情報は暗号化を前提
5	不正監視	ログ取得	<ul style="list-style-type: none"> 実施する 	—
6		ログ保管期間	<ul style="list-style-type: none"> 5年 	<ul style="list-style-type: none"> 一般的に5年間保管が多いため
7		不正監視対象	<ul style="list-style-type: none"> システム全体 	<ul style="list-style-type: none"> クラウド利用前提のため コスト最適化観点でご提案頂ければよい

大規模な仮名化データを扱うプラットフォームサービスとして、「セキュリティ・バイ・デザイン」の考え方をシステム開発に取り入れ、セキュリティ担保とデータ提供者への説明責任を果たす

5. セキュリティーセキュリティ品質の作り込みとセキュリティ運用



情報連携基盤および電子カルテ情報DB（仮称）のセキュリティ要件を定義する際には、下記のような関連する規定・ガイドラインの遵守事項も取り込むこととする

5. セキュリティー 情報連携基盤運用に要するセキュリティ要件前提条件となる規定・ガイドライン

#	文書名	発表
1	医療情報システムの安全管理に関するガイドライン第6版	令和5年5月厚生労働省
2	医療情報を取り扱う情報システム・サービスの提供事業者における安全管理ガイドライン第1.1版	令和5年7月総務省・経済産業省
3	政府機関等の情報セキュリティ対策のための統一基準群（令和5年度版）	令和3年7月7日、内閣官房 内閣サイバーセキュリティセンター
4	デジタル・ガバメント推進標準ガイドライン	2024年5月31日各府省情報化 統括責任者(CIO) 連絡会議決定

情報連携基盤運用に要する非機能要件（重要項目抜粋）の考え方（6／6）

6. システム環境・エコロジー

	分類	項目	要件	考え方・背景
1	制約条件	構築時制約条件	・ 制約有り(全ての制約を適用)	・ 法令、条例などの制約に基づき適切に構築する
2		運用時制約条件	・ 制約有り(全ての制約を適用)	・ 法令、条例などの制約に基づき適切に構築する
3	システム特性	ユーザ数	・ 不特定多数のユーザが利用 （2.性能・拡張性 1.ユーザ数同様）	・ システムに習熟していないユーザも容易に操作を習得 できるようなUI設計やマニュアル準備を行う
4		拠点数	・ 単一拠点	・ 情報連携基盤というシステム特性上、単一拠点となる 想定
5		展開地域	・ 国内	－

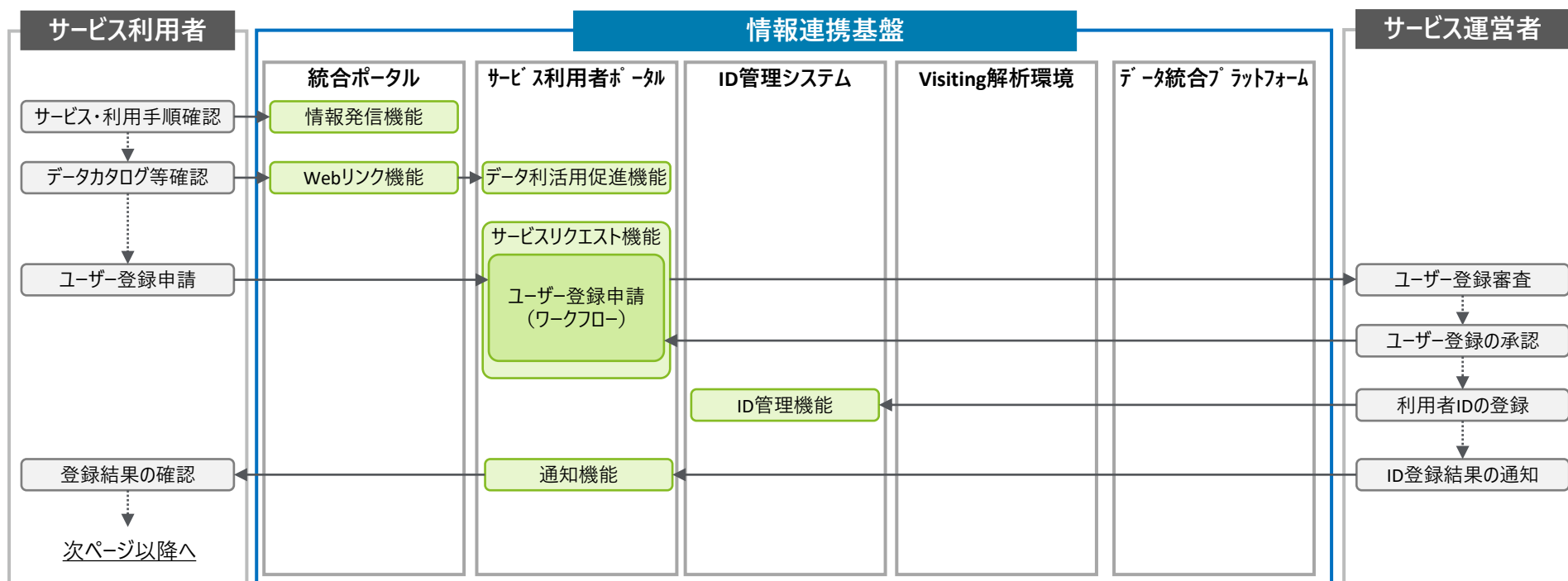
Appendix-1. サービス利用フロー詳細

サービス利用者を初期登録する場合のフロー概観は下記の通り

サービス・機能利用フロー（サービス利用者の初期登録）

: 機能 \longrightarrow : 要求・アクセス
 : 運用 $\cdots\cdots\longrightarrow$: 運用フロー

- サービス利用者の認証、トレーサビリティ確保、申請ワークフローのデジタル化等のため、情報連携基盤の全てのサービス利用者（政府・自治体、医療機関、国立機関・研究機関、製薬・医療機器会社等）は、利用開始時にユーザー登録する
- ユーザー登録では、ユーザーの属性情報（政府・自治体関係者、医療機関職員等）を管理し、利用可能なサービスを特定する

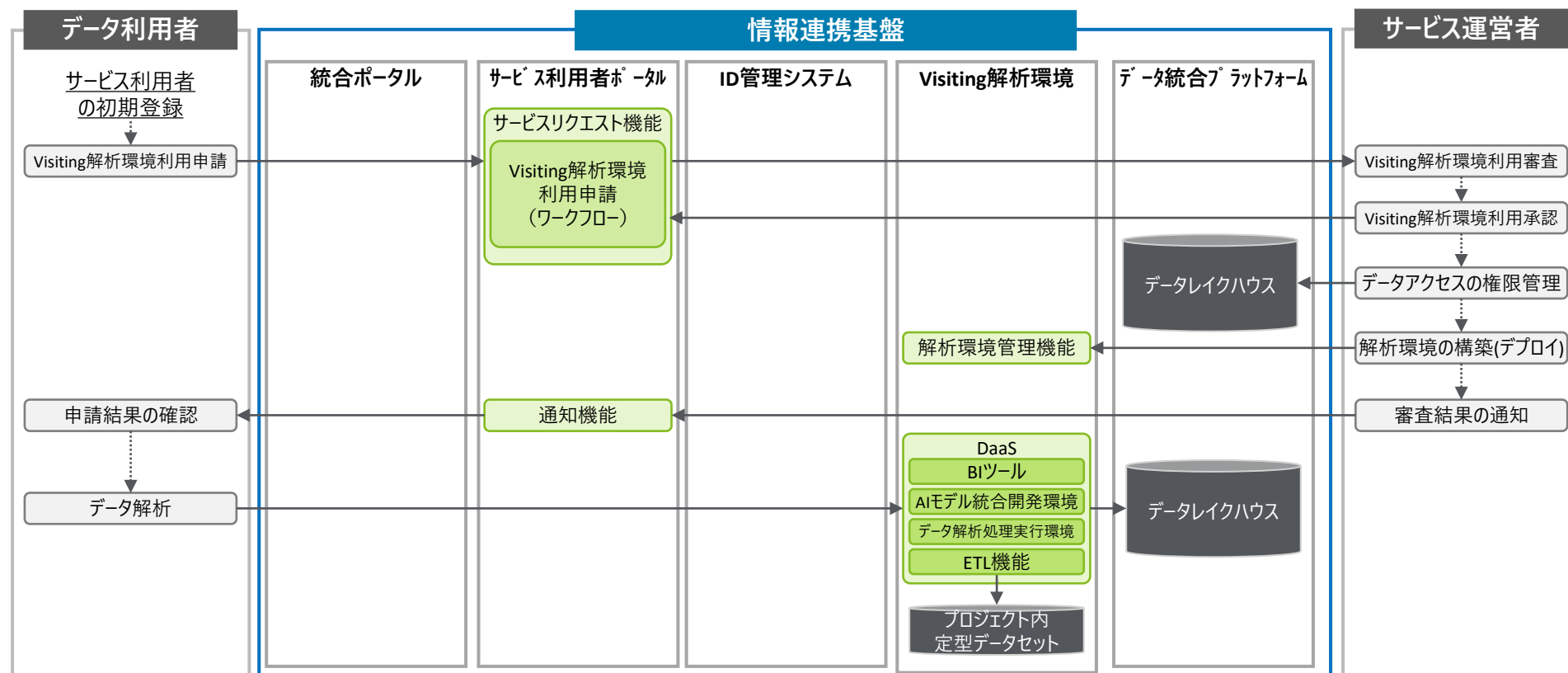


データ利用者がVisiting解析環境を利用する場合のフロー概観は下記の通り

サービス・機能利用フロー（Visiting解析環境利用）

 ：機能 \longrightarrow ：要求・アクセス
 ：運用 $\cdots\cdots\longrightarrow$ ：運用フロー

- データ利用者がVisiting解析環境を利用する場合は、Visiting解析環境の利用審査を行い、承認後にサービス運営者がデータ利用者専用の解析環境を構築（デプロイ）する
- データ利用者はサービス運営者が構築したDaaSからWebブラウザ等でBIツール、AIモデル統合開発環境等を利用してデータレイクハウスのデータを用いた分析を行うと共に、パフォーマンス確保のためにETL機能でプロジェクト内データセットの作成も可能とする

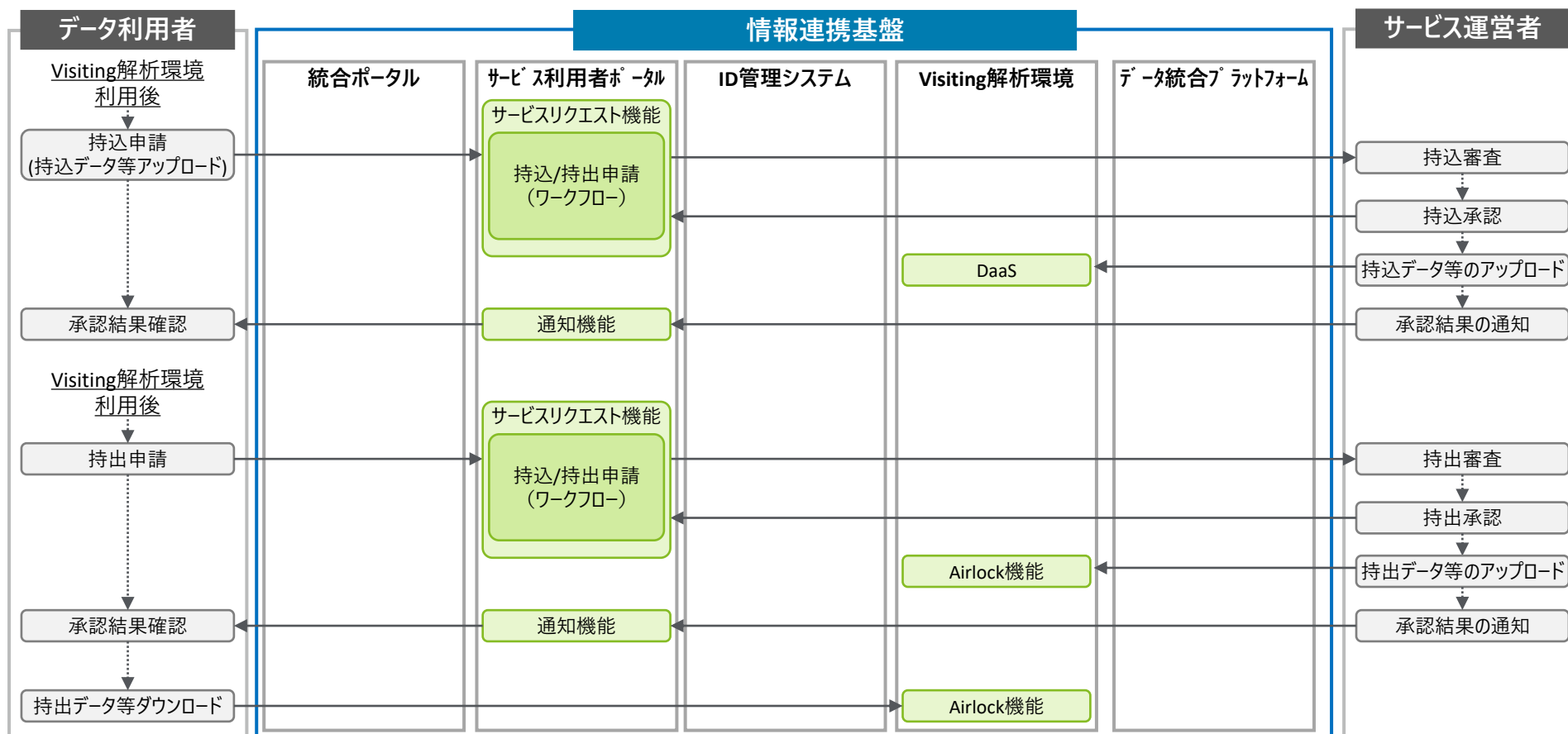


データ利用者がVisiting解析環境でデータやツール等の持出/持込を行う場合のフロー概観は下記の通り

サービス・機能利用フロー（データやツール等の持込/持出）

: 機能 \longrightarrow : 要求・アクセス
 : 運用 $\cdots\cdots\longrightarrow$: 運用フロー

- データ利用者がデータやツールの持込を行う場合は、データ利用者が持ち込みたいデータ・ツールをサービス利用者ポータルへアップロードし、サービス運営者が審査・承認後にDaaS環境にアップロードする
- データ利用者が成果物の持出を行う場合は、データ利用者が持ち出したい成果物をサービス利用者ポータルから申請し、サービス運営者が審査・承認後にAirlock機能へ当該成果物をアップロードする。データ利用者はAirlock機能を利用して成果物をダウンロードする



データ利用者がデータやツール等を持込／持出する場合の補足

データやツール等の持込/持出のレビューに関する補足

- データやツール等の持込／持出の申請・審査は、ワークフローで管理する
- データの持込／持出に当たっては、ウイルスチェック等の基本的なセキュリティチェックを共通で実施した上で、持込／持出する内容に応じて個人情報保護、セキュリティリスク、技術的な提供可否等の観点からレビューし、承認後にデータ利用者-Visiting解析環境間で授受する

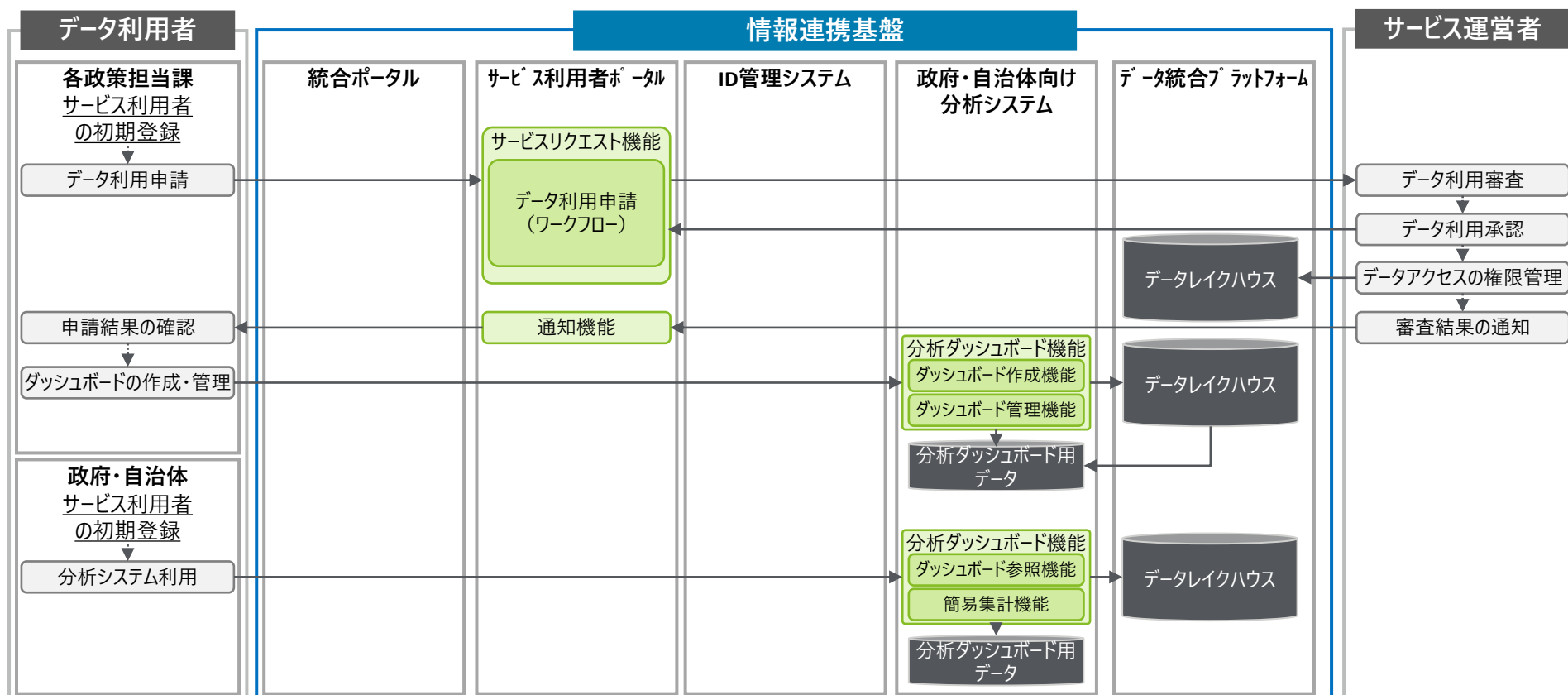
分類	区分	対象	レビュー観点
持込	データ	<ul style="list-style-type: none"> データ利用者が保有するデータ データ利用者が収集・作成したデータ 	<ul style="list-style-type: none"> ✓ 個人特定に繋がるようなデータは無いか ✓ データ量、データ形式（様式、圧縮等）がVisiting解析環境で取り扱い可能か
	ツール	<ul style="list-style-type: none"> データ利用者が利用したい解析用ソフトウェア データ利用者が開発した解析用スクリプト データ利用者が利用したい表計算、文書作成等の汎用ソフトウェア 	<ul style="list-style-type: none"> ✓ ホワイトリストに適合するソフトウェアか ✓ 不正なプログラムが含まれていないか ✓ Visiting解析環境の実行環境に悪影響を与えないか
持出	成果物	<ul style="list-style-type: none"> データ利用者がVisiting解析環境上で作成した解析用スクリプト データ利用者がVisiting解析環境上で解析した結果（集計表、グラフ等） データ利用者がVisiting解析環境上で作成したレポート・文書等 	<ul style="list-style-type: none"> ✓ 個人特定につながるような情報が含まれていないか ✓ 提供可能なデータ容量、データ形式か

政府・自治体が分析システムを利用する場合のフロー概観は下記の通り

サービス・機能利用フロー（政府・自治体向け分析システム利用時）

: 機能 \longrightarrow : 要求・アクセス
 : 運用 $\cdots\cdots\longrightarrow$: 運用フロー

- サービス運営者は、各政策担当課のデータ利用申請に基づき、データのアクセス権限を設定する
- 各政策担当課は、政府・自治体向け分析システム上でダッシュボードを作成・管理する。その際、性能面を考慮し、必要に応じて分析ダッシュボード用のデータ領域にデータを保管することができる
- 政府・自治体は、各政策担当課が作成した分析ダッシュボードや、簡易集計機能を使用してデータ分析を行うことができる（サービス利用者の初期登録時に利用者属性も併せてシステムに登録するため、政府・自治体向け分析システムの利用申請は不要とする）

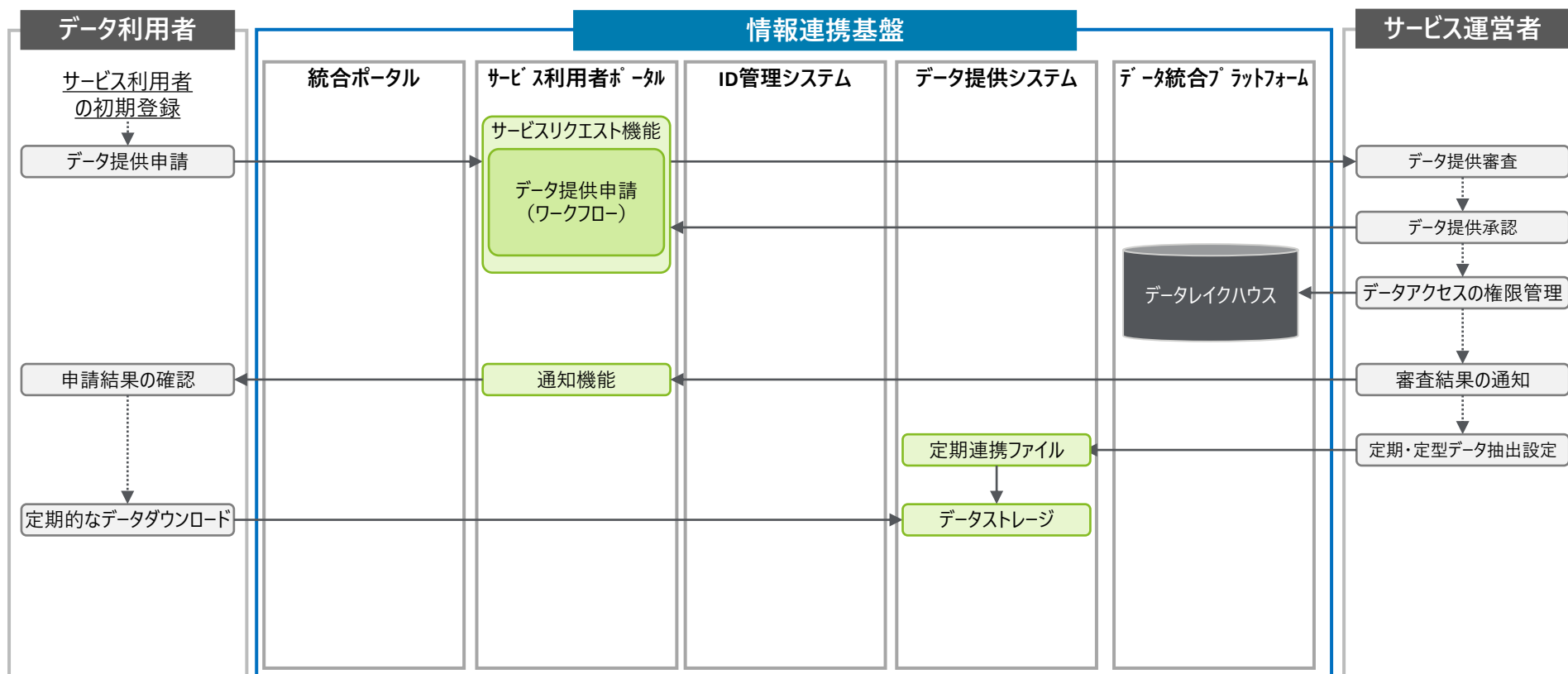


データ利用者がデータ提供サービスを利用する場合のフロー概観は下記の通り

サービス・機能利用フロー（データ提供サービス利用）

 ：機能 \longrightarrow ：要求・アクセス
 ：運用 $\cdots\cdots\longrightarrow$ ：運用フロー

- データ提供サービスは、特定の機関／法人が有するシステム環境に定期的に定型データを提供する
- サービス運営者は、データ提供サービスの利用申請に基づき、データ提供システムに定期ファイル配信を設定する。設定後、データ提供システムが自動でデータ提供システムに配信用ファイルを作成する
- データ利用者は、定期的にデータ提供システムにアクセスしてデータをダウンロードする



Appendix-2. ユースケース例

情報連携基盤サービスでは下記のユーザーによる利用を想定している

情報連携基盤サービスの利用ユーザーの種別とユースケース例

利用ユーザー		ユースケース例	
政府の政策担当者	EBPM による政策立案と検証・改善	生活習慣病予防対策推進	✓ 生活習慣病予防KPIの可視化 ✓ 生活習慣病発症リスクの予測
		重症患者診療体制整備事業	✓ 重症患者診療体制整備事業への活用
		予防接種データ分析	✓ 予防接種データの可視化および施策への活用
		疾病別(がん、難病等)医療費分析	✓ 疾病に関する医療費等の分析
		医療計画策定	✓ 5疾病6事業分析
		医薬品の安全対策	✓ 医薬品の安全性サーベイランス
		希少疾患	✓ 希少疾病患者の統計調査等
		介護データ分析	✓ 地域包括支援センター等におけるデータ共有 ✓ 高額介護給付費の削減
		感染予防や流行時の危機管理対策の強化	✓ 新型コロナ薬剤処方に係る公費支援の在り方の検討 ✓ 感染症罹患状況の可視化 ✓ 感染症のリアルタイムデータ分析（平時・有事）
		医療費適正化	✓ 医療費適正化計画（都道府県データブック） ✓ 地域医療計画 ✓ 行政施策の評価
自治体の政策担当者	地域医療の最適化	地域医療計画	✓ 再編統合等の医療資源の最適化
		行政施策の評価	✓ 感染症等のワクチン接種者数や感染者数の分析
学術・研究機関の研究者	研究・論文の増加	臨床研究	✓ 患者レベルの予測(発症リスク、薬剤の副作用、治療効果、再入院リスク等) ✓ 疾病のコホート分析 (治療法の効果、予防策の有効性、生存率分析等) ✓ LLMを活用した新規医療や診断補助アプリケーションの開発
		手法評価	✓ 研究方法論やデータ解析手法の開発、評価、改善
	国際共同研究の促進	各国の国際比較	✓ 連合学習を活用した医療・健康データの国際共同研究 ✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等 ✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析 ✓ 臨床情報の解析による患者ニーズの特定 ✓ 費用対効果の評価、既存薬剤との比較、分配の適正評価等
民間企業 (製薬)	治験、臨床試験、市販後調査等	医薬品の有効性・安全性評価 臨床情報の解析による市販後調査等、	✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等 ✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析 ✓ 臨床情報の解析による患者ニーズの特定 ✓ 費用対効果の評価、既存薬剤との比較、分配の適正評価等
民間企業 (医療機器)	治験、臨床試験、市販後調査等 AI機器開発の推進	医療機器の有効性・安全性評価 臨床情報の解析による市販後調査等、 AI搭載機器の機械学習	✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等 ✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析 ✓ AI搭載機器やそのための機械学習の研修開発
民間企業 (再生医療等製品)	治験、臨床試験、市販後調査等	再生医療等製品の有効性・安全性評価 臨床情報の解析による市販後調査等、	✓ 治験参加者及びコントロール群の要件検討、試験対照群としての評価等 ✓ 市販後の安全性評価 ✓ 上記に関するAIを活用した調査分析
医療機関	臨床教育、診断補助	過去症例の参照や診断補助等	✓ 過去症例にアクセスし、現実に即したケーススタディや教育 ✓ LLMを用いた診断補助など
患者・国民	医療・健康への理解向上	自身の健康増進や予防への理解促進	✓ 医療・健康等の情報参照等

公的DBの患者情報および人口動態データ等进行分析することで、地域ごとの医療需要やニーズを把握・予測し、地域医療資源の最適化を実現することができる

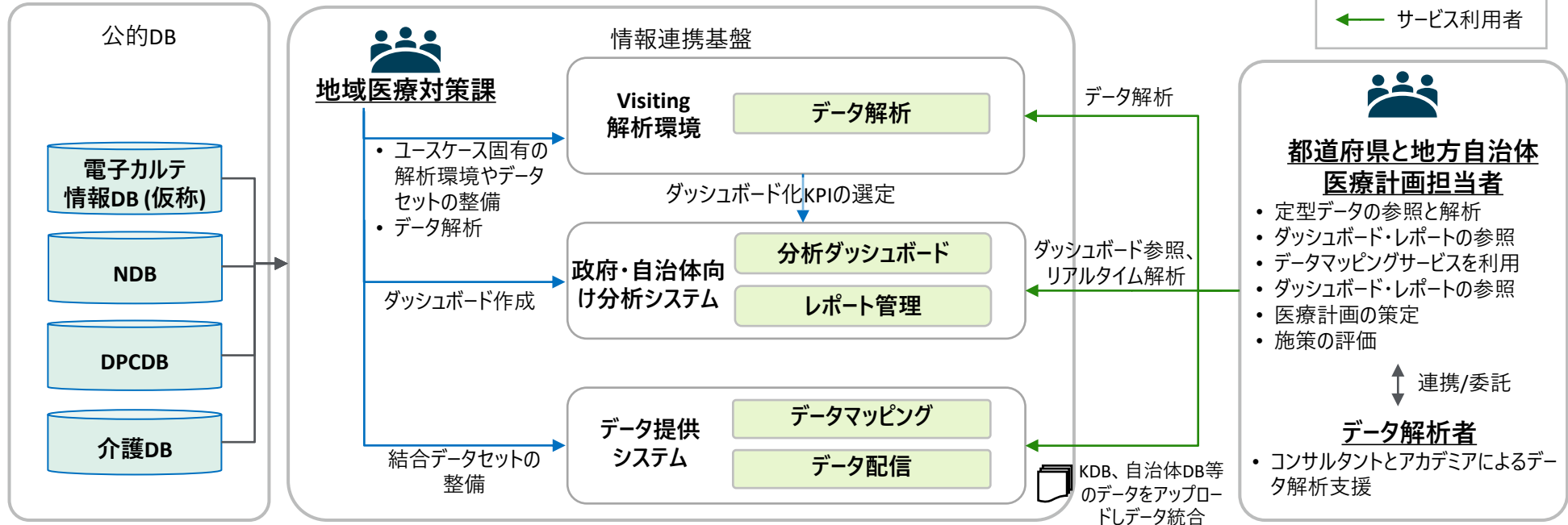
ユースケース：地域医療資源の最適化

ユーザー像	<ul style="list-style-type: none"> 厚生労働省医政局・地域医療対策課 都道府県と地方自治体の医療計画担当者
背景	<ul style="list-style-type: none"> 地域ごとの医療リソースの偏在が進んでおり、医療サービスの均等化が課題になる 高齢化や人口減少により、医療ニーズの変化が生じ、資源の最適配分が求められる
目的	地域ごと医療資源分布の偏在及び医療ニーズを可視化し、医療資源配分及び医療費を適正する

利用するDBとデータ項目

- 電子カルテ情報DB
 - ✓ 3文書6情報
- NDB:
 - ✓ 診療行為の内容、頻度、医療機関別の患者数、請求点数等
- DPCDB
 - ✓ 入院情報、退院情報、患者情報、診断情報等
- 介護DB
 - ✓ 要介護認定情報、介護レセプト情報等

<データフロー>



公的DBの患者情報および人口動態データ等进行分析することで、地域ごとの医療需要やニーズを把握・予測し、地域医療資源の最適化を実現することができる

ユースケース：地域医療資源の最適化

区分	ユースケース	概要	利用DB				機能		アウトプット
			NDB	感染症DB	DPCDB	電カルDB	ダッシュボード	Visiting解析環境	
現状把握	医療資源の地域分布状況及び格差・偏在分析	地域別医療資源マップを作成し、各地域における医療機関、医療従事者、医療機器などの資源分布を可視化し、医療資源の偏在を特定する	○		○		○		<ul style="list-style-type: none">医療資源分布レポート医療資源配分計画
	患者の流入・流出状況分析	患者の地域間移動や受診傾向を分析し、地域ごとの医療需要の充足度を評価できる患者流入・流出分析システムを構築し、必要な医療サービスの提供や地域医療の充実に役立てる情報と提供する	○		○		○	○	<ul style="list-style-type: none">患者流入・流出レポート地域間の需要ギャップ分析地域医療サービス改善提案
	地域別の疾病発生率と医療ニーズの分析	地域ごとの疾病発生率や患者属性（年齢、性別など）を分析し、地域別疾病ニーズを特定し、地域ごとに必要とされる医療サービスの分布を明確にする	○	○	○	○	○	○	<ul style="list-style-type: none">地域別の疾病分析レポート地域別の医療ニーズレポート医療サービス提供方針提案
改善対応	地域医療機関の役割分担と連携強化による医療提供体制の最適化	地域ごとに異なる医療ニーズや資源の分布に応じて、各医療機関の役割を明確にし、地域間での連携を強化することで、効率的な医療提供体制による資源偏在の緩和を実現する			○	○	○	○	<ul style="list-style-type: none">医療機関ごとの役割分担リスト医療資源調整計画
	地域の医療サービス供給やニーズに合わせた病院の再編統合	地域の医療ニーズや供給の状況に基づき、病院の再編や統合を実施し、効率的かつ持続可能な医療提供体制を構築する。各病院の診療報酬や医療収益も考慮し、財政面でも安定した運営が可能な医療リソース配分を実現する	○		○			○	<ul style="list-style-type: none">地域病院再編統合計画診療科・病床数の再配置計画診療報酬改善による財務シミュレーション
	地域別医療費の適正化	地域別に診療報酬や医療費の傾向を分析し、不均衡のある領域を特定する。不要な診療や検査を削減するための基礎データを提供し、地域別医療費の適正化の実現する	○	○		○	○	○	<ul style="list-style-type: none">地域別医療費分析レポート医療費の削減提案書
将来予測	地域別疾病発生予測と医療資源配分計画	各地域の疾病発生率や人口動態を基に将来の医療需要を予測し、病床や医療機器、人員など必要な医療資源の配分計画を策定する	○	○	○	○		○	<ul style="list-style-type: none">地域別疾病発生予測レポート医療資源配備計画将来医療ニーズに基づくリソース最適化計画

電子カルテ情報や公的DB等の大規模データを活用した学習により、医療用LLMを構築し、臨床での診断補助を目的としたアプリケーション開発が期待される

ユースケース：LLMを活用した電子カルテ情報のリアルタイム分析による診断補助アプリケーション開発

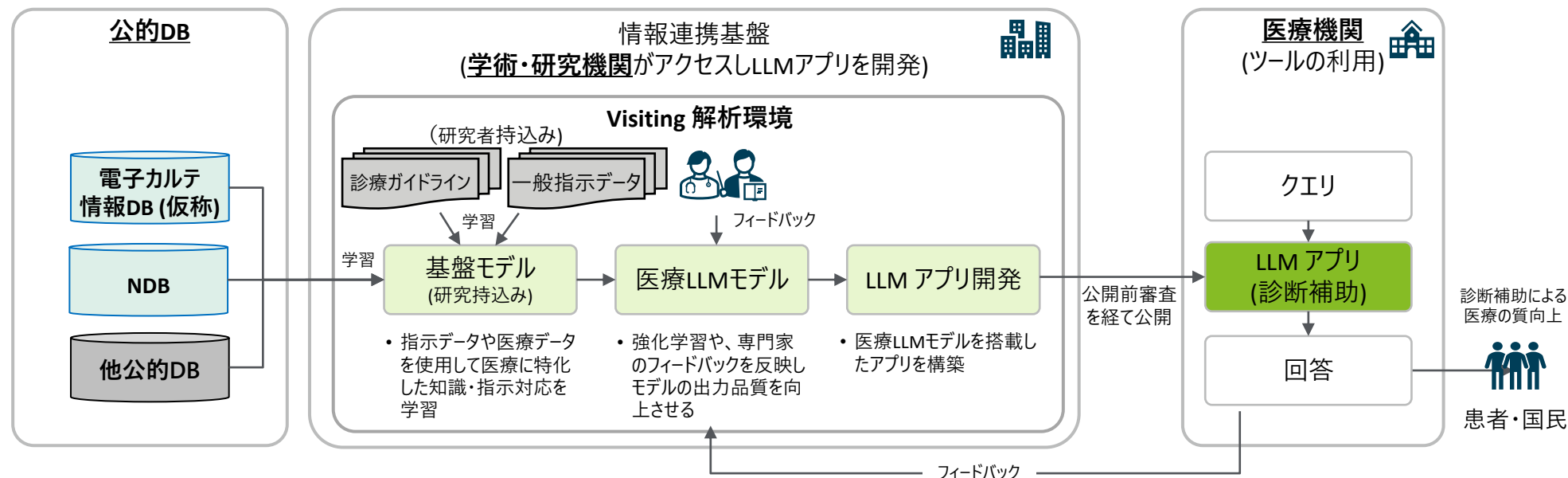
ユーザー像	学術・研究機関
利用目的	診断補助
背景	<ul style="list-style-type: none"> 診断の効率化と質の向上により、医療全体の質向上が期待される 地域医療における人材不足の対策として、AIアシストの導入が有効と考えられる 2021年以降、LLM関連の医学研究が急増しており、米国、インド、中国などが積極的に取り組んでいる

利用するDBとデータ項目(案)

- 電子カルテ情報DB
 - ✓ 3文書：診療情報提供書、退院時サマリー、健康診断結果報告書
 - ✓ 6情報：傷病名、アレルギー情報、感染症情報、薬剤禁忌情報、検査情報、処方情報
- NDB
 - ✓ 診断アウトカムの時系列情報
- その他対象とする疾患領域に該当するDB
 - ✓ 診断情報
- 上記の他、診療ガイドラインや一般指示データなど

※臨床に影響するため、データの品質やバイアスの管理、有効性の検証等が重要課題となる

<データフロー>



4.3 次年度に向けた論点項目の洗い出し

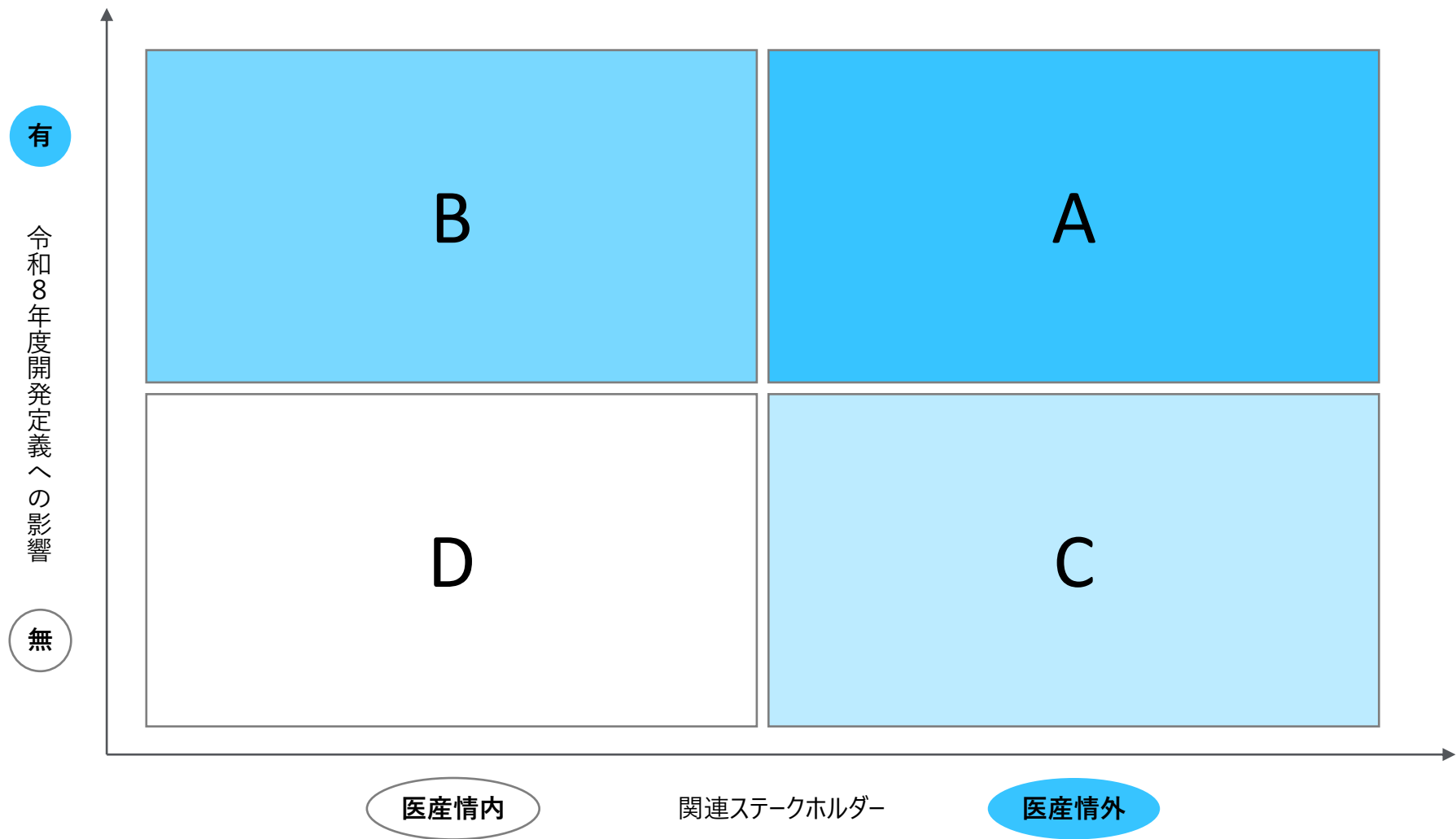
二次利用基盤開発における論点について、下記のWP案毎に整理した

Work Package (案)

共通 テーマ	[WP1] プログラム管理	<ul style="list-style-type: none"> ✓ WP設計とWP推進体制（コンソーシアム）の構築 ✓ WP全体進捗の管理、WP間整合や横断課題の管理、及び本プログラム外部との折衝
	[WP2] 規制・法令遵守	<ul style="list-style-type: none"> ✓ 関連法規制の改正、オプトアウトに係る対応など
	[WP3] ガバナンス構築	<ul style="list-style-type: none"> ✓ 本サービスの運営状況を監督し、医療等情報の二次利用に係る適切なサービス提供・安全対策等を実現するガバナンスルール・体制・プロセスの構築
	[WP4] PoC／パイロット企画推進	<ul style="list-style-type: none"> ✓ PoC及びパイロットの企画及び推進を実施
事業 テーマ	[WP5] ユースケース／サービス利用フロー設計	<ul style="list-style-type: none"> ✓ ユーザー像の具体化と、各ユーザーのユースケースの具体化 ✓ 各ユーザーのサービス利用フローの設計
	[WP6] ビジネスモデル／サービス運営体制構築	<ul style="list-style-type: none"> ✓ ビジネスモデル（サービスメニュー、マネタイズプラン、等）の策定と収支計画の策定 ✓ サービス運営業務の設計及び体制の構築
	[WP7] 情報発信／普及活動	<ul style="list-style-type: none"> ✓ データ利用者及び国民・患者向け情報発信コンテンツの検討 ✓ 普及活動の推進
技術 テーマ	[WP8] データ品質向上／データマネジメント構築	<ul style="list-style-type: none"> ✓ [WP5]を踏まえたデータ品質目標の定義、品質課題の抽出、品質向上施策の立案と推進 ✓ データ標準（メタデータ、データモデル、コード）策定、データマネジメントプロセス・体制構築
	[WP9] プラットフォーム開発	<ul style="list-style-type: none"> ✓ 情報連携基盤、電子カルテ情報DB（仮称）の要件定義・調達・設計・開発・テスト ✓ プラットフォームアーキテクチャ標準ルールの策定
	[WP10] インテグレーション	<ul style="list-style-type: none"> ✓ 公的DB等、外部システムとのインテグレーション機能の要件定義・調達・設計・開発・テスト ✓ インテグレーションアーキテクチャ標準ルールの策定

関連ステークホルダーと、令和 8 年度開発定義への影響有無により、優先度を定義付けた

二次利用基盤開発における論点の優先度定義



共通テーマ（WP1:プログラム管理、WP2 :規制・法令遵守、WP3 :ガバナンス構築、WP4 :PoC/パイロット企画推進）に関する論点

共通テーマ（WP1:プログラム管理、WP2 :規制・法令遵守、WP3 :ガバナンス構築、WP4 :PoC/パイロット企画推進）に関する論点は下記のとおり

WP1～WP4に関する論点（1/2）

：優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP1	1. 他プロジェクト整合	✓ 他プロジェクトを通して見た時に標準化すべき成果物は何か？	厚労省各担当課	有	A
		✓ 他プロジェクトを通して見た時に成果物の依存関係はどうなるか？	厚労省各担当課	有	A
		✓ 他プロジェクト間の整合を図るためにどのような会議体の場でどのような情報共有・討議を行うべきか？	厚労省各担当課	有	A
	2. 連携可能性	✓ R10年度のサービス開始までに連携可能な公的DBは何か？	厚労省各担当課	有	A
WP2	1. 個人情報の取り扱い	✓ 非構造化データに混入し得る顕名データの取り扱いについてどう定義するか？ （法規制上は仮名化、匿名化データを扱うが、混入した場合の対応）	医産情内	有	B
	2. オプトアウト	✓ オプトアウト申請があった際の保護されるデータの範囲は？ （過去に提供済みのデータ等どこまで追跡する必要があるか）	医産情内	有	B
		✓ オプトアウトを実現する業務フロー、各公的DBとの役割分担、情報の流れはどうあるべきか？	医産情内	有	B
	3. 仮名化定義	✓ データ項目レベルでの仮名化データの定義とは？ ID5付与と、情報を丸める処理があり、何をどこまでやれば良いのか？	医産情内	有	B
WP3	1. ガバナンス対象	✓ データガバナンスをどのような目的・スコープ・体制で実施するべきか？	医産情内	無	D
		✓ AIガバナンスをどのような目的・スコープ・体制で実施するべきか？	医産情内	無	D
		✓ セキュリティガバナンスをどのような目的・スコープ・体制で実施するべきか？	医産情内	無	D
		✓ テクノロジーガバナンスをどのような目的・スコープ・体制で実施するべきか？	医産情内	無	D
		✓ サービス利用ガバナンスをどのような目的・スコープ・体制で実施するべきか？	医産情内	無	D
		✓ サービス運営ガバナンスをどのような目的・スコープ・体制で実施するべきか？	医産情内	無	D

共通テーマ（WP1:プログラム管理、WP2 :規制・法令遵守、WP3 :ガバナンス構築、WP4 :PoC/パイロット企画推進）に関する論点は下記のとおり

WP1～WP4に関する論点（2/2） ：優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP3	2. 法的要件	✓ 分析手法の共有にあたり守るべきレギュレーションはあるか？（著作権等）	厚労省各担当課	有	A
	3.ライセンス、契約管理	✓ 連携対象のデータで利用の制限（有償ライセンスが必要なマスタ等）があるものをどうするか？	厚労省各担当課	有	A
WP4	1. パイロットリリース設計要素	✓ パイロットリリースの企画はどうあるべきか？（検証する事項、参加するユーザ、収集するフィードバック等）	厚労省各担当課、ユースケースに基づく利用ユーザ	有	A
	2.本番環境・PoC環境の役割	✓ 本番環境のリリースに伴ってPoC環境の役割をどのように定義し、運用していくか？	厚労省各担当課	無	B

事業テーマ（WP5: ユースケース／サービス利用フロー設計、WP6: ビジネスモデル／サービス運営体制構築、WP7: 情報発信/普及活動）に関する論点

事業テーマ（WP5: ユースケース／サービス利用フロー設計、WP6: ビジネスモデル／サービス運営体制構築、WP7: 情報発信/普及活動）に関する論点は下記のとおり

WP5～WP7に関する論点（1/3） : 優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP5	1. ユースケース	✓ AI利用に関わるユースケースをどのように整理するか？	医産情内	有	B
		✓ 政府、企業、アカデミアの具体的なユーズおよび分析テーマをどのように整理するか？	医産情内	有	B
	2.業務フロー	✓ ユーザー利用申込～データ提供プロセスにおけるAIの利用価値・利用可能箇所はどこか？	医産情内	有	B
WP6	1. バリュープロポジション	✓ 既存公的DBの第三者データ提供サービスとの利用者への提供価値についての差別化はどうするか？	厚労省各担当課	有	A
		✓ HIC / TRE（全ゲノム）との棲み分けはどうするか？ 誰とどのように合意形成していくべきか？	厚労省各担当課	有	A
	2. 事業目標達成指標（KGI/KPI）	✓ 民間DBの第三者データ提供サービスとの利用者への提供価値についての差別化はどうするか？	医産情内	有	B
		✓ 電子カルテ等情報分析サービスの事業としてのKGI/KPIはどうするか？	医産情内	有	B
		✓ 利用者獲得と研究成果の促進のための戦略はどうするか？	医産情内	有	B
	3. サービスメニュー（Visiting環境）	✓ BIツールをどのように選定するか？	医産情内	有	B
		✓ Visiting解析環境で利用するツール群をどのように選定するか？	医産情内	有	B
		✓ Visiting解析環境の標準スペックをどのように決定するか？	医産情内	有	B
		✓ セルフサービス化を実現する上での各政策部局とのあるべき役割分担は？	医産情内	有	B

事業テーマ（WP5: ユースケース／サービス利用フロー設計、WP6: ビジネスモデル／サービス運営体制構築、WP7: 情報発信/普及活動）に関する論点は下記のとおり

WP5～WP7に関する論点（2/3）

：優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP6	4.サービスメニュー (データ提供方法)	✓ 利用者からのオーダーに基づくデータセットの作成（特別抽出）を行うか？	医産情内	有	B
		✓ 提供するデータをどう管理するか？ （スナップショットか？ 共通データセットへの権限付与か？）	医産情内	有	B
		✓ データ提供サービスが具備すべき機能は何かがあるか？ （物理メディアでのデータの提供をするか？）	医産情内	有	B
	5.利用形態 ・契約	✓ 利用者をどのような単位で管理するか？（組織、個人等）	医産情内	有	B
		✓ 包括契約を導入するべきか？	医産情内	有	B
		✓ プロジェクト単位での利用申請と、ユーザー単位のアクセス権限をどのように紐付けるべきか？	医産情内	有	B
		✓ サービス利用料をどのように設定するか？	医産情内	有	B
	6.サービス 運営業務	✓ 審査窓口の一元化を行う場合の審査プロセスはどうあるべきか？ 既存の審査プロセスは残すべきか？	厚労省各担当課	有	A
		✓ Visiting解析環境の適切な利用状況を保つために何を管理すべきか？	医産情内	有	B
		✓ 継続的サービス改善を実現するための情報／データフローをどうするか？ またそれを支える仕組みはどうデザインできるか？	医産情内	有	B
		✓ 医療等情報二次利用を運営するために必要な業務ドメインは？	医産情内	有	B
		✓ 各業務ドメインで遂行すべき業務の一覧と、大まかの業務フローはどう描けるか？	医産情内	有	B

事業テーマ（WP5: ユースケース／サービス利用フロー設計、WP6: ビジネスモデル／サービス運営体制構築、WP7: 情報発信/普及活動）に関する論点は下記のとおり

WP5～WP7に関する論点（3/3） : 優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP6	7.サービス運営組織・体制	✓ 公的DBと情報連携基盤の責任分界点（費用、データ品質、データ説明責任）をどう定義するか？	厚労省各担当課	有	A
		✓ 情報連携基盤の運用主体は誰か？	医産情内	有	B
		✓ 情報連携基盤のデータへのアクセスについての審査の主体は誰か？	医産情内	有	B
		✓ 電子カルテ情報DBの運用主体は誰か？	医産情内	有	B
WP7	1.情報発信のステークホルダー	✓ 医療機関や国民・患者等に、いつからどのように情報発信を行うか？	厚労省各担当課	有	A
	2.コンテンツの分類	✓ 国民・利用検討者・利用者・ヘビーユーザー毎に提供するコンテンツの分類はどうか？	医産情内	有	B
		✓ WEB上のユーザー導線はどうあるべきか？	医産情内	有	B
	3.情報発信の運営体制・方法	✓ 各公的DBとの情報発信の統合・連携分担はどうあるべきか？	厚労省各担当課	有	A
		✓ 既存の厚労省サイトとのすみ分けをどうするか？	厚労省各担当課	有	A
	4.効果測定	✓ 情報発信の成果、効果をどのように計測するか？	医産情内	有	B

**技術テーマ（ WP8: データ品質向上／データマネジメント構築、WP9:プラットフォーム開発に関する論点、WP10:インテグレーション） に関する
論点**

技術テーマ（ WP8: データ品質向上／データマネジメント構築、WP9:プラットフォーム開発に関する論点、WP10:インテグレーション） に関する論点は下記のとおり

WP8～WP10に関する論点（1/4） : 優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP8	1.データ収集	✓ 公的DBが保持するデータを踏まえて情報連携基盤が収集、保持すべきデータは何か？	厚労省各担当課	有	A
		✓ データソース毎にユーザニーズを踏まえた、あるべきデータ更新頻度は？	厚労省各担当課	有	A
		✓ 収集元との整合を担保したデータ収集方式は？	医産情内	有	B
	2.データ結合	✓ 各データソースとの名寄せに使用するキー情報は何か？	厚労省各担当課	有	A
		✓ データ結合は情報連携基盤側で行うとした場合、結合検証はどのように行うべきか？	医産情内	有	B
		✓ 各公的DBのデータを連結可能とし、かつ個人再識別のリスクを可能な限り低減し、オプトアウトにも対応可能なハッシュロジックはどうあるべきか？	医産情内	有	B
	3.データモデル	✓ メダリオンアーキテクチャ等、情報連携基盤に具備するデータモデルを検討する際の参照すべきモデルや先行研究は？	医産情内	有	B
		✓ 電子カルテ情報DBのJSON構造のデータモデルを分析しやすいRDB型に変換するにあたりどのようなデータモデルを採用するべきか？	医産情内	有	B
		✓ 標準で用意しておくデータセットをどう設計するか？/需要のあるデータセットをどう特定していくか？	医産情内	有	B
		✓ 情報連携基盤として採用すべきCDMは？（CDM = Common Data Model）	医産情内	有	B
	4.データ保存	✓ 各公的DBと異なる期間のデータ保持が許容されるか？	厚労省各担当課	有	A
		✓ 公的DBとの整合性も含めて、情報連携基盤上の保存期間をどう定めるか？	医産情内	有	B
		✓ 情報連携基盤上の保存期間を超過したデータについては削除するべきか？	医産情内	有	B

技術テーマ（WP8: データ品質向上／データマネジメント構築、WP9:プラットフォーム開発に関する論点、WP10:インテグレーション）に関する論点は下記のとおり

WP8～WP10に関する論点（2/4）

：優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP8	5. データ加工・変換	✓ 公的DBから情報連携基盤にデータを連携するためのあるべきデータフローは？	厚労省各担当課	有	A
		✓ データ加工の際、どのようにPIIを検出し、どのように処理するか？			
	6. データ品質保証・品質管理	✓ 公的DBのデータセットと合わせて提供する場合、品質管理や説明責任の分界点をどう定義するか？	厚労省各担当課	有	A
		✓ 電子カルテ情報DBとして具備するべきデータバリデーションは？	医産情内	有	B
		✓ 公的DBからコピーし変換を行ったデータについて両者の整合性についてどう考えるか？（SSOTの定義）	医産情内	有	B
		✓ データ品質管理を行う上で監視すべき項目は何か？	医産情内	有	B
	7. マスタデータ管理	✓ 公的DBと情報連携基盤を接続するにあたり、標準で具備するべきマスタはどのようなものがあるか？	医産情内	有	B
		✓ マスタデータは中央集権的に管理して同期を取るか、各公的DBで分散管理して変換し合う形が良いか？	医産情内	有	B
	8. メタデータ管理	✓ 情報連携基盤として、既存公的DB含めたディクショナリ管理を行うか？	厚労省各担当課	有	A
		✓ メタデータ管理運用方法を公的DB側とどう連携して実現するか？（データの更新・メタデータ品質担保等）	厚労省各担当課	有	A
		✓ 利用者視点及び管理者視点で具備するべきデータカタログの表示項目・検索項目は？	医産情内	有	B
		✓ メタデータ標準を設計する場合、アカデミア等、誰とどの様な場で合意形成していくべきか？	アカデミア、標準化関連団体	無	C
		✓ メタデータについて日本の二次利用で共通して利用可能な標準設計とするべきか？	医産情内	無	D

技術テーマ（ WP8: データ品質向上／データマネジメント構築、WP9:プラットフォーム開発に関する論点、WP10:インテグレーション） に関する論点は下記のとおり

WP8～WP10に関する論点（3/4）

：優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP8	9.データ安全性	✓ 非構造化データ(コード化されていないフリーテキストカラムのイメージ)の取り扱い範囲をどうすべきか？	医産情内	有	B
		✓ 仮名化、匿名化したデータの結合による再識別リスクの評価をどのように行うか？	医産情内	有	B
	10.データマネジメント業務・体制	✓ データマネジメントを実施するにあたって必要な業務・役割は？	医産情内	有	B
WP9	1.ソリューション選定	✓ 匿名化・仮名化された医療等情報を安全に取り扱い可能なAIサービスの選定方法は？	医産情内	有	B
		✓ ガバメントクラウドを採用すべきか？	医産情内	有	B
		✓ 情報連携基盤の要件を実現するためのクラウドサービスやソリューションは？またその選定方法は？	医産情内	有	B
		✓ 情報連携基盤、電子カルテ情報DBに求められるサービスレベル（非機能要件）は？	医産情内	有	B
	2.認証認可	✓ ユーザロール（利用者、運用者、審査会など）毎の権限やアクセス可能なりソーススコープをどう設計し実装するか？	医産情内	有	B
		✓ 各政策部局等向けのセルフサービス化を実現する権限をどう設計するか？	医産情内	有	B
	3.セキュリティ	✓ 情報連携基盤に求められるセキュリティレベルとは？またそのセキュリティレベルをどのような仕組み・プロセスで満たしていくべきか？	医産情内	有	B
	4.サービス管理	✓ 従量課金であるクラウドサービスを利用する際のサービス利用コスト管理・分析する仕組みをどう実現するか？	医産情内	有	B

技術テーマ（ WP8: データ品質向上／データマネジメント構築、WP9:プラットフォーム開発に関する論点、WP10:インテグレーション） に関する論点は下記のとおり

WP8～WP10に関する論点（4/4） : 優先度A

論点カテゴリ		論点	関連ステークホルダー	R8開発影響	優先度
WP10	1.他プロジェクト 整合	✓ 公的DBなどのデータソース群とのインターフェースの仕様の整合性をどのように担保するか？	厚労省各担当課	有	A
		✓ 公的DBと情報連携基盤を接続するにあたり、公的DB側にどのような改修が発生するか？	厚労省各担当課	有	A
		✓ 公的DBとのインターフェースをどの程度標準化することができるか？	医産情内	有	B

5. 医療等情報の二次利用推進に向けた海外における データ活用基盤及び技術動向調査について

諸外国における代表的なデータ利活用基盤の技術やビジネスモデルを調査し、情報連携基盤の開発のための検討資料とする

データ活用基盤等調査事業

調査目的	以下の表に記載の項目に関して、海外における主要なデータ活用基盤での活用技術を明らかにし、今後の情報連携基盤の開発に向けた検討資料として活用する
調査手法	デスクトップ調査を通じて公開情報を収集する。なお、必要に応じて、公開情報で得られない情報については有識者へのヒアリングを検討する
調査対象	英・米・仏・フィンランドにおける、主要な公的データベース (例) 英国：NHS、米国：CMS、フランス：HDH、フィンランド：Findata

調査項目	調査内容
① データ利活用基盤・データベースの概要	<ul style="list-style-type: none">国外の分野別（医薬品医療機器開発、疾患研究、疫学分析）の代表的なデータ利活用基盤・データベースに関する利用状況各データ利活用基盤の構築費用及び運営費用（初期費用と維持費用）各データ利活用基盤の利用料や政府補助等の収益構造・ビジネスモデル
② インターフェース・システムアーキテクチャ等	<ul style="list-style-type: none">各データ利活用基盤のシステムアーキテクチャ医療機関から二次利用として活用されるまでのデータ収集フローコード変換、マスター/データモデルの詳細
③ 多国間利用	<ul style="list-style-type: none">超大規模データを多国間で扱っているデータ利活用基盤の事例の収集（国際的なデータ共同利用と、データの直接的な以降を行わない形の利活用基盤の事例を調査）上記における、各種基盤のアーキテクチャの推定

医療情報の二次利用に関連する先進事例について、既存の報告書を参照しながら各事例の詳細を調査した

技術動向調査（案）

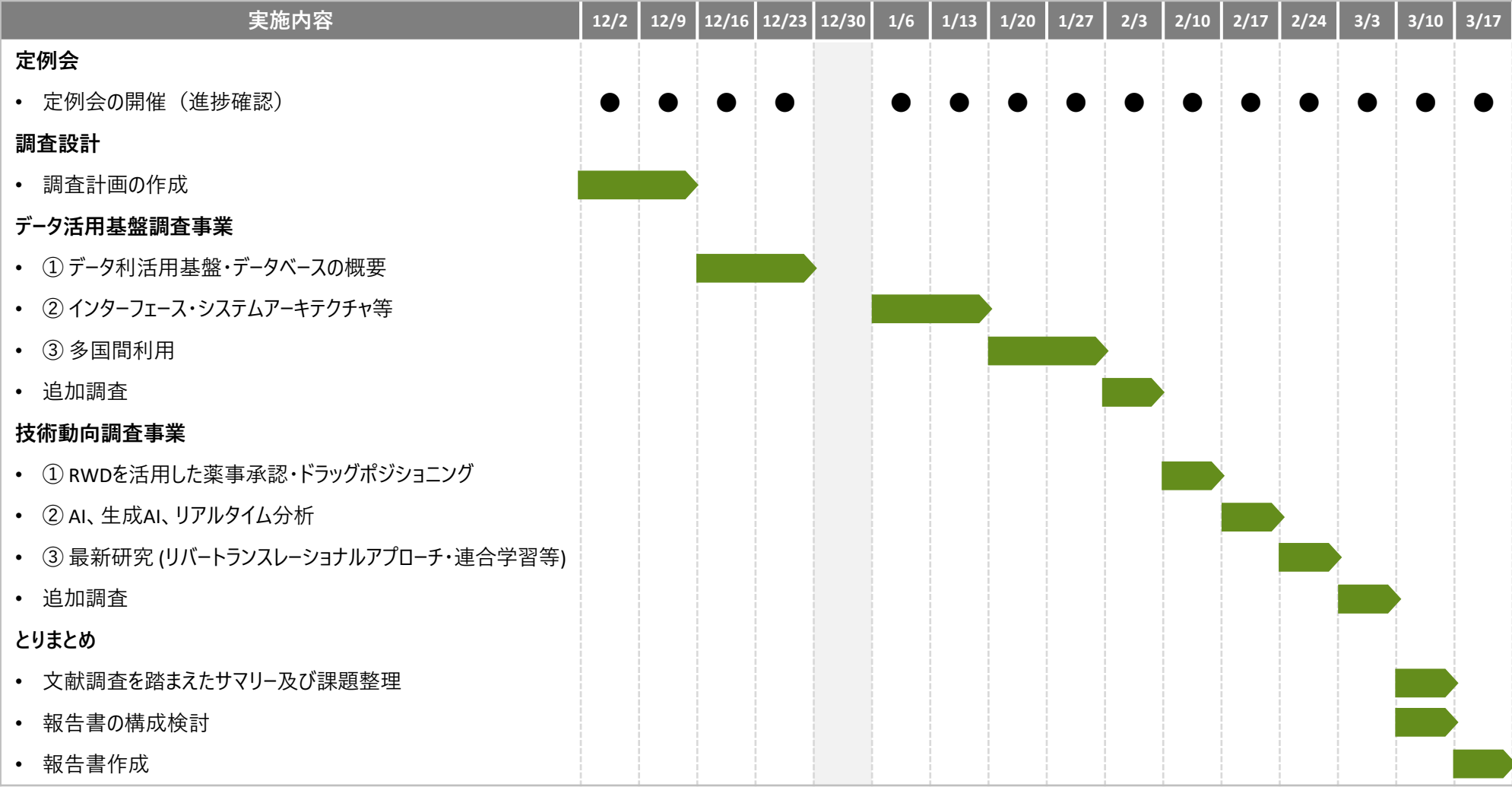
- 既存の調査報告書やシステマティックレビューに掲載されている先進事例を抽出し、それらに基づきデスクトップ調査で詳細（活用基盤・利用データ、データフロー、実績、国民のメリット、課題等）を確認する。なお、既存の報告書に掲載されていない事例に関してはデスクトップ検索で追加調査する
- 国外の調査報告書やガイドラインも参考とする（例） NIH（国立衛生研究所）、FDA（米国食品医薬品局）、EMA（欧州薬品庁）

（例）CRDS報告書に掲載の先進事例抜粋

報告書	分類	事例
CRDS 国立研究開発法人科学技術振興機構 研究開発戦略センター (2023) 健康・医療トランスフォーメーション 科学技術・イノベーションの潮流	RWDの活用	<ul style="list-style-type: none">• FDA（米）は、電子医療記録および登録データからのデータの仕様に関するガイドライン “Framework for FDA’s Real-World Evidence Program”を医薬品開発者に提供している。RWEによりRCT（Randomized Controlled trial：ランダム化比較試験）の問題点を解決することが挙げられており、新しい承認プロセスのパイロットモデルが提示されている• Flatiron Health社（米）は、腫瘍に特化した電子カルテシステムの開発提供を行い、そこで得た匿名化臨床情報を、AIに加え、1000人以上の専門人材により構造化している。このデータを数百カ所のがんクリニックや主要研究機関、腫瘍薬物治療を扱う各種企業と連携している。Flatiron Health 社は他に英国のNICE（National Institute for Health and Clinical Excellence）と2020年より共同で、臨床データを通して医療技術の臨床有用性と医療経済性を評価する研究を開始した
	AI	<ul style="list-style-type: none">• Medidata社（米）は、臨床試験の対照群データを過去の臨床試験またはRWDを活用して合成する技術を持つ。同社プラットフォームでは700万を超える匿名患者データベースを用い、高度なマッチングアルゴリズムによって、臨床研究の目的に最適となる対象群を合成する• Unlearn.AI社（米）は患者のデジタルツインを作成するために、AI、新規統計手法を組み合わせ、より小規模で効率的な臨床試験を可能にする試みを行っている。過去の患者データに基づく機械学習モデルのトレーニングと検証、患者訪問時に収集されたデータがモデルに入力され、デジタルツインが生成される

本調査は以下のスケジュールで実施した

調査スケジュール



5.1 データ活用基盤調査

米国、英国、フランス、フィンランドの公的データ活用基盤は以下の通りである。医療情報等の二次利用を促進し、多くの研究やイノベーションに利用されている

各国比較（概要）

項目		米国 (CMS)	英国 (NHS England)	フランス (HDH)	フィンランド (Findata)
概要	運営主体	ResDAC (ミネソタ大学)、 SDRC (DOBSON DAVANZO & ASSOCIATES, LLC、ACUMEN, LLC) に委託	NHS England	Health Data Hub	フィンランド社会保健省、 フィンランド国立保健福祉研 究所(THL)
	運用開始年	2007年	2010年	2019年	2020年
	データ種類	<ul style="list-style-type: none">Medicare:レセプトと入院から退院までの在院日数をベースとした情報Medicaid:患者サマリ、入院患者、処方箋、長期治療、その他の治療ファイルをベースとした情報	<ul style="list-style-type: none">臨床データ(外来患者ケア、入院患者ケア、救急医療等)疾病別データ(COVID-19、がん登録データ、糖尿病等)社会福祉データ(社会福祉調査、育児、医療スタッフの労働に関するデータ等)死亡登録データ(国家統計局(ONS)より取得)	<ul style="list-style-type: none">医療記録(入院報告,診断書等)検査結果処方箋レセプト情報予防接種受診履歴データ 等	<ul style="list-style-type: none">家系図、親族情報、居住履歴患者データリポジトリ出生児情報、予防接種、社会福祉データ処方箋データがん関連データ死因データ健康保険スキーム下で補償対象となる薬剤、薬剤費用補償資格等
	データ量	米国人口の30%程度 Medicare: 5,680万人 (2016) Medicaid: 6,660万人 (2018)	英国全人口の約90%以上 (データセットにより異なる場合がある)	フランス全人口(約6,804万人,2023年) ※データカタログに応じて規模は異なる	未公開
	研究数	未公開	<ul style="list-style-type: none">530件 (2022年)526件 (2023年)558件 (2024年11月末)	<ul style="list-style-type: none">1,565件(2021年)1,429件(2022年)1,337件(2023年)	<ul style="list-style-type: none">262件(2021年)284件(2022年)351件(2023年)

各国の公的データ活用基盤の収益構造は、公益補助と利用料から成り立っている

各国比較（収益構造）

※1ユーロ=163.6円,2024.12

項目		米国 (CMS)	英国 (NHS England)	フランス (HDH)	フィンランド (Findata)
収益構造	利用料	年間のアクセス料金、1プロジェクトあたりのデータ容量、更新料、追加データの抽出等に対し料金を設定し徴収	データ利用に係る契約、データセット抽出・配信・結合などに対し料金を設定し徴収	<ul style="list-style-type: none"> アクセスは無料 プロジェクトに対しては特定の料金が発生することがある 	<ul style="list-style-type: none"> 決定手数料 データ管理者によるデータ抽出費 データ処理と転送費 Kapselの利用料
	収入源	公的補助及び利用料	公的補助及び利用料	公的補助（欧州委員会および政府）及び利用料	公的補助（欧州委員会および政府）及び利用料
	構築費用	非公開	非公開	初期費用として4年間で 7,600万ユーロ (約121.6億円)	2019～2021年一時的な投資資金:580万ユーロ (約9億4,888万円)
	運営費用	非公開	非公開	2022年より:1783万ユーロ/年 (約28.5億円)	2022年より:220万ユーロ/年 (約3億5,992万円)
	内訳	—	【2021年】 二次利用に限定されないものの、財務諸表上の「データサービス」「IT運用」「プラットフォーム」「サイバー運用」の項目において以下の費用が計上 <ul style="list-style-type: none"> 人件費：8,757万£ (約167億円) 非人件費：2億1,684万£ (約415億円) 	【2022年】 <ul style="list-style-type: none"> 人件費：667万ユーロ (約10.5億円) 運営費（クラウドインフラ等）：1,115万ユーロ (約18億円) 	【2023年】 <ul style="list-style-type: none"> 運営費：120万ユーロ (約1億9632万円) 一般コスト：100万ユーロ (約1億6360万円)

各国標準化団体による定義と周知およびSNOMED CTによるマップ化が進められており、基本的に医療機関における独自コードの使用はないことが想定される

各国比較（データ収集、コード変換）

項目		米国 (CMS)	英国 (NHS England)	フランス (HDH)	フィンランド (Findata)
データ 収集 フロー	標準化に関する ルール	<ul style="list-style-type: none"> ONCが標準規格を定義 NLMC等が定義・配布 	<ul style="list-style-type: none"> NHS等が定義・配布 	<ul style="list-style-type: none"> 保健省およびCNAM（国民健康保険基金）等が定義・配布 	<ul style="list-style-type: none"> 国立保健福祉研究所（THL）等が定義・配布
	医療機関における 独自コードの有無	<ul style="list-style-type: none"> 基本、独自コードなし 政府によるデータの標準化および標準化団体による定義・配布が進んでおり、EHRベンダーがマッピング等の対応を実施 	<ul style="list-style-type: none"> 基本、独自コードなし ITK（相互運用性ツールキット）によるヘルスデータおよびソフトウェア開発等の共通仕様に基づき標準化を推進 NHSが電子カルテベンダー向けにコンフォーマンステスト（適合テストであり、製品が仕様/規格に準拠していることを証明する）を実施し、適合証明書を発行 	<ul style="list-style-type: none"> 基本、独自コードなし 元々、ICD-10のみコード化されていたが、英国を参考にSNOMED CTを導入し標準化を推進中 	<ul style="list-style-type: none"> 基本、独自コードなし 法規制に基づき、THLはコードサーバーおよび外部システムと連携するためのガイドラインを作成
	タイミング	<ul style="list-style-type: none"> 医療機関で実施 	<ul style="list-style-type: none"> 医療機関で実施 	<ul style="list-style-type: none"> 医療機関で実施 	<ul style="list-style-type: none"> 医療機関で実施
	コード の取り 扱い	タイミング コード付 与の状況 <ul style="list-style-type: none"> ④SNOMED CTを入力すると自動的にICD-10付与 ③検査項目を入力するとLOINCコードが付与されるなどコーディング作業が簡素化されている 	タイミング コード付 与の状況 <ul style="list-style-type: none"> ④SNOMED CTを入力すると自動的にICD-10等のコードを付与 	タイミング コード付 与の状況 <ul style="list-style-type: none"> ③ICD-10のみコード化 ②2022年からSNOMED CTの導入を推進中 	タイミング コード付 与の状況 <ul style="list-style-type: none"> 詳細なコード入力およびマッピング方法は不明

- | | |
|--------------------------------|---------------|
| ① マスターが定義・配布できていない | (例：アレルギーコード) |
| ② マスターかできているが、標準コードとマップ化できていない | (例：JLAC10等) |
| ③ マスターかできているが、標準コードとマップ化できている | (例：ICD-10) |
| ④ 標準コードでの運用できている | (例：SNOMED CT) |

各国多国間利用のルールを設け、共同研究を実施している。利用する場合、既存のVisiting解析環境を利用しアクセスを行っている

各国比較（多国間利用）

項目			米国 (CMS)	英国 (NHS England)	フランス (HDH)	フィンランド (Findata)
多国間利用	実績		有	有	有	無（ただしルールはあり）
	研究手法	海外機関の単独研究	－	－	－	－
		共同研究	○	○	○	○
	データ共有	移転	－	－	－	－
		ビジティング解析環境へのアクセス	○ 米国内にいる場合のみVRDCにアクセスしデータ利用が可能である	○ SDEにデータを格納し、海外の大学・民間企業がSDE上にアクセスし分析を行った研究事例あり	○ EU/EEAのデータの個人情報保護ルールまたは拘束的企業規則(BCR)によってデータが保護される	○ GDPR第5章に準拠して「EU/EEA外地域用データ許可」を取得する必要がある

医療情報の一元管理とデータ利活用基盤の整備が必要であり、公的DB等においてもカバレッジを確保し、ユースケースに基づく今後の拡充等についても検討が必要である

海外調査を踏まえた日本への示唆（1/5）

① データ利活用基盤・データベース概要

本調査で得られた情報		日本への示唆
国外の代表的なデータ利活用基盤の利用状況	利用状況	<ul style="list-style-type: none">米国、英国、仏国、フィンランドでは国が一元的にデータを管理し、二次利用のための基盤が整備され、行政、アカデミア、企業等によるデータ利活用が促進されている米国CMS、英国NHSでは以前より長年にわたり二次利用に取り組んでいる。フランス、フィンランドにおいては2019～2020年にかけて公的な基盤を構築し二次利用に注力している
	データ種類	<ul style="list-style-type: none">各国豊富な医療情報を利用できるよう収集・蓄積している<ul style="list-style-type: none">請求情報：レセプト情報臨床情報：外来・入院データ、患者サマリ疾患情報：がんデータ、他その他：死亡情報、予防接種情報、社会福祉データ等
	データ量	<ul style="list-style-type: none">米国では、MedicareとMedicaidがそれぞれ5,680万人および6,660万人をカバーし、英国、フランスでも、全人口を対象としたデータベースが存在する各国のデータ利活用基盤ではカバレッジを拡大し、データ連結解析を進めている
		<ul style="list-style-type: none">日本においても、二次利用を推進するためには<u>医療情報を一元管理しデータ利用を推進する公的なデータ利活用基盤の整備</u>が必要日本の公的DBでは、主にレセプトデータ、DPCデータが主流である。今後、予定している3文書6情報、がん登録DBや難病・小慢DB、感染症DB、予防接種DB等との連結解析が進むことでデータ種類は充実すると思われるただし、臨床情報は現在6情報（アレルギー、感染症、薬剤禁忌、検査、処方、傷病名）に限られているため、<u>ユーズニーズに基づき今後の整理・拡充が必要</u>であるレセプトやDPCデータに加えて、がん登録データベース、難病・小児慢性特定疾病データベース、感染症データベース、予防接種データベースなどの<u>カバレッジを確保し、多様なデータソースを統合</u>することで、データの充実度と利活用の幅を広げる必要がある

データ利活用基盤の構築・運用費用の確保に加え、基盤にデータを収集できるようデータソース側への補助や改修費も組み込む必要がある

海外調査を踏まえた日本への示唆（2/5）

① データ利活用基盤・データベース概要

	本調査で得られた情報	日本への示唆
データ利活用基盤の構築・運用費用	<div>構築・運用費用</div> <ul style="list-style-type: none">構築・運用費用に係る詳細な内訳は未公開であるが、技術費用の他、人件費にも多くの資金が投じられている・英国：人件費約167億円/年、他約415億円/年（ただし、二次利用に限定されていない可能性あり）・フランス：初期費用 4 年間で約120億円、運営費用約28億円/年	<ul style="list-style-type: none">国の規模が類似している英国やフランスにおいても、データ利活用基盤の構築と運用には、多くの公的資金が必要されているデータ利活用基盤の構築と運用費用以外にも、データソースとなる公的DBやその他DBのデータ提供に係る改修費用などの予算確保が必要である
	<div>その他費用</div> <ul style="list-style-type: none">フランスでは、データソースとなるデータ保有者への助成金（年間13～19億円）を確保しており、基盤にデータを収集・蓄積できるよう予算確保を行っている	<ul style="list-style-type: none">データ利活用基盤の構築と運用費用以外にも、データソースとなる公的DBやその他DBのデータ提供に係る改修費用などの予算確保が必要であるそのため、接続する対象スコープを整理し、公的DB側でこういった調整が必要となるか議論が必要である
収益構造	<ul style="list-style-type: none">米国、英国、仏国、フィンランドで利用料に加えて、公的資金による運営がされているデータ自体に費用は発生しないが、データの抽出、提供、解析環境へのアクセスなどに関しては、プロジェクトベースで料金が請求される	<ul style="list-style-type: none">研究者や民間企業からの費用請求の在り方を検討する必要があるその際、データ利活用の妨げとならないよう、適正な価格設定、費用負担の透明性・公平性を確保することが求められる

ユーザニーズを踏まえたデータモデルやシステムアーキテクチャの検討が必要である。また、データ利活用基盤における共通仕様の検討とデータ品質向上が必要となる

海外調査を踏まえた日本への示唆（3/5）


② インターフェイス・システムアーキテクチャ

海外調査で得られた情報		日本への示唆
インターフェイス・システムアーキテクチャ	システムアーキテクチャ	<ul style="list-style-type: none">集中管理方式には、データの標準化やクレンジングの統一が図れること、提供スピードの向上などのメリットがあると考えられる。しかし、その一方で、費用負担の増大も懸念されるため、<u>研究ニーズに合わせた円滑なデータ利活用を実現できるようデータモデルや利用までのバランスを考慮した検討が必要</u>
	技術・相互運用性の推進	
データ収集フロー	技術・相互運用性の推進	<ul style="list-style-type: none">日本の医療情報システム等における相互運用性は低く、<u>データ利活用基盤におけるシステムやデータに関する共通仕様の検討が必要</u>である
	データ収集フロー	

日本では、傷病名（ICD-10）の国際標準化は進んでいるものの、検査や薬剤は国内コードの利用に留まっている。今後、国際標準コードやCDMの導入検討を進めていく必要がある

海外調査を踏まえた日本への示唆（4/5）

② インターフェイス・システムアーキテクチャ

	海外調査で得られた情報	 日本への示唆
標準化	医療機関における独自コードの有無 <ul style="list-style-type: none">米国、英国、仏国、フィンランドにおいて基本的に独自コードの使用はなく、医療機関側で標準化を実施している	<ul style="list-style-type: none">データの標準化は、二次利用の効率化や品質向上に重要である異なるデータベース間の結合や利活用を進めるためには、<u>独自コードの使用を減らすか、他のコードとのマッピングが必要</u>となる
	標準コードの利用状況 <ul style="list-style-type: none">ICD-10、SNOMED、LOINCなどの国際標準コードの利用が主流となっている	<ul style="list-style-type: none">日本では、ICD-10は普及しているが、検査や薬剤情報においては国内コードが主流であり、かつ普及率は伸び悩んでいる今後、<u>国際標準コードの導入の必要性と導入にかかるプロセスや労力などを詳細に整理</u>していく必要がある
	CDMへの対応 <ul style="list-style-type: none">英国やフランスでは、OMOP CDMへの変換対応が進められている	<ul style="list-style-type: none">今後の国際共同研究の可能性を考慮すると、国際標準への対応が求められる<u>研究ニーズに基づいたCDM導入の検討</u>を進めるべきである

海外では、多国間利用に関するルールを定め運用している。日本においても、国際共同研究利用を想定しているため、GDPRやHIPAA等の法規制を念頭に検討が必要である

海外調査を踏まえた日本への示唆（5/5）

③多国間利用

	海外調査で得られた情報	日本への示唆
海外実績とルール化	<ul style="list-style-type: none">米国、英国、仏国は多国間利用に関するルールと実績あり海外から利用する場合、明確なルールを設けて運用を行っている<ul style="list-style-type: none">EUの場合、GDPRに基づく十分性認定を受けているか英国の場合、特別な理由がない限り国内にデータの共有を留める方針であり、英国外での利用は、適切な根拠を示す必要がある米国の場合、国内外からのアクセスかつ外国人が利用等のケースに応じてルール化	<ul style="list-style-type: none">国際共同研究等による基盤利用を想定し、<u>多国間利用のルールを整備する必要がある</u>その際にEU法のGDPRや米国のHIPAA等との整合性についても確認が必要となる
データ共有方法	<ul style="list-style-type: none">データの移転はなく、ビジティング解析環境を通じたデータアクセス・分析が基本となっている	<ul style="list-style-type: none">日本においても、国際的なデータ利活用を視野に入れつつ、データを移転せずにビジティング解析環境を通じて提供する仕組みが求められる国内向けのルールに加えて、<u>多国間利用の際には、セキュリティ基準や運用管理の再検討が必要</u>

データ活用基盤調査

- 米国
- 英国
- 仏国
- フィンランド

CMSは米国の医療制度の強化と近代化を目指し、低コストで質の高い医療と健康増進へのアクセスを提供するためのデータ統合プラットフォームである

CMS (Medicare・Medicaid・サービスセンター)概要

① データ利活用基盤・データベースの概要

名称	CMS (Medicare・Medicaid・サービスセンター)
運営主体	<ul style="list-style-type: none">• ResDAC：研究データ支援センター (ミネソタ大学)• State DRC：州レベルでのデータアクセス管理センター(DOBSON DAVANZO & ASSOCIATES, LLC、ACUMEN, LLC) に委託
運用開始年	2007年
関連法	FISMA、The Privacy Act of 1974、HIPAA
設置目的	国の医療制度の強化と近代化を目指し、低コストで質の高い医療と健康増進へのアクセスを提供
利用者	<ul style="list-style-type: none">• 研究者またはイノベーター (営利組織に所属する研究者、または販売する製品やツールを作成する目的で研究を行う人、企業も含まれる)• 政府機関(連邦捜査局、司法省(DOJ)等

データソース	<ul style="list-style-type: none">• Medicare:レセプトと入院から退院までの在院日数をベースとした情報• Medicaid:患者サマリ、入院患者、処方箋、長期治療、その他の治療ファイルをベースとした情報• これらのデータを統合し、PUF (Public Use Files)、LDS (Limited Data Set)、IDF (Identifiable Data Files) の3種類のデータセットを作成<ul style="list-style-type: none">➢ PUF：定型化された全国及び各州の集計データ➢ LDS：Medicare又はMedicaidの入院・外来・介護に関する 16 種類のデータセット➢ RIF：Medicare又はMedicaidの入院・外来・介護に関する 54 種類のデータセット
データ規模	<ul style="list-style-type: none">• Medicare: 5,680万人 (2016年)• Medicaid: 6,660万人 (2018年)
利用実績	非公開 ※CMSおよび委託先のResDACのサイトにおいてデータ提供実績を公開していない

IDRは主にCMSおよび政府機関が利用するDB、CCWはCMSデータを外部の研究者が利用できるよう設計されたDBである

データベースの概要

① データ利活用基盤・データベースの概要

項目		IDR (The Integrated Data Repository : 統合データリポジトリ)	CCW (Chronic Conditions Data Warehouse : 慢性疾患データウェアハウス)
データベースの役割		✓ リレーショナル・データベース (複数のデータを関連付けて格納したデータベース)	✓ Medicare、Medicaid、処方薬イベント(PDE : Prescription Drug Event※)のデータをより容易に利用できるようにするために設計された研究用データウェアハウス ✓ 仮想研究データセンター (VRDC) を通じて利用
利用対象者		✓ CMS →データ分析等 ✓ 政府機関(連邦捜査局、司法省(DOJ)等 →MedicareとMedicaidの不正、浪費等の調査や訴訟 関連のために利用	✓ 研究者、企業など
保有データ		✓ <u>2007年から現在まで数十億のデータを統合し管理</u> <ul style="list-style-type: none">Medicare、MedicaidのパートA(入院患者)、パートB(外来)、パートC(給付金)、パートD(処方薬) 等基本情報：氏名、生年月日、電話番号、医療記録番号(単一の患者を参照するためにプロバイダーによって割り当てられる一意の識別子)、金融口座情報など	✓ <u>予備的な連結作業はすでに完了しており、CCWには研究者に送られるデータファイルとして提供される</u> <ul style="list-style-type: none">Medicare、MedicaidのパートA、パートB、パートD処方薬のデータ基本情報：BENE_ID、州・郡コード、郵便番号、生年月日、死亡日、性別、人種、年齢など
データベースの更新	パートA、B	毎週	毎週
	パートD(処方)	毎日	年間
	Medicare、Medicaidの資格情報	毎日	毎月

※PDE：処方データセットであり、処方日、製品ID、分量、処方日数、料金等を含む。処方薬を含む研究や治療結果を比較する研究に利用

出所: <https://www.census.gov/fedcas/cfc2017/ppt/tdEvron.pdf>, <https://support.sas.com/resources/papers/proceedings20/4285-2020.pdf>, <https://s8637.pcdn.co/wp-content/uploads/2017/04/RESDAC-CMS-Data-Data-Science-WHI-Data-Science-Wiki.pdf>, <https://www.hhs.gov/sites/default/files/cms-ccio-customer-relations-management-system.pdf>, <https://resdac.org/cms-data/variables/ccw-beneficiary-id>

CMSが保有するMedicareデータおよびMedicaidデータは、双方のデータを統合できること、データベースの規模が大きい (人口の3割程度) ことなど、有用な点が多い

① データ利活用基盤・データベースの概要

CMS保有データの概要

項目	内容		
保有データ	Medicare	標準分析ファイル (SAFs) ※レセプトベース	外来の場合の治療日・入退院日、誕生日 (年齢)、人種、性別、診断名 (ICD-10)、治療内容 (ICD-10)、治療内容に対する支出、治療を受けた医療施設ID、治療を受けた医師のUPI (Unique Physician Identifier)、主治医のUPI 等
		医療供給者分析ファイル (MedPAR) ※入院から退院までの在院日数ベース	
	Medicaid	患者サマリファイル (Personal Summary file)	生年月日、性別、人種、郵便番号、Medicaid入会日、managed care plan加入時期
		入院患者ファイル(Inpatient file)	医療施設ID、入退院日、転帰、診断名 (9個までICD-9-CM)、治療 (6個までICD-9-CM)、入院費用
		処方箋ファイル(Prescription Drug file)	処方箋コード (NDC) : 製造販売業者、力価、剤型、処方量 等
		長期治療ファイル(Long Term Care file)	長期治療記録: 施設規模、医療サービス提供日、診断、転帰 等
		その他の治療ファイル(Other Therapy file)	入院を除く医療サービスの記録・診療、臨床検査、放射線検査 (検査記録は、実施日・検査項目のみであり、結果データは得られない)
	これらのデータを統合し、PUF (Public Use Files) 、LDS (Limited Data Set) 、IDF (Identifiable Data Files) の3種類のデータセットを作成		

CCWでは以下の3種類のデータセットを利用可能であり、データの形式に応じて審査や手続き内容が異なる

CCWデータの概要

① データ利活用基盤・データベースの概要

利用可能データ		①Public Use Files (非識別情報)	②Limited Data Set File (リミテッドデータセット)	③Research Identifiable Files (研究用特定ファイル)
概要		編集され、個人が識別される可能性のあるすべての情報が削除されている	16種類の直接識別情報を除外した PHI(保護対象医療情報)。RIFとLDSの違いは、LDSファイル内の選択された変数が空白または範囲指定されていることである	DUALに加えCMSプライバシー委員会の審査を受け、患者のプライバシーが保護され、必要な最小限のデータのみを使用することを正当化する必要がある
審査・契約	CMS privacy board (CMSプライバシー委員会) の審査	不要	不要	必要
	研究機関のIRBの審査	不要	不要	必要
	DUA(データ使用契約)	不要	必要	必要
	患者情報を含むか	含まない	含む	含む
	患者の承認	不要	DUAを締結することで不要	IRB審査による本人承認の免除が可能
	研究者はカスタマイズされたコホート (例:ミネソタ州在住の糖尿病患者) をリクエスト可能か	不可	不可	可能
	患者識別子を用いてCMS以外のデータにリンク可能か	不可	不可	可能 ※外部データとリンク可能な患者識別子を含めるには、CMSの承認が必要
利用データ	固有の患者識別子	-	暗号化	暗号化
	HIC (保険者番号)、SSN(社会保障番号)		含まない	特別な許可を得た場合のみ含まれる
	日付		四半期と年のみ	含まれる
	郵便番号		州・群のみ	含まれる
	死亡日		一部含まれる	含まれる
	医師固有の識別番号		含まれる	含まれる
	施設番号		含まれる	含まれる

CCWのVRDCでは年間のアクセス料金、プロジェクト単位の料金をベースとし、データ量や追加データの抽出条件などに基づき利用料金が掛かる

データ利用料金

① データ利活用基盤・データベースの概要

- CCWの仮想研究データセンター (VRDC)は、仮想空間を通じて研究データを安全にアクセスする方法である
 - Public Use Files のみダウンロード可能
 - イノベーターがLimited Data Set File、Research Identifiable Filesを利用する場合、必ずVRDCを利用しなければならない
- ※1 米ドル = 155.92 日本円

項目		単位	研究者		イノベーター ※営利組織に所属、または製品やツールを販売する目的で研究を行うもの	
			オプション1 「SASのみ」	オプション2 「Full VROC」	オプション1 「SASのみ」	オプション2 「Full VROC」
年間のアクセス料金 ・ CMSオンボーディング(研修等) ・ シートライセンス ・ サポート ・ 管理手数料		1プロ ジェクト	3,118,400円	3,430,240円	3,118,400円	3,430,240円
1プロジェクトあたりのデータ容量料金 (2TBまで)			2,338,800円	2,806,560円	3,898,000円	5,457,200円
2年目以降の 更新料	ユーザーアクセス 「シート」の更新	1年	2,026,960円	2,338,800円	2,026,960円	2,338,800円
追加料金	追加スペース (1TB)	1回	2,338,80円			
	追加データの抽出	1回	311,840円			

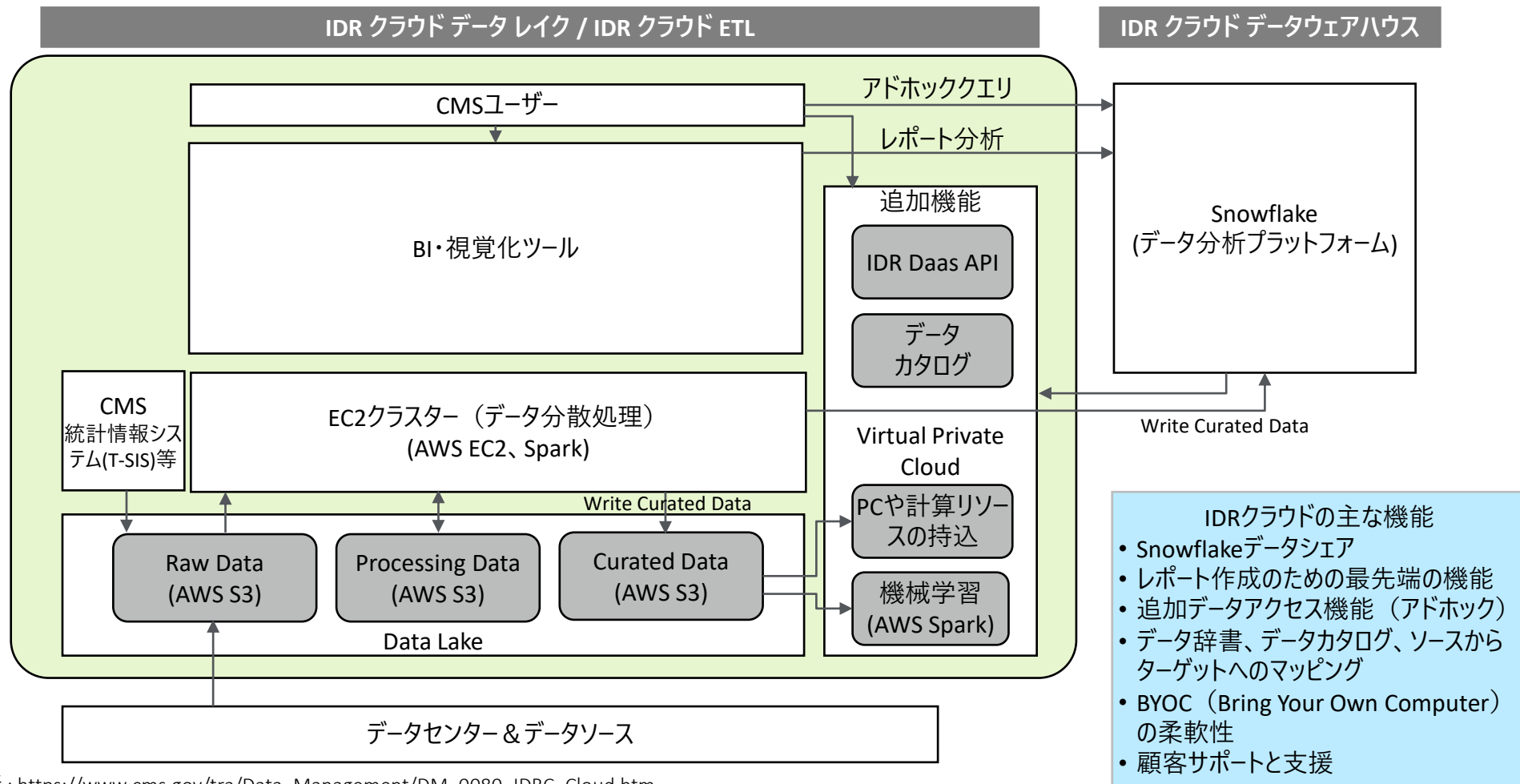
オプション1 SASのみ = SAS形式ファイル
オプション2 Full VROC = 生データ、分析ファイル

IDRにおけるデータレイクとETLのシステムアーキテクチャは以下の通りである

② インターフェース・システムアーキテクチャ等

IDRの概念アーキテクチャ (1/3)

✓ IDR（統合データリポジトリ）は、Medicare請求（パート A、B、C、D）およびプロバイダーのデータソース等を統合する大容量データウェアハウスである



IDRでは5つのデータ可視化ツールが使用されており、データの簡素化や重要なインサイトの提供を行っている

IDRの概念アーキテクチャ（2/3） 使用されているデータ可視化ツール

② インターフェース・システムアーキテクチャ等

ツール	Tableau	Looker	Power BI	QlikView	QuickSight
ベンダー	Salesforce	Google	Microsoft	Qlik	AWS
概要	ドラッグ & ドロップで、データを直感的かつ視覚的に表現するクラウドベースの分析プラットフォームである	顧客が利用しているデータベースやデータウェアハウス(DWH)に接続することを前提とし、企業が利用するデータを一元管理し、データ分析と可視化を行う	ソフトウェア サービス、アプリ、コネクタのコレクションを組み合わせることにより、関連性のないデータソースから、まとまりがあり、実体験的な対話型の分析情報を提供する	大規模で複雑なデータセットを処理している場合でも、全体像や関係性の理解を助ける分析プラットフォームである	ユーザーが、自然言語を使用して有意義なインサイトの構築、発見、共有を数秒間で行える
使用目的	ユーザーへの以下の機能を提供する： <ul style="list-style-type: none">主要なデータに簡単にアクセスして分析し、革新的なレポートやビジュアライゼーションを作成し、重要なインサイトを機関全体で共有する	－	全てのアプリケーションデータソースに接続することで、データの正確性、一貫性、セキュリティを維持しながら、スケーラブルなダッシュボード、対話型レポート、埋め込みビジュアルなどを使用して、データの評価と共有を簡素化する	－	インタラクティブなビジュアライゼーションとダッシュボード、アドホック分析、APIとSDKによる組み込み分析を行う

IDRでは4つのエンタープライズBIツールが使用されており、詳細は以下のとおりである

IDRの概念アーキテクチャ（3/3）

② インターフェース・システムアーキテクチャ等

IDRで使用されているエンタープライズBIツール

ツール	SAS-EBI	MicroStrategy	BusinessObjects	ArcGIS
ベンダー	SAS	MicroStrategy	SAP	Esri
最適なユースケース	<ul style="list-style-type: none"> データ管理 分析 多変量解析 予測分析のための統計ソフトウェアスイート 	<ul style="list-style-type: none"> インタラクティブダッシュボード スコアカード フォーマット済みレポート しきい値とアラート 自動レポート配布 	<ul style="list-style-type: none"> SAPエコシステム内でのレポート作成 視覚化 シームレスな統合ネイティブ接続 	<ul style="list-style-type: none"> 地理空間マッピング 解析
想定ユーザー	データアナリスト	開発者とレポート/ダッシュボードの利用者	開発者、データアナリスト、レポート/ダッシュボードの消費者	マップとダッシュボードの開発者とユーザー
ユーザー容量	<ul style="list-style-type: none"> ユーザー(Web、Citrix Microsoftクライアント)：約1,200人 同時ユーザー：200人 SASライセンス：制限あり 	<ul style="list-style-type: none"> ユーザー：約3,500人 同時ユーザー：100人 ユーザーのライセンス割り当て：制限あり 	<ul style="list-style-type: none"> ユーザー：約1,400人 同時ユーザー：200人 	<ul style="list-style-type: none"> Creator ユーザー：150人 視聴者：無制限 Amazon Workspaces：2つのArcGIS Proライセンス 5 ArcGIS Desktop Pro (CPI ユーザー)

※エンタープライズBIツールとは、大規模な企業や組織がデータを収集、統合、分析し、意思決定を支援するためのソフトウェア

米国では、ONCが標準化ルールを調整、EHRベンダーがサポートを行うことで、医療機関でのコード入力および変換が実施されている

データ収集フロー

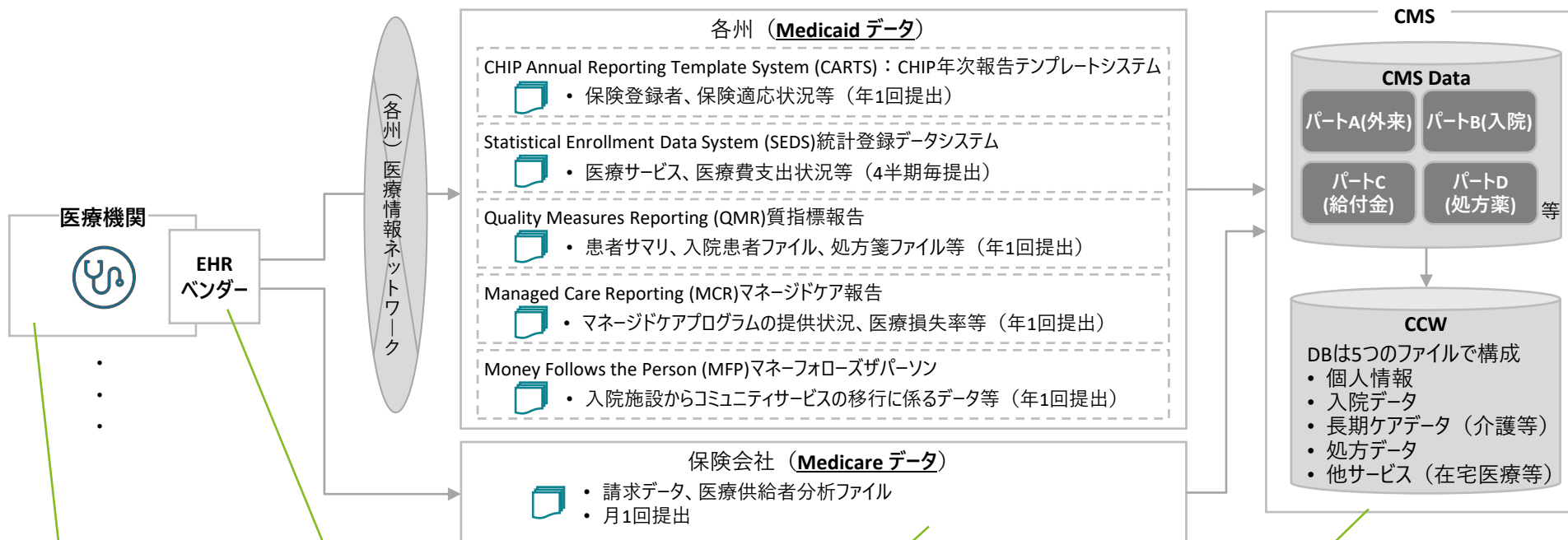
② インターフェース・システムアーキテクチャ等

データガバナンス（ONC：米国保健福祉省の医療IT全米調整官室）

データ標準化（NLMC：米国国立医学図書館）
※SNOMED CTの例

- ・ 米国版SNOMED CTの定義・配布

- ・ 関連法規制の調整、ガバナンスの策定・更新
→EHR基準の設定（SNOMED CTやLOINC等の標準化を推進）



- ・ SNOMED CTコードの入力
- ・ 検査項目からLOINCコードのマッピング
- ・ SNOMED CTコードを元に、ICD-10コードの自動生成
- ・ 医療者によるコードのレビュー

- ・ SNOMED CTやLOINCを含む標準化された医療用語集をシステムに統合
- ・ 医療者がデータ入力時に適切なコードを容易に選択できるようサポート

- ・ Medicareエンティティ（医療機関、保険会社等）はCMS請求データ変換ツールを用いてCMSの指定ファイルを作成後、データを提供

- 【ガバナンス】
- ・ データ収集フローガイドライン、二次利用ユーザーガイドの策定・周知
- 【データ管理】
- ・ データ処理/品質管理の実施

SNOMED Internationalは、包括的な医療用語であるSNOMED CTと他標準コードとのマップ化を推進している

② インターフェース・システムアーキテクチャ等

【参考】SNOMED CTマップ

- SNOMED CT は、他の国際的な用語、分類、コード システムに相互マップできる用語である
- SNOMED International（英国にあるSNOMED CT を所有、管理、開発する非営利団体）は、SNOMED CT から他のコードシステム、分類、または用語へのマップを作成し、SNOMED CTを使用するメンバーやその他の関係者をサポートしている

相互マップの例

コード	コード概要	マップ方法
ICD-10	国際疾病分類 (ICD) は、WHO (世界保健機関)が発行する、世界規模の健康傾向と統計の特定、および疾病と健康状態の報告に関する国際標準の基礎	<ul style="list-style-type: none">✓ SNOMED CTで臨床的にエンコードされた臨床記録からICD-10統計分類データを半自動でコーディング✓ レジストリや診断グループ化に使用するために、SNOMED CT エンコードされたレコードから ICD-10 分類コードを開発するのに役立つ✓ WHOおよびSNOMED International によって検証された SNOMED CT から ICD-10 へのマップとして、また加盟国によって開発および管理されている ICD-10 拡張へのマップの開発のソースとして機能する✓ SNOMED CTはICD-10バージョン 2016 にマッピングされている
ICD-O	国際腫瘍分類 (ICD-O) は、国際疾病統計分類の腫瘍疾患に対するドメイン固有の拡張	<ul style="list-style-type: none">✓ SNOMED CTからICD-O へのマップは、SNOMED CT 国際リリースのシンプルなマップ参照セットとしてリリースされている✓ マップのSNOMED CTソース ドメインは、ICD-O形態コード、サブタイプに限定されている
GMDN	グローバル医療機器命名法(Global Medical Device Nomenclature) は、すべての医療機器製品を識別するために使用される一般名のリスト	<ul style="list-style-type: none">✓ GMDNコードを使用する予定で、SNOMED CTからGMDNへのシンプルマップを取得したいユーザーは、標準リンク テーブルライセンスフォームをダウンロードして記入し、GMDNに送信する必要がある✓ SNOMED Internationalは、ライセンスが発行されたことを当局が確認すると、マップをダウンロードできるようにする（このプロセスにより、マップは 2015年から一般に公開されている）
Orphanet	希少疾患患者の診断、ケア、治療を改善するユニークなリソースであり、Orphanetはフランス国立保健医療研究所 (INSERM) で設立	<ul style="list-style-type: none">✓ OrphanetとSNOMED International は、Orphanet の希少疾患コンテンツを SNOMED CTと整合させるために、SNOMED CTとOrphanet 間のマップを作成✓ 2022年と2023年のマップリリースには追加のマップが含まれている

米国では、ONCがデータ標準化を促進している

CMSで使用されているコード一覧

② インターフェース・システムアーキテクチャ等

- 米国では、ONC(保健福祉省傘下の医療IT全米調整官室)がEHR認証基準を設定し、コモンルール(45 CFR)の追加・更新の調整を実施している
- 2015年の更新において、SNOMED CTやLOINC(Logical Observation Identifiers Names and Codes®)などの標準用語への関連付けが進められている

分類	コード(略語)	説明	利用範囲
臨床診断	ICD-10	WHOが定めた国際疾病分類として、病気やその他の健康状態を分類するため使用されている	世界標準
医療行為 医療機器	HCPCS I / CPT	主に医師やその他の医療専門家が提供する医療サービスと手順を識別するために使用される	米国
	HCPCS II	救急車サービスや耐久性のある医療機器、補綴物、装具、消耗品(DMEPOS)など、CPTコードに含まれていない製品、消耗品、サービスを主に識別するために使用される	米国
医療行為	SNOMED CT	状態、診断、症状、徴候など、多領域にわたる臨床概念を表す。他にも手順、観察、および一部の臨床検査、薬物、およびデバイスも表す	世界標準
医薬品	RxNorm	医療保険で払い戻し可能な医療機器およびサービスのリストである	世界標準
検査	LOINC	臨床観察および検査観察、ヘルスケアスクリーニング/調査機器、およびドキュメントタイプ識別子のためのコードシステムである	世界標準

AcademyHealthの研究コミュニティがデータアクセスのVRDC統一化に懸念を示した結果、2026年まで統一化の対応は見送りとなった

② インターフェース・システムアーキテクチャ等

AcademyHealthの研究コミュニティのVRDCアクセス統一化に係る懸念への対応

- CMSの提案（2024年2月12日）
 - ✓ CMSがデータセキュリティの強化を理由に、物理データの授受を終了し、CCW VRDCを介したデータアクセスのみを許可することを提案している。
これにより、2024年8月19日以降のすべての新しいデータ使用契約にCCW VRDCを利用することが必要になる
- CMSの提案に対してAcademyHealthの研究コミュニティがCMSに示した懸念点（2024年2月26日）
 - ✓ 医療サービスの研究や医療政策の進歩に対する重大な障壁となる可能性として以下 6 つの懸念を挙げている
- AcademyHealthの懸念点に対するCMSの対応（2024年4月15日）
 - ✓ ステークホルダーからの意見を聞き(2025年4月15日まで)今後の対処を検討するために、少なくとも2026年までは新たな変更を実施しないと発表

【 AcademyHealthがまとめるCCW VRDC移行に係る懸念点】※1 米ドル =155.92 円

懸念点	内容
高額なコストの発生	<ul style="list-style-type: none">• 現在のモデル（物理データの授受）よりも研究者に発生する費用が高価(1人あたり、プロジェクトあたり、年間30,000ドル以上)である• 画面共有を禁止しており、データを複数のプロジェクトや助成金で再利用することができない• 一部の研究者においては、VRDCへの移行に数十万ドルの新たな資金を確保する必要があり、非現実的である
システム環境が不十分	<ul style="list-style-type: none">• 利用可能なプログラミング言語、ツール、ライブラリなどの制限がある• 研究者が現在使用しているツールよりもはるかに技術的能力が低く、機械学習やマイクロシミュレーションモデリングなどの高度なコンピューティング技術のサポートが不足しており、RやPythonなどの統計言語を大幅な追加コストなしで使用できない
利用負担	<ul style="list-style-type: none">• 現在物理データの授受を行っていた研究者にとってどのCCW VRDCオプションがニーズに最も適しているのかが不明である• 新しいシステムを学ぶ負担がある• 最適でないユーザーエクスペリエンスである可能性がある
既存資産の取り扱い	<ul style="list-style-type: none">• 既存のプロジェクトでは、研究者は現在のCCW VRDCにアップロードする許可を得ない独自のデータセットを利用している• 多くの機関が既に資金、時間、労力に多額の投資をして、独自のデータ資産を確立している
移行時間が短い	<ul style="list-style-type: none">• 提案されている6か月の移行期間は、研究を中断し、新しいデータセットと方法論を見つけるのには短すぎる
その他	<ul style="list-style-type: none">• 新しい政策により、研究者がニーズに合わない環境を使用することを余儀なくされるのではないかと懸念がある

CMSは米国を拠点としない研究者に対して、アクセス可能なデータとデータへのアクセス条件に制限を設けている

③ 多国間利用

CMSにおける多国間利用ポリシー

対象		自国の研究者	外国人の研究者	
概要		①米国に住んでいない (現在 or 一時的に)	②米国に住んでいる (現在 or 一時的に)	③米国に住んでいない (米国市民権または米国居住権を持たない)
基本ルール		<ul style="list-style-type: none">RIF、LDSファイルは仮想プライベートネットワーク(VPN)を含むいかなる手段でも、アクセスすることは許可されていない	—	現在住んでいる場所に関係なく、以下へのアクセスが許可されていない <ul style="list-style-type: none">連邦システムCMS Virtual Research Data Center内のデータ
アクセス可能なデータ	概要	CMS cell suppression policyとHIPAAの匿名化基準を満たすCMSデータ	米国の大学/非営利団体のCMSデータ	CMS cell suppression policyの基準を満たすCMSデータ
	種類	PUF	RIF、LDS	PUF
	条件	—	<ul style="list-style-type: none">米国の大学/非営利団体は、CMSデータ使用契約(DUA)に基づくデータの要求者および管理者である研究者は、米国の大学/非営利団体(客員教授など)に所属している	—
例		<p>■ アクセス可能な場合：</p> <p>フランスの教授が米国の大学(USU)を訪れており、USUにいる間にMedicareまたはMedicaidのデータを使用して研究を行うことを望んでいる。客員教授は、USUデータカストディアンの監督の下で、実際に米国にいる間のみデータにアクセス可能</p> <p>■ アクセス不可能な場合：</p> <p>USUの研究者は、スイスの大学の客員教授であり、スイスにいる間も研究を続けたいと考えている。研究者は、USUにVPN接続して、集計されていないデータにアクセスすることはできない。しかし、米国外にいる間、研究者はCMS cell suppression policyを満たす集計結果にアクセスすることができる</p>		

※多国間利用の場合、VRDC内のデータを利用することが原則となっているため、既存のVRDC環境以外に特別なシステム利用はないと想定

CMS cell suppression policyとは個人情報を守るために定められたCMSデータの表記に関するルールである

③ 多国間利用

CMS cell suppression policyの概要

ポリシー名	CMS cell suppression policy
概要	CMSデータの表記に関するルールを定めており、1～10の数字を記載することが出来ない
目的	MedicareやMedicaid使用者（受益者）の 個人情報を守るために 、誰がどのような治療を受けたかが特定されないようにする
基本ルール	<ul style="list-style-type: none">1から10までの数字を表記してはいけない（0は記載可能） 例）入院した人の数や退院した人の数が1から10の間の場合、その数字は記載不可である他の方法を使用した場合でも、1から10の数字がわかるようなデータの表記を行ってはいけない 例）パーセンテージの記載など、計算することで数字がわかる記載方法など
対象	<ul style="list-style-type: none">受益者、手続き、診断についてのデータが記載されている表と文RIF、LDS のデータを使用時CMSデータを使用して作成されたすべてのドキュメント(原稿、表、チャート、研究、レポートなど)
回避方法	セルの折りたたみ、データの粗大化、セルの抑制など

CMS cell suppression policyの例

【違反する例】

- 85+歳のデータに 6 と表示されている
- 85+歳のデータの値が計算可能である
** = 2690 - (1900+400+290+94) = 6

受益者の年齢層	人数(%) ※N=2,690
60～69歳	1900 (71)
70～74歳	400 (15)
75～79歳	290 (11)
80～84歳	94 (3.5)
85+歳	6 (0.22)

受益者の年齢層	人数(%) ※N=2,690
60～69歳	1900 (71)
70～74歳	400 (15)
75～79歳	290 (11)
80～84歳	94 (3.5)
85+歳	**

【違反しない例】

- 85+歳のデータを80～84歳のデータに統合する
- 80～84歳と85+歳のデータをマスキングしている

受益者の年齢層	人数(%) ※N=2,690
60～69歳	1900 (71)
70～74歳	400 (15)
75～79歳	290 (11)
80+歳	100 (3.7)

受益者の年齢層	人数(%) ※N=2,690
60～69歳	1900 (71)
70～74歳	400 (15)
75～79歳	290 (11)
80～84歳	**
85+歳	**

データ活用基盤調査

- 米国
- 英国
- 仏国
- フィンランド

NHS England では、Data Access Request Service (DARS) が NHSが保有する医療・社会福祉データの利用承認・提供するための仕組みとして機能している

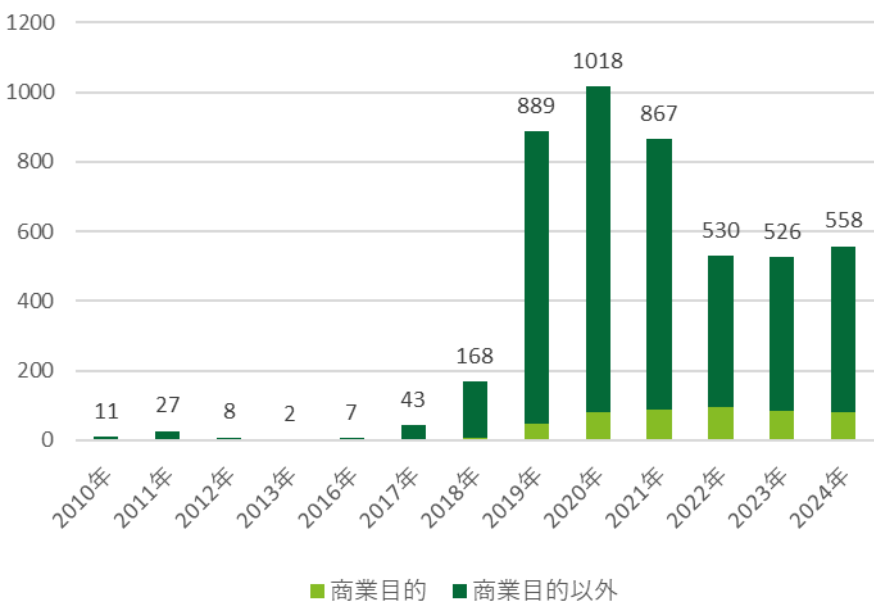
NHS England Data Access Request Service (DARS) 概要

① データ利活用基盤・データベースの概要

名称	NHS England DARS
運営主体	NHS England
運用開始年	2010年
関連法	NHS Act 2006 Health and Social Care Act 2012
設置目的	<ul style="list-style-type: none">医療機関のデータを収集・管理し、それを分析・提供することで、医療・社会福祉・政策立案に役立てるDARSは、外部研究者を含む利用者がNHSの医療・社会福祉データを申請できるようにするために設置され、データ利用プロセスをサポートしている
利用者	DARS の利用申請で承認を得た研究者 (NHS内外可)
データソース	<ul style="list-style-type: none">NHSが保有する情報やその他の情報保有主体がNHSへ提供した情報に基づき作成されたテーマ別のデータセット39種類 (2024年12月時点)NHS外の主な情報提供機関はPublic Health England 公衆衛生局 (がんデータ)、Department of Health and Social Care 社会保険福祉省 (医療従事者勤務データ)、Office for National Statistics (国家統計局) 等が挙げられる
データ規模	<ul style="list-style-type: none">全人口の90%以上 (データセットにより異なる)※ 2021年時点、英国心臓財団データサイエンスセンターは、NHSのデータを活用し約5,440万人(人口の96%)のデータ解析を行いCovid-19と心疾患の関連性に関して研究を発表している

データ連結手法	NHS番号、生年月日、性別等の識別子を複合的に使用して連結
標準化	各DBで利用マスターが異なり、データセットに応じて連結可能と不可なものがある。2023年8月時点、NHSでOMOP-CDMの導入に向けて合意が図られている
利用実績 (新規データ共有契約数)	530件、うち商業目的 95件(2022年) 526件、うち商業目的 86件(2023年) 558件、うち商業目的 80件 (2024年11月末)

<年毎の新規データ共有契約(DSA)数>



出所 : <https://www.england.nhs.uk/>

NHS Englandはデータ利用に係る契約やデータセットの抽出・配信・結合などに対し、料金を徴収している

① データ利活用基盤・データベースの概要

NHS Englandのデータ利用に係る料金表

	サービスの種類	単価 (£)	単価 (円)	備考
申請	新規申請料	1,030	202,704	
	申請内容見直し料	820	161,376	・ 申請内容の修正・更新、期間延長に対して徴収
	年間レビュー料	500	98,400	・ 1年以上の契約に対し、契約の確認目的で徴収 (例)3年契約の場合、2年目と3年目に徴収
データセット抽出等	データセット配信料	930	183,024	・ データセット配信毎に徴収 (例)四半期毎に1回、3年間送信した場合、4×3=12回徴収
	データセット抽出料	320	62,976	・ データセット毎に徴収 (例1)1年毎のデータセットを3年分の場合、1×3=3回徴収 (例2)月毎のデータセットを3年分の場合、12×3=36回徴収
	データセット結合料	2,060	405,408	・ NHS Digitalが保有する患者情報とデータセットの結合ごとに徴収 (例)患者情報を結合した四半期毎のデータセットA・Bを3年間要求した場合、4×2×3=24回徴収
その他	データセットの初期トレーシング	420	82,656	・ データセットの質を担保する場合に実施する
	データアクセス環境の構築	3,200	629,760	・ 利用者専用のアクセス環境を構築する場合に毎年徴収
	データセットの集計	800	157,440	・ データテーブル毎に徴収
	サブライセンス	10,000	1,968,000	・ サブライセンス毎に毎年徴収

出所 :<https://digital.nhs.uk/services/data-access-request-service-dars/data-access-request-service-dars-charges>

3年間のDSA契約において、過去21年分の5つのデータセットを利用する場合、総額費用は約556万円である

① データ利活用基盤・データベースの概要

NHS Englandのデータ利用に係る料金例

【申請内容】

- 新規に3年間のDSAを締結
- 2000～2021年の21年間の①入院患者ケア、②外来患者ケア、③アクシデント・救急ケアに関する病院統計(HES)データセットと2000～2021年に渡る④死亡データの病院統計(HES)データセット、⑤死亡登録データセットを要求し、配信は1回のみ

1イギリスポンド=196.8円 (2024年12月23日時点)で換算

サービスの種類		単価 (£)	単価 (円)	回数 データ数	合計 (円)	
申請	新規申請料		1,030	202,704	1	202,704
	年間レビュー料		500	98,400	2	196,800
データ セット 抽出等	データセット配信料		930	183,024	1	183,024
	データセット抽出料	①入院患者ケア(HES)	320	62,976	21	1,322,496
		②外来患者ケア(HES)	320	62,976	21	1,322,496
		③アクシデント・救急ケア(HES)	320	62,976	21	1,322,496
		④死亡データの病院統計(HES)	320	62,976	1	62,976
		⑤死亡登録データセット	320	62,976	1	62,976
	データセット結合料	①②③④(HES)と患者情報を結合	2,060	405,408	1	405,408
		⑤と患者情報を結合	2,060	405,408	1	405,408
その他	データセットの初期トレーニング		420	82,656	1	82,656
合計						5,569,440

NHS Englandの収入の大部分は保健省からの助成金だが、データ利用について料金を徴収し収入を得ている

① データ利活用基盤・データベースの概要

NHS England の収入と支出

	2020年-2021年	2021年-2022年	2022年4月 -2023年1月
	£1,000 (百万円)	£1,000 (百万円)	£1,000 (百万円)
収入			
保健省からの助成金	483,554 (95,163)	612,574 (120,555)	498,203 (98,046)
収入	45,298 (8,915)	44,863 (8,829)	31,754 (6,249)
助成金と収入	528,852 (104,078)	657,437 (129,384)	529,957 (104,296)
支出			
人件費	-207,651 (-40,866)	-263,142 (-51,786)	-244,016 (-48,022)
営業支出	-296,045 (-58,262)	-380,765 (-74,935)	-284,227 (-55,936)
固定資産除去損	-2,759 (-543)	-2,330 (-459)	-1,922 (-378)
予備金の移動	2,065 (406)	-377 (-74)	-10 (-2)
支出計	-504,390 (-99,264)	-646,614 (-127,254)	-530,175 (-104,338)
収支	24,462 (4814)	10,823 (2130)	-218 (-43)

NHS England の収入の内訳

◆「プログラムおよびサービスの提供(緑部分)」には、データ利用に係る申請料、データ配信料、データ結合料等が含まれ、「調査およびデータ収集」には、国民の健康調査やその他データ収集に関連する費用が含まれる

	2020年-2021年	2021年-2022年	2022年4月 -2023年1月
	£1,000 (百万円)	£1,000 (百万円)	£1,000 (百万円)
契約収入			
プログラムおよびサービスの提供	35,203 (6,928)	36,920 (7,266)	27,145 (5,342)
調査およびデータ収集	1,208 (238)	2,495 (491)	1,264 (249)
助成金収入	1,679 (330)	1,302 (256)	0 (0)
手数料など	3,489 (687)	2,857 (562)	2,237 (440)
契約収入計	41,579 (8,183)	43,574 (8,575)	30,646 (6,031)
契約外収入			
プログラムおよびサービスの提供	293 (58)	234 (46)	239 (47)
物品の販売	2,175 (428)	2,175 (428)	730 (144)
営業外収益	1,087 (214)	1,087 (214)	139 (27)
実習生納付金の利用	164 (32)	189 (37)	1,108 (218)
契約外収入合計	3,719 (732)	1,289 (254)	1,108 (218)
収入合計	45,298 (8,915)	44,863 (8,829)	31,754 (6,249)

医療・社会福祉情報の二次利用に関して、「データサービス」「IT運用」「プラットフォーム」「サイバー運用」等が中心的な支出であり、高額な人件費や情報技術費が投じられている

① データ利活用基盤・データベースの概要

分野別支出 (2021年-2022年)

単位: 百万円

活動	保証・リスク管理	コーポレートサービス	データサービス	IT運用	プラットフォーム	サイバー運用	製品提供	戦略、政策 およびガバナンス	COVID-19対応	中央（セグメント未 割り当て）	合計
収入	0	-4	-489	-133	-59	0	-2884	-161	-226	-430	-8829
人件費	71	381	1299	8745	1472	1896	13660	3869	4757	-127	51787
専門家報酬	18	140	651	2651	5926	2331	4926	439	14375	-142	38597
情報技術（IT）	1	23	151	12475	1112	2224	2406	138	12868	-1	32968
宿泊費用	0	197	0	5	1	0	2	0	0	-8	1974
旅費・日当	0	9	1	10	1	4	17	5	2	0	130
マーケティング、研修、イベント、広報	0	44	11	3	11	68	15	27	13	-1	687
オフィスサービス	0	26	5	60	4	15	14	32	3	0	442
その他	0	9	0	34	0	0	0	0	0	7	136
固定資産除却損	0	16	6	74	1	0	35	0	141	-11	459
減価償却および償却	0	30	206	2107	56	1296	2947	112	4602	-21	13464
減損損失 – 設備および機械	0	0	0	25	5	0	0	0	0	0	31
減損損失 – 無形資産	0	0	0	0	0	0	2	0	7908	0	7910
中央コストの再配分	6	-551	225	814	-1334	148	1894	251	1432	0	0
非人件費	26	-56	1255	18257	5782	6087	12257	1005	41343	-179	96798
純支出	97	321	2065	26869	7195	7984	23033	4713	45874	-736	139755

<分野詳細>

- データサービス：医療・福祉システムのデータ管理、集計、統計提供を担当。データ品質の向上を図り、政策立案や医療研究、ライフサイエンスの発展に貢献できるように取り組む。また、データプライバシーが厳守されるように情報管理を行う
- IT運用：NHSデジタルが提供する医療・福祉向けの全システムとサービスが、安全かつ信頼性を維持して運用されるよう管理
- プラットフォーム：医療・福祉システム全体のデジタルサービスプロバイダーを結ぶ中核的なインフラとプラットフォームを提供。データサービスを支える基盤構築・運営費用も含む
- サイバー運用：医療・福祉システムがサイバー攻撃に対して強靱で、安全性を保てるよう運用を担保

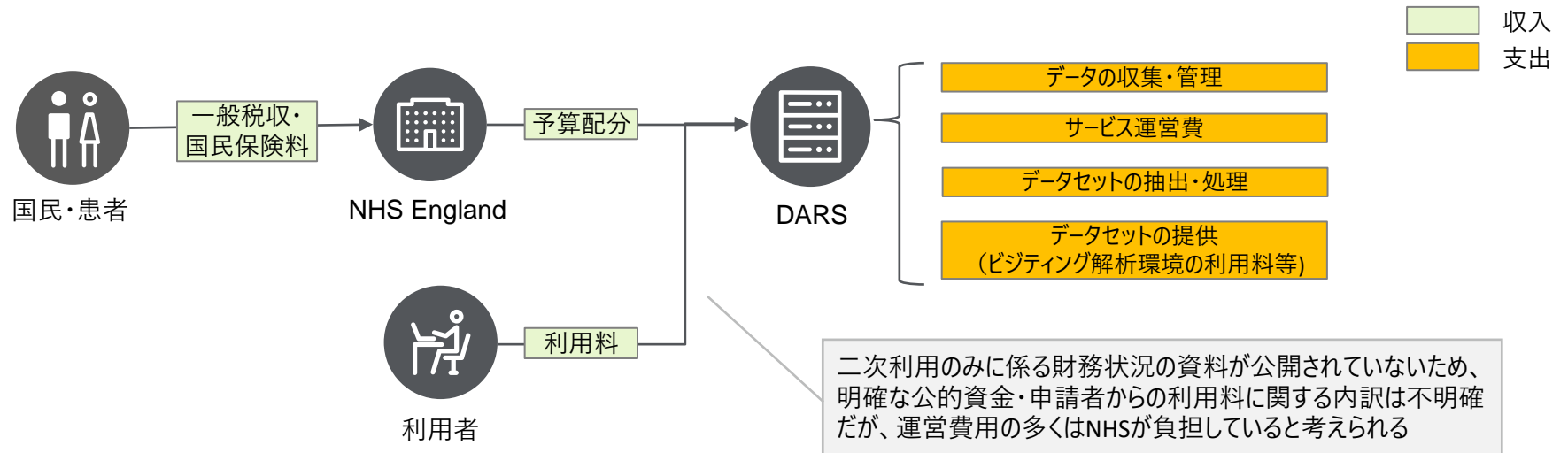
DARSは主に公的資金で運営され、データそのものに関する費用は発生しないが、抽出・提供にかかる費用が利用条件に応じて請求されるコスト回収型の仕組みを採用している

NHS DARS のビジネスモデル

① データ利活用基盤・データベースの概要

- DARSは公的資金で運営されており、コスト回収ベースで運営されている
- 利用者には、データそのものに対する費用は請求されないが、サービス提供に必要なデータ抽出・提供に関する費用は徴収される。
なお、利用者に請求する料金は以下の要素に基づいて計算されており、利用申請の段階で合意が図られる。
 - ・ データ量やデータ結合に係る費用
 - ・ データの提供回数
 - ・ ビジティング解析環境の利用
 - ・ その他要望の内容の複雑さ、追加的なデータ処理作業が生じた場合等には追加料金が発生する

NHS DARS の収支フロー



NHSの医療情報の二次利用は主に学術・公的機関による観察研究が中心で、そのインパクトの定量化が難しいことから、今後は評価方法と還元方法の検討が求められている

① データ利活用基盤・データベースの概要

英国の医療情報のデータフローのマッピング及びその評価に関する文献概要

タイトル	Mapping and evaluating national data flows: transparency, privacy, and guiding infrastructural transformation (英国のデータフローのマッピング及びその評価；透明性、プライバシー、インフラ改革の指針)
著者	Joe Zhang (Imperial College London), Jess Morley (University of Oxford), Jack Gallifant (Imperial College London) 他
発行日	2023年10月 (The Lancet Digital Health に掲載)
目的	<ul style="list-style-type: none">NHSにおける患者の医療情報のデータフローを全国規模でマッピングし、包括的な利用状況を明らかにする。更に、データの利用状況を可視化することでデータインフラの改革に関する示唆を得ることを目的とした
手法	<ul style="list-style-type: none">NHSの一次(GP)及び二次医療機関(病院)から抽出される患者の電子化された医療情報を対象に、2021年4月～2022年4月の期間でデータの流れを追跡データフローの特定にあたり、次の情報源を参考に行っている：NHSや学術機関の年次報告書やデータ戦略に関する文書、GDPR関連の法的文書、医療機関および臨床委員会グループへの情報提供依頼を通じて得た患者データの共有や二次利用に関する情報。また、MEDLINEデータベースを用いてNHS関連データベースに関するスコーピングレビューを実施し、研究におけるデータベース利用状況を調査明らかになったデータの起点と提供先に対して、データの内容、量、提供手法等を特定し主要なデータフローを視覚化している
結果	<p><データフローの全体像></p> <ul style="list-style-type: none">英国では6,544の一次(GP)と216の二次医療機関(病院)が約56万人の患者データを収集し、460以上の組織(学術機関、企業、非営利団体等)がそのデータを二次利用しているデータ利用者の多く(約78%)はビジティング解析環境等を介さずにデータを受取っている可能性がある <p><データの価値と国民への還元></p> <ul style="list-style-type: none">NHSはデータインフラ投資として、ビジティング解析環境の開発に最大2億ポンド、さらに国データプラットフォームに4億8千万ポンドを割り当てている大規模データベースは、データアクセスに係るコストが高額であり特に学術機関等の研究者にとっては負担が大きいと考えられるデータの大部分は観察研究に使用されるため、長期的な公共の利益をもたらすが、直接的なインパクトの定量化は困難。ただし、今後はデータフロー全体での金銭的価値の評価方法を検討し、データ提供者への利益還元を実現するモデルを構築する必要がある同様のデータベースが重複する場合、新たな価値創出には繋がらない。また、NHSのビジティング解析環境を通じたデータ提供に統一することで、複数のデータインフラの構築・維持にかかるコストを削減し、共同研究の効率化や質の向上が期待できる

学術・研究機関等	<ul style="list-style-type: none">NHSデータの最大の利用者は、観察研究やサービスの評価や全国監査を行う260の学術機関、研究機関、非営利団体や臨床機関であった
一般企業	<ul style="list-style-type: none">合計106社の製薬、ライフサイエンス、データ分析企業が創薬やワクチン・治療に関する評価にデータを使用していた
政策・コンサル等	<ul style="list-style-type: none">この層では、NHS・英国政府が情報を利活用するほか、コンサルティング企業等との共同調査による医療政策の検討等に利用されている
その他データ抽出機関	<ul style="list-style-type: none">NHS以外に、37の学術機関、一般企業、公的機関等が独自でデータを抽出・管理し第三者に提供している。更に、分析サービスを提供する場合もある
NHSによるデータ抽出	<ul style="list-style-type: none">全国規模の医療情報はNHSイングランド及び地域のNHSTラスト等によって収集されている
医療機関	<ul style="list-style-type: none">電子化された医療情報は、NHSイングランド全域の一次医療機関 (GP) (n=6544) 及び二次医療機関 (病院) (n=216) から抽出され、5500万人の患者データが該当する

NHSデータを活用した研究事例及び期待される成果は、公式ウェブサイトで紹介されている

研究事例 (1/2)

① データ利活用基盤・データベースの概要

研究分類	プロジェクト名	実施者	概要	期待されるアウトカム
疫学研究	ASCEND PLUS ケーススタディ	<ul style="list-style-type: none"> NHS DigiTrials オックスフォード大学 	<ul style="list-style-type: none"> 2型糖尿病患者は、心血管疾患のリスクが従来2倍とされてきた 本臨床試験は、<u>2型糖尿病患者における心臓発作、脳卒中、その他心血管疾患の問題を予防する可能性のある経口薬を検証するもの</u> 英国全土の心臓発作や脳卒中の既往がない55歳以上の2型糖尿病患者20,000人を対象とした 	<ul style="list-style-type: none"> 臨床試験が成功すれば、<u>世界中の約4億人の糖尿病患者に対し、心疾患リスクの予防医療に寄与</u>することが見込まれる NHS DigiTrials の活用により、臨床試験の分散化が可能となり、参加者の多様性が向上し、研究の妥当性が強化された。また、参加者募集期間を6年から2年に短縮することで、臨床試験のスピードが大幅に加速
	Our Future Health ケーススタディ	<ul style="list-style-type: none"> NHS DigiTrials Our Future Health Partnerships 	<ul style="list-style-type: none"> 健康寿命の短縮や65歳以上の54%が複数の疾患を抱える現状が、人口動態や経済に影響を与えており、公衆衛生施策の強化が求められている NHS DigiTrialsは、約500万人のボランティアを募集し、血液サンプル、医療記録、身体測定データ、アンケート情報を収集し、<u>データをSDE上で利用可能にすることで、英国の疾患予防、早期発見、治療に関する多数の研究を支援</u> 	<ul style="list-style-type: none"> 当該プロジェクトは、英国政府から1億3,000万ポンド、ライフサイエンス企業や慈善団体から1億8,000万ポンドの資金提供を受けて開始され、<u>複数の大規模な研究が進行中</u>である 現在行われている研究例は以下のとおり： <ul style="list-style-type: none"> 非喫煙者における電子タバコのリスク評価 心代謝の遺伝リスク要因の解析 免疫機能障害と遺伝子・環境要因の関連
	INFORM ケーススタディ	<ul style="list-style-type: none"> NHS England AstraZeneca Evidera 	<ul style="list-style-type: none"> <u>免疫抑制状態の人々におけるCOVID-19の重症化リスクの特定・評価を目的とした研究</u> NHSの安全データ環境(SDE)上で、一次・二次医療機関から収集された英国全人口の約25%のCOVID-19、ワクチン接種状況、処方薬、死亡率などのデータを結合し、ほぼリアルタイムで解析を行った 	<ul style="list-style-type: none"> 3回以上のワクチン接種を受けたにもかかわらず、<u>免疫抑制状態の人々は一般人口と比較して重症COVID-19のリスクが13倍高いことが判明</u>し、さらなる予防策の必要性が示された 高リスクの免疫抑制群を特定することで、予防的介入やワクチンプログラムへの適用が期待される。その他<u>治療方針の改善や資源配分の最適化を通じて、高リスク患者の治療成果の向上が見込まれる</u>

NHSデータを活用した研究事例及び期待される成果は、公式ウェブサイトで紹介されている

研究事例 (2/2)

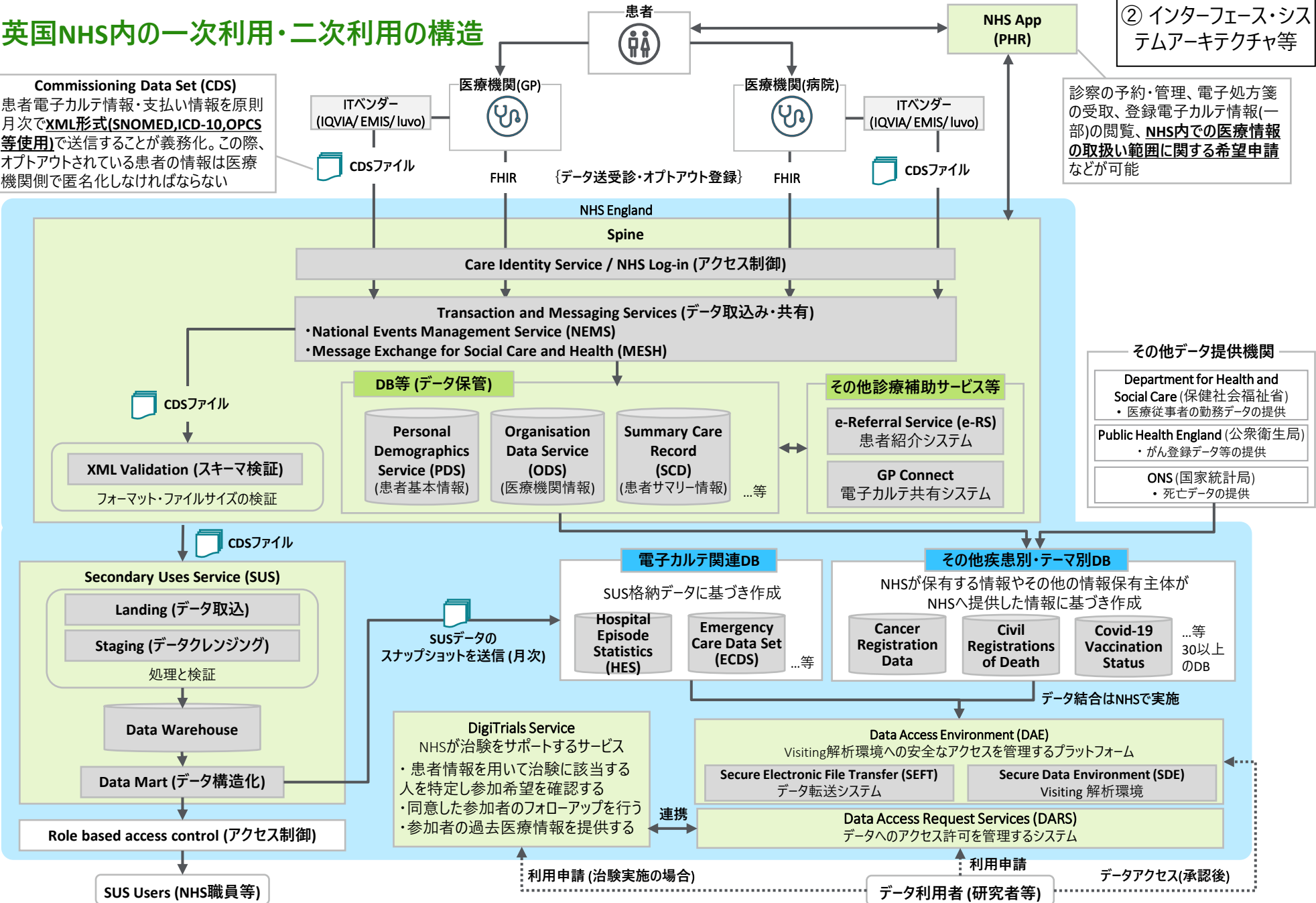
① データ利活用基盤・データベースの概要

研究分類	プロジェクト名	実施者	概要	期待されるアウトカム
疫学研究	C型肝炎の特定に係るケーススタディ	<ul style="list-style-type: none">Imperial College Healthcare NHS TrustDiscover-Now London SDE	<ul style="list-style-type: none">英国では、2022年時点で70,000人以上がC型肝炎に罹患しており、そのうち80%以上が貧困地域に集中している本研究は、GPデータを活用して<u>リスクの高い患者を特定し、一次医療の場でC型肝炎の診断と治療を行い、新たな感染を防ぐことを目的としている</u>	<ul style="list-style-type: none">研究調査により479,471人のリスクの高い患者が特定され、高リスク患者が多い診療所での戦略的な介入が可能となったC型肝炎の高発生地域や地域ごとの特性のマッピングにより、さらに効果的な対策が期待されているC型肝炎撲滅に向けた政策立案や公衆衛生活動の強化にも寄与することが期待される
AI・アルゴリズム研究	TRICORDER ケーススタディ	<ul style="list-style-type: none">Imperial College LondonDiscover-Now London SDEDiscover NowBetsi-Cadwaladr Health BoardNational Institute for Health and Care Research	<ul style="list-style-type: none">心不全は英国全人口の2%以上に影響を与え、NHS予算の4%（心不全による入院だけで年間約20億ポンド）を消費する疾患であり、多くのがんよりも死亡リスクが高い本研究は、AI搭載の聴診器を活用及びGPの診療記録からリスクの高い患者を特定し、一次医療における心不全の診断と治療を支援することを目的としているAI搭載聴診器「Eko-DUO」を200以上のGP診療所で使用し、250万人以上の患者データを臨床試験データと統合することで、AIツールの有効性を評価し、さらに心不全や関連する心血管疾患の早期発見におけるコスト効果を検証することを目指している	<ul style="list-style-type: none">心不全診断の精度と効率を高めることで、医療コストの削減と患者の予後改善に大きく寄与すると期待されている現在、患者の5人中4人が緊急入院時に診断されており、この遅れが早期死亡リスクを大幅に増加させている早期診断により救急外来の利用を削減し、患者の治療効果を高めると同時に医療コストを削減できる可能性がある「Eko-DUO」のAI技術を英国の導入することで、患者1人当たり約2,400ポンドの医療費削減が期待されており、全国導入によりNHSで約1億ポンドの医療費削減が見込まれている

英国NHS内の一次利用・二次利用の構造

一次利用

二次利用



NHSではデータ一次利用や二次利用の利用場面に応じて、各種システム、機能を提供している

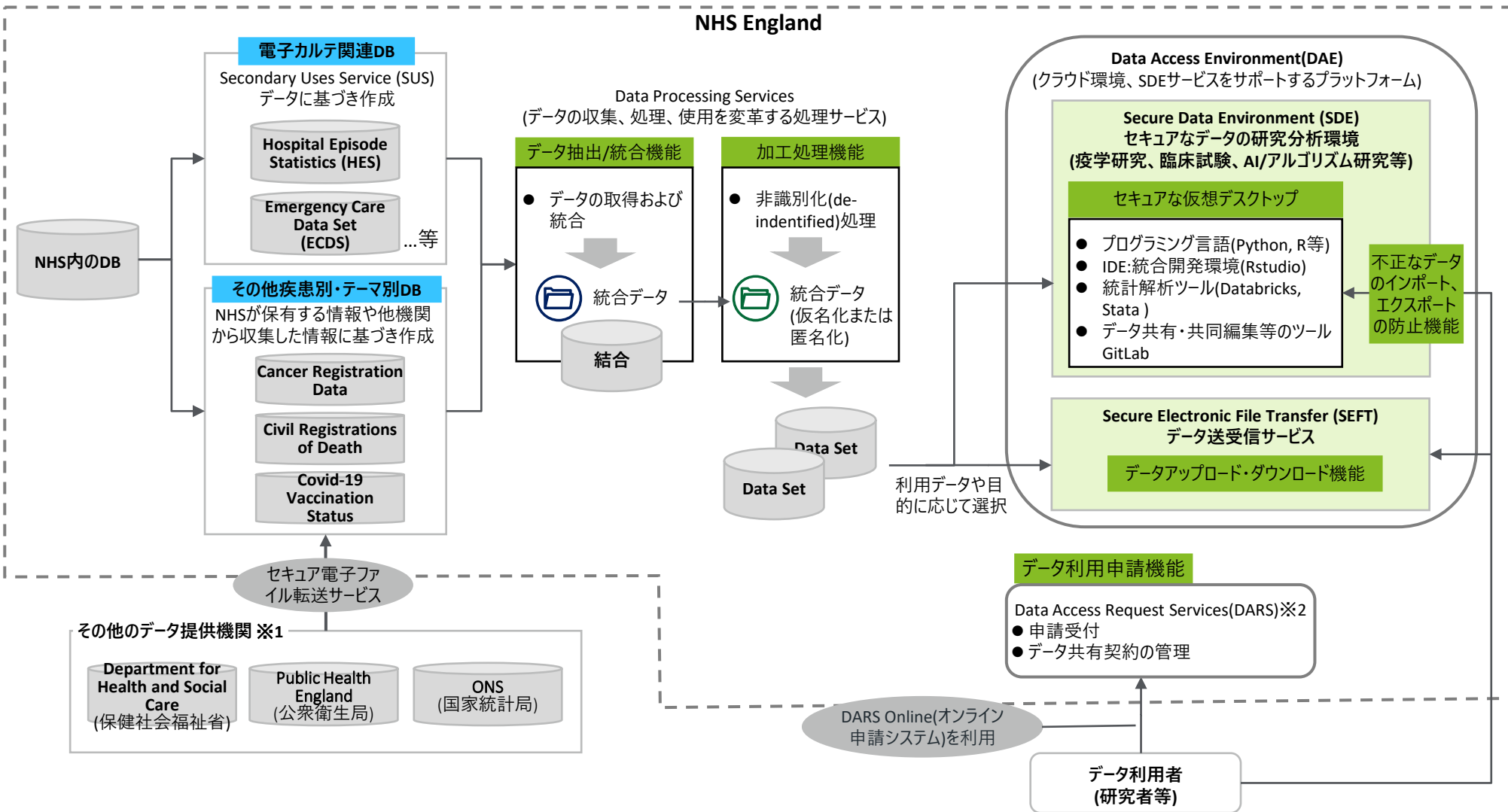
② インターフェース・システムアーキテクチャ等

英国NHSの一次利用・二次利用に係る主なシステム・機能

データ利用場面	名称	略語	概要
一次利用	SPINE	—	医療機関やNHS職員の間で診療に関する医療情報の交換を可能とするプラットフォーム
	E-Referral Service	e-RS	GPが患者を病院へ紹介し、患者が希望する医療機関をリアルタイムで予約できる電子紹介プラットフォーム
	GP Connect	—	英国の主要GPプロバイダシステムの情報を統合し、外部医療従事者に患者の電子カルテ情報を共有するプラットフォーム
	Care Identity Service	—	スマートカードやその他の認証手段を用いて患者データへのアクセスを管理する機能
	National Events Management Service	NEMS	イベントが発生した際に、その情報をプッシュ通知の方式で閲覧者に自動的に送信する機能※イベントの例：患者の死亡、新生児の予防接種など
	Message Exchange for Social Care and Health	MESH	ヘルスケア関連組織間で安全かつ直接的にデータを共有できるデータ共有方式/機能（ファイルでの共有が可能）
二次利用	Secondary Uses Services	SUS	NHS内で患者情報を治療以外（サービスの向上、施策検討等）、二次利用のための情報を提供するシステム
	Data Access Environment	DAE	Visiting解析環境への安全なアクセスを管理するプラットフォーム
	Data Access Request Services	DARS	研究者等のデータ利用者がデータの申請を行い、申請を承認するシステム
	DigiTrials Service	—	治験参加者の過去の医療情報を研究者に提供し、治験の成果向上をサポートするサービス

データソース

情報連携基盤



※1 英国ではHealth and Social Care Act 2012の第254条に基づき、国務大臣は医療サービスまたは社会福祉の提供に関連する職務の遂行に関連して、情報の収集が適切である場合、NHS Englandに情報システムの構築および運用を指示することができる

※2 今後、DAEはDARSと統合され、リクエストとデータ利用がより迅速かつ容易になる予定(2024年5月時点)

出所: <https://www.gov.uk/government/publications/nhs-englands-protection-of-patient-data/nhs-englands-protection-of-patient-data#:~:text=section%20of%20the%202012,the%20provision%20of%20health%20services>

No.	データセット名	概要	データ収集方法	コーディング	データ提供方法	データ結合	保管形態	更新頻度
1	Adult Psychiatric Morbidity Survey (APMS)* 成人精神疾患調査	英国の成人人口 (16歳以上)における治療済及び未治療の精神疾患の有病率に関するデータ	NHSが定期的に全国調査を実施	－	抜粋送付	－	－	7年毎に調査実施
2	Breast and Cosmetic Implant Registry (BCIR) 乳房・美容インプラント登録	NHS・私立医療機関で乳房・美容インプラント手術を受けた患者の詳細を含むデータセット	処置を実施した医療機関が報告	無	抜粋送付	不可	顕名	日次
3	Cancer Registration Data (Cohort based studies only) がん登録データ (コホート研究のみ)	がんの種類、部位、登録日、登録場所等の情報が含まれるデータセット	医療機関から月次で収集	・ICD10	抜粋送付	可	顕名	月次
4	Civil Registrations of Death 死亡登録データ	国家統計局 (ONS) が提供する、1993年以降にイングランドとウェールズで登録されたすべての死亡者情報	自治体から国に登録されONSに情報提供される	・ICD9, ICD10	SDE・ 抜粋送付	可	顕名	週次
5	Community Services Data Set (CSDS) コミュニティ・サービス・データセット	子供、若者、成人向けの公的資金によるコミュニティサービスに関する患者レベルのデータセット	公的資金で運営によるケア事業者から Strategic Data Collection Service Cloud (SDCS Cloud) でデータ提出	・ICD10, ・SNOMED-CT ・OPCS-4 ・Read Codev2 ・Read Codev3	SDE・ 抜粋送付	可	仮名	月次
6	COVID-19 Electronic Prescribing and Medicines Administration (ePMA) in Secondary Care COVID-19 病院電子処方・医薬品管理データ	コロナ禍 (2019-2023年) において患者に処方・投与された医薬品に関する情報を電子処方・医薬品管理 (ePMA) システムから抽出したデータセット	医療機関の電子処方・医薬品管理 (ePMA) システムから情報抽出	・SNOMED-CT ・dm+d ・UCUM	SDE・ 抜粋送付	可	顕名	日次
7	COVID-19 Ethnic Category Data Set COVID-19 民族カテゴリーデータセット	NHSが患者の民族カテゴリーのカバー率を高めるためにGPデータと病院統計データを組合わせて作成したデータセット	NHSが保有するデータセットに基づいて作成	無	SDE・ 抜粋送付	可	顕名	月次
8	COVID-19 General Practice Extraction Service (GPES) Data for Pandemic Planning and Research (GDPPR) COVID-19 パンデミック計画・研究用GP診療抽出データ	コロナ禍流行への対応を支援するために、GP患者データを収集したもの	General Practice Extraction Service (GPES) 経由でGPデータ抽出	・SNOMED-CT	SDE	可	顕名	隔週

*No.1 APMS は 2024年4月以降 UK Data Service 経由の申請・利用へ移行
出所: <https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue>

No.	データセット名	概要	データ収集方法	コーディング	データ提供方法	データ結合	保管形態	更新頻度
9	COVID-19 Second Generation Surveillance System (SGSS) COVID-19第2世代監視システム	新型コロナウイルス感染症(COVID-19)の抗原検査レポートに関する人口統計および診断情報	SPINEのTransaction Messaging Services 経由でNHSが情報収集	無	SDE・ 抜粋送付	可	顕名	日次
10	COVID-19 UK Non-hospital Antigen Testing Results (Pillar 2) 病院外のCOVID-19抗原検査結果	検査機関、ドライブスルー検査、サテライト検査センター、ケアホーム等病院以外の抗原検査結果に関するデータ	検査機関がNHSに報告	無	SDE・ 抜粋送付	可	顕名	日次
11	COVID-19 Vaccination Adverse Reactions COVID-19 ワクチン接種時副反応	COVID-19ワクチン接種15分以内に発生した副反応に関するデータ	ワクチン実施機関使用のアプリや電子カルテシステム経由でNHSに送信される	無	SDE・ 抜粋送付	可	顕名	日次
12	COVID-19 Vaccination Status COVID-19 ワクチン接種状況	英国内でCOVID-19ワクチンを受けた患者の詳細データ	ワクチン実施機関使用のアプリや電子カルテシステム経由でNHSに送信される	・SNOMED-CT ・dm+d	SDE・ 抜粋送付	可	顕名	日次
13	Demographics (Cohort based studies only) 人口統計データ (コホート研究のみ)	英国で登録されている全患者の人口統計に関するデータ	SPINEに格納されているPersonal Demographics Service(PDS)に基づき作成	無	抜粋送付	可	顕名	日次
14	Diagnostic Imaging Data Set 画像診断データセット	NHSプロバイダーの放射線情報システムから抽出された、MRI検査やX検査等に関するデータ	放射線情報システム経由でNHSに情報送信	・NCIP ・SNOMED-CT	抜粋送付	可	仮名	月次
15	Emergency Care Data Set (ECDS) 救急医療データセット	NHS病院における救急受診の詳細に関するデータ	医療機関が定期的にCDSファイルを送信しSUSで処理の上データセット構築	・SNOMED-CT ・TFCs	抜粋送付	可	仮名	月次
16	General Practice Workforce Data GP労働データ	GPで働く医師、看護師、スタッフに関する業務の詳細が含まれるデータ	医療機関が Strategic Data Collection Service Cloud (SDCS Cloud) でデータ提出	無	抜粋送付	可	顕名	月次
17	Hospital Episode Statistics 病院エピソード統計	NHS病院において、救急外来、急性期ケア、入院、外来診療に関する患者10億件を超えるデータ	医療機関が定期的にCDSファイルを送信しSUSで処理の上データセット構築	・ICD9, ICD10 ・SNOMED-CT ・OPCS-4 ・TFCs	SDE・ 抜粋送付	可	仮名	月次

No.	データセット名	概要	データ収集方法	コーディング	データ提供方法	データ結合	保管形態	更新頻度
18	Improving Access to Psychological Therapies (IAPT) 心理療法へのアクセス改善	うつ病・不安等のために心理療法へ紹介された患者に関するデータ	医療機関等が Strategic Data Collection Service Cloud (SDCS Cloud) でデータ提出	・SNOMED-CT ・ICD10	SDE・ 抜粋送付	可	仮名	月次
19	Maternity Services Data Set マタニティー・サービス・データセット	マタニティーサービスによって提供された産科医療に関する母子のデータ	医療機関等が Strategic Data Collection Service Cloud (SDCS Cloud) でデータ提出	・ICD9, ICD10 ・OPCS-4 ・SNOMED-CT	SDE・ 抜粋送付	可	顕名	月次
20	Medicines Dispensed in Primary Care (NHSBSA) プライマリケアで使用された医薬品	GP、地域クリニック、看護ケアサービス、歯科医等のプライマリケアで使用された医薬品に関するデータ	払戻しのためNHSに提出される電子処方箋・紙処方箋のデータに基づく	・dm+d	SDE・ 抜粋送付	可	顕名	月次
21	Mental Health data (MHMDS, MHLDDS, MHSDS) メンタルヘルス・データ	NHSが提供する精神保健サービスや学習障害サービスを受けている成人に関するデータ	医療機関等が Strategic Data Collection Service Cloud (SDCS Cloud) でデータ提出	・ICD10 ・SNOMED-CT	SDE・ 抜粋送付	可	顕名	月次
22	Mental Health of Children and Young People (MHCYP) こどもと若者のメンタルヘルス調査	こども・若者のメンタルヘルス傾向について全国調査したデータ	NHSが定期的に全国のこども・若者を対象に調査を実施しデータ収集	・ICD10	抜粋送信	不可	匿名	不定期
23	National Diabetes Audit 全国糖尿病監査	プライマリケア及びセカンダリーケアの両方において、糖尿病医療の有効性を測定したもの	General Practice Extraction Service (GPES) 及びClinical Audit Platform 経由で情報収集	・ICD10 ・SNOMED-CT	抜粋送信	可	顕名	年次
24	NDRS Cancer Consolidated Data Set NDRS 統合がんデータセット	National Cancer Registrations, Cancer Pathways, Rapid Cancer Registrations を統合したがん登録データ	NHSの全国疾病登録サービス (NDRS)から情報抽出	・ICD10 ・OPCS-4 ・SNOMED-CT ・NICIP	SDE・ 抜粋送付	可	顕名	月次
25	NDRS Cancer Registration (pre-1995) NDRS がん登録 (1995年以前)	1985年-1995年の間に登録されたがん患者の情報を含むデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	・ICD10 ・OPCS-4	抜粋送付	NDRSデータのみ可	顕名	年次

No.	データセット名	概要	データ収集方法	コーディング	データ提供方法	データ結合	保管形態	更新頻度
26	NDRS Lung Cancer Data Audit (LUCADA) NDRS 肺がんデータ監査	肺がん・中脾腫と新d何された患者に対して病院紹介、診断、治療、治療結果等に関するデータ。2005-2014年のデータを含む	NHSの全国疾病登録サービス (NDRS)から情報抽出	・ICD10 ・OPCS-4 ・SNOMED-CT ・UICC	抜粋送付	NDRSデータのみ可	顕名	不定期
27	NDRS National Cancer Diagnosis Audit (NCDA) NDRS 全国がん診断監査	がんと診断された患者の診断経路に関するデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	無	抜粋送付	NDRSデータのみ可	顕名	不定期
28	NDRS National Cancer Patient Experience Survey (CPES) NDRS 全国がん患者体験調査	がんと診断された患者の治療経験等に関する情報を収集するために全国調査を実施したもの	NHSの全国疾病登録サービス (NDRS)から情報抽出	無	抜粋送付	NDRSデータのみ可	顕名	年次
29	NDRS National Lung Cancer Audit (NLCA) NDRS 全国肺がん監査	NDRS肺がんデータ監査に続く、2015年以降のデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	・ICD02 ・OPCS-4 ・SNOMED-CT ・UICC ・NICIP ・IMD	抜粋送付	NDRSデータのみ可	顕名	不定期
30	NDRS National Radiotherapy Data Set (RTDS) NDRS 全国放射線治療データセット	2009年以降の放射線治療やその他の処置に関するデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	・dm+d ・ICD10 ・OPCS-4 ・SNOMED-CT	抜粋送付	NDRSデータのみ可	顕名	月次
31	NDRS Quality of Life of Cancer Survivors in England 英国がん生存者QOL調査	2011-2012年にがん生存者約5000人を対象に実施したQOL調査に関するデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	無	抜粋送付	NDRSデータのみ可	顕名	不定期
32	NDRS Quality of Life of Colorectal Cancer Survivors in England: Patient Reported Outcome Measures Survey (PROMS) 大腸がん生存者のQOL調査	2013年に大腸がん生存者約34000人を対象に実施したQOL調査に関するデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	無	抜粋送付	不可	匿名	不定期
33	NDRS Somatic Molecular Data Set NDRS 体細胞分子データセット	体細胞分子検査で遺伝変異の種類、検査日、検査数等のデータが含まれる	NHSの全国疾病登録サービス (NDRS)から情報抽出	無	抜粋送付	NDRSデータのみ可	顕名	年次

No.	データセット名	概要	データ収集方法	コーディング	データ提供方法	データ結合	保管形態	更新頻度
34	NDRS Systemic Anti-Cancer Therapy data set (SACT) NDRS 抗がん剤治療データセット	2012年4月以降に化学療法を受けた患者の臨床詳細に関するデータ	NHSの全国疾病登録サービス (NDRS)から情報抽出	無	抜粋送付	NDRSデータのみ可	顕名	月次
35	Personal Social Services – Adult Social Care Survey (ASCS) 個人福祉サービス-成人福祉ケア調査	Councils with Adult Social Services Responsibilities (成人福祉サービス担当評議会) が毎年実施する成人福祉サービスに関する全国調査のデータ	NHS 成人福祉サービス担当評議会からデータ収集	無	抜粋送付	不可	－	年次
36	Personal Social Services – Survey of Adult Carers in England (SACE) 個人福祉サービス-英国成人介護者調査	Councils with Adult Social Services Responsibilities (成人福祉サービス担当評議会) が隔年で実施する成人介護者に関する全国調査のデータ	NHS 成人福祉サービス担当評議会からデータ収集	無	抜粋送付	不可	－	隔年
37	Sentinel Stroke National Audit Programme (SSNAP) センチネル脳卒中全国監査	脳卒中患者に提供されるケアやサービスについて調査し、全国臨床ガイドライン基準に照らし合せ評価するためのデータ	NHSが大学に委託し定期的に調査を実施	無	SDE・ 抜粋送付	可	顕名	四半期毎
38	Summary Hospital-level Mortality Indicator (SHMI) 病院レベル死亡率指標サマリー	各医療機関で入院後に死亡した患者数や治療を受けた患者の特徴を考慮し、英国の平均値と比較したデータ	NHSがHES、死亡統計、各医療機関のデータに基づいて調査を実施	・IMD	抜粋送付	不可	－	四半期毎
39	Uncurated Low Latency Hospital Data Sets for Admitted Patient Care, Outpatient Care, Critical Care & Emergency Care 入院患者ケア、外来患者ケア、クリティカルケア、救急ケアの未キュレーション・低遅延データセット	入院患者ケア、外来患者ケア、クリティカルケア、救急ケアのデータセットを、ニーズに応じて利用できるよう未キュレーションの状態・低遅延で提供するもの	医療機関が定期的にCDSファイルを送信しSUSで処理の上データセット構築	・ICD9, ICD10 ・OPCS-4 ・SNOMED-CT ・TFCs	SDE・ 抜粋送付	可	仮名	月次

- ✓ DARSで利用可能なDBの主なデータ収集経路は、「NHSによる直接調査」、「SPINE経由」「Secondary Uses Service 経由」「Strategic Data Collection Service 経由」「GPES 経由」「全国疾病登録サービス経由」などがあり、NHSが研究者のニーズに応じてDBを構築している
- ✓ DBのコーディングは「ICD9,10」「SNOMED-CT」「OPCS-4」「Read Code」等が主である
- ✓ NHSが管理するDBの多くは顕名で保管されており、連結可能である

英国では主要な臨床コードが各領域ごとに体系化されており、NHS全体で国家基準として電子カルテに導入されている

② インターフェース・システムアーキテクチャ等

NHS England 保有DBにおけるコーディング

症状	観察	検査	治療	処置	処方	医療機器
SNOMED CT						
Read / CTV3（2020年まで採用、以降SNOMED CTに移行）						
ICD-10						
			OPCS-4			
			National Codes			
					dm+d	

<各コードの概要>

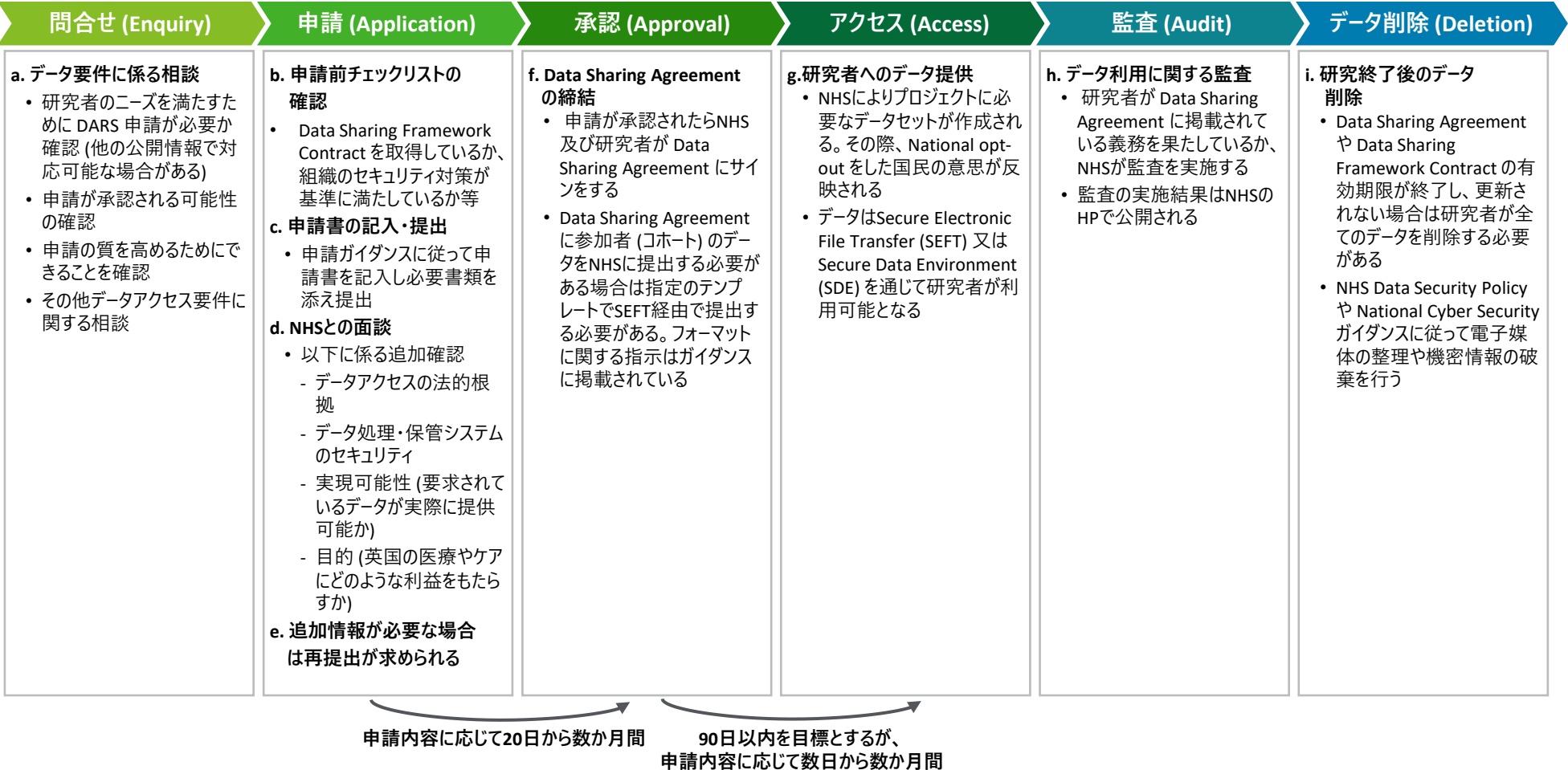
- SNOMED CT：診断や処置、症状、処方などを分類する国際的なコード。患者ケアに特化しており、すべてのNHS医療機関で電子カルテシステム内での使用が義務付けられている
- READ / CTV3：2020年頃まで英国で使用されていたコードだが、現在は非推奨となっておりSNOMED CTへ移行している
- ICD-10：WHOの国際疾病分類で診断情報をコーディングしている。2026年以降にICD-11導入予定
- OPCS-4：英国で治療や処置に特化したコード体系で、他の分類システムと補完的に使用される。3年毎に更新され、最新バージョンは2023年に実装
- National Codes：患者が受けた専門医療サービスを記録するために使用されるNHS特有のコード
- dm+d：NHSが提供する医薬品や医療機器に関するコードで、SNOMED CTの形式に基づいて構築されている

DARSの利用手順は段階的に整備されており、技術的なセキュリティ対策のみならず、データ共有契約やガイドライン等に基づいた安全なデータ提供が担保されている

② インターフェース・システムアーキテクチャ等

DARSの利用手順

■ NHS は 審査段階で、データを利用する研究組織が適切な法的根拠に基づいて研究が行われるか、又データを安全に扱うための保護措置が整備されているか確認する。更に、データ共有後は Data Sharing Agreement 等の契約に掲載されている義務を遵守しているか、監査を実施している



NHSのデータ収集・改善サービスは、データ提出対象となる医療機関や組織と連携し、健康や社会福祉に係るデータの収集や品質改善に取り組んでいる

② インターフェース・システムアーキテクチャ等

データ収集・改善サービス (Data Collection and Improvement Service)

■ データ収集・改善サービスは2023年に構築され、健康・社会福祉に係るデータの収集や品質改善に取り組むことで、NHSが提供するサービスの改善に貢献している。データ収集、データカバレッジ、データ改善、マネジメントの4つのチームから構成されている

NHS England : Data Collection and Improvement Services			
Data Collection Team	Data Coverage Team	Data Improvement Team	Management Information Team
<ul style="list-style-type: none">医療や社会福祉サービスを提供する医療機関や組織からのデータ収集、提出管理を行うチーム100以上のデータセットのデータ収集に関する運用管理を一元化	<ul style="list-style-type: none">データの提供者と協力して、提出状況や提出内容を監視しながら、データのカバー率向上に取り組むチーム未提出の医療機関や組織が抱える課題を確認しながら、提出に向けた課題解決に取り組む	<ul style="list-style-type: none">データの提供者と信頼関係を構築し長期的に連携しながら、提出データの有効性の改善に取り組むチーム地域別のデータマネジャーが医療機関や組織とコミュニケーションを取りながらデータ品質の改善サポートを行う	<ul style="list-style-type: none">Data Collection and Improvement Service が提供するサービスの運用管理やサービスの向上に取り組むチーム

< 取組の例 >

CDSファイルの提出支援	<ul style="list-style-type: none">急性期トラストに対してデータ収集・提出に関するサポートを実施新規のデータ提出項目 (例:退院予定日など) の提出に対し 1:1 のサポートを実施
救急ケアデータの提出支援	<ul style="list-style-type: none">各地域に専任のデータリアゾンマネジャーを配置して、定期的なサポートを提供無効な提出データに対し内訳レポートを作成して、提出者と協力してデータ改善を支援
社会ケアデータの提出支援	<ul style="list-style-type: none">データ提出の対象となる組織の登録や初回のデータ提出までサポート特定の政策に係るデータの品質改善に取り組む
全国監査・レポートニング	<ul style="list-style-type: none">全国のデータ提出状況を監視し、調査レポート作成電子カルテを新しく導入又は移行した医療機関・組織が中断・遅延なくデータの提出が行えるようにサポート

NHSデータの二次利用は原則英国国内に限定されるが、ICOの十分制認定を受けた国であり、DARSの申請要件を満たす機関であれば、海外利用が認められる場合がある

③ 多国間利用

NHSデータの多国間利用

■ NHS DARS 申請にあたっての海外利用に関する要件

- ・ 利用申請時に、データ利用者やデータ保存方法を明確にする必要がある
- ・ NHSは、特別な理由がない限りデータの共有・処理を英国国内に留める方針を取っている。そのため、英国外でのデータ利用は、適切な根拠を示す必要がある
- ・ データ利用者がEU圏内またはICO (Information Commissioners Office) が英国と同等のデータ保護があると認めた十分制認定(Adequacy Decision)を受けた国であることを確認する必要がある

■ 一般要件

- ・ セキュリティ基準の遵守を証明 (NHS DSPT (Data Security and Protection Toolkit)、ISO7001認証、SLSP (System Security Level Policy) のいずれか)
- ・ ICO で Data Protection Act の登録を実施
- ・ 関与組織・データフローに関する証明を提出
- ・ 研究プロトコルや倫理審査に関する証明を提出
- ・ データの利用目的、想定のアウトカム・利益、法的根拠を明示
- ・ 商業利用・外部資金提供がある場合、成果物の用途と公的利益への貢献を明示

Data Sharing Agreement の締結・データ提供

ICO 十分制認定 (Adequacy Regulations) 対象国・地域

- ✓ 欧州経済領域 (EEA) 諸国
- ✓ EUまたはEEA諸国の機関、団体、事務所など
- ✓ ジブラルタル、大韓民国、アンドラ、アルゼンチン、フェロー諸島、ガンジー島、マン島、イスラエル、ジャージー島、ニュージーランド、スイス、ウルグアイ

< 部分的な十分制認定を受けている国 >

- ✓ 米国：UK Extension to EU-US Data Privacy Framework
対象データのみ
- ✓ 日本：民間企業向け個人情報保護法対象データのみ
- ✓ カナダ：PIPEDA (個人情報保護法) 対象データのみ

(参考)

欧州経済領域(EEA)諸国

EU加盟国	オーストリア、ベルギー、ブルガリア、クロアチア、キプロス、チェコ共和国、デンマーク、エストニア、フィンランド、フランス、ドイツ、ギリシャ、ハンガリー、アイルランド、イタリア、ラトビア、リトアニア、ルクセンブルク、マルタ、オランダ、ポーランド、ポルトガル、ルーマニア、スロバキア、スロベニア、スペイン、スウェーデン
EFTA加盟国	アイスランド、ノルウェー、リヒテンシュタイン

データ活用基盤調査

- 米国
- 英国
- 仏国
- フィンランド

HDHは2019年からフランスの公衆衛生法典のもとで運用され、フランス全人口をカバーする健康データの二次利用を通して市民の健康改善に貢献する者をサポートする

情報連携基盤HDHの概要

① データ利活用基盤・データベースの概要

名称	Health Data Hub
運営主体	56組織で構成されている公益法人
運用開始年	2019年
関連法	公衆衛生法典(CSP)
設置目的	市民の健康改善に貢献する者をサポート <ul style="list-style-type: none">プロジェクトリーダーのサポートツールを提供するための技術プラットフォームの構築と運用研究の可能性が高いデータの収集と形成
利用者	HDHの審査(CESREES)およびCNIL(データ保護機関)の承認を受けアクセス権が付与された研究者
データソース	<ul style="list-style-type: none">SNDSボルドー大学病院AP-HP(フランス国内に38の病院を持つ大学病院組織)ロシュ社（民間製薬企業）Unicancer(がんの治療・研究等に関わる医療機関のネットワーク)CRCDC-OC(オクシタニー地域がん検診コーディネーションセンター)フランス公衆衛生省保健総局国民健康保険基金
規模	フランス全人口(約6,804万人,2023年) ※データカタログに応じて規模は異なる

データ連結手法	<ul style="list-style-type: none">直接ペアリング：NIR(登録番号または社会保障番号)で連結間接ペアリング
標準化	OMOP-CDM
利用実績 (年間) ※複数年に跨るプロジェクトもあり、年単位でカウント	1,565件(2021年) 1,429件(2022年) 1,337件(2023年)
収入源	欧州委員会 (フランス国内)公共活動変革基金（FTAP）、国民健康保険支出目標 (ONDAM)

利用実績内訳(全実績(約8,500件)分を集計)

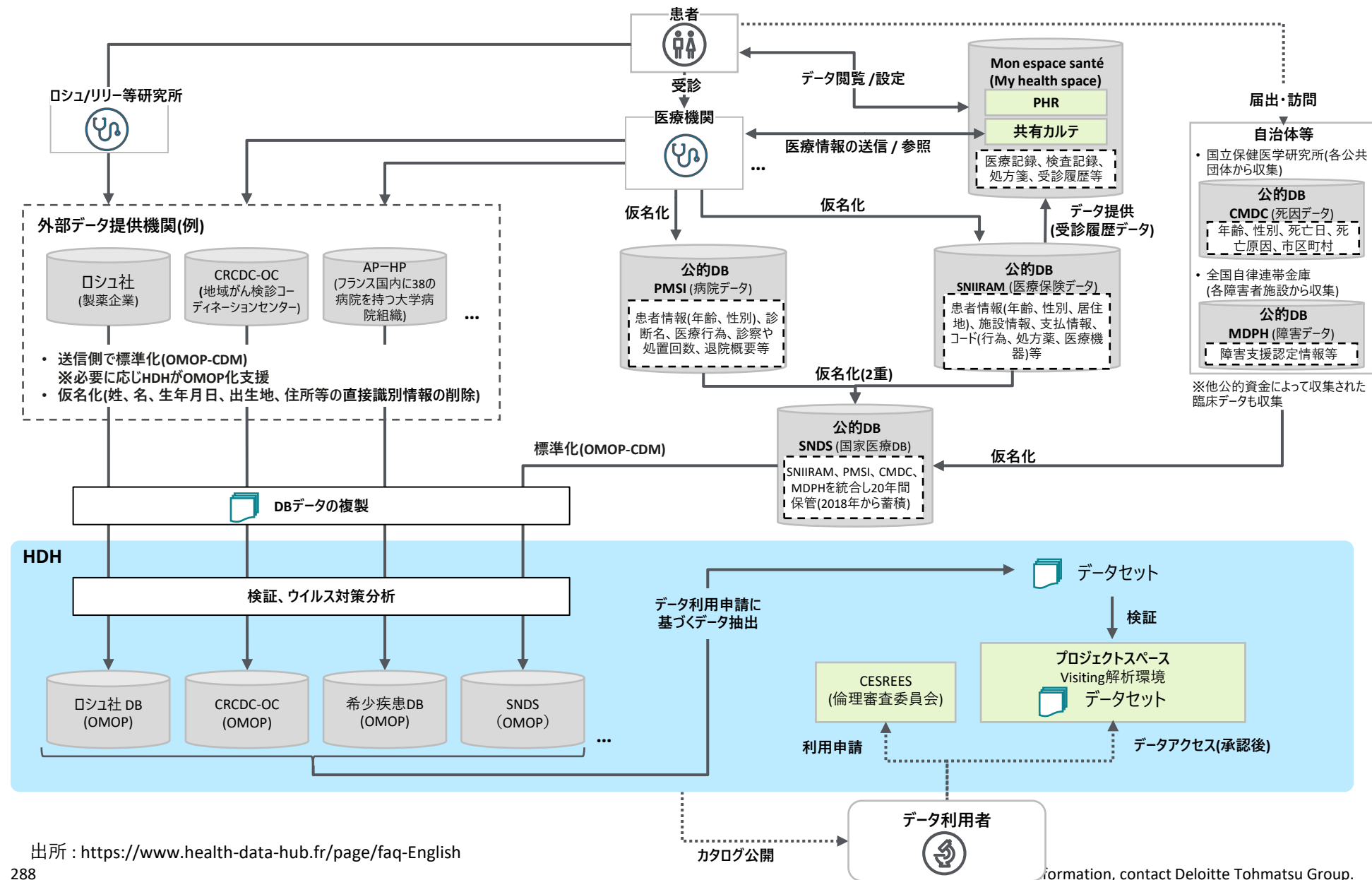
利用者別		目的別	
種類	件数	利用目的	件数
公的医療機関	4,204	患者ケア管理	2,899
製薬会社	1,282	予防と治療	2,153
学校・研究機関	1,063	病気の理解	1,291
私立医療機関	617	診断・医学研究	1,181
その他企業	534	患者の安全	740
政府機関	34	病気の要因分析	587
患者協会	17	公衆衛生政策	449
医療機器メーカー	16	医療施設の運営	448
その他	1,129	その他	4,149

出所: <https://www.health-data-hub.fr/>、<https://www.health-data-hub.fr/projects>

※累計実績数は8,593件(2018～2024年6月時点)。研究範囲に応じて目的が複数あるため、重複カウントが含まれる

システム全体像・フランスの一次利用・二次利用の構造

① データ利活用基盤・データベースの概要



No.	データセット名	提供者	提供者属性	分野	収集期間	更新頻度	データ規模	収集範囲	NIR(登録番号または社会保障番号)での連結
1	アルツハイマー病および関連疾患の決定因子と進展 (MEMENTO)	ボルドー大学病院	大学病院	神経学、老年医学、その他	2011年-2019年	収集完了のため更新なし	2,323人	フランス全土	NIRが含まれていないが、SNDSと間接的なペアリングが可能である
2	国立希少疾患データバンク (BNDMR)	AP-HP	政府	希少疾患	2004年-	更新中であるが記載なし	80万人 (2021年6月まで)	政府に認定された希少疾患管理ネットワークに受診した者	NIRが含まれていないが、SNDSと間接的なペアリングが確立されている
3	小細胞肺癌腫瘍学コホートの一時的な使用承認(ATUC CBPC)	ロシュ社	民間企業	呼吸器学、腫瘍学	2019年5月6日-2020年3月20日	収集完了のため更新なし	1,402人	フランス全土	NIRが含まれていないが、SNDSと間接的なペアリングが可能である
4	疫学・医療・経済戦略 転移性乳がん(ESME CSM)	Unicancer	がんセンターネットワーク	腫瘍学	2008年-	年次	1.9万人 (2015年まで)	フランスのがんセンターで治療を受けた患者	NIRが含まれていないが、SNDSと間接的なペアリングが可能である
5	e-SIS - 2004年から2019年までのGardおよびLozèreにおける組織的乳がんスクリーニング(DOCS)のデータベース	CRCDC-OC (オクシタニー地域がん検診コーディネーションセンター)	検診団体	腫瘍学	2004年-2019年	収集完了のため更新なし	15万人	GardおよびLozèreのみ	NIRが含まれており、NIRでのペアリングが可能である
6	救急監視調整組織(OSCOUR)	フランス公衆衛生省	政府	救急医学	2010年-	年次	1億3,100万人 (2018年まで)	マルティニーク島を除くすべてのフランスの県	NIRを使用し、DBの相互ペアリングが可能だが、SNDSと連結することを想定されていない
7	攻撃や例外的な健康状態 (Covid-19)の被害者を監視するための情報システム(SI-VIC)	保健総局	政府	感染症	記載なし	記載なし	記載なし	記載なし	記載なし
8	病院で Covid-19 と診断された患者のための国民健康データシステムからのデータ(SNDS Fast Track)	国民健康保険基金	政府	感染症	記載なし	記載なし	記載なし	記載なし	記載なし

HDHデータカタログ(SNDSのデータ種類)

① データ利活用基盤・データベースの概要

DB名称	管轄 (組織名 / 組織概要)	概要	設立年	データの種類	データ量
SNDS	公的機関	CNAM(経済財務省、連帯保健省の共同監督下にある国家公共機関)	2017年	<ul style="list-style-type: none"> • SNIIRAM • PMSI • CMDC • MDPH • 補完的なデータ(公的資金によって収集された臨床データ等) 	<ul style="list-style-type: none"> • 12億件のレセプトデータ(年間) • 1,100万件の入院データ(2018年時点)
SNIIRAM		健康保険制度の管理に責任を負う機関による任務の遂行および公衆衛生法の第L.1461-1条に記載の国民健康データシステム	2002年	<ul style="list-style-type: none"> • 患者情報(年齢、性別、居住地、保険情報、長期疾病(ALD)、職業病の有無、死亡日) • 医療機関情報(施設情報、専門分野、診療形態) • 医療費情報(受診日、医療費、処方内容、処置等の医療行為、医療機器、生物学的サンプル等のコード) 	
PMSI		ATIH(連帯保健省の大臣の監督下に置かれた公的機関)	2009年	<ul style="list-style-type: none"> • 患者情報(年齢、性別、地理コード) • 施設の識別子 • 入院情報(診断名、実施された医療処置、入院期間、入退院の形態、死亡情報等) • 該当入院情報に付随する外来情報(実施された診察および外来処置の数、薬剤および植込み型医療機器(人工器官等)の消費量等) 	2,500万件 (1968年以降、年間約50万人の死因データ)
CMDC		Inserm(連帯保健省と高等教育研究イノベーションの共同監督下にある公的な科学技術施設)	2000年	<ul style="list-style-type: none"> • 患者情報(年齢、性別、市区町村) • 死亡の医学的原因 • 死亡場所等 	
MDPH		CNSA(全国自律連帯金庫)	2005年	<ul style="list-style-type: none"> • 障害支援認定情報 • 補償給付情報等 ※ただし、他のDBと異なりSNDSとの接続試験中(2022年) 	

出所: <https://www.health-data-hub.fr/>、<https://www.snds.gouv.fr/SNDS/Accueil>

HDHは4年間の初期資金として公的資金から合計7600万ユーロ(約121.6億円)が提供された。また、運営費用として、2022年に一年間に約63.52億円がかかると推測された

HDHの構築費用及び運営費用

① データ利活用基盤・データベースの概要

■構築費用

✓ 4年間の初期資金として公共活動変革基金 (FTAP)から3,600万ユーロ(約57.6億円)、国民健康保険支出目標 (ONDAM)から 4,000 万ユーロ(約 64.0億円)、合計約7,600万ユーロ(約121.6億円)が提供された

■運営費用(2018年アニュアルレポートの推測による)

※1ユーロ=160.0円,2025.2.17で円換算

費用項目※ (億円)		2018 第4四半期	2019	2020	2021	2022
導入された地域ハブの数		0.00	0.00	3.20	6.40	8.00
1. 技術プラットフォーム		0.00	4.48	8.00	9.92	10.24
2. 中央ハブ(全体の管理、地域ハブの支援)		0.48	8.00	5.92	5.44	5.44
	2.1 技術チーム、セキュリティおよびデータ管理	0.16	4.64	3.20	2.72	2.72
	2.2 法律専門チーム	0.00	0.32	0.32	0.32	0.32
	2.3 医療専門チーム	0.00	0.16	0.48	0.48	0.48
	2.4 プログラム管理、評価、改善チーム	0.32	2.56	1.60	1.60	1.60
	2.5 施設料金	0.00	0.16	0.32	0.32	0.32
3.地域ハブ(地域のデータ所有者の申請、データセット準備と提供、データ活用支援)		0.00	0.00	3.20	7.36	10.40
	3.1 技術チーム、セキュリティおよびデータ管理	0.00	0.00	1.76	4.00	5.60
	3.2 法律専門チーム	0.00	0.00	0.32	0.64	0.80
	3.3 医療専門チーム	0.00	0.00	0.32	0.64	0.96
	3.4 プログラム管理、評価、改善チーム	0.00	0.00	0.80	1.76	2.40
	3.5 施設料金	0.00	0.00	0.16	0.32	0.48
4. プロジェクト支援		0.00	3.84	10.56	18.24	18.24
	4.1 中央ハブ	0.00	3.84	7.68	8.64	8.64
	4.2 地域ハブ	0.00	0.00	2.88	9.60	9.60
5. データ保有者への助成金		0.00	12.80	16.00	19.20	19.20
合計		0.48	29.12	43.68	60.16	63.52

HDHは2022年に1,783万ユーロ(約28.5億円)の支出費用が計上された。職員数の予測分布において、部門が総局、技術、戦略、支援に分けられている

HDHの支出費用及び職員数予測分布

① データ利活用基盤・データベースの概要
※1ユーロ=160.0円,2025.2.17で円換算

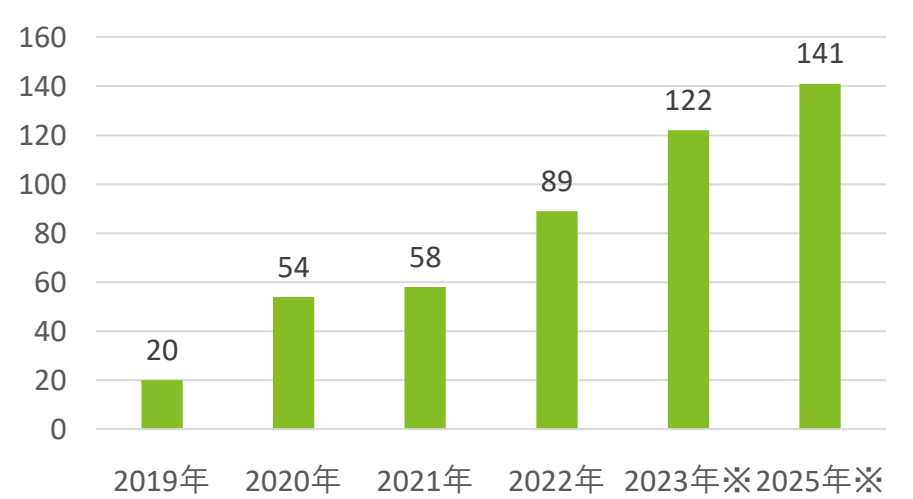
■ HDHの支出費用(2019年～2022年)

費用(万ユーロ)	2019	2020	2021	2022
人件費	12.7	318.9	518.7	667.4
その他運営費 (クラウドインフラ ストラクチャ等)	146.4	800.1	917.2	1,115.5
合計	159.1	1,118.9	1,435.9	1,783

■ 2023年末～2025年末のHDHの職員数の予測分布

部門の種類	部門	2023	2025
総局	管理および機関関係	3	4
	技術およびプラットフォーム開発	20	23
技術	パートナーシップ	13	18
	データ	24	27
	データアクセス	7	8
	プロジェクトおよびユーザーサービス	8	10
	科学	13	15
戦略	医療	3	3
	市民	3	3
	秘書、HR、財務	9	11
支援	法務	10	11
	ITセキュリティ	3	4
	コミュニケーション	6	6
	合計	122	143

■ HDH職員数推移(2019年～2025年) ※予測値を表す

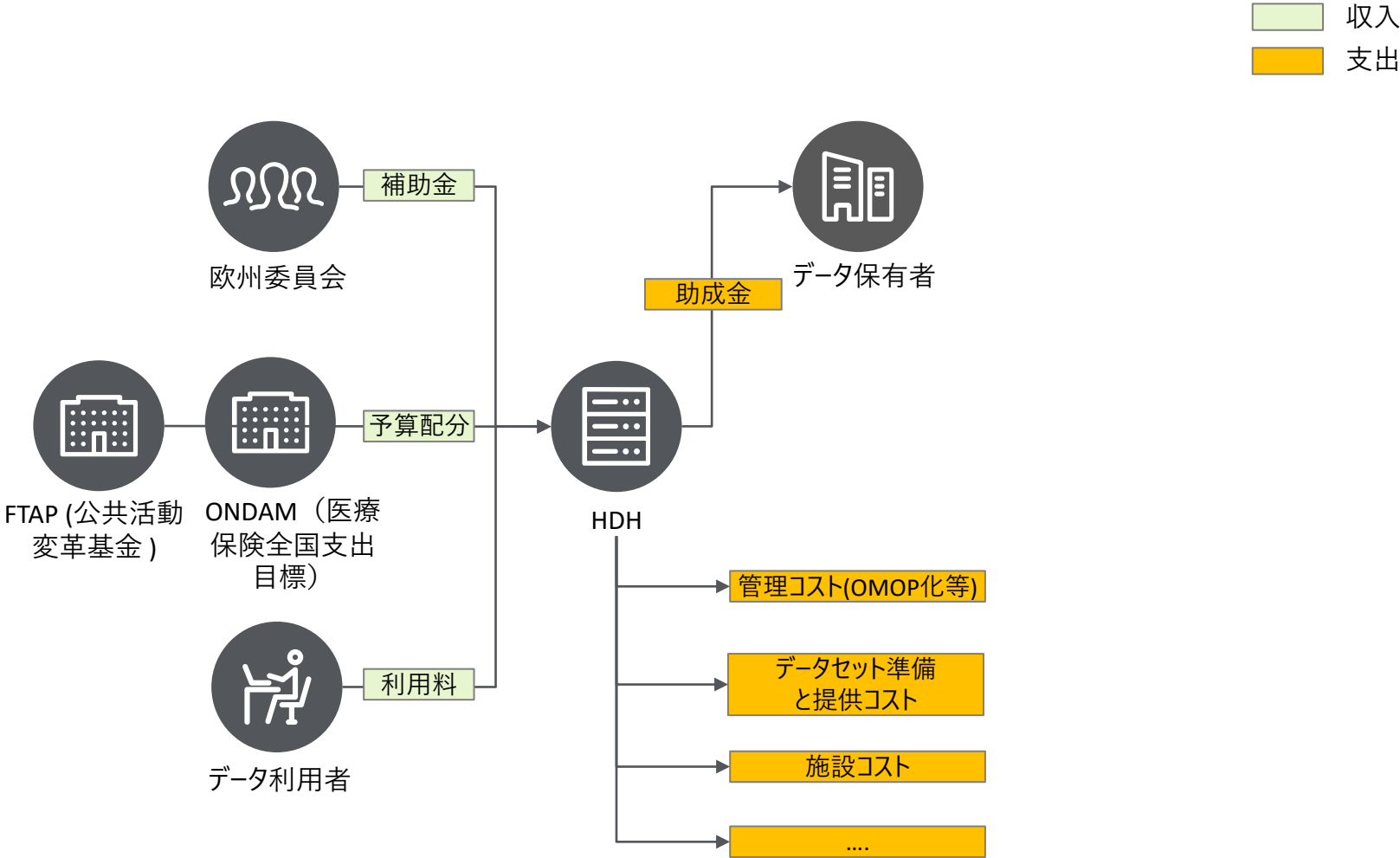


出所： <https://tehdas.eu/app/uploads/2023/09/tehdas-sustainability-plan-for-ehds.pdf>

HDHは、欧州委員会、フランス政府及び利用者からの利用料を財源とし、データ管理組織への助成金やOMOP化、データセット準備と提供及び施設にコストを要している

HDHのビジネスモデル

① データ利活用基盤・データベースの概要



HDHに掲載されたフランスのコンサル会社が実施した調査によると、3種類の経済効果の算出方法が提示されている

① データ利活用基盤・データベースの概要

HDHにおける健康データの二次利用の経済効果の算出

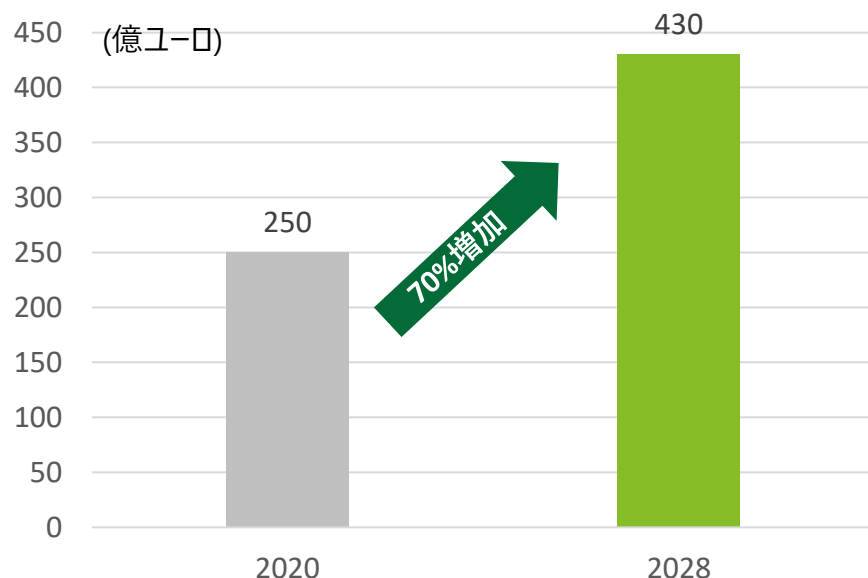
※1ユーロ=160.0円, 1ドル=152.3円, 1ポンド=192.1円, 2025.2.17で円換算

Veltys(フランスコンサルティング会社)が実施し、2022年7月にHealth Data Hubに提出した調査で、3種類の算出方法が提示されている

① 欧州全体のヘルスデータの二次利用の年間経済効果に基づいた推測

EHDSに関する評議会の資料によると、2028年に欧州全体のヘルスデータの二次利用の年間経済効果が**430億ユーロ(約6.9兆円)**である。2019年の欧州全体の名目GDPにおいて、フランスが17%を占めているため、**年間約73億ユーロ(約1.2兆円)**である

—— 欧州全体のヘルスデータの二次利用の年間経済効果推測 ——



② 英国に関するE&Yの推計に基づいた推測

NHS(国民保健サービス)にとって50億ポンド(約9,605億円)、患者にとって46億ポンド(約8,837億円)の経済的利益が生じるとされている。この数値に基づき、フランスでは、政府にとって年間46億ユーロ(約8,837億円)、患者にとって42億ユーロ(約8,069億円)の利益が見込まれる

③ IQVIAによる製薬業界におけるリアルワールドデータの活用による利益

IQVIAによると2021年から2023年の間に、リアルワールドデータ（RWD）の活用による利益は、製薬業界全体で40億ドル(約6,092億円)に達すると予想されている

- ・ イノベーションフェーズ：17.5億ドル（うち5億ドルはコスト削減）
- ・ 承認フェーズ：6億ドル（うち半分はコスト削減）
- ・ アクセスと保険適用フェーズ：3億ドル
- ・ 診断および治療フェーズ：7.5億ドル
- ・ 商業化フェーズ：6億ポンドの利益（うち1億ドルはコスト削減）

NHS のヘルスデータは、ビッグデータ、AI、個別化医療の応用を通じて、NHSに50億ポンド(約9,605億円)、患者に46億ポンド(約8,837億円)の経済的利益をもたらす可能性がある

NHSのヘルスデータの二次利用による経済効果(収益ベースアプローチ)

① データ利活用基盤・データベースの概要

※1ポンド=192.1円,2025.2.17で円換算

EYが2019年に発表された「医療データの価値を実現する：未来のためのフレームワーク」によると、NHSのヘルスデータの二次利用による経済効果は、「収益ベースアプローチ」及び「マーケットベースアプローチ」によって測定する

■ 収益ベースアプローチ

- ・ 推定方法
データセットの経済的利益に基づいて経済価値を推定する。経済的価値の向上と商業的機会の2種類に分けられる
 - ① 経済的価値の向上
患者への利益には、より迅速かつ正確な診断、改善された治療結果、新製品の開発、改善されたケアパス、および運用効率の向上が含まれる。英国経済への利益には、社会経済的利益や付加価値の増加が含まれる
 - ② 商業的機会
データの活用によりビジネスを展開することで得られる追加の利益やライセンス収入を指す。例えば、ライフサイエンス企業がデータを活用して薬の開発期間の短縮や新しい治療法の発売につながる可能性がある
- ・ 計算方法
ビッグデータ、AI、個別化医療の応用を通じて、NHS の運営コストの大幅な削減、患者の治療成果の向上、そしてより広範な経済的利益をもたらす可能性があると考えられるため、それぞれの短期的、中期的、長期的価値を推定する

	ビッグデータ	AI	個別化医療の応用
短期的価値 (3年未満)	NHSは必要な技術を調達する	データセットが作成され、ベストプラクティスに関する有益な洞察	新薬開発プロセスの開始
中期的価値 (3年～9年)	ベストプラクティスと学習がNHSを通じて広まる	短期的なベストプラクティスをNHS全体で実施	新薬の臨床試験
長期的価値 (10年以上)	潜在的なメリットの実現	有益な洞察の継続的な発見	新薬の登場
NHS年間経済効果	27億ポンド (5,186.7億円)	17億ポンド (3,265.7億円)	6億ポンド (1,152.6億円)
患者年間経済効果	－	25億ポンド (4,802.5億円)	21億ポンド (4,034.1億円)

NHS のヘルスデータの二次利用による経済効果は、マーケット取引からそれぞれのデータセットの価値を評価倍率とし、データ数にかけることでデータベースの価値を推定できる

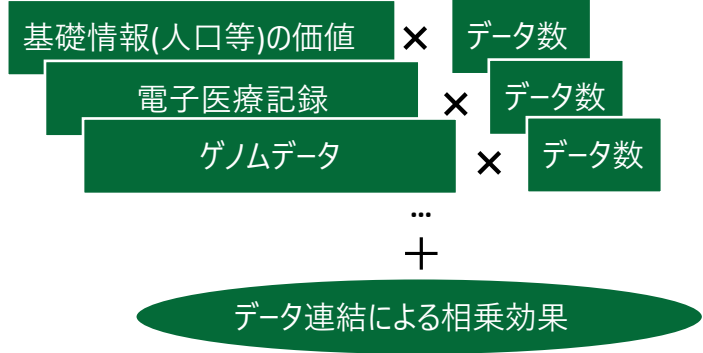
(参考)NHSのヘルスデータの二次利用による経済効果(マーケットベースアプローチ) ① データ利活用基盤・データベースの概要

EYが2019年に発表された「医療データの価値を実現する：未来のためのフレームワーク」によると、NHSのヘルスデータの二次利用による経済効果は、「マーケットベースアプローチ」と「収益ベースアプローチ」によって測定する

■ マーケットベースアプローチ

- 推定方法
データセットの価値をそのプロファイル（サイズ、内容、性質、利用可能性、成熟度、独自性及びデータの品質）と、比較可能なデータ資産の市場価格に基づいて推定する
- 分析方法
Step1
データセットを整理する
Step2
カテゴリーの定性フレームワークを使用し、データセットを評価する
Step3
類似マーケット取引分析し、評価倍率を選択する
Step4
データセットに評価倍率を適用する
Step5
データセットの価値を推定し、限界を特定する

• 計算方法



※下記のデータセットの価値や属性も考慮に入れて計算する必要がある

項目	内容
データ属性	データ種類、データの可用性、排他性または希少性、粒度、データ保有者の属性
データ品質	生データ品質、データ整理状況、追跡可否、洞察や分析状況
データ収集の複雑さ	データの生成元、データのアクセス可能性、データ収集方法
データ利用可能性	使用と潜在的な影響、独占性、使用上の制限、他の企業や競合他社による使用

NHSのエピソード記録が6.9億ポンド(1,325.5億円)～142.37億ポンド(2.27兆円)、ゲノムデータが2,080万ポンド(40億円)～2億240万ポンド(388億円)の経済効果があると推定される

① データ利活用基盤・データベースの概要

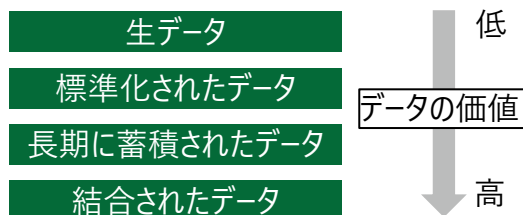
(参考)NHSのヘルスデータの二次利用による経済効果計算例(マーケットベースアプローチ)

※1ポンド=192.1円,2025.2.17で円換算

Step1 NHSのデータセットを整理する

- ✓ PCR：5,500万件
- ✓ エピソード記録/電子医療記録(EMR)：約2,300万件
- ✓ ゲノムデータ：約10万件
- ✓ その他：処方箋データ、運用及びプロセスデータ、医薬品の支出、財務データ、社会福祉情報

Step2 カテゴリーの定性フレームワークを使用し、データセットを評価する



Step3 類似マーケット取引分析し、評価倍率を選択する

患者データを含む資産を有する企業を調査する

公開されている取引データベース等を詳細に調査する

各種類のデータの1データあたりの価値を算出する

※注意：①医療データを広範囲に取得するケースが多いため、データの価値を独立して切り離して評価することが課題である。②各データセットや取引ごとにデータの性質、品質、複雑さ、適用方法が異なるため、注意深く評価する必要がある

Step4 NHSデータセットに評価倍率を適用する

データ種類別1データあたりの価値

上場企業の時価総額

電子健康記録(EHR)

£21-£192

エピソード記録

£30-£619

ゲノミクス

£2,354-£4,773

非上場企業および製薬企業の提携の資産後評価

ゲノミクス

£208-£2,024

がん関連データ

£40-£5,549

一般

£8-£486

薬剤開発

£0.20-£35

人口データ

£3-£21

Step5 NHSデータセットの価値を推定し、計算

方法の限界を特定する

エピソード記録

£30-£619 × 2,300万件 = 6.9億ポンド～142.37億ポンド

ゲノムデータ

£208-£2,024 × 10万件 = 2,080万ポンド～2億240万ポンド

本計算方法の限界

- 結合されたデータは単独なデータより価値が上がるので特別に計算する必要がある
- データの可用性や属性により価値が変動する
- ベンチマークとなる評価倍率変動することがある
- データセットサイズの増加に伴う収益が減少することがある

出所： <https://docslib.org/doc/1592334/realising-the-value-of-health-care-data-a-framework-for-the-future-realising-the-value-of-health-care-data-a-framework-for-the-future>

2015年～2018年のがん関連データを保有する民間企業の投資・買収取引を12件を選択し、1データあたり金額が£40(7,684円)～£5,549(106.6万円)を計算した

(参考)(Step3)類似マーケット取引分析

① データ利活用基盤・データベースの概要

※1ポンド=192.1円,2025.2.17で円換算

■ がん関連データを保有する民間企業の投資・買収取引一覧

	企業名	買収・投資元	買収・投資元の属性	取引の種類	年	取引額 (£)	取引比率 (%)	取引後評価額 / 買収価格 (£)	データ数	1データあたりの評価額 (£)
1	Flatiron	Roche	製薬企業	買収	2018	£1,498,834,000	100%	£1,498,834,000	2,200,000	£681
2	Cota Health	IQVIA	ヘルスケアデータ企業	投資	2018	£31,554,400	18%	£172,831,337	n/a	n/a
3	Tempus	Baillie Gifford, T. Rowe Price, Revolution Growth, NEA	投資会社	投資	2018	£87,563,460	6%	£1,577,720,000	220,000	£7,171
4	Precision Health AI	SymphonyAI Group	AI・データ分析企業	投資	2017	£15,777,200	21%	£75,841,000	n/a	£76
5	Paige.AI	Breyer Capital	投資会社	投資	2018	£19,721,500	21%	£93,251,141	25,000,000	£4
6	Freenome	Andreessen Horowitz	投資会社	投資	2017	£56,797,920	32%	£175,868,448	n/a	n/a
7	Foundation Medicine	Merck	製薬企業	買収	2015	£1,893,264,000	45%	£4,180,958,000	200,000	£20,905
8	Cambridge Cancer Genomics Inc.	AME Cloud Ventures	投資会社	投資	2018	£3,549,870	15%	£23,665,800	n/a	n/a
9	Mendel Health Inc.	Bootstrap Labs	投資会社	投資	2017	£1,577,720	11%	£13,954,933	n/a	n/a
10	Oncora Medical Inc.	非公開投資家	不明	投資	2017	£1,490,945	9%	£16,700,166	n/a	n/a
11	PathAI Inc	8VC	投資会社	投資	2017	£11,785,568	17%	£70,445,198	n/a	n/a
12	SkinVision BV	PHS Capital B.V.	投資会社	投資	2018	£5,995,336	19%	£31,775,281	1,156,359	£27

四分位数	取引額 (£)	データ数	1データあたりの評価額 (£)
第1四分位数 (25%)	£29,747,911	415,000	£40
中央値 (50%)	£84,546,071	1,078,180	£379
第3四分位数 (75%)	£506,609,836	1,939,090	£5,549



がん関連データ
1 データあたり£40～£5,549

出所 : <https://docslib.org/doc/1592334/realising-the-value-of-health-care-data-a-framework-for-the-future-realising-the-value-of-heath-care-data-a-framework-for-the-future>

HDHでは、CESREESによる審査およびCNILによる承認を経て、HDH内のプロジェクトスペースにアクセスする形式でデータを利用できる

HDHにおけるデータの利用の流れ

① データ利活用基盤・データベースの概要

- 【公共の利益】 公的機関、民間機関ともに、潜在的な商業目的で研究開発プロジェクトを進める権利を有するが、研究が公共の利益となることを、つまり市民の利益となる一般的な利益であることを倫理科学委員会 (CESREES) に示さなければならない
- 【販促目的の禁止】 公衆衛生法典L.1461-1条は、商業目的で医療専門家に製品の販売促進をするために健康データを使用することを明示的に排除し、犯罪としている。個人または集団に対する保証対象の除外や保険料の変更となるおそれのある使用についても、同様である。プロジェクト実施者によるいかなる侵害も刑事訴追に繋がる
- 【販売の禁止】 データの販売をしてはならない

- ①

 - データ利用者はHDHへのアクセスをリクエストする

*CESREES: Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé

 - ✓ 委員会の事務局はHealth Data Hub
 - ✓ 健康、疫学、遺伝学、生物統計学、人間社会科学、倫理、法律、データのアルゴリズム処理、健康分野の個人情報情報の専門家約20名が法により任命を受ける
 - ✓ 月次で委員会を開催
- ②

 - データ申請が独立した倫理科学委員会 (CESREES*)に送られ、審査を受ける
- ③

 - CNILが研究を承認
 - 研究過程での目的変更等で再度承認の必要あり
- ④

 - HDHのチームが依頼されたデータをまとめ、「プロジェクトスペース」としての研究のためのプラットフォームを作成する
- ⑤

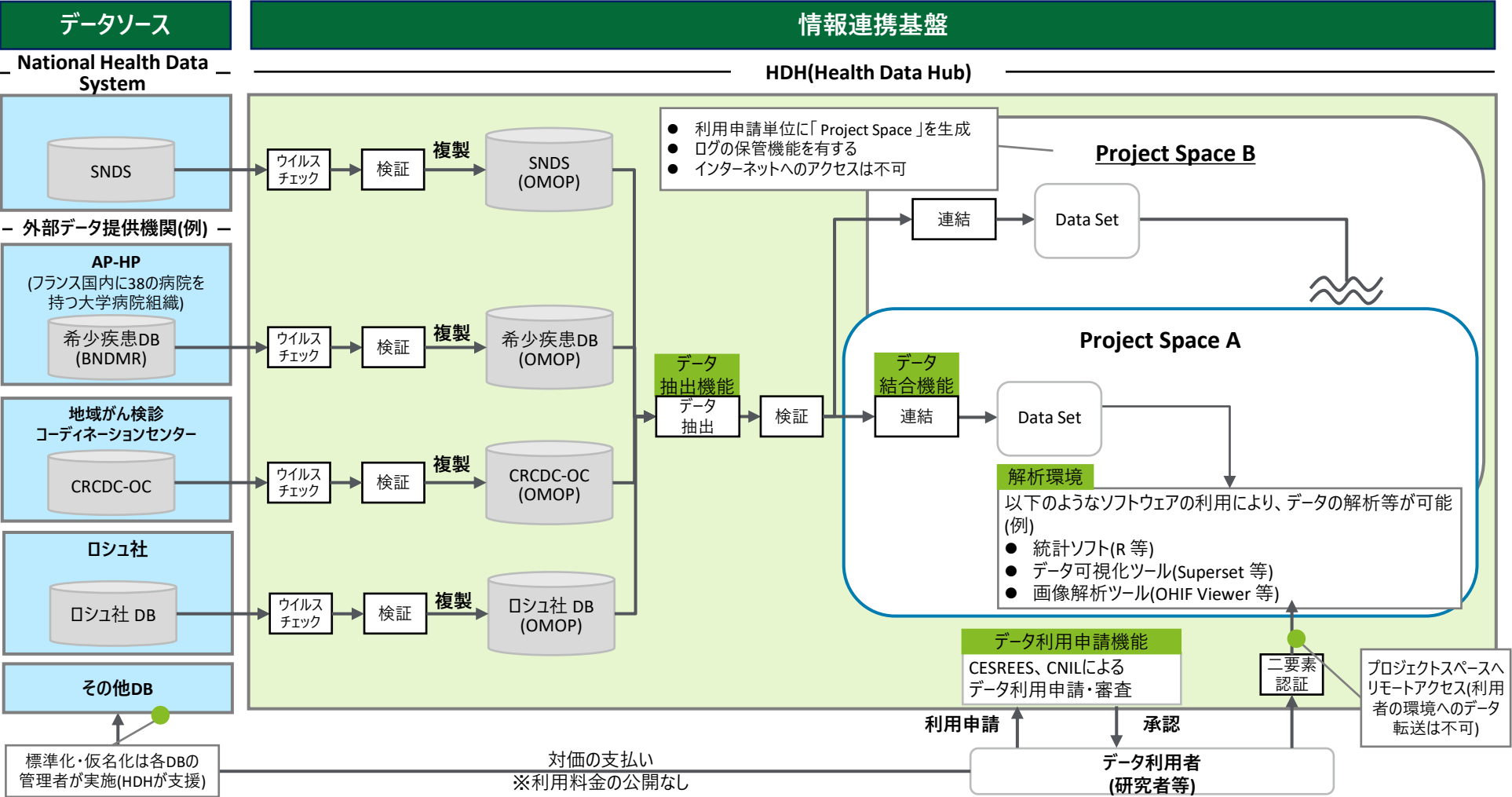
 - データ利用者は、自信の研究のために設置された「プロジェクトスペース」に遠隔でアクセスし、データを処理する
 - HDHはデータへのアクセスに対して請求をかける
- ⑥

 - 研究結果は学術・産業競争を尊重する範囲でHDHのホームページ上で公開される

HDHの情報連携基盤ではOMOP化されたデータが格納され、利用申請ごとにインターネットへアクセスできない「Project Space」が生成される

情報連携基盤の構成

② インターフェース・システムアーキテクチャ等



HDHでは、研究者が医療情報を安全に受信・閲覧・保存・解析できる環境として「Project Space」を提供している

Visiting解析環境「Project Space」

② インターフェース・システムアーキテクチャ等

概要	HDH内の「Project Space」は、権限が付与された研究者等が安全に医療情報を受信・閲覧・保存・解析できる環境。リモートアクセスが可能だが、外部環境から遮断されており、ローカルPCにデータを転送することはできない	
Visiting解析環境のメリット	<ul style="list-style-type: none">データ整理・分析・可視化のための幅広いツールを用意チームメンバー間のコラボレーションが可能SNDSのセキュリティ基準への準拠(GDPR基準への準拠は現在対応中)HDHによるユーザーのログ管理 等	
提供データ	HDHのデータカタログに基づく	
利用者	HDHの審査(CESREES)およびCNIL(データ保護機関)の承認を受けアクセス権が付与された研究者	
利用環境	プログラミング言語	<ul style="list-style-type: none">PythonR
	IDE(統合開発環境)	<ul style="list-style-type: none">JupyterLabRStudio
	ソフトウェアインフラ	<ul style="list-style-type: none">Apache Spark
	データ可視化	<ul style="list-style-type: none">Apache Superset
	データストレージ	<ul style="list-style-type: none">MySQLPostgreSQL
	画像解析	<ul style="list-style-type: none">CytomineOHIF Viewer
	データ共有・共同編集のための開発支援ツール	<ul style="list-style-type: none">GitGitLab

SNDSのデータと研究者等の元データとの連結を提供しており、連結にはCNILからの事前承認が必要である

HDHにおけるデータ連結

② インターフェース・システムアーキテクチャ等

- ✓ マッチングは、主に2つのタイプ（直接ペアリング、間接ペアリング）があり、連結にはCNILからの事前承認が必要
- ✓ 利用者は、研究者や民間企業等（ただし、民間企業の場合、データアクセスはより厳しく規制されており、CNILの定める機密性、専門性、独立性の基準に遵守しなければならない（なお、このコミットメントを行った組織の一覧は、CNILによって公開される）

① 直接ペアリング

【手法】

- 利用者(研究者)は、「NIR（社会保障番号） / 生年月日、性別」の患者情報を知らせる必要がある
- この手順では、社会保障番号-生年月日-性別情報を不可逆的に暗号化して、暗号化されたSNDS識別子に変換する
- CNAM（国民健康保険基金）はSNDS（国立医療データベース）上のSNDS識別子に関連付けられる被験者識別子（NS）（例えば、NS_000001-> NS_999999）を検索し、SNDSからデータを抽出し、申請者に配信する

※NIRがない場合、特性（名前、性別、日付、出生地）から最も近いNIRを探し、連結を行う

【例】

- 研究者は、NIRを持っている患者の追加情報を入手したい
- CNAMと協力し、研究者は、NIR、性別コード、および生年月日から、必要な手順に従い、データの要求を実行する
- CNAMは、SNDSで保管されている暗号化された識別子を回復し（NIRに明確に戻すことはできない）、被験者識別子（NS）からデータを抽出し、研究者に配信する

② 間接ペアリング

【手法】

- 識別子がなくても、定義された情報に従って患者を選択する方法であり、必要な情報を確率的に一致させる

【例】

- 研究者は、選択された10人の医師リストから1,000人の患者に関する情報を取得したいと考えている（医師1人あたり100人）
- 研究者は、自身の研究のすべての患者の共通変数のリストを提供する（例えば、年齢、性別、処方薬、治療日など）
- CNAMは必要な情報を取得し、SNDSテーブルで基準に確率的に対応できる患者を検索する
- 次に、これらの患者に対してSNDSデータの抽出が実行される

匿名データの連結

- 匿名データの照合は、薬、医療行為、診断、部門のリストなどのデータをSNDSポータルに（インポート/エクスポートツールを使用して）インポートすることで構成できる
- これらのデータとSNDSのデータの間で連結を実行し、次に、対象患者のデータを抽出する

出所: https://documentation-snds.health-data-hub.fr/fiches/appariement_donnees_externes.html#definition
<https://documentation-snds.health-data-hub.fr/introduction/03-acces-snds.html#les-acces-sur-projet>

SNDSおよびHDH OMOPで採用されているコードは以下になる

SNDSで使用されている主なコード

② インターフェース・システムアーキテクチャ等

分類	コード (略語)	正式名称	説明	利用範囲
臨床診断	ICD-10	International Statistical Classification of Diseases and Related Health Problems 10th Revision	WHOが定めた国際疾病分類として、病気やその他の健康状態を分類するため使用されている。フランスでは入院診断で使用されている	世界標準
医療行為	CCAM	Classification Commune des Actes Medicaux	医療行為を表す医療分類である	フランス
	CSARR	Catalogue Spécifique des Actes de Rééducation et Réadaptation	2012年から適用されるリハビリテーション及び機能回復の行為特有のカタログである	フランス
医薬品	ATC	Anatomical Therapeutic Chemical	WHOによって管理される医薬品分類法の一つで、5水準からなる医薬品コードが主成分の薬効別に付与される	世界標準
	CIP	Code Identifiant de Présentation	医薬品を識別する7桁または13桁のコードである	フランス
	UCD	Unités Communes de Dispensation	医薬品医療機関で使用されている医薬品を表すコード。週次で更新が行われている	フランス
生物学的行為	NABM	Nomenclature des Actes de Biologie Médicale	生物学的行為をコード化および分類するために使用される。健康保険のtable nationale de codage de biologie (全国生物学コーディング表)にまとめられている	フランス
医療機器	LPP	Liste des Produits et des Prestations	医療保険で払い戻し可能な医療機器およびサービスのリストである	フランス

OMOPで使用されている主なコード

※実態として、ICD-10の変換は医療機関の専門スタッフによって実施されている。一部の医療機関では、医師によって行われることもある

分類	コード (略語)	正式名称	説明	利用範囲
医療行為	SNOMED-CT	Systematized Nomenclature of Medicine-Clinical Terms	IHTSDO (International Health Terminology Standards Development Organization) の9か国からなる国際機関によって管理されている構造化された医療用語集であり、臨床所見や病名等が含まれる。現在30万8,000概念、77万7,000用語からなる	世界標準
医薬品	RxNorm	RxNorm	米国国立医学図書館(NLM)によって開発された医薬品を識別するコードである	米国
検査	LOINC	Logical Observation Identifiers Names and Codes	1994年に米国の非営利医学研究機関によって開発された医療検査室の観察を識別するためのコードである	世界標準

出所 : https://www.health-data-hub.fr/sites/default/files/2021-05/HDH_StarterKit_Projet_Guide_Pedagogique_EDB_SNDS.pdf; https://www.health-data-hub.fr/sites/default/files/2024-05/HDH_RA_2023%20%281%29.pdf



SNDSからの変換として、直接移行するデータもあれば、OMOPのコードマッピング表に基づく変換もある

(ご参考) SNDSからOMOPの変換例：来院

② インターフェース・システムアーキテクチャ等

ER_PRS_F 外来治療に対する保険給付			
変数名	概要	記入例	
ben_nir_psa	SNIIRAM特有の患者識別子	XX	
prs_nat_ref	給付コード	1130 担当医治療	1331 放射線診断
exe_soi_dtd	処置開始日	2013-03-04	2013-03-05
exe_soi_dtf	処置終了日	2013-03-04	2013-03-05
pse_spe_cod	処置を行う医療従事者の分類	1	6
pse_act_nat	処置を行う医療従事者の活動タイプ	0	0
etb_pre_fin	医療施設のFINESS番号	NULL	750300360

VISIT_OCCURENCE			
変数名	概要	記入例	
Visit_occurrence_id*	来院番号	1	2
Person_id*	患者識別id	1	
Visit_source_value	来院目的	1130	1331
Visit_concept_id*	来院種類id (外来、入院等)	9202	9202
Visit_start_date*	医療機関訪問開始日	2013-03-04	2013-03-05
Visit_end_date*	医療機関訪問終了日	2013-03-04	2013-03-05
Provider_id	プロバイダー ID	4	5
Care_site_id	ケアを実施する医療施設のID		1
Visit_type_concept_id	来院記録の情報源	32810	

-  データを直接に移行する
-  OMOPのコードマッピング表に基づき変換

出所 : <https://www.youtube.com/watch?v=WGxfReMbHil>

データ活用基盤調査

- 米国
- 英国
- 仏国
- フィンランド

フィンランドでは国内法に基づきFindataを設立しており、データの二次利用に関する許可や個人データの保護などを行っている

① データ利活用基盤・データベースの概要

Findata概要

当名称	Findata Social and Health Data Permit Authority
運営主体	フィンランド社会保健省 フィンランド国立保健福祉研究所(THL)
運用開始年	2020年
関連法	健康・社会データの二次利用に関する法律 (Laki sosiaali- ja terveystietojen toissijaisesta käytöstä)
設置目的	<ul style="list-style-type: none">データセキュリティと個人のデータ保護の改善社会福祉および医療データリソースの利用のスピードアップおよび合理化二次利用の許可処理での作業の重複を減らすことデータ保有者と社会福祉およびヘルスケアセクターのデータ概要の作成
利用者	データ利用許可を得た利用者

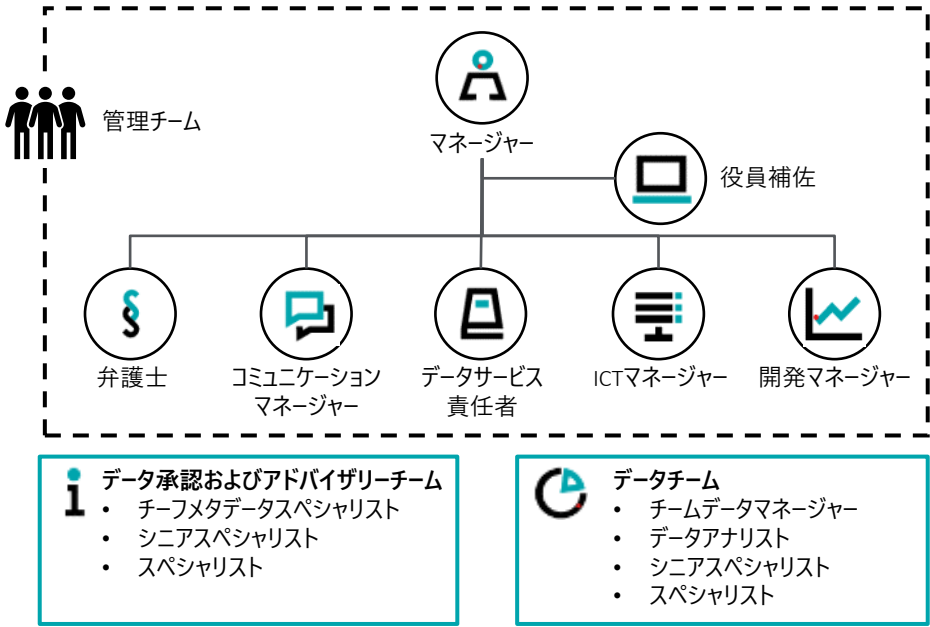
データソース	<ul style="list-style-type: none">Kantaサービスデジタルおよび人口データサービス庁フィンランド年金センターフィンランド国立保健福祉研究所(THL)フィンランド職業健康研究所フィンランド医薬品庁 (Fimea)福祉・健康監督局 (Valvira)フィンランド社会保険機構 (Kela)フィンランド統計局、地方の国家行政機関等社会福祉、医療の公的・私的サービス提供者
利用実績 (新規データ共有 契約数)	2023年:新規申請296件、許可決定351件 2022年:新規申請270件、許可決定284件 2021年:新規申請312件、許可決定262件

Findataは3つのチームからなり、審査基準は申請段階によって異なる

① データ利活用基盤・データベースの概要

組織図

- Findataは「管理チーム」、「データチーム」、「データ承認およびアドバイザリーチーム」の3つのチームで運営されている
- 二次利用のデータ利用許可は、「データ承認およびアドバイザリーチーム」で処理する
- 運営グループのメンバーは社会保健省、デジタル・人口データサービス庁、フィンランド年金センター、フィンランド医薬品庁、フィンランド社会保険協会(Kela)、フィンランド健康福祉研究所、フィンランド統計フィンランド、フィンランド労働衛生研究所、国家福祉衛生監督局バルビラ、社会福祉および医療サービスプロバイダの代表者からなっている



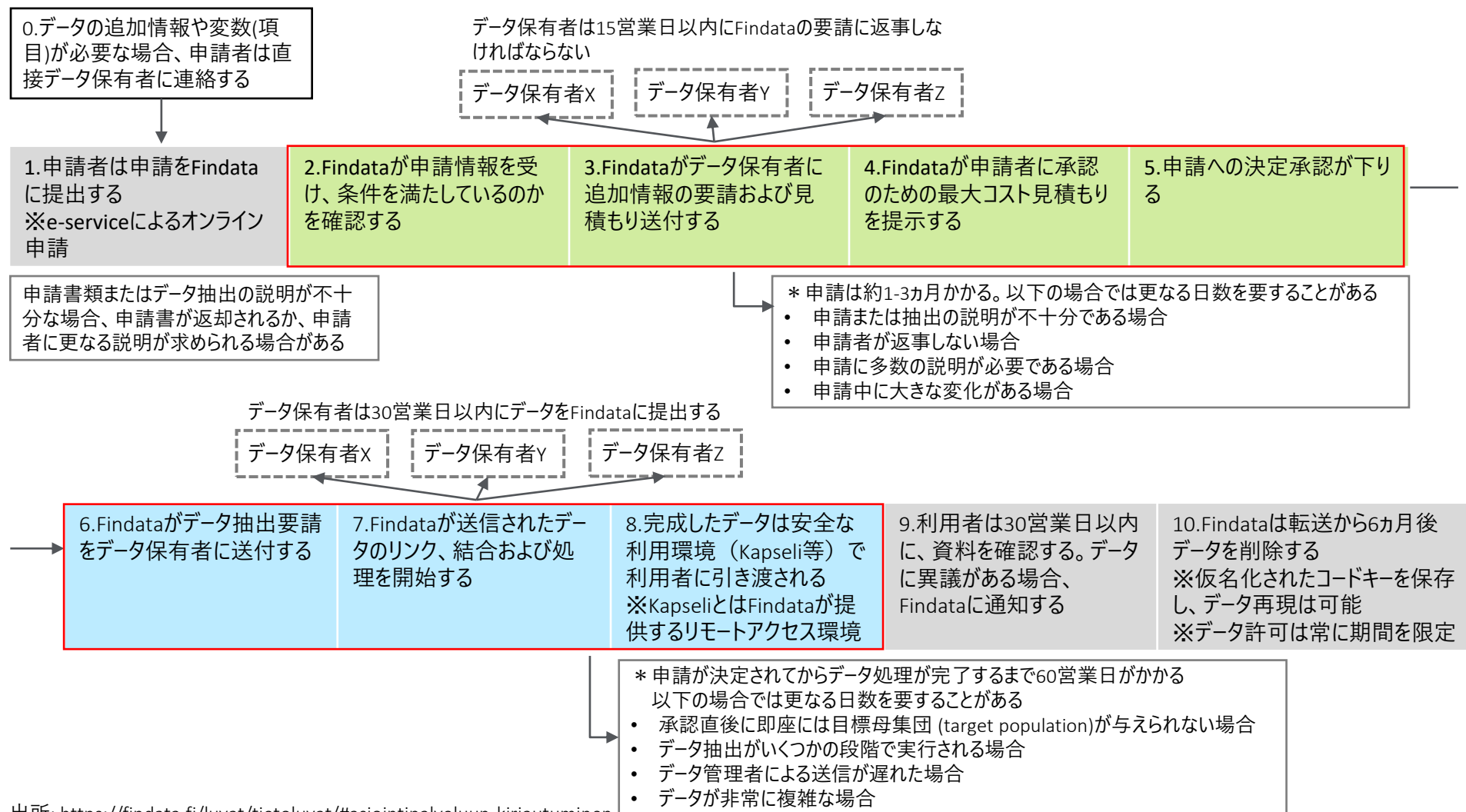
Findataでの審査の手順および基準（概要）

No	手順	基準（概要）
1	申請情報	<ul style="list-style-type: none">申請情報が正確であり、矛盾がない申請者の連絡先が明確に記載されている
2	請求書	<ul style="list-style-type: none">請求の詳細は適切に記載されている請求書の内容が申請時の情報と合致している
3	データの利用目的	<ul style="list-style-type: none">健康社会データの二次利用に関する法律に準拠したデータの使用の全体像を申請書に記載される申請するデータ情報は、データの利用目的と一致する論文は、少なくとも学士号の最終課題であるデータの使用に関する説明にミス、矛盾がない
4	抽出基準の定義	<ul style="list-style-type: none">ターゲットグループが明確に指定されている
5	抽出条件の定義	<ul style="list-style-type: none">対象被験者および相対抽出の基準は明確である
6	抽出する情報の要求	<ul style="list-style-type: none">抽出するレジスタ情報を明確にする
7	結合するその他のデータ	<ul style="list-style-type: none">Findataから申請されたデータが結合される他のデータを明確に指定する
8	データ処理	<ul style="list-style-type: none">FindataのKapseli以外での開示を希望する場合、理由を明確に示す
9	データ活用計画	<ul style="list-style-type: none">必要な添付ファイルが添付されているデータ活用計画には明らかな矛盾がない
10	追加情報と添付ファイル	<ul style="list-style-type: none">必要な添付ファイルが提出されている必要なすべての情報を一覧表示にされている

二次利用の申請者は以下の手順に沿って医療情報データにアクセスすることが可能である

① データ利活用基盤・データベースの概要

Findataの二次利用の申請プロセス



出所: <https://findata.fi/luvat/tietoluvat/#asiointipalveluun-kirjautuminen>

2023年の利用申請数が296件、過去申請を含めて351件が承認されている。利用目的として科学研究が最も多く、申請者のうち半分以上は公的機関である

① データ利活用基盤・データベースの概要

Findataのデータ利用状況

項目		2020	2021	2022	2023
新規申請数		319件	312件	270件	296件
許可決定数		175件	262件	284件	351件
申請処理時間		/	80日	年初88日：年末66日	年初78日：年末72日
データの 利用目的	科学研究	299	294	257	270
	統計学	10	8	10	8
	行政機関の計画・報告	5	4	1	2
	開発・イノベーション	4	3	0	2
	トレーニング	1	3	0	0
	その他	0	0	2	0
申請者	公的機関	246	206	192	170
	民間組織	48	91	69	105
	その他(個人等)	25	14	9	21
申請件数が最も多い データ管理機関 (一部抜粋)		/ <ul style="list-style-type: none">フィンランド国立保健福祉研究所 (THL)フィンランド社会保険機構 (Kela)フィンランド統計局 (Statistics Finland)デジタル・人口情報サービス庁 (DVV)ヘルシンキ・ウーシマー病院			

出所:<https://findata.fi/en/permits/issued-permits/>

Findataの申請者のうち、半数以上は公的機関であるが、民間からの申請も増加している。
また、データ許可の利用概要については公開されている

① データ利活用基盤・データベースの概要

Findataのデータ利用例(2023年)

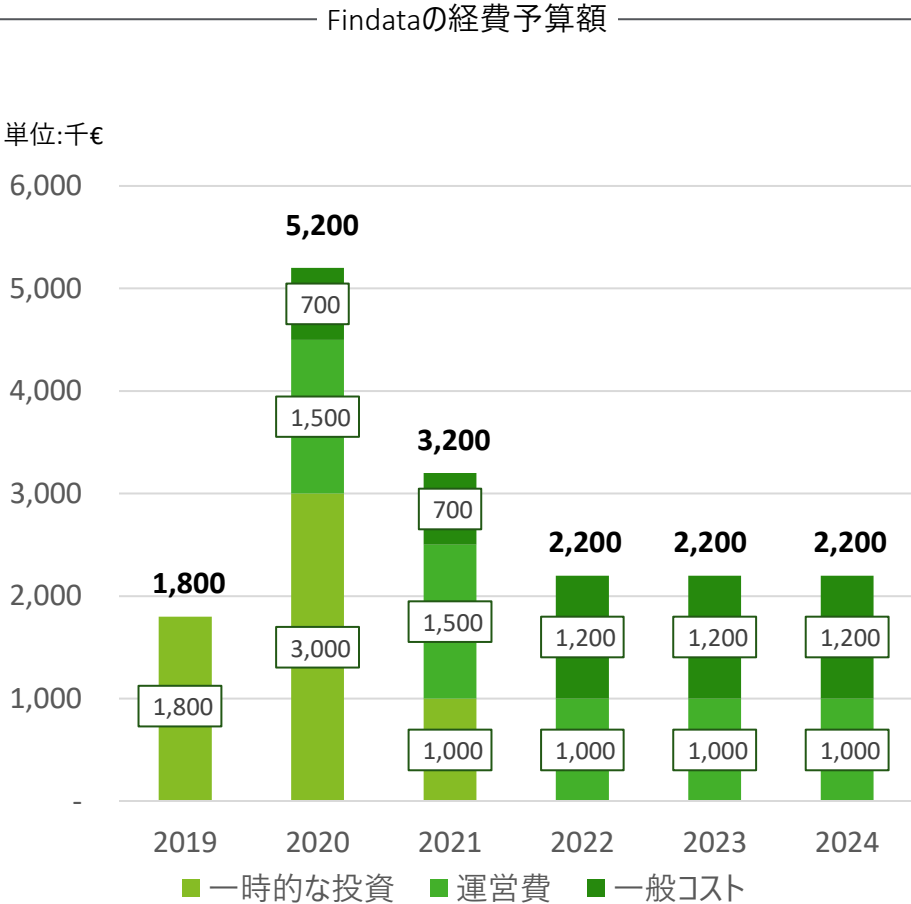
分野	利用者	データ利用例
医薬品・医療機器開発	IQVIA Finland Oy	<ul style="list-style-type: none">• 多国籍コホート研究として、40歳以上の慢性閉塞性肺疾患（COPD）患者を対象に、吸入粉末薬剤（TRIMBOW）の使用を開始した場合の心血管疾患および脳血管疾患の発生リスクを、吸入スプレー薬剤と比較して評価する。ヨーロッパのデータベースから収集された二次データを用い、後ろ向きに実施される
疫学分析	ヘルシンキ大学中央病院	<ul style="list-style-type: none">• 月状骨軟化症（Lunatummalasia）の疫学および病因を広範に調査し、Hilmoデータを用いてフィンランドにおける月状骨軟化症の発生率および疾患の発症や進行に影響を及ぼす要因を明らかにする。さらに、フィンランド社会保険機構(Kela)および年金保険センターのデータに基づき、全国的な登録データを活用して、月状骨軟化症がどの程度病欠や障害年金、さらには社会的費用を引き起こしているかを検討する
疾患研究	MedEngine Oy	<ul style="list-style-type: none">• 2型糖尿病（T2D）の早期段階で患者を特定し、初期治療を最適化することは、治療成績の向上や医療費削減につながると期待されている。日常的に収集される健康データを用いて訓練された機械学習（ML）モデルは、集団レベルでT2Dの発症を正確に予測することが可能である。本レジストリは、MLベースのアルゴリズムを検証し、T2Dの進行とそれに伴う医療利用、コスト、死亡率への影響をフィンランドにおいて予測する• 膿疱性乾癬（Pustular psoriasis, PP）は、繰り返し発生する肉眼的に観察可能で痛みを伴う膿疱が特徴の、重度の炎症性皮膚疾患群である。PPは異なる臨床的表現型として現れるが、その中でも最も一般的なのは全身性膿疱性乾癬（GPP）と手掌足底膿疱症（PPP）である。本研究は、フィンランドの全国的なレジストリから後ろ向きに収集したデータに基づき、GPPおよびPPPの疫学、臨床的特徴、医療資源の利用状況、そして医療費について特性評価を行う

出所:<https://findata.fi/en/permits/issued-permits/>

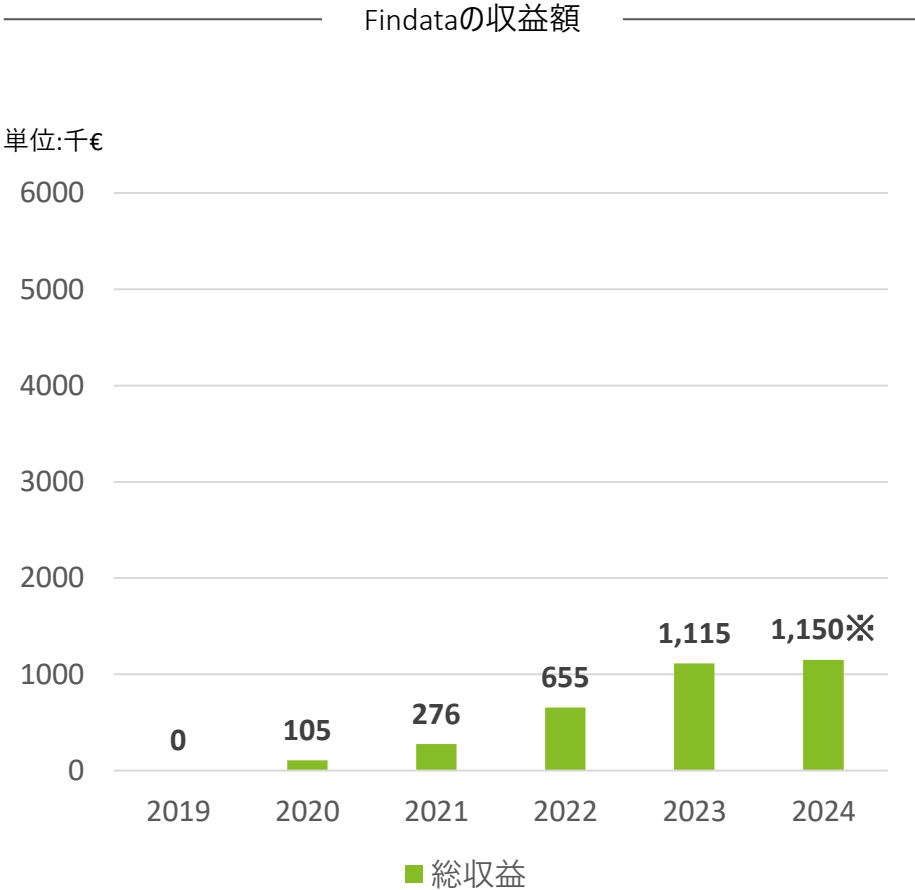
Findataでは、フィンランド社会保健省によって予算が配分され、2019年から2021年にかけては初期構築のための投資資金が割り当てられた

① データ利活用基盤・データベースの概要

Findataの経費予算と収益



出所:<https://budjetti.vm.fi/index.jsp>



※：2024年の総収益は予算額になる

Findataでは、2019年から2021年にかけては初期構築のための9億2,742万円の一時的な投資資金が割り当てられた

Findataの経費予算と収益(円換算)

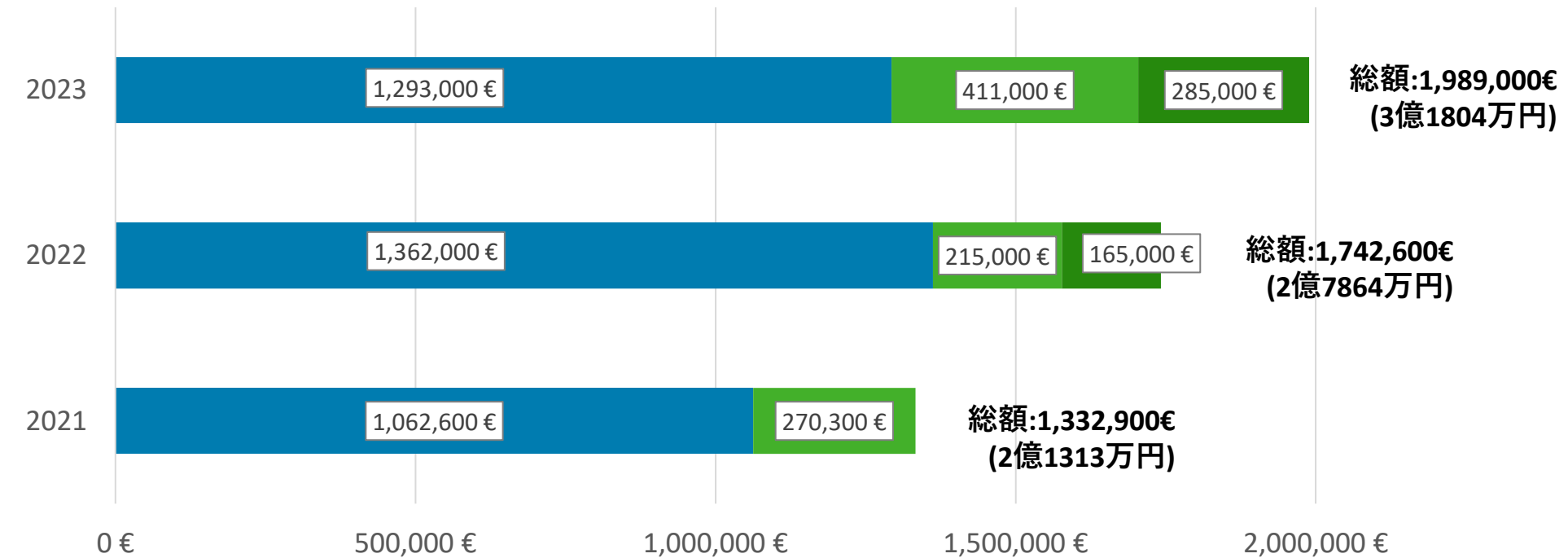
		2019	2020	2021	2022	2023	2024
経費 予算額	一時的な 投資	1,800,000ㄱ-□ (2億8782万円)	3,000,000ㄱ-□ (4億7970万円)	1,000,000ㄱ-□ (1億5990万円)	-	-	-
	運営費	-	1,500,000ㄱ-□ (2億3985万円)	1,500,000ㄱ-□ (2億3985万円)	1,200,000ㄱ-□ (1億9188万円)	1,200,000ㄱ-□ (1億9188万円)	1,200,000ㄱ-□ (1億9188万円)
	一般 コスト	-	700,000ㄱ-□ (1億1193万円)	700,000ㄱ-□ (1億1193万円)	1,000,000ㄱ-□ (1億5990万円)	1,000,000ㄱ-□ (1億5990万円)	1,000,000ㄱ-□ (1億5990万円)
総収益		0	105,000ㄱ-□ (1678万円)	276,000ㄱ-□ (4413万円)	655,000ㄱ-□ (1億473万円)	1,115,000ㄱ-□ (1億7828万円)	1,115,000ㄱ-□ (1億7828万円)

※2025.02.15 13:30の為替レートにより計算 1ㄱ-□ = 159.90円

Findataのデータ二次利用にかかるコストは、データ管理者によるデータ抽出コスト、Findataのデータ処理コスト、およびFindataの意思決定コストで構成されている

① データ利活用基盤・データベースの概要

Findataのデータ二次利用にかかるコスト



※2025.02.15 13:30の為替レートにより計算 1ユーロ = 159.90円

- データ管理者によるデータ抽出コスト
- Findataのデータ処理コスト(データ結合、仮名化・匿名化等のコスト)
- Findataの意思決定コスト(データ許可の発行にかかる意思決定コスト)

出所:<https://budjetti.vm.fi/index.jsp>

Findataの利用料金はフィンランド社会保健省の政令に基づき、以下の4つで構成されている

① データ利活用基盤・データベースの概要

Findataの利用料金の構成

a. 決定手数料

- 社会保健省の政令「社会保健分野のデータ許可機関の業務に関する手数料について」に基づいて設定されている固定価格である
- 失効申請の処理費用については例外であり、時間単位の料金に基づいて徴収
- 公法上のサービスおよび決定手数料には付加価値税（VAT）は課されない



b. データ管理者によるデータ抽出費

- 費用は各データ管理者の規定に基づくものである
- データ管理者は最終的なデータ利用者を基にして、抽出費用に付加価値税（VAT）を加算するか否かを判断する



c. データ処理と転送費

- データ処理における時間単価は固定されており、社会保健省の政令「社会保健分野のデータ許可機関の業務に関する手数料について」に基づくものである
- データ処理と転送の費用に付加価値税（VAT）は課されない



d. Kapseli料金

- データ分析にKapseli環境を使用する場合、月額料金が発生し、付加価値税（VAT）が加算される



Findata利用料

Findataの意思決定料は社会保健省の政令に基づいて設定されている固定価格である

① データ利活用基盤・データベースの概要

Findataの利用料金

a.意思決定料

	料金カテゴリ	条件	料金
データ許可	卒業論文用データ許可	申請者がフィンランドまたは他のEU/EEA諸国に居住している場合である。プロジェクトが複数の論文や、論文に関連しない成果物を生じる場合には、別の許可料金カテゴリが適用される	250 € (40,743円)
	限定データ許可	申請処理が7時間未満であり、申請者の事業所がEUまたはEEAに所在する場合である。例えば、データや新しい抽出内容に変更がない新しい研究質問などが該当する。	300 € (48,891円)
	通常データ許可	1～15のデータ管理者に関するデータが対象であり、申請者がEUまたはEEAに所在する場合である。申請処理の途中で申請者の主導で抽出内容が変更され、14時間以上かかる場合には「拡張データ許可」が適用される	1,400 € (228,158円)
	研究者主導の研究用通常データ許可	フィンランドまたはEU/EEA内に事業所を持つ申請者が対象であり、処理時間が7時間以上14時間未満の研究である。外部資金を伴わない、または公的医療機関、大学、研究所、非営利団体から資金提供される研究が対象である	700 € (114,079円)
	拡張データ許可	以下のいずれかを満たす場合に必要である： － 15以上のデータ管理者に関するデータを含む場合 － 抽出内容が申請者の主導で変更され、複数回のコスト見積もりが必要な場合	2,000 € (325,940円)
	研究者主導の研究用拡張データ許可	フィンランドまたはEU/EEA内に事業所を持つ申請者が対象であり、処理時間が14時間以上の研究である。外部資金を伴わない、または公的医療機関、大学、研究所、非営利団体から資金提供される研究が対象である	1,000 € (162,970円)
	EU/EEA外地域用データ許可	申請者の事業所がEUまたはEEA以外の国に所在する場合である	3,000 € (488,910円)

2024/12/23 17:45の為替レートにより計算されている 1ユーロ=162.97円

出所： <https://findata.fi/en/pricing/>

Findataの意思決定料は社会保健省の政令に基づいて設定されている固定価格である

① データ利活用基盤・データベースの概要

Findataの利用料金

a.意思決定料

	料金カテゴリ	条件	料金
変更許可	軽微な変更または修正	申請が以下の変更または修正を求める場合である： – 許可期間の延長 – データ処理を許可された人物の変更 – データとともに公開された資料の管理者の変更 – FindataのKapseliリモートアクセスシステムから他のサービスプロバイダーの安全な処理環境への変更	300 € (48,891円)
	その他の変更または修正	申請が以下の1つ以上の変更または修正を求める場合である： – Findataによって付与された許可の監視期間の延長（調査対象者と抽出データが同一である場合） – EUまたはEEA以外の国での許可から逸脱したデータ処理の延長 – 抽出後のデータ収集への変数の追加 – 付与された許可から逸脱して、他の法律の下でデータを追加	600 € (97,782円)
データ要求	卒業論文用データ要求	卒業論文に関連するデータ要求であり、申請者がフィンランドまたは他のEU/EEA諸国に居住している場合である。プロジェクトが複数の論文や論文に関連しない成果物を生じる場合、通常または拡張情報許可が適用される	250 € (40,743円)
	研究者主導の研究用データ要求決定	外部資金を伴わない、または公的医療機関、大学、研究所、その他の公共または非営利団体から資金提供を受ける研究である。申請者はEUまたはEEA内に拠点を持ち、研究資金の説明を添付する必要がある	1,250 € (203,713円)
	その他のデータ要求	その他すべてのデータ要求	2,500 € (407,425円)

出所： <https://findata.fi/en/pricing/>

Findataの意思決定料は社会保健省の政令に基づいて設定されている固定価格である

① データ利活用基盤・データベースの概要

Findataの利用料金

a.意思決定料

	料金カテゴリ	条件	料金
失効申請	無料の失効申請	申請者がFindataによる処理開始前に申請を取り下げた場合である	0 €
	許可料金の約20%を請求	申請者がFindataによる処理開始後、データ抽出のコスト見積もり要求が管理者に送信される前に申請を取り下げた場合	88 €/h (14,341円/h)
	許可料金の約60%を請求	データ抽出のコスト見積もり要求が管理者に送信された後、コスト見積もりの概要が申請者に送信される前に申請を取り下げた場合	88 €/h (14,341円/h)
	許可料金の約75%を請求	コスト見積もりの概要を受け取った後に申請を取り下げた場合	88 €/h (14,341円/h)
既存データセット許可	Findataの既製データセット	他のデータが結合されていないFindataの既製データセットに対するデータ許可である • データ許可料金に加えて、Findataの抽出費用が作業量に基づいて課される。 • 他のサービスプロバイダーの処理環境への提供には、2時間の作業が含まれ、料金は294 €（VAT 0%）である。見積もりが2時間を超える場合、顧客にその旨が通知される	300 € (48,891円)

出所： <https://findata.fi/en/pricing/>

2023年1月1日から2024年10月15日までのデータ管理者によるデータ抽出の平均コストは1,600ユーロである

① データ利活用基盤・データベースの概要

Findataの利用料金

b.データ管理者によるデータ抽出料金

データ管理者		平均見積もり(€)	平均コスト(€)
② データ管理者によるデータ抽出費	フィンランド健康福祉研究所	2,300 (374,831円)	1,000 (162,970円)
	フィンランド社会保険機構 (Kela)	3,900 (635,583円)	1,800 (293,346円)
	フィンランド統計局	600 (97,782円)	300 (48,891円)
	デジタル・人口データサービス機関 (DVV)	5,100 (831,147円)	3,800 (619,286円)
	Fフィンランド年金センター (ETK)	1,700 (277,049円)	1,000 (162,970円)
	HUS	5,200 (847,444円)	3,100 (505,207円)
	南西フィンランド福祉サービス郡 (Varha)	5,100 (831,147円)	3,400 (554,098円)
	ピルカンマー福祉サービス郡	6,200 (1,010,414円)	2,000 (325,940円)
	北オストロボスニア福祉サービス郡 (Pohde)	2,600 (423,722円)	1,300 (211,861円)

出所： <https://findata.fi/en/pricing/>

データ処理における時間単価は社会保健省の政令により固定されている

① データ利活用基盤・データベースの概要

Findataの利用料金

c.データ処理と転送費

	料金カテゴリ	条件	料金
③ データ 処理 と 転 送 費	データ処理	データの結合、前処理、仮名化、匿名化に対する時間料金である。最低1時間が必要である	147€ / h (23,957円/h)
	データ転送	転送するZIPフォルダの数に基づき料金が設定される。1つのZIPフォルダを転送するのに2.5時間がかかる ・ 1フォルダあたりの料金は以下の通りである：2.5 × 147 EUR = 367.50 EUR ・ 各ZIPフォルダのサイズは最大4GBとする	367.50 € /ZIPフォルダ (59,891円/ZIPフォルダ)

出所： <https://findata.fi/en/pricing/>

データ分析にKapseli環境を使用する場合、月額料金が発生し、付加価値税（VAT）が加算される

① データ利活用基盤・データベースの概要

Findataの利用料金

d.Kapseli利用料金

	製品	特徴	Windows (VAT抜き)	アカデミックWindows (VAT抜き)	Linux (VAT抜き)
基本料金	Kapseli S	- 8GB RAM - 4コア	1,770 €/年 (288,457円/年)	1,640 €/年 (267,271円/年)	1,305 €/年 (212,676円/年)
	Kapseli M	- 16GB RAM - 6コア	2,390 €/年 (389,489円/年)	2,120 €/年 (345,496円/年)	1,660 €/年 (270,530円/年)
	Kapseli L	- 32GB RAM - 8コア	3,560 €/年 (580,173円/年)	2,935 €/年 (478,317円/年)	2,345 €/年 (382,165円/年)
	Kapseli XL	- 64GB RAM - 8コア	5,825 €/年 (949,399円/年)	4,790 €/年 (780,626円/年)	3,300 €/年 (537,801円/年)

※：1つのマシンパッケージ（S/M/L/XL）は、約10人のユーザー向けに設計されている。料金には最大10人までのユーザー利用権が含まれる。加入者が10人以上を希望する場合、1ユーザーごとに月額 16.16 € + 25.5% VAT が追加料金として課される。

	料金カテゴリ	条件	料金
カスタマイズ料金	処理システムの変更	作業時間のコストに基づく固定価格	470 €/変更 (76,596円/変更)
	技術的カスタマイズ	リモート環境のカスタマイズにかかる作業時間のコスト (例：Kapseliパッケージへのユーザー独自ソフトウェアのインストール（ユーザー自身のライセンスを使用）、その他の技術的カスタマイズ)	別途見積
ストレージサービス	ストレージサービス処理費用	処理費用と最小12か月分の注文が新規ストレージ注文時に請求される	460 € (74,966円)
	ストレージサービス（月額料金）	6か月ごとに請求される	7 €/月（1,141円/月）

Findataは、欧州委員会、フィンランド社会保健省、利用者からの資金と利用料を財源とし、データ抽出や内部処理、意思決定にコストを要している

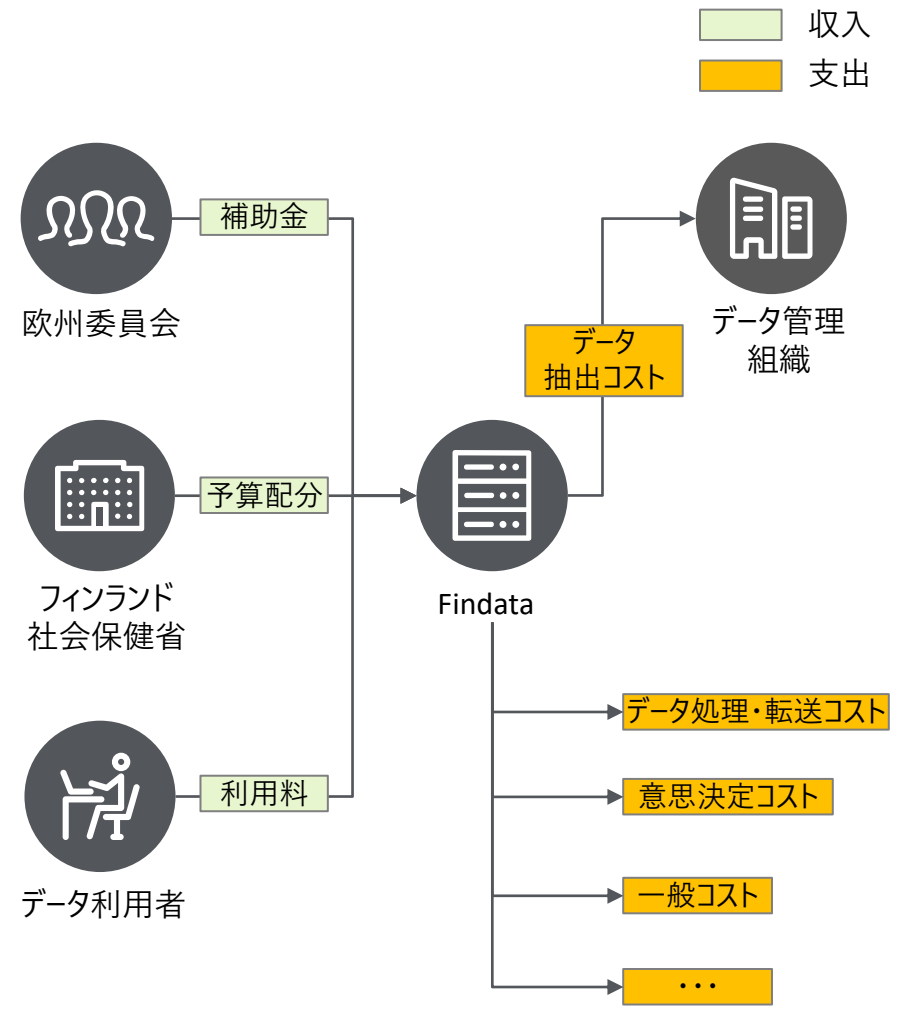
① データ利活用基盤・データベースの概要

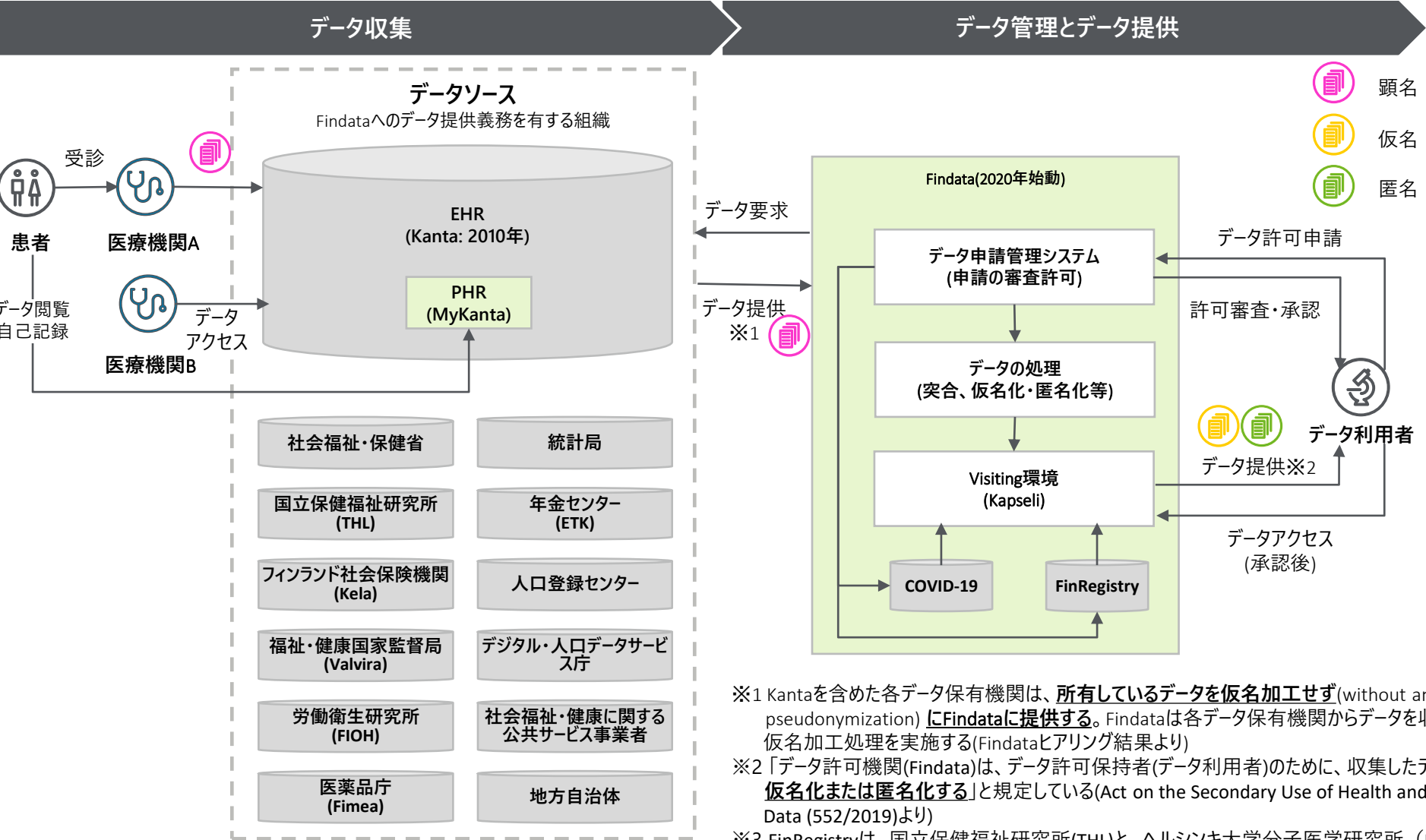
Findataの補助金

項目	概要
提供元	欧州委員会
補助金額	250万ユーロ
自己負担額	40%
目的	FinHITSプロジェクトを通じた健康データの二次利用促進および欧州健康データ空間（EHDS）の実現準備
プロジェクト期間	4年間
資金用途	<ul style="list-style-type: none">・ eサービスの開発・ 全国データカタログの開発・ Kapseli環境の開発・ 情報システムの構築・ HealthData@EUパイロットプロジェクトおよび既存データセット関連作業の継続

出所： <https://findata.fi/en/news/findata-receives-2-5-million-euros-in-eu-funding/>

Findataのビジネスモデル





※1 Kantaを含めた各データ保有機関は、**所有しているデータを仮名加工せず**(without any pseudonymization) **にFindataに提供する**。Findataは各データ保有機関からデータを収集後、仮名加工処理を実施する(Findataヒアリング結果より)

※2 「データ許可機関(Findata)は、データ許可保持者(データ利用者)のために、収集したデータを**仮名化または匿名化する**」と規定している(Act on the Secondary Use of Health and Social Data (552/2019)より)

※3 FinRegistryは、国立保健福祉研究所(THL)と、ヘルシンキ大学分子医学研究所 (FIMM) のデータサイエンス遺伝疫学研究グループが共同で実施する研究プロジェクトである Findataは、利用者がFinRegistryプロジェクトのデータをより迅速・直接に利用できるように、**2024春から既製データセットを順次作成している**

フィンランドのKantaは社会保険庁管轄のもと、医療情報や社会福祉データを蓄積したデータベースである

② インターフェース・システムアーキテクチャ等

Kanta概要

DB名称		Kanta	
管轄	組織名	社会保険庁 (フィンランド語: Kansaneläkelaitos, Kela)	
	組織概要	国内の社会保障の運用を担う組織	
設置目的		社会福祉およびヘルスケア分野においてのデジタルサービスを通じて関連情報に容易にアクセス可能となることで、市民と社会福祉・ヘルスケアサービス提供者双方にとって利益をもたらすという目的で設立	
データの種類		➤ 処方箋センター	• 処方箋、リフィル処方箋(renewals)、調剤情報(dispensation)、予約情報
		• 医薬品データベース	• 医薬品情報、代替調剤規則(substitution rules)、薬価情報、償還規則(reimbursement rules)
		• 患者データレポジトリ	• 外来診察データ(Encounter data)、構造化済み医療記録(structured health records)、画像データ、過去の健康記録
		• データ管理サービス	• 情報通知、同意、同意の制限(consent restrictions)、リビング・ウィル、臓器提供意思表明書、所見(problems)、諸手続き(procedures)、検査結果(lab results)、ワクチン接種履歴、画像検査結果、リスク所見(risks)、生理的所見、社会的サービスについてのサマリ(customer summary for social services)
		• 社会サービスデータレポジトリ	• サービス項目(Service items)、顧客の文書(customer documents)、過去データ(legacy data)
		• Personal Health Records	• アンケート回答結果、その他市民によって保存されたデータ
データ量 (2024末)		680万人累計41億件健康診療記録 ➤ 社会福祉データ: 累計1億1,338万件 ➤ 電子処方箋: 累計8,384万件	

Findataではデータアクセス許可証を必要とする個人データ処理用の安全なオペレーティング環境kapseliを提供している

② インターフェース・システムアーキテクチャ等

FindataのVisiting解析環境kapseli

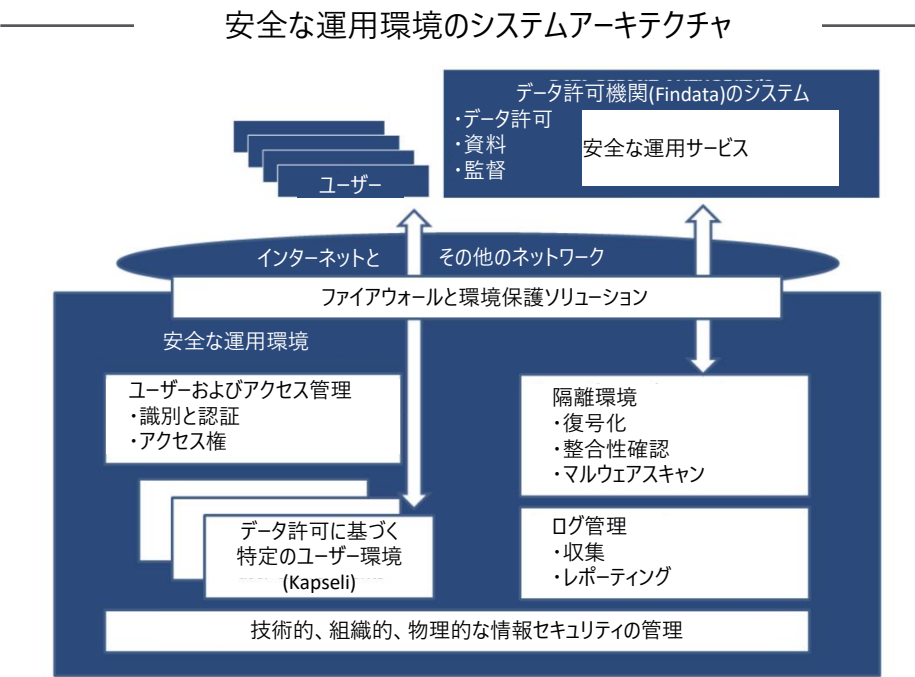
概要	データ使用許可を得た人が個人に関するデータを処理する際に利用するFINDATAが提供している安全なオペレーティング環境	
Visiting解析環境のメリット	<ul style="list-style-type: none">セキュリティが担保されていること(例)データ使用許可に応じてプロジェクトが構築され、異なるデータ使用許可では他のプロジェクトにアクセスできない	
利用者	FINDATAの審査によって許可された利用者	
利用環境	OS	<ul style="list-style-type: none">Linux、Windows
	プログラミング言語	<ul style="list-style-type: none">Golang、Python、R
	IDE(統合開発環境)	<ul style="list-style-type: none">Anaconda3、RStudio
	統計解析ツール	<ul style="list-style-type: none">SAS、SPSS、Stata
	その他、使用できるツール	<ul style="list-style-type: none">Acrobat Reader DC7-Zip、Pigz(ファイル圧縮・解凍)Emacs、Notepad++、VSCode(テキストエディタ)GeoDa、QGIS(空間データ分析、地理視覚化等)GitGoogle Chrome(HTML 形式のドキュメントを開く等、ローカルでのみ使用可能)LibreOfficeLyXMiKTeXNextflowWinPython

Findataを用いてデータ分析を行うには、規制の要件を満たす環境でのみ許可され、データセキュリティ評価機関が評価を行った証明書が必要になる

② インターフェース・システムアーキテクチャ等

Findataの安全運用のためのシステムアーキテクチャ

- 「健康・社会データの二次利用に関する法律」第20条第2項および第21条から第29条により**個人レベルのデータ分析は、規制の要件を満たす環境でのみ許可される**
- そのため、Findataは安全な運用環境に関する要件を定めた規制を発行し、技術、運用者の信頼性評価等に関する要件等を掲載した
- これらの要件は、Findata自身の運用環境と同等の情報セキュリティレベルを必要とするものである**



項目	技術要件
識別	<ul style="list-style-type: none">初回識別は強力な電子手段（例: Suomi.fi）を使用二要素認証を必須化特定の条件下で二要素認証を省略可能だが、代わりにユーザー本人確認が必要
ユーザーアクセス権管理	<ul style="list-style-type: none">アクセス権はデータ許可に基づき制限され、データセット間の移動は禁じられる必要最低限のアクセス権を割り当てデータ許可が失効すると自動的にアクセス権がロックされるアクセス権が失効した後、データは6か月以内に自動削除
環境の保護	<ul style="list-style-type: none">ユーザー環境をインターネットから分離し、ファイアウォールを適用ユーザー環境内のデータ転送は許可された範囲に限定端末デバイスの外部接続禁止（例: USB使用禁止）定期的なセキュリティスキャンを実施し、脆弱性を管理
ログ管理	<ul style="list-style-type: none">使用者、目的、使用時間などを記録技術ログデータは5年間保管定期的なログデータの分析と不正行為の監視
環境の管理・監視	<ul style="list-style-type: none">環境は24時間7日自動監視され、対応責任が明確化されている管理者の作業はログに記録されるシステム管理には暗号化された通信接続を使用する
データ削除	<ul style="list-style-type: none">➤ データ許可が失効した6か月以内にデータを削除➤ 削除は「KATAKRI 2020」I-19※の基準に基づき実施

出所 : <https://findata.fi/en/kapseli/regulation-on-secure-operating-environments/>

※KATAKRI（Kansallinen Turvallisuusauditointikriteeristö）：フィンランド国防省が策定した「国家セキュリティ監査基準」のこと

Findataのデータ管理者はCSV形式のデータをFindataの暗号化ツールにて暗号化し、Nextcloudにて転送する

② インターフェース・システムアーキテクチャ等

Findata管理者のデータ転送ステップ

Findataへ担当者
情報を通知

- 下記の情報をFindataに送信する
 - データ許可ジャーナル番号
 - データを提供する組織名
 - 対象グループを受け取る担当者名
 - データをFindataに引き渡す担当者名

Nextcloud ID
の取得

- Findataのカスタマーサービスにログイン
- 「NextcloudユーザーIDの注文」を選択し、フォームを記入して提出
- 提出後、data@findata.fi に通知
- CSCから接続完了の連絡が来たら、次に進む

データを暗号化して
Nextcloudに転送

- 必要なファイルをダウンロード
 - Findata Encryption Tool
 - csc.pub
 - journal_number.txt
- CSV形式のデータを一つのzipファイルにまとめる**(データ説明ファイルを含む)
- Findata Encryption Toolで.zipファイルを暗号化**し、以下の2ファイルを生成
 - .c4ghファイル
 - .jsonファイル
- Nextcloud**内のジャーナル番号フォルダに以下の3つをアップロード
 - journal_number.txt
 - .c4ghファイル
 - .jsonファイル
- Findataから転送成功の確認を受けた後、作業フォルダ内の個人データを完全削除**

FinOMOPではフィンランド全体を対象とした包括的な人口ベースの医療データネットワークを構築することを目的とし、OMOP CDMへの用語マッピングを実施している

② インターフェース・システムアーキテクチャ等

FinOMOP：プロジェクト概要

プロジェクト名	OHDSI Finland（FinOMOP）
背景・目的	<ul style="list-style-type: none">フィンランドの医療データは一次医療から三次医療まで複数システム・レジストリ・ゲノムプロジェクトに分散しており、相互運用性の確保が課題FinOMOPプロジェクトは、主要な医療データソース（EMR〈電子カルテ〉、レジストリ、ゲノムシーケンスプロジェクト、一次・二次医療など）をOMOP CDMへ高品質かつ詳細にマッピングし、<u>フィンランド全体を対象とした包括的な人口ベースの医療データネットワークを構築することを目的としている</u>
データパートナー	<ul style="list-style-type: none">病院データ<ul style="list-style-type: none">✓ ヘルシンキ・ウーシマー病院地区(HUS)✓ ピルカンマー福祉サービス郡(PIRHA)✓ 南西フィンランド福祉サービス郡(VARHA)✓ 北サヴォ福祉サービス郡(PSSHP)✓ オウル大学病院(PPSHP)レジストリ<ul style="list-style-type: none">✓ FinnGen✓ 国立保健福祉研究所(THL)
概要・方法	<ul style="list-style-type: none">OMOP CDM変換プロジェクト<ul style="list-style-type: none">✓ 2019年に開始され、EU IMI2 EHDENデータパートナーシップおよび地域資金からの支援を受けている一次医療<ul style="list-style-type: none">✓ フィンランド保健福祉研究所(THL)が管理する法令で義務付けられた人口レジストリを通じてカバーされている二次・三次医療<ul style="list-style-type: none">✓ フィンランド国内の5大学病院が担っており、このうち3院がOMOPへの変換を完了している(人口の70%以上をカバー)。2025年までに完全カバレッジを目指している✓ FinnGenプロジェクト全国規模の官民共同ゲノムプロジェクトであり、50万人のバイオバンク参加者から得られる生殖細胞系ゲノムデータを対象としている用語マッピング<ul style="list-style-type: none">✓ Githubリポジトリによって管理され、各国内用語ごとのUSAGIファイルをバージョン管理している。これらは定期的に変換され、OMOPの概念テーブルおよび概念間関係テーブルに取り込まれているETLのコーディング<ul style="list-style-type: none">✓ EHDEN認定の中小企業(SME)に委託されている
結果(2023)	フィンランド人 410万人分のプライマリOMOPマッピングが完了 し、すべての主要OMOP臨床データテーブルを詳細かつ包括的に網羅している

Findataは、「健康・社会データの二次利用に関する法律」に基づき、データ管理者&利用者に対して、データ記述、安全な運用環境、データ許可申請に関する規則を発布している

② インターフェース・システムアーキテクチャ等

Findataによるガバナンス体制

番号	規則名	規制対象	発効日	法律根拠	概要
①	『データ記述のデータ内容、概念、およびデータ構造に関する規則』 Regulation of the Health and Social Data Permit Authority: Data contents, concepts and data structures for data descriptions	• データ管理者	2021/02/01	「健康・社会データの二次利用に関する法律」第12条第2項	二次利用法に定義されるデータ管理者は、 <u>データ内容に関するデータ記述(カタログ)を作成しなければならない</u>
②	『安全な運用環境に関する要件』 Regulation by the Finnish Social and Health Data Permit Authority Findata: Requirements for other service providers' secure operating environments	• データ利用者 • データ管理者	2022/01/19	「健康・社会データの二次利用に関する法律」第22条第3項および第24条第2項	個人レベルのデータ分析は、本規則の <u>要件を満たす環境でのみ許可される</u> 。これらの要件は、Findataの安全な運用環境と同等レベルの情報セキュリティを確保することを求める
③	『データ許可申請およびデータ利用計画に関する規則』 Regulation on data content and data structures of data permit applications and data utilisation plans	• データ利用者	2022/11/01	「健康・社会データの二次利用に関する法律」第46条第2項	データ利用者は規則に従い、 <u>所定の項目を記載したデータ許可申請およびデータ利用計画を提出しなければならない</u>

出所： <https://findata.fi/en/services-and-instructions/regulations/>

Findataでは、データを一括してカタログ化するのではなく、データ規則を發布し、それに基づいて各データ提供組織が規則に従ったデータカタログを作成することで統一性を図っている

② インターフェース・システムアーキテクチャ等

① Findataにおけるデータカタログ作成に関する規則

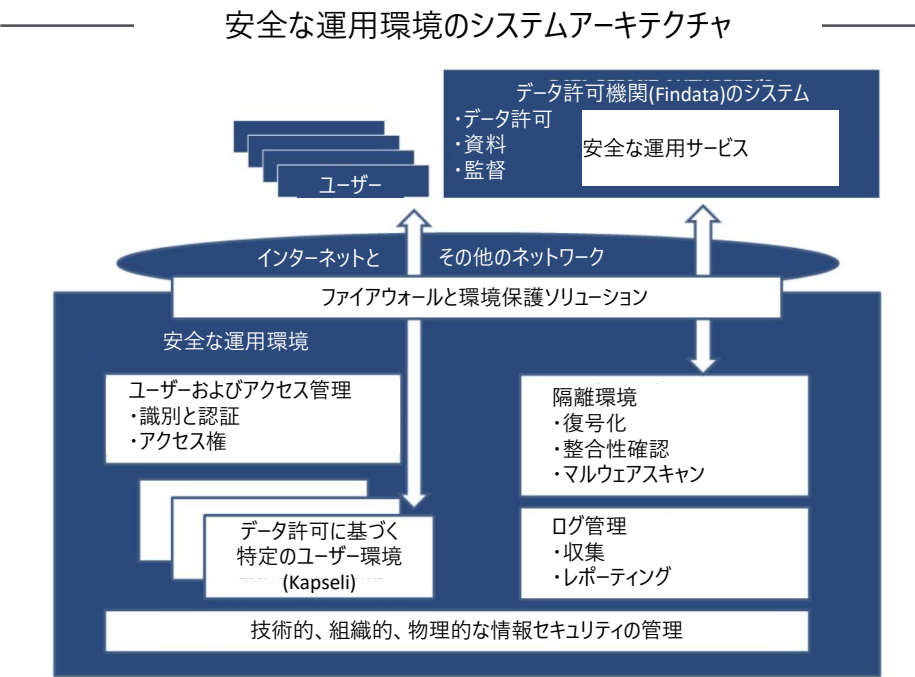
規制名	「データ説明に関する規制」 Regulation on data contents, concepts and data structures of data descriptions	
発行日	2021年2月1日	
政令根拠	「登録者およびデータ許可当局のデータ記述作成義務に関する規則」 Sosiaali- ja terveysministeriön asetus rekisterinpitäjän ja Tietolupaviranomaisen velvollisuudesta laatia aineistokuvaus	
適用対象	「健康・社会データの二次利用に関する法律」によりFindataへのデータ提供義務を有する組織(データ管理者)	
目的	<ul style="list-style-type: none">健康および社会データリソースの記述に関する包括的な全国カタログを作成するデータリソースの統一された高品質なデータ説明を確保する健康および社会データソースの安全かつ効果的な利用を推進する相互運用性を促進する	
データ記述の範囲	データリソース、データセット、変数、(必要に応じて)コードリストレベルまでの記述	
データ記述の項目	<ul style="list-style-type: none">データ名説明管理組織データリソースに関連する人物/連絡先データ関連リンク利用条件対象集団地域カバレッジ参照期間の開始・終了日データソースのライフサイクル段階データキーワードデータ資源の種類	
利用ツール	データ説明編集ツール	Aineistoeditori(国立保健福祉研究所(THL)管理)
	データカタログ公開サイト	Aineistokatalogi(国立保健福祉研究所(THL)管理)
	コードリストツール/サーバー	Koodistot(デジタル・人口データサービス庁管理) koodistopalvelu(国立保健福祉研究所(THL)管理)

Findataを用いてデータ分析を行うには、規制の要件を満たす環境でのみ許可され、データセキュリティ評価機関が評価を行った証明書が必要になる

② インターフェース・システムアーキテクチャ等

②Findataの安全な運用環境に関する規則

- 「健康・社会データの二次利用に関する法律」第20条第2項および第21条から第29条により**個人レベルのデータ分析は、規制の要件を満たす環境でのみ許可される**
- そのため、Findataは安全な運用環境に関する要件を定めた規制を発行し、技術、運用者の信頼性評価等に関する要件等を掲載した
- これらの要件は、Findata自身の運用環境と同等の情報セキュリティレベルを必要とするものである**



項目	技術要件
識別	<ul style="list-style-type: none">初回識別は強力な電子手段（例: Suomi.fi）を使用二要素認証を必須化特定の条件下で二要素認証を省略可能だが、代わりにユーザー本人確認が必要
ユーザーアクセス権管理	<ul style="list-style-type: none">アクセス権はデータ許可に基づき制限され、データセット間の移動は禁じられる必要最低限のアクセス権を割り当てデータ許可が失効すると自動的にアクセス権がロックされるアクセス権が失効した後、データは6か月以内に自動削除
環境の保護	<ul style="list-style-type: none">ユーザー環境をインターネットから分離し、ファイアウォールを適用ユーザー環境内のデータ転送は許可された範囲に限定端末デバイスの外部接続禁止（例: USB使用禁止）定期的なセキュリティスキャンを実施し、脆弱性を管理
ログ管理	<ul style="list-style-type: none">使用者、目的、使用時間などを記録-技術ログデータは5年間保管-定期的なログデータの分析と不正行為の監視
環境の管理・監視	<ul style="list-style-type: none">環境は24/7自動監視され、対応責任が明確化されている管理者の作業はログに記録されるシステム管理には暗号化された通信接続を使用する
データ削除	<ul style="list-style-type: none">➤ データ許可が失効した6か月以内にデータを削除➤ 削除は「KATAKRI I 19」基準に基づき実施

出所 : <https://findata.fi/en/kapseli/regulation-on-secure-operating-environments/>

データ利用者は規則に従い、規定の言語で所定の項目を記載したデータ許可申請およびデータ利用計画をFindataに提出しなければならない

② インターフェース・システムアーキテクチャ等

③ データ許可申請に関する規則

データ許可申請

データ許可申請には以下の情報を含める必要がある：

- 申請者および連絡担当者の情報（氏名、住所、連絡先）
- データ管理者の特定
- データ使用目的（科学研究、統計、教育、行政計画・報告義務）
- 取得するデータの詳細（データの出所、対象期間、収集方法）
- 申請するデータと他のデータセットの統合の有無
- データのセキュリティ対策
- 取得データの保管および破棄の計画

データ利用計画

データ利用計画には以下の内容を含める必要がある：

- データ利用目的の詳細
- データを処理する担当者およびその組織
- GDPRおよびフィンランドのデータ保護法に準拠した法的根拠
- データ処理環境
- データの識別可能性
- データ処理期間と保管場所
- データ主体（個人）の権利行使の可否（例：削除要求の受け入れ可否）
- データ移転の有無（例：EU域外へのデータ転送を行う場合、GDPR第5章に基づく正当性を説明）

顧客データ法により、フィンランド保健福祉研究所（THL）はデータ構造の統一化を推進する義務を負っている

② インターフェース・システムアーキテクチャ等

フィンランドにおける医療情報の標準化

法律根拠

Sosiaali- ja terveydenhuollon asiakastietojen
käsittelystä(703/2023)
社会福祉および医療分野の顧客データの取り扱いについて

第4章 顧客文書の取り扱いに関する原則

第20条

顧客データのデータ構造は、電子的な顧客データおよび顧客データに対するアクセス権限の割り当て、利用、提供、保存および保護、ならびに活用を可能にしなければならない。また、当該データ構造は、第65条に規定する全国的な情報システムサービスおよびサービス提供者のその他の情報システムを通じて、顧客データを二次的利用目的のためにも活用できるようにしなければならない。

国立保健福祉研究所(THL) は、顧客データのデータ構造およびデータ内容、並びに当該データ構造において全国的に活用されるコード体系に関する規定を定めるものとする。

THLの対応

ガイドライン等の作成

社会福祉・医療サービスを提供する機関、薬局、顧客・患者データ情報システム事業者等の対象に対して、顧客情報のデータ構造および社会福祉・医療サービスの外部へ情報を提供する際の手順や要件等にかかわるガイドラインを作成

コードサーバーの運営

社会福祉・医療分野における電子患者記録・文書システムに必要な、コード体系、分類、用語集、および登録データを含む全国統一のデータ構造を管理するコードサーバーを運営し、無償で提供する

THLでは、社会福祉・医療サービスを提供する機関、顧客・患者データ情報システム事業者等の対象に対して、顧客情報のデータ構造等にかかわるガイドラインを発布している

② インターフェース・システムアーキテクチャ等

フィンランドにおける医療情報の標準化：データ構造

データ項目	使用されるコード体系
診断情報(Diagnoosit)	<ul style="list-style-type: none">• ICD• ICPC
受診理由(Käyntisyyt)	<ul style="list-style-type: none">• ICD-10• ICPC-2
リスク情報(Riskitiedot)	<ul style="list-style-type: none">• 診断や処置に関連するコードを利用• 薬剤禁忌：ATCコードを利用
処置(Toimenpiteet)	<ul style="list-style-type: none">• THL処置分類コード• 合併症：ICD-10
画像診断(Kuvantamistutkimukset)	<ul style="list-style-type: none">• 放射線診断コード• THL処置分類コード• 測定値や検査結果の異常度分類コード
臨床検査(Laboratoriotutkimukset)	<ul style="list-style-type: none">• FinLOINC• 臨床検査名リスト• 医療機関の独自コード
生理学的測定 (Fysiologiset mittaukset)	<ul style="list-style-type: none">• FinLOINC
投薬情報(Lääkitys)	<ul style="list-style-type: none">• 投薬リスト（導入中）

THLはコードサーバーKoodistopalvelinを運営し、患者情報システムにおけるデータの記録や保存に使用される統一されたデータ定義と構造を提供している

② インターフェース・システムアーキテクチャ等

フィンランドにおける医療情報の標準化：データ構造

名称	Koodistopalvelin
運営管理	THLが運営、Kelaが技術運用保守
概要	<ul style="list-style-type: none">フィンランドの社会福祉・医療分野において使用される標準化されたコードを一元的に管理・配布するためのオンラインサービスであるコード体系、分類、登録データ、用語集、データ定義等Koodistopalvelinにて公開されている
利用方法	<ul style="list-style-type: none">WebServicesインターフェースを提供しており、コードやデータ定義の更新情報を自動的に取得できる。変更点や新規に追加された定義を常に最新の状態で取り込むことで、システムを最新状態に保つことが可能ブラウザを介して直接にアクセスし、コードの一覧や個別項目を検索・閲覧できるExcelやテキスト、XMLなどの形式でダウンロードすることも可能であるため、ローカル環境に保存し独自に加工することも容易
メリット	<ul style="list-style-type: none">システム開発ベンダー：承認済みで品質が保証されたデータ定義を無料で入手できる。新規開発や既存システムの更新時に、正確かつ最新のコードを容易に組み込むことが可能社会福祉・医療の専門職：全国的に統一されたデータ定義を使用することで、構造化記録やデータ保存の質を向上させることができる

出所:<https://thl.fi/aiheet/tiedonhallinta-sosiaali-ja-terveysalalla/koodistopalvelu/jakelualustat>

フィンランドでは、社会福祉・医療分野の情報管理における相互運用性を実現するため、標準規格の活用が推進されている

② インターフェース・システムアーキテクチャ等

フィンランドにおける医療情報の標準化：標準規格

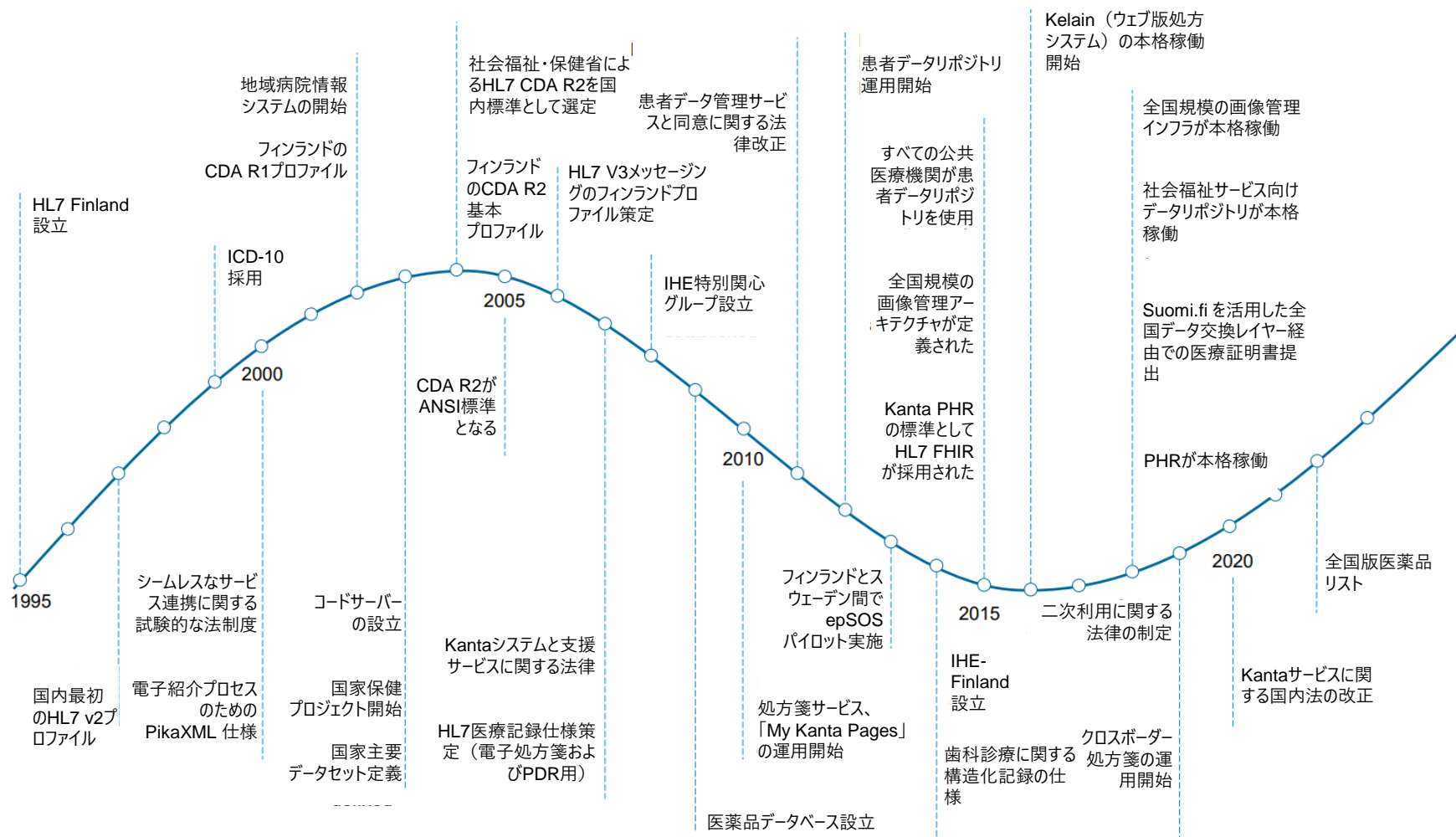
分類	タイプ	用途
DICOM	内容標準 データ転送標準	患者データのアーカイブと処方箋センターのメタデータおよびデータ内容はCDA R2形式で処理される。社会福祉データのアーカイブのメタデータおよび内容フレームワークもCDA R2形式で扱う。地域の画像処理ソリューションに対応。
HL7 CDA R2	内容標準	<ul style="list-style-type: none">• Kanta PDR(Patient Data Repository)や処方センター(Prescription centre)のデータを電子文書（CDA形式）として記録・交換• 社会福祉データも同様に管理する
HL7 FHIR	内容標準・サービスAPI標準	<ul style="list-style-type: none">• Kanta PHRや地域の患者情報システムとの連携• Web API を使ってデータをリアルタイムに交換
HL7 V2	メッセージ標準	<ul style="list-style-type: none">• IHEプロフィールに基づく画像診断や臨床検査のデータ送受信（HL7 Finlandの仕様）• 患者情報・検査結果をリアルタイムにメッセージ形式でやり取り
HL7 V3 Medical Records	メッセージ標準	<ul style="list-style-type: none">• 患者データアーカイブ• 社会福祉データアーカイブ• 処方センターなどとのインターフェース
IHE ITI-I	実施モデル・内容仕様・データ転送仕様・標準の適用ガイド	<ul style="list-style-type: none">• 画像データのアーカイブに関する技術仕様• 地域レベルの画像診断ソリューション• 欧州各国の相互運用性を構築するためのNCP(National Contact Point)のインターフェース
ISO IDMP	医薬品情報および薬事活動の標準	<ul style="list-style-type: none">• EU加盟国、製薬会社、行政機関が適用する国際的なISO IDMP標準は、薬事活動および医薬品関連情報の分類、検索、提示、リスク評価、電子交換、通知に使用される• 標準の運用は欧州委員会の実施規則によって規定され、EMA（欧州医薬品庁）は段階的に適用を進めている

出所： <https://yhteistyotilat.fi/wiki08/display/SYPLJULK/2.+Periaatetaso>

フィンランドでは、1995年から長年にわたり医療データの標準規格化を推進している

② インターフェース・システムアーキテクチャ等

フィンランドにおける標準規格の推進スケジュール



出所： <https://www.kanta.fi/documents/20143/106828/Kanta-arkkitehtuuri+ja+terveydenhuollon+yhteentoimivuuden+ja+IT-standardoinnin+aikajana+EN.pdf/6d93272e-b29b-1f86-423d-796ab8ee5e45>

EEA域外へのデータ移転または処理を行うには、GDPR第5章に定められた以下のいずれかの法的根拠を満たす必要がある

③ 多国間利用等

EU/EEA域外へのデータ移転について

第45条 十分性認定に基づく移転

- 十分性認定基準に基づき移転要件が満たされる場合、他の措置を講じる必要はなく、当局はこの基準をデータ許可審査における正当化要素として考慮する

第46条 適切な保護措置に従った移転 第2項

- EU/EEA地域外へのデータ移転は、第46条の条件を満たし、申請者または許可保持者がFindataに署名済みの標準契約条項を提出した場合に限り認められる。また、標準条項の使用に際しては、データ保護の十分性に関する追加の審査が必要となる。なお、標準契約条項は改変や追加を行わず、承認された原文のまま使用しなければならない

第47条 拘束的企業準則

- 拘束的企業準則（BCR）とは、企業グループ内または共同経済活動を行う企業グループ間において、個人データを第三国に移転する際に適用される共通かつ拘束力のある規則である
- この規則をデータ移転の根拠として用いる場合、当局はデータ許可審査において当該規則を考慮する。データをEU/EEA地域外へ移転するには、申請者または許可保持者が行動規範をFindataに通知し、その規範がEU加盟国のデータ保護監督官により正式に承認されていることが条件となる。

第49条 特定の状況における例外

- 例外的な状況においてデータ移転を認める条件および保護措置を定めるものである。たとえば、移転に伴うリスクについて説明を受けたうえでデータ主体が移転に明示的に同意した場合などが該当する
- この法的根拠を基に許可対象データを利用する場合、管理者または処理者は、移転に関する評価および適切な保護措置を記載した処理活動報告書を提出しなければならない

フィンランド参考資料

フィンランドでは、データの二次利用を効率的かつ安全に促進するとともに、EHDSへの円滑な統合を実現することを目指し、FinHITSプロジェクトを立ち上げた

【参考】FinHITSプロジェクトの概要

プロジェクト名	FinHITS（Finnish Health Data ICT Strengthening）
期間	2023年～2026年（4年間）
予算金額	<ul style="list-style-type: none">250万ユーロ(約3億9,975万円)EUより提供され、そのうち40%はFindataの自己資金で賄われている
目的	<ul style="list-style-type: none">フィンランドの健康データICTインフラを強化し、二次利用を効率的かつ安全に促進するEHDSへの円滑な統合を実現する
着目分野	<ul style="list-style-type: none">データアクセス申請管理システムの能力向上健康データの全国的なデータセットカタログの開発健康データの安全な処理環境「Kapseli」の強化HealthData@EUインフラストラクチャと接続するための国境を越えたゲートウェイの構築健康データの品質向上

※2025.02.15 13:30の為替レートにより計算 1ユーロ = 159.90円

FinHITSプロジェクトは9つのワークパッケージに分かれ、データ申請効率化、カタログの改善、国境を超えるデータ転送技術の開発とデータ品質向上等により二次利用を促進している

② インターフェース・システムアーキテクチャ等

【参考】FinHITSプロジェクトのワークパッケージ(WP)

	番号	WP名	概要
プロジェクト管理WP	1	管理と調整 Management and Coordination	プロジェクトの全体的な調整と実施を管理し、モニタリングや報告を行う
	2	普及活動、トレーニング、およびサポート Dissemination, training and support	プロジェクトの成果を普及活動とトレーニング等を通じて伝える。国際的なデータ利用者や利害関係者に対し、サービスデスクを開発
	3	評価 Evaluation	パフォーマンス指標を設定・モニタリングし、目標達成状況を評価
	4	サステナビリティ Sustainability	プロジェクト終了後もデジタル能力が維持されるよう、実装と保守を確実にを行う
中核WP	5	データアクセス申請管理システム（DAAMs） Data Access Applications Management System	Findataのe-サービスを全国的なデータセットカタログと統合する形で開発し、申請者が必要なデータや抽出内容を記述し、データセットを選択できるようにする。また、申請者と複数のデータ管理者間の効率的な連携を支援
	6	全国的な健康データセットカタログ National dataset catalogue for health data	国際的な利用者に焦点を当て、健康データの全国的なデータセットカタログを改善
	7	健康データの安全な処理環境 Secure Processing Environment for health data	計算能力とライセンスソフトウェアの使用を強化し、安全な処理環境を改善。結果の匿名化や匿名性の確認を容易にする自動化ツールを開発
	8	国境を越えたゲートウェイ Cross-border gateway	健康データの二次利用に向けた国境を越えたデータ転送の技術的ソリューションを開発。この際、HealthData@EU Pilotの成果を活用
	9	健康データの品質向上 Health data quality enhancement	健康データの品質向上を図るためのソリューションや指針を提供。2つの既製データセットを作成し、必要に応じてデータ記述に関する全国的なガイドラインや指針を更新

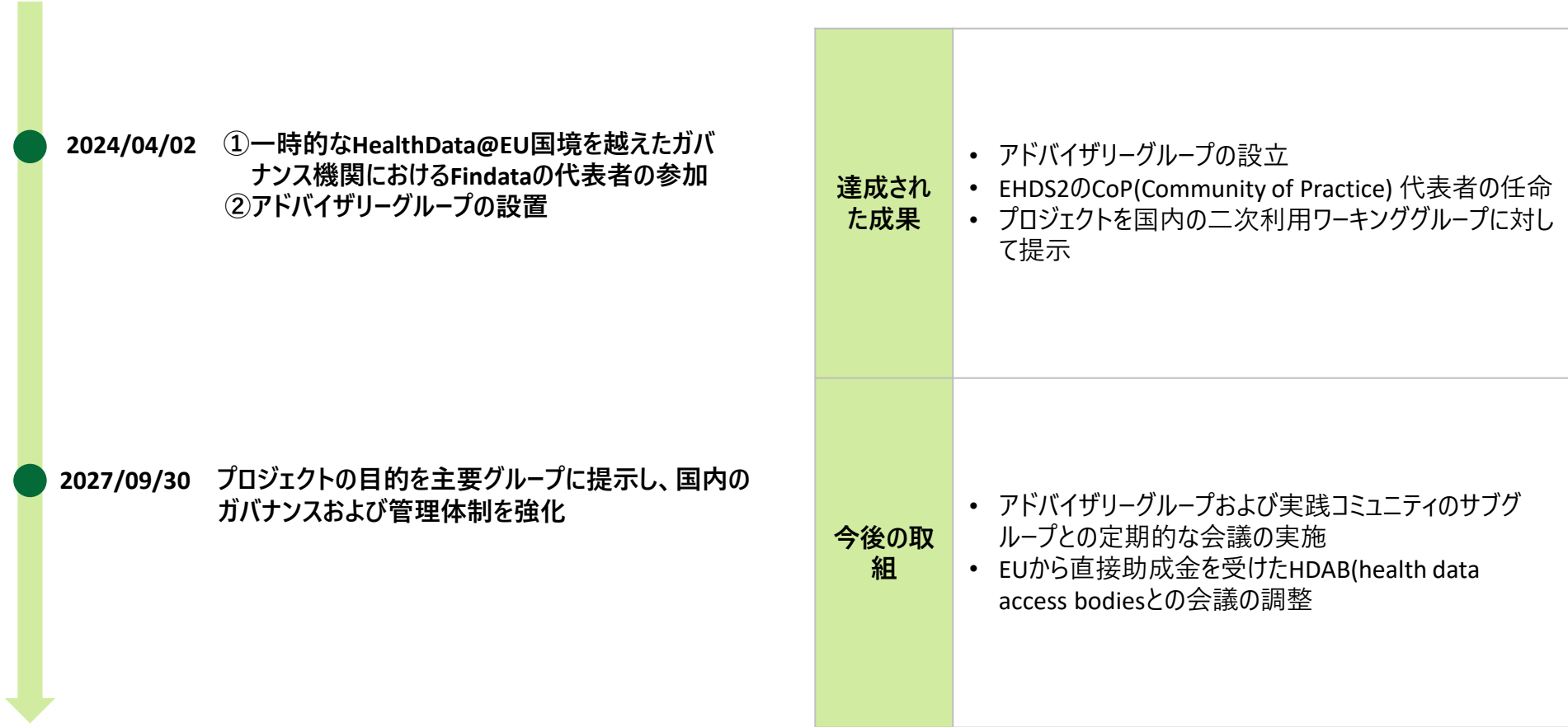
出所： <https://findata.fi/en/finhits/work-packages/>

FinHITSのWP1では、プロジェクトの全体的な調整と実施を管理し、モニタリングや報告を行う

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP1: 管理と調整

WP名	WP1:管理と調整(Management and Coordination)
目的	プロジェクト目標を達成するための効果的なプロジェクト調整と協力の確保



出所： <https://findata.fi/en/finhits/results/>

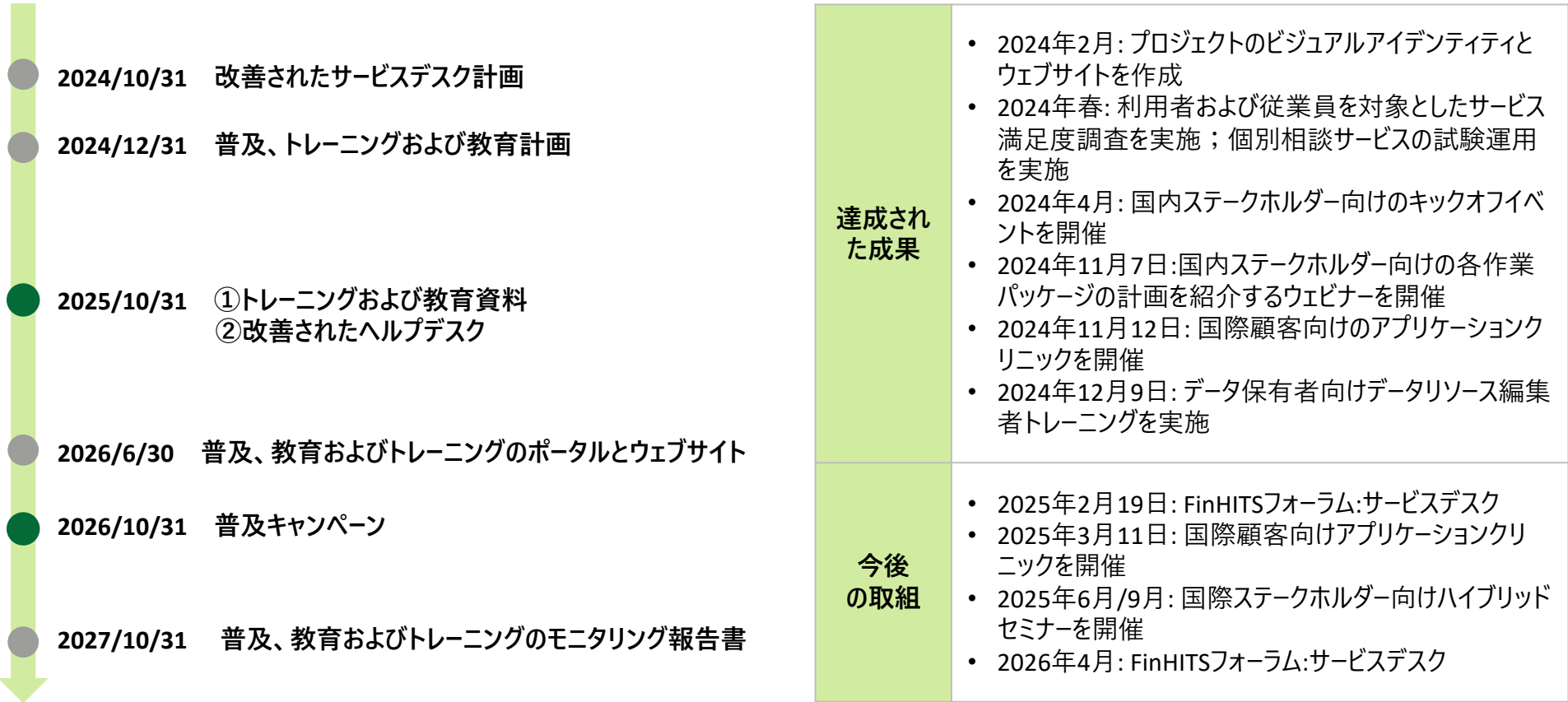
● マイルストーン ● 成果物

FinHITSのWP2では成果を普及活動とトレーニング等を通じて伝える。国際的なデータ利用者や利害関係者に対し、サービスデスクを開発する

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP2:普及活動、トレーニング、およびサポート

WP名	WP2:普及活動、トレーニング、およびサポート (Dissemination, training and support)
目的	ステークホルダーへの認知向上、トレーニング提供、HDAB、データ保有者、データ利用者の支援を行い、サービスデスクを改善する



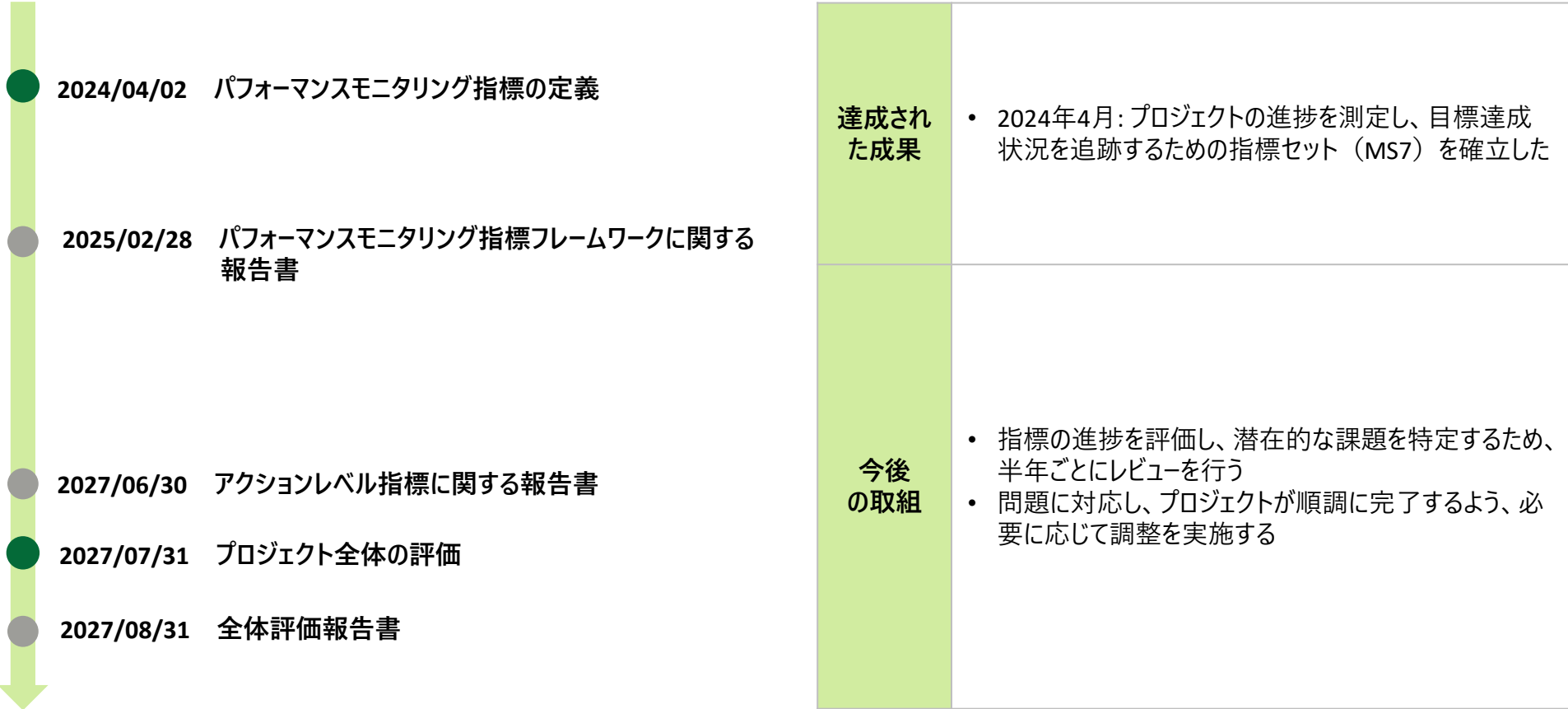
出所： <https://findata.fi/en/finhits/results/>

FinHITSのWP3では、パフォーマンス指標を設定・モニタリングし、目標達成状況を評価する

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP3:評価

WP名	WP3:評価(Evaluation)
目的	プロジェクトの進捗を追跡し、活動が計画通りに実施されていることを確実にするため、パフォーマンスをモニタリングする



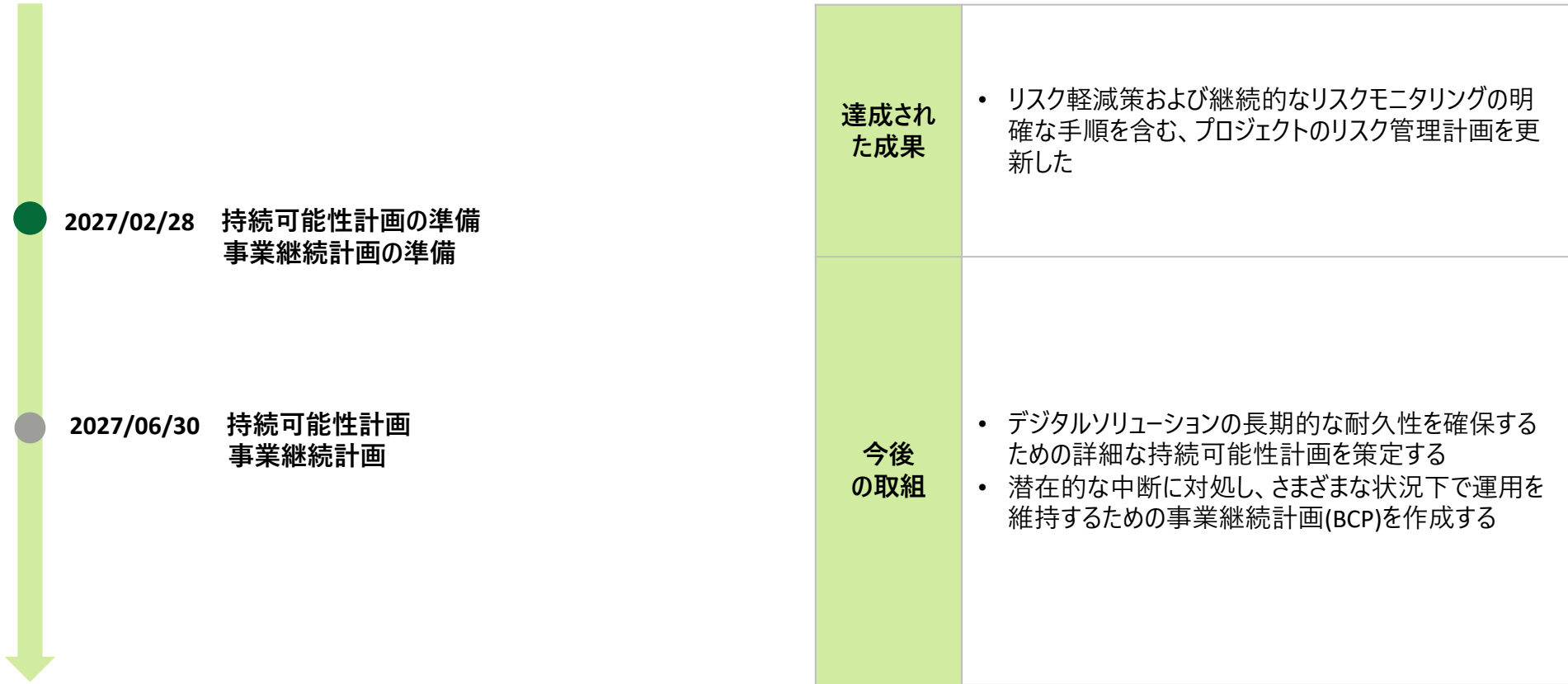
出所： <https://findata.fi/en/finhits/results/>

FinHITSのWP4では、プロジェクト終了後もデジタル能力が維持されるよう、実装と保守を確実に行う

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP4:持続可能性

WP名	WP4:持続可能性 (Sustainability)
目的	長期的なデジタルビジネス能力の確保



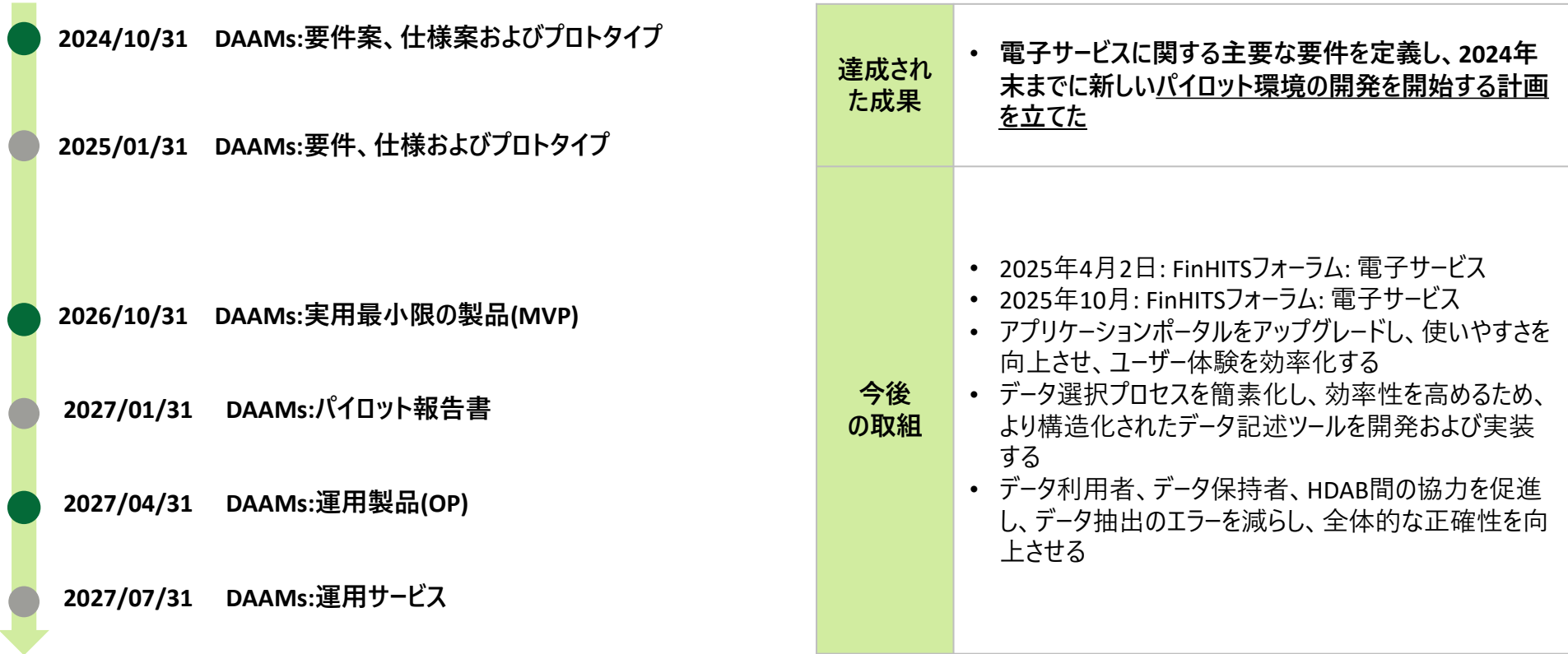
出所： <https://findata.fi/en/finhits/results/>

FinHITSのWP5では、データアクセスアプリケーション管理システム(DAAMs)のアプリケーションポータルをより使いやすく効率的なものにする

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP5:データアクセス申請管理システム(DAAMs)

WP名	WP5:データアクセス申請管理システム(DAAMs) (Data Access Applications Management System)
目的	データアクセスアプリケーション管理システム（DAAMs）の実務経験、ステークホルダーのニーズ、およびEHDSによる今後の要件に基づき改良を進め、アプリケーションポータルをより使いやすく効率的なものにする



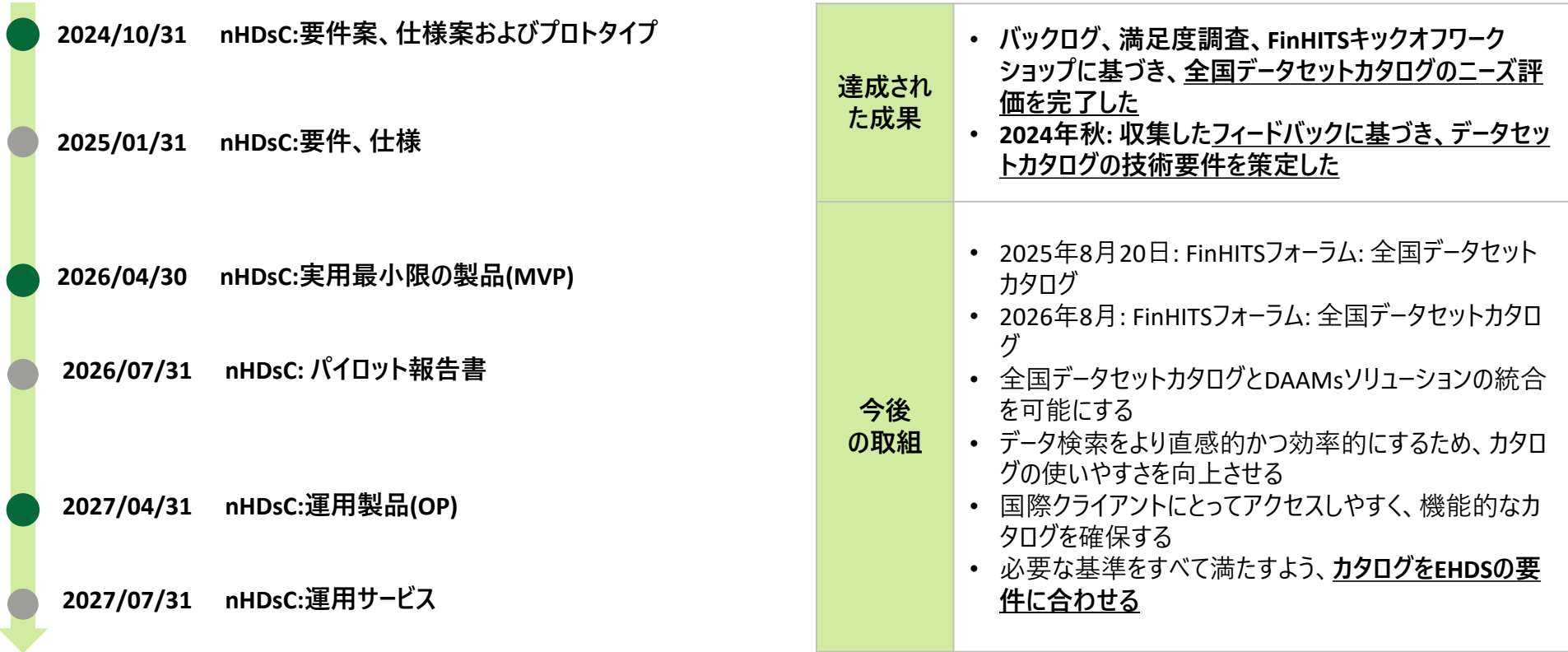
出所： <https://findata.fi/en/finhits/results/>

FinHITSのWP6 では、国際的な利用者に焦点を当て、ヘルスデータの全国的なデータセットカタログを改善する

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP6:全国的なヘルスデータセットカタログ

WP名	WP6:全国的なヘルスデータセットカタログ(National dataset catalogue for health data)
目的	<ul style="list-style-type: none">全国ヘルスデータセットカタログ（nHDsC）の改良に取り組み、国内外の利用者にとってより使いやすいものとする 것을 目指しているEHDSが定める基準を満たすため、フィンランド健康福祉研究所（THL）と協力している



出所： <https://findata.fi/en/finhits/results/>

FinHITSのWP7では、計算能力とライセンスソフトウェアの使用を強化し、安全な処理環境を改善する

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP7:ヘルスデータの安全な処理環境(Kapseli)

WP名	WP7:ヘルスデータの安全な処理環境(Kapseli)(Secure Processing Environment for health data)
目的	<ul style="list-style-type: none">セキュア処理環境（SPE）であるKapseli®の機能拡張に注力しているヘルスデータ処理や複雑なデータプロジェクトの取り扱いをより効果的にサポートするため、その容量、使いやすさ、およびセキュリティを強化する



達成された成果	<ul style="list-style-type: none">ライセンス管理ツールの開発を開始したスポットチェックツールの開発を開始した2024年6月6日: <u>フェデレーテッド分析に関するステークホルダー会議を開催した</u>
今後の取組	<ul style="list-style-type: none">2025年春: フェデレーテッド分析に関するステークホルダー向けワークショップを実施2025年6月4日: FinHITSフォーラム: SPE2026年6月: FinHITSフォーラム: SPE大規模データセットを処理可能とするため、データ転送サービスを拡張するGPU機能を追加し、計算能力を向上させるフェデレーテッド分析のプロセスを開発し、複数のデータセット間でデータを安全に統合するライセンス管理を強化し、ユーザー体験を改善する

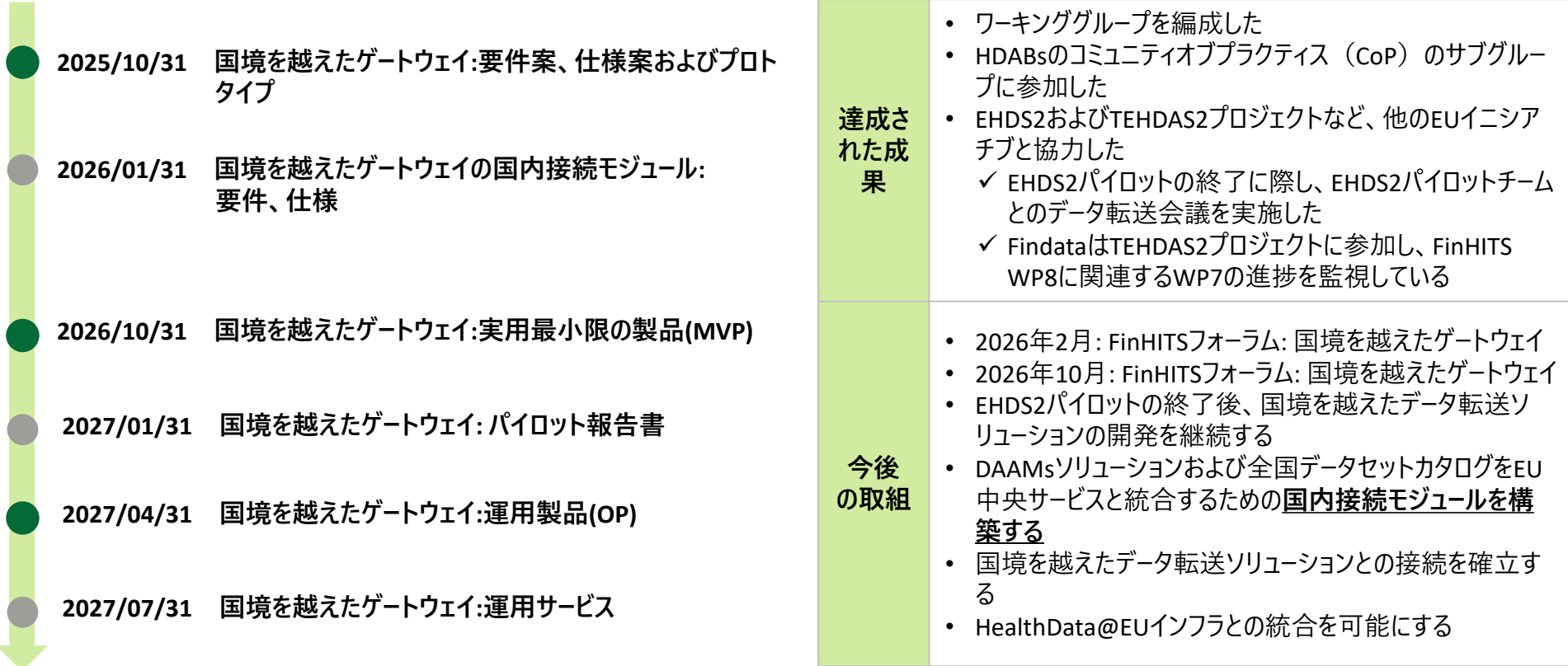
出所： <https://findata.fi/en/finhits/results/>

FinHITSのWP8では、ヘルスデータの二次利用に向けた国境を越えたデータ転送の技術的ソリューションを開発する

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP8:国境を越えたゲートウェイ

WP名	WP8:国境を越えたゲートウェイ(Cross-border gateway)
目的	・ EHDS規制に従い、国境を越えたデータ転送ソリューションへの接続を確立し、HealthData@EUインフラとの統合を進める



出所：<https://findata.fi/en/finhits/results/>

FinHITSのWP9では、ヘルスデータの品質向上を図るためのソリューションや指針を提供する

② インターフェース・システムアーキテクチャ等

【参考】FinHITS WP9:ヘルスデータの品質向上

WP名	WP9:ヘルスデータの品質向上 (Health data quality enhancement)
目的	・ ツールの開発、支援の提供、ガイドラインの作成、既製データセットの作成を通じてデータ品質の向上に取り組んでいる



出所 : <https://findata.fi/en/finhits/results/>

● マイルストーン ● 成果物

5.2 技術動向調査

医学系データベースを活用した諸外国の最新研究の調査報告を収集・分析し、日本の情報連携基盤構築に必要な要素と方針を整理する

技術動向調査の方針

調査内容

- ① 公的・民間を問わず、医学系データベースを活用した諸外国の最新研究や動向を取りまとめた調査報告書を収集
- ② 調査報告書に掲載された研究事例を深掘りし、詳細を確認・整理
(事例名、取組主体、目的、事例概要、活用データ内容、情報の流れ、インターフェイス、実用化に向けた取組など)
- ③ 収集・整理した情報をもとに、日本の情報連携基盤構築に向けた必要な要素や検討すべき方針を整理

(A) RWDの活用

例) リアルワールドデータ (RWD) を活用した薬事承認・ドラッグリポジショニング研究

(B) AI・生成AI

例) AI、リアルタイム分析、生成AI (LLM: Large Language Model /LMM: Large Multimodal Model) を活用した診断補助や予測モデル

(C) 最新研究

例) リバートランスレーショナルアプローチの研究や連合学習等の新たなデータフローを活用した研究、リアルタイム解析

左記項目の(A)～(C)に関し、事例を収集し以下の項目をまとめる：

- ・ 事例名
- ・ 取組主体
- ・ 目的
- ・ 事例概要
- ・ 活用データ
- ・ 実用化段階
- ・ 情報の利活用フローや技術的インターフェイス ...等

医学研究の
最新動向を
踏まえた示唆
・
報告書作成

信頼できる政府機関、研究機関等に掲載されている先進事例を抽出し、それらに基づきデ スクトップ調査で詳細を整理した

活用検討中の調査報告書・文献等

発行国	発行元	出所	発行年	概要	事例番号	RWD	AI	最新研究
日本	CRDS 国立研究開発法人科学技術振興機構研究開発戦略センター	健康・医療トランスフォーメーション 科学技術・イノベーションの潮流	2023	リアルワールド研究、創薬、大規模言語モデルの活用、分散臨床試験など、諸外国における最新の研究事例を幅広く取り上げ、最新動向を報告書としてまとめたもの	事例①：米国施策	○		
					事例⑧：米国事例		○	
					事例⑨：米国事例		○	
		健康・医療リアルワールドデータ利活用基盤の構築と生成 AI への展開	2023	RWDの収集・活用および医療版LLM・基盤モデルの開発に焦点を当て諸外国の事例を取上げている	事例⑤：ドイツ事例	○		
		研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野	2023	ライフサイエンス・臨床医学分野における、グローバルな社会的課題、経済的課題を整理し、主要国の科学技術政策動向や先端の研究開発潮流をまとめたもの	事例⑪：EU事例			○
英国	Medicines and Healthcare Products Regulatory Agency	MHRA guidance on the use of real-world data in clinical studies to support regulatory decisions	－	英国におけるRWDを活用した医薬品・医療製品の規制判断に関するガイドライン。現在、パイロットプロジェクトが進行中	事例②：英国施策	○		
	NHS	EHS England	－	NHS Englandのデータ利用に関するHP	事例⑫：英国事例			○
		NHS DigiTrials	－	NHS HP上で DigiTrials を利用した研究事例を公開	事例③：英国事例	○		
		NHS AI Lab roadmap	－	AI技術の開発・検証・導入を促進するため、AI 技術の安全で倫理的な開発・導入のための規制の整備やデータの適正な活用を推進しておりそのための様々な取組がまとめられている	事例⑥：英国事例		○	
	Department of Health and Social Care	Foresight—a generative pretrained transformer for modelling of patient timelines using electronic health records: a retrospective modelling study	2024	NHSが保有する医療情報を活用し、OpenGPTモデルを深層学習させ医療分野向けのLLMモデルおよびそれを基盤とした対話型アプリケーションを開発した経緯を報告した個別研究事例。調査・開発は現在も進められている	事例⑦：英国事例		○	
		Real-time Assessment of Community Transmission (REACT) Programme	2020	NHSの患者データを活用した世界最大規模のCOVID-19観察研究。感染モニタリングから長期的健康影響の解明へ発展し、リアルタイム解析やパンデミック対応の事例として参考にする	事例⑩：英国事例			○
EU	EHDEN	EHDEN	2022	EHDENが日常的な医療データを活用し、有害事象の検証や優先順位付けを実施する	事例④：EU施策	○		
米国	FDA	BenevolentAI社	2020	COVID-19に対する効果的な治療薬を既存の承認薬の中から、見つけることを目指した	事例⑬：米国事例		○	○
	NIH	アリゾナ大学	2023	乳がん治療薬(EMT)がアルツハイマー病のリスクに与える影響を明らかにし、そのメカニズムを解明することを目指した	事例⑭：米国事例			○

米国FDAは医療製品に関する規制判断において、従来の厳密な臨床試験データだけでなく、RWDの活用も認める方針を示し、実証プロジェクトや詳細な規定策定を進めている

諸外国の先進事例【事例①: RWD】(米国・施策)

事例名	Framework for FDA’s Real-World Evidence Program
取組主体	U.S. Food & Drug Administration (FDA)
国	米国
取組時期	2018年 ～
目的	医薬品の承認・評価の加速化
事例概要	<ul style="list-style-type: none">2016年に医薬品や医療製品の開発を加速化を目的として制定された21st Century Cures Actを踏まえ、<u>FDAは、従来の厳密な臨床試験データのほか、日常的に医療現場で収集されるRWDを活用した規制上の判断・追跡研究に関する枠組みを策定した</u>本取組により、医薬品等に関する承認・評価のプロセスにおいて、ハイブリッド試験やプラグマティック試験、観察研究など、従来より幅広い試験デザインが認められるようになる
対象データ	<ul style="list-style-type: none">電子カルテ (EHR) から得られるデータ請求データ疾患レジストリ等健康情報を取得するモバイルデバイス等から得られるデータ ...等
出所	https://www.fda.gov/media/120060/download?attachment

実用化段階

RWEに関するガイダンスの整備、承認プロセスの見直し、実証プロジェクトへの助成金提供などの取組は進行中であり、今後の評価を経て、RWEを活用した詳細な規制決定がさらに進む見込みである

<タイムライン>

- 2020年:** RWDを用いてRWEを生成し、規制決定に貢献する実証プロジェクトへ助成金提供開始
- 2021年:** FDAがRWE活用に関するガイダンス文書の発行・公開を開始 (今後さらに増やす予定)
- 2023年:** 新たに4プロジェクトに助成金を提供し合計 10プロジェクト進行中
- 2024年:** 米国の医薬品利用者負担金法 (PDUFA) において、RWEの受容性拡大を図る方針を表明

英国MHRAにおいても、医薬品・医療製品の規制判断におけるRWD活用を推進する方針を示した上でガイドラインを策定し、2025年よりパイロットプロジェクトを実施する

諸外国の先進事例【事例②: RWD】(英国・施策)

事例名	MHRA guidance on the use of real-world data in clinical studies to support regulatory decisions	対象者	RWDを活用した臨床試験 (観察試験・ランダム化比較試験) を実施する研究機関・民間企業など
取組主体	Medicines and Healthcare Products Regulatory Agency (MHRA) - 医薬品・医療製品規制庁	対象データ	臨床試験外で収集された患者の健康状態や治療に関するデータ (例：電子カルテ、疾病レジストリ・DB、ウェアラブルデバイス等など)
国	英国		
取組時期	2020年～		
目的	<ul style="list-style-type: none">医薬品・医療製品の承認・評価プロセスの加速化およびコスト削減臨床試験よりも実際の治療環境を反映したデータを反映したエビデンス創出臨床試験の実施における地理的な制約を排除する	実用化段階	<ul style="list-style-type: none">2025年2月より “RWE Scientific Dialogue Programme” の名のもとでパイロットプロジェクトを行う研究機関・民間企業の募集を開始本プロジェクトでは、医薬品の規制判断、有効性・安全性評価、プラグマティック研究、外部対照群を用いた介入・非介入研究などを実施し、RWEのエビデンス基準や評価基準の検討に活用する
事例概要	<ul style="list-style-type: none"><u>米国FDAのReal World Evidence Program、EU EMAのBig Data Taskforceなどの動向を考慮し、英国もRWDを医薬品・医療製品の規制判断において活用する際のガイドラインを策定した</u>RWDの活用においてデータ品質 (収集方法、正確性・網羅性、管理方法など) が厳密に評価される指針が示されており、研究プロトコルに明記する必要があるガイドラインは一般的な原則を示すものであり、個別な研究デザインに関する指導はMHRAの Scientific Advice Meeting を踏まえて提供される	事例の特徴	<ul style="list-style-type: none"><u>RWDを活用した臨床試験は、適切に設計され、質の高いデータが確保されていれば、従来の臨床試験と同等のエビデンスを有するとしている</u>RWDの品質が不十分であれば信頼性のある判断ができないため、<u>データの品質についてMHRAで厳密に評価する</u>
出所	https://www.gov.uk/government/publications/mhra-guidance-on-the-use-of-real-world-data-in-clinical-studies-to-support-regulatory-decisions		

Heartburn Health Programme は、NHSのDigiTrialsサービスを活用して被験者を募集し、承諾を得た参加者のRWDをNHS経由で収集し、研究に利用する

諸外国の先進事例【事例③: RWD】(英国・事例)

事例名	Heartburn Health Programme
取組主体	<div><div>• ケンブリッジ大学 & ケンブリッジ大学病院</div><div>• Cancer Research UK Cancer Prevention Trials Unit</div></div>
国	英国
取組時期	2024年～
目的	胸やけ・逆流性食道炎などの消化器症状の管理・治療やがんの早期発見に関する研究の推進
事例概要	<div><div>• <u>NHS EnglandのDigiTrialsサービスを活用し、研究における被験者のリクルーティングとRWDの収集を実施</u></div><div>• NHSが研究対象となりうる18歳以上のランダムサンプルを選定し、SMSで招待状を送信。少なくとも12万人の被験者を募集予定</div><div>• SMSに掲載の登録フォームに記入することで治療に参加でき、NHSから消化器症状に関するRWDが最低12年間収集・SDEに格納される</div><div>• データは Heartburn Health Programme の複数研究に使用される</div></div>
提供情報	<div><div>• NHSが保有する消化器関連の健康情報</div><div>• 特定の研究で追加データが必要な場合は、被験者に対しアンケート、血液サンプル、検査などの案内が行われる</div></div>
出所	https://www.heartburnhealth.org/about-the-programme/how-the-programme-works/

実用化段階	被験者募集の段階
概要図	
事例の特徴	<div><div>• 被験者のリクルーティングや追跡などには個人情報を使用されるが、研究には非識別化されたRWDが用いられ、SDE上で厳重に管理される</div><div>• 患者の参加は自由で、メールの受信拒否や途中脱退も可能</div><div>• 参加者には研究の最新情報や生活習慣に関するアドバイスが提供されるなどのメリットもある</div></div>

EHDENは欧州におけるリアルワールドデータを用いたネットワークで、多くの国の様々な関係者が参画し、医療データのOMOP化や分散型研究を行っている

諸外国の先進事例【事例④: RWD】(EU・施策)

名称	EHDEN (European Health Data and Evidence Network)
設立年	2018年11月
設立背景	欧州委員会が代表を務める欧州連合(EU)と欧州製薬団体連合会(EFPIA)の間の官民パートナーシップであるIMI (Innovative Medicines Initiative)の一環として設立された
目的	欧州の医療データをフェデレーテッド(分散型)の研究ネットワークに統合して、患者の健康や医療の質の向上に貢献することである
活動内容	<ul style="list-style-type: none">OMOPへの変換や管理等医療データの標準化医療データの所有者や管理者が自主的に参加する分散型研究ネットワークによる医療データの共有と分析
データパートナー	29か国の187のデータパートナー
中小企業との連携	22か国の64か国のパートナーと協力してデータのOMOP化の変換を支援している
データの範囲と規模	500万人の患者データを含む医療記録
ユースケース例	<ul style="list-style-type: none">エストニアの健康データを観察医療アウトカムパートナーシップ(OMOP)共通データモデルに変換する日常的に収集される健康データの分析による薬品の安全性評価と優先度の決定(EHDENネットワーク研究から学んだ教訓)ワクチンに対するAESIのバックグラウンド率医療技術評価 (HTA) のユースケース: 慢性閉塞性肺疾患 (COPD)

ドイツBfArMはDiGA Fast-Track Processを創設することで、デジタルヘルスアプリの迅速な導入、RWDの収集とRWDを基にしたエビデンスの創出により動的な医療技術評価を図る

諸外国の先進事例【事例⑤: RWD】(ドイツ・事例)

事例名	The Fast-Track Process for Digital Health Applications (DiGA※)		
取組主体	BfArM(ドイツ連邦医薬品医療機器庁)		
国	ドイツ		
取組時期	<ul style="list-style-type: none">2019年DVG (Digitale-Versorgung-Gesetz)法制定2020年運用開始～現行運用中		
目的	<p>デジタルヘルスアプリの迅速な導入とエビデンス創出の両立</p> <ul style="list-style-type: none">安全性・品質等の要件を満たす有望なデジタルヘルスアプリを迅速に保険適用し、患者が早期に利用できるようにする保険償還下でRWDを収集して有効性を検証し、エビデンスに基づき恒久的な承認可否と適正な価格設定につなげる	対象データ	<ul style="list-style-type: none">収集データ<ul style="list-style-type: none">✓ 患者のアプリ利用データ（アプリの使用頻度、入力された症状スコアや日記データなど）✓ 患者報告アウトカム（PRO）✓ 電子カルテ（EHR）✓ 健康保険請求データ✓ 生体データRWDの利用<ul style="list-style-type: none">✓ DiGA申請時のエビデンス文書の一部として提出される✓ 市販後調査や臨床フォローアップ等にも活用される
		出所	<ul style="list-style-type: none">https://www.bfarm.de/SharedDocs/Downloads/EN/MedicalDevices/DiGA_Guide.htmlhttps://pubmed.ncbi.nlm.nih.gov/36059268/
事例概要	<ul style="list-style-type: none">2019年のDVG法により、医療目的のデジタルアプリを試行的に保険収載するDiGA Fast-Track制度が創設されたメーカーがBfArMに申請し、安全性・機能性・データ保護の要件を満たせば、<u>有効性のエビデンスが不十分でも最大12か月の償還リストへの試行的収載が認められる</u>。この期間中、アプリが医師による処方が可能となり、公的医療保険の償還対象となるメーカーは<u>実際の患者利用データ(RWD)を収集し</u>、期間終了時までに臨床評価結果をBfArMに提出し、審査の結果、<u>有効性が確認されれば正式収載</u>(恒久的保険償還)となり、その後保険者と提供企業間での償還額設定交渉が行われる。一方、エビデンスが不十分な場合はDiGAリストから削除されるこの制度により、デジタルヘルスアプリを迅速に導入し、<u>RWDを基にした動的な医療技術評価(HTA)が可能となった</u>	DiGA Fast-Track におけるRWD活用の実例	<p>【糖尿病管理アプリ「Esysta」】</p> <ul style="list-style-type: none">インスリン治療中の糖尿病患者向けに処方可能なDiGAとして、自動インスリン記録機能とデータ共有プラットフォームを備えている。患者の血糖値・インスリン投与量をBluetooth対応デバイスからリアルタイム収集し、医療者と患者双方が閲覧できる電子糖尿病日誌を生成するメーカーはRWDを活用し、健康保険組合およびドレスデン工科大学との協同研究を実施した。その結果、200名超の実患者データの解析により平均HbA1c値が最大2%低下し、低血糖リスクの増加なく血糖コントロールが有意に改善することが示されてるこのエビデンスは日常臨床での有効性を裏付けるRWDとして評価され、BfArMの審査で血糖コントロールの医学的ベネフィットを有するDiGAと認められ、正式収載された

※DiGA(Digitale Gesundheitsanwendungen): ドイツのデジタルヘルスアプリケーションの精度であり、医療保険でカバーされるデジタルヘルスアプリを指す。
医師が処方することで患者がDiGAを無料で利用できる。

NHS AI Lab では、AI技術の開発・検証・導入を促進するため、資金提供、AIトレーニング用SDEの整備、画像データベースの整備、規制強化など多方面での取組を進めている

諸外国の先進事例【事例⑥: AI】(英国・施策)

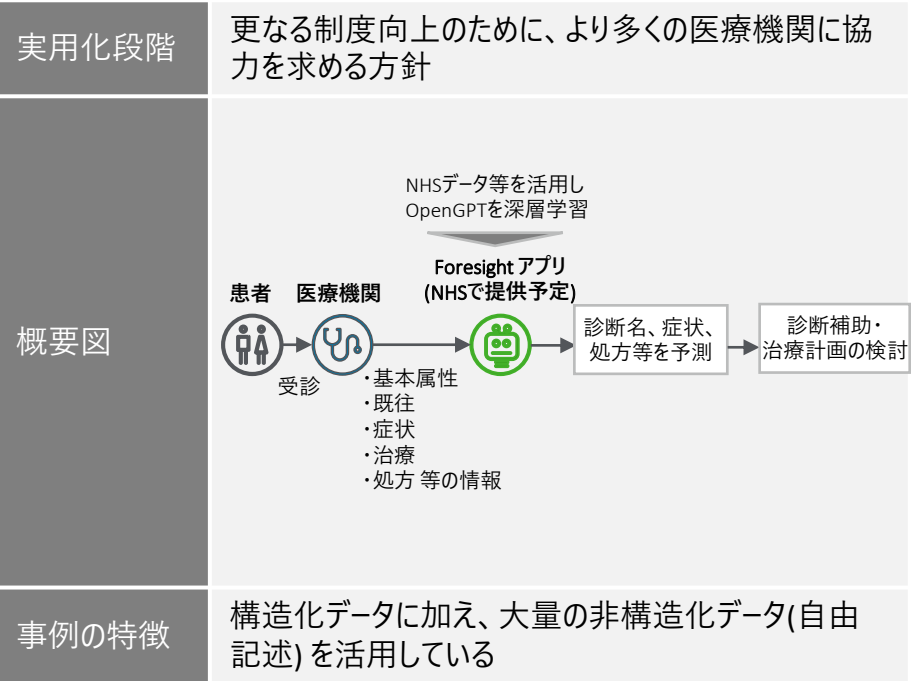
事例名	NHS AI Lab
取組主体	NHS
国	英国
取組時期	2019 年～継続中
目的	AI 技術の安全で倫理的な開発・導入のための規制の整備やデータの適正な活用を推進する
事例概要	<p>NHS はAI技術の開発・検証・導入を促進するため、多様な取組を進めている</p> <ul style="list-style-type: none">• 経済的支援として、1億1300万ポンド (約215億円) を投資し、86のAI開発プロジェクトの支援を実施。さらに、21億ポンドのAI Diagnostic Fundを設立し、NHSの医療機関におけるAI診断ツールの導入を推進• データ利活用推進のため、英国医療研究機関(HRA) がAIトレーニングに適した Secure Data Environment (SDE) の整備を進め、AI開発企業にもデータアクセスを許可する方針。また、画像データベースの構築などを通じて、医療AIの活用を支援• その他、AI規制やガイドラインの策定、AI認証プロセスの見直し、Algorithmic Impact Assessment(AIA)の試験運用を実施による医療におけるAI技術の影響評価などを進めている
出所	https://transform.england.nhs.uk/ai-lab/nhs-ai-lab-roadmap/

実用化段階	NHS AI Labは今後、医療機関へのAI技術の導入を促進する “AI Deployment Platform” (AIDP) の試験運用を開始する予定。AIDPの導入により、規制に準拠したAIツールのみが掲載され、安全な技術をNHS内で普及させることが可能
事例の特徴	<ul style="list-style-type: none">• NHSは、AI技術の開発促進のため多額の予算を確保し、資金提供やデータ利活用体制の整備を積極的に推進している• 画像データベースの構築を推進• 技術の承認後に専用プラットフォームへ掲載することで、技術の監視と普及促進を図っている

キングス・カレッジ・ロンドンは、NHSが保有する医療情報をOpenGPTに学習させ、医療分野向けのNHS-LLMを基盤とした対話型アプリケーションを開発している

諸外国の先進事例【事例⑦: AI】(英国・事例)

事例名	Foresight (NHS情報を活用した疾病予測モデル)
取組主体	King's College London King's College Hospital NHS Foundation Trust など
国	英国
取組時期	2022年～
目的	医療用AIモデルの開発による疾病予測・リスク評価を通じた医療の質の向上
事例概要	<ul style="list-style-type: none">NHSが保有する医療情報を活用し、OpenGPTモデルを深層学習させ医療分野向けのLLMモデルを構築し、この技術を基盤とした対話型アプリケーションを開発研究では、該当技術の予測精度を検証するため、2つのNHSトラストのEHRデータを用いて患者の疾患進行を予測し、アプリの精度は68～88%と評価された。さらに、医師の評価を加えた結果、約93%の正確性が確認された
対象データ	NHSの診療ガイドラインに加え、Foresightの学習・検証には以下の3データセットが活用された <ul style="list-style-type: none">King's College 病院 NHS財団 (1999年～2021年間の約145万人の構造化データ・自由記述データ)South London & Maudsley NHS財団 (2019年8月以降に重度精神疾患と診断された約150万人の構造化・自由記述データ)米国MIMIC-III 公開データ(マサチューセッツ工科大学が開発した約46万人のICU入院患者のデータ)
出所	https://pmc.ncbi.nlm.nih.gov/articles/PMC11220626/



Unlearn.AI社はアルツハイマー病および軽度認知障害の予測モデルを作成するために、患者レベルのデータを持つ大規模な遺伝子データベースを使用し、デジタルツインを作成した

諸外国の先進事例【事例⑧: AI】(米国・研究)

事例名	Modeling Disease Progression in Mild Cognitive Impairment and Alzheimer’s Disease with Digital Twins
取組主体	Unlearn.AI, Inc.
国	米国
取組時期	2020年 (本論文の発行日)
目的	アルツハイマー病（AD）および軽度認知障害（MCI）の患者の病気の進行をデジタルツイン*を用いてモデル化し、臨床試験の設計や外部対照群**として使用すること
事例概要	<ul style="list-style-type: none">条件付き制限ボルツマンマシン（CRBMs）***を使用して、<u>AD患者のデジタルツインを作成し、臨床試験において、標準的なAD治療を受けた患者の状態を予測するためのモデルを作成した</u>作成されたデジタルツインが実際の患者のデータと一致するかを評価した。具体的には、ADとMCIの進行度合いを評価する変数（AD評価尺度などの認知機能テストの各項目）が評価に使用された結果として、デジタルツインと実際の患者の特性が一致しており、臨床試験の設計や外部対照群としてデジタルツインが使用可能であることが判明した

* Unlearn.AI社のデジタルツインは、患者レベルのデータでトレーニングされており、個人の健康状態が時間の経過とともにどのように変化するかを予測する

出所： <https://www.unlearn.ai/digital-twin-generators>、<https://c-path.org/tools-platforms/critical-path-for-alzheimers-disease-cpad-database/>、<https://adni.loni.usc.edu/>

出所	https://arxiv.org/pdf/2012.13455
対象データ	<ul style="list-style-type: none">ADの観察研究および臨床試験のプラセボ群からの大規模な記録セット<ul style="list-style-type: none">➢ Critical Path for Alzheimer’s Disease (CPAD) は非営利団体であるC-Pathによって運営されており、ADおよびMCIの<u>12,811人の患者からの患者レベルのデータ</u>が入っている。データの種類としては<u>人口統計情報、APOE4遺伝子型、併用薬、認知尺度</u>である➢ Alzheimer’s Disease Neuroimaging Initiative (ADNI) データベースは官民パートナーシップを組んでいるADNIが運営を行っている。データセットは、縦断的な臨床、イメージング、遺伝、およびその他のバイオマーカーデータが集まっており、これには、構造的・機能的・分子的な<u>脳イメージング、生体流体バイオマーカー、認知評価、遺伝データ、人口統計情報</u>など、さまざまなデータタイプがっている
実用化段階	記載なし

**臨床試験において新しい治療法の効果を評価するために使用される、既存の治療法やプラセボと比較するためのグループ

***アルツハイマー病（AD）の進行をモデル化するために使用される生成モデルの一種。このモデルは、観察データや臨床試験のプラセボ群から得られたデータを用いて訓練された。CRBMは、患者の健康データの変化を時系列データとして扱い、これを基に将来の健康状態を予測する

Medidata社は700万件超の匿名化患者データと高度なマッチング技術を駆使し、合成対照群(SCA)を作成することで対照群の症例数を削減し、治験の迅速化・効率化を図る

諸外国の先進事例【事例⑨: AI】(米国・事例)

事例名	再発膠芽腫の第3相臨床試験における合成対照群 (Synthetic Control Arms/SCA※)の利用
取組主体	Medidata Solutions, Inc.(米) Medicenna Therapeutics Corp.(カナダのバイオ創薬企業)
国	米国、カナダ
取組時期	2020年
目的	標準治療が確立していない重篤疾患でランダム化対照試験(RCT)の設定が非倫理的または困難な状況において、 <u>過去データ由来の合成対照群(SCA)で補完することで、対照群症例数の削減によって治験を迅速化・効率化する</u>
事例概要	<ul style="list-style-type: none"> Medicenna社は、再発膠芽腫(rGBM)を対象とした第2相単群試験において、治験薬MDNA55の有効性を評価するため、<u>Medidata社が保有する過去の臨床試験データおよびRWDにより構築された合成対照群(SCA)を利用した</u> <u>FDAは、Medicennaによる第3相rGBM試験へのSCA導入方針に合意し、従来の1:1ランダム化対照を用いてきた適応疾患領域において、SCA由来の患者とランダム化患者を組み合わせたハイブリッド外部対照を第3相試験で採用することを初めて認めた</u> 結果として、対照群として新たに割り付ける必要のある患者数を約100人削減でき、被験者の実験薬受療機会が高まるとともに、登録や追跡の効率化、さらに開発期間の短縮が期待されている

SCAの構築方法	<ul style="list-style-type: none"> データソース： <ul style="list-style-type: none"> ➢ Medidata社は、23,000本以上の過去臨床試験と700万例以上の患者データを保有し、これらの<u>過去臨床試験データとRWDを統合することで、Integrated Evidence</u>を提供 データ抽出・処理： <ul style="list-style-type: none"> ➢ <u>人工知能（AI）を活用してデータベースを検索し、特定の疾患の治療に関する過去の試験で対照群として使用された患者を見つけ、標準化、統合、クリーニング、および非識別化を経て過去のデータソースからデータを取得する</u> データマッチング： <ul style="list-style-type: none"> ➢ <u>患者レベルのデータマッチングや傾向スコア（propensity score）等のバイオスタティクス手法を利用して、現在試験の実験群と過去の試験データで組成された外部対照群をマッチングする</u>
出所	https://www.medidata.com/wp-content/uploads/2021/04/Medidata-Ebook_jpn.pdf https://www.medidata.com/wp-content/uploads/2021/09/SCA-Whitepaper.pdf

※合成対照群(Synthetic Control Arm, SCA):ランダム化されていない臨床試験において、外部の患者データ(外部臨床試験データ、RWD等)を統計的に選択・調整することで、対照群として機能させる外部対照群の一種である。特に、ランダム化対照試験(RCT)の実施が困難な場合に、試験結果の解釈を向上させる手段として活用される

英国では、約300万人の自己検査・アンケート結果とNHSデータを統合し、COVID-19の感染動向をリアルタイム解析する大規模調査を実施し、公衆衛生施策に貢献した

諸外国の先進事例【事例⑩: リアルタイム解析】(英国・事例)

事例名	Real-time Assessment of Community Transmission (REACT) Study
取組主体	<ul style="list-style-type: none">Imperial College LondonIpsos MORI (調査企業)Department of Health and Social Care (保健福祉省)
国	英国
取組時期	2020年～
目的	新型コロナウイルスの感染状況をリアルタイムに把握し、政府の公衆衛生対策の基盤となるデータ提供を行うこと
事例概要	<ul style="list-style-type: none">COVID-19の感染動向をリアルタイムで追跡し、感染拡大の傾向、ワクチンの有効性、健康格差、リスク因子の評価などを明らかにし、公衆衛生施策の科学的根拠を提供REACT研究はウイルスの流行の波を特定し、公衆衛生施策に活用されたほか、早期に感染者の約3分の1が無症状でありながら他者に感染を広げる可能性があることを解明本データを用いたCOVID-19に関する追跡調査は現在も継続されており、2024年時点で少なくとも80本の研究論文が発表され、感染拡大のパターンや長期的な影響についての研究が進められている研究参加に同意した被験者のデータは、NHSが保有する医療情報(受診情報、診断名、COVID-19感染歴、ワクチン接種状況など)と統合され、長期的な健康影響やアウトカムに関する追跡調査が進行中
出所	https://www.imperial.ac.uk/medicine/research-and-impact/groups/react-study/

研究対象者	英国全315地域から選ばれた合計約300万人の住民(英国内で最も大規模なCOVID-19疫学調査) <ul style="list-style-type: none">被験者は地域、年齢、性別、人種等の要素が考慮され、英国の人口構成を反映するサンプルを確保
提供情報	<ul style="list-style-type: none">Ipsos MORIがNHSのデータを活用し、Imperial College Londonおよび保健福祉省の代理で被験者を募集研究参加に同意した被験者は、自己検査キットの使用とアンケート調査への協力が求められる収集されたデータは、NHS Englandが保有する医療情報と統合される
研究段階	研究初期はリアルタイム解析を通じた感染状況の把握が中心だったが、現在は被験者データを基に長期的な健康アウトカムや持続的な症状の解明が進められている
事例の特徴	<ul style="list-style-type: none">政府、学術機関、民間企業が連携し、COVID-19流行下の公衆衛生施策に貢献するデータを収集自己検査キットやアンケートを活用したりリモート調査と、NHS保有データの統合により、迅速なデータ収集・分析を実現

MELLODDYプロジェクトは、連合学習プラットフォームを構築し、10社の製薬企業が保有するデータを相互に開示することなく活用可能にすることで、創薬プロセスの効率化を図る

諸外国の先進事例【事例⑪: 連合学習】(EU・事例)

事例名	Machine Learning Ledger Orchestration for Drug Discovery (MELLODDY) project	取組主体	<ul style="list-style-type: none">・ コーディネーター：Owkin France・ 参加者：AstraZeneca AB、NVIDIA Switzerland AG等の製薬企業10社、及び学術研究機関とIT企業を含む計17パートナーで構成・ 資金提供：IMI2 JU(Innovative Medicines Initiative 2 Joint Undertaking)
国	EU8か国（ドイツ、フランス、ベルギー、イギリス、スウェーデン、スイス、オランダ、ハンガリー）	対象データ	<ul style="list-style-type: none">・ 2.6億件以上の実験データポイント・ 2,100万以上の化学化合物・ 40,000以上のアッセイデータ
取組時期	2019.06.01～2022.05.31	出所	https://cordis.europa.eu/project/id/831472 https://pubmed.ncbi.nlm.nih.gov/37642660/
目的	競争相手のデータも含めた創薬関連情報を、各企業が所有する化学・生物データを相互に開示することなく活用できるようし、創薬プロセスの効率化を図る	データフロー	<ul style="list-style-type: none">・ データ準備と転送<ul style="list-style-type: none">✓ 前処理：各製薬企業より実施✓ 転送：データセットは手動でAWS MELLODDYクラスタのS3バケットに転送・ 計算プランごとの処理<ul style="list-style-type: none">✓ 学習：各製薬パートナーの環境で、最新のヘッドとトランクモデルを使用してデータを学習。トランクモデルの更新情報を中央MLに送信し、統合・平均化された更新版が各パートナーに返される✓ 評価：最新のモデルを用いてパフォーマンス評価を実施し、中央オーケストレーターに報告✓ 出力データの転送：モデル・予測ファイル・各指標のパフォーマンスデータをAWSから自社環境へ転送
事例概要	<ul style="list-style-type: none">・ MELLODDYプロジェクトでは、データ前処理の標準化ツール「MELLODDY TUNER」を開発・導入し、<u>各製薬企業が保有する化学構造データを標準化されたフィンガープリント表現に変換し、各社のアッセイデータを予測タスクとしてニューラルネットワーク(Neural Network)に統合する</u>・ このニューラルネットワークは、全企業で共通の「トランク(Trunk)」と、各企業固有のアッセイデータに対応する「ヘッド(Head)」から構成されており、<u>トランク部分はすべての企業で共同学習される一方、ヘッド部分は各企業ごとに個別に学習される</u>・ この連合学習の仕組みにより、MELLODDYプロジェクトは、製薬企業10社がデータを共有することなく共同訓練を実現した。その結果、<u>単独学習と比較してモデルの予測性能が向上し、特に分類タスクや適用可能範囲の拡張において有益であることが示された</u>。さらに、<u>ADMEやパネルアッセイにおいても顕著な改善が実証されている</u>。本手法では、データのやり取りを伴わずにモデルの重みのみを共有することで、知的財産の保護と競合リスクの回避を両立しながら、高精度な創薬予測を可能にした。		
		今後の活用	<ul style="list-style-type: none">✓ MELLODDYコンソーシアムの成功を継続・拡大するための活用戦略に関する議論が進行中であり、これらの活動は2023年に開始される予定である(※現時点での公開情報がない)

NHSはFederated Data Platformを設立することで、地域ごとに分散した医療データを安全かつ効率的に統合・活用し、患者中心の医療と医療資源の最適管理を図る

諸外国の先進事例【事例⑫: 連合学習】(英国・事例)

事例名	NHS Federated Data Platform(FDP)		
取り組み主体	NHS England		
国	英国		
取組時期	2022年：パイロット 2023年：調達 2024年4月：各NHストラストへ展開 11月：87の病院，28の統合ケアボードに導入された		
目的	患者ケアを向上し、医療資源の管理を効率化する		
事例概要	<ul style="list-style-type: none">NHS Federated Data Platform（FDP）は、各医療機関の分散していたデータをインスタンス※1に統合し、安全かつ効率的に管理・活用するためのデジタルプラットフォームである。各NHS病院トラストや統合ケアシステム（ICS）の単位で、「ローカルインスタンス」を持ち、インスタンスごとにデータ管理者が存在する。NHS Englandは「全国インスタンス」を管理する。「ローカルインスタンス」を「全国インスタンス」と連携することで、データの安全な管理と共有が可能になり、地域ごとの課題に応じた柔軟な運用を実現する当初は①手術待機リスト削減 ②ケア調整 ③予防接種・免疫 ④人口健康管理 ⑤サプライチェーン管理の五つのユースケースでの活用にフォーカスしており、将来的にNHSのサービス提供を変革し患者転帰を向上させる土台と位置付けられている	対象データ	<ul style="list-style-type: none">主に各医療機関の運用・臨床関連データ<ul style="list-style-type: none">➢ 病院・医療資源：病床数、手術室スケジュール、医療用品の在庫状況、医療スタッフの勤務表等➢ ケア連携情報：退院先の介護施設空き状況等➢ 人口健康データ：地域住民の健康状態や医療サービスのニーズ等* GP由来の生データは全国インスタンスに直接は取り込まれず、必要時も個人識別情報を除去して活用する
出所	https://www.england.nhs.uk/digitaltechnology/nhs-federated-data-platform/	データフロー	<ul style="list-style-type: none">各組織ごとの独立インスタンスと中央インスタンスから成る連合型アーキテクチャを採用するローカルインスタンスまでデータフロー：<ul style="list-style-type: none">病院ネットワークからデータコネクタを通じてFDPに取り込まれ、データはローカルFDPテナント内で整備(取り込み・クレンジング・統合・Canonical Data Modelにマッピング)され、FDPアプリケーション(ダッシュボード、ケア管理など)で活用される全国インスタンスまでのデータフロー<ul style="list-style-type: none">地域処理センター(Regional Processing Centre)に送られ、必要な前処理が行われたのち、UDAL(Universal Data Access Layer)という中間レイヤーを経由して処理される。その後、データはNHS-PET※2 (Privacy Enhancing Technology)に登録され、必要に応じて非識別化処理が行われる。最終的に、処理されたデータは全国FDPインスタンスに取り込まれ、NHS Englandによる政策立案、全国的な医療計画、パフォーマンス評価等に活用される
		※1 インスタンス: 各組織が利用する独自のデータプラットフォーム環境を指す ※2 NHS-PET(Privacy Enhancing Technology)：IQVIAより提供され、データの流れを記録し、合法かつ安全に個人データを非識別化するためのプライバシー保護プラットフォームである	

Baricitinibは、関節リウマチの薬として承認されていたが、AIと専門知識の組み合わせにより、COVID-19の治療にも有望であることが特定され、パンデミック解消に貢献した

諸外国の先進事例【事例⑬:ドラッグリポジショニング】(米国・事例)

事例名	Expert-Augmented Computational Drug Repurposing Identified Baricitinib as a Treatment for COVID-19
取組主体	BenevolentAI社
国	米国
取組時期	2020年
目的	COVID-19に対する効果的な治療薬を迅速に特定すること。具体的には、 <u>既存の承認済み薬の中から、抗炎症作用と抗ウイルス特性を持つ薬を見つける</u> ことを目指した
事例概要	<ul style="list-style-type: none">右記載の使用データと科学文献のデータを基に、<u>BenevolentAI社のAI搭載バイオメディカル知識グラフ*で遺伝子、タンパク質、化合物、疾患などの関係図を構築した</u>また、データから学習し、どの薬剤が特定のタンパク質や経路に対して効果を持つ可能性が高いかを予測したさらに、タンパク質相互作用ネットワーク**を使用し、COVID-19に関連する重要なタンパク質やウイルスの感染メカニズム、ウイルスに対する体の反応に関与するタンパク質を特定したその後、細胞培養や動物実験を通じて、薬剤が実際に標的タンパク質や経路に作用するかを検証したその結果、Baricitinibが抗ウイルスおよび抗炎症作用を持つことが確認され、有効性は臨床試験で実証された

*複数のデータを使って新薬発見を支援するために設計されたツールである。病気やタンパク質などの情報をつなげて、データベースと文献から得た情報を基に構築されている

**このツールは、遺伝子やタンパク質とその相互作用を理解するために作られており、特定の生物学的な働きに関与する経路(一連の流れ)を見つけることができる

使用データ	<p>主にタンパク質の相互作用に関するDB 6 つが使用された：</p> <ul style="list-style-type: none">BioGRID(バイオバンク)：国立衛生研究所(NIH)とカナダ保健研究機構からの助成金によって運営。モデル生物やヒトの遺伝的・タンパク質的相互作用に関するDBであるSTRING：学術コンソーシアムによって運営されている既知・予測されるタンパク質間相互作用のデータベースであり、12,535生物の59,309,604個のタンパク質をカバーしているKEGG(日本のマルチオミクスDB)：金久ラボラトリーズが運営しており、細胞・個体・生物界がもつ生命の情報システムを分子間相互作用・反応ネットワークで表現し、コンピュータ上にモデル化したDBであるSiGNor：ローマ大学とローマ・ラ・サピエンツァ大学の共同協力で運営されており、タンパク質と生物学的に関連する他のエンティティ(刺激、表現型、酵素阻害剤、複合体、タンパク質ファミリーなど)との間の因果関係に関する情報を持つDBであるOmniPath：Saez 研究室と Korcsmaros 研究室で開発された分子生物学の事前知識のデータベースであり、シグナル伝達ネットワーク、酵素とPTMの関係、タンパク質複合体、タンパク質の注釈、細胞間コミュニケーションの役割についての 5 つのDBを構築しているPubMed：米国国立医学図書館 (NLM)の国立生物工学情報センター (NCBI)が管理するデータベースで、生物医学文献の3,800万件以上の引用と要約が含まれている
出所	https://www.fda.gov/science-research/science-and-research-special-topics/critical-path-initiative

乳がん治療に使われるエストロゲン調節療法（EMT）は、アルツハイマー病のリスクを低減する可能性があることが、72万人以上のデータと実験で示された

諸外国の先進事例【事例⑭:リバーストランスレーショナルリサーチ】(米国・事例)

事例名	Breast cancer therapies reduce risk of Alzheimer’s disease and promote estrogenic pathways and action in brain	使用データ	Marinerデータベースからの保険請求データ ➢ <u>PearlDiver社(民間)が運営するアメリカ合衆国のすべての州および地域の患者集団を対象とした保険請求データセット</u> である。このデータベースには、患者の人口統計学的特徴、処方記録等が含まれている。2010年から2018年第2四半期までの期間にわたる請求データを含み、1億2200万人の患者を対象としている
取組主体	アリゾナ大学		
国	米国		
取組時期	2022年～2023年		
目的	乳がん治療薬(EMT)がアルツハイマー病(AD)のリスクに与える影響を明らかにし、そのメカニズムを解明すること		DrugBankデータベース ➢ <u>OMx Personal Health Analytics Inc(民間)が運営する医薬品と医薬品の特定の分子(標的)に関するデータベース</u> である。約4,900の医薬品、1,565対524の非冗長タンパク質または薬物標的配列が含まれている
事例概要	<ul style="list-style-type: none">リバーストランスレーショナルリサーチでは民間企業のMarinerデータベース上にある72万人以上の乳がん患者のデータを分析し、患者集団の交絡変数を最小限に抑えながらEMTを受けた患者と受けていない患者のADリスクが比較したEMTは、元々乳がんの治療のためにエストロゲンの作用を抑える目的で開発されたが、脳では逆にエストロゲンの役割を果たす可能性があることが示唆されたDrugBankデータベースを使用し、EMTがどのタンパク質と相互作用するかを特定した。これにより、EMTが脳内でエストロゲンのような作用を持つ可能性が示されたラットにEMTを投与し、脳組織の変化を観察したところ、EMTがニューロン（情報を伝達する役割を持つ細胞）のエネルギー生成を促進することが判明した結果として、EMTを受けた患者はアルツハイマー病のリスクが低くなることが示された	出所	https://www.sciencedirect.com/science/article/pii/S2589004223023933

*エストロゲンとは体内で作られるホルモンの一つで、脳の神経伝達細胞に良い影響を与える

情報連携基盤に組み込む技術の検討

治験リクルーティングとRWDの抽出を行う場合、顕名情報の保有および取り扱いに関する法規制がないため、法規制、運用面での検討が必要である

事例③：RWD

分類	RWD
事例名	英国：Heartburn Health Programme
テーマ	治験リクルーティングとRWDの収集
概要	<ul style="list-style-type: none"> 胸やけ・逆流性食道炎などの消化器症状の管理・治療やがんの早期発見に関する研究プロジェクト NHS EnglandのDigiTrialsサービスを活用し、12万人以上被験者のリクルーティングとRWDの収集を実施 個人に送付されたSMSに掲載の登録フォームに記入することで治験に参加でき、その後NHSから消化器症状に関するRWDが最低12年間収集され解析環境(SDE)に格納される

公益性	英国では、胸やけや逆流性食道炎などの症状に苦しみ、長期的な投薬を必要とする数百万人の患者がおり、これらの症状ががんなどの深刻な健康問題につながることもあるため、その治療やケアの改善に寄与する可能性があるとしている
国民・患者への還元	<ul style="list-style-type: none"> 消化器疾患の治療・ケアの改善を通じて医療の質向上を図り、その実現に貢献できる可能性がある 研究を通じて新たな検査や治療が提供される機会の可能性がある 定期的なニュースレターを通じて最新の健康情報や生活習慣に関するアドバイスを受けることができる
インパクト	12万以上の被験者を対象に12年以上のデータを集積し、さらに従来の厳密な臨床試験とは異なり、日常的に医療現場で収集されるRWDを活用することで、より実世界に近い環境での観察研究が可能となり、エビデンスレベルの向上に寄与すると考えられる

【フィジビリティ（政策および基盤の技術検討との整合性）】			
分類	要件	整合性	理由
運用	顕名情報を保有し、データベースと紐づけが可能	×	顕名情報の保有予定がないため
	顕名情報を第三者に提供する仕組み	×	
法的	母集団から特定されず、招待を受けないオプアウトの仕組みを導入	×	現行法ではオプアウト等がないため
法的	研究の途中で参加を拒否できる仕組み	×	
技術	抽出条件に基づき対象群を作成する機能	○	対応可能
技術	参加同意者のRWDを抽出し、集積する技術	△	RWDの収集・利用範囲によるため

LLMモデルを用いた診断補助アプリ等の開発を行う場合、標準化やデータクレンジング手法の統一化や臨床ガイドラインの管理などが必要となる

事例⑦：AI

分類	AI
事例名	英国：Foresight
テーマ	RWDを活用した疾病予測モデル及び診断補助アプリ開発
概要	<ul style="list-style-type: none">OpenGPTモデルにNHSが保有する医療情報を深層学習させ医療用LLMモデルを構築し、当該技術を基盤とした対話型アプリケーションを開発技術の予測精度を検証するため、2つのNHSTラストのEHRデータを用いて患者の疾患進行を予測し、アプリの精度は68～88%と評価された。さらに、医師の評価を加えた結果、約93%の正確性を確認

公益性	<ul style="list-style-type: none">医療現場で診断補助として活用することで、診断の正確性を向上させ、見落としを防いだり、疾患の進行を予測することで早期介入が可能となり、医療の質の向上に貢献するデジタルヘルスツインの生成などを通じて、研究や医療教育への活用が可能となる
国民・患者への還元	<ul style="list-style-type: none">診断補助として活用することで、医療の質を向上させ、患者満足度の向上に貢献する患者の既往・症状を踏まえた疾病予測や健康リスク分析を行うため、個別化医療が推進するデータが蓄積されるほどに予測精度が向上する
インパクト	<ul style="list-style-type: none">医療従事者の診療負担を軽減し、正確性の向上による業務効率化を実現デジタルツインを活用した仮想臨床試験により、研究や新たな治療法の開発を促進する教育分野での活用を通じて、次世代の医療従事者の育成に貢献する

【フィジビリティ（政策および基盤の技術検討との整合性）】

分類	要件	整合性	理由
運用	膨大な量の医療情報を用いたAI学習の際に、個人情報保護される仕組み	△	匿名・仮名化情報であっても、多量の情報を学習させることで個人特定リスクあり
技術	構造化データ(年齢、性別、疾患、処方)と非構造化データ(自由記述情報等)の双方が必要となる	△	電子カルテ情報サービスの3文書6情報の活用
	複数のデータソース間で統一された標準化とデータクレンジング手法の適用	×	現段階統一されていない
	治療や処方などの医療行為およびアウトカムを時系列で整理・管理する仕組み	△	一部時系列管理されていない
	年齢、性別、地域、疾患背景など多様な患者データを網羅することが重要	○	公的DBは網羅性が一つの強みである
	最新の臨床ガイドラインを適用し、継続的に更新する必要がある	×	ガイドライン管理まで及んでいない
	医療従事者のフィードバックを反映し、モデルの精度向上を図ることが求められる	△	研究者の協力が必要
	今後画像データベースの重要性が高まる可能性がある	要検討	今後検討が必要

感染症等のリアルタイム解析を行う場合、顕名情報の保有および特定の研究データとの突合等の検討が必要である

事例⑩：リアルタイム解析

分類	最新研究（リアルタイム解析）
事例名	英国：Real-time Assessment of Community Transmission (REACT) Study
テーマ	COVID-19のリアルタイム解析及びその後の追跡調査
概要	<ul style="list-style-type: none">COVID-19の感染動向のリアルタイム追跡を通じて感染拡大の傾向やワクチンの有効性を評価し、公衆衛生施策の根拠を提供流行の波を特定し、公衆衛生施策に活用。感染者の約3分の1が無症状でも感染を広げる可能性があることを早期に解明研究参加者のデータはNHSの医療情報と統合され、長期的な健康影響の追跡が継続している

公益性	<ul style="list-style-type: none">COVID-19の感染状況を基に、感染症対策の強化や医療資源の適切配分に役立てるCOVID-19の長期的な影響を調査し、将来の治療や健康管理に活用するとともに、リスクの高い人々に適切なケアを提供する疾患への理解を深めることで、医療ガイドラインの策定や医療従事者の教育に貢献する
国民・患者への還元	<ul style="list-style-type: none">COVID-19の感染動向の把握による感染症対策強化及び医療の質向上リスクの高い人々への理解が進むことで個別化医療の発展に寄与自己検査キットやアンケート実施後に、結果の要約レポートを受け取ることができる
インパクト	<ul style="list-style-type: none">2024年時点で80本以上の論文が発表され、感染拡大のパターンや長期的影響の研究が進行中研究成果は医療ガイドラインや政策に活用される今後のパンデミックへの対応力強化

【フィージビリティ（政策および基盤の技術検討との整合性）】

分類	要件	整合性	理由
運用	顕名情報の第三者提供の仕組み	×	顕名情報保有予定ではない
	顕名情報を使った患者連絡の仕組み	×	
法的	オプトアウト・途中離脱が可能な仕組み	△	オプトアウトの仕組みなし
技術	全国的に幅広いサンプルを抽出し、年齢・性別・地域など多様な背景の対象者を含め、無症状者も対象として抽出する	△	公的DBは網羅性が一つの強みである
	該当する患者の仮名化されたRWDを抽出・集積する機能	△	患者紐づけが必要
	データ形式・標準化の統一的な仕組み	△	現在統一されていない
	官民連携によるデータ共有・解析システム	○	ビジティング解析環境で可能
	検査・アンケート結果のRWD統合技術	×	患者紐づけが必要
	リアルタイム解析・可視化ダッシュボード	○	今後構築
	今後、デジタルツインや予測モデルを活用した感染シナリオのシミュレーション技術の重要性が高まる可能性がある	×	今後検討が必要

連合学習による医療資源管理を行う場合、連合学習用のデータ項目の整理と標準化の推進が必要である

事例⑫：連合学習

分類	最新研究（連合学習）
事例名	英国：NHS Federated Data Platform(FDP)
テーマ	連合学習プラットフォームの設立による医療資源管理の効率化
概要	<ul style="list-style-type: none">組織ごとの独立インスタンスと中央インスタンスから成る連合型アーキテクチャを採用し、各医療機関のローカルインスタンスで分散していたデータを安全に全国インスタンスに統合・連携することで、安全かつ効率的なデータ管理と活用を図る①手術待機リスト削減 ②ケア調整 ③予防接種・免疫 ④人口健康管理 ⑤サプライチェーン管理の五つのユースケースでの活用及び、将来的にNHSのサービス提供を変革し患者転帰を向上させることが期待される

公益性	<ul style="list-style-type: none">医療資源の最適配分、地域間での連携強化と格差是正等を実現することにより公的利益に資する
国民・患者への還元	<ul style="list-style-type: none">患者の待機時間を短縮する分断のない連携したケアにより、個別の状況に応じたサポートが可能となるワクチン接種への公平なアクセスが保障され、健康格差を是正する地域に根ざしたサービス計画によって、患者のニーズに即した医療が提供され、効率的な資源配分は医療の質と持続可能性を向上する
インパクト	<ul style="list-style-type: none">2024年11月時点で87の病院と28の統合ケアボードがFDPを導入したFDPを導入した病院トラストでは、ツールの導入後、毎月平均で114人多くの入院患者を手術室で治療できるパイロット病院の一つでは、患者が退院可能となった後の在院日数が37%削減された

【フィージビリティ（政策および基盤の技術検討との整合性）】

分類	要件	整合性	理由
運用	フェデレーション運用のためのガバナンス体制	×	十分なデータガバナンスがない
	医療機関オペレーティングデータの所持	×	連合学習用のデータ項目がない
法的	直接診療以外の目的での組織間でデータを共有できる仕組み	×	診療目的を超えた利用は本人同意の検討が必要
技術	クラウドベースの分散型アーキテクチャ	△	医療機関ではオンプレミス中心のシステム運用
	標準化されたデータモデル	△	標準化が十分進んでいない
	統合解析・ダッシュボードのリアルタイム可視化	○	情報連携基盤での機能に含まれる
	データ照合・結合・連携できる仕組み	○	

リバーストランスレーショナルリサーチではレセプト情報と医薬品DBを連結解析する必要があり、本技術を適用する場合には、医薬品関連DBのデータソースが必要となる

事例⑭：リバーストランスレーショナルリサーチ

分類	最新研究（リバーストランスレーショナルリサーチ）
事例名	米国：Breast cancer therapies reduce risk of Alzheimer’s disease and promote estrogenic pathways and action in brain
テーマ	既存薬の転用
概要	<ul style="list-style-type: none">乳がん治療薬(EMT)がアルツハイマー病(AD)のリスクに与える影響を明らかにし、そのメカニズムを解明することを目的とした研究であるMarinerデータベース上にあるアメリカ合衆国のすべての州および地域の患者集団を対象とした保険請求データセットから72万人以上の乳がん患者のデータを分析した薬物の詳細情報とタンパク質についての情報を所有するDrugBankデータベースを使用し、EMTがどの脳内タンパク質と相互作用するかを特定した

公益性	<ul style="list-style-type: none">米国では乳がんとアルツハイマー病(AD)のリスクは、女性の高齢化に伴い増加している。EMTは、乳がんの治療に広く使用されており、神経機能や健康に対する潜在的なリスクや利益が懸念されている
国民・患者への還元	<ul style="list-style-type: none">研究を通じて新たな治療が提供される機会の可能性がある既存薬へのリスクを懸念し使用を避けていた乳がん患者に対し、既存薬のリスクの低減を示すことで乳がん治療に対する新たな選択肢を与えることになる
インパクト	<ul style="list-style-type: none">72万以上の乳がん患者のデータ分析にタンパク質の相互作用に関するデータを組み合わせることで、従来では知り得なかった既存薬と人間の体の細胞との関係性を洗い出すことが可能となる。これにより、既存薬の多様な可能性が可視化され、多種多様な病に対する解決策の獲得に寄与すると考えられる

【フィージビリティ（政策および基盤の技術検討との整合性）】

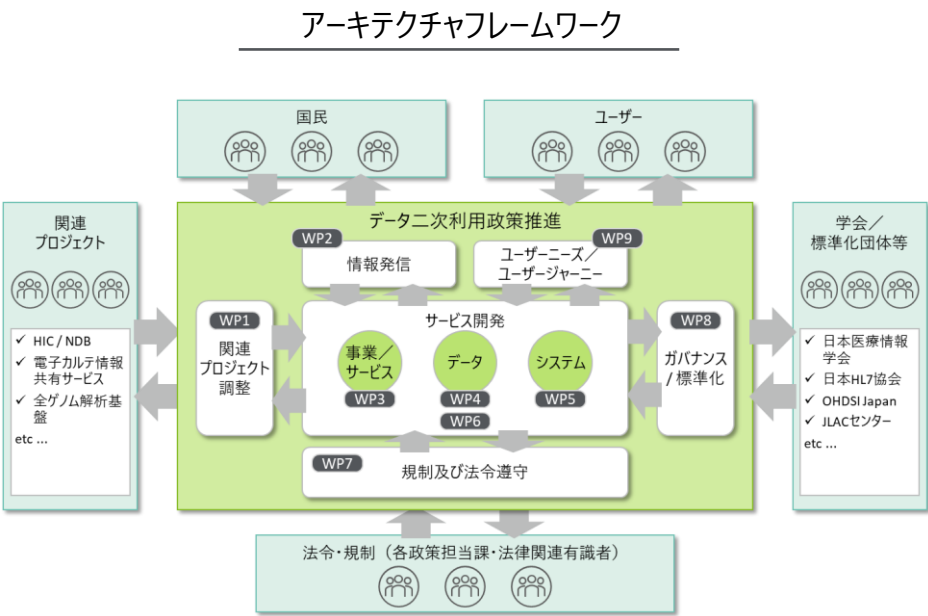
分類	要件	整合性	理由
運用	保険請求データの利用	○	NDBでレセプトデータを保有している
	医薬品と医薬品の性質を持つDBの利用	×	医薬品と医薬品の性質を持つDBを所持していない
技術	条件に基づき対象群（薬剤の詳細情報に基づく対象者特定）を抽出する機能	○	対応可能

6. 総括

事業総括として、令和10年度のサービス開始に向けたフレームワークおよびロードマップを整理した

総括

- ✓ 今年度の事業総括として、情報連携基盤および電子カルテ情報データベース（仮称）に係る医療等情報の二次利用のシステムアーキテクチャのフレームワークを整理
- ✓ EHDS規則のWP（ワークパッケージ）構成を参考にしながら、事業計画・事業運営面でのWPを補強した「日本版のWP」を作成
- ✓ WPを踏まえた、R10年度に向けてロードマップ案を作成



ロードマップ

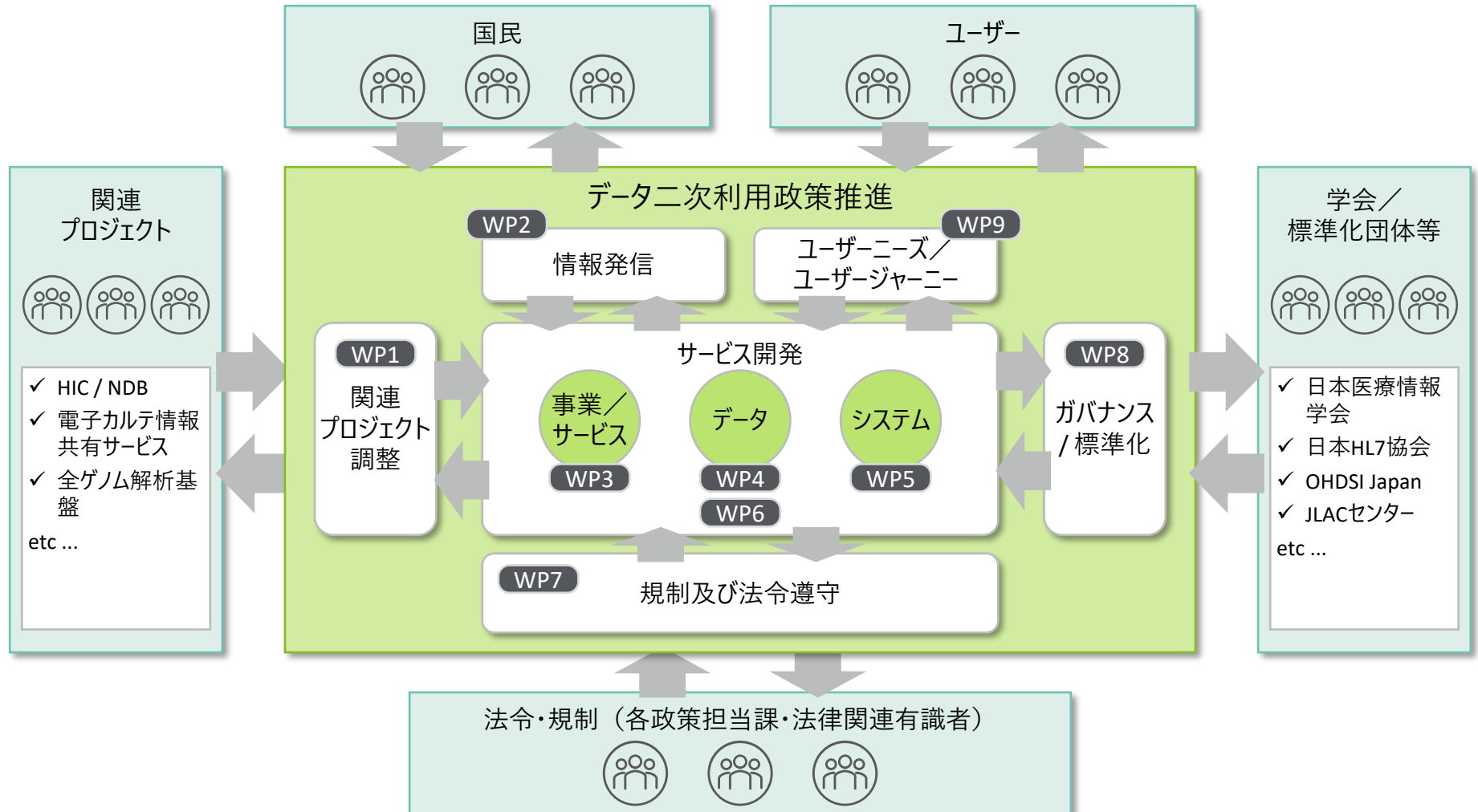
実施タスク	R6	R7年度				R8年度				R9年度				R10
	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q
マイルストーン				★ PoC 実証	★ 調査仕様 書作成					★ パイロット検証開始				★ 全利用者への公開
[WP1] プログラム管理	WP設計 関連PR 整理	推進体制/コンソ 設計・構築				WP全体進捗の管理、WP間整合や横断課題の管理、及び本プログラム外部との折衝								
[WP2] 規制・法令遵守	法規制の改正に係る検討・対応													
[WP3] ガバナンス構築						ガバナンス方針／ スコープ定義				規程類の作成、ガバナンスプロセス設計・体制構築				訓練
[WP4] PoC／パイロット企画推進	PoC企画	PoC準備	PoC 実施	PoC 検証	パイロット企画	パイロット準備				パイロット実施				
[WP5] ユースケース／サービス利用フロー設計	ユースケース整理				サービス利用フロー 設計				サービス利用マニュアル開発				訓練	
	ユースケース具体化								ユースケース具体化					
[WP6] ビジネスモデル／サービス運営体制構築	ビジネスモデル/サービスモデル 初期案検討				サービス運営 業務フロー設計				サービス運営オペレーション設計・構築					
					情報発信内容定義				情報発信コンテンツ開発					
[WP7] 情報発信／普及活動					普及活動計画				普及活動推進					
[WP8] データ品質向上／データマネジメント構築	データ提供 調整				データ品質分析・改善施策立案				データマネジメントルール/プロセス・体制構築					
	標準化対象定義・ 初案作成				データ標準策定（メタデータ、データモデル、コード）				実装					
[WP9] プラットフォーム開発	プラットフォーム 企画構想				要件定義				システム 開発調達				総合 テスト	
					インテグレーション要件定義				プラットフォーム設計・開発・テスト					
[WP10] インテグレーション	関連シス テム調査				標準ルール開発				インテグレーション機能設計・開発・テスト					

アーキテクチャフレームワークの整理

データ二次利用サービスの検討を行うにあたり、データ二次利用サービスの事業・サービス／データ／システムの構造に加え、外部要因を捉えながら検討を行う必要がある

データ二次利用政策推進体制と関連するステークホルダー

WP-X 対応するEHDS WP



内外の要因により論点の複雑化が想定される本サービスの検討推進において、 エンタープライズアーキテクチャのフレームワークを利用した論点の構造化が有用と推察される

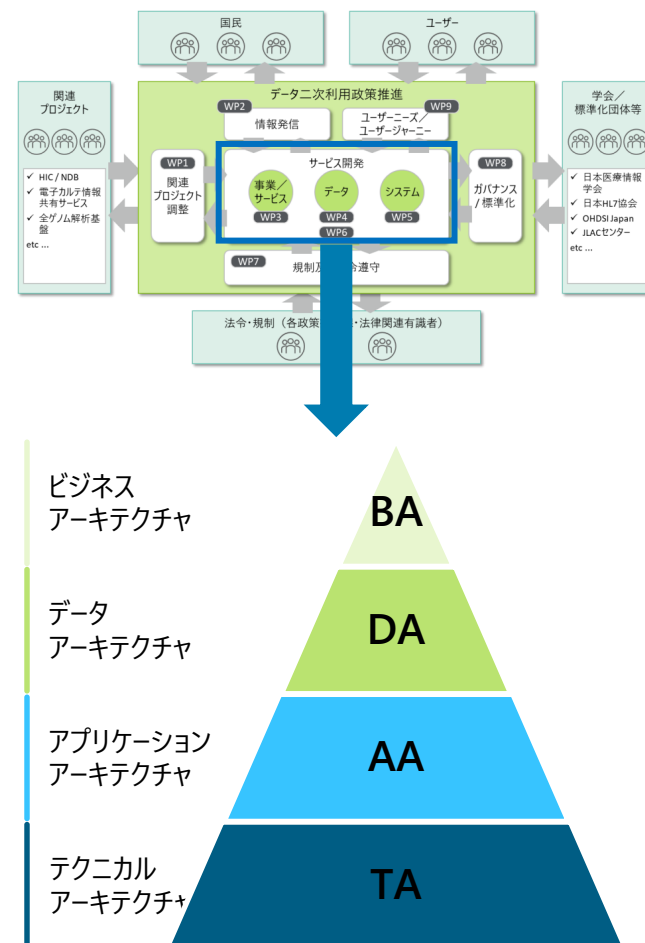
エンタープライズアーキテクチャのフレームワークの活用理由

エンタープライズアーキテクチャ（EA）とは？

- 組織全体のITインフラとビジネスプロセスを統合・最適化するためのフレームワークである
- ビジネス戦略と情報システムの整合性を図ることで、組織目標の確実な達成につなげるためのアプローチである

なぜ本サービスの論点整理にEAのフレームワークを活用すべきか？

- 本サービスは治療法・創薬・医療機器開発のために、研究機関や製薬企業・医療機器メーカーのような多様なバックグラウンドを持つステークホルダーが存在する。よって、公益性に基づくサービスガバナンスを確立し、情報連携基盤として実現すべき目標を明確に示す必要がある
- 目標の確実な実現のため、ビジネス戦略と情報システムが整合した全体構想（Blueprint）の策定、それを受けたプロセス・ルールの定義/最適なIT環境の定義へブレークダウンしながら進めることが肝要である



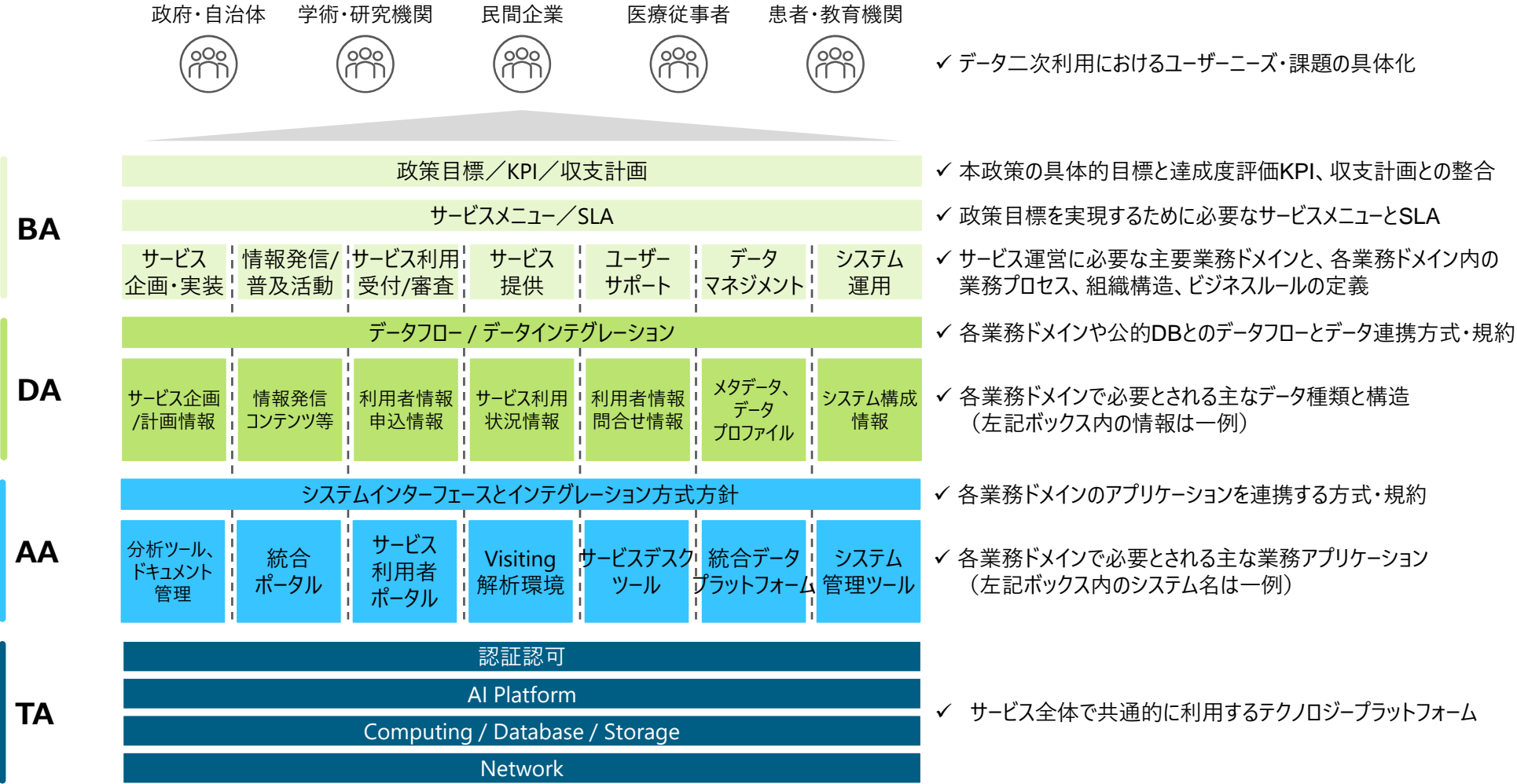
EAの各レイヤにおける目的と主要論点、主な成果物は下記の通り

(参考) EAの各レイヤにおける目的・主要論点・主要成果物

	主要目的・主要論点	主要成果物
BA	<p>目的：業務内容、実施主体等で実現すべき姿を体系的に示す</p> <p>主要論点：</p> <ul style="list-style-type: none">✓ ①ビジネスプロセスの観点でベースラインとターゲットアーキテクチャ定義✓ ②ベースラインとターゲットアーキテクチャとのギャップ分析✓ ③ターゲット実現に求められる対応方針検討	<ul style="list-style-type: none">✓ ビジネス目標・戦略定義書、KPI定義書✓ ビジネスプロセス定義書（AsIs/ToBe）✓ ビジネスルール定義書✓ 業務機能一覧 など
DA	<p>目的：業務・システムで利用するデータ内容/各データ間の関連を体系的に示す</p> <p>主要論点：</p> <ul style="list-style-type: none">✓ ①情報システム（データ）の観点でベースラインとターゲットアーキテクチャ定義✓ ②ベースラインとターゲットアーキテクチャとのギャップ分析✓ ③ターゲット実現に求められる対応方針検討	<ul style="list-style-type: none">✓ データモデル定義書（ER）✓ メタデータ定義書✓ データフローダイアグラム定義書（DFD）✓ データガバナンスポリシー定義書✓ データアクセス権限方針書 など
AA	<p>目的：業務処理に最適なシステムの形態を体系的に示す</p> <p>主要論点：</p> <ul style="list-style-type: none">✓ ①情報システム（アプリ）の観点でベースラインとターゲットアーキテクチャ定義✓ ②ベースラインとターゲットアーキテクチャとのギャップ分析✓ ③ターゲット実現に求められる対応方針検討	<ul style="list-style-type: none">✓ アプリケーションポートフォリオ（AsIs/ToBe）✓ アプリケーション関係図（AsIs/ToBe）✓ システム機能一覧✓ インターフェース一覧✓ アプリケーションガバナンスポリシー定義書
TA	<p>目的：システムを構築する際に利用する技術的構成要素を体系的に示す</p> <p>主要論点：</p> <ul style="list-style-type: none">✓ ①テクノロジーの観点でベースラインとターゲットアーキテクチャ定義✓ ②ベースラインとターゲットアーキテクチャとのギャップ分析✓ ③ターゲット実現に求められる対応方針検討	<ul style="list-style-type: none">✓ テクノロジーロードマップ（AsIs/ToBe）✓ インフラストラクチャダイアグラム（AsIs/ToBe）✓ 技術標準定義書/ガイドライン✓ セキュリティポリシー定義書✓ テクノロジーガバナンスポリシー定義書

EAを踏まえ、本サービスの提供価値と提供サービス、必要な業務ドメイン、各業務ドメインを支えるデータ・アプリケーション・テクノロジーについて全体整合を取りながら具体化を行う

EAフレームワークを活用した本サービスの論点全体構造（案）



R10からの電子カルテ情報等分析関連サービスの開始に 向けた二次利用領域の整理について

R10からの電子カルテ情報等分析関連サービスの開始に向けて、R7 ～ R9の約3カ年で実施すべき事項について整理する

今後の年間スケジュール

○：翌年度に向けて対応する事項
●：当該年度の実施事項



本事業においては、EHDSのWP構成を参考にしつつ、事業計画・事業運営面でのWPを補強した下記のWP構成が想定される

Work Package (案)

EHDS WP
との関連性

共通 テーマ	[WP1] プログラム管理	<ul style="list-style-type: none"> ✓ WP設計とWP推進体制（コンソーシアム）の構築 ✓ WP全体進捗の管理、WP間整合や横断課題の管理、及び本プログラム外部との折衝 	WP1 WP3
	[WP2] 規制・法令遵守	<ul style="list-style-type: none"> ✓ 関連法規制の改正、オプトアウトに係る対応など 	WP7
	[WP3] ガバナンス構築	<ul style="list-style-type: none"> ✓ 本サービスの運営状況を監督し、医療等情報の二次利用に係る適切なサービス提供・安全対策等を実現するガバナンスルール・体制・プロセスの構築 	WP4
	[WP4] PoC／パイロット企画推進	<ul style="list-style-type: none"> ✓ PoC及びパイロットの企画及び推進を実施 	—
事業 テーマ	[WP5] ユースケース／サービス利用フロー設計	<ul style="list-style-type: none"> ✓ ユーザー像の具体化と、各ユーザーのユースケースの具体化 ✓ 各ユーザーのサービス利用フローの設計 	WP9
	[WP6] ビジネスモデル／サービス運営体制構築	<ul style="list-style-type: none"> ✓ ビジネスモデル（サービスメニュー、マネタイズプラン、等）の策定と収支計画の策定 ✓ サービス運営業務の設計及び体制の構築 	—
	[WP7] 情報発信／普及活動	<ul style="list-style-type: none"> ✓ データ利用者及び国民・患者向け情報発信コンテンツの検討 ✓ 普及活動の推進 	WP2
技術 テーマ	[WP8] データ品質向上／データマネジメント構築	<ul style="list-style-type: none"> ✓ [WP5]を踏まえたデータ品質目標の定義、品質課題の抽出、品質向上施策の立案と推進 ✓ データ標準（メタデータ、データモデル、コード）策定、データマネジメントプロセス・体制構築 	WP6
	[WP9] プラットフォーム開発	<ul style="list-style-type: none"> ✓ 情報連携基盤、電子カルテ情報DB（仮称）の要件定義・調達・設計・開発・テスト ✓ プラットフォームアーキテクチャ標準ルールの策定 	WP5
	[WP10] インテグレーション	<ul style="list-style-type: none"> ✓ 公的DB等、外部システムとのインテグレーション機能の要件定義・調達・設計・開発・テスト ✓ インテグレーションアーキテクチャ標準ルールの策定 	WP8

各WPに対して、R7 ～ R9の約3カ年のマイルストーンおよび実施すべき事項について整理する

Work Packageに基づくロードマップ

実施タスク	R6	R7年度				R8年度				R9年度				R10	
	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	
マイルストーン				★ PoC 実証	★ 調達仕様 書作成					★ パイロット検証開始				★ 全利用者への	
[WP1] プログラム管理	WP設計 関連PRJ 整理	推進体制/コンソ 設計・構築			WP全体進捗の管理、WP間整合や横断課題の管理、及び本プログラム外部との折衝										
[WP2] 規制・法令遵守	法規制の改正に係る検討・対応														
[WP3] ガバナンス構築				ガバナンス方針／ スコープ定義		規程類の作成、ガバナンスプロセス設計・体制構築							訓練		
[WP4] PoC／パイロット企画推進	PoC企画	PoC準備		PoC 実施	パイロット企画	パイロット準備			パイロット実施						
[WP5] ユースケース／サービス利用フロー設計	ユーザ像整理			サービス利用フロー 設計		サービス利用マニュアル開発								訓練	
	ユースケース具体化					ユースケース具体化									
[WP6] ビジネスモデル／サービス運営体制構築	ビジネスモデル 初期案検討	ビジネスモデル 精緻化		サービス運営 業務フロー設計		サービス運営オペレーション設計・構築									
[WP7] 情報発信／普及活動				情報発信内容定義		情報発信コンテンツ開発									
				普及活動計画		普及活動推進									
[WP8] データ品質向上／データマネジメント構築	データ提供 調整	データ品質分析・改善施策立案			データマネジメントルール・プロセス・体制構築										
		標準化対象定義・ 初案作成			データ標準策定（メタデータ、データモデル、コード）				実装						
[WP9] プラットフォーム開発	プラットフォーム 企画構想		要件定義			システム 開発調達	プラットフォーム設計・開発・テスト						総合 テスト		
[WP10] インテグレーション		関連シス テム調査	インテグレーション要件定義			システム 開発調達	インテグレーション機能設計・開発・テスト								
			標準ルール開発												

WS3では、サービス利用、サービス運営、データ、AI、プラットフォーム、インテグレーションに対するルール作りと浸透、レビュー、改善に向けた提言を行う体制・プロセスを構築する

[WP3] ガバナンス構築の目的と実施事項の概要（例）

WP3の目的

- ✓ 利用価値が高く、安心・安全にヘルスケアデータの2次利用が可能サービスを実現するために、サービス利用、サービス運営、データ、AI、プラットフォーム、インテグレーションに対するルール作りと、ルールの浸透、レビュー、ルール適合のための改善提言を行う体制・プロセスを構築すること

ガバナンスモデルのイメージ

