

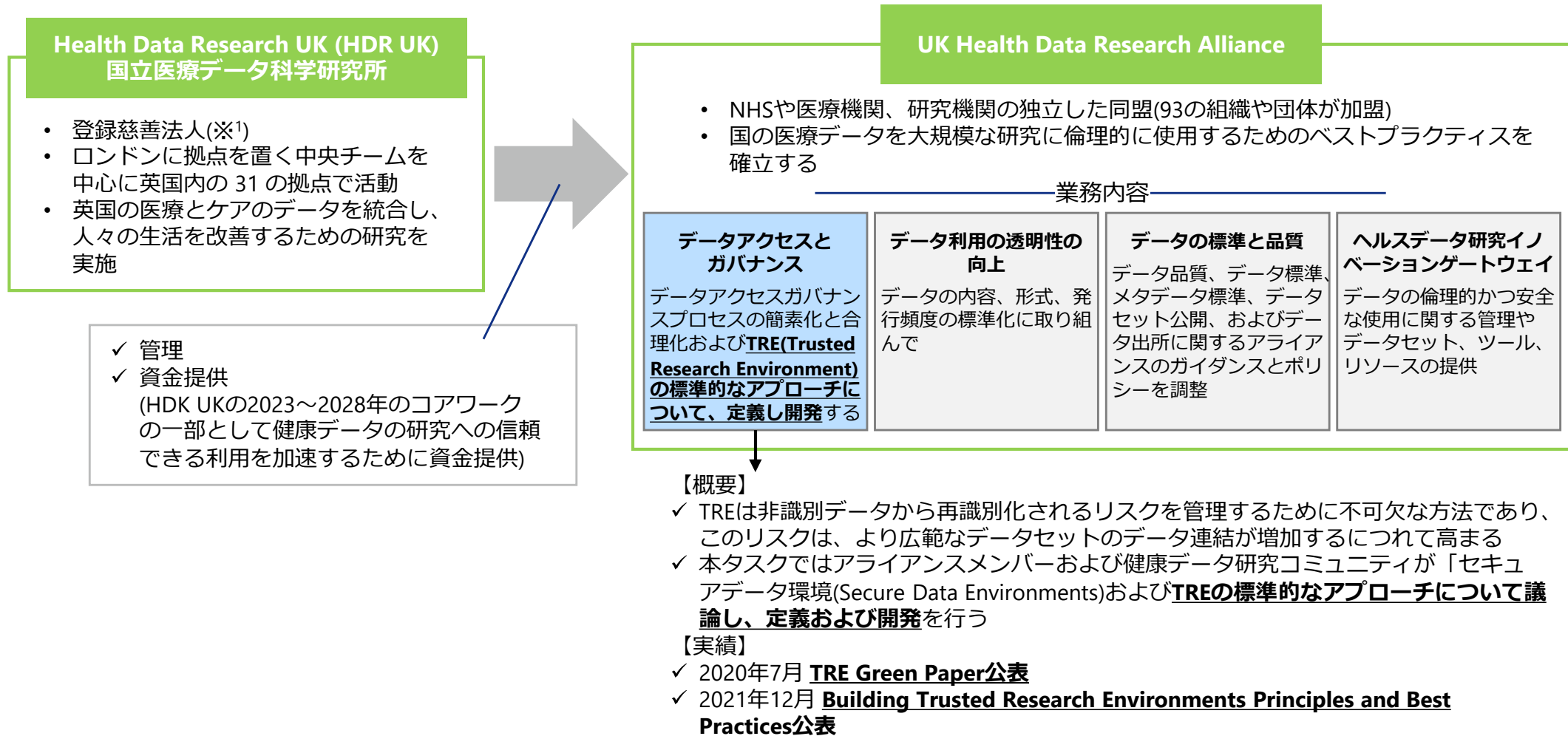
諸外国における取組について

厚生労働省 医政局 特定医薬品開発
支援・医療情報担当参事官室

Ministry of Health, Labour and Welfare of Japan

英国の医療データの利活用に関する組織構成

TREに関する文書を公表しているUK Health Data Research Allianceは、HDR UK管理の元、NHSや医療機関、研究機関などの組織によるアライアンスによって運営されている



※¹ 慈善委員会 (the Charity Commission : 政府が管理) によって規制される規約を持つ法人

TREの原則とベストプラクティス策定までの流れ

TREの標準化に向け、UK Health Data Research Allianceは、TREの検討に関する専用チームを設置し、2020年に「Green Paper」、2021年に「原則とベストプラクティス」を発表した



Green PaperおよびPrinciples and Best Practicesの関係性

政策提案文書Green Paperでの患者や市民の広範な支持に基づき、よりTREの利用を拡大させるため Principles and Best Practices(原則とベストプラクティス)が公表された

Green Paper



Green Paper v2.0 dated 15 July 2020	
Table of Contents	
Executive Summary	3
Status of the document	3
Overview	4
Purpose	4
Background	4
The case for TREs providing access to health data through safe settings	4
Requirements for a Trusted Research Environment	10
High profile	10
High integrity	10
High security	10
Open computing – an extension of safe setting	10
High data	10
High quality	10
High value – delivering the Trusted Research Environment objectives	10
Figure 1. Lifecycle example of a TRE environment	10
Responsible requirements for a TRE environment and for data providers	10
Accreditation of TREs	10
Communication, engagement and involvement	10
Public Trust	10
Conclusions	10
Next steps	10
Appendix A: Consultation Questions	10
Appendix B: Summary of changes to document	10

Building Trusted Research Environments Principles and Best Practices ; towards TRE ecosystems



Building Trusted Research Environments v1.0 dated 08 December 2021	
Foreword	3
Executive Summary	4
Alliance vision	4
Overview	7
Purpose	7
Reasons for existence	7
Communication, Engagement and Public Trust	7
Principles and best practice for a Trusted Research Environment	13
High profile	13
High integrity	13
High security	13
High data	13
High quality	13
High value – delivering access to the TRE objectives	13
Principles and examples towards a federated TRE ecosystem	13
Foundation for high profile	13
Foundation for high integrity	13
Foundation for high security	13
Foundation for high data	13
Foundation for high quality	13
Conclusions and next steps	13
Appendix A: The case for TREs providing access to health data through safe settings	13
Appendix B: Timeline of TRE ecosystem	13

項目	内容
公表日	2020年7月
作成者	UK Health Data Research Alliance
概要	・ 公衆の信頼を構築し、変化する健康データサイエンスのニーズに対応する戦略立案としてTREの要件を作成 ・ 全31ページ
章立て	✓ エグゼクティブサマリー、概要 ✓ TREの要件 ✓ TREの環境とデータフェデレーションに関する研究者の要件 ✓ TREの認定 ✓ コミュニケーション、エンゲージメント、関与

項目	内容
公表日	2021年8月
作成者	UK Health Data Research Alliance
概要	・ Green Paperに関する協議では、特に患者や一般市民の代表からこの方向性に対する広範な支持が示された ・ 本文書は医学研究の公益を促進する一方で、研究でデータを利用される人々のプライバシーやその他の利益を適切に保護するため、TREのアプローチを発展させる原則とベストプラクティスを成文化し、TREを拡大するものである ・ 全34ページ
章立て	✓ エグゼクティブサマリー、概要 ✓ TREの原則とベストプラクティス ✓ 連携したTREエコシステムに向けた原則と事例 ✓ 結論と次のステップ

Principles and Best Practices (原則とベストプラクティス) の概要

Principles and Best Practices では、「TREの実施ケース(ベストプラクティス)」や「一般市民等へのアプローチ」、「TREの要件」などが示されている

Principles and Best Practices (原則とベストプラクティス)の主な内容

TREの実施ケース	TREに関するパブリック及びその他のステークホルダーとの関わり	TREの要件	連携したTREエコシステムに向けた原則と事例
TREは健康データを保有する個人のプライバシーを保護する一方で、疾病の理解を深め、医療とケアの改善を図る大規模なデータ解析を促進する。TREの実施ケースとして付録Aに以下の例が示されている 【Data Safe Havenのケース】 ✓ 英国スコットランドではNHSのデータを研究や医薬品安全性監視の利用において分析プラットフォームを持つシステムを導入しており、Safe Haven憲章に記載されているように安全な設定である 【Genomics England】 ✓ Genomics England Research Environment (GERE)も安全な環境であり、10万ゲノムプロジェクトからの大規模な20Pb(ペタバイト)を超えるゲノムデータを運営している ✓ GEREには2,000人以上の研究者が登録し、様々なツールを利用できる解析環境を提供している 付録A実施ケースより一部抜粋	広範な協議の結果、特に患者代表や一般市民代表からこのアプローチに対する広範な支持について示されている。具体的に一般市民等へのアプローチとして、付録Bに以下が示されている ✓ 2019/09: 公正なパートナーシップに関する患者データ理解市民審査委員会へのプレゼンテーション ✓ 2019/11: Pistoria Alliance(ライフサイエンス領域で200社超の会員企業を擁する非営利団体)にてTRE導入に関するプレゼンテーション ✓ 2020/02: ロンドン市民サミットでのTREに関するセッション ✓ 2020/06: NHSフォーラムでTRE戦略について発表 ✓ 2020/10: 医療政策フォーラムでTREについて講演 ✓ 2020/11: Genomics EnglandによるTREのデモと討論集会 ✓ 2020/11: マイデータイベント、TRE紹介 ✓ 2021/02: TREを含む公共政策プロジェクトについて講演	✓ UK Health Data Research AllianceのTREの「参加のための原則」には、以下の5つの安全設定が盛り込まれている ✓ 国民の信頼を確実に維持するためには、TREが共通の原則を採用し、異なる組織によって運営される異なるシステムであるにもかかわらず、認定を受けるために5つの要件を実装する必要がある 【5つの安全設定】 ① Safe people ② Safe projects ③ Safe setting ④ Safe data ⑤ Safe output	TREを使用する可能性を最大限に高めるためには、研究者の利用プロセスを簡素化し、複数のTREにアクセスするための障壁を下げるための共通仕様とシステムが必要である。そのためのTREの追加的な検討事項および事例が示されている 【HDR Gateway】 ✓ HDR UKが運用するユーザログイン機能を備えたGatewayを用いた接続方法であり、複数のアライアンスがTREを利用する場合のサポートサービスとして実装を検討 ✓ 複数のTREにアクセスしたい個人の手続きを簡素化することで、より多くのデータ間での分析を容易にサポートできるようになる 事例より一部抜粋

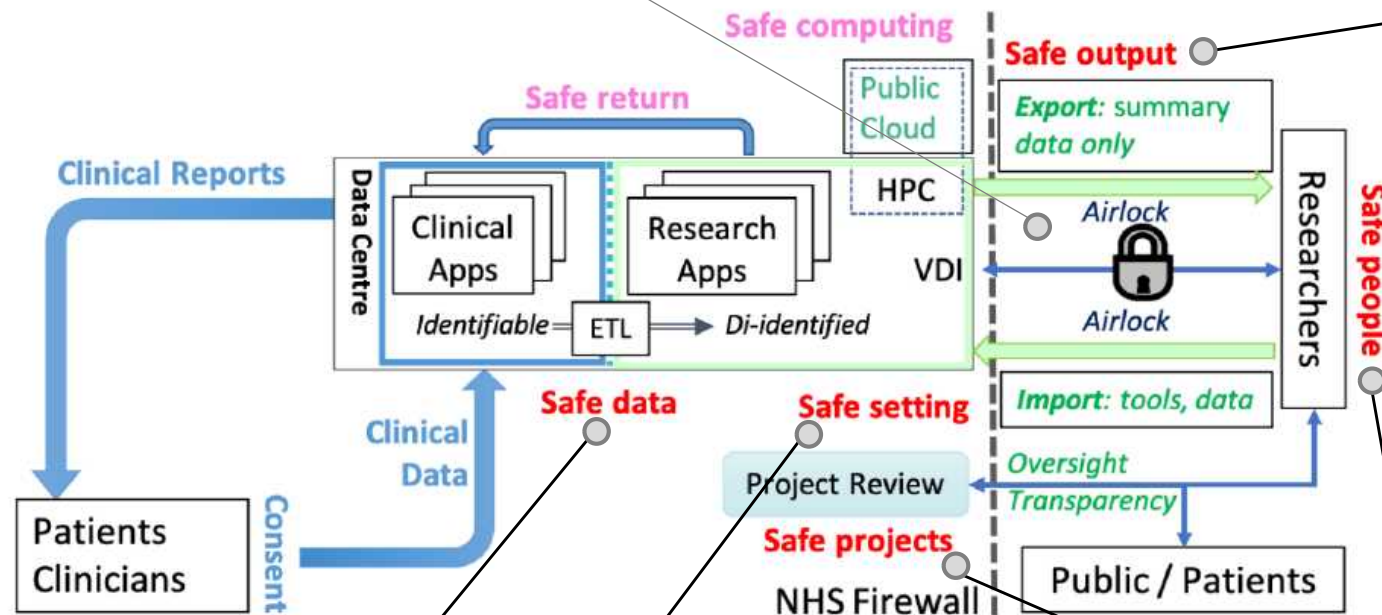
TREにおける5つの安全設定

【参考】Principles and Best Practicesにおける5つの安全設定は以下の通り

- Airlockとは**不正なデータのインポートやエクスポートを防止**することを目的とし設置されるアプリケーションである
- Airlockの機能として、出入りメカニズム (ingress/egress Mechanism)、データ移動 (Data movement)、セキュリティゲート (Security gates)、承認メカニズムと通知 (Approval mechanism and Notifications)がある
- 研究者が必要なデータ項目を申請し、**Airlock マネージャーの確認ののち、必要であればAirlock委員会で協議**される (例：Genomics EnglandのAirlock委員会は6人からなる委員会)

赤字 = 5つの安全設定

ETL：データ収集・加工ツール
HPC：ハイパフォーマンスコンピューティング
VDI：デスクトップ仮想化



- TREは無許可のデータエクスポート(またはインポート)を防止するために、**セーフセッティングと外部との間にバリア(または「エアロック」)**を設置しなければならない
- TREは、承認されたデータがこのバリアを通過できるような**プロセスとシステムを導入**しなければならない

- データは偶発的な再識別を防ぐために、全てのデータから**全ての直接識別子を削除し、情報のない仮名に置き換えることが重要**である
- TREプロバイダーは、データ管理者によって排他的に制御されるデータ暗号化キーを使用して保存時に暗号化されたデータを**安全に保持および管理するプロセスおよびシステムを実装**する必要がある

- データ利用者にTREへのリモートアクセスを提供するために、**仮想デスクトップインターフェース(VDI)またはReproducible Analytical Pipelines(RAP:自動化された統計、分析プロセス)を使用**する

- データ利用者はプロジェクトは**公共の利益**の可能性があることを証明する必要がある
- アクセスの申請は重要な作業であるため、**データ管理者はそのプロセスを可能な限り明確かつ透明にする必要がある**

- TREへのアクセスを許可される個人は、**アナリストや研究者**であるべきである
- **適切な資格を証明し、データセットへのアクセス要件に合致した承認されたプロジェクトを実施**しているべきである

Visiting解析環境の各国比較

未定稿

○：デフォルト △：オプション

			英国	フランス	フィンランド	米国
組織名			NHS England	Health Data Hub	FINDATA	CMS
Visiting解析環境の名称			Secure Data Environment	Project Space	Kapseli	VRDC(Virtual Research Data Center)
利用環境	プログラミング言語	Python	○	○	○	○
		R	○	○	○	○
		Golang			○	
		SQL	○			○
	IDE(統合開発環境)	JupyterLab		○		不明
		RStudio	○	○	○	
		Anaconda3			○	
	ソフトウェアインフラ	Apache Spark		○		
	統計解析ツール	SAS			○	○
		SPSS			○	
		Stata	△(追加料金)		○	○
		Databricks	○			○
	データ可視化	Apache Superset		○		
	データストレージ	MySQL		○	△(追加料金)	
		PostgreSQL		○		
	その他	Cytomine		○		
		OHIF Viewer		○		
		GeoDa			○	
		QGIS			○	
	データ共有・共同編集等のツール	Git		○	○	
		GitLab	○	○		
独自ソフトウェアの使用可否			不明	不明	使用可 (審査が必要)	不明

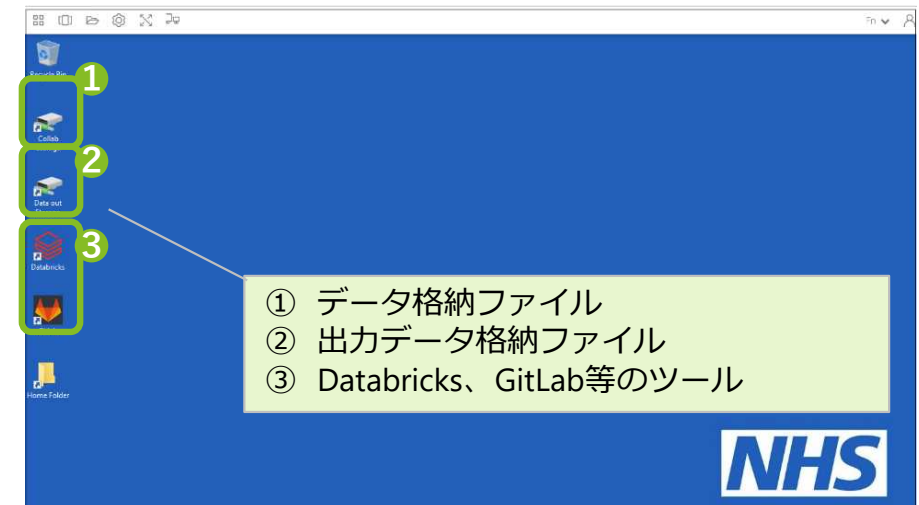
英国NHSのVisiting解析環境 (Secure Research Environment) について

未定稿

Secure Research Environment (SDE) はアクセス権限が付与された研究者が安全にNHSの医療データの受取り及び統計分析が可能なプラットフォームである

概要	研究者が患者の個人情報侵害することがない安全な環境の基で、必要なデータにアクセス・解析すること可能なプラットフォーム。SDEに格納される患者情報は全て仮名化されている。	
TRE利用のメリット	TRE利用の主なメリットとして以下の項目の担保が挙げられる 【データ品質】 → データソースを大規模にリンクできるようにし、リンクされたデータを標準化された方法で分析するためのツールを提供 【セキュリティ】 → 権限が付与された研究者のみがTREを利用する → 一次データはTREからダウンロードされることはない 【透明性】 → 研究者のデータ解析活動を記録・監視することができる 【プライバシー】 → プロジェクトに必要なデータのみが提供される	
利用者	アクセス権が付与された研究者	
提供データ	研究者はData Sharing Agreement (データ共有契約) に従い、研究に必要と判断され承認されたデータへのアクセスが与えられる	
利用環境	プログラミング言語	Python, R, SQL
	IDE(統合開発環境)	RStudio
	統計解析ツール	Databricks, Stata (オプション)
	データ共有・共同編集等のツール	GitLab

SDE利用画面



NHS England では、SDEユーザーのために利用方法やツールの使い方をサポートするデータスペシャリストによるドロップインセッションを定期定期に開催している

フランスのVisiting解析環境「Project Space」について

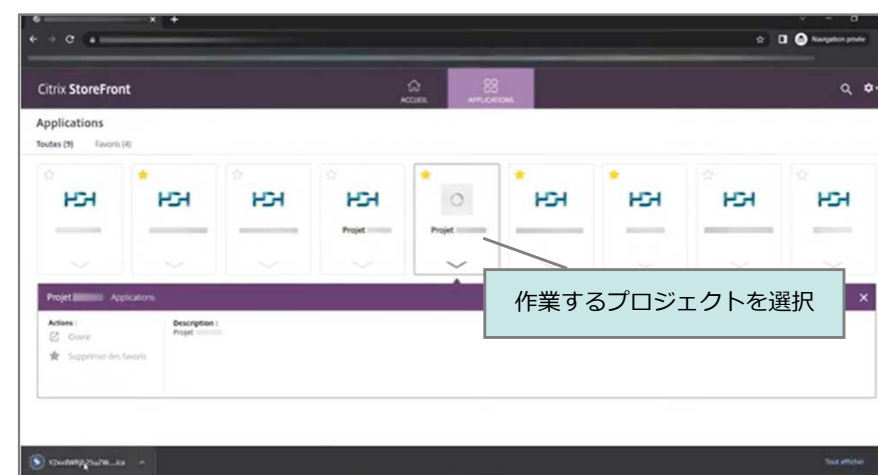
未定稿

HDH (Health Data Hub) 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

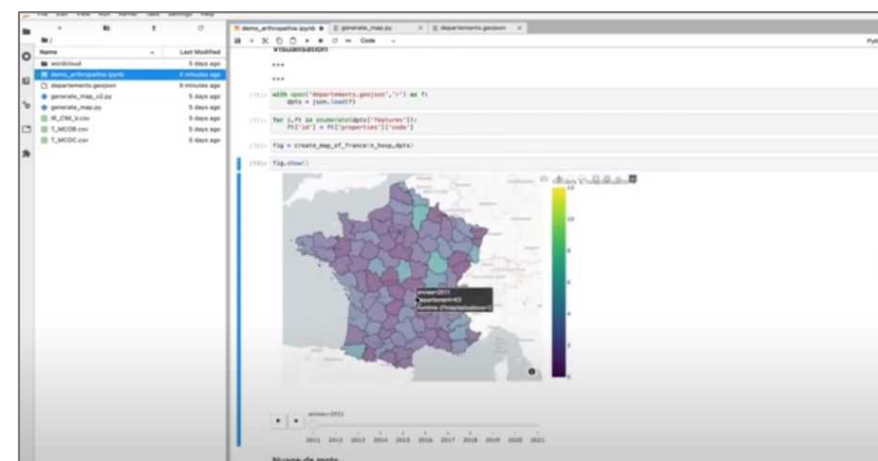
概要	HDH内の「Project Space」は、権限が付与された研究者等が安全に医療情報を受信・閲覧・保存・解析できる環境。リモートアクセスが可能だが、外部環境から遮断されており、ローカルPCにデータを転送することはできない	
Visiting解析環境のメリット	- データ整理・分析・可視化のための幅広いツールを用意 - チームメンバー間のコラボレーションが可能 - SNDSのセキュリティ基準への準拠（GDPR基準への準拠は現在対応中） - HDHによるユーザーのログ管理 等	
提供データ	HDHのデータカタログに基づく	
利用者	HDHの審査(CESREES)およびCNIL(データ保護機関)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	- Python - R
	IDE(統合開発環境)	- JupyterLab - RStudio
	ソフトウェアインフラ	Apache Spark
	データ可視化	Apache Superset
	データストレージ	- MySQL - PostgreSQL
	画像解析	- Cytomine - OHIF Viewer
	データ共有・共同編集のための開発支援ツール	- Git - GitLab

Project Space利用画面例

■ プロジェクト選択画面



■ JupyterLab解析画面(Project Space内)



フィンランドのVisiting解析環境(FINDATA kapseli)について

未定稿

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を提供している

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	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	
提供データ	健康・社会データの二次利用等に関する法律の対象となる管理者(Kanta、健康福祉研究所(THL)、医薬品庁(Fimea)、年金センター等)が管理するデータ	
利用者	FINDATAの審査によって許可された利用者	
利用環境	OS	Linux、Windows
	プログラミング言語	Golang、Python、R
	IDE(統合開発環境)	Anaconda3、RStudio
	統計解析ツール	SAS、SPSS、Stata
	その他、使用できるツール	- Acrobat Reader DC 7-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) - Git - Google 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) - LibreOffice - LyX - MiKTeX - Nextflow - WinPython

ログイン画面

FINDATA

Findata – Finnish Social and Health Data Permit Authority

Welcome to Findata's e-service!

By logging in to the service, you can

- Apply for a data permit for individual-level data
- Make a data request for statistical data
- Apply for changes to previous, currently valid permits
- Review the applications you have submitted and the permit decisions you have received, as well as their attachments
- Register as a Nextcloud transfer service user
- Order a Kapseli remote access environment and make changes to your order

See technical releases related to e.g. system faults and maintenance outages here: <https://www.findata.fi/en/login/>

Please always log in using the same authentication method to see the applications you have submitted and the decisions you have received.

Login

アメリカのVisiting解析環境 (CMS VRDC:Virtual Research Data Center)について

未定稿

CMS VRDC(Virtual Research Data Center)はCMSが保有するデータへのアクセスを許可された研究者が安全にデータにアクセスするための仮想環境である

概要	CMSが保有するデータへのアクセスを許可された研究者が安全にデータにアクセスするための仮想環境	
Visiting解析環境のメリット	• CMSのプライバシーおよびセキュリティの要件を全て満たしている • Research Identifiable Files(RIF)にタイムリーにアクセスできる • コスト効率が高い	
提供データ	• CMSが保有するデータ ただし、ダウンロードできるのは統計情報のみで、個人を特定できる情報や保護された健康情報はダウンロードできない	
利用者	DUAを結び、データへのアクセスを許可されたアメリカに居住している研究者	
利用環境	OS	• Windows
	プログラミング言語	• Python、R、SQL
	IDE(統合開発環境)	• 不明
	統計解析ツール	• Databricks*、SAS、Stata ※ Databricksはクラウド上でデータ保存、データ統合、データ分析、AI活用できる統合分析プラットフォーム
	ユーザー側で必要な環境、ツール	• ブラウザ Edge、Firefox、Chrome • VMware Horizon Client 仮想デスクトップ化(VDI化)するためのツール



【参考】 Research Identifiable Filesについて

未定稿

CMS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

利用可能データ		Public Use Files (非識別情報)	Limited Data Set File (リミテッドデータセット)	Research Identifiable Files (研究用特定ファイル)
概要		編集され、個人を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	DUA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

医療情報の標準化動向（英国、フランス）

仕様およびコードの説明			英国	フランス	米国
医療情報標準化の制定・管理組織			- 保健福祉省 - NHS - Professional Record Standards Body (PRSB)※	- ANS(デジタルヘルス庁) —2019年に設立(以前は保健省が管理) - ANSM(国立医薬品健康製品安全性局)	- ANSI(米国規格協会)がSDO(認定標準化団体)を通じて実施 - 医療情報分野のSDO: HL7、NCPDP(全国処方薬プログラム評議会)、CDISC(臨床データ交換標準コンソーシアム)等
標準仕様	HL7 FHIR	Web通信の一般的技術であるRESTを使用して、可読性が高く取り扱いがし易いJSON/XML形式のデータの集合をやり取りする医療情報交換標準規格	○	○	○
	HL7 CDA	電子カルテ等の診療文書を電子的に交換する際の主としてXMLによる表現を定めた医療情報交換標準規格	○	○	○
	IHEプロファイル	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	○	○	○
病名	ICD10、ICD11	WHOが管理する国際疾病分類	○	○	○
	ICPC-2	プライマリケア国際分類第2版	—	○	—
検査	LOINC	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	○	○	○
	DICOM	医用画像の国際規格	○	—	○
薬剤	ATC	WHOが管理する薬剤の解剖治療化学分類法	—	○	—
	EDQM	欧州医薬品品質部門)標準用語	—	○	—
他	SNOMED CT	医療分野で最も包括的な臨床用語集	○	—	○
その他国内基準			dm+d(NHS Dictionary of Medicines and Devices) NHSによって使用されている医薬品と医療機器の説明とコードを含む辞書(医薬品はSNOMED CTと連携)	- CCAM(医療行為コード) - ADICAP(細胞学および解剖病理学) - CIS、CIP(薬剤製品の国家識別コード)	- USCDI(相互運用可能な医療情報交換の標準化されたデータ要素のセット) - C-CDA(医療情報交換に最も利用されている臨床文書アーキテクチャ) - NCPDP(薬剤情報の標準コード)

※PRSB: 独立した非営利団体(専門職記録標準団体)であり、医療およびケア記録の臨床標準を開発するために2013年に設立

EU：電子医療記録交換フォーマットに関する推奨事項（1/2）

EU内での電子医療情報の交換を行うため、欧州委員会は交換フォーマットの推奨事項として、「原則」と「相互運用性仕様」を公開している

原則

- a. **市民中心の設計**：国民は制度設計の中心となるべきである。デフォルトでGDPRの要件を満たすように設計される必要がある
- a. **包括性と機械可読性**：電子医療記録はEU全体の医療及びケアサービスをサポートするために包括的である必要がある。また、電子医療記録に導入される健康データは、それらのデータの合理的な意図された再利用に必要な範囲で機械可読である必要があり、構造化されるべきである
- b. **データ保護と機密保持**：電子医療記録システムと相互運用性ソリューションは、個人の健康データの機密性を保証し、設計段階からデータ保護法のあらゆる側面に準拠する必要がある
- c. **同意またはその他の法的根拠**：健康データの処理は、GDPR第6条および第9条に従って、関係する国民の明示的な同意またはその他の法的根拠に基づいていなければならない
- d. **監査**：健康データの処理はすべて、電子記録へのアクセス、その交換、またはその他の処理操作の正確な記録を保持するために、ロギング（データを記録すること）や監査追跡などの適切な技術を使用して監査目的で登録および検証される必要がある
- e. **セキュリティ**：GDPRおよびNIS(ネットワークおよび情報システム)指令に従って、電子医療記録システムの安全性を確保するために、適切な技術的および組織的対策を実施する必要がある
- f. **識別と認証**：国民電子身分証明書「eID※」の使用は、国境を越えた国民の識別と認証をサポートし、健康データに完全な安全性と利便性でアクセスできるようにするだけでなく、そのようなデータの出所と完全性を保証する「否認防止」の原則をサポートする。EU電子識別規則 No 910/2014第6条に従い、電子識別保証を必要とするオンライン公共サービスは、他の加盟国の通知された電子識別スキームを受け入れなければならない
- g. **サービスの持続性**：電子医療記録交換サービスの継続性と可用性は、治療の継続性を保証するために不可欠である。サービスの使用中に発生する可能性のあるインシデントや中断には、定義された事業継続計画に従って速やかに対処する必要がある

※eID：国境を越えた国家電子識別スキーム (eID) の相互認識を可能にするために欧州委員会によって提供される一連のサービス。これにより、欧州国民は他の欧州諸国からオンライン サービスにアクセスする際に自国のeIDを使用できるようになる

EU：電子医療記録交換フォーマットに関する推奨事項（2/2）

EU内での電子医療情報の交換を行うため、欧州委員会は交換フォーマットの推奨事項として、「原則」と「相互運用性仕様」を公開している

相互運用性仕様

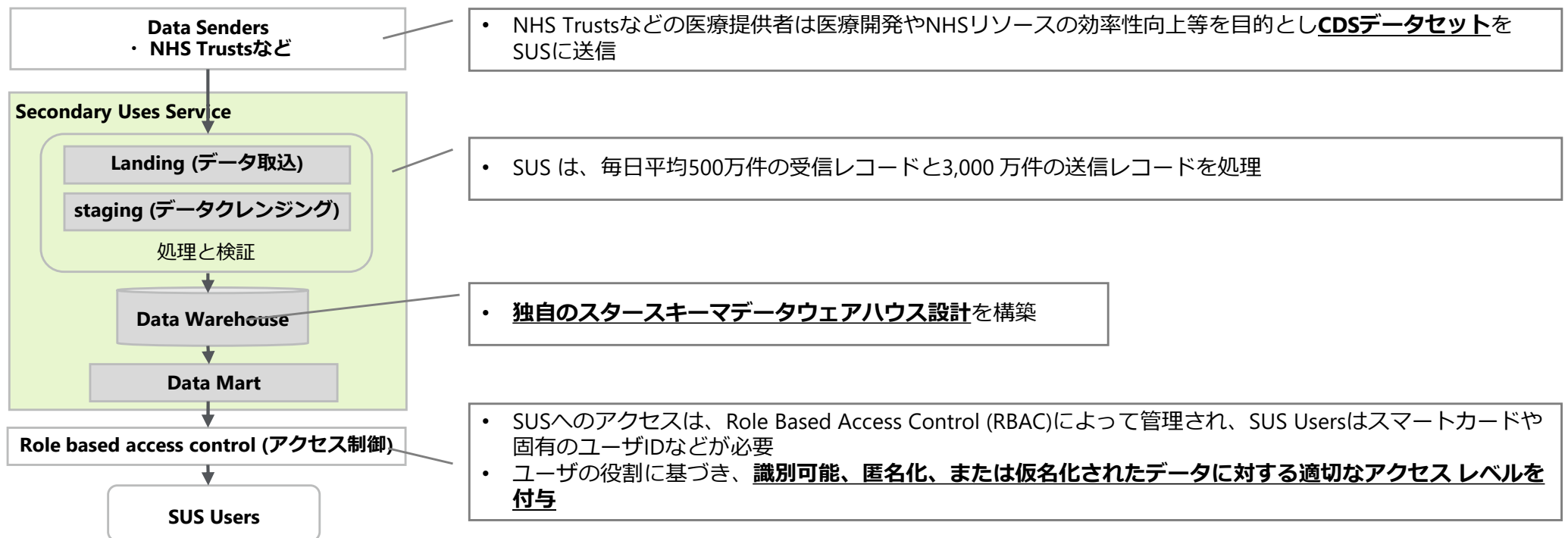
分類	推奨される仕様/コード
Patient Summary (患者概要)	✓ HL7 CDA Release 2 ✓ Level 3 and Level 1(PDF 7/A)
ePrescription / eDispensation (電子処方箋/電子調剤)	✓ HL7 CDA Release 2 ✓ Level 3 and Level 1(PDF 7/A)
Laboratory results (検査結果)	【検査結果、医療画像レポート、退院レポート用】 ✓ HL7 CDA Release 2 ✓ Level 3 and Level 1(PDF 7/A) 【医用画像用】 ✓ DICOM
Medical imaging and reports (医用画像とレポート)	
Hospital discharge reports (退院報告書)	

英国：二次利用の特徴（全体像）

英国では患者データを収集し二次利用を行うSecondary Uses Servicesを導入することで、データの二次利用を促進している

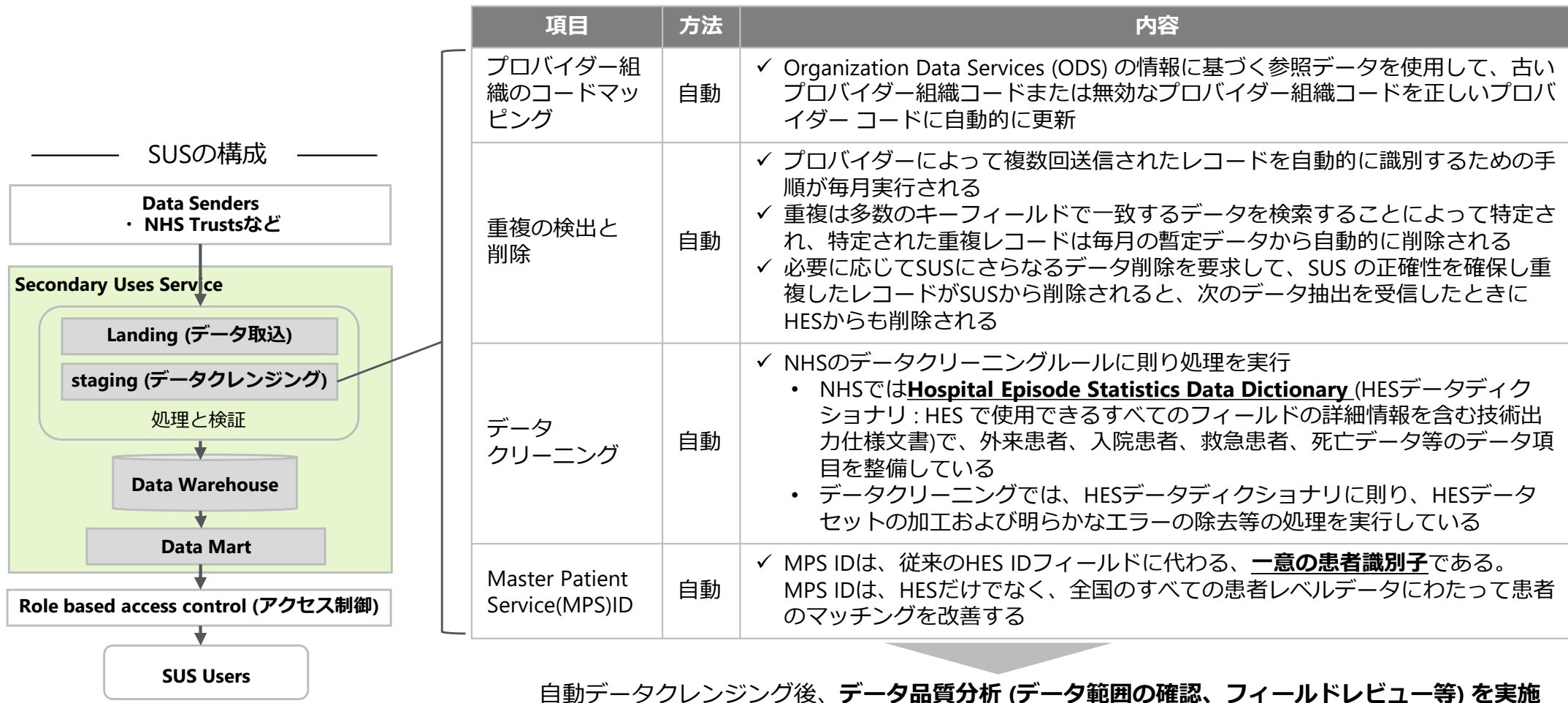
- ✓ Secondary Uses Services (SUS) は、**NHSが提供する医療データの二次利用サービス**であり、患者データを医学研究、統計分析、公衆衛生、国家政策策定などに活用することを目的としている
- ✓ 英国における医療データの単一の包括的なリポジトリであり、NHS による医療サービスの提供をサポートするための様々なレポートと分析が可能

SUSの構成



英国：二次利用の特徴（データクレンジング・検証）

NHSのSecondary Uses Servicesでは、自動でのデータ処理や組織および患者IDとのマッピングサービスを実施している



Organization Data Services (ODS)：NHSとやり取りする外部組織向けに、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

HES：Hospital Episode Statisticsの略で外来や入院等の病院データセット