

2021 年度成果報告書

戦略的イノベーション創造プログラム(SIP)第2期／ビッグデータ・AIを活用したサイバー空間基盤技術／分野間データ連携基盤技術に関する国際動向調査

2021 年 12 月

国立研究開発法人新エネルギー・産業技術総合開発機構

委託先:PwC コンサルティング合同会社

目次

目次	2
1.研究開発の成果と達成状況	3
1.1 要約	3
(1) 和文要約	3
(2) 英文要約	5
1.2 本文	7
1.2.1. 本調査の概要	7
1.2.1.1. 本調査の背景/目的	7
1.2.1.2. 本調査におけるアプローチ	9
1.2.2. 調査結果サマリ	11
1.2.2.1. 調査方法	11
1.2.2.2. 調査観点	12
1.2.2.2.1 技術要素調査(一次調査)	12
1.2.2.2.2 事業形態調査(一次調査)	18
1.2.2.3. グローバルベンチマーク調査の全体像	21
1.2.2.4. グローバルベンチマーク調査(事前調査)	21
1.2.2.5. グローバルベンチマーク調査の全体像(一次調査)	22
1.2.2.6. 一次調査対象の51団体の分類結果	28
1.2.2.7. グローバルベンチマーク調査の全体像(深掘調査)	30
1.2.2.8. 深掘調査結果を踏まえた概念アーキテクチャ分類	36
1.2.3. 国際動向を踏まえた国内データ連携基盤の立ち位置	37
1.2.3.1. CADDE(仮称)の機能要素	37
1.2.3.2. 日本にとって参考となりうる事例の調査	40
1.2.3.3. 競合となりうる団体の検討	44
1.2.3.4. 深掘調査対象との連携課題	45
1.2.4. 本調査の総括・まとめ	47
別紙	49
(別紙1) 事前調査結果のロングリスト	49
(別紙2) 一次調査対象の51団体調査結果	52
(別紙3) 一次調査対象の51団体分類結果	56
(別紙4) CADDE(仮称)の国際比較結果	58
(別紙5) 各団体調査結果詳細	63
参考文献	113

1.研究開発の成果と達成状況

1.1 要約

(1) 和文要約

総合科学技術・イノベーション会議の SIP 第 2 期「ビッグデータ・AI を活用したサイバー空間基盤技術」における「分野間データ連携基盤技術」では、分野を越えたデータ基盤連携の実現 (DATA-EX) を掲げ、日本が海外各国の急速な整備対応に遅れをとることなく連携し、国際競争力を維持した安定かつ信頼性のある持続的な運営を目標としている。

上記目標の実現に向けて、海外データ連携基盤の取組について、事前調査、一次調査、深掘調査の 3 段階の調査を実施した。

- 事前調査では、PwC がリストアップした海外データ連携基盤 120 団体に GAFAM やデータ連携基盤関連の各領域リーディングカンパニーを加えた、計 174 団体の簡易調査を実施した。
- 一次調査では、事前調査対象であった 174 団体の内、Data を提供する Provider を中心に計 51 団体を一次調査対象として選定し、アーキテクチャ (集中型、分散連邦型、ビジネスモデル (収益性)、政府の関与度、コア技術 (認証・認可、ID など) の観点で、デスクトップ調査を実施した。
- 深掘調査では、一次調査の対象であった 51 団体の内、① メタデータ基盤、② GAFAM、③ インド/中国の観点で、計 10 団体を深掘対象として選定し、組織目的、技術要素、コンプライアンス、事業形態・政府関与度、連携条件の観点で、エキスパートインタビュー含め深掘調査を実施した。

上記 3 段階の調査の結果として、CADDE (仮称) の国際比較、日本にとっての参考事例、競合可能性とその要素、日本との連携課題の 4 つを整理した。

- CADDE (仮称) の国際比較
 - 海外団体が既にデータ取込～管理までの一連の機能をサービス提供できている一方で、CADDE (仮称) は既存のデータ連携基盤同士を連携する思想であり、現状ユーザ側で機能実装することを要求している。既に機能実装も完了している **Data Catalog** や **Privacy** 関連の機能については海外団体と同水準と見受けられる。
- 日本にとっての参考事例
 - プラットフォーム利用者拡大の為、データストレージとサードパーティサービスの導入を参考事例として調査した。
- 競合可能性
 - **GAIA-X/IDSA/HL7 FHIR** は、グローバル範囲でサービス提供しており他団体との連携意思が強いため、優先的に協調していくべき対象といえる。**Google/Amazon/Microsoft** は、CADDE (仮称) との重複部分が多く連携意思が薄いため、競合対象として注視していくべき対象といえる。ただし、技術的には、活用の可能性が存在するとみられる。
- 日本との連携課題

- 制度面では、主に他団体より CADDE(仮称)のデータ提供者やデータ自体への追加審査が必要な課題があると分かった。技術面では、主にデータ・メタデータ基準や連携コネクタ整備が必要な課題があるとわかった。

(2) 英文要約

As a background to this theme, **Ver 1.0 of "CADDE (tentative name)" was released in 2020**. CADDE (tentative name) is part of the "Technology for cross-sectional data federation" among the "Big-Data and AI-Enabled Cyberspace Technologies", which is one of the themes listed in the second term of Cross-ministerial Strategic Innovation Promotion Program (SIP), which is part of the in the Council for Science, Technology and Innovation.

From now on, while Ver 3.0 of CADDE (tentative name) is planned to be released in 2022, cross-sectional data federation is planned (DATA-EX), so as to **achieve data federation which keeps up with the rapid development of foreign data federation services**, and realize stable, trustworthy and sustainable operation, while maintaining international competitiveness

In this project, to realize the above objectives, we conducted research on foreign data federation services in the following three steps; preliminary research, primary research, and deep-dive of each services.

- The preliminary research
 - we conducted **a quick research on 174 organizations**, which consists of 120 foreign data federation services (listed by PwC), along with leading companies of data federation related fields (GAFAM etc.)
- In the primary research, among the 174 organizations, 51 organizations, which mainly consists of Data Providers, were selected. We conducted **desktop research**, mainly based on the following viewpoints; architecture (centralized/ decentralized), business model, government involvement, core technology (certification/ identification, ID etc.).
- In the deep-dive, among the 51 organizations, 10 organizations were selected, based on the following three criteria; A) "Meta" data federation service, B) GAFAM services, C) India or China. We conducted a deeper research, including expert interviews, based on the following viewpoints; organizational objectives, technical components, compliance, business form/ governmental involvement, requirements for cooperation.

As a result of research, we summarized the following four points; international comparison of CADDE (tentative name), potential competitors and its elements, challenges in cooperation with Japan, and case studies.

- International comparison of CADDE (tentative name)
 - While foreign organizations already started providing services from data ingestion ~ management, **CADDE (tentative name) currently requires the functionalities to be implemented by each participant, because the architecture based on the connecting existent data providers. Data Catalog and Privacy-related functionalities, which have been implemented, seem to have features equivalent to foreign services.**
- Reference cases for Japan
 - **In order to expand platform users**, we conducted research on data storage and the third party service introductions.
- Competitiveness
 - **GAIA-X/IDSA/HL7 FHIR** are considered as **potential partners**, as they provide services globally, and have strong intent on collaboration with

external organizations. **Google/Amazon/Microsoft** are considered as **potential competitors to monitor**, as they have multiple functional overlaps with CADDE (tentative name), with low interest in collaboration. Still, in terms of technology, there seems to be a margin for cooperation.

- Issues of cooperation with Japan
 - From institutional aspects, challenges mainly lie in CADDE (tentative name)'s **additional examination on data providers/ data itself**, rather than challenges in other organizations. From technical aspects, challenges lie mainly in **preparing standards on data / metadata, federation connectors**.

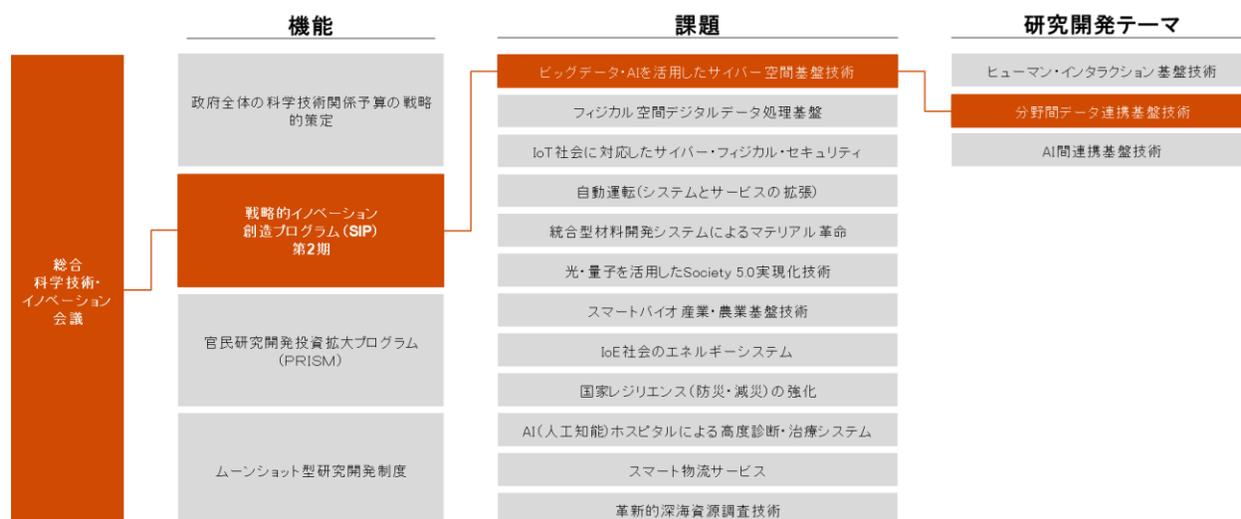
1.2 本文

1.2.1. 本調査の概要

総合科学技術・イノベーション会議の SIP 第 2 期「ビッグデータ・AI を活用したサイバー空間基盤技術」における「分野間データ連携基盤技術」に係る調査を実施。国内で分野を越えたデータ基盤連携(DATA-EX)の実現に向け、海外のデータ連携基盤を調査した。調査は事前調査、一次調査、深掘調査の 3 段階実施し、海外団体との比較や連携可能性などを整理した。

1.2.1.1. 本調査の背景/目的

本調査は、総合科学技術・イノベーション会議における戦略的イノベーション創造プログラム (SIP) の第 2 期で選定された、12 課題中の「ビッグデータ・AI を活用したサイバー空間基盤技術」に関する研究開発テーマ「分野間データ連携基盤技術」に係る調査という位置づけとなる。



(図 1.) SIP 第 2 期における本テーマの位置づけ

本調査の目的¹²は、現時点での海外の分野間データ連携基盤動向を踏まえつつ、今後の海外との分野間データ連携を国内で実現するための方向性や進め方を取りまとめることである。

背景	目的
<ul style="list-style-type: none"> • ここ数年のビッグデータとAI技術の利活用は、ネット上の膨大なデータを囲い込み、その利活用で成功した米国のIT企業が主導。現在では、海外の企業等による激しい覇権争いが繰り広げられ、我が国は研究論文数やビジネスへの導入等で後れを取っている状況¹⁾ • 総合科学技術・イノベーション会議が、社会的に不可欠で、日本の経済・産業競争力にとって重要な課題として、ビッグデータ・AIを活用したサイバー空間基盤技術を選定し、その中から産業を選定し、その中、産業を横断して情報連携させる分野間データ連携基盤技術を研究テーマとして設定 <ul style="list-style-type: none"> ➢ 総合科学技術・イノベーション会議²⁾ <ul style="list-style-type: none"> 平成13年1月、内閣府設置法に基づき、「重要政策に関する会議」の一つとして内閣府に設置。我が国全体の科学技術を俯瞰し、各省より一段高い立場から、総合的・基本的な科学技術政策の企画立案及び総合調整を検討 ➢ 戦略的イノベーション創造プログラムSIP²⁾ <ul style="list-style-type: none"> 総合科学技術・イノベーション会議が府省・分野の枠を超えて自ら予算配分して、基礎研究から出口(実用化・事業化)までを見据えた取組を推進するためのプログラム 生産性革命への貢献等を目指し、生産性の抜本的向上に必要な分野(農業、物流等)を含む12課題をSIP第2期の検討対象として選定 ➢ ビッグデータ・AIを活用したサイバー空間基盤技術²⁾ <ul style="list-style-type: none"> SIP第2期の課題として選定 本分野における国際競争力を維持・強化するため、世界最先端の、実空間における言語情報と非言語情報の融合によるヒューマン・インタラクション基盤技術(感性・認知技術開発等)、分野間データ連携基盤技術、AI間連携基盤技術を確認し、社会実装 ➢ 分野間データ連携基盤技術²⁾ <ul style="list-style-type: none"> ビッグデータ・AIを活用したサイバー空間基盤技術の課題における研究テーマの1つとして、分野を越えたデータ共有と利活用のための技術開発とプラットフォーム整備 	<ul style="list-style-type: none"> • ビッグデータと人工知能(AI)の活用から生まれたイノベーションにより、新たなサービスやビジネスモデルが誕生し、様々な分野で新たな価値が創出され、経済社会システムのパラダイムシフトが起こること <ul style="list-style-type: none"> ➢ ビッグデータ・AIを活用したサイバー空間基盤 <ul style="list-style-type: none"> 本課題において、関連の基盤技術を確認し、生産性(作業時間・習熟速度等)を10%以上向上させる実用化例を20以上創出 ➢ 分野間データ連携基盤技術 <ul style="list-style-type: none"> 産官学でバラバラに保有するデータを連携し、AIにより活用可能なビッグデータとして供給するプラットフォームである「分野間データ連携基盤」を、3年以内に整備し、5年以内に本格稼働させ、実用化例を創出 本案件「分野間データ連携基盤技術」に係る調査目標は以下: <ul style="list-style-type: none"> ✓ 介護、教育等のAIによるデータ利活用に必要な海外との分野間データ連携基盤の実現に向けて、海外にてデータ連携に関わる活動、特に政府主導での活動団体は積極的に抽出 ✓ データ連携活動における重要性を踏まえ深堀すべき団体を選定し、具体的なデータ連携に関わる活動内容、成功事例やユースケースを調査 ✓ 得られた調査結果より、本件、SIP第2期/ビッグデータ・AIを活用したサイバー空間基盤技術プロジェクト、研究開発内容、活動内容の妥当性を分析 ✓ あるべき分野間データ連携の姿を描くと共に、出口戦略も踏まえた海外との分野間データ連携実現に向けた今後の方向性や進め方、実現に向けての規模/コスト感を取りまとめ

1) NEDO「戦略的イノベーション創造プログラム(SIP)ビッグデータ・AIを活用したサイバー空間基盤技術研究開発計画」より抜粋
 2) 内閣府「戦略的イノベーション創造プログラム(SIP)概要」より抜粋

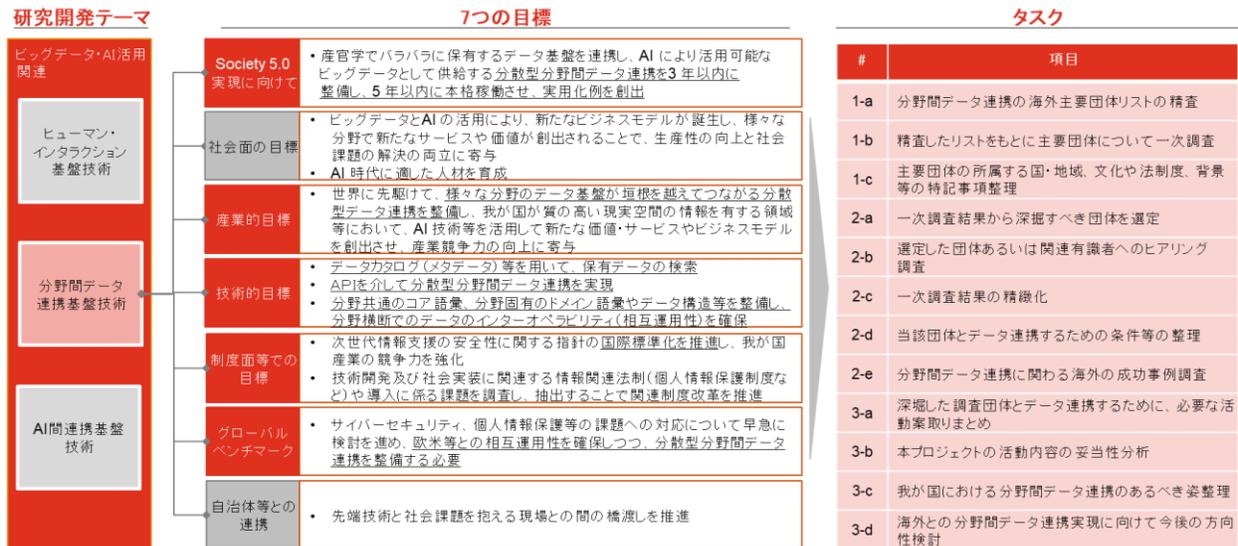
(図 2.) 本調査の背景・目的

¹ NEDO「戦略的イノベーション創造プログラム(SIP)第2期/ビッグデータ・AIを活用したサイバー空間基盤技術/分野間データ連携基盤技術に関する国際動向調査」(2021年) (<https://www.nedo.go.jp/content/100931755.pdf>)

² 内閣府「戦略的イノベーション創造プログラム(SIP)概要」(2021年) (<https://www8.cao.go.jp/cstp/gaiyo/sip/sipgaiyou.pdf>)

1.2.1.2. 本調査におけるアプローチ

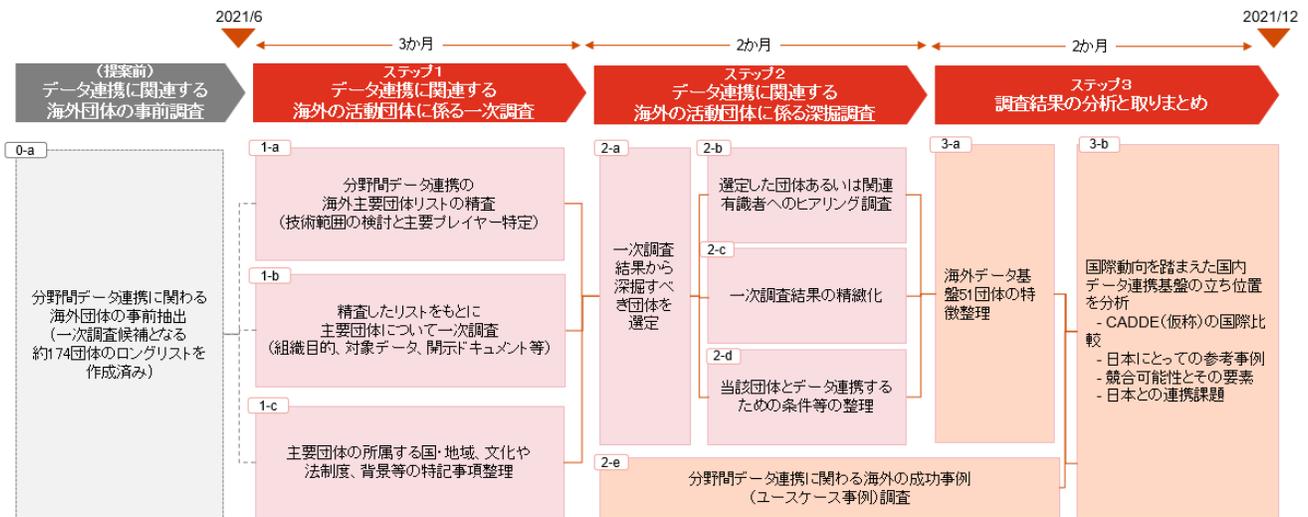
本テーマにて実現したい7つの目標をうけて、本調査における実施事項を整理した。



(図 3.) 本テーマにおける目標と本調査における実施事項

PwC における調査アプローチは次の通りである。

- ①174 の候補団体に対して一次調査/深堀調査による詳細把握
- ②PwC グローバル知見を活用した海外成功事例調査
- ③国際動向を踏まえた CADDE (仮称) の立ち位置の分析



(図 4.) PwC における調査アプローチ

参考として、調査タスクの各ステップにおけるタスクの詳細や、必要インプット情報、成果物/アウトプットを紹介する。

(表 1.) 各ステップの詳細 (ステップ 1)

タスク	タスク詳細	必要インプット情報	成果物/アウトプット
分野間データ連携の海外主要団体リストの精査	海外主要団体リストをベースに Web・文献等のオープン情報も必要に応じて追加しつつ、一次調査の対象となる団体を選定	<ul style="list-style-type: none"> ● PwC 保有のデータ連携に関わる海外主要団体リスト ● Web・文献等の公開情報 	一次調査対象団体
精査したリストをもとに主要団体について一次調査	<ul style="list-style-type: none"> ● 一次調査対象団体における、組織としての目的、主な実施者、取扱対象データ開示ドキュメント、活動内容、他機関との連携状況等を調査 ● 一次調査結果として取りまとめ 	<ul style="list-style-type: none"> ● PwC 保有のデータ連携に関わる海外主要団体リスト ● Web・文献等の公開情報 ● PwC の海外におけるデータ関連のアドバイザー実績 ● 一次調査対象団体 	一次調査結果:各団体の組織としての目的、主な実施者、取扱対象データ開示ドキュメント、活動内容、他機関との連携状況等
主要団体の所属する国・地域、文化や法制度、背景等の特記事項整理	<ul style="list-style-type: none"> ● 一次調査対象となった団体の所属している国・地域において、特筆すべき文化や法制度等の背景があれば調査 ● 一次調査結果として取りまとめ 	<ul style="list-style-type: none"> ● PwC 保有のデータ連携に関わる海外主要団体リスト ● Web・文献等の公開情報 ● 一次調査対象団体 	一次調査結果:団体が所属している国・地域における特筆すべき文化や法制度等の背景

(表 2.) 各ステップの詳細 (ステップ 2)

タスク	タスク詳細	必要インプット情報	成果物/アウトプット
一次調査結果から深掘すべき団体を選定	一次調査対象団体について、データ連携に関する具体的な取組内容、取扱データ種類、データ連携分野での影響力の大きさ、関連政策の重要性等から深掘団体を総合的に判断	<ul style="list-style-type: none"> ● 一次調査結果 	深掘調査対象団体
選定した団体あるいは関連有識者へのヒアリング調査	深掘対象団体に対して PwC AI Lab アカデミア有識者や Global ネットワーク等のコネクションを活用し、一次調査にて整理した内容を詳細化するための有識者ヒアリングを実施	<ul style="list-style-type: none"> ● データ連携における有識者知見 ● PwC 保有の Global エキスパートコンタクトリスト ● 深掘調査対象団体 	有識者ヒアリング結果
当該団体とデータ連携するための条件等の整理	有識者ヒアリング結果を踏まえ、深掘調査した団体やヒアリングから得られたその他団体含めデータ連携を実現するための方式、システムや法的制約等の条件を整理	<ul style="list-style-type: none"> ● 有識者ヒアリング結果 	深掘調査結果:有識者ヒアリング結果をベースにデータ連携に関わる詳細取組内容、データ連携実現のための方式、システムや法的制約等の条件
分野間データ連携に関わる海外の成功事例(ユースケース事例)調査	PwC のナレッジを活用して海外での分野間データ連携に関わるプロジェクト事例や成功事例、ユースケース事例を調査	<ul style="list-style-type: none"> ● PwC 保有のナレッジ(プロジェクト、成功実績等) ● 深掘調査対象団体 	深掘調査結果:海外でのプロジェクト事例、成功事例、ユースケース事例

(表 3.) 各ステップの詳細 (ステップ 3)

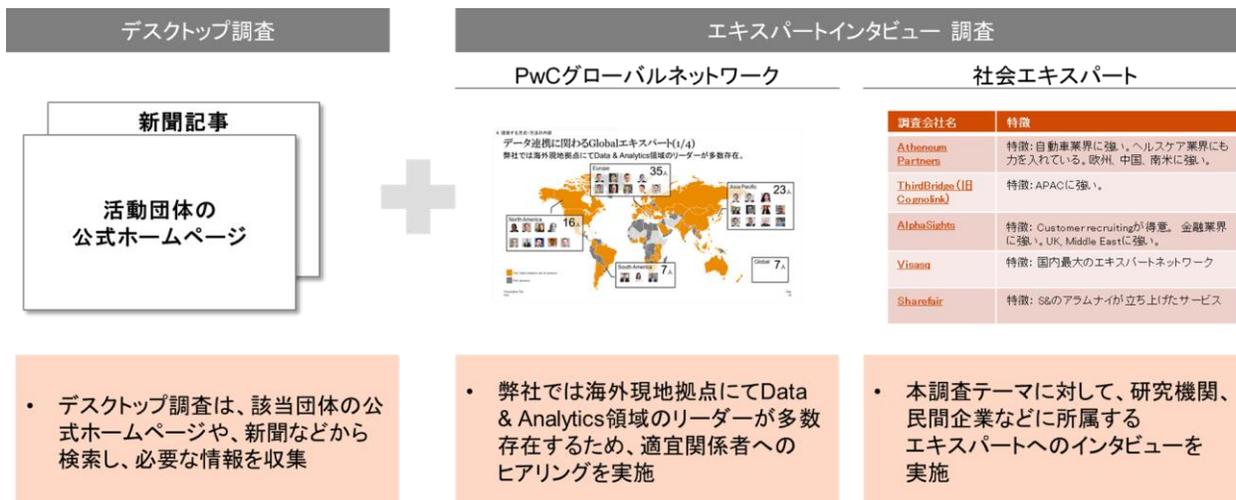
タスク	タスク詳細	必要インプット情報	成果物/アウトプット
海外データ基盤 51 団体の特徴整理	データ関連団体(51 団体)を、サービスモデル、アーキテクチャ、連携対象・データ種別、使用しているテクノロジーの分類軸に沿って分類	一次調査結果	海外データ基盤 51 団体の分類結果
国際動向を踏まえた国内データ連携基盤の立ち位置を分析	CADDE(仮称)の国際比較:中央が仕様や機能を提供する・しないの観点で CADDE(仮称)と他団体を比較することで、CADDE(仮称)が現在開発中の機能状況を把握	深掘調査結果	CADDE(仮称)と深掘調査団体の機能比較
	日本にとっての参考事例:深掘調査団体の取組から CADDE(仮称)の参考となりそうな事例を紹介	深掘調査対象団体	日本にとっての参考事例 分析結果
	競合可能性とその要素:サービス提供内容や提供地域、連携意思から深掘調査団体の CADDE(仮称)との競合可能性を分析	深掘調査対象団体	深掘調査団体の競合可能性 分析結果
	日本との連携課題:CADDE(仮称)が深掘調査団体と連携する際の課題を制度面と技術面から分析	深掘調査対象団体	日本との連携課題 分析結果

1.2.2. 調査結果サマリ

事前調査、一次調査、深掘調査の 3 段階の調査では調査団体を 174 団体→51 団体→10 団体と絞り込んでいった。アーキテクチャや機能要素、政府の関与度、ビジネスモデルを各団体に対し公開情報を基に調査した。そのうえで、深掘調査団体に対しては有識者へのインタビュー等を通じて詳細情報を確認し、CADDE(仮称)が国際競争力を持った基盤となるために参考となる示唆を抽出した。また、51 団体についてはアーキテクチャや採用技術等から任意の分類ができるように整理した。

1.2.2.1. 調査方法

デスクトップ調査とエキスパートインタビューを通じて、調査対象となる活動団体の設立目的などに対して調査を実施した。



(図 5.) 本調査における調査方法

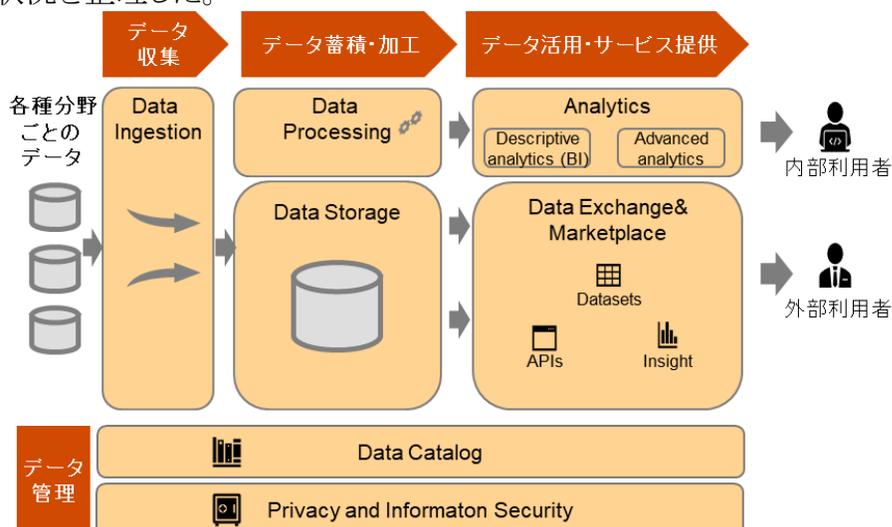
1.2.2.2. 調査観点

一次調査では技術要素と、事業形態の 2 つの観点より調査を実施した。

1.2.2.2.1 技術要素調査(一次調査)

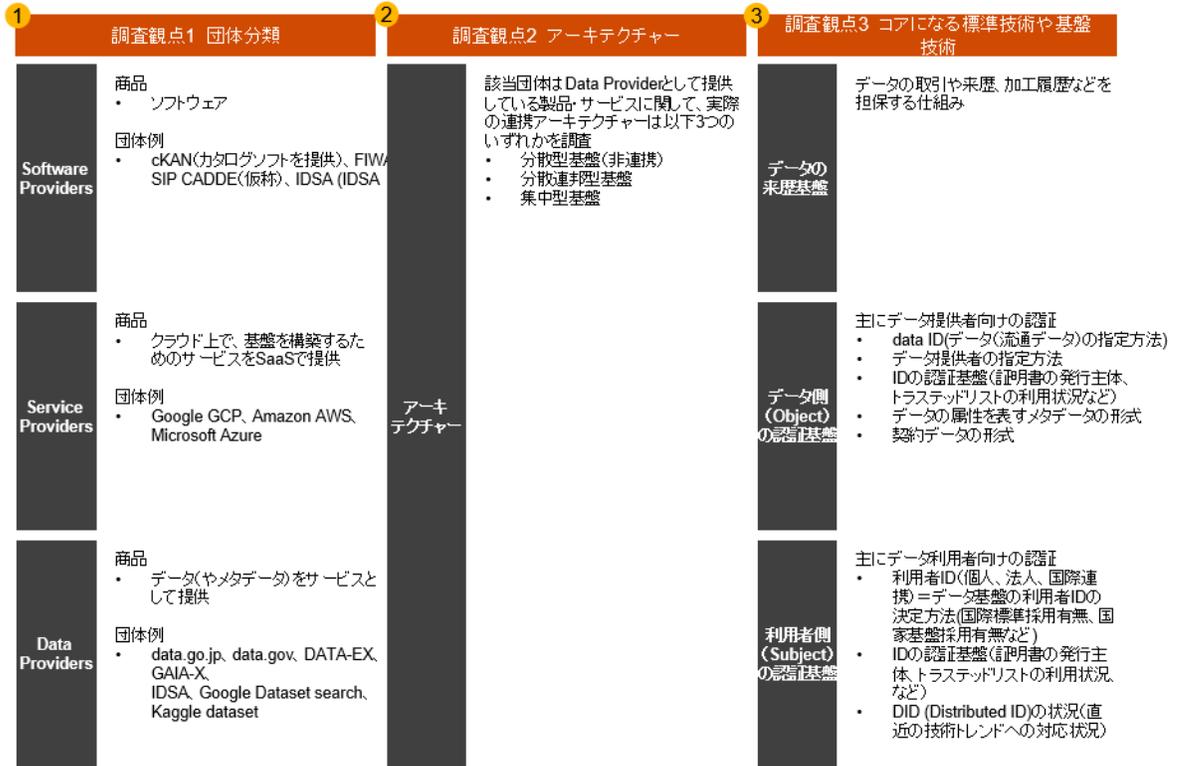
技術要素においては、データ連携基盤に必要な機能要素をベースに、本テーマを担当するサブプログラムディレクター (PD) である東京大学の越塚教授等との打合せを踏まえ、データ連携基盤の技術要素を精査した。

- 機能要素においては、調査対象となる各団体に対して、データ収集⇒データ蓄積・加工⇒データ活用・サービス提供⇒データ管理の一連の流れに必要な機能要素を洗い出し、各調査対象の対応状況を整理した。



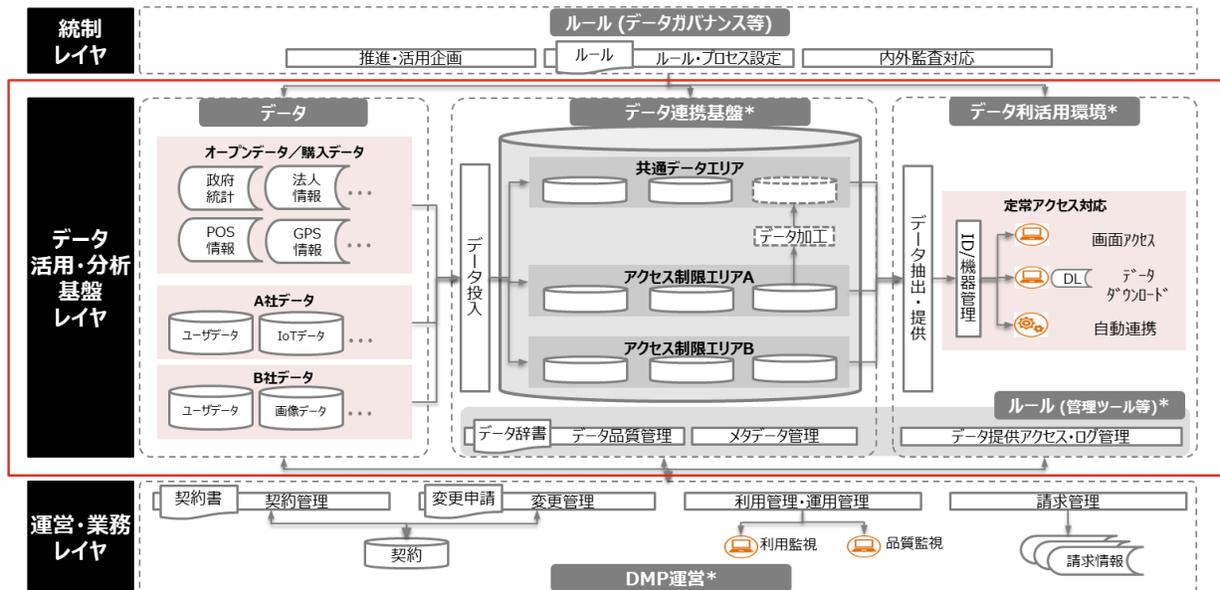
(図 6.) 本調査における対象となる機能要素

- 上記機能要素を調査する観点としては、下記 3 つの観点に沿って調査を実施した。



(図 7.) 機能要素における調査観点

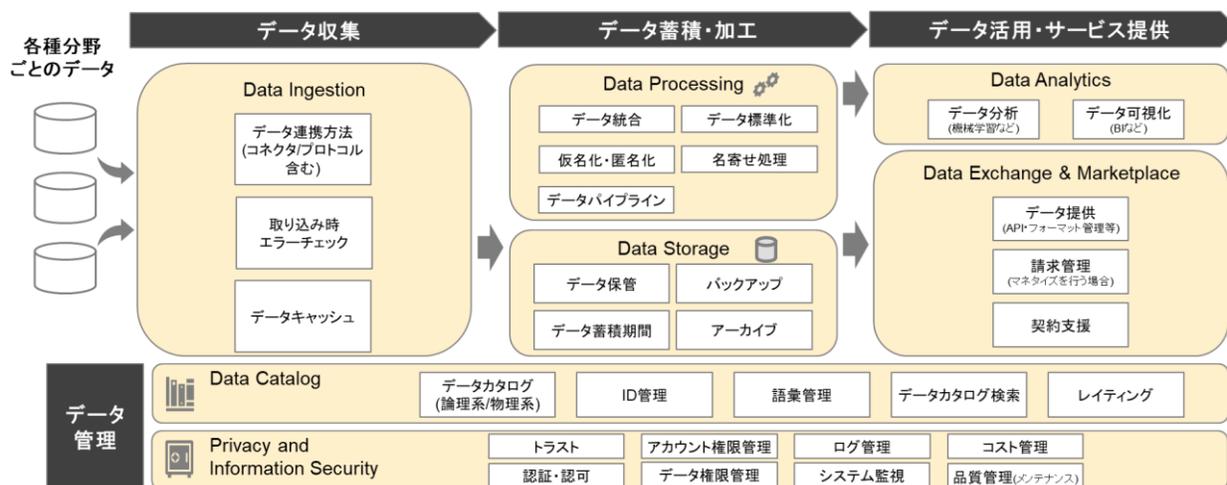
なお、データ連携基盤においては、基盤そのもの以外にも、統制や運営・業務観点も考慮する必要があるが、本調査では、機能面にフォーカスしている。



*がついた項目については、実現のためのインフラ検討が必要
 **上記に加えて検討すべき要素として、人材・セキュリティ、組織・ビジネスなどが考えられる

(図 8.) データ基盤におけるレイヤと本調査における対象範囲

本調査では、Data Ingestion/ Data Processing/ Data Storage/ Data Analytics/ Data Exchange & Marketplace/ Data Catalog/ Privacy and Information Security の 7 機能要素ごとに、対象となるデータ連携基盤の技術要素を定義した。



(図 9.) データ連携基盤の技術要素

図 9.の技術要素の前提となる、各機能要素の定義は以下の通り。

(表 4.) データ基盤の機能ごとの機能要素と定義

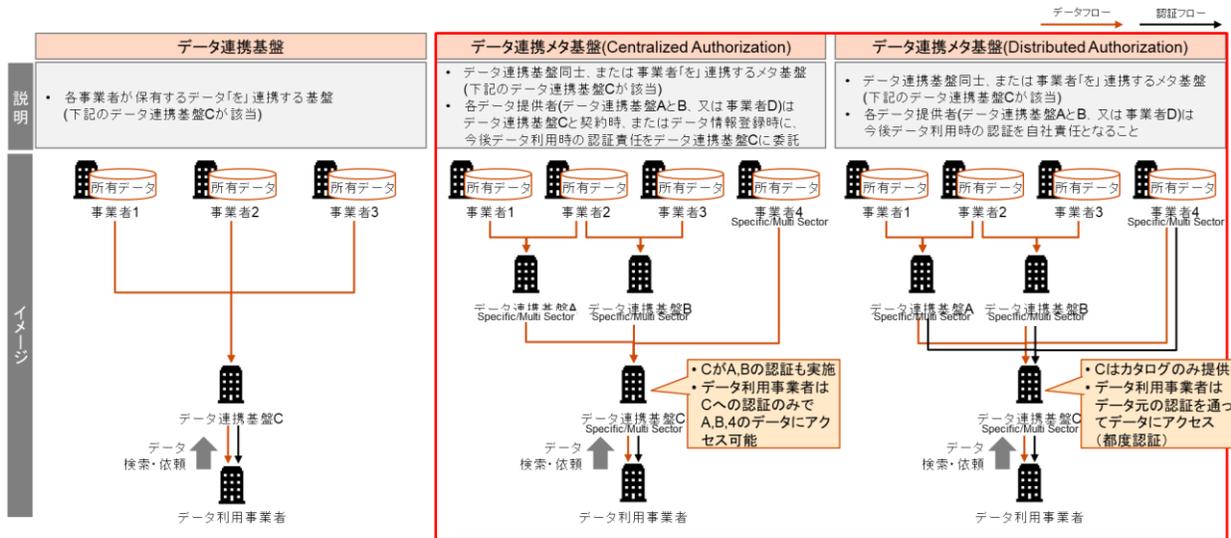
#	機能要素	機能名	定義・概要
1	Data Ingestion	データ連携方法 (コネクタ/プロトコル含む)	データソースとの連携方法、特に連携の際に使用するコネクタ (API) やプロトコルなどをさす。文書データや IoT データ等、取得するデータにより異なる想定。自動取得が不要な場合は手動での取得なども対象
		取り込み時エラーチェック	データ取込時に実施する、ファイル名やスキーマ(csv ファイルなど)等のエラーチェック
		データキャッシュ	一時的に分析に使用する等のデータがアップロードされる場合の処理 (格納先や保存期間、削除方法など)
2	Data Processing	データ統合	多様なデータソースから、一元的にデータベースに格納するための統合処理
		データ標準化	データの標準化や構造化など、データを連携基盤に格納するための変換(規定のスキーマに合わせてデータ変換、文字コード変換、データ数に合わせてデータ内 0 埋め、日付フォーマット統一など)
		名寄せ処理	同一の名称等、同一とみられるデータを同一 ID と紐づける処理をさす。運用の一環として、自動化や、ルール変更に対応していることが望ましい
		仮名化・匿名化	必要に応じた仮名化・匿名加工化などの個人情報の加工
		データパイプライン	データ加工プロセスや、(分析データマート等への)データ出力等、他の機能との連携を行う際の実現方法、特に、連携の際に使用するコネクタ (API) やプロトコルなど
3	Data Storage	データ保管	文書やメディア情報などを構成するビット・バイト情報を保存する機能
		バックアップ	蓄積期間を超えるデータの保管方針やデータベースのバックアップ方針の定義
		データ蓄積期間	データベース上にデータを保持する期間の定義
		アーカイブ	蓄積期間を超えるデータの保管方針やデータベースのアーカイブ方針の定義
4	Data Analytics	データ分析 (機能学習など)	分析基盤上で分析を実施する場合の分析環境など
		データ可視化 (BI など)	分析基盤上で BI によるデータの可視化を実施する場合のツールや保存期間・対象データソースなど
5	Data Exchange / Data Marketplace	データ提供 (API・フォーマット管理等)	データを外部に提供する際の提供方法や、提供する際のフォーマット、実現するためのコネクタ、通信プロトコルなど
		請求管理 (マネタイズを行う場合)	データ活用の目的でデータを抽出するユーザに対して、自動的に請求先・請求額を管理する機能
		契約支援	契約する為のコミュニケーション機能、契約内容を管理する機能

(表 5.) データ基盤の機能ごとの機能要素と定義 (続き)

#	機能要素	機能名	定義・概要
6	Data Catalog	データカタログ (論理系/物理系)	データに関する情報の定義 (論理的/物理的なテーブル構成、データソース、定義、算出方法等) ※ メタデータ管理も本項目に含まれる想定
		ID 管理	データセットの ID 管理
		語彙管理	特に各業界に固有の語彙の管理方法(データカタログのルール定義)
		データカタログ検索	特定の条件に基づいたデータの検索を行う機能
		レイティング	データごとのレイティングを行う機能
7	Privacy and Information Security	トラスト	データの存在証明 (原本性保証含む) や、タイムスタンプ付与により、データの信頼性を担保
		品質管理 (メンテナンス)	定期的なデータの重複性チェックや、エラーチェックを通した、データ品質の担保を行う機能 (過去データにも適用されるルール変更があった場合に、そのルールが正しく反映されているかのチェック等)
		アカウント/権限管理	業務役割に応じたアカウント/権限の管理、ユーザ認証方法など
		データ権限管理	データごとにアクセス可能な権限の設定 (競合である等の理由でデータへのアクセスを制限する場合など)
		ログ管理	ユーザへのデータ提供記録を含む、ログの保管対象・管理方法など
		認証・認可	特定業界のデータを取り扱う際に認証・認可が必要な場合等に考慮が必要
		システム監視	ジョブやインスタンス等の死活監視およびパフォーマンス監視および可視化
		コスト管理	データ連携基盤クラウドを利用する場合の運用コストの管理および可視化

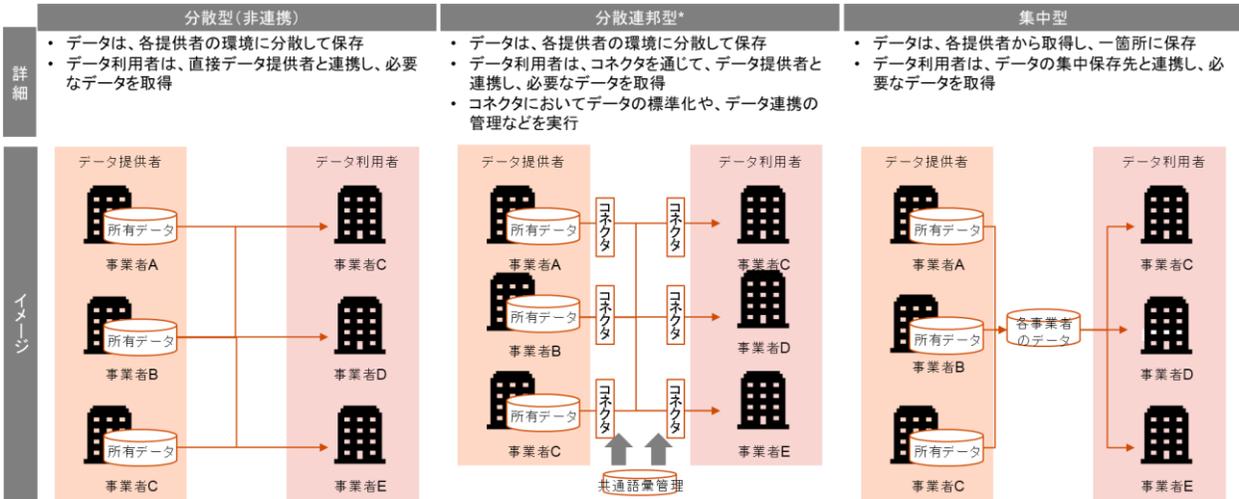
なお、実際はこれらの内容に加え、一般的なシステム同様、ジョブ管理 (各データを効率よく、かつ確実に連携するための管理) の検討も必要となることに留意が必要である。

図7の調査観点にある、①団体の提供する商品の種別(ソフトウェア/SaaS/Data)、②団体のアーキテクチャ(集中型、分散型)、③データの来歴基盤などのコア技術を踏まえ、一次調査を実施した。加えて、データ連携基盤同士を連携するメタ基盤において、利用されている認証技術要素・方法を調査した。



(図 10.) メタ基盤における認証技術の考え方

なお上記②については、データ連携基盤におけるアーキテクチャとして、分散型(非連携)、分散連邦型、集中型データ連携基盤の3種を定義し、この3種類による分類を実施した³。



(図 11.) データ連携基盤におけるアーキテクチャの分類

³DATA-EX「データ基盤システム DATA-EXと海外の取組」(2021年)
(https://www.kantei.go.jp/jp/singi/it2/dgov/data_strategy_tf/dai6/siryou4.pdf)

1.2.2.2.2 事業形態調査(一次調査)

事業形態の調査は、ビジネスモデル、政府の関与度合いの2つの観点から実施した。

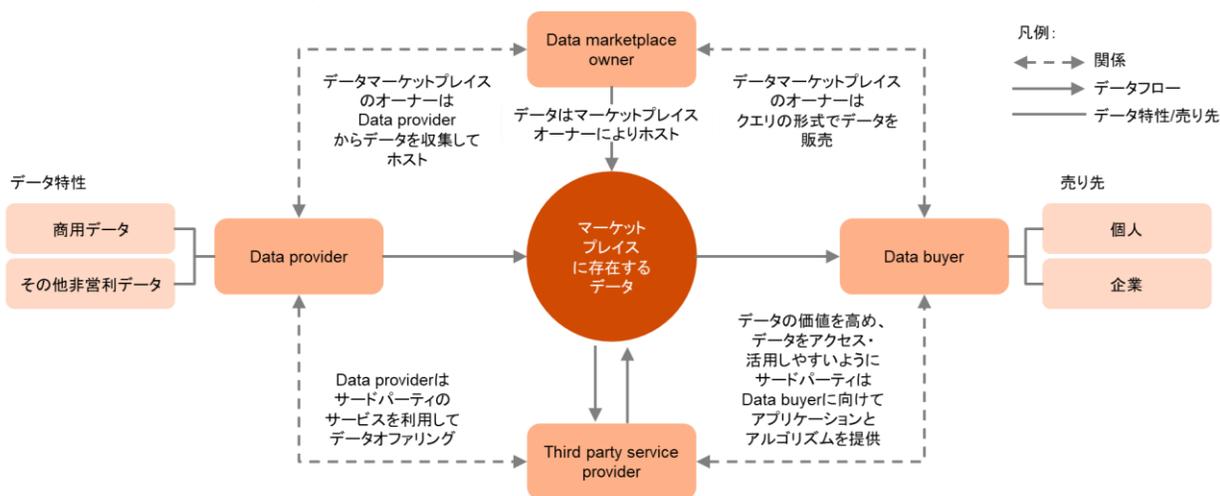
- ビジネスモデル: 該当団体のデータ連携基盤の構築目的の調査
例: データ販売 / 公的支援 / 産業界支援 / 他のビジネスとの垂直統合など
- 政府の関与度合い: 該当データ連携基盤にとって政府の役割などの調査

ビジネスモデルの対応事例として、データ取引市場、PDS&情報銀行、データ担保融資、データ証券化、データ信託の5つを想定した。特に、調査着手時点で情報の少なかったデータ担保融資とデータ証券化について追加調査を実施した。

(表 6.) データ基盤のビジネスモデル

ビジネスモデル	説明	主なプレイヤー	調査優先度
データ取引市場	データ保有者と当該データの活用を希望する者を仲介し、売買等による取引を可能とするプラットフォーム	Snowflake, Dawex	低 (既に多く対応されていて、または日本国内での検討が多い)
PDS & 情報銀行	<ul style="list-style-type: none"> ● PDS(Personal Data Store)は、他者保有データの集約を含め、個人が自らの意思で自らのデータを蓄積・管理するための仕組み ● 情報銀行(情報利用信用銀行)とは、個人とのデータ活用に関する契約等に基づき、PDS等のシステムを活用してデータを第三者に提供する事業 	Datos, Facebook	
データ担保融資	データの所有権を抵当に入れて、銀行やほかの金融機関から融資を取得	中国中関村データアセットイニシアチブプラットフォーム	高 (データ提供者自社で所有権を持つデータを対象として対応する為、データの活用にあたりリスクが低い、かつ海外事例が多い)
データ証券化	<ul style="list-style-type: none"> ● データを資産として適正評価 ● データを証券化商品として、特別目的事業体を通じて投資家へ売却 	デロイト	
データ信託	<ul style="list-style-type: none"> ● データの所有権を外部企業に委託することで、キャッシュと交換 ● データのマネタイゼーションは外部企業による推進 	現状特になし	低 (法規整備が十分でなく現状の進出企業も少ない)

ビジネスモデルについて補足する。データ取引市場のビジネスモデルでは、データの提供者/利用者に加え、取引市場の所有者や、サードパーティーのサービスプロバイダがプレイヤーとして存在し、下図のような関係となる⁴。



(図 12.) データ取引市場のビジネスモデル

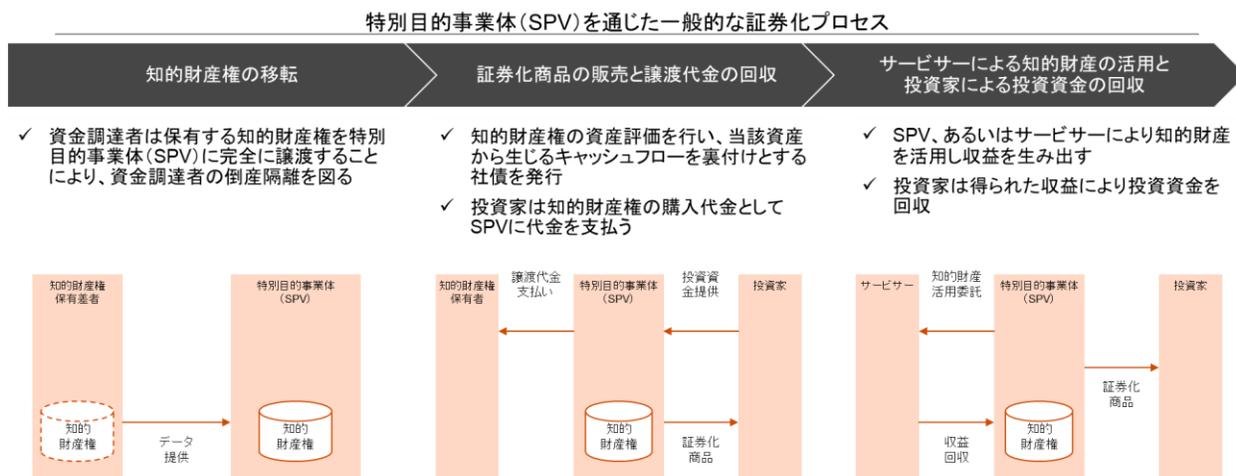
データ取引市場の調査においては、バリュープロポジション (Value proposition)、付加価値 (Value added)、アーキテクチャ (Architecture)、ビジネスモデル (Business Model) の 4 つの観点 を盛り込んで調査した。

メタディメンション	プラットフォームの キーファクター	マーケットプレースの属性	特性
Value Proposition	Function of the platform	Value Proposition	<input checked="" type="checkbox"/> Transaction-centric <input checked="" type="checkbox"/> Data-centric
	Independence of the platform	Market position	<input checked="" type="checkbox"/> Data supplier <input checked="" type="checkbox"/> Neutral
	Openness of the platform	Market Access	<input checked="" type="checkbox"/> Closed <input checked="" type="checkbox"/> Open <input checked="" type="checkbox"/> Hybrid
		Integration	<input checked="" type="checkbox"/> Domain-specific <input checked="" type="checkbox"/> Domain-unspecific
Value added	Dynamic Strategy	Transformation	<input checked="" type="checkbox"/> Raw data <input checked="" type="checkbox"/> Aggregation <input checked="" type="checkbox"/> Normalization <input checked="" type="checkbox"/> Quality Assurance
Architecture		Platform architecture	<input checked="" type="checkbox"/> Central <input checked="" type="checkbox"/> Open <input checked="" type="checkbox"/> Hybrid
Business model	Sales and revenue concepts	Price model	<input checked="" type="checkbox"/> Free <input checked="" type="checkbox"/> Flat rate <input checked="" type="checkbox"/> Freemium <input checked="" type="checkbox"/> Fee
		Revenue model	<input checked="" type="checkbox"/> Listing fee <input checked="" type="checkbox"/> Service fee <input checked="" type="checkbox"/> Transaction fee/commission <input checked="" type="checkbox"/> Storage fee

(図 13.) データ取引市場において考慮すべき観点

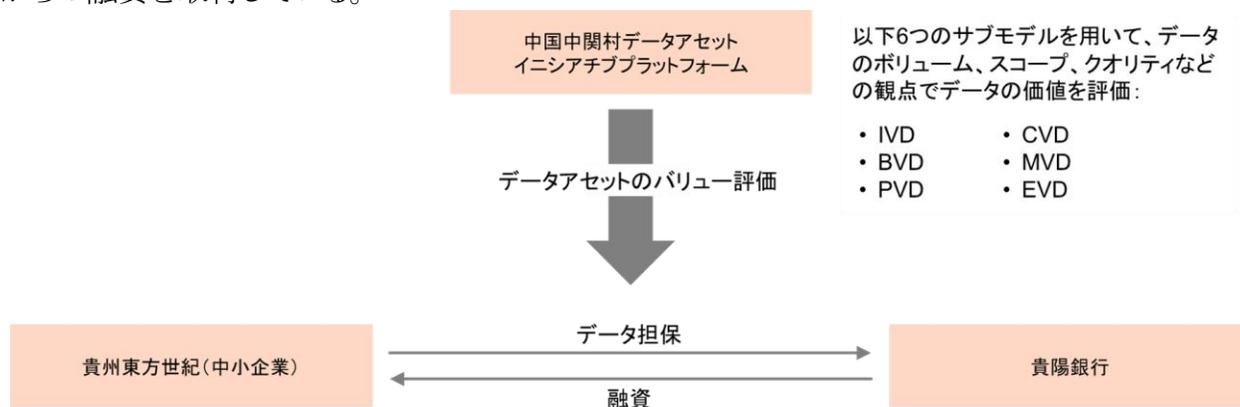
⁴Markus Spiekermann「Data Marketplaces: Trends and Monetisation of Data Goods」(2019年)
 (<https://www.intereconomics.eu/contents/year/2019/number/4/article/data-marketplaces-trends-and-monetisation-of-data-goods.html>)

追加調査したデータ証券化とは、知的財産権を証券商品として資金調達を行う比較的新しいビジネスモデルである。知的財産権の証券化が1997年に米国で初めて行われた事例が確認されている。



(図 14.) 知的財産権の証券化プロセス⁵

データ担保融資 & データ証券化のビジネスモデルの事例として、中国において、中小企業の融資問題を解決する為に、中国中関村データアセットイニシアチブプラットフォームがデータアセットの価値評価モデルを2016年に公開した。本モデルを活用して、貴州東方世紀(中小企業)は貴陽銀行からの融資を取得している。

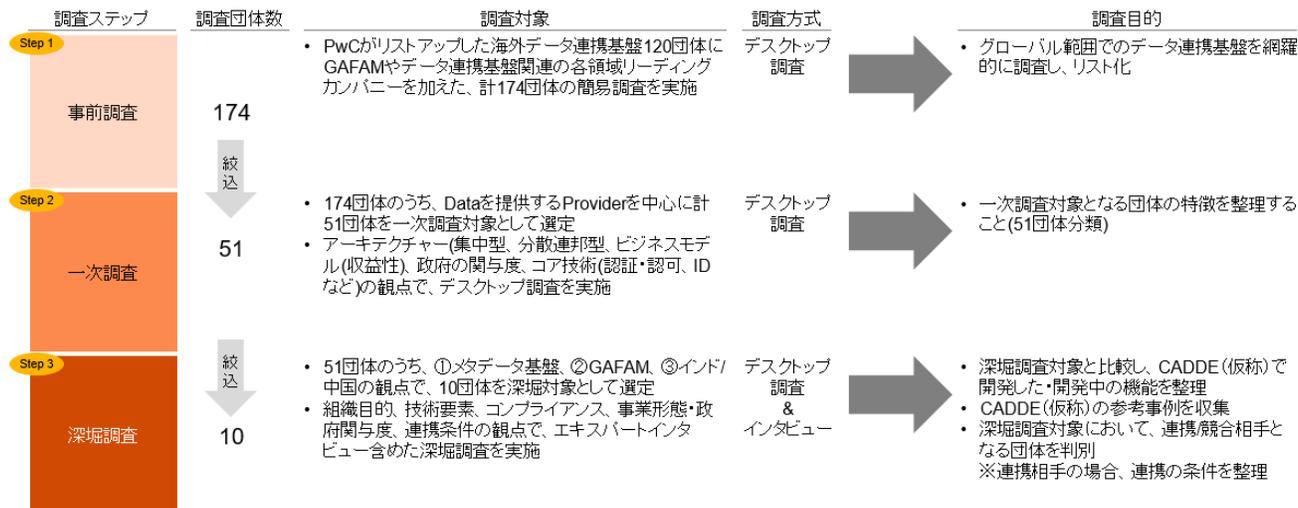


(図 15.) データ担保融資&データ証券化の概念図

⁵流通流動化検討委員会「資金調達手法～証券化を中心に」(2007年)
https://system.jpaa.or.jp/patents_files_old/200704/jpaapatent200704_083-089.pdf

1.2.2.3. グローバルベンチマーク調査の全体像

グローバルベンチマーク調査は事前調査、一次調査、深掘調査の3ステップで実施した。

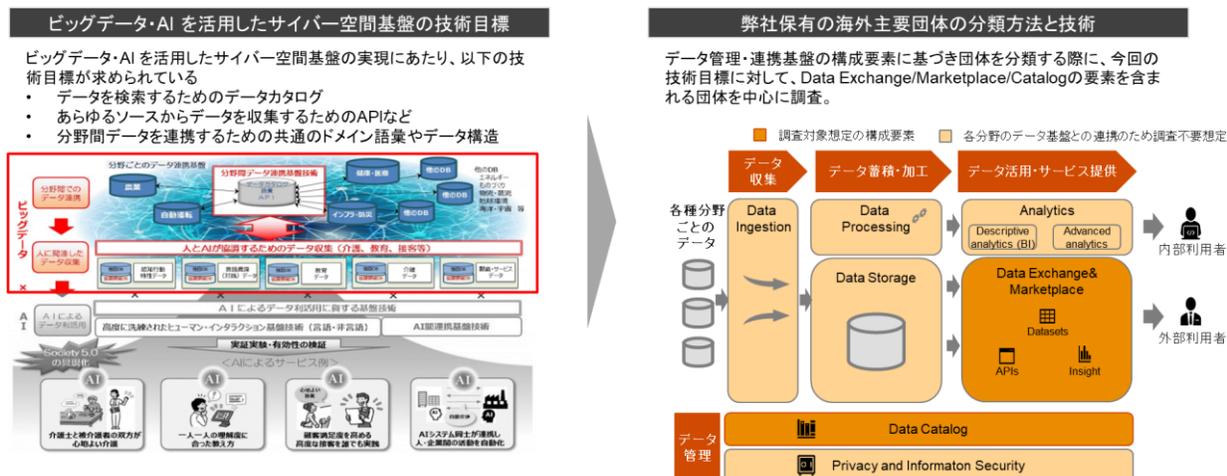


(図 16.) グローバルベンチマーク調査の全体像

1.2.2.4. グローバルベンチマーク調査(事前調査)

事前調査の対象となる活動団体の分類方法を説明する。

- ①データ管理・連携基盤の構成要素に基づき団体をリストアップ
- ②リストアップした団体について、Data Exchange&Marketplace、Data Catalogの機能を中心に調査。



(図 17.) 活動団体分類の考え方⁶

各団体におけるデータ連携基盤に関わる情報は、リスト化して整理した。詳細は、別紙1を参照されたい。

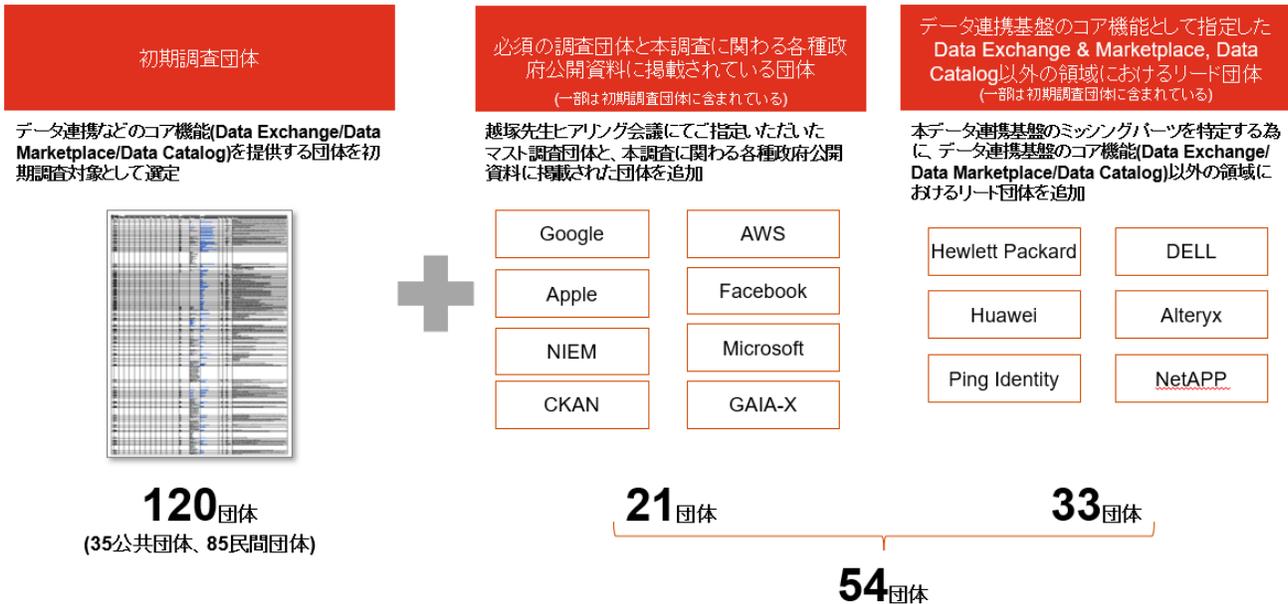
⁶ NEDO「戦略的イノベーション創造プログラム(SIP)第2期/ビッグデータ・AIを活用したサイバー空間基盤技術」(2021年)(https://www.nedo.go.jp/activities/ZZJP_100158.html)

約120団体 (公共35件、民間85件) を対象に初期調査済

団体分類 公共/民間団体	データ基盤構成要素 データ基盤の8つの構成要素に関する保有状況確認					提供内容 サービス/製品/ソリューション		取扱データ Multi / Specific Sector		団体基本情報 団体名・年報発表日・設立日・所属国/地域など									
	データ提供/共有	Data Ingestion	Transformation	Data Storage	Data Storage	Analytics	Data Exchange	Data Marketplace	Data Catalog	Privacy and Information Security	サービス提供	ソリューション	取扱データ分類	取扱データ領域	団体・組織名	年報発表日	設立日	所属国/地域	備考としての目的
1 公共団体													Multi Sector	データ交換	Substantia Asia	-	2015	Asia	Categorical area aims at facilitating an Asian market.
2 公共団体													Specific Sector	Personal Data	Data Exchange	-	2014	Australia	Australian Data Exchange provides a Government or business directly with data.
3 公共団体													Multi Sector	データ交換	Canadian Open Data	-	2014	Canada	Facilitate an open data market to help data.
4 公共団体													Specific Sector	農業	EU Space Data Hub	-	2020	EU	The F4G7 digital service platform will aid developers of digital solutions.
5 公共団体													Multi Sector	データ交換	European Data Hub	-	-	EU	The official portal for European data.
6 公共団体													Specific Sector	Biological Data	European Biobank	-	1984	EU	The European Biological Data Hub
7 公共団体													Multi Sector	Various open data	GA4 - Interop.eu	-	2016	EU	The GA4 Programme supports the 66 cross-border and cross-sector public.
8 公共団体													Specific Sector	Astronomical Data	INSPIRE ASTRO-EU	-	-	France	The Paris astronomical Data Centre compliant simulation codes. data site.
9 公共団体													Multi Sector	人、法人、組織のインフラ、業務、活動など	ベータ・データ	-	2020	Japan	ベータ・データは、政府・企業・市民のデータを統合し、データ連携基盤の中心となるプラットフォーム。
10 公共団体													Multi Sector	データ交換	The Luxembourg	-	2010	Luxembourg	In a general effort to strengthen data openness and transparency. The first Luxembourg initiative. This area is identified as a key factor, not only for
11 公共団体													Specific Sector	Personal Data	Singapore Finance	-	2020	Singapore	Singapore Finance Data Exchange enables individuals to access through the world's leading open source data management system.
12 公共団体													Multi Sector	-	CKAN	-	2008	UK	The world's leading open source data management system.
13 公共団体													Multi Sector	275,027 datasets	Data site	-	2008	US	Here you will find data, tools, and resources.
14 公共団体													Specific Sector	America's education data	Department of Ed	-	2010	US	Data set gov in the U.S. Department of Education's data assets in a single place.
15 公共団体													Specific Sector	Humanitarian Data Exchange	Humanitarian Data	-	2014	US	The Humanitarian Data Exchange (HDX)
16 公共団体													Specific Sector	General Justice Records	National Data Hub	-	2008	US	The NDCU system provides online

(図 18.) 事前調査における各団体の整理方法

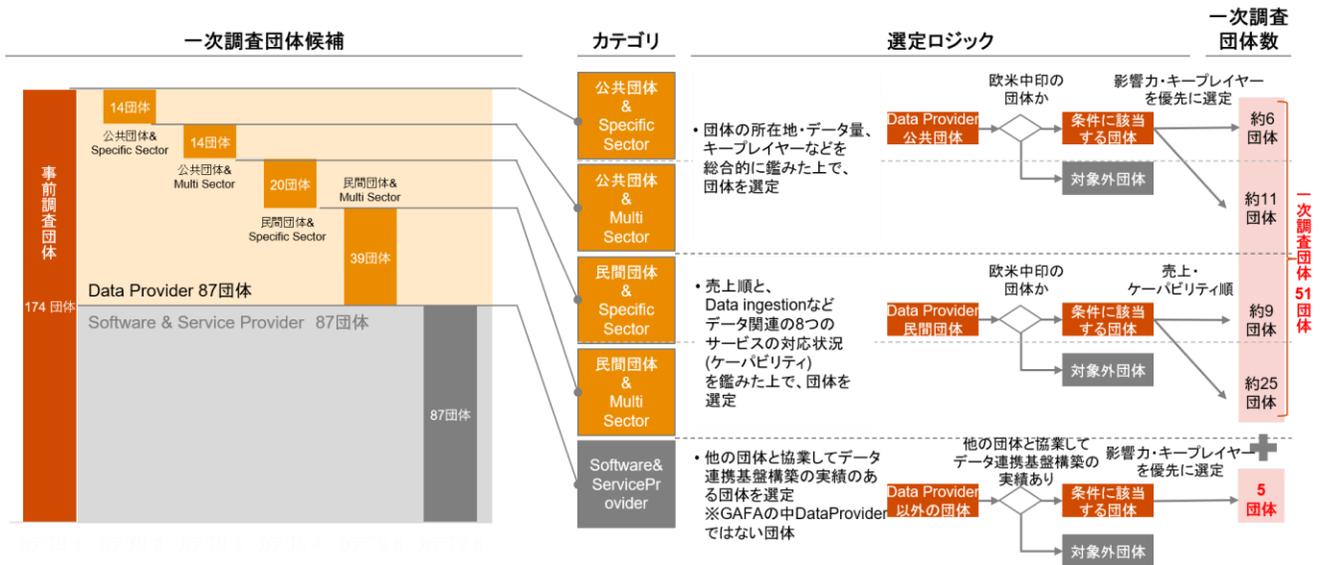
なお、事前にリスト作成した 120 団体に加えて、GAFAM 等の必須の調査団体と、本調査に関わる各種政府公開資料に掲載されている団体、データ連携基盤のコア機能として指定した Data Exchange & Marketplace, Data Catalog 以外の領域におけるリード団体などの合計 54 団体を加えた計 174 団体を調査候補の対象に追加している。



(図 19.) 調査対象団体の整理

1.2.2.5. グローバルベンチマーク調査の全体像 (一次調査)

前節まで整理していたロングリストを基に、Data をサービスとして提供する Data Provider に該当する団体を中心に一次調査団体を選定した。



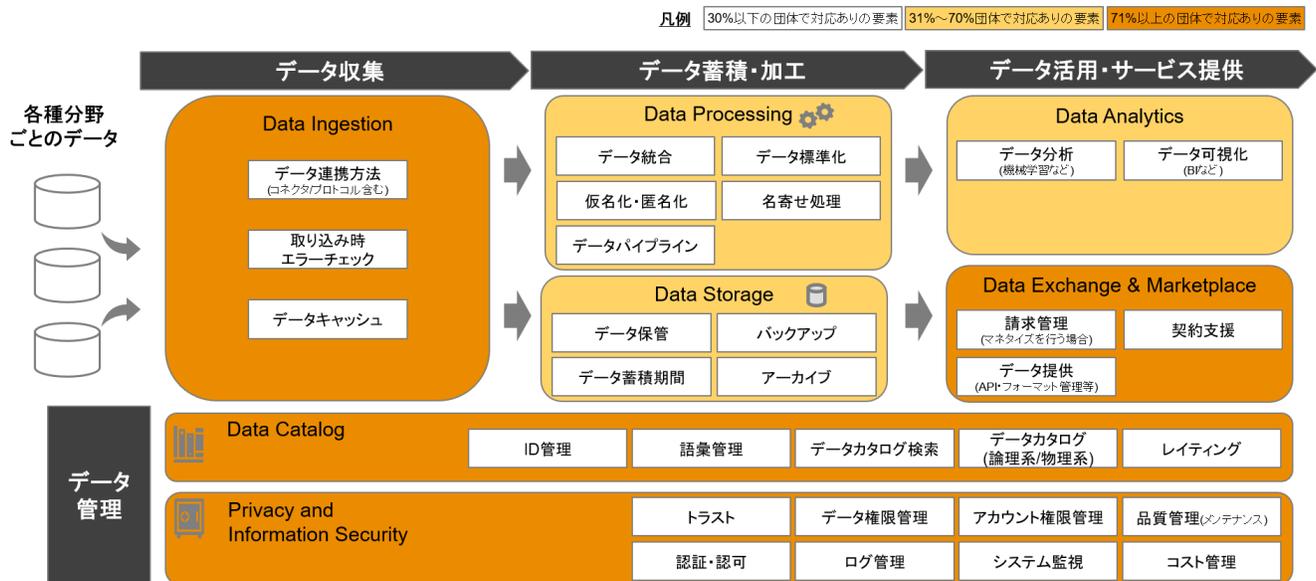
(図 20.) 一次調査の選定ロジック

上記ロジックに基づき、一次団体の対象として 51 団体(追加調査の FHIR 含む)を選定し、デスクトップベースの調査を実施した。

カテゴリ	#	会社・組織名	Software Provides	Service Provider	Data Provider
公共団体 & Specific Sector	1	HDX			✓
	2	digi.me w/Australia Gov			✓
	3	FaST			✓
	4	Blue Button	Blue Button Platform	✓	✓
	5	CNGBdb			✓
	6	India Stack	Cashless/Paperless Layer		
公共団体 & Multi Sector	7	data.gov.uk.		✓	✓
	8	EU Data Portal		✓	✓
	9	GAIA-X		✓	✓
	10	IDSA		✓	✓
	11	Data.gov	✓		✓
	12	Luxembourgish data platform		✓	✓
	13	Dataportal.asia		✓	✓
	14	NDAP		✓	✓
	15	data.gov.sg		✓	✓
	16	ODX		✓	✓
	17	Shanghai Data Exchange Corp.		✓	✓
民間団体 & Specific Sector	18	Facebook	✓	✓	✓
	19	DATOS		✓	✓
	20	Esri	ArcGIS Open Data	✓	✓
	21	Caruso	Caruso dataplace	✓	✓
	22	Here	HERE Data Hub	✓	✓
	23	Wejo	Marketplace	✓	✓
	24	Otonomo	Otonomo Automotive Data Services Platform	✓	✓
	25	Quandi		✓	✓
	26	FHIR		✓	✓
民間団体 & Multi Sector	27	Microsoft	Research Open Data	✓	✓
	28	Microsoft	Azure Open Datasets	✓	✓
	29	Microsoft	AIDC Data Cloud	✓	✓
	30	Alibaba	Marketplace	✓	✓
	31	Google	Dataset search		✓
	32	Google	GCP		✓
	33	SAP SE	SAP Data Warehouse Cloud	✓	✓
	34	Snowflake	Data Marketplace		✓
	35	Dawex	Data Marketplace		✓
	36	data intelligence hub	Marketplace		✓
	37	IBM Corporation	Data Asset eXchange	✓	✓
	38	Salesforce	Audience studio		✓
	39	Adobe	Audience Marketplace		✓
	40	Data Republic	Data market		✓
	41	Databroker DAO	Data Marketplace		✓
	42	Streamr	Marketplace		✓
	43	Axiom	ACXIOM DATA		✓
	44	Precisely	Precisely Data Experience		✓
	45	owner IQ	-		✓
	46	Harbr	Exchange Platform		✓
	47	data.world	Community		✓
	48	Openprise Data Marketplace	-		✓
	49	Crux Informatics	Crux Platform		✓
	50	Narrative.io	Narrative Data Stream		✓
	51	AWS	AWS Data Exchange		✓

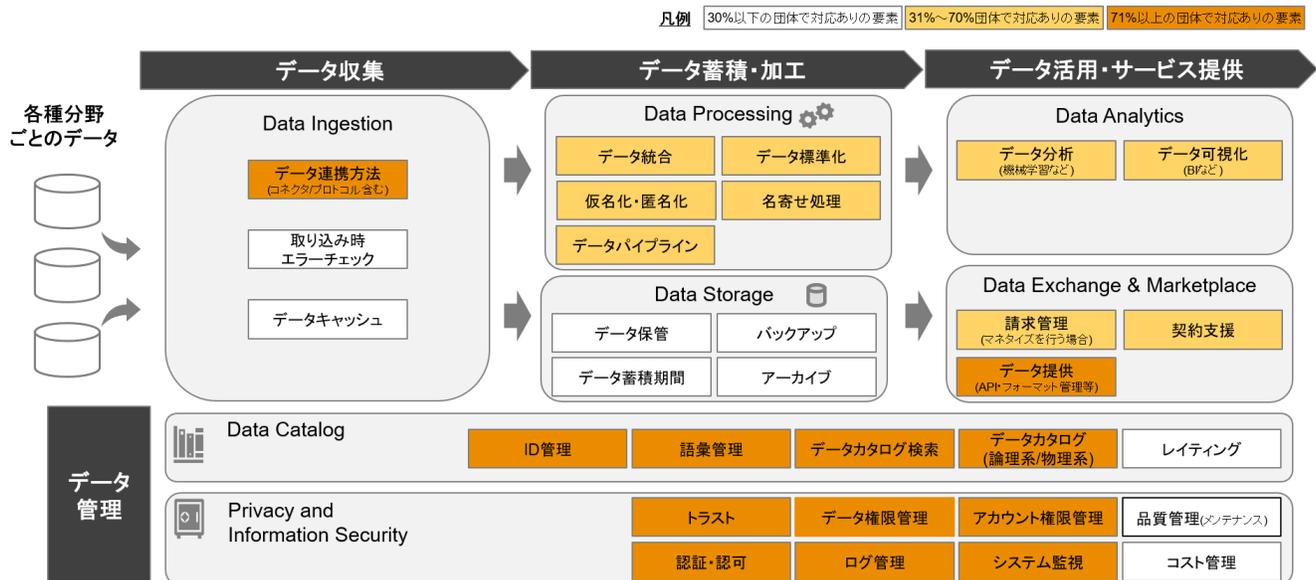
(図 21.) 一次調査対象団体

技術面での調査結果を述べる。海外データ連携基盤団体の 7 割以上の企業が、7 つの主要技術すべてに対応している。一方、Data Processing、Data Storage、Data Analytics については対応がまばらな状態である。



(図 22.) 海外データ連携基盤の機能要素ごとの対応状況

詳細な機能要素でみると、全体で 30 機能が存在する中で、21 個の機能が多くの団体で網羅されており、9 個については対応がまばらな状況である。



(図 23.) 海外データ連携基盤の詳細な機能要素ごとの対応状況

調査観点ごとの調査結果を述べる。調査観点 1 (団体分類) については、Data Provider となる団体は、Data 以外にも関連の Service または Software も追加で提供するものが多くみられる。特に民間団体においては、すべての団体がこのパターンに該当した。

(表 7.) 一次調査の対象団体の分類結果

団体分類		公共団体	民間団体
Data Providers	Data Providers のみ	6 団体 データ共有の為にプラットフォームで、データ連携関連の機能を用いて、データ提供者から収集されたローデータをそのままデータ利用者に提供するケースが多い	該当なし
	Data Providers 以外の機能も提供 (Service/ Software も提供)	11 団体 公共団体にも一部分析・データ表現サービスを提供する団体が存在するが、分析ツールは民間より初歩的な技術を使う傾向(簡単な集計または可視化のみを対応)	33 団体 <ul style="list-style-type: none"> Google, Amazon, Facebook などを含み、民間団体は全て本枠に該当 データ連携基盤関連のサービス(データマネジメント)から、データ連携基盤を活用してデータの価値を引き出す為の分析サービスなどを合わせて提供
標準策定団体		該当なし	1 団体

調査観点 2 (アーキテクチャ) については、分散連邦型と集中型が主流であった。また、あらゆるケースに対応する為に、分散連邦型と集中型の混合形態で構築するケースが多く存在した。

	公共団体	民間団体	計
分散型(非連携)	IDSA	Google	2社
分散連邦型	digi.me w/Australia Gov, Blue Button, European Data Portal, FaST, GAIA-X	DATOS, Streamr, Harbr, Adobe, Wejo, Alibaba, Openprise Data Marketplace, data.world, Databroker DAO, Crux Informatics	15社
	集中型と分散連邦型の混合形態 HDX, India Stack, Dataportal.asia, Data.gov, Luxembourgish data platform, data.gov.uk.	DataRepublic, Google, IBM Corporation, Esri, Dawex, SAP SE, Acxiom, Precisely, Microsoft, Data Intelligence Hub, owner IQ, Snowflake, Microsoft, Microsoft	20社
集中型	CNGBdb, data.gov.sg, NDAP, ODX, Shanghai Data Exchange Corp.	Facebook, Otonomo, Salesforce, Narrative.io, Caruso, Quandl, AWS, Here	13社
標準策定団体		FHIR	1社
計	17社	34社	51社*

*: Googleは「分散型(非連携)」と「集中型と分散連邦型の混合形態」に重複計上された為、総計から除外

(図 24.) 一次調査の対象団体のアーキテクチャによる分類結果

調査観点3(コア技術)については、来歴基盤・データ認証・ユーザ認証の3つの観点から整理した。

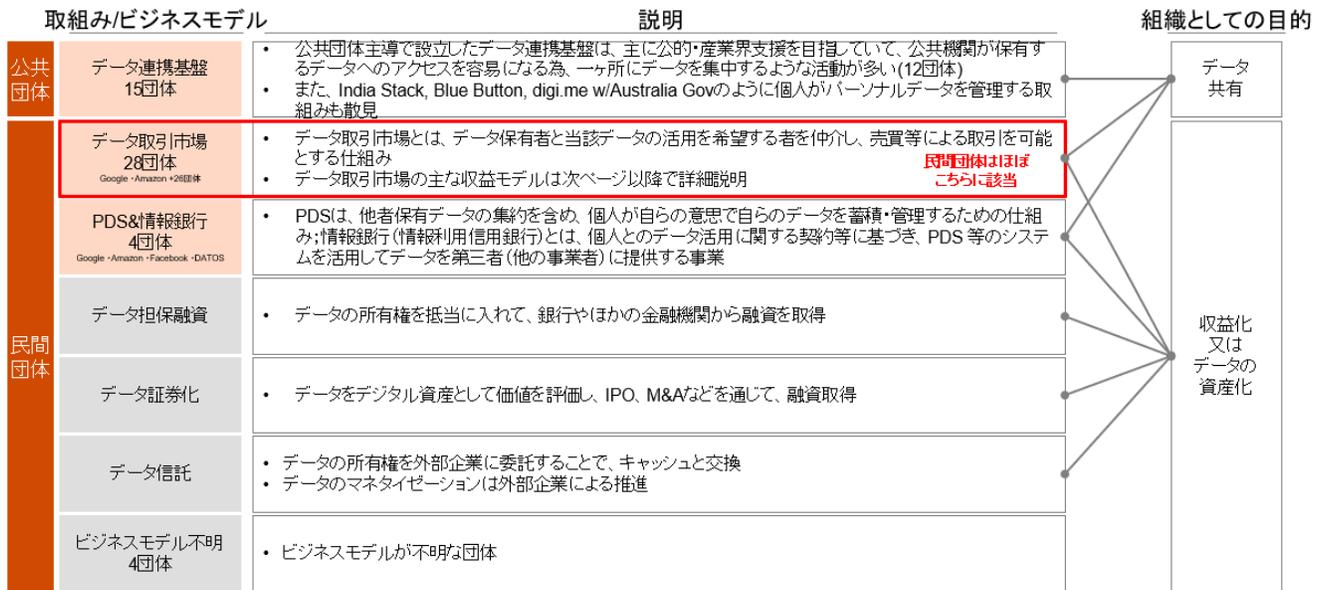
- 来歴基盤
 - データ来歴は基本的に登録必須となっている一方で、データ加工来歴はデータのクエリ履歴、更新タイムスタンプ情報のみの管理が主流となっていた
- データ認証
 - データIDやデータ提供者の認証は各団体独自での管理、ID認証はアカウントまたはAPIによる認証、契約データ形式はCSV等データファイル、API両方に対応しているパターンが多く見られた
- ユーザ認証
 - 利用者認証はアカウント管理の団体が主流であった一方、DID等の先端技術は、一部の団体での検討にとどまるなど少数であった

	来歴基盤	データ認証	ユーザ認証
データ来歴	<ul style="list-style-type: none"> ● 原則データのメタ情報として、基本登録必須 ● 一部の団体はデータ品質報告や、データ説明の機能を搭載 	<ul style="list-style-type: none"> ● データID指定方法 <ul style="list-style-type: none"> ○ 原則、各団体独自でユニークなIDを付与 ● データ提供者認証 <ul style="list-style-type: none"> ○ 多くは団体発行のアカウントにて管理 ○ 個人番号などで管理するケースも存在 	<ul style="list-style-type: none"> ● 利用者認証 <ul style="list-style-type: none"> ○ アカウントで管理する団体が多い(データ提供者と同様) ○ 団体によってはIdPによる認証を要求する場合がある ○ 完全にオープンなデータ連携基盤の場合は認証不要のものも存在
データ加工来歴	<ul style="list-style-type: none"> ● 主にデータのクエリ履歴・更新のタイムスタンプの情報のみを管理 <ul style="list-style-type: none"> ※ Data Providerとなる各団体は、データの所有権を持っていない為データの加工を行えない場合が多い ● 一部の団体はデータ追跡機能等を保持 	<ul style="list-style-type: none"> ● ID認証基盤 <ul style="list-style-type: none"> ○ 主にアカウント登録またはAPI関連の認証 <ul style="list-style-type: none"> ※ 認証方法は活用する技術によって異なる ● 契約データ形式 <ul style="list-style-type: none"> ○ CSV等のデータファイル・API両方対応の場合が多い(片方のみ対応の団体もある) 	<ul style="list-style-type: none"> ● DID等の先端技術対応 <ul style="list-style-type: none"> ○ EU Data Portal等、一部の団体においてブロックチェーンの活用が検討されている ○ GAIA-XIにおいては分散ID管理や、分散公開鍵が検討されている

(図 25.) コア技術の整理結果

事業形態の調査結果を述べる。ビジネスモデルについては、民間団体はデータ収益化、またはデータの資産化を実現する為に、多くがデータ取引市場に該当した。

PDS & 情報銀行に該当する団体も一部存在したが、更なる収益化に向けては、前述のとおり、これらに加えてデータ担保融資、データ証券化、データ信託といった3つのビジネス形態も存在している。



(図 26.) ビジネスモデルの整理結果

政府関与度については、公開情報からは関与が確認できなかった民間団体、官民連合型で政府から資金を受け取っている団体、公共団体が主体で開発している団体、といった3つのパターンが存在している。



(図 27.) 政府関与度の整理結果

個別の一次調査結果については、別紙 2 を参照されたい。

1.2.2.6. 一次調査対象の 51 団体の分類結果

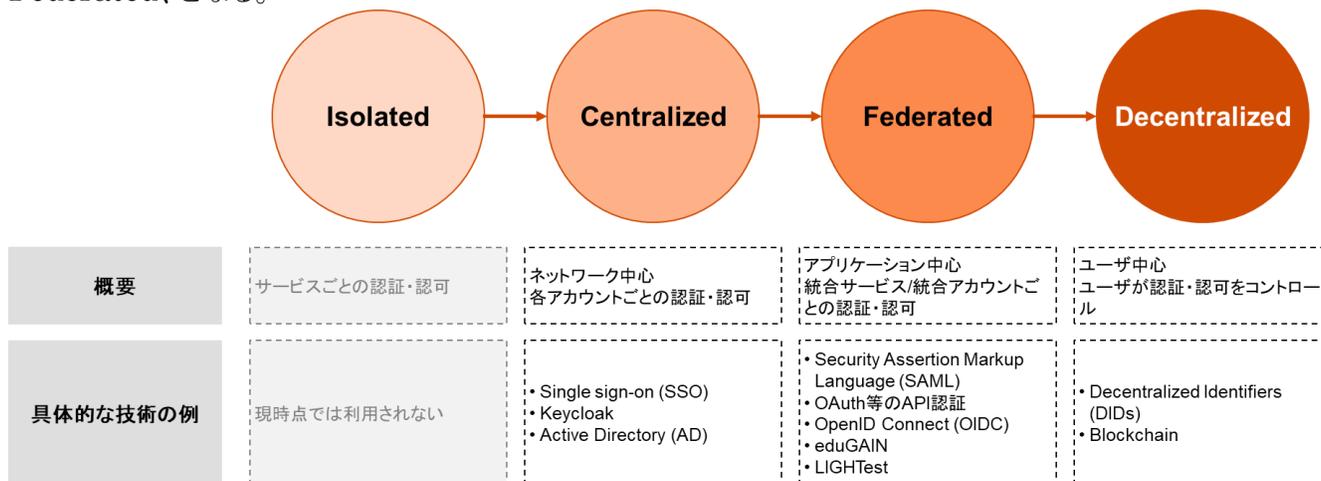
一次調査の対象となった 51 団体について、似たような団体を類型化し比較・参照しやすくするために分類を行った。

サービスのモデル、アーキテクチャ、連携対象・データ種別、使っているテクノロジーの 4 種類の分類軸に沿って分類した。それぞれ軸ごとに任意の分類が可能であるが、本節では認証・認可、公共団体の語彙標準への準拠状況および民間団体のサービスモデルによる分類結果を説明する。(詳細な結果は別紙 3 を参照)

分類軸	分類項目	説明	分類詳細			
サービスのモデル	1-1 提供するサービスの内容	提供しているサービスの目的に従って団体を分類	データ公開	収益を上げたい	国家単位でデータを囲い込みたい	
アーキテクチャ	2-1 サービス提供の手段・手法	サービス提供の方法に従って団体を分類	GAIA-X型 (連邦型)	GAFAM型 (中央集権型)		
連携対象・データ種別	3-1 取り扱うデータ	提供するデータの種別に従って団体を分類	メタデータ	業界特有or 複数業界データ	オープンデータ	
使っているテクノロジー	4-1 メタデータ検索	メタデータ検索機能の有無に従って団体を分類	機能有		機能無	
	4-2 認証・認可	認証・認可の方法に従って団体を分類	isolated	centralized	federated	decentralized
	4-3 データID	データIDの付与方式に従って団体を分類	自社定義	標準定義	その他方法	
	4-4 ドメイン語彙	ドメイン語彙の定義に従って団体を分類	自社定義	標準定義		

(図 28.) データ関連団体分類の考え方

なお、4-2. 認証・認可については、Isolated, Centralized, Federated, Decentralized の4つのレベルに分類可能⁷である。なお SSO や AD は Centralized、一般的な OAuth 等の認証は Federated、となる。



(図 29.) 認証・認可の分類の考え方

図 28 で整理した観点に基づき、公共団体の語彙標準への準拠の傾向を加味し、分類を実施した。民間団体については、標準への準拠が見られないため、各団体の方針に応じた分類を実施した。

公共団体 (半官・半民の団体を含む)	民間団体
<p>17 団体</p> <p>(Data.gov, data.gov.uk., EU Data Portal, GAIA-X, IDSA, Dataportal.asia, data.gov.sg, Luxembourgish data platform, NDAP, ODX, HDX, FHIR, India Stack, FaST, Blue Button, CNGBdb, Shanghai Data Exchange Corp.)</p>	<p>34 団体</p> <p>(Microsoft (Azure Open Datasets, AIDC Data Cloud, Azure Open Data), SAP SE, IBM Corporation, Data.world, AWS (AWS Data Exchange), Alibaba, Google (GCP, Google Dataset Search), Snowflake, Dawex, data intelligence hub, Salesforce, Adobe, Data Republic, Databroker DAO, Streamr, Acxiom, Precisely, owner IQ, Harbr, Openprise Data Marketplace, Crux Informatics, Narrative.io, digi.me w/ Australia Gov., Facebook, DATOS, Esri, Caruso, Here, Wejo, Otonomo, Quandl)</p>

<p>語彙標準への準拠状況による分類を実施し、SIP/ CADDE (仮称)として従うべき方向性があるかを調査</p>	<p>語彙標準への準拠が見られないため、団体意図による分類を実施</p>
--	---

上記の通り、語彙標準への準拠状況を見ると、公共団体の対象 17 団体のうち、複数セクター対応の団体は、既存の標準に基づいて語彙を定義する機会が多いことが確認された。

⁷ ミュンヘン連邦軍大学 サイバー防御研究機関 (CODE)「IMC: A Classification of Identity Management Approaches」(2020 年) (<https://detips2020.github.io/PreProceedings/DETIPS2020-paper-01.pdf>) をもとに PwC が作図

	標準に基づいて語彙を定義	自社にて語彙を定義	定義なし ※ユーザが各自で定義する	記載なし
複数セクター対応 (11団体)	6団体 (Data.gov, data.gov.uk., EU Data Portal, GAIA-X, IDSA, Dataportal.asia)	1団体 (data.gov.sg)	0団体	4団体 (Luxembourgish data platform, NDAP, ODX, Shanghai Data Exchange Corp.)
特定セクター対応 (6団体)	1団体 (HDX)	1団体 (FHIR)	1団体 (India Stack (Authentication Layer, Paperless Layer, Cashless Layer))	3団体 (FaST, Blue Button, CNGBdb)

(図 30.) 公共団体の分類結果

また 34 団体が対象となった民間団体においては、基本的にはデータの囲い込みや自社他サービスへの誘導を意図した団体が多いことが分かる。特定セクター向けの団体では、この傾向がより顕著になる。

	ニュートラルな団体	ニュートラルでない(※) 団体 ※データの囲い込み、自社他サービスへの誘導等の意図がみられる団体
複数セクター対応 (25団体)	6団体 (Databroker DAO, Streamr, Harbr, data.world, Data Republic, Openprise Data Marketplace)	19団体 (*Microsoft (Azure Open Datasets, AIDC Data Cloud, Azure Open Data), *AWS (AWS Data Exchange), Alibaba, *Google (GCP, Google Dataset Search), SAP SE, IBM Corporation, Snowflake, Dawex, data intelligence hub, Salesforce, Adobe, Acxiom, Precisely, owner IQ, Crux Informatics, Narrative.io)
特定セクター対応 (9団体)	1団体 (Caruso)	8団体 (digi.me w/ Australia Gov., Facebook, DATOS, Esri, Here, Wejo, Otonomo, Quandl)

(図 31.) 民間団体の分類結果

1.2.2.7. グローバルベンチマーク調査の全体像 (深掘調査)

深掘調査団体の選定の考え方を最初に述べる。① データ基盤を連携している団体(「メタ基盤」に該当するか)、② GAFAM のように自社でプラットフォームを持っている団体、③ インド/中国の団体⁸ という 3 つの観点から整理を行った。

⁸ SIP の研究実施者は欧米の調査は行っていたが、インド/中国の調査は不十分であった。

判定条件	選定理由	欧州・公共団体	欧州・民間団体	北米・公共団体	北米・民間団体	印中・公共団体	印中・民間団体
① データメタ連携基盤の調査 (GAIA-Xなど)	<ul style="list-style-type: none"> ✓ データ基盤を連携するメタ基盤 (GAIA-Xなど) ✓ SIPが目指すデータ連携基盤間をつなぐメタ連携基盤を調査 ✓ 認証やID付与に利用している技術を把握する 	GAIA-X IDSA	Data Intelligence Hub Dawex	Data.gov* *US Data Federation メタ連携基盤としての機能はないが、USの国家データ戦略に関わる団体になる為、要検討	Snowflake Google* Data.world Openprise	<div style="border: 1px solid black; padding: 5px;"> 凡例: 候補団体 (ニータル) 候補団体 深掘対象外 </div>	
② GAFAMの調査	<ul style="list-style-type: none"> ✓ GAFAM ✓ クラウドサービスのプロバイダとして状況を理解する ✓ 今後の競争相手として調査 (連携拡大意図があった場合に勝てる見込みがあるのか) 				Google Microsoft Amazon Facebook Apple	データ連携に積極的な活動をしていない	
③ インド/中国の団体調査	<ul style="list-style-type: none"> ✓ インド/中国におけるデータ連携基盤の関連取組 ✓ 欧州・米国に比べまだ調査が進んでいないので状況を把握する 			立ち上げたばかりで実績なし メタ連携基盤としての機能がない		<div style="border: 1px solid black; padding: 5px;"> India Stack Shanghai Data Exchange NDAP Beijing Data Exchange CNGBdb Alibaba </div>	

(図 32.) 深掘対象団体の整理の考え方

上記①～③の観点に基づき、GAIA-X, IDSA, Data.gov, Data.world, Google, Microsoft, Amazon, India Stack, Shanghai Data Exchange, さらに追加調査として FHIR の合計 10 団体を深掘対象とした。

有識者へのインタビューや追加のデスクトップ調査を実施した。

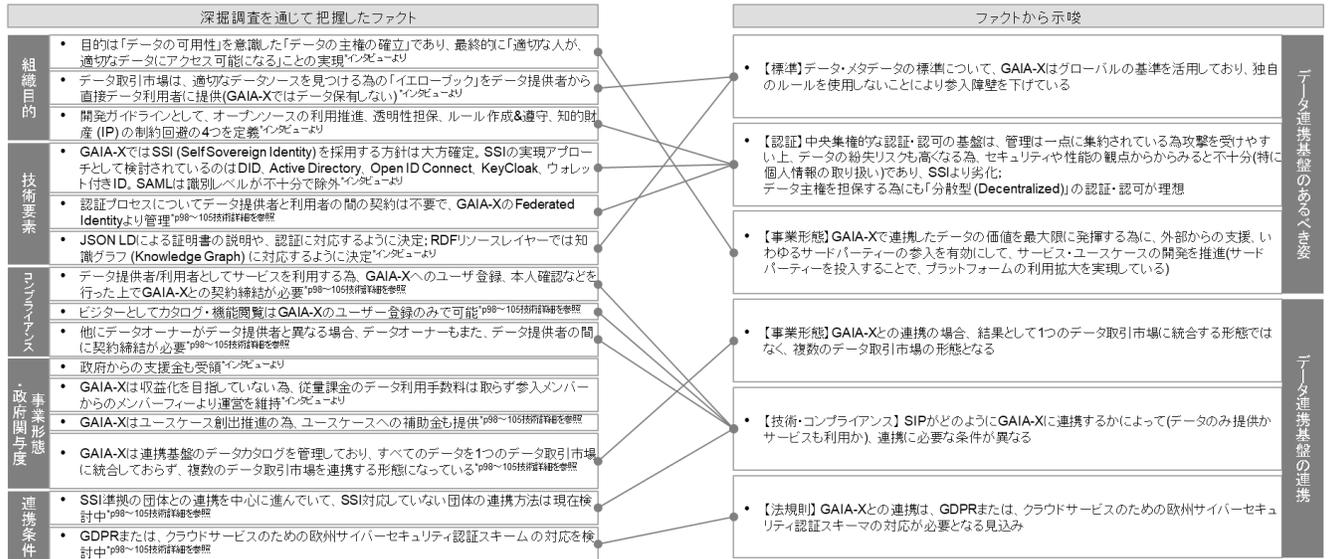
深掘条件	#	団体名	調査方法
① データメタ連携基盤の調査 (GAIA-Xなど)	1	GAIA-X	有識者ヒアリング
	2	IDSA	有識者ヒアリング
	3	Data.gov	デスクトップ調査
	4	Data.world	有識者ヒアリング (Chatベース)
② GAFAM調査	5	Google	有識者ヒアリング
	6	Amazon	有識者ヒアリング
	7	Microsoft	有識者ヒアリング
③ インド/中国の団体調査	8	India Stack	有識者ヒアリング
	9	Shanghai Data Exchange	有識者ヒアリング
追加調査	10	HL7 FHIR	Communityにて問い合わせ、回答受領

(図 33.) 深掘団体の調査方法

各団体の調査結果について、以下に要約する。各団体の詳細な調査結果については別紙を参照されたい。なお深堀対象 10 団体のうち、Data.gov と Data.world は示唆を得られる内容が無かったため、要約はなく調査結果の別紙のみで取りまとめを行っている。

● GAIA-X

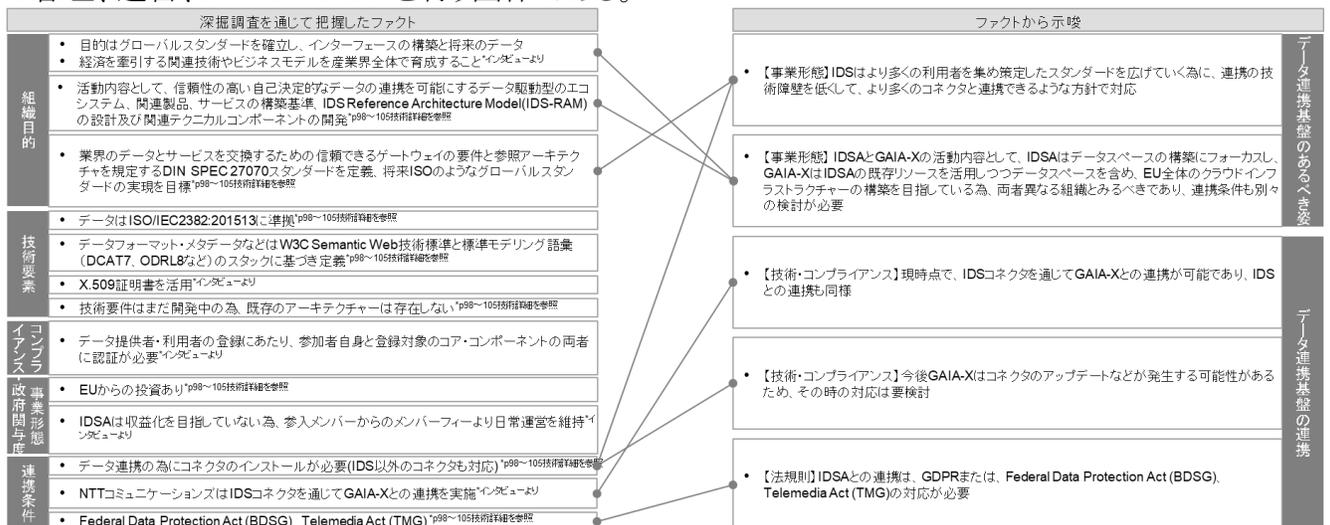
EU や世界各国の事業者、政府、および研究機関と連携して、安全かつデータ主権を担保可能なデータスペース・インフラの構築を目標とした団体である。



(図 34.) GAIA-X の概要

● IDSA

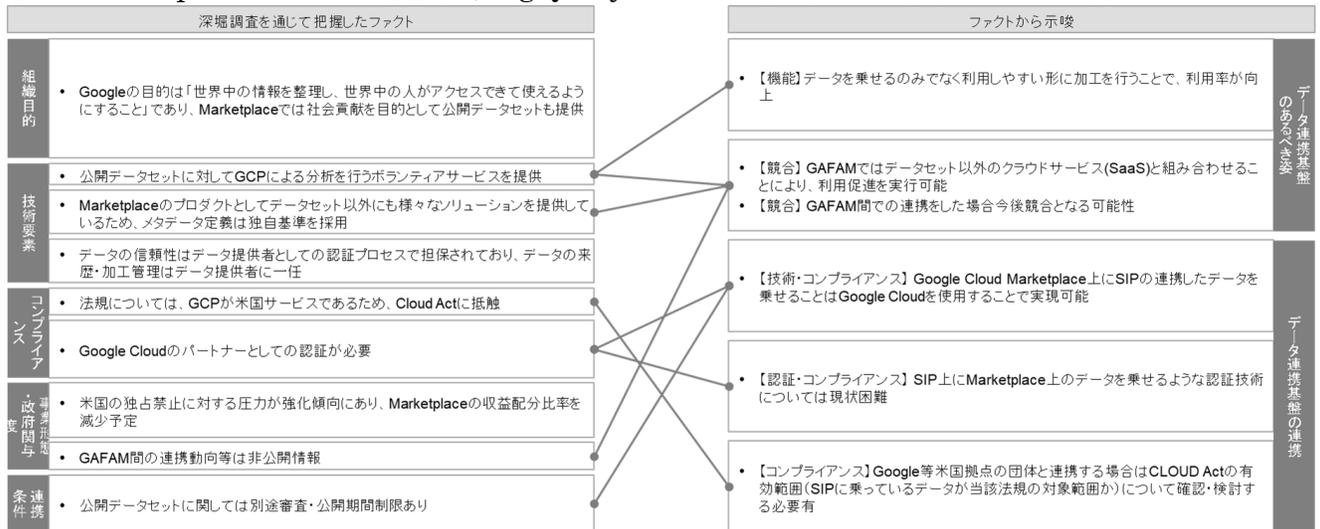
IDS プロジェクトで開発したデータ主権を担保するデータ連携の取組 (RAM、スタンダード) の管理、運営、プロモーションを行う団体である。



(図 35.) IDSA の概要

● Google Cloud Platform

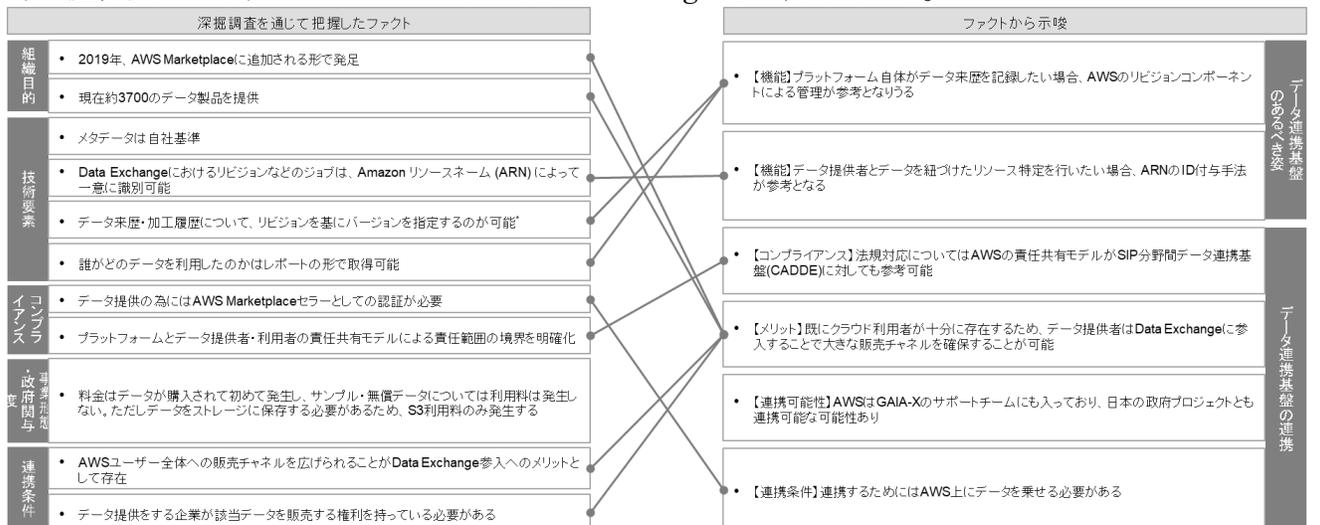
GCP クラウドサービスを通じて、ビルド済みのデータソリューションやデータセットをユーザに提供(Marketplace でデータを検索; BigQuery でデータを抽出・利用など)している。



(図 36.) Google Cloud Platform の概要

● AWS Data Exchange

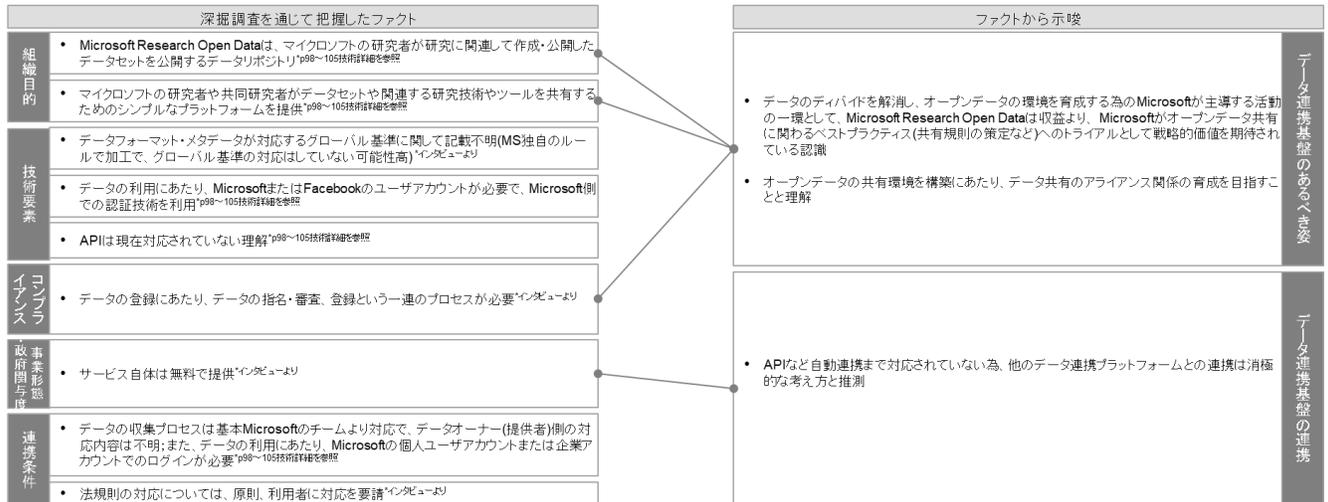
クラウドプロダクトの共有プラットフォームである AWS Marketplace に追加する形でデータセット提供用のプラットフォーム AWS Data Exchange を公開している。



(図 37.) AWS Data Exchange の概要

● Microsoft Research Open Data

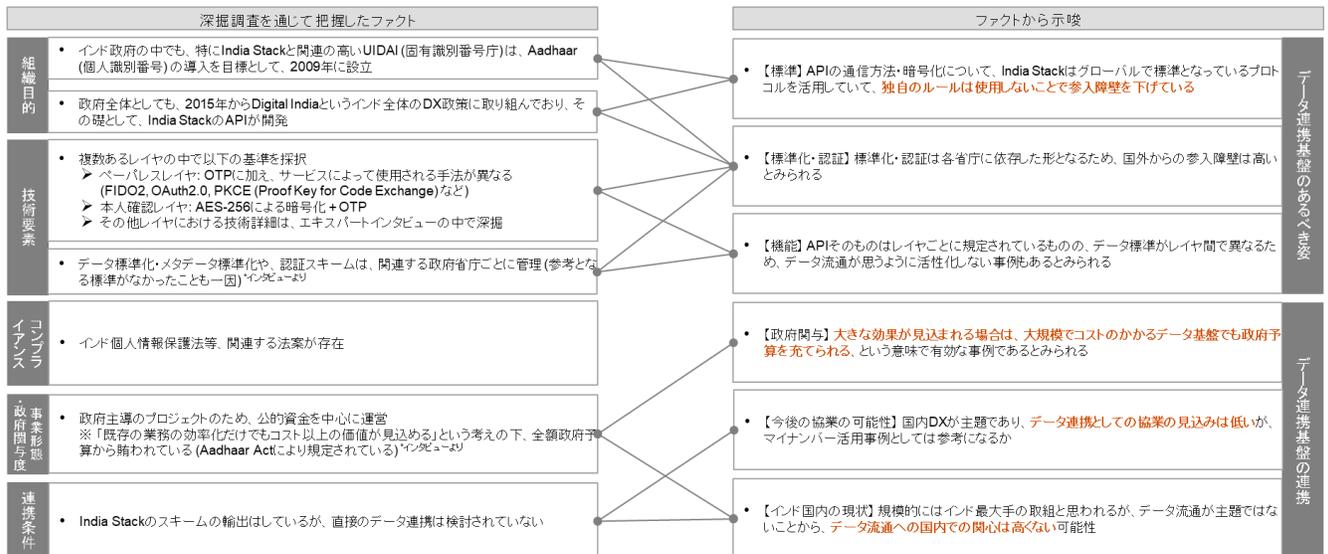
Microsoft のオープンデータ共有環境構築の推進する為の一環として、科学研究に必要なデータを無料提供しつつ、オープンデータの共有に必要な利用規約を策定する為の模索をしている。



(図 38.) Microsoft Research Open Data の概要

● India Stack

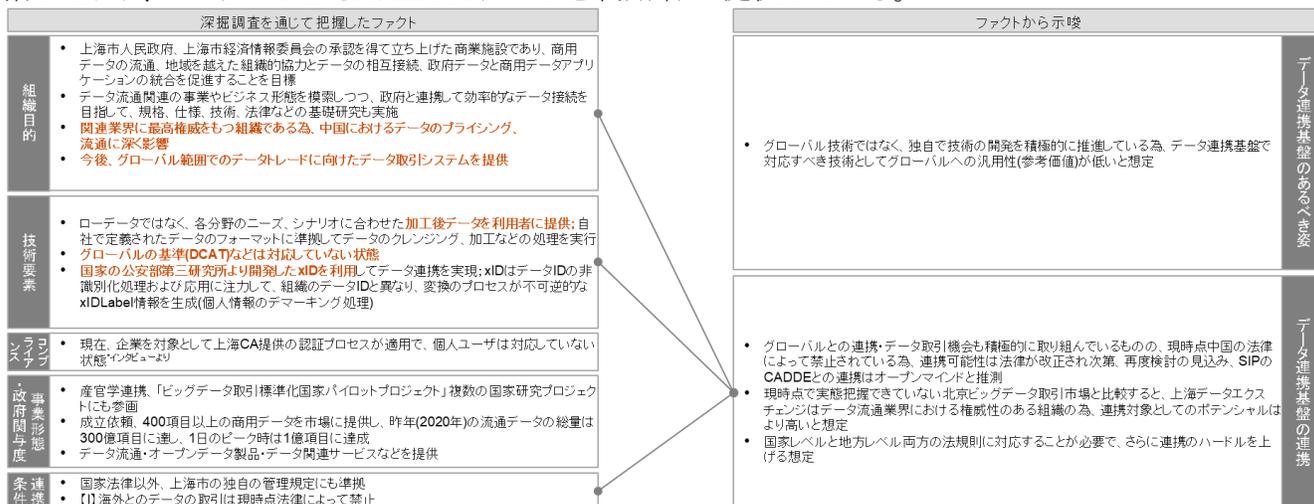
国内のDX化を目指しており、政府主導でレイヤという機能単位でデータ連携基盤を構築し、DXを推進している。



(図 39.) India Stack の概要

● 上海データエクステンション

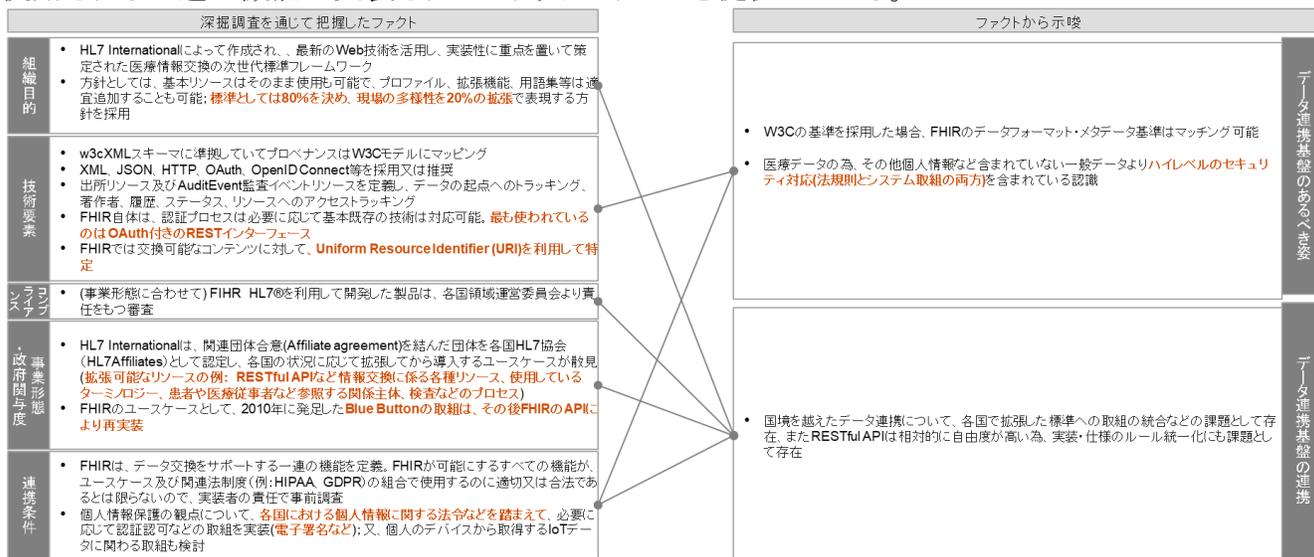
上海データエクステンションは中国国内だけではなく、グローバルとのデータ取引の実現を目指しており、ローデータから加工したデータを利用者へ提供している。



(図 40.) 上海データエクステンションの概要

● FHIR

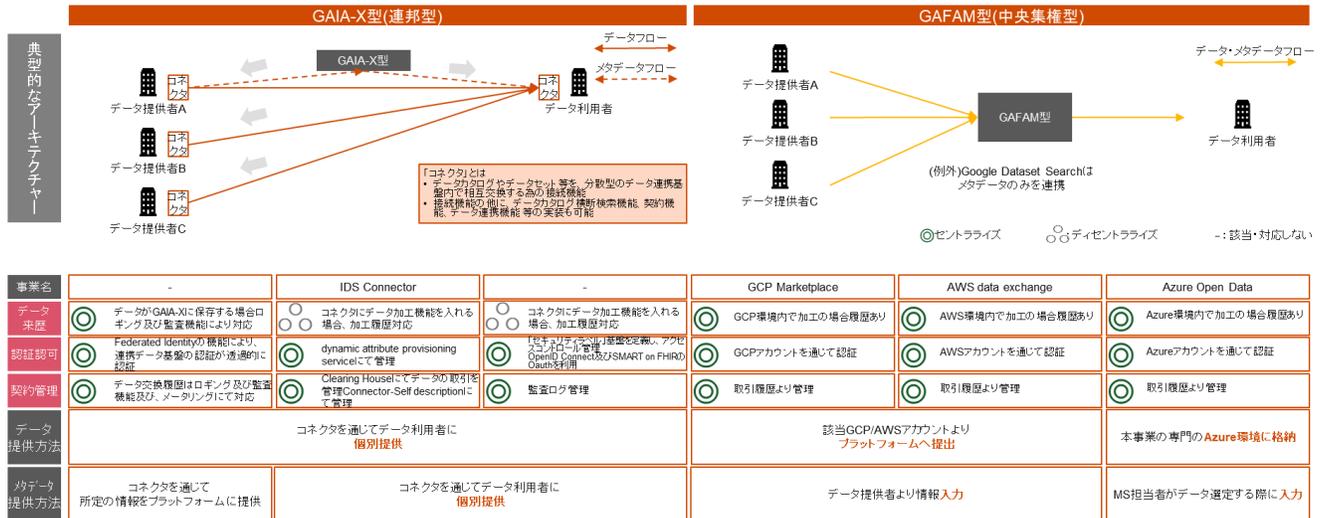
HL7 が公開している最新の医療データ交換のための標準規格であり、医療プロセス全般で使用される一連の機能が実装されたプラットフォームを提供している。



(図 41.) FHIR の概要

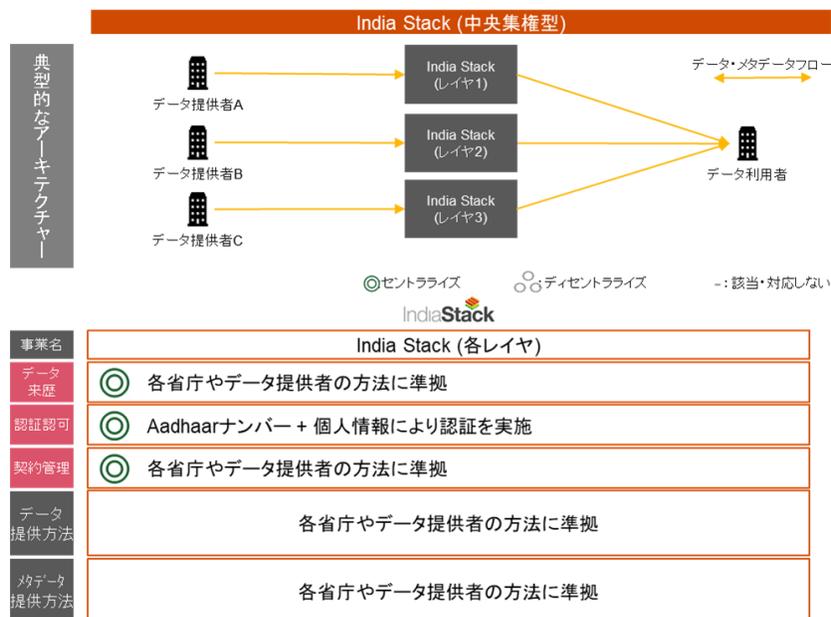
1.2.2.8. 深掘調査結果を踏まえた概念アーキテクチャ分類

データ連携の代表的な取組の種類として、GAIA-Xに代表される連邦型、GAFAMに代表される中央集権型が見受けられた。連邦型では、既存のデータ基盤同士をつなげて幅広くデータ連携基盤を繋げていく。一方、中央集権型ではデータを集約して、自社のサービスの活用促進を目指していると想定される。



(図 42.) 代表的な取組の概念図

なお、India Stackについては、アーキテクチャ構造としては各サービスレイヤの中央集権型をとる。しかし、レイヤとして定義される個別機能は各データ提供者や関連省庁により個別に基準が定義されているなど自律しており、中央集権と連邦のハイブリットとなる。



(図 43.) India Stack の概念図

1.2.3. 国際動向を踏まえた国内データ連携基盤の立ち位置

これまでの調査結果をもとに、海外団体とCADDE(仮称)の要素技術を比較した。国際競争力を持ちつつ海外団体との連携を実現するために、海外団体の参考事例や競合可能性、連携に向けて解決すべき課題・法規制を整理した。

1.2.3.1. CADDE(仮称)の機能要素

海外団体との対応状況比較を述べる。ただし、海外の諸団体は分野ごとのデータ連携基盤を意識せずに構築された一方で、CADDE(仮称)は既存の分野内データ連携基盤を連携する為に立ち上げたデータ連携基盤であるといった設立背景の違いがある。そのため、CADDE(仮称)は既存インフラである分野ごとの団体と共存しながら連携を進めていかなければならないことに留意が必要である。

	日本	米国	欧州	インド	中国
背景	<ul style="list-style-type: none"> 既に構築されている分野内データ連携基盤を連携することで分野間のデータ連携を実現 各分野におけるデータ基盤は既にできており、それらの分野内データ連携基盤を築くことが求められている状況 	<ul style="list-style-type: none"> 巨大IT企業がデータ流通を主導。政府は公的分野のデータ流通活用推進 民間分野のデータ活用促進は民間大企業が主導し、政府は強く介入していない。公的分野に関しては連邦・州政府レベル双方が積極的な取組を推進 民間団体が自由放任的にデータ流通を進めているが、プライバシー保護などの為、国家的にデータ戦略を立ち上げ、ガバナンス体制の構築なども急速に推進している状況 	<ul style="list-style-type: none"> 欧州委員会が中心となり、国際的なルール形成を主導。関連機関にてデータ連携基盤を構築・利用促進 欧州委員会が中心となり、デジタル分野での国際的なルール形成を主導 連邦型の欧州クラウド/データインフラ構想であるGAIA-Xプロジェクトの動きが加速していて、技術アーキテクチャや参加者に求める共通ルールの策定を推進 	<ul style="list-style-type: none"> ボトムアップによる分野別のレイヤー構成で基盤構築。レイヤーごとにデータ流通を促進 ボトムアップの急速な最先端技術の受容を進んでおり、政府によるデジタル公共インフラを推進 	<ul style="list-style-type: none"> 国家権力を背景としたトップダウン型のデータ連携を実現 中国のデータ活用政策は、市街地における顔認証機能付き監視カメラの配備等に象徴される、トップダウンの強固な国民監視が特徴
団体・関連サービス設立目的	<ul style="list-style-type: none"> 既存の分野毎のデータ連携基盤を繋いで、分野間のデータ連携基盤を実現し、最終的にグローバルのあらゆる分野やステークホルダーと連携し、信頼性のある自由なデータ流通 DFFT: Data Free Flow with Trust を実現する為CADDE(仮称)を開発 	<ul style="list-style-type: none"> GAFAMは自社のクラウドサービス(分析など)の利用者を拡大する為、データセットを商品としてユーザに提供 MicrosoftはData.govなど政府機関と協力し、OpenDataの流通を推進する為、関連標準化の取組を推進 	<ul style="list-style-type: none"> 蓄積・処理・活用されるデータの管理を欧州外の企業に依存せず、欧州内で実行できる技術環境整備 欧州域内外の企業のさまざまなクラウドサービスを単一のシステム上で統合し、業界をまたがるデータ交換を容易に行える標準的な認証の仕組みを通じて、インターオペラビリティ(相互運用性)を実現 	<ul style="list-style-type: none"> 国民IDシステム「Aadhaar」を基盤とした、デジタル公共インフラ(IndiaStack)を次々と創造 中央政府に限らず多くの人に設計図やプロセスを共有することで、各州政府、各省、ITベンダーなど様々なアクターが協働可能なプラットフォームを実現 	<ul style="list-style-type: none"> 国家主導によるデジタル社会基盤の実現
深掘団体	<ul style="list-style-type: none"> SIP CADDE(仮称) 	<ul style="list-style-type: none"> Google Microsoft Amazon 	<ul style="list-style-type: none"> GAIA-X IDSAs 	<ul style="list-style-type: none"> IndiaStack 	<ul style="list-style-type: none"> Shanghai Data Exchange Corp.

(図 44.) 各国におけるデータ連携基盤の設立背景

機能比較の観点として、

- ・プラットフォームが機能提供するのか
- ・利用者が機能を用意する必要があるのか

という観点でCADDE(仮称)と他団体を比較することで、CADDE(仮称)が現在開発中の機能状況を把握することが可能となる。

マーク	分類名	分類基準
◎	完全中央	・ 中央が機能提供し、中央の機能しか利用できない
○	中央提供&利用者実装必須	・ 中央が仕様や機能・コネクタを提供し、利用者側が実装
△	中央提供&利用者実装任意	・ 中央にも機能がおり、利用者側は必要に応じて実装 (例: GAIA-Xのストレージ)
□	完全利用者	・ 利用者側で実装。中央は仕様や機能を提供しない
-	機能対応状況非公開	・ -
×	機能開発なし	・ -

(図 45.) 機能比較の分類基準

分類結果を述べる。現時点での CADDE(仮称)は、主に Data Storage と Data Analytics 関連の機能には利用者側が対応している。

機能要素	#	機能名	CADDE(仮称)	GAIA-X	IDSA	Google	Microsoft	Amazon
Data Ingestion	1	データ連携方法(データ提供者側コネクタ)	○	△	△	△	◎	◎
	2	取込み時エラーチェック	×	◎	◎	◎	-	◎
	3	データキャッシュ	×	◎	-	◎	◎	◎
Data Processing	4	データ統合	△(開発中)	△	△	△	△	△
	5	データ標準化	△(開発中)	△	△	△	△	△
	6	名寄せ処理	△(開発中)	△	△	△	△	△
	7	仮名化・匿名化	△(開発中)	△	△	△	△	△
	8	パイプライン	△(開発中)	△	△	△	△	△
Data Storage	9	データ保管	□	△	□	△	◎	◎
	10	バックアップ	×	△	-	△	◎	◎
	11	データ蓄積期間	×	△(開発中)	-	△	◎	◎
	12	アーカイブ	×	△	-	△	◎	◎
Data Analytics	13	データ分析	□	△(開発中)	△(開発中)	△	△	△
	14	データ可視化	□	△(開発中)	△(開発中)	△	△	△
Data Exchange& Marketplace	15	データ提供	○	△	△	◎	◎	◎
	16	請求管理	◎	◎(開発中)	◎(開発中)	◎	◎	◎
	17	契約支援	◎	◎(開発中)	◎(開発中)	◎	◎	◎
Data Catalog	18	データカタログ	◎	◎	◎	◎	◎(開発中)	◎
	19	ID管理	◎	◎	◎	◎	◎	◎
	20	語彙管理	◎	◎	◎	◎	◎	◎
	21	データカタログ検索	◎	◎	◎	◎	◎	◎
	22	レイティング	×	-	◎	-	-	-
Privacy and Information Security	23	トラスト	◎	◎	◎	◎	-	◎
	24	品質管理	×	◎	◎	◎	-	◎
	25	アカウント権限管理	◎	◎	◎	◎	◎	◎
	26	データ権限管理	◎	◎	◎	◎	◎	◎
	27	ログ管理	◎	◎	◎	◎	◎	◎
	28	認証・認可	◎	◎	◎	◎	◎	◎
	29	システム監視	◎	◎	◎	◎	◎	◎
	30	コスト管理	×	◎	-	◎	◎	◎

(図 46.) 機能比較の分類結果

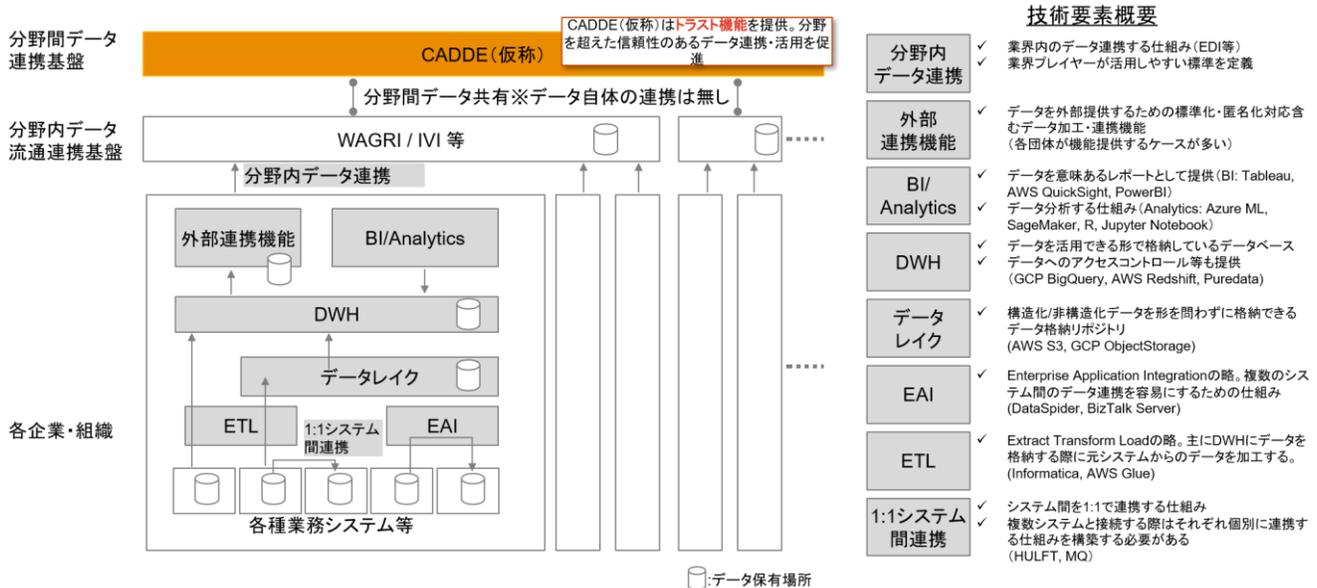
その他の団体側で機能実装している箇所を詳細に比較した結果、CADDE(仮称)で対応済の重要機能はグローバルの取組と比較して大きく劣後していない。また、メタデータ管理に AI を活用する等、先進的な取組も開発中である。

	SIP	GAIA-X	IDSA	Google	Microsoft	Amazon
	CADDE(仮称)	-	-	GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange
16 請求管理	・ 課金の仕組みあり	・ 個別契約の中で課金ルール設定(請求機能を開発中)	・ 個別契約の中で課金ルール設定(請求機能を開発中)	・ 課金の仕組みあり ・ クラウドサービス機能	・ 課金の仕組みあり ・ クラウドサービス機能	・ 課金の仕組みあり ・ クラウドサービス機能
17 契約支援	・ 契約管理の仕組みあり(契約ポータル)	・ 現状個別に契約締結、契約管理の仕組み構築中(自動審査の構想)	・ 現状個別に契約締結、契約管理の仕組み構築中(Blockchain活用によるスマートコントラクトの構想)	・ 契約管理の仕組みあり ・ データ利用者はGCPアカウント ・ データ提供者はパートナー契約	・ 契約管理の仕組みあり ・ データ利用者はAzureアカウント ・ データ提供者は現状オープンデータのみであるため契約管理不要	・ 契約管理の仕組みあり ・ データ利用者はAWSアカウント ・ データ提供者はパートナー契約
19 ID管理	・ リテラル(ユニークコード) ・ 一意の識別子(URI)で識別(データのIDもURLを使うと越塚先生からのコメント)	・ リテラル(ユニークコード) ・ 一意の識別子(URI)で識別	・ リテラル(ユニークコード) ・ 一意の識別子(URI)で識別	・ リテラル(ユニークコード)で識別 ・ 一意の識別子(URI)で識別	・ リテラル(ユニークコード)で識別 ・ 一意の識別子(URI)で識別	・ リテラル(ユニークコード)で識別 ・ 一意の識別子(URI)で識別
20 語彙管理	・ ISA Core Vocabulary, NIEMを参考にしたIMI共通語彙基盤採用 ・ AI技術を用いてメタデータを構造化	・ グローバル基準(W3C)採用	・ グローバル基準(W3C)採用	・ 自社独自の基準採用	・ 自社独自の基準採用	・ 自社独自の基準採用
28 認証・認可	・ Keycloakを実装済み(DID検討中と理解)	・ SSIコンセプトに基づき、DID, Active Directory, Open ID Connect, KeyCloak, フォレット付きID, などの実装を検討中	・ Dynamic Attribute Provisioning Service (DAPS)を活用(Oauth 2.0のアクセストークンを発行)	・ Google アカウントでの認証認可	・ Azure アカウントでの認証認可	・ AWS アカウントでの認証認可

(図 47.) 詳細機能の比較結果

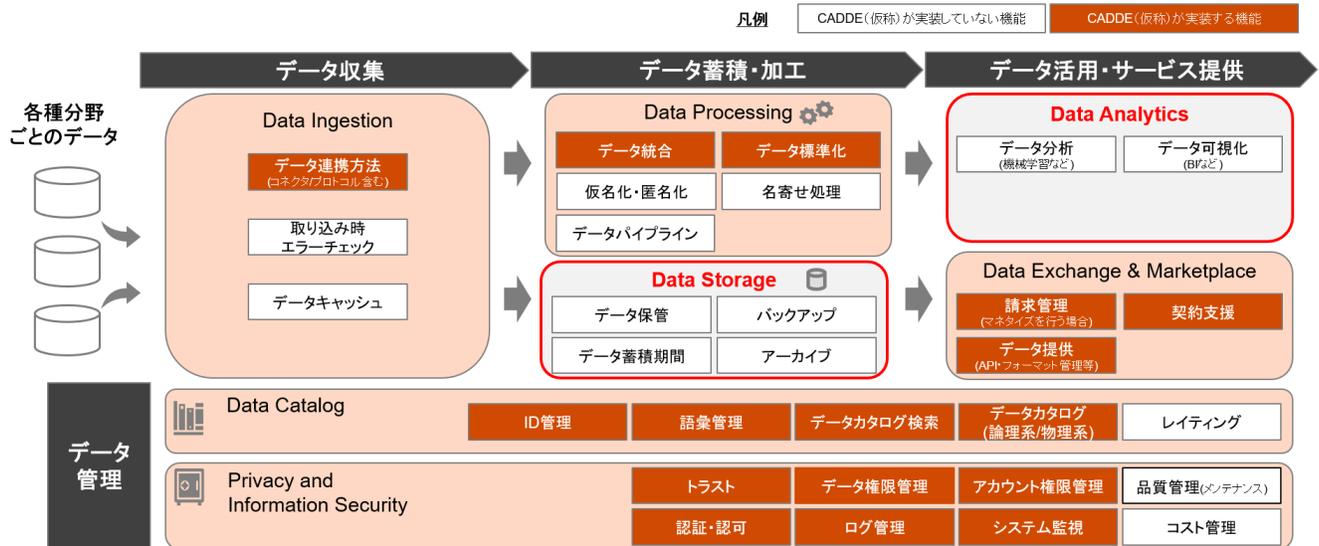
機能比較の詳細については、別紙 4 を参照されたい。

また、参考として、データ連携における代表的な技術構成要素を整理している。



(図 48.) データ連携における代表的な技術構成要素

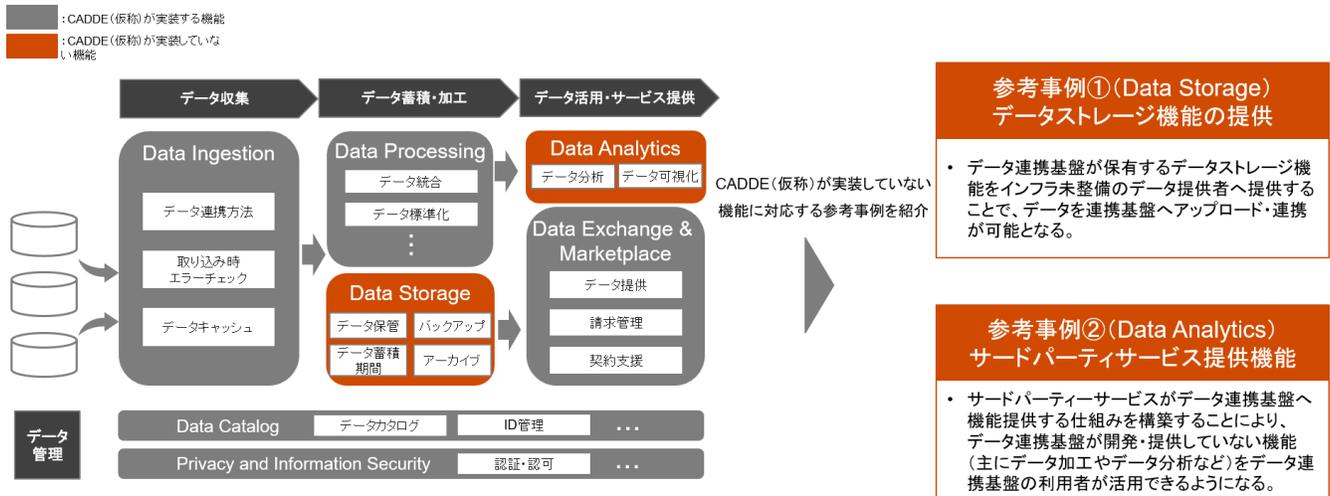
これまで整理した機能要素に基づき、CADDE(仮称)のデータ連携基盤が対応する技術的観点を図 32 のようにマッピングした。なお Data Storage, Data Analytics については団体として実装を行っていない(利用者側で実装している)ことを示している。



(図 49.) CADDE(仮称)の機能要素ごとの対応状況

1.2.3.2. 日本にとって参考となりうる事例の調査

これまでの調査結果に基づき、CADDE(仮称)の普及を加速する方針を提案する。なお、検討にあたり、CADDE(仮称)が実装していない Data Analytics と Data Storage 機能に注目し、海外団体の取組を分析した。



(図 50.) 参考事例の考え方

(1) データストレージ機能の提供

CADDE(仮称)においてデータストレージ機能は現時点では未実装である。しかし、データ連携基盤が保有するデータストレージ機能を、未だデータ連携のインフラが整備されていない分野へ提供することが参入障壁を下げ、結果として分野間連携の早期拡大につながることを期待される。

<p>概要</p>	<p>データ連携基盤が保有するデータストレージ機能をインフラ未整備データ提供者へ提供することで、データを連携基盤へアップロード・連携が可能となる。</p>
<p>機能提供例</p>	<p style="text-align: center;">インフラ未整備の一部利用者へデータストレージ機能を提供</p>
<p>対応する団体 (サービス名)</p>	<p>GAIA-X、Google※ (Google Cloud Storage)</p>
<p>期待される効果</p>	<ul style="list-style-type: none"> 未だインフラが整備されていない分野へデータストレージ機能を提供することにより、データストレージを持たずコネクタを介したデータ連携が行えなかったデータ提供者・分野との連携拡大が期待される。 データストレージでデータを保管する地理的位置をデータ連携基盤が管理することで、国際的な法規制 (Cloud Actなど) への対応も期待される。

※ Googleは従量課金によるストレージ利用料獲得も存在

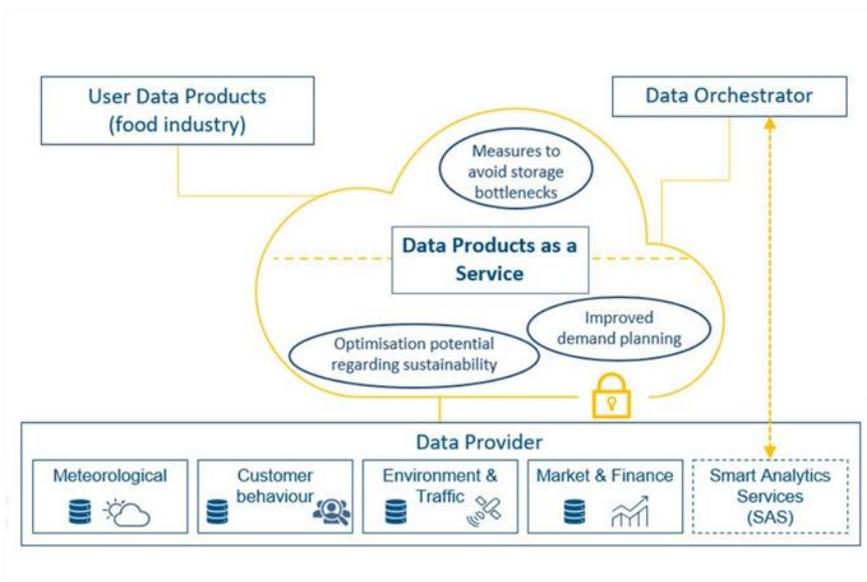
(図 51.) データストレージの提供事例

特に GAIA-X によるデータストレージ機能提供事例として、データストレージが整備されていない団体(機能を保持していない、セキュリティ要件を満たさない等)へ、データストレージ機能を提供することにより、分散型のデータ連携を拡大した事例を紹介する。

① Data Interoperability and Data Sovereignty for Food Production⁹(食品加工分野)

食品の原材料から標準化された製品へ加工し、品質管理するためには大量のデータの収集・活用が必要である。しかし、場所やメーカーを跨いだデータ活用は一部に留まっており、継続的なデータ収集と分析は行えていない状況であった。

そこで GAIA-X は、現場が保持していない安全なストレージと様々なソースのデータと連携機能を提供する。そのうえで、コンピューティングパワーを提供することで食品業界のデータエコシステムを実現し、分野外と連携するためルールと基準を設定した。



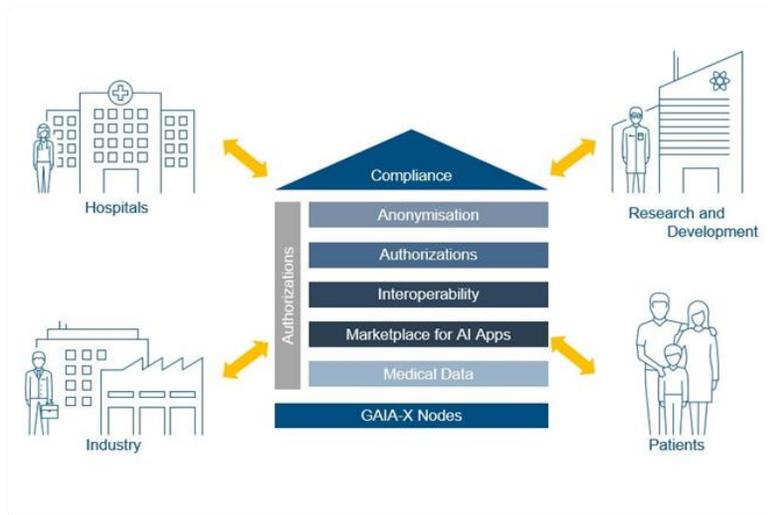
(図 52.) 食品分野におけるデータストレージ提供のイメージ

⁹GAIA-X「Data Interoperability and Data Sovereignty for Food Production」(2021 年)
(<https://www.bmwi.de/Redaktion/EN/Artikel/Digital-World/GAIA-X-Use-Cases/fast.html>)

② AIQNET – Medical Data Ecosystem¹⁰(医療分野)

背景として、医療製品の性能・安全性データの継続的なデータ収集と共有は、ヘルスケア分野の病院、産業、患者、研究機関など様々なステークホルダーへ利益をもたらす。しかし、機微情報となる医療分野のデータは、商用クラウドへのデータ提供に抵抗があり連携が進んでいない状況となっていた。

そこで GAIA-X は、国際的な法規制、サイバーセキュリティなどの要件を備えた安全なストレージとコンピューティング機能を提供した。これにより、機密性の高い分野で活動する病院や企業がデータ連携・活用する恩恵を得られるようにした。

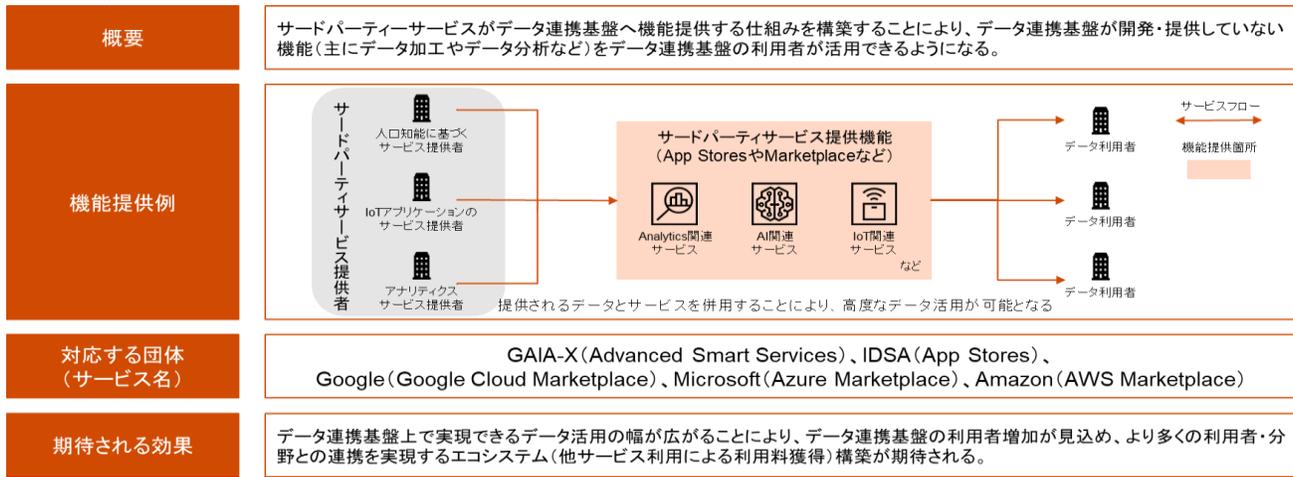


(図 53.) 医療分野におけるデータストレージ提供のイメージ

¹⁰GAIA-X「AIQNET – Medical Data Ecosystem (previously KIKS)」(2021 年)
(<https://www.bmwi.de/Redaktion/EN/Artikel/Digital-World/GAIA-X-Use-Cases/aiqnet.html>)

(2) サードパーティサービス提供機能

CADDE(仮称)においてデータアナリティクス領域は現時点では未実装である。サードパーティサービスを導入することにより、未対応機能の補完によるエコシステム構築と、機能利用を契機とした利用者・連携分野増加が期待される。



(図 54.) サードパーティサービスの提供事例

1.2.3.3. 競合となりうる団体の検討

CADDE(仮称)の普及にあたり競合となりうる団体の検討結果を報告する。提供データの分野ならびにデータの収集・提供地域、連携意思から、深掘団体とCADDE(仮称)の競合可能性を分類し、協調/独立共存/競合の3段階で整理を行った。GAIA-X/IDSA/HL7 FHIRは、グローバル範囲でサービス提供しており他団体との連携意思が強いため、優先的に協調していくべき対象といえる。

Google/Amazon/Microsoft/Data.worldは、CADDE(仮称)との重複部分が多く連携意思が薄いため、競合対象として注視していくべき対象といえる。ただし、技術的には、活用の可能性が存在するとみられる。

#	団体名	事業名	提供データの分野	主なデータ収集元地域	主なデータ提供先地域	連携意思	分類
1	GAIA-X	GAIA-X	複数業界	欧州	全世界(欧州中心)	有(NTT Comと連携実績あり)	協調 連携意思有
2	IDSA	IDSA	複数業界	(プラットフォーム未完成)	(左記同様)	有(NTT Comと連携実績あり)	
3	HL7 FHIR	HL7 FHIR	特定業界(医療)	全世界	全世界	有	
4	Shanghai Data Exchange	Shanghai Data Exchange	複数業界	中国	中国	将来的に可能性あり(グローバルでのデータ取引システムを今後提供見込み)	独立共存 データの分野、地域がCADDEと重複無 連携意思無
5	data.gov系列	Data.gov.uk, data.gov, data.gov.sg and etc.	複数業界	英/米/シンガポールなど	左記同様	グローバル連携する仕様ではない(米国政策IDを入れる仕組みが必要)	
6	India Stack	India Stack	特定業界(レイヤごとに業界が確定)	インド	インド	無(データ輸出は想定していない)	競合 データの分野、地域がCADDEと重複有 連携意思無
7	data.world ²⁾	Community	複数業界	米国	全世界	無	
8	Microsoft ²⁾	Azure Open Datasets	複数業界	全世界	全世界	検討されてない	
9	Amazon	AWS Data Exchange	複数業界	全世界	全世界	検討されてない	
10	Google	Google Cloud Marketplace	複数業界	全世界	全世界	検討されてない	

1) 立ち上げ期の為、中国国内のデータを中心に提供

2) 競合団体の中でも、data.world CommunityとMicrosoft Azure Open Datasetsは、主にオープンデータのみ収集・提供するため競合の程度はGoogle、Amazonよりも低い

(図 55.) 競合になりうる団体の分類結果

1.2.3.4. 深掘調査対象との連携課題

深掘調査対象と連携する際の課題・懸念事項を報告する。

- 制度面での連携課題
 - 他団体より CADDE(仮称)のデータ提供者・一つ一つのデータへの追加審査が必要となる。
- 技術面での連携課題
 - 主にデータ・メタデータ基準の標準規格への変換対応(深掘団体は CADDE(仮称)と異なる基準を準拠している)、または連携コネクタの整備が必要となる。

深掘調査対象と連携する際の課題・懸念事項		
制度	データ提供側の信頼性評価	<ul style="list-style-type: none"> ● GAIA-Xを含む6団体*は連携基盤の先のデータ提供者まで審査が求められる。連携の際にはCADDE(仮称)のみならず各データ提供者の審査が必要 ※:GAIA-X, IDSA, FHIR, Shanghai Data Exchange, Amazon, Google
	データ自体の信頼性確保	<ul style="list-style-type: none"> ● GAIA-Xを含む3団体*は個別データの審査が求められる。連携の際にはデータ一つ一つの審査が必要 ※:GAIA-X, IDSA, FHIR
技術	データ・メタデータ基準	<ul style="list-style-type: none"> ● CADDE(仮称)が準拠している標準はIMI共通語彙基盤と文字情報基盤であり、GAIA-Xを含む8団体*と異なる為、他団体が準拠している標準規格への変換対応が必要 ※:GAIA-X, IDSA, FHIR, Shanghai Data Exchange, Data.gov, IndiaStack, Microsoft, Google ● Amazonは標準規格対応していない為、CADDE(仮称)に取り込む際のデータ標準化対応が必要(IMI共通語彙基盤と文字情報基盤への対応)
	連携コネクタ	<ul style="list-style-type: none"> ● IDSA以外は、各団体が要求している連携方式にCADDE(仮称)のコネクタが適用できるかの個別調整が必要 ● CADDE(仮称)のデータをMicrosoft, Amazonと連携する場合、それぞれのアカウントとデータ保存領域をデータ提供者が個別に有する必要がある ● Microsoft, AmazonのデータをCADDE(仮称)に提供し保存する場合は、CADDE(仮称)側でMicrosoft, Amazonのデータ保存領域を持つことが必要
	データID	<ul style="list-style-type: none"> ● 団体ごとにデータID重複させないようにすることが必要
	認証・認可	<ul style="list-style-type: none"> ● 特になし (公開されている団体に関しては、CADDE(仮称)がすべて対応可能)

(図 56.) 深掘調査対象と連携する際の課題・懸念事項

参考として、団体ごとに制度面・技術面での連携要件を整理すると、下図のとおりであった。

#	団体名	事業名	制度		技術			
			データ提供側信頼性評価	データの信頼性確保	データ・メタデータ基準	連携方式(コネクタ/API)	データID	認証・認可
			各深掘調査対象のデータ提供者になる為、初期に必要な審査	各深掘調査対象のプラットフォームで提供するデータへの初期審査	各深掘調査対象が対応するデータ基準	各深掘調査対象のプラットフォームにデータを連携する手法	各深掘調査対象が利用するデータID	各深掘調査対象が利用する認証・認可手法
1	GAIA-X	GAIA-X	<ul style="list-style-type: none"> GAIA-Xの各エンティティに検証可能な資格(Credentials)を発行 本資格は、WSCの Verifiable Credentials Data Modelに準拠した、第三者によって検証・署名されたクレーム(Claim) 	<ul style="list-style-type: none"> 「クラウドサービスのための欧州サイバーセキュリティ認証スキーム」の採用を検討。共通の評価プロセスとスキームで定義された要件を考慮してICT成果物の適合性を評価<small>ICT Information and Communication Technologies</small> 	<ul style="list-style-type: none"> DCATなどのグローバル基準を採用 	<ul style="list-style-type: none"> GAIA-X又は第三者提供のネクタを通じて連携 	<ul style="list-style-type: none"> リテラル(ユニークコード) 一意の識別子(URI) 	<ul style="list-style-type: none"> SSI (DID, Open ID Connectなど)
2	IDSA	IDSA	<ul style="list-style-type: none"> 参加者とコア・コンポーネントの両方を評価機関より評価 評価後、Certification Authority(CA)より、ペア(参加者とコンポーネント)に固有のIDを生成し、デジタル証明書(X.509)を発行 		<ul style="list-style-type: none"> DCATなどのグローバル基準を採用 	<ul style="list-style-type: none"> IDSコネクタ、第三者提供のネクタ、または自社開発のコネクタを通じて連携 	<ul style="list-style-type: none"> リテラル(ユニークコード) 一意の識別子(URI) 	<ul style="list-style-type: none"> Dynamic Attribute Token (DAT) にて分散型の認証手法を利用
3	HL7 FHIR	HL7 FHIR	<ul style="list-style-type: none"> 各国独自の審査体制を設けて、HL7 FHIRを利用する製品の適合性を確認 		<ul style="list-style-type: none"> 独自定義するデータフレームワークを採用 	<ul style="list-style-type: none"> RESTfulインターフェースを通じて連携可能 	<ul style="list-style-type: none"> リテラル(ユニークコード) 一意の識別子(URI) 	<ul style="list-style-type: none"> Open ID Connect
4	Shanghai Data Exchange	Shanghai Data Exchange	<ul style="list-style-type: none"> 上海CAより企業を対象としてデータ提供者の資格を審査(インタビューより) 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> グローバル基準は参考にしていない独自基準を採用 	<ul style="list-style-type: none"> 非公開 データ提供者より直接データ利用者として連携。データを提供(インタビューより) 	<ul style="list-style-type: none"> 非公開 ただし、データ提供者により設定したIDを流用(特殊な処理をしていない)(インタビューより) 	<ul style="list-style-type: none"> 非公開
5	data.gov	data.gov	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> DCAT-US/Schemaなどのグローバル基準を採用 	<ul style="list-style-type: none"> 非公開 CKANのソースコードを利用する為、直接アップロードとAPIなどを通じてデータ連携方式が可能と推測 	<ul style="list-style-type: none"> リテラル(ユニークコード) ※data.govが提供している実データを確認した結果 	<ul style="list-style-type: none"> API key/Oauth
6	India Stack	India Stack	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 各省庁が独自で定義した基準を採用 	<ul style="list-style-type: none"> 該当なし 外部とのデータ連携は想定していない(インタビューより) 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 各省庁が独自で定義した基準を採用する為、認証スキームは異なる可能性が高い(インタビューより)
7	data.world	Community	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> アカウント認証, API key/Auth認証あり
8	Microsoft	Azure Open Datasets	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> 非公開 	<ul style="list-style-type: none"> MSが定義したコンマデータモデルのような独自基準を採用している為、CADDE(仮称)のデータ・メタデータ基準と異なる 	<ul style="list-style-type: none"> Microsoftより選定し、自社Azure環境に保存 	<ul style="list-style-type: none"> リテラル(ユニークコード) 一意の識別子(URI) 	<ul style="list-style-type: none"> Azure AD
9	Amazon	AWS Data Exchange	<ul style="list-style-type: none"> 米国またはEU加盟国に所在する有効な法人を利用し、有効な銀行および納税者番号を提供して、かつ、AWS Data Exchangeビジネスオペレーションチームによる認定が必要 	<ul style="list-style-type: none"> 該当なし データの信頼性担保は全部データ提供者の責任とする方針(インタビューより) 	<ul style="list-style-type: none"> 該当なし データ基準の対応が要求していない 	<ul style="list-style-type: none"> データ提供者のAWS環境に保存 	<ul style="list-style-type: none"> リテラル(ユニークコード) 一意の識別子(URI) 	<ul style="list-style-type: none"> AWSアカウントを利用した認証
10	Google	Google Cloud Marketplace	<ul style="list-style-type: none"> GCPのPartnerとしてオンボード(インタビューより) 	<ul style="list-style-type: none"> データを商品として登録するたびに、データオーナーシップとビジネスオペレーションを提示して、USの本社より審査(インタビューより) 	<ul style="list-style-type: none"> 自社データをアップロードする際のみDCATとSchema.orgに準拠 他社データに対して基準の準拠は要求されていない 	<ul style="list-style-type: none"> データ提供者のGCP環境に保存 API, FTP経由で外部DBと連携 	<ul style="list-style-type: none"> リテラル(ユニークコード) 一意の識別子(URI) 	<ul style="list-style-type: none"> Googleアカウントを利用した認証

(図 57.) 深掘対象の各団体の連携要件

また、参考として、データ連携にあたり対応が懸念される法規則を整理した。

名称	国・地域	概要	対応が懸念される団体
個人情報保護法	日本 	2020年6月に改正個人情報保護法案が国会で可決。報告義務や規制の厳格化を追加。2022年4月1日までに完全施行される。個人情報扱う全ての事業者が適用対象となる。	Google、Amazon、Microsoft
個人情報保護法 (PIPL)	中国 	2021年8月20日に成立し、同年11月1日から施行。個人情報の収集、使用、保管に関する基本原則ならびに中国国外に拠点を置く企業がデータ処理を行う際の要件も規定している。	Shanghai Data Exchange
データセキュリティ法 (DSL)		2021年9月1日から施行。国家によるデータセキュリティ業務の強化。データの分類と管理、データセキュリティの審査とリスク評価などを確立している。また、厳格なデータ輸出管理と関連主体の義務と責任も明確化されている。	Shanghai Data Exchange
サイバーセキュリティ法		2017年6月に施行。中国国民の個人情報の取り扱いを制限する内容を含む。	Shanghai Data Exchange
個人情報保護法案 (PDPB)	インド 	インド議会合同委員会 (Joint Parliamentary Committee) が審査中。インド国内における個人情報の取り扱い、ならびにインドの公的機関、会社、個人などが基本的な適用対象となる。	India Stack
一般データ保護規則 (GDPR)	EU 	2018年5月から施行。EU域内の個人データ保護を規定する法。2021年6月4日より新たな標準契約条項 (SCC) が発表されており、第三国への国境を越えたデータ移動に関する条項などが更新されている。	GAIA-X、IDSA、HL7 FHIR、data.gov.uk、Google、Amazon、Microsoft
eプライバシー規制		2021年2月10日にEU理事会で合意。通信コンテンツに加えてメタデータもプライバシー保護の対象となる。	GAIA-X、IDSA、HL7 FHIR、data.gov.uk、Google、Amazon、Microsoft
Cloud Act	米国 	2018年3月に可決。米国内に本拠地を持つ企業に対して、米国外に保存されているデータであっても令状なしでデータの開示を要求することが可能となる。	data.gov、Google、Amazon、Microsoft

(図 58.) データ連携にあたり対応が懸念される法規則の一覧

1.2.4. 本調査の総括・まとめ

「分野間データ連携基盤技術」領域において、日本が海外各国の急速な整備対応に遅れをとることなく連携し、国際競争力を維持した安定かつ信頼性のある持続的な運営を目標とし、本国際動向調査を実施した。

< CADDE (仮称) との国際比較 >

CADDE (仮称) は既存の分野内データ連携基盤を連携することを目的としている。そのため、海外団体が既にデータ取込～管理までの一連の機能をサービス提供しているが、CADDE (仮称) は現時点ではユーザ側で機能実装することを要求している。既に CADDE (仮称) で機能実装も完了している Data Catalog や Privacy 関連の機能については海外団体と同水準の機能を提供している。

< 日本にとっての参考事例 >

CADDE (仮称) が実装していない Data Storage/Data Analytics 機能について、他海外団体は実装しており、団体の参入障壁を下げ、サービス利用を契機としたプラットフォーム利用者拡大に繋げている。

< 競合可能性 >

GAIA-X/IDSA/HL7 FHIR は、グローバル範囲でサービス提供しており他団体との連携意思が強いため、優先的に協調していくべき対象といえる。

Google/Amazon/Microsoft/Data.world は、CADDE (仮称) との重複部分が多く連携意思が薄いため、競合対象として注視していくべき対象といえる。ただし技術的には、活用の可能性が存在するとみられる。

<日本との連携課題>

海外団体との連携において、制度面では主に他団体より CADDE(仮称)のデータ提供者やデータ自体への追加審査、技術面では主にデータ・メタデータ基準や連携コネクタ整備に課題感が存在する。

別紙

(別紙 1) 事前調査結果のロングリスト

#	会社・組織名	団体分類	データ基盤構成要素								
			Data Ingestion	Transformation	Data Storage	Analytics	Data Exchange	Data Marketplace	Data Catalog	Privacy and Information Security	Summary
1	GAIA-X	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	8
2	Hesedata	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	8
3	IDS(Industrial)	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	8
4	ISA*	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	8
5	NDS	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	8
6	WEF	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	8
7	FAST	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	7
8	IDS	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	7
9	midata	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	7
10	NIEM	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	7
11	SEMIC	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	7
12	Luxembourgish data platform	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	6
13	Data Exchange	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	5
14	HDX	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	5
15	NDAP	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	5
16	SGFinDex	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	5
17	US Data Federation	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	5
18	Core Vocabulary	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	4
19	Data.gov	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	4
20	data.gov.sg	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	4
21	India Stack	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	4
22	ODX	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	4
23	UNSD	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	4
24	CNGBdb	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
25	data.gov.uk	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
26	Dataportal.asia	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
27	EU Data Portal	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
28	EURDEP	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
29	Green Button	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
30	N-DEX	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
31	PADC	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
32	Qatar Open Data	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
33	SPARQL	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	3
34	Blue Button	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
35	DECODE	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
36	Department of ED Open Data Platform	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
37	DoD	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
38	INITIATIVES	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
39	My Student Data(My Data Button)	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
40	UCORE	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
41	UNCEFACT	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	2
42	Data Catalog Vocabulary	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	1
43	Dublin Core	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	1
44	GLDwg	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	1
45	NIEM-UML profile	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	1
46	中国電子技術標準化研究所	公共団体	✓	✓	✓	✓	✓	✓	✓	✓	1
47	1010DATA	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
48	Acxiom	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
49	AWS	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
50	Bloomberg	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
51	burstIQ	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
52	DATOS	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
53	Datex	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
54	DealCloud	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
55	Equifax	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
56	esri	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
57	EVERRISE	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
58	FactSet	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
59	Farmobile	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8
60	Google	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	8

#	会社・組織名	団体分類	データ基盤構成要素								Privacy and Information Security	Summary
			Data Ingestion	Transformation	Data Storage	Analytics	Data Exchange	Data Marketplace	Data Catalog			
61	IBM Corporation	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
62	Lotame	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
63	LunaPBC	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
64	owner IQ	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
65	Salesforce	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
66	Snowflake	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
67	Streamr	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
68	Thomson Reuters	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
69	wejo	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
70	Alation	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
71	CSC	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
72	Data Republic	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
73	data.world	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
74	DataSwift	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
75	DemystData	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
76	Deutsche Telekom	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
77	Eagle Alpha	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
78	Facebook	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
79	Fixate	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
80	here	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
81	Informatica	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
82	IOTA	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
83	MesInfos	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
84	Morningstar	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
85	ODI	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
86	Oracle	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
87	Precisely	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
88	S&P Global	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
89	SAS	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
90	Turbine	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
91	Adobe	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
92	Alex Solutions	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
93	Apple	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
94	BDX	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
95	CKAN	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
96	Collibra	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
97	Denodo	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
98	erwin	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
99	Microsoft	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
100	otonomo	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
101	Qlik DataMarket	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
102	SAP SE	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
103	Semarchy	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
104	Smartlogic	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
105	TIBCO Software	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
106	Unisys Data Exchange	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
107	Adeptia	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
108	Advaneo	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
109	Alibaba	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
110	Apheris	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
111	ASG	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
112	BDEX, LLC	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
113	Caruso	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
114	Crux Informatics	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
115	digi.me	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
116	Dow Jones	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
117	InfoSum	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
118	InsightSoftware	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
119	Kill	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
120	Mozah	民間団体	✓	✓	✓	✓	✓	✓	✓	✓	✓	5

#	会社・組織名	団体分類	データ基盤構成要素								Privacy and Information Security	Summary
			Data Ingestion	Transformation	Data Storage	Analytics	Data Exchange	Data Marketplace	Data Catalog			
121	Optn Insight	民間団体				✓		✓	✓	✓	✓	5
122	Riversand	民間団体	✓	✓		✓			✓	✓	✓	5
123	Talend	民間団体	✓	✓			✓		✓	✓	✓	5
124	Treasure Data	民間団体	✓	✓		✓				✓	✓	5
125	x-Road	民間団体	✓	✓	✓		✓				✓	5
126	CODY	民間団体	✓	✓			✓			✓	✓	4
127	Databroker DAO	民間団体					✓		✓	✓	✓	4
128	Datarade	民間団体					✓		✓	✓	✓	4
129	Erveil	民間団体	✓	✓			✓		✓	✓	✓	4
130	EvenSense	民間団体					✓		✓	✓	✓	4
131	Harbr	民間団体					✓		✓	✓	✓	4
132	NetApp	民間団体		✓	✓					✓	✓	4
133	Planetwv	民間団体	✓	✓			✓			✓	✓	4
134	Quandl	民間団体	✓	✓							✓	4
135	Comarc n SA	民間団体					✓	✓	✓	✓	✓	3
136	Complementics	民間団体					✓		✓	✓	✓	3
137	Crunchbase	民間団体					✓		✓	✓	✓	3
138	DataProvider	民間団体	✓						✓	✓	✓	3
139	Dataspark	民間団体					✓		✓	✓	✓	3
140	Eight Wire	民間団体					✓				✓	3
141	iGrant.io	民間団体							✓	✓	✓	3
142	MasterCard	民間団体					✓		✓	✓	✓	3
143	Moody's	民間団体					✓		✓	✓	✓	3
144	Narrative	民間団体					✓		✓	✓	✓	3
145	Openrise Data Marketplace	民間団体	✓						✓	✓	✓	3
146	Anxway Software	民間団体						✓			✓	2
147	DataMotion Inc.	民間団体						✓			✓	2
148	DBpedia	民間団体							✓	✓	✓	2
149	egress	民間団体						✓			✓	2
150	MK Data Hub	民間団体							✓	✓	✓	2
151	qDatum	民間団体						✓			✓	2
152	Spring Labs	民間団体					✓				✓	2
153	UBI TECH	民間団体						✓			✓	2
154	Validis	民間団体						✓			✓	2
155	VERACITY by DNV	民間団体							✓	✓	✓	2
156	ADMS	民間団体								✓	✓	1
157	Alteva	民間団体					✓					1
158	Commvault	民間団体									✓	1
159	Datameer	民間団体					✓					1
160	Datajiku	民間団体					✓					1
161	Dell EMC	民間団体			✓							1
162	Element42	民間団体									✓	1
163	HPE	民間団体			✓							1
164	Huawei	民間団体			✓							1
165	Infinidat	民間団体			✓							1
166	LuxTrust S.A.	民間団体									✓	1
167	Mathworks	民間団体					✓					1
168	Okta	民間団体									✓	1
169	Ping Identity	民間団体									✓	1
170	Pure Storage	民間団体			✓							1
171	schema.org	民間団体								✓		1
172	Thales	民間団体									✓	1
173	Veeam	民間団体									✓	1
174	W3C	民間団体								✓		1

(別紙 2) 一次調査対象の 51 団体調査結果

#	深堀調査対象	会社・組織名	サービス・事業名	調査観点1			調査観点2			調査観点3 ※事業別に対応が異なる場合があります。									
				Software Provider	Service Provider	Data Provider	アーキテクチャ	ビジネスモデル	政府関与状況	データ来歴	データ加工履歴	data ID指定方法	データ提供者指定方法	ID認証基盤	メタデータの形式	契約データの形式	利用者ID	DIDなど技術の対応	
1		HDX	-			✓	集中型と分散連邦型の混合形態	Free	政府主導	メタデータで管理	-	-	-	ユーザアカウント/IPアドレス	API key/Auth認証あり	CKAN	データファイル/APIなど	ユーザアカウント/IPアドレス	-
2		digi.me w/Australia Gov	-			✓	PDS (分散連邦型)	個人データにアクセスする企業からアクセスフィーを徴収	Australia政府主導のプロジェクト参加あり	-	-	-	ユーザごとに連携するデータと、第三者への提供可否を設定	ユーザデバイスごとの鍵認証、非対称暗号化・対称暗号化など	-	-	-	データ連携先 (Facebook, Fitbitなど) のID	-
3		FaST	FaST Platform			✓	分散連邦型	Free	EU主導	PostgreSQL 12の機能に準拠	PostgreSQL 12の機能に準拠	PostgreSQL 12の機能に準拠	ユーザアカウント	API key/Auth認証あり	PostgreSQL 12の機能に準拠	API	国ごとに指定されているIdPで認証を取得	不明	
4		Blue Button	Blue Button Platform		✓	✓	分散連邦型	Free	政府主導	-	-	API提供のみでID付与は各種サービスに一任	FHIRに準拠した健康データ提供者	-	-	FHIR準拠健康データ	各種データ提供者のもとで付与	-	
5		CNGBdb	-			✓	集中型	Free	政府主導	メタデータで管理	-	-	ユーザアカウント	FTPで接続の為、関連の認証あり	-	-	データファイル	ユーザアカウント	-
6	✓	Data.gov	-	✓		✓	集中型/分散連邦型	Free	政府主導	メタデータで管理	-	Metadata Catalog ID付与	-	API key/Auth認証あり	DCAT	データファイル/APIなど	2FA活用	-	
7		data.gov.uk.	-		✓	✓	集中型/分散連邦型	Free	政府主導	メタデータで管理	-	-	-	API key/Auth認証あり	ISO 19139/GEMIN 12.3	データファイル/APIなど	User Identification Documents database	-	
8		EU Data Portal	-		✓	✓	分散連邦型	Free	政府主導	データ品質報告あり	-	-	-	API key/Auth認証あり	CKAN	APIなど	Licensing Assistantでユーザ管理	Blockchain活用の検討あり	
9	✓	GAIA-X	-		✓	✓	分散連邦型	Free	政府主導	メタデータで管理	-	-	IdP登録/ユーザアカウント	ISO/IEC24760-1/SSI利用 (DID, Open ID Coonect など)	-	データファイル	IdP登録/ユーザアカウント	分散ID管理や分散公開鍵インフラ (DPKI) 検討	
10	✓	IDSa	-				分散連邦型	Free	-	-	-	X.509	Dynamic Attribute Token (DAT)	W3C Semantic Web技術標準と標準モデリング語彙を活用	-	-	-	-	
11	✓	India Stack	Cashless/Paperless Layer			✓	分散連邦型	Free	政府主導	-	-	Aadhaar番号 (マイナンバーに相当)	-	OTP認証または生体認証	-	API	-	-	
12		Luxembourgish data platform	-		✓	✓	集中型と分散連邦型の混合形態	Free	政府主導	メタデータで管理	-	-	ユーザアカウント	API Key 活用 X-API-KEYは HTTPヘッダーから呼び出す	-	-	データファイル/APIなど	ユーザアカウント	-

#	深掘調査候補	会社・組織名	サービス・事業名	調査観点1			調査観点2			調査観点3 ※事業別に対応が異なる場合があります。								
				Software Provider	Service Provider	Data Provider	アーキテクチャ	ビジネスモデル	政府関与状況	データ来歴	データ加工履歴	data ID指定方法	データ提供者指定方法	ID認証基盤	メタデータの形式	契約データ形式	利用者ID	DIDなど技術の対応
13		Dataportal.asia	-	✓	✓	集中型と分散連邦型の混合形態	Free	-	メタデータで管理	-	-	「ビジネスユーザ」として登録し、ユーザアカウントで管理	ユーザ登録API認証	DCAT Ver2の利用が推奨(またCKANも利用可能)	csv/pdf/xlsx/xls/Json/XML/HTML/ZIP/SH/P7z/png/etc.	公開データの為、登録なしでデータ利用可能API認証あり	-	
14		NDAP	-	✓	✓	集中型	Free	-	メタデータで管理	-	-	ユーザアカウント	API認証	-	データファイル/APIなど	ユーザアカウント	-	
15		data.gov.sg	-		✓	集中型	Free	政府主導	メタデータで管理	メタデータで管理	-	シンガポール政府および法定機関	Public API gatewayを使用	GovTech定義	データファイル/APIなど	オープンデータのため認証不要	-	
16		ODX	-	✓	✓	集中型	Free	政府主導	メタデータで管理	-	-	ユーザアカウント	API key/Auth認証あり	-	データファイル/APIなど	公開データの為、登録なしでデータ利用可能API認証ありAPI認証あり	-	
17		Facebook	-	✓	✓	集中型	-	-	-	-	Facebook指定	ユーザアカウント	API key/Auth認証あり	Facebook指定	データファイル	ユーザ	-	
18		DATOS	-	✓	✓	分散連邦型?	Service fee (おそらく)	なし	-	-	-	-	-	-	-	-	-	-
19		Esri	ArcGIS Open Data	✓	✓	集中型と分散連邦型の混合形態	Service/Storage fee	-	メタデータで管理	-	ObjectIDは設定可能X、Y座標値	ユーザアカウント	API key/Auth認証あり	-	データファイル/APIなど	Open dataの為、ユーザアカウント不要	-	
20		Caruso	Caruso dataplace	✓	✓	集中型	Transaction fee/commission	-	メタデータで管理	-	-	Vehicle Identification Number (VIN)	API key/Auth認証あり	-	データファイル/APIなど	ユーザアカウント	ブロックチェーン技術対応あり	
21		Here	HERE Data Hub	✓	✓	集中型	Transaction fee/commission	-	メタデータで管理	-	-	ユーザアカウント	API key/Auth認証あり	-	データファイル/APIなど	ユーザアカウント	ブロックチェーンベースの許諾管理機能	
22		Wejo	Marketplace	✓	✓	分散連邦型	Service/Transaction fee/commission	-	メタデータで管理各OEMで管理	-	-	OEM関連情報自動車関連情報	API key/Auth認証あり	-	データファイル/APIなど	ユーザアカウント	-	
23		Otonomo	Otonomo Automotive Data Services Platform	✓	✓	集中型	Transaction fee/commission	-	メタデータで管理	とくに記載なしデータは自社で処理する為、加工プロセスは自社で把握	-	OEM関連情報自動車関連情報	API key/Auth認証あり	-	データファイル/APIなど	Client ID, Client Secret and Service IDにて管理	-	
24		Quandl	-	✓	✓	集中型	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録FTP対応ありAPI認証(無料のAPIキー)	-	tables and time-series CSV, XML and JSON	ユーザアカウント	-	
25	✓	AWS	AWS Data Exchange	✓	✓	集中型	Service/Storage Fee/Commission	-	メタデータで管理	Amazon QLDB	-	ユーザアカウント	API key/Auth認証あり	-	データファイル	ユーザアカウント	Blockchain関連技術活用	
26		Microsoft	Research Open Data	✓	✓	分散連邦型	Service Fee	-	メタデータで管理	サイト上で管理	-	Microsoft社内	-	-	データファイル	-	-	

#	深掘調査候補	会社・組織名	サービス・事業名	調査観点1			調査観点2			調査観点3 ※事業別に対応が異なる場合があります。								
				Software Provider	Service Provider	Data Provider	アーキテクチャ	ビジネスモデル	政府関与状況	データ来歴	データ加工履歴	data ID指定方法	データ提供者指定方法	ID認証基盤	メタデータの形式	契約データ形式	利用者ID	DIDなど技術の対応
27	✓	Microsoft	Azure Open Datasets	✓	✓	✓	分散連邦型	Service/Storage Fee/Commission	-	メタデータで管理	対応あり/仕組みは不明	Microsoft指定あり	ユーザアカウント	Directory連携	-	データファイル	ユーザアカウント/Directory連携	分散型IDソリューション活用
28		Microsoft	AIDC Data Cloud	✓	✓	✓	分散連邦型	Commission Fee	-	メタデータで管理	メタデータで管理	-	ユーザアカウント	アカウント認証	-	データファイル	ユーザアカウント	分散型IDソリューション活用
29		Alibaba	Marketplace	✓	✓	✓	分散連邦型	-	-	データ説明	-	-	ユーザアカウント	-	Alibaba指定	APIなど	ユーザアカウント	-
30	✓	Google	Dataset search			✓	分散型	-	-	メタデータで管理	メタデータで管理	Google指定	メタデータ提供	-	-	データファイル/APIなど	-	-
31		Google	GCP		✓	✓	分散連邦型	Service/Storage Fee/Commission	-	メタデータで管理	Cloud Data Fusion	Google指定	ユーザアカウント	API key/Auth認証あり	Schema.org/W3C	データファイル/APIなど	ユーザアカウント	Blockchain関連技術活用
32		SAP SE	SAP Data Warehouse Cloud	✓	✓	✓	集中型/分散連邦型	Service/Storage Fee	-	メタデータで管理	変更履歴など対応	-	ユーザアカウント	API key/Auth認証あり	-	データファイル/APIなど	ユーザアカウント	Blockchain活用の検討あり
33		Snowflake	Data Marketplace		✓	✓	集中型/分散連邦型	Service/Storage Fee/	-	メタデータで管理	クエリ監視機能	-	フェデレーション認証を活用	API key/Auth認証あり	-	データファイル/APIなど	フェデレーション認証を活用	Blockchain活用の検討あり
34		Dawex	Data Marketplace		✓	✓	集中型/分散連邦型	Service/Storage Fee/	-	メタデータで管理	対応あり	-	ユーザアカウント/2FA活用	API key/Auth認証あり	-	データファイル/APIなど	ユーザアカウント/2FA活用	Blockchain関連技術活用
35		data intelligence hub	Marketplace		✓	✓	集中型と分散連邦型の混合形態	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録API認証	-	CSV, geoJSON, JSON, XML, zip, pdf, etc	ユーザアカウント	-
36		IBM Corporation	Data Asset eXchange	✓	✓	✓	集中型と分散連邦型の混合形態	Free	-	メタデータで管理 ※データソース信頼性保証あり	-	-	Githubより共有の、IBM側認証不要	API key/Auth認証あり	-	データファイル	Openの為、認証不要	-
37		Salesforce	Audience studio		✓	✓	集中型	Service/Storage Fee	-	メタデータで管理	データ項目変更履歴/ログ	-	Partner IDにより認証	API key/Auth認証あり	-	データファイル	Salesforce ユーザアカウントID	-
38		Adobe	Audience Marketplace		✓	✓	分散連邦型	Service fee	-	メタデータで管理	記載なし	-	ユーザアカウント	アカウント認証	Adobe定義	データフィード	Adobe ID	不明
39		Data Republic	Data market		✓	✓	集中型と分散連邦型の混合形態	Service/Storage fee	スタートアップ企業の為、政府からの資金援助あり	メタデータで管理	データ追跡機能対応	-	ユーザアカウント	ユーザ登録/SSO対応API認証	-	csv/tsv PDFs / Coding scripts	ユーザアカウント	Privacy-Preserving Matching活用 Blockchain関連技術検討あり
40		Databroker DAO	Data Marketplace		✓	✓	分散連邦型	Transaction fee/commission	-	メタデータで管理	-	-	ユーザアカウント	API key/Auth認証あり	-	APIなど	ユーザアカウント	ブロックチェーン対応あり
41		Streamr	Marketplace		✓	✓	分散連邦型	Transaction fee/commission	-	メタデータで管理	-	-	Ethereum addressを元に認証 StreamID活用	Ethereum keyを利用して、API認証	-	データファイル/APIなど	Ethereum addressを元に認証 StreamID活用	ブロックチェーン技術対応あり
42		Acxiom	ACXIOM DATA		✓	✓	集中型と分散連邦型の混合形態	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録API認証	-	TAR,CSV,TIFF	ユーザアカウント	-

#	深掘調査候補	会社・組織名	サービス・事業名	調査観点1			調査観点2			調査観点3 ※事業別に対応が異なる場合があります。								
				Software Provider	Service Provider	Data Provider	アーキテクチャ	ビジネスモデル	政府関与状況	データ来歴	データ加工履歴	data ID指定方法	データ提供者指定方法	ID認証基盤	メタデータの形式	契約データの形式	利用者ID	DIDなど技術の対応
43		Precisely	Precisely Data Experience		✓	✓	集中型と分散連邦型の混合形態	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録API認証	-	CSV, tab, shapefile, vectorfile	ユーザアカウント	-
44		owner IQ	-		✓	✓	集中型と分散連邦型の混合形態	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録API認証	-	データファイル	ユーザアカウント	-
45		Harbr	Exchange Platform		✓	✓	分散連邦型?		-									
46	✓	data.world	Community		✓	✓	分散連邦型	Service fee	なし	データ加工監視機能あり	データアップロード時明記	-	ユーザアカウント	アカウント認証, API key/Auth認証あり	Data.world定義	データファイル	ユーザアカウント	不明
47		Openprise Data Marketplace	-		✓	✓	分散連邦型	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録API認証	-	データファイル	ユーザアカウント	-
48		Crux Informatics	Crux Platform		✓	✓	分散連邦型	Service Fee	なし									
49		Narrative.io	Narrative Data Stream		✓	✓	集中型	Service/Storage Fee/Commission	-	メタデータで管理	-	-	ユーザアカウント	ユーザ登録API認証	-	csv/text/json and etc.	ユーザアカウント	-
50	✓	Shanghai Data Exchange Corp.	-			✓	-	-	官民結合	データソース説明	ログ記録など対応あり	xIDLabelを用いて個人情報削除	元IDを暗号化したexIDを利用	-	-	-	元IDを暗号化したexIDを利用	-
51	✓	FHIR	-							メタデータで管理				Open ID Connect+OAuth2.0	自社定義	電子文書RESTなど		

(別紙 3) 一次調査対象の 51 団体分類結果

#	深堀調査対象	会社・組織名	サービス・事業名	①サービスのモデル(提供するサービスの内容)			②アーキテクチャ (サービス提供の手段・手法)		③連携形態・データ原則			④使っているテクノロジー									
				中央集約(ニーストラル)or 分散	情報としての状態 (匿名性)	(匿名性の高い団体の場合) 団体によるもの	(収集する団体の場合) 収集元のポイント データ集約に伴うフォーマット	その取り方のサービスへの依存	API	オープンデータ	各種特許や 特許権	ほかデータ収集	認証方式	データ形式	管理方法	ドメイン署名	採用標準				
1		HDX		○	×	民間へのデータ流通			GAFAM	○	○	特許権あり	○	Federated	API key/Auth認証あり				標準定義	CKAN	
2		digi.me w/Australia Gov		○	○		×		GAFAM	○	○	特許権あり	○	Federated	API key/Auth認証あり	ユーザサービスごとの 認証、非対称暗号化・特許権等を含む					
3		FAST	FAST Platform	○	×	民間へのデータ流通			GAFAM	○	○	特許権あり	○	Federated	API key/Auth認証あり						
4		Blue Button	Blue Button Platform	○	×	民間へのデータ流通			GAFAM	○	×	特許権あり	○					API提供のみで他件は 各種サービスに一任			
5		CNGdb		○	×	民間へのデータ流通			GAFAM	○	○	特許権あり	○	Federated		FTPで接続のみ、開 通の認証あり					
6	✓	Data.gov		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	API key/Auth認証あり	自社定義	Metadata Catalog ID付 与		標準定義	DCAT	
7		data.gov.uk		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	API key/Auth認証あり				標準定義	ISO 19139/GEMINI 2.3	
8		EU Data Portal		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	API key/Auth認証あり					標準定義	CKAN
9	✓	GAIA X		○	×	民間へのデータ流通・国 レベルでのデータ思い込み			GAIA X	○	○	複数定義あり	○	Decentralized	ISO/IEC24760-1/SSI利 用(DDI, Open ID Connectなど)	標準定義	SSI (Self Sovereign Identity)		標準定義	W3C	
10	✓	IDSA		○	×	民間へのデータ流通			GAIA X	○	○	複数定義あり	○	Decentralized	Dynamic Attribute Token (DAT)	標準定義	X.509		標準定義	W3C Semantic Web + 標準モ ダリング認定	
11	✓	India Stack	Cashless/Paperless Layer	○	×	民間へのデータ流通			GAFAM	○	×	特許権あり	○	Federated	DTT認証または登録 認証	自社定義	Aadhaar Number / Aadhaar Token、トラン ザクションIDなど			India Stackとして定義ではな く、関連者が定義	
12		Luembourgish data platform		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	API Key 活用 & API KEY はHTTPヘッダから 呼び出す						
13		Dataportal.asia		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	ユーザ登録API認証				標準定義	標準 DCAT Ver2 (CKANも利用 可能)	
14		NDAP		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	API認証						
15		data.gov.sg		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	Public API gatewayを 使用				自社定義		
16		ODX		○	×	民間へのデータ流通			GAFAM	○	○	複数定義あり	○	Federated	API key/Auth認証あり						
17		Facebook		×	○		×		GAFAM	○	○	特許権あり	○	Federated	API key/Auth認証あり	自社定義		Facebook指定		自社定義	
18		DATOS		×	○		×		GAFAM	○	×	特許権あり	○								
19		Esri	ArcGIS Open Data	×	○				GAFAM	○	○	特許権あり	○	Federated	API key/Auth認証あり		その他	ObjectIDは読取可能 & Y 座標読			
20		Caruso	Caruso dataplace	×	○				GAFAM	○	×	特許権あり	○	Federated	API key/Auth認証あり						
21		Here	HERE Data Hub	×	○				GAFAM	○	×	特許権あり	○	Federated	API key/Auth認証あり						
22		Wejo	Marketplace	×	○				GAFAM	○	○	特許権あり	○	Federated	API key/Auth認証あり						
23		Otonomo	Otonomo Automotive Data Services Platform	×	○				GAFAM	○	×	特許権あり	○	Federated	API key/Auth認証あり						
24		Quandl		×	○				GAFAM	○	○	特許権あり	○			ユーザ登録+FTP認証 あり API認証無料の APIキー					
25	✓	AWS	AWS Data Exchange	×	○				GAFAM	○	○	複数定義あり	○	Federated	API key/Auth認証あり						
26		Microsoft	Research Open Data	×	○		×		GAFAM	○	○	複数定義あり	○								

#	採掘調査対象	会社・組織名	サービス・事業名	①サービスのモデル(提供するサービスの内容)			②アーキテクチャ (サービス提供の手段・手法)		③連携形態・データ連携			④使っているAPI/ログイン				ドメイン署名			
				中央集約(ニユートラル/cloud)	組織としての統合 担当者	(収益性のない)団体の場合 団体によるもの	収益化する団体の場合 データ取引に特化フェーズ (手数料等)	収益化する団体の場合 収益化のポイント データの取付サービスへの 取付	API データ	オープン データ	連携形態 連携形態	APIデータ 連携	認証認可	認証方法	データID	管理方法	ドメイン署名	利用標準	
27	✓	Microsoft	Azure Open Datasets	○	○	-	×	○	GAFAM	○	○	複数連携	○	Federated	OAuth等基本的な技術に依存	自社定義	Microsoft指定あり	-	-
28		Microsoft	AIDC Data Cloud	○	○	-	○	○	GAFAM	○	○	複数連携	○	Federated	アカウント認証	-	-	-	-
29		Alibaba	Marketplace	×	○	-	×	○	GAFAM	○	○	複数連携	○	-	-	-	-	自社定義	-
30	✓	Google	Dataset search	×	○	-	×	○	GAFAM	○	○	複数連携	○	-	-	自社定義	Google指定	-	-
31		Google	GCP	×	○	-	○	○	GAFAM	○	○	複数連携	○	Federated	API key/Auth認証あり	自社定義	Google指定	標準定義	Schema.org/W3C
32		SAP SE	SAP Data Warehouse Cloud	○	○	-	?	?	GAFAM	○	○	複数連携	○	Federated	API key/Auth認証あり	-	-	-	-
33		Snowflake	Data Marketplace	×	○	-	○	○	GAFAM	○	○	複数連携	○	Federated	API key/Auth認証あり	-	-	-	-
34		Dawex	Data Marketplace	×	○	-	○	○	GAFAM	○	○	複数連携	○	Federated	API key/Auth認証あり	-	-	-	-
35		data intelligence hub	Marketplace	×	○	-	○	○	GAIA	×	○	複数連携	○	Federated	ユーザ登録API認証	-	-	-	-
36		IBM Corporation	Data Asset eXchange	○	○	-	×	○	GAFAM	○	○	複数連携	○	Federated	API key/Auth認証あり	-	-	-	-
37		Salesforce	Audience studio	×	○	-	×	○	GAFAM	○	○	複数連携	○	Federated	API key/Auth認証あり	-	-	-	-
38		Adobe	Audience Marketplace	×	○	-	×	○	GAFAM	○	×	複数連携	○	Centralized	アカウント認証	-	-	-	自社定義
39		Data Republic	Data market	×	○	-	○	○	GAFAM	○	○	複数連携	○	Centralized	ユーザ登録/SSO対応API認証	-	-	-	-
40		Databroker DAO	Data Marketplace	×	○	-	○	×	GAIA	×	○	複数連携	○	Federated	API key/Auth認証あり	-	-	-	-
41		Streamr	Marketplace	×	○	-	○	×	GAIA	×	○	複数連携	○	Decentralized	Ethereum keyを利用して、API認証	-	-	-	-
42		Axiom	AXIOM DATA	×	○	-	○	×	GAFAM	○	○	複数連携	○	Federated	ユーザ登録API認証	-	-	-	-
43		Precisely	Precisely Data Experience	×	○	-	○	○	GAFAM	○	○	複数連携	○	Federated	ユーザ登録API認証	-	-	-	-
44		owner IQ	-	×	○	-	○	×	GAFAM	○	○	複数連携	○	Federated	ユーザ登録API認証	-	-	-	-
45		Harbr	Exchange Platform	×	○	-	○	○	GAFAM	○	○	複数連携	○	-	-	-	-	-	-
46	✓	data.world	Community	○	○	-	○	×	GAFAM	○	○	複数連携	○	Federated	アカウント認証 API key/Auth認証あり	-	-	-	自社定義
47		Openprise Data Marketplace	-	×	○	-	○	×	GAFAM	○	○	複数連携	○	Federated	ユーザ登録API認証	-	-	-	-
48		Cruze Informatics	Cruze Platform	×	○	-	○	○	GAFAM	○	○	複数連携	○	-	-	-	-	-	-
49		Narrative.io	Narrative Data Stream	×	○	-	○	×	GAFAM	○	○	複数連携	○	Federated	ユーザ登録API認証	-	-	-	-
50	✓	Shanghai Data Exchange Corp.	-	○	○	-	○	○	GAFAM	○	○	複数連携	○	-	-	標準定義	xIDLabelを用いて個人情報削除	-	-
51	✓	FHIR	-	○	×	民間へのデータ流通	-	-	GAIA	×	○	特定連携	○	Federated	Open ID Connect+OAuth2.0	-	-	-	自社定義

(別紙 4) CADDE(仮称)の国際比較結果

※ 各種公開情報より PwC にて整理、以下同じ

		定義	SIP		GAIA-X		IDSA		Google		Microsoft		Amazon	
			CADDE				GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange					
Data Ingestion	1	データ連携方法 データソースとの連携方法。特に連携の際に使用するコネクタ (API) やプロトコルなど。文書データや101データ等、取得するデータにより異なる想定。自動取得が不要な場合は手動での取得なども対象	○	CADDEのコネクタ	△	コネクタを通じて連携 IDSコネクタ Eclipse Data Connector Dataspaces Connector (DSC)	△	コネクタを通じて連携 IDSコネクタ 独自開発のコネクタ、既存のコネクタ (Dataspaces Connector, Trusted Connector, Enterprise Integration Connector など)	△	データ提供者所有のGCP環境を通じて連携 外部からのデータはプログラムによってアクセスすることも可能 (API/FTP) Google パートナーとして契約締結し、パートナーポータルにてデータセット製品を作成・管理	◎	Microsoftより選定した専用の Azure環境に保存し、データ利用者に提供	◎	データ提供者所有のAWS環境にアップロード AWSアカウントを作成し、アカウント内のS3ストレージにデータをアップロード
	2	取込み時エラーチェック データ取込時に実施する、ファイル名やスキーマ (csvファイルなど) 等のエラーチェック	-	(「分野間データ連携基盤技術全体説明書」により)	◎	メタデータチェックあり、データ自体のチェックは対応不明 Self-description (メタデータ関連のチェックは対応しているが、データ自体のチェックに関して関連の記載がない。)	◎	メタデータチェックあり、データ自体のチェックもあり data integrity checkでデータクラスなどをチェックし、データクォリティーを評価する為のデモンショも定義しており、データの品質を確認	◎	(データ提供者からのデータ)メタデータチェックあり、データ自体のチェックは該当しない GCP上でデータ製品を登録する際に、詳細情報によってメタデータを登録する為、登録時のチェックはあり； データに確定のスキーマなど規定されていない為、データスキーマのチェックは対応されていない想定 ※ Googleより公開したデータが schema.org の Dataset マークアップ、または W3C の Data Catalog Vocabulary (DCAT) 形式に準拠	-	メタデータチェック、データ自体のチェックは不明 Microsoftの Azure Open Dataのオプションデプロイプロセスは完全に自社内で対応し、公開されていない為、対応状況は不明(インタビューにより)	◎	メタデータチェックあり、データ自体のチェックは該当しない AWS上でデータ製品を登録する際に、ローデータのスキーマと同様なデータプロダクトを作成したため、データのチェックは対応していない想定
	3	データキャッシュ 一時的に分相に使用する等のデータがアップロードされる場合の処理 (供給先や保存期間、削除方法など)	-	(「分野間データ連携基盤技術全体説明書」により)	◎	データキャッシュ対応あり DIDの認証情報、カタログ情報などのキャッシュ機能を提供、	-	対応状況不明 IDS RAMの技術資料に「Temporal storage」、「Temporal data」、「Cache」関連の記載はない為、一時保存空間として管理する仕組みの開発は対応していない想定	◎	データキャッシュ対応あり GCPの他のサービスによりデータのキャッシュ機能を提供	◎	データキャッシュ対応あり Azureの他のサービスによりデータのキャッシュ機能を提供	◎	データキャッシュ対応あり AWSの他のサービスによりデータのキャッシュ機能を提供

		定義	SIP	GAIA-X		IDSA		Google		Microsoft		Amazon	
				CADDE				GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange			
Data Processing	4	データ統合	多様なデータソースから、一元的にデータベースに格納するための統合処理	△ (開発中)	△ (日立検閲結果により) 表形式データ抽出・加工、意味ノリション機能をもつデータ統合支援プラットフォーム	△	データ統合関連機能を提供	△	データ統合関連機能を提供	△	データ統合関連機能を提供	△	データ統合処理が可能
	5	データ標準化	データの標準化や構造化など、データを連携基盤に格納するための変換・規定のスキーマに合わせてデータ変換、文字コード変換、データ数に合わせてデータ内0埋め、日付フォーマット統一など	△ (開発中)	△ (日立検閲結果により) 表形式データ抽出・加工、意味ノリション機能をもつデータ統合支援プラットフォーム	△	データ変換関連機能を提供	△	データ変換関連機能を提供	△	データ変換関連機能を提供	△	データ変換処理が可能
	6	名寄せ処理	同一の名寄せ、同一とみられるデータを同一IDと紐づける処理。運用の一環として、自動化や、ルール変換に対応していることが望ましい	△ (開発中)	- (「分野間データ連携基盤技術全体説明」により)	△	データ加工関連機能を提供	△	データ加工関連機能を提供	△	データ加工関連機能を提供	△	データ加工処理が可能
	7	仮名化・匿名化	必要に応じた仮名化・匿名加工化などの個人情報加工	△ (開発中)	- (「分野間データ連携基盤技術全体説明」により)	△	データ名寄せ関連機能を提供	△	データ名寄せ関連機能を提供	△	データ名寄せ関連機能を提供	△	データ名寄せ処理が可能
	8	データバイブライン	データ加工プロセスや、(分析データマート等への)データ出力等、他の機能との連携を行う際の实现方法、特に、連携の際に使用するコネクタ(API)やプロトコルなど	△ (開発中)	- (「分野間データ連携基盤技術全体説明」により)	△	コネクタを通じて連携 IDSCコネクタ Eclipse Data Connector Dataspaces Connector (DSC)	△	中央のデータストレージ関連の機能を提供していない IDSAで定義したアーキテクチャにおいて、中央におけるデータストレージが必要ではない為、本機能は対応していない想定	△	データ提供者所有のGCP環境を通じて連携 Google パートナーとして契約締結し、パートナーポータルにてデータセット製品を作成・管理	△	Microsoftより選定した専用のAzure環境に提供 AWSアカウントを作成し、アカウント内のS3ストレージにデータをアップロード

		定義	SIP	GAIA-X		IDSA		Google		Microsoft		Amazon			
				CADDE				GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange					
Data Storage	9	データ保管	文書やメディア情報などを構成するビット・バイト情報を保存する機能	□	- (「分野間データ連携基盤技術全体説明」により)	△	データ状態の管理機能を提供 データの格納、更新、出力など一連のプロセスをモニタリングして、ログ情報として残す機能	□	中央のデータストレージ関連の機能を提供していない IDSAで定義したアーキテクチャにおいて、中央におけるデータストレージが必要ではない為、本機能は対応していない想定	△	GCPの他のサービスによりデータ処理の機能を提供	◎	Azureの他のサービスによりデータ処理の機能を提供	◎	AWSの他のサービスによりデータ処理の機能を提供
	10	バックアップ	蓄積期間を超えるデータの保管方針や、データベースのバックアップ方針の定義	-	- (「分野間データ連携基盤技術全体説明」により)	△	自動的なバックアップ・アーカイブ機能を提供 自動的なバックアップルーティン、または誤って削除してから回復が可能になるようにデータのバックアップ機能を開発	-	中央のデータストレージ関連の機能を提供していない IDSAで定義したアーキテクチャにおいて、中央におけるデータストレージが必要ではない為、本機能は対応していない想定	△	GCPの他のサービスによりデータ処理の機能を提供	◎	Azureの他のサービスによりデータ処理の機能を提供	◎	AWSの他のサービスによりデータ処理の機能を提供
	11	データ蓄積期間	データベース上にデータを保持する期間の定義	-	- (「分野間データ連携基盤技術全体説明」により)	△ (開発中)	必要に応じて今後対応可能 今後、ユーザの要望に応じてデータオプティミゼーションの際に蓄積期間の指定が可能になるような取組を検討	-	中央のデータストレージ関連の機能を提供していない IDSAで定義したアーキテクチャにおいて、中央におけるデータストレージが必要ではない為、本機能は対応していない想定	△	GCPの他のサービスによりデータ処理の機能を提供	◎	Azureの他のサービスによりデータ処理の機能を提供	◎	AWSの他のサービスによりデータ処理の機能を提供
	12	アーカイブ	蓄積期間を超えるデータの保管方針や、データベースのバックアップ方針の定義	-	- (「分野間データ連携基盤技術全体説明」により)	△	自動的なバックアップ・アーカイブ機能を提供 自動的なバックアップルーティン、または誤って削除してから回復が可能になるようにデータのバックアップ機能を開発	-	中央のデータストレージ関連の機能を提供していない IDSAで定義したアーキテクチャにおいて、中央におけるデータストレージが必要ではない為、本機能は対応していない想定	△	GCPの他のサービスによりデータ処理の機能を提供	◎	Azureの他のサービスによりデータ処理の機能を提供	◎	AWSの他のサービスによりデータ処理の機能を提供

			定義	SIP	GAIA-X	IDSА	Google	Microsoft	Amazon
				CADDE			GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange
Data Analytics	13	データ分析	分析基盤上で分析を実施する場合の分析環境など	□ - (分野間データ連携基盤技術全体説明)により)	△ データアナリティクス関連のデータ分析機能を提供(サードパーティーサービス) Analyticsやアプリケーションを提供 ※ユースケース構築・サービス開発中	△ データアナリティクス関連のデータ分析機能を提供(サードパーティーサービス)	△ GCPの他のサービスによりデータ分析の機能を提供 又、データを入力して分析することも可能	△ Azureの他のサービスによりデータ処理の機能を提供 又、データを入力して分析することも可能	△ AWSの他のサービスによりデータ処理の機能を提供 又、データを入力して分析することも可能(一部データ提供者の指定によってAWS環境内の活用のみが可能)
	14	データ可視化	分析基盤上で図によるデータの可視化を実施する場合のツールや保存期間、対象データソースなど	□ - (分野間データ連携基盤技術全体説明)により)	△ データアナリティクス関連のデータ表現機能を提供(サードパーティーサービス) サービスとしてデータビジュアライゼーションを提供 ※ユースケース構築・サービス開発中	△ データアナリティクス関連のデータ表現機能を提供(サードパーティーサービス) サードパーティーからのService Providerによりデータ統合関連のサービスを提供 ※サービス開発ステータス不明	△ GCPの他のサービスによりデータ分析の機能を提供 又、データを入力して可視化することも可能	△ Azureの他のサービスによりデータ処理の機能を提供 又、データを入力して可視化することも可能	△ AWSの他のサービスによりデータ処理の機能を提供 又、データを入力して分析することも可能(一部データ提供者の指定によってAWS環境内の活用のみが可能)

			定義	SIP	GAIA-X	IDSА	Google	Microsoft	Amazon
				CADDE			GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange
Data Exchange& Marketplace	15	データ提供	データを外部に提供する際の提供方法や、提供する際のフォーマット、実現するためのコネクタ、通信プロトコルなど	○ CADDEのコネクタ	△ コネクタを通じて連携 IDSコネクタ Eclipse Data Connector Dataspaces Connector (DSC)	△ コネクタを通じて連携 IDSコネクタ 独自開発のコネクタ、既存のコネクタ(Dataspaces Connector, Trusted Connector, Enterprise Integration Connector など)	◎ データ利用者所有のGCP環境を通じて連携 GCPのBigQueryというサービスを通じてデータを取得	◎ データ利用者所有のAzure環境を通じて連携 ※一旦Azure環境からSDKファイルをダウンロード後、自由に利用可能 Azure Machine Learning UIとSDK(Pythonなど)を通じて利用	◎ データ利用者所有のAWS環境を通じて連携 同じくAWSのクラウド環境内で、データ提供者のアカウントからデータ利用者のアカウントへ複製 Amazon Redshift向け AWS Data Exchange では、お客様は ETL を必要とせず、AWS Data Exchange で見つけたサードパーティーのデータと、自社のファーストパーティーのデータを Amazon Redshift のクラウドデータウェアハウスで組み合わせることが可能
	16	請求管理	データ活用の目的でデータを抽出するユーザに対して、自動的に請求先・請求額を管理する機能	△ 契約管理の仕組みあり(契約管理ポータル、課金)	◎ 個別契約の中で課金ルール設定(請求機能を開発中) 現在は個別契約の中で課金ルール設定するような形で機能する状態 正式な請求機能は開発中、今後 Gaia の課金サービス(Billing Service)よりData Exchange Logging Serviceのデータ送込ログを構築し、自動的に請求を開始	◎ 個別契約の中で課金ルール設定(請求機能を開発中) Cleaning&Billing機能より提供 Financialは現在開発中の為、実態は把握できていない	◎ 課金の仕組みあり クラウドサービス機能 クラウドサービスCloud Billingより管理 Billingの情報をBigQueryに出すことも可能	◎ 課金の仕組みあり クラウドサービス機能 クラウドサービスCost Management + Billing機能より管理 Azure上利用するサービスに関するコスト分析機能もあり	◎ 課金の仕組みあり クラウドサービス機能 クラウドサービスBilling and Cost Management consoleより管理
	17	契約支援	契約するためのコミュニケーション機能、契約内容を管理する機能	△ 課金の仕組みあり(契約管理ポータル、課金)	◎ 現状個別に契約締結 契約管理の仕組み構築中 Data Contract Service という契約管理の仕組みを構築中(自動審査の構想)で、機能としてはO/RLに基づいて、人間と機械両方が読める形式でバックジャージ化することが可能	◎ 現状個別に契約締結 契約管理の仕組み構築中 Cleaning&Billing機能より提供 Financialは現在開発中の為、実態は把握できていない + Blockchain活用によるスマートコントラクトについて検討中	◎ 契約管理の仕組みあり データ利用者はGCPアカウント データ提供者はパートナー契約	◎ 契約管理の仕組みあり データ利用者はAzureアカウント データ提供者は現状オープンデータのみであるため契約管理不要	◎ 契約管理の仕組みあり データ利用者はAWSアカウント データ提供者はパートナー契約

		定義	SIP	GAIA-X	IDSA	Google	Microsoft	Amazon
			CADDE			GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange
Data Catalog	18	データカタログ データに関する情報の定義 (論理的物理的なモデル構成、データソース、定義、算出方法等) ※ メタデータ管理も本項目に含まれる想定	- (「分野間データ連携基盤技術全体説明」により)	グローバル基準採用 W3C Organization Ontology, community-maintained schema.org vocabulary, W3C Data Catalog Vocabulary DCAT, W3C Open Digital Rights Language, the International Data Spaces Information Model	グローバル基準採用 W3Cモデルを利用(DCAT7, ODR1.8等) ※ 国際標準を参考にIDSAでデータカタログ内に記載すべき項目の基準を定義	メタデータ項目は自社定義、データ提供者などにより手動入力 料金プランと詳細、提供されるサービスの説明、SaaSページに表示されるアイコンなど	メタデータ項目は自社定義、データ提供者などにより手動入力 データ請求時メールにて以下の項目を記入 Dataset name/Dataset description/Dataset size/Dataset file formats/License/Use cases	メタデータ項目は自社定義、データ提供者などにより手動入力
	19	ID管理 データセットのID管理	リテラル(ユニークコード)一意の識別子(URI)で識別(データのIDもURLを使うと越塚先生からのコメント) IMI共通語彙基盤、文字情報基盤	データID: リテラル(ユニークコード)一意の識別子(URI)で識別 Self-Descriptionsよりデータ、サービスまたは参加者のIDを管理し、データIDはリテラル(ユニークコード)又は一意の識別子(URI)で識別	データID: リテラル(ユニークコード)一意の識別子(URI)で識別 Identity Managementより管理IDとなるものは、異なる使用シナリオやドメインでリネーム/共有/共有できる、再利用可能な概念に似て、規範となるリテラル(コード)または一意の識別子(URI)で識別される各用語となる	データID: リテラル(ユニークコード)で一意の識別子(URI)で識別(公開データにアクセスするには)	データID: リテラル(ユニークコード)で一意の識別子(URI)で識別(一部のみ)	データID: リテラル(ユニークコード)で一意の識別子(URI)で識別(一部のみ)
	20	語彙管理 特に各業界に固有の語彙の管理方法(データカタログのルール定義)	IMI共通語彙基盤、文字情報基盤	語彙: ※データカタログと同様	語彙: ※データカタログと同様	語彙: 他社データの場合、データ提供者からのデータそのままだ連携する為、語彙の管理は対応していない可能性が高い 自社(Google)データの場合、Schema.org/DCATに準拠	語彙: ※データ提供者からのデータそのままだ連携する為、語彙の管理は対応していない可能性が高い	語彙: ※データ提供者からのデータそのままだ連携する為、語彙の管理は対応していない可能性が高い
	21	データカタログ検索 特定の条件に基づいたデータの検索を行う機能	検索エンジン開発	検索対象はメタデータ検索 データ検索は検索エンジンのサービスとしてGAIA-Xに実装し、カタログ(メタデータ)を対象として検索	検索対象はメタデータ検索 Data Brokerより実現Data Brokerでは、データのメタデータを保存する為、メタデータを対象として検索	検索対象はメタデータ検索 メタデータにある文字の検索が可能、ワイルド検索 ※メタデータに検索エンジン最適化用の検索メタデータと検索キーワードを追加することも可能	検索対象はメタデータ検索 ※ Azure Open Dataは対応していないが 現在開発中のAzure Data Marketplaceにおいて、メタデータにある文字の検索が可能、ワイルド検索	検索対象はメタデータ検索 メタデータにある文字の検索が可能、ワイルド検索
	22	レーティング データごとのレーティングを行う機能	- (「分野間データ連携基盤技術全体説明」により)	関連記事のない為対応状況が不明 GAIA-Xの技術資料に「Rating」関連機能の記事はない為、対応されていない想定、	データレーティング関連機能を提供 CommunityからのFeedBack、又はQuality of Service(QoS)にてレーティングを管理	公開されたデータ詳細情報がない為対応状況が不明	公開されたデータ詳細情報がない為対応状況が不明	公開されたデータ詳細情報がない為対応状況が不明
		定義	SIP	GAIA-X	IDSA	Google	Microsoft	Amazon
			CADDE			GCP Data Marketplace	Azure Open Data/Marketplace	AWS Data Exchange
Privacy and Information Security	23	トラスト データの存在証明(原本性保証含む)や、タイムスタンプ付与により、データの信頼性を担保	-	トラスト関連機能を提供 「ビット(参加者)及び「モノ」データなどのアセット」の両方からトラストを提供 The Gaia-X Trust Service各参加者のDID又はパブリックキー	トラスト関連機能を提供 参加者及び参加者が利用するコンポーネント(コネクタなど)にユニークなIDを作成し、デジタル証明書としてX.509を発行	トラスト関連機能を提供 信頼性を確保する為、GCPのセキュリティモデル(暗号化など)、コンプライアンスと規則対応(認証/法令など)、セキュリティ向上するためのプロジェクトを合わせて提供	トラスト関連機能を提供 トラストセンターよりコンプライアンスや法令規則対応の状況を公開	トラスト関連機能を提供 AWS Data Exchangeでは、IAMアクセス許可ポリシーの利用が可能 IAMアクセス許可ポリシーは、アイデンティティ(ユーザー、ユーザーのグループ、ロール)やリソースに関連付けて、これらのアクセス許可を定義
	24	品質管理 定期的なデータの重複性チェックや、エラーチェックをした、データ品質の担保を行う機能 (過去データにも適用されるルール変更があった場合に、そのルールが正しく反映されているかのチェック等)	- (「分野間データ連携基盤技術全体説明」により)	品質管理関連機能を提供 データの品質を担保する為にデータ品質を評価する機能	品質管理関連機能を提供 コネクタの拡張機能として自動化されたデータ品質チェック基盤を構築	品質管理関連機能を提供 GCPの他のサービスによりデータ処理の機能を提供	対応状況不明 Microsoft認定のOpen Datasetsに関して、データ品質に関する対応有無、対応詳細は非公開の為対応状況不明	品質管理関連機能を提供(同社の他サービスより) データの正確性ができず、自動的データ品質確認が対応していない想定
	25	アカウント権限管理 業務役割に応じたアカウント/権限の管理、ユーザー認証方法など	-	ロールを2つ区分して権限管理対応 GAIA-X参加者を分けて権利と義務を定義: 提供者/利用者/デジタル/アイデンティティ提供者/データオーナー	ロールを12つ区分して権限管理対応 参加者を分けて権利と義務を定義: Provider/Data Owner/Service Provider/Broker Service Provider/Cleaning House/Identity Provider/ App Store Provider/Software Provider/App Provider/Vocabulary Provider/Certification Authority	ロールを区分して権限管理対応 Identity and Access Management (IAM)より管理・プリンシパル/プリンシパルは、リソースへのアクセスが許可 ・役割:ロールは権限のコレクションです ・ポリシー:IAM ポリシーは、1つ以上のプリンシパルを特定のロールにバインドするロールバインディングの集合	ロールを7つ区分して権限管理対応 ※ Open datasets関連のロールではなく、Microsoft全製品のロール数 Azure ADより管理 Microsoftのすべてのサービスをカバーする為、全部で約7種のRoleを定義	ロールを主に2つ区分して権限管理対応 AWSのアカウントとしてアクセス可能な機能として、AWS License Managerでマネージド使用権限を管理が可能 主にデータ提供者とデータ利用者に分けられる認証

(別紙 5) 各団体調査結果詳細

※ 各種公開情報やヒアリングを通じ、PwCにて整理、以下同じ



技術 アーキテクチャー

GAIA-Xはデータとインフラストラクチャーエコシステム、それを支えるポリシー・標準で構成。大きく5つの技術要素でEU向けのサービスを提供

アーキテクチャー	技術要素	説明	ユーザ要件
<p>データエコシステム</p> <p>アドバンスドスマートサービス</p> <p>データスペース</p> <p>ソフトウェア分断 ・ 多リージョン環境 ・ クラウド間の移動性 ・ 仮想化コンテナ</p> <p>アイデンティティ・アクセス ・ フェデレーション環境 ・ トラスト管理 ・ フェデレーションアクセス</p> <p>フェデレーション 制御 ・ 自己管理 ・ サービスカタログ ・ 通知と検閲</p> <p>コンプライアンス ・ キー・プロセス ・ コンプライアンス ・ 権利者の権利と保護 ・ オンボード・脱退</p> <p>インフラストラクチャーエコシステム</p> <p>ポリシー・標準 (標準のアーキテクチャ (Architecture of standards), 投票)</p> <p>参加者 Provider Consumer Both</p> <p>物理ストレージ データベース Physical Logical</p>	Advanced Smart Service	<ul style="list-style-type: none"> 分野を越えたイノベーション/マーケットプレイス/アプリケーション AI、分析、またはビッグデータ関連の技術を活用し、GAIA-Xエコシステムのイノベーションを促進 次世代のデジタルインフラストラクチャのための持続可能なビジネスとイノベーションのエコシステムの構築を促進 	
	Data Spaces	<ul style="list-style-type: none"> 相互運用可能なポータブルな分野間データセットとサービス デジタルサービスとメカニズムに基づく価値創造を促進するために、ドメイン共通の基盤となるアーキテクチャが必要 EUデータ戦略に従って、分野内と分野間のデータ連携のための相互運用性とAPIのオンロジーを提供 データ主導のバリューチェーンにおけるコラボレーションを促進 	
	GAIA-X Federation services	<ul style="list-style-type: none"> 相互運用性の信頼と主権を担保する上で分散するシステムを統合 フェデレーションGAIA-Xエコシステムを運用するために必要な最小限の技術要件とサービスを特定 これらのサービスの間隔は、Security-by-Designの原則に従い、Privacy-by-Designの概念も必要 	
	Portability, Interoperability & Interconnectivity	<ul style="list-style-type: none"> 技術上：標準のアーキテクチャー コマース上：ポリシー ネットワークおよび相互接続プロバイダー、クラウドソリューションプロバイダー (CSP)、ハイパフォーマンスコンピューティング (HPC)、セクター固有のクラウドおよびエッジシステムの間の連携 ユーザーフレンドリーなインフラストラクチャーエコシステムを実現するための連携サービス 分散のユースケースを支援/完全な技術スタックを実現 	
	Compliance	<ul style="list-style-type: none"> リーガル：レギュレーション & ポリシー - 	

技術 データフォーマット・メタデータ対応基準

GAIA-Xはグローバルで定められているデータ・メタデータの標準を採用。独自のルールは使用しない方針

データフォーマット	<p>Json-LD</p> <ul style="list-style-type: none"> JSON-LDは、軽量のLinkedData形式 人間は読み書きが簡単 JSONデータをWebスケールで相互運用するのに役立つ方法を提供 JSON-LDは、プログラミング環境、REST Webサービス、およびApache CouchDBやMongoDBなどの非構造化データベースにとって理想的なデータ形式
データ・メタデータ基準	<p>DCAT*</p> <ul style="list-style-type: none"> DCATは、Web上で公開されているデータカタログ間の相互運用性を促進するように設計されたRDF語彙 DCATを使用すると、発行者は、複数のカタログからのメタデータの消費と集約が容易 標準モデルと語彙を使用して、カタログ内のデータセットとデータサービスを記述することが可能 データセットとデータサービスの発見可能性を向上

* DCAT以外にも RDF, OWL, VoID という標準（いずれも W3C にて策定）への対応もドキュメントに記載あったが、詳細不明

技術 データ来歴、加工記録技術

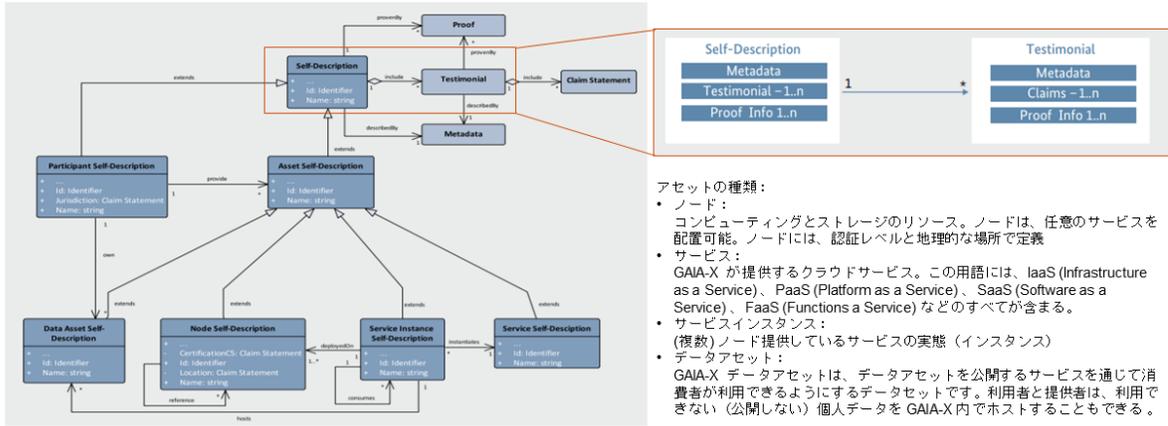
GAIA-Xに登録されたデータ等のアセットはデータエコシステムの開発を可能にするために、監視のための標準的なメカニズムを提供 ※GAIA-Xに登録するまでの記録対応状況は公開されていない

目的	<ul style="list-style-type: none"> GAIA-Xは法規対応、運用、透明性の向上などを以下の3つの機能により対応 		
ロギングと監査	モニタリングとアラート送信	メータリング(Metering)	
<ul style="list-style-type: none"> ロギングは、GAIA-Xのサービスまたはノードによって生成されるランタイムログ情報へのアクセスを指す GAIA-Xのサービスやノードの中には、法律や契約上の要求から、監査を必要とするものがあり、多くの場合、監査のためのログは、別のシステムに転送され、改ざん防止された方法で保存される必要がある 透明性を高め、多くのベンダーのサービスを統合するために、標準的なインターフェースを提供 	<ul style="list-style-type: none"> GAIA-Xにおけるモニタリングは、サービスやノードのステータス情報へのアクセス、およびアラートの送信を指す モニタリングは大規模な分散アプリケーションの運用に不可欠 GAIA-Xはモニタリングのための標準的なメカニズムとインターフェースを定義 GAIA-Xのモニタリング定義は、監視のための技術的なインターフェースと、監視レベルや監視対象の分類などの概念的な定義の2つのレベルで構成 	<ul style="list-style-type: none"> 一般的に、メータリングはモニタリングと似ており、特にパフォーマンス指標や消費統計へのアクセスを意味。メータリングは、課金の透明性を高めるだけでなく、大規模なクラウドアプリケーションのリソース効率の高いスケーリングと運用にも重要 GAIA-X自体は、課金プロバイダやクリアリングハウスとして機能しないものの、消費者やプロバイダーが使用する計測のための標準的なインターフェースやメカニズムを定義 	

技術 ID関連技術

GAIA-Xにおける各種アセット（ノード、サービス、サービスインスタンス、またはデータアセット*）と参加者のそれぞれのIDと、他の特徴に合わせてSelf-Descriptionにて管理

トップレベルのSelf-Descriptionのスキーマ



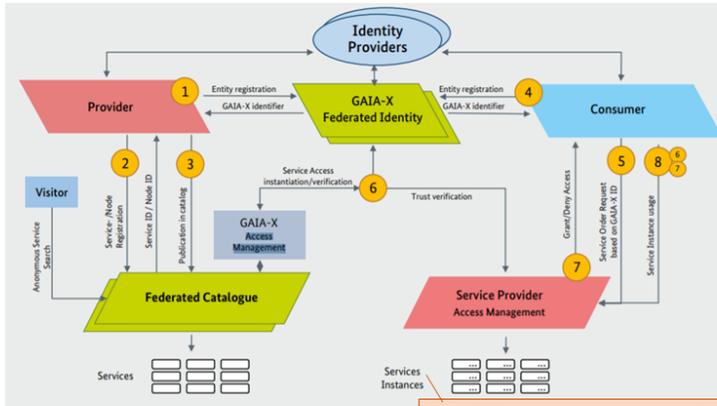
技術 ID関連技術 参加者ID

GAIA-Xにおいて、各種ユーザ及び実行可能な権限は5つ定義されている。各種ユーザごとに権限を振り分けることにより、権限管理をシンプルに設定する意向

参加者	権利と義務
提供者 (Provider)	<ul style="list-style-type: none"> 提供者は、GAIA-Xの登録手続きを行い、本人確認が必要 提供者は、GAIA-Xサービス契約の履行が必要 登録後の提供者の責務は、提供したGAIA-X資産のコミッションングとデコミッションであり、各アセットには説明文があり、その内容の確認が必要 提供された情報は十分かつ正確であることが必要 提供者は提供するGAIA-X資産の機能的責任を担当 提供者は、データ所有者および消費者との間で取り決められた契約の遵守が必要
利用者 (Consumer)	<ul style="list-style-type: none"> 利用者は、GAIA-Xの登録手続きを行い、本人確認を行うのが必要 利用者は、GAIA-Xサービス契約を履行が必要 全ての利用者はGAIA-Xによって平等に対応 利用者は自分の要求に応じてアセットを検索することが可能 契約交渉は提供者と利用者間で実行 GAIA-Xは仲介の役割を果たさず、両者に信頼性を提供
ビジター (Visitor)	<ul style="list-style-type: none"> ビジターはGAIA-Xのアセットを制限なく閲覧、ナビゲート、検索することが可能 ビジターがアセットを利用する場合は、利用者として登録またはログインする必要
IDプロバイダー (Identity Providers)	<ul style="list-style-type: none"> IDプロバイダー (IdP) は、GAIA-Xの法的要求事項の遵守が必要 IdPは、GAIA-Xの資産および参加者のアイデンティティの保証、およびアイデンティティ・ライフサイクル・プロセスを担当
データオーナー (Data Owner)	<ul style="list-style-type: none"> データ所有者は、サービスのプロバイダー (サービスとしてのデータ) に通じてデータを提供 (Data as a Service) データ所有者は、プロバイダーと交渉した契約の遵守が必要

技術 認証

GAIA-XはFederated Identityの機能により、連携データ基盤の認証が透過的に認証されるように実行（個別連携先ごとの認証は不要）



認証の中核として、OpenID ConnectやOAuth2などの承認フレームワークなどの既存の十分に活用された認証プロトコル、またはX.509ベースの相互TLS認証の有効化との互換性を保証

内容

- 匿名サービス検索

1 提供者エンティティの登録

2 サービス登録

3 カタログ掲載

4 利用者登録

5 サービスオーダーリクエスト

6 信頼性検証

7 許可/拒否 アクセス拒否

8 サービス利用状況

(参考) 技術 認証プロセス詳細

# 内容	詳細なステップの説明
- 匿名サービス検索	提供者はGAIA-X統合IDにアクセスし、GAIA-Xカタログを閲覧し、サービス検索クエリを開始。サービス検索条件に一致する可能性のあるサービスのリストがビジターに表示
1 提供者エンティティの登録	提供者のエンティティをGAIA-Xに登録。必須項目の一つにIdPの入力があり、IdPは提供者エンティティの身分証明を行う必要がある。提供者にはGAIA-X ID（識別子）が提供され、GAIA-Xに認証・登録
2 サービス登録	提供者はGAIA-X統合カタログにサービスを登録することが可能。サービスIDはGAIA-Xが生成し、提供者が取得
3 カタログ掲載	登録されたサービスは、GAIA-Xの統合カタログに掲載され、検索アルゴリズムで公開
4 利用者登録	利用者をGAIA-Xに登録。その際、必須項目の一つとしてIdPの入力を行う。IdPは利用者エンティティの身分証明を行う必要があり、GAIA-X自身が検証することが可能。GAIA-X ID（識別子）は、利用者エンティティに提供され、利用者はGAIA-Xで検証された後登録
5 サービスオーダーリクエスト	登録された利用者は、特定のサービスを注文するためにサービス提供者に連絡
6 信頼性検証	サービス提供者は、提供者の信頼性を確認。GAIA-X Federated Identity は IdP を介して身分証明を行う。GAIA-X 統合アイデンティティはサービスアクセス（利用者の属性→健康データ）を検証し、結果をサービス提供者に返送。サービス提供者は受信した属性を検証し、利用者の識別子を作成
7 許可/拒否 アクセス拒否	サービス提供者は、その結果を利用者に提供。許可されると、サービス提供者は、サービスオーケストレーションエンジンを起動して、利用者向けのサービスを作成（=サービスインスタンス化プロセス）。サービス提供者はサービスインスタンス識別子を利用者に転送
8 サービス利用状況	利用者は、注文されたサービス・インスタンスを使用することが可能。サービス利用中、サービス提供者AMは、アクセスごとに利用者の身元を確認し、利用者の属性が必要なものと一致していることを保証（ステップ6/7参照）

技術 認証の規格

GAIA-Xは分散型の認証・認可を実現する為にSSI (Self Sovereign Identity) を採用
SSIを採択しない団体と連携する為には、SSIとのインターフェースが必要であり、具
体的な対処方法は検討中

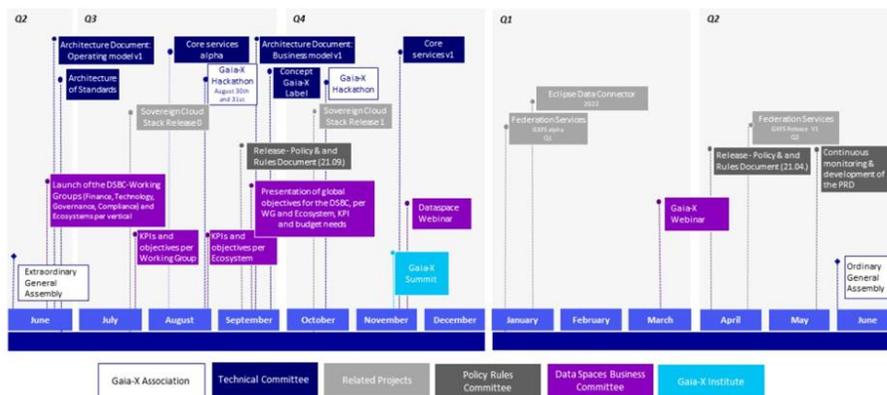
SSIの実現にむけて、GAIA-Xで検討中の認証規格

	DID (W3C)	Active Directory (Microsoft)	Open ID Connect	KeyCloak	SAML	ウォレット付きID
詳細	DIDは、検証可能な分散型デジタル・アイデンティティを実現する新しいタイプの識別子	Windows Serverにおいて、ID管理に関連するすべてのものが、Active Directoryによって管理可能	オープンスタンダードで分散型の認証プロトコル	オープンソースのIDおよびアクセス管理ソリューション	IDプロバイダー (IdP) がサービスプロバイダー (SP) に認証情報を渡すためのオープンスタンダード	個人々のウォレットに加えて、組織では証明書やマネージャーがわかる構成 GAIA-Xのレジストリが統合サービスの管理等を実行
GAIA-Xでの検討方針	<ul style="list-style-type: none"> DIDは9月にドラフトペーパーにリグレードし、検討を進めていく方針 W3Cは一部企業の反対があったため、引き続き検討しながら、DID以外の代替手法も検討 年末に向けて、より議論を成熟させていく 	<ul style="list-style-type: none"> 前向きに対応を検討 	<ul style="list-style-type: none"> 前向きに対応を検討 	<ul style="list-style-type: none"> 前向きに対応を検討 	<ul style="list-style-type: none"> 識別レベル (Level of Identification) が不十分であり、不適合として、GAIA-Xのロードマップから外 	<ul style="list-style-type: none"> 現状のコンセプトとして推進

SSI採択背景・特記事項	<ul style="list-style-type: none"> ✓ SSIの採択背景として中央集権的な基盤はセキュリティや性能観点から失敗になる為、分散型(Decentralized)の認証・認可が求められている状態 ✓ SSIを採択しようとしても、現状すべての参加団体が採択しているわけではない ✗ 例えばBMWなどは、SSIの採択を強制化していない。そのため、既存の認証サービスとSSIとのインターフェースが必要 ✓ Googleが採択しているような現状のレベルよりも、より高い識別レベルが求められる
--------------	---

技術 GAIA-Xのロードマップ

公開されているGAIA-Xロードマップによると、来年以降の戦略として、データスペース、インフラストラクチャ、フェデレーションサービスの技術分野、および共通のルールとガイドラインの体系を設立することを予定

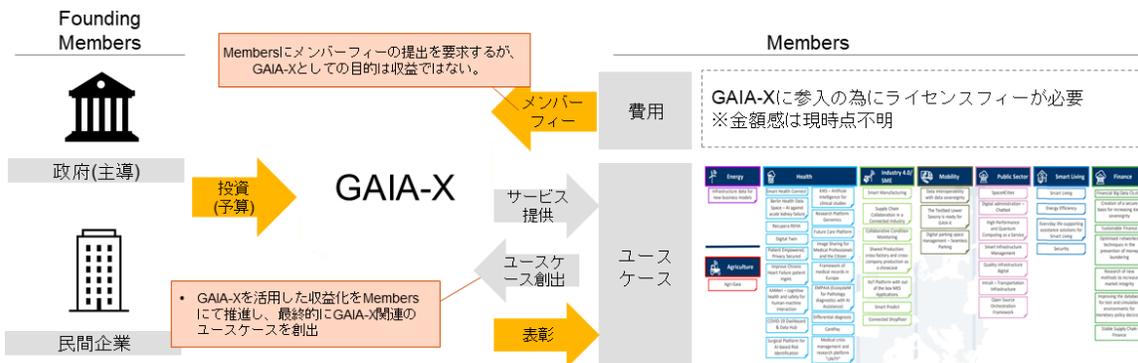


技術 GAIA-Xの今後の展望

#	カテゴリ	検討内容
1	標準化	<ul style="list-style-type: none"> 欧州の標準化団体であるETSIやCEN-CENELEC等、関連するICT Roadmapとの同期 GAIA-Xのコアとなるコンセプト(統合的なアプローチ)が、IDSAやOCEANプロトコルなどのデータ共有の取り組みとのインキュベートを実現する為の検討 Tier IVなどのカタログ機能のプロバイダーとの議論も進めており、技術的な判断が必要
2	認証	<ul style="list-style-type: none"> DAO (Decentralized Automated Organization) というアプローチも検討していたが、関連組織がこのアプローチに対応できないことが分かってきているため、トーンダウンしている
3	インフラストラクチャー	<ul style="list-style-type: none"> Distributed Ledger Technology (DLT) 技術に関しても、今後継続適用していく為に、エネルギーや、サステナビリティの観点の検討が必須 ※欧州でSDGsの検討が義務化されている

ビジネスモデル 全体図

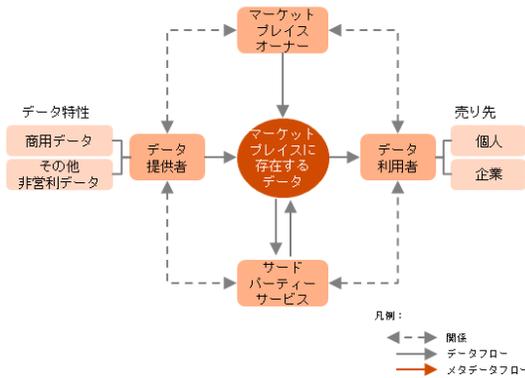
GAIA-X自体はデータ利用料徴収などの収益確保が主な目的ではなく、各領域におけるユースケースの創出が目的



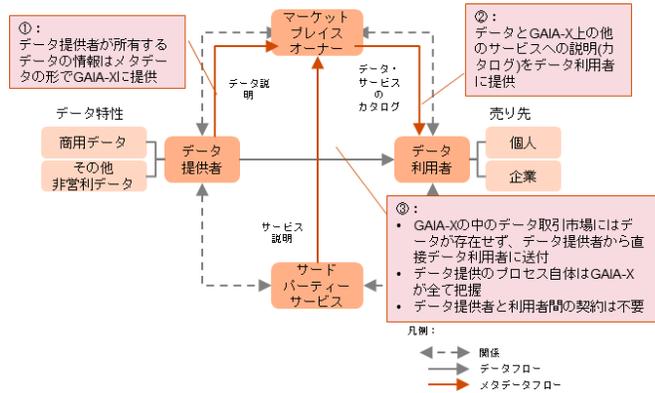
ビジネスモデル GAIA-Xでのデータ取引市場について

GAIA-Xの中のデータ取引市場は、適切なデータソースを見つける為の「イエローブック」であり、データのやり取りはデータ提供者と利用者の間で連携
GAIA-Xは、データと他のサービスに合わせてカタログベースでデータ利用者に提供

一般的なデータ取引市場の形態



GAIA-Xの中のデータ取引市場の形態



ビジネスモデル データ取引市場について

GAIA-X自体はデータリソースの情報をカタログとして管理しつつ、Catena XのようにGAIA-Xに実装したモビリティ分野専用のサードパーティーのデータ取引市場も存在

GAIA-X



メルセデスベンツ
シーメンス
ポッシュ
SAP等が参画

Catena-Xネットワークの実現に重要な要素の1つとして、データスペース技術、特にIDS規格であるDIN Spec 27070に基づいて構築した汎欧州的なGAIA-Xデータインフラがある

GAIA-Xの様に、全体的にデータのカatalogを管理した上ですべてのデータを1つのデータ取引市場に統合する必要のない、複数のデータ取引市場という形態も可能

データ提供者/利用者として連携

利用者やサービスなど、一部認証関連の要件はまだ完全に確定されていない状態である為、今後連携の際に連携の要件を踏まえて検討が必要

提供者:

- ノードとサービスのすべてのプロバイダーはプロバイダーとして登録する必要
- Self-Descriptionに必要な属性カテゴリは、ID、連絡先情報、および認証

ノード:

- ノードのSelf-Descriptionは、ノードの関連機能属性と非機能属性を記述
- 属性カテゴリには、可用性、接続性、ハードウェア、監視、物理的セキュリティ、および持続可能性の記載が必要

利用者 (optional):

- 消費者のSelf-Descriptionはオプションであるが、重要なデータセットや特定の領域にアクセスするために必要になる場合が存在
- 消費者のSelf-Descriptionに必要な属性カテゴリは議論中であり、**まだ確定されていない状態***

サービス:

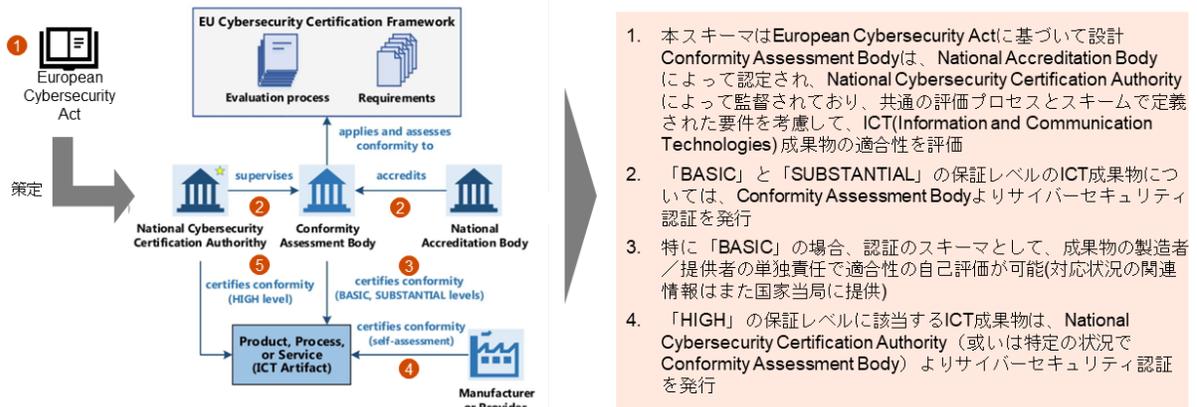
- サービスのSelf-Descriptionは、サービス関連の必要な情報を記述
- サービスのSelf-Descriptionに必要な属性カテゴリは議論中であり、**まだ確定されていない状態**

*GAIA-X: Technical Architecture, June, 2020

法規則対応

GDPRの他に、現在GAIA-Xは欧州連合サイバーセキュリティ機関（ENISA）が策定している「クラウドサービスのための欧州サイバーセキュリティ認証スキーム」の採用を検討
GAIA-Xとの連携にあたり、本認証スキームの対応要否は要確認

クラウドサービスのための欧州サイバーセキュリティ認証スキームーオーバービュー (ENISA)

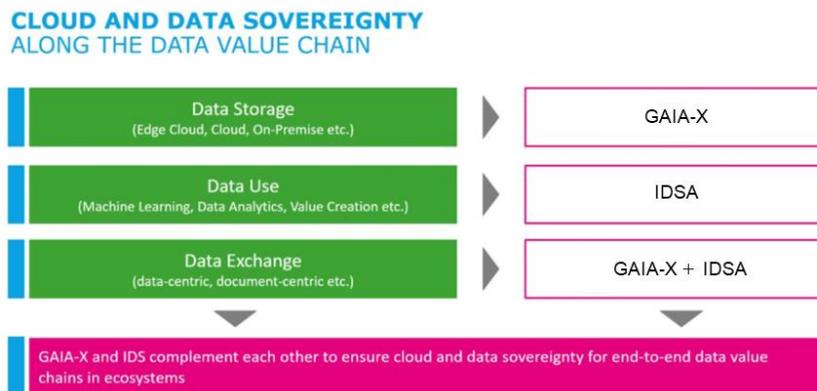


The International Data Spaces Association (IDSA)

IDSA

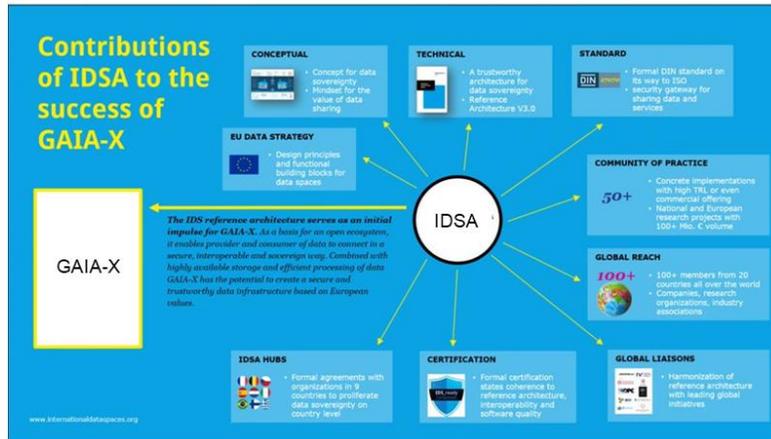
背景 IDSA&GAIA-X それぞれの役割定義

IDSAはデータにフォーカスするものの、データ連携にあたり必要なインフラストラクチャーはGAIA-Xにて構築する為、両団体に合わせてデータ連携のエコシステムを構築



背景 IDSAのGAIA-Xに対する貢献 Overview

GAIA-XのData Spaceまたは各種連携機能はIDSAの既存のスタンダードまたは技術を活用し、更にGAIA-Xのデータ連携を実現する目的の為に開発

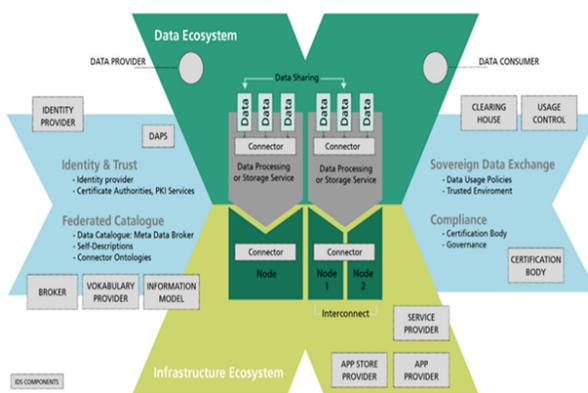


背景 IDSAからGAIA-Xへの貢献 詳細

GAIA-Xの構築にあたり、Data SpaceやFederation ServiceはIDSAの既存のコンポーネントを活用；また、GAIA-XはIDSのConnectorを利用する為、GAIA-Xと連携する際にIDS Connector連携の技術要件検討やコンプライアンス上の要件検討が必要

※GAIA-Xでの開発は検討中の為、今後変わることも可能

IDS コンポーネントとGAIA-Xのマッピング



IDSコンポーネント

コンポーネント説明

IDSコンポーネント	コンポーネント説明
Identity Provider	✓ID プロバイダは、すべての IDS 参加者に対して認証サービスを提供
DAPS	✓ダイナミックアトリビュートプロビジョニングサービスを提供
Broker	✓データプロバイダーはメタデータをIDS Brokerを介して利用可能
Vocabulary Provider	✓アノテーションのためにデータセットをアノテーションしたり記述したりするために、特定のボキャブラリが提供
Information Model	✓IDS情報モデルは、データ中心の視点で作成
Clearing House	✓コネクタはデータ・トランザクションを記録し、データ・レコードをクリアリング・ハウスに送信
Usage Control	✓データ主権の重要な要素は、使用ポリシーの実施
Certification Body	✓独立した第三者が評価し、正しく実装されていることを確認することで、エコシステム全体の信頼性を確立
Connector	✓DSコネクタは、データの交換を実現
Service Provider	✓IT企業がIDS参加者にソフトウェアやサービスを提供することも可能
App Store Provider	✓これらのデータアプリは、App Storeで販売
App Provider	✓データアプリは、連携後データをさらに処理することが可能

全体図 IDS&IDSAの活動内容

IDSの活用内容は主に3つに分けられ、その内IDSAのメインの役割は、IDSより考案されたデータ連携に関わる国際基準を確立する為のIDS-RAM普及活動

活動団体	役割	具体的な活動内容	
IDS	マーケット	IDSの基準を用いて開発したサービス・製品の評価・認証を担当 ※これらのサービス・製品は最終的にIDSのエコシステムの一部となる為	商用ソフトウェア・データマーケット・技術開発・セントラルサービスの提供（認証など） ・ロールアウトとスケールアップの活動・プロフェッショナル・サービス プロフェッショナル・サービス・ドメインに特化した（パーティカル）インプレメンテーション...
	非営利活動法人 (IDSA)	IDSの国際基準を確立するためにIDS-RAMの普及に必要な活動内容を担当	リファレンス・アーキテクチャー・モデルのメンテナンス・要件管理・標準化活動・標準化活動・セントラルサービスに関する仕様とRIG知識の移転、国際化、ドメイン固有の活動のためのプラットフォーム・ドメインスペシフィックな活動
	研究	IDSの設立者であるFraunhoferより運営。主にIDS-RAMのコア・プリンシプルの設計と継続的な開発を担当	リファレンス・アーキテクチャー・モデル（初期バージョン）・プロトタイプの実装ユースケースでのプロトタイプの実装IDSコンポーネントの基本バージョン（知識の移転（研究成果の提供と支援センター）・技術革新（利用制御、信頼できるコネクタなど） 技術革新(利用制御、トラステッドコネクタなど)・標準化活動の支援 など

全体図 IDSの成果物

IDSは、データ主権を確保した上での安全かつ信頼可能なデータ共有を促進するための、オープンスタンダードの確立を目指してコアのIDS-RAM及び関連スタンダードを定義

目的	<ul style="list-style-type: none"> IDSのグローバルスタンダードを確立し、インターフェースを構築すること 将来のデータ経済を牽引する関連技術やビジネスモデルを産業界全体で育成すること 											
Reference Architecture Model (RAM)	<ul style="list-style-type: none"> ISAは、信頼性の高い自己決定的なデータの連携を可能にするデータ駆動型のエコシステム、関連製品、サービスの構築基準となるのは、IDS Reference Architecture Model(IDS-RAM) 構築プリンシプル <ul style="list-style-type: none"> Role model Data sovereignty Information model Usage policy enforcement IDS-RAMは、世界中の組織に、データエコシステムやマーケットプレイスを低コストかつ参入障壁を低くして開発・利用できるように支援 <div style="border: 1px solid gray; padding: 5px; margin-top: 10px;"> <p>IDS Component</p> <table border="1"> <tr> <td rowspan="2">Three major technical component</td> <td>Connector</td> <td>Broker</td> <td>App Store.</td> </tr> <tr> <td>Identity Provider</td> <td>Vocabulary Hub</td> <td>Update Repository</td> </tr> <tr> <td>Additional components</td> <td></td> <td></td> <td>Trust Repository</td> </tr> </table> </div>	Three major technical component	Connector	Broker	App Store.	Identity Provider	Vocabulary Hub	Update Repository	Additional components			Trust Repository
Three major technical component	Connector		Broker	App Store.								
	Identity Provider	Vocabulary Hub	Update Repository									
Additional components			Trust Repository									
IDS Certification	<ul style="list-style-type: none"> IDSのエコシステムとなるのは信頼であり、システムに加入する組織・製品に対して厳格で透明性のある認証プロセスが必要 IDS基準を利用した組織とその基準を元に開発した製品に対して、IDSより内容の評価を行った上、IDS公式からのcertificationを発行し、該当組織・製品はデータの主権とパートナー間の安全なコラボレーションを保証できるとことを証明 											
DIN SPEC	<ul style="list-style-type: none"> DIN SPEC 27070は、相互に合意されたルールと制約に基づき、業界のデータとサービスを交換するための信頼できるゲートウェイの要件と参照アーキテクチャを規定 次のステップは、このDIN SPECを国際標準規格であるISO規格にすること 											

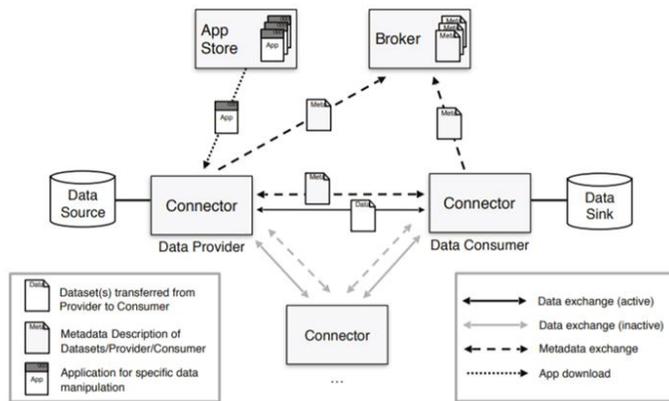
全体図 IDS Component

機能レイヤで特定された要件を踏まえ、IDSはConnector、Broker、App Storeの3つのコンポーネント、及びそれらを支える追加の4つのコンポーネントが存在
IDSは環境に合わせて常に進化する為に、自身のアップデート管理を意識しつつ、機能を拡充する為のサードパーティーのサービスを取り入れる取組みが存在

	Connector	Broker	App Store	
Three major technical component	<ul style="list-style-type: none"> IDSコネクタは、データの交換を機能し、内部のデータリソースやエンタープライズシステムとの間で参加組織の内部データ・リソースやエンタープライズ・システムとの間で完全なデータ交換プロセスを実行 	<ul style="list-style-type: none"> メタデータの管理 	<ul style="list-style-type: none"> これらは、参加者がInternational Data Spacesに参加するために必要なコア技術コンポーネントであるConnector内に展開できるアプリケーション 	
Additional components	<ul style="list-style-type: none"> IDSエコシステムの参加メンバーのアイデンティティ情報の作成、維持、管理、検証など一連のサービスを提供 	<ul style="list-style-type: none"> ポキャブラリーと関連ドキュメントの編集、閲覧、ダウンロードのためのメンテナンス機能を提供 シームレスかつ可用性を保证する為、外部サードパーティーのポキャブラリーセットも参照 	<ul style="list-style-type: none"> コネクタ内で利用した技術種別ごとに、更新する為のリソース集合 	<ul style="list-style-type: none"> 信頼性の高いソフトウェアスタックとフィンガープリント、およびリモート認証チェックのソース集合

(参考)全体図 IDS Componentの関係(主要)

3つの主要コンポーネントの関係図



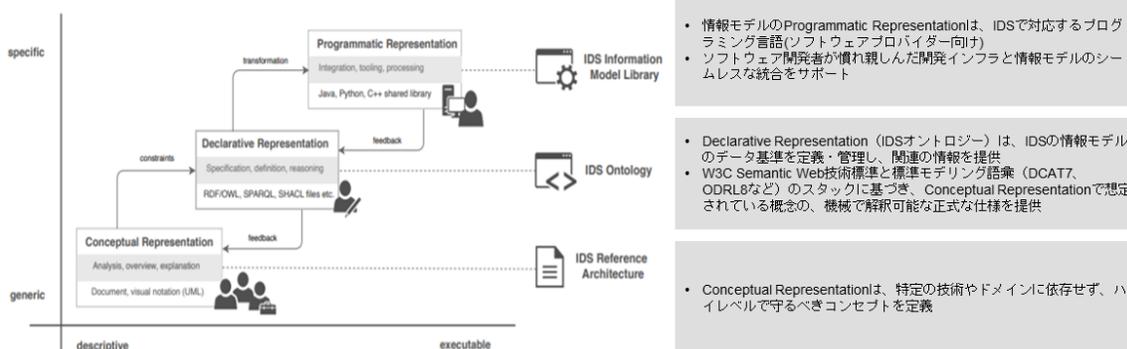
(参考) 各コンポーネントの開発状態

各コンポーネントは主にIDS-Gにて管理され、Githubで公開
IDSにおける各種コンポーネントはおおよそ開発完了しているが、一部開発継続中

Component	Version	location	Approved	Status
Certificate Authority	n/a	IDS-G	no	Work in Progress
Dynamic Attribute Provisioning Service DAPS	V1.0	IDS-G	✓ yes	stable
Participant Information Service PartS	V1.0	IDS-G-pre	no	under review
Meta Data Broker	V1.0	IDS-G	✓ yes	stable
Clearing House	V1.0	IDS Homepage	✓ yes	work in progress towards new version
App Store	n/a	IDS-G	no	Work in Progress
Connector and Broker Criteria Catalogue	V2.1.2	IDS jive, available on demand	✓ yes	Continuous improvement by IDS WG Certification
Connector REST interaction	V1.0	IDS-G	✓ yes	stable
Connector IDSCP2 interaction	V1.0	IDS-G	no	under internal review
Connector Multipart interaction	V1.0	IDS-G	no	work in progress
Connector Communication	V1.0	IDS-G	no	work in progress
Operation Environment Criteria Catalogue	V1.0	IDS jive, available on demand	✓ yes	Continuous improvement by IDS WG Certification
Informationmodel	V4.0.0	github	✓ yes	ongoing development

技術 データ・メタデータ対応基準

IDSの情報モデルは3つのレイヤーに分けられ、IDSエコシステム内の（半）自動的なデータ連携を可能にするための各種リソースのフォーマットを定義
IDSにおいて、データはISO/IEC2382:201513に準拠して、主に機械処理を目的とした構造化されたフォーマットで提供



- 情報モデルの Programmatic Representation は、IDS に対応するプログラミング言語 (ソフトウェアプロバイダー向け) ソフトウェア開発者が慣れ親しんだ開発インフラと情報モデルのシームレスな統合をサポート
- Declarative Representation (IDS オントロジー) は、IDS の情報モデルのデータ基準を定義・管理し、関連の情報を提供
- W3C Semantic Web 技術標準と標準モデリング語彙 (DCAT7, ODRL8 など) のスタックに基づき、Conceptual Representation で想定されている概念の、機械で解釈可能な正式な仕様を提供
- Conceptual Representation は、特定の技術やドメインに依存せず、ハイレベルで守るべきコンセプトを定義

技術 データ来歴、加工記録技術

データの出所に関する正しい情報（メタデータ）を提供することで、個々の参加者の責任を定義。IDSはClearing機能を提供することでデータをトラッキング

Data Source Description	Data provenance tracking
<ul style="list-style-type: none"> IDSではデータソースの関連情報を取得する為に、データソースの異なるバージョンのメタデータを記述、公開、維持、管理する権限を各参加者に付与 メタデータは、データソースのセマンティクスだけでなく、シンタックスやシリアライゼーションを記述も必要 更にメタデータは、データソースのアプリケーションドメインも記述 Connector のオペレータは、価格、価格モデル、および使用ポリシーを定義することができなければならない。 価格、価格モデル、および特定のデータに関する使用ポリシーを定義できなければならない。 	<ul style="list-style-type: none"> 透明性を高め、クリアリング機能を提供することで、International Data Spacesはデータの出所と系譜を追跡する方法を提供 データの出所の追跡は、IDSコネクタに統合されたローカルな追跡コンポーネントと、クリアリングハウスに接続された中央の出所保存コンポーネントで実装可能 データの実績追跡は、分散型データ利用制御と密接に関連していますが、補完関係ではない。データ・プロバンス・トラッキングとはデータがいつ、どのように、誰によって変更されたか、また、どのデータが新しいデータアイテムの作成プロセスに影響を与えたかを知ることができます。 データ・プロバンス・トラッキングは、さまざまな目的に合わせて編成することができます。IDSに関して、最も重要な目的は、透明性を確立し、契約、協定、または法的規制へのコンプライアンスを証明することです。コンテナの信頼性は、IDSにおけるデータ実績追跡の二次的な目標です。

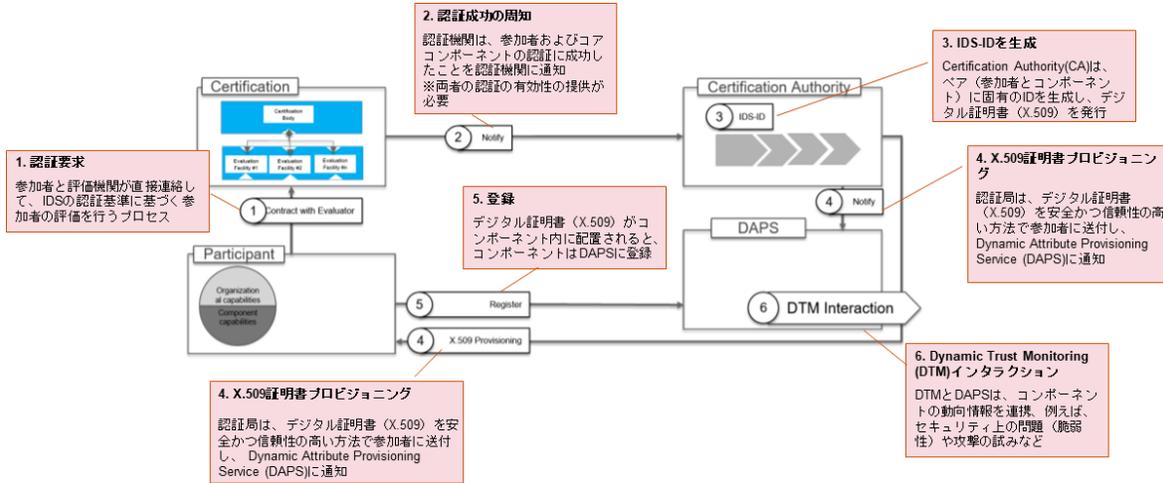
技術 ID関連技術 IDS参加者ロール定義

IDSに参加希望の組織に対する認証は、すべての参加者の間で（特にIDSの全体的な機能にとって重要なロールに関して）信頼を確立するための基本的な手段であり、IDSは各参加者に適切な認証スキームを開示

#	参加者ロール	ロール定義
1	Data Owner	データ所有者がデータ提供者として同時に行動しない場合、データ所有者の唯一の活動は、データ提供者がデータ消費者に自分のデータを利用させることを許可すること
2	Data Provider	データプロバイダは、データオーナーとデータコンシューマーの間でデータの交換を可能※データプロバイダは、ほとんどの場合、データオーナーと同一であるが、必ずしもそうではない
3	Data Consumer	データ消費者は、データ提供者からデータを受領。ビジネス・プロセス・モデリングの観点からは、データ消費者はデータ提供者のミラー・エンティティであり、データ消費者が実行するアクティビティは、データ提供者が実行するアクティビティに類似
4	Data User	データ所有者がデータを法的に管理する法人であるのと同時に、データ利用者はデータ所有者のデータを利用ポリシーに基づいて利用する法的権利を有する法人
5	APP Provider	アプリ提供者は、International Data Spacesで使用するデータアプリを開発
6	Broker Service Provider	Broker Service Providerは、International Data Spacesで利用可能なデータソースに関する情報を保存・管理する仲介者
7	Clearing House	クリアリングハウスは、すべての金融取引やデータ交換取引の清算・決済サービスを提供する仲介業者
8	Identity Provider	IDプロバイダは、国際データスペースの参加者のID情報を作成、維持、管理、監視、および検証するサービスを提供
9	APP Store Provider	App Storeでは、Data Appsを提供。これは、参加者がInternational Data Spacesに参加するために必要なコア技術コンポーネントであるConnectorの内部に配置できるアプリケーション
10	Vocabulary Provider	Vocabulary Providerは、データセットのアノテーションや記述に使用できるボキャブラリー（オントロジー、参照データモデル、メタデータ要素など）を管理・提供する。特に、Vocabulary Providerは、International Data Spacesの情報モデルを提供
11	Service Provider	サービスプロバイダには、International Data Spacesで交換されるデータの品質を向上させるための追加データサービス（データ分析、データ統合、データクレンジング、セマンティックエンリッチメントなど）を提供するプロバイダも含む
12	Software Provider	ソフトウェア提供者は、国際データスペースが要求する機能を実現するためのソフトウェアを提供

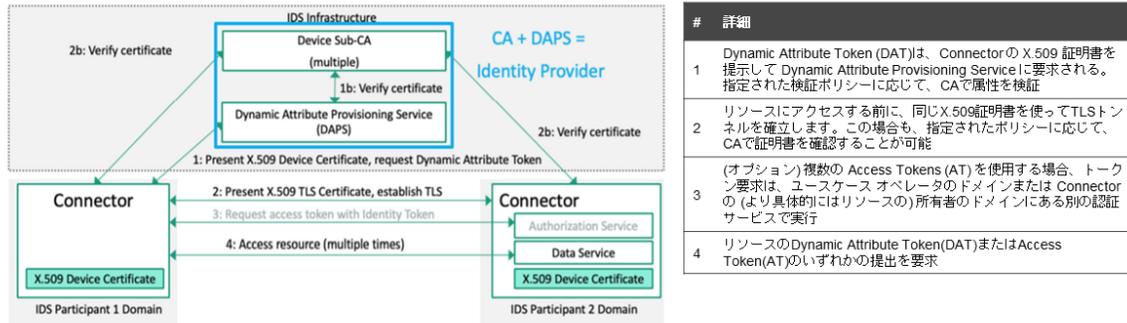
技術 ID関連技術 認証発行

認証の対象は参加者とコア・コンポーネントで、両方の認証が必要
参加者とコア・コンポーネントの両方の認証を取得した参加者には、識別、認証、暗号化のためのデジタルX.509証明書を発行



技術 認可プロセス

認証発行後、参加者とコア・コンポーネント関連の認証情報を元に、各参加者(コネクタ)の間での連携が可能で、参加者同士間の個別契約は不要



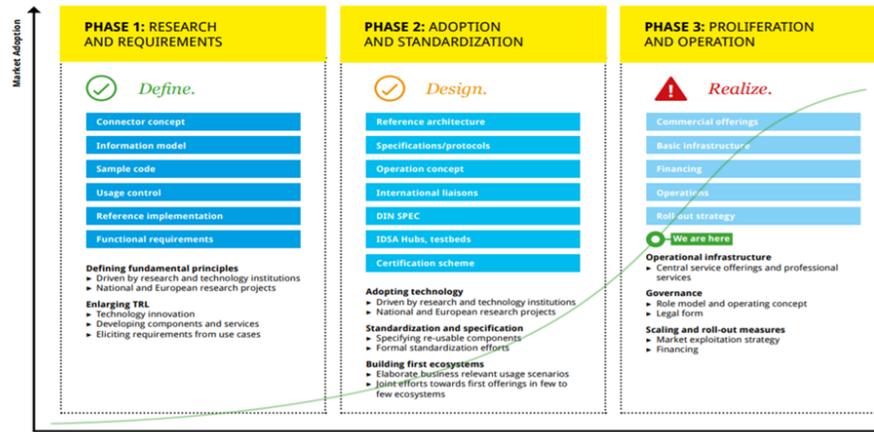
#	詳細
1	Dynamic Attribute Token (DAT)は、Connectorの X.509 証明書を提示して Dynamic Attribute Provisioning Serviceに要求される。指定された検証ポリシーに応じて、CAで属性を検証
2	リソースにアクセスする前に、同じX.509証明書を使ってTLSトンネルを確立します。この場合も、指定されたポリシーに応じて、CAで証明書を検証することが可能
3	(オプション) 複数の Access Tokens (AT) を使用する場合、トークン要求は、ユースケース オペレータのドメインまたは Connector の (より具体的にはリソースの) 所有者のドメインにある別の認証サービスで実行
4	リソースのDynamic Attribute Token(DAT)またはAccess Token(AT)のいずれかの提出を要求

技術 今までの成果と今後のロードマップ

IDSはリサーチ・要件定義から、採用・標準化のプロセスを通じて最終的にIDSのコンセプトの実現と運用を目標

現在のステータスはフェーズ3で、インフラストラクチャーの開発最中

※IDSは独自のインフラストラクチャーは未完成 (2021/9時点)



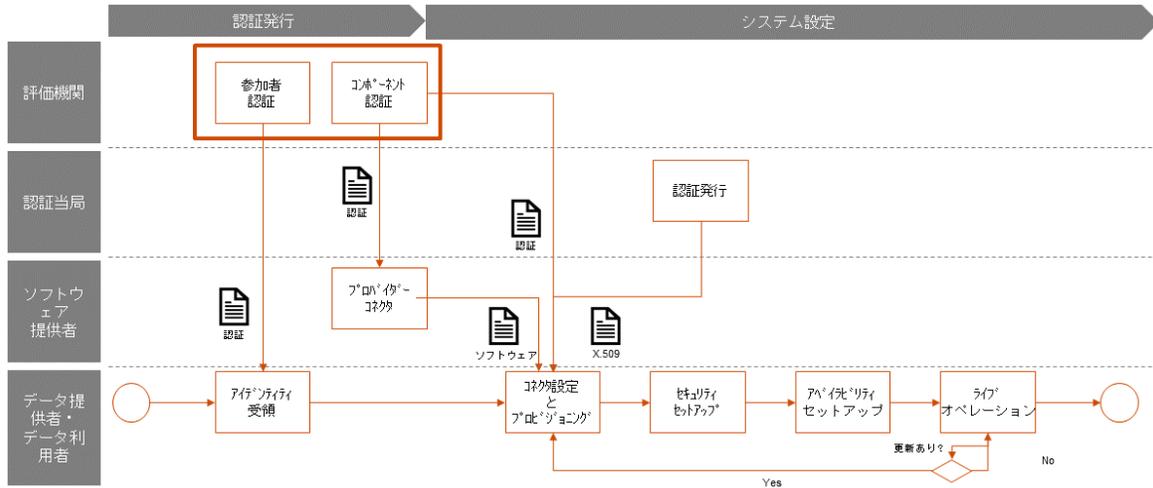
ビジネスモデル 全体図

IDSАは日常運営を維持する為IDSに参入するメンバーにメンバーフィーの納付を求めている、参入メンバーの性質(企業、スタートアップ、大学など)と年間売上によって金額が異なるプランを設定

Form of organization	Yearly turnover (Mio. EUR) (group related)	Annual fee (EUR)
Company	From 10,000	35,000
	From 2,500 to under 10,000	25,000
	From 500 to under 2,500	15,000
	From 50 to under 500	7,500
	Under 50	2,500
Start-ups	-	1,000
Universities, non-commercial institutions, associations, etc.	-	1,000

データ提供者/利用者として連携 プロセス

IDSのデータ提供者・利用者としてオンボーディングするにあたり、IDSの参加者としての認証、及びIDSの基準に遵守するデータ連携用コネクタの2つの認証が必要



データ提供者/利用者として連携 技術条件

IDSエコシステムに参加するためには、IDSに準拠したデータ交換を担当する企業のゲートウェイとして、少なくとも1つのコネクタのインストールが必要

IDSコネクタの利用方式は、自社システムに統合・コネクタ開発・既存のコネクタ活用の3つのバリエーションが存在

<p>IDSAのConnectorと自社システムのランドスケープへの統合</p>	<ul style="list-style-type: none"> IDSコネクタはモジュラーデザインを採用しており、多くの企業の既存のITアーキテクチャに対応するため、自社のITインフラを利用して、IDSに対応させることも選択肢 International Data Spacesを自社のインフラに統合するためには、3つの必須コンポーネントの利用が必要 (IDSインターフェース (httpsまたはIDSCP)、情報モデル、利用制御) 設定完了後、システムの認証が必要
<p>独自のConnectorを開発する場合</p>	<ul style="list-style-type: none"> 独自のコネクタをモジュラーなソフトウェアとして実装することも可能で、IDSフレームワークのサポートを活用してJAVAコネクタの簡単な実装が可能だが、独自開発されたシステムを認証する必要 既存の他のConnectorを利用する場合： <ul style="list-style-type: none"> 既存のコネクタを使用するのが最も簡単な方法で、最も先進的な3つのコネクタは、Dataspacesコネクタ、Trustedコネクタ、およびEnterprise Integrationコネクタで、その他のコネクタは、さまざまなアプリケーション、さまざまなハードウェア、さまざまなプログラミング言語にも対応
<p>既存のConnectorを利用する場合</p>	<ul style="list-style-type: none"> 既存のコネクタを使用するのが最も簡単な方法で、最も先進的な3つのコネクタは、Dataspacesコネクタ、Trustedコネクタ、およびEnterprise Integrationコネクタで、その他のコネクタは、さまざまなアプリケーション、さまざまなハードウェア、さまざまなプログラミング言語にも対応 この場合、コネクタ以外にもいくつかのセントラルインフラサービスが必要で、特に、コネクタのIDを定義して配布するためには、IDプロバイダが必要

法規則

個人情報の保護に関して、IDSAはGDPRの他に、連邦データ保護法（BDSG）およびテレメディア法（TMG）の規定に準拠して、どのようなデータをいつ保存し、どのように使用しているのかを管理

Federal Data Protection Act (BDSG)	Telemedia Act (TMG)
<ul style="list-style-type: none"> • ドイツのBDSG（Bundes Daten Schutz Gesetz）は、連邦データ保護法であり、ドイツ連邦州のデータ保護法やその他の地域固有の規制とともに、手動で処理されたりITシステムに保存されたりしている個人データの公開について規定 • 本規則に違反になった場合の罰金はMAX 30万ユーロ 	<ul style="list-style-type: none"> • テレメディア法（Tele Medien Gesetz, TMG）は、ISPなどの電子情報通信サービスを提供するすべての事業者に適用されるが、電気通信サービスを提供していない場合には不適用 • EUのE-Commerce指令（Directive 2000/31/EC）をベースにしたテレメディア法では、一定の要件（セーフハーバー）の下で潜在的な責任を免除される3種類のサービス・プロバイダーを挙げている。 <ol style="list-style-type: none"> 1. 単なるアクセス・プロバイダー（第8条TMG） 2. キャッシング・プロバイダー（第9条TMG） 3. ホスティング・プロバイダー（第10条TMG）

data.gov

4. data.gov 概要

data.govは米国政府のProject Open Dataの一部として、データをオープン化してサービスとして提供するプラットフォーム

基礎情報	
設立された経緯	<ul style="list-style-type: none"> data.govは2009年に立ち上げられ、米国一般調達局のテクノロジー・ライセンス・フォーメーションサービスがオーナーになり、サイトを運営・管理 データは、米国政府、各パートナー、および一般市民にとって貴重な国のリソースであり、戦略的資産です。data.govは米国政府Project Open Dataの一環として、データを資産として管理し、アクセス・発見・活用できるように、データをオープン化することで、民主主義を強化し、政府の効率と有効性を促進するだけでなく、経済的機会を創出し、市民の生活品質向上を目標
組織としての目的	
団体タイプ	公共団体
所属国/地域	US
設立日	2009年
売上	不明

機能提供状況	
Data Ingestion	提供対象
Processing	提供対象
Data Storage	提供対象
Analytics	提供対象
Data Exchange & Data Marketplace	提供対象
Data Catalog	提供対象
Privacy and Information Security	提供対象

■ : 提供対象 □ : 提供対象外

団体分類			
Software Providers	Service Provider	Data Provider	
✓	x	✓	
各機能を対応する製品/サービス例			
	Software Providers	Service Provider	Data Provider
Data Ingestion			
Processing			
Data Storage			
Analytics	Data Visualization Wizard		
Data Exchange & Data Marketplace			Datasets
Data Catalog			
Privacy and Information Security			

対象としているデータ						
ライセンスによる分類	個人データ対応	分野による分類	用途による分類	性質による分類	構造による分類	データサイズによる分類
オープンデータ 非公開データ	個人データ	公共データ 産業データ 学術データ	研究用データ 商用データ	静的データ 動的データ	構造化データ 非構造化データ	ビッグデータ ピッグデータ 以外

data.govのメタデータ管理

DCAT-US Schemaに則った、メタデータ管理が要求されている。データセット名等の一般的なメタデータの内容に加えて、米国政府内で管理されるIDの付与などが求められる

DCAT-US Schemaの概要	<ul style="list-style-type: none"> Project Open Data Metadata Schemaのガイドラインであり、公的データセットやAPIを文書化・リスト化する際に準拠すべき規格を定義している data.govにおいて使用されるだけでなく、agency.gov/dataにおいてデータをホストする際も求められる規格 2014年時点で制定されたv1.1が現在も利用されている
データ提供時に求められる文書など	<ul style="list-style-type: none"> メタデータの標準語彙 データセット・WebAPIに関する説明文書 メタデータそのものはJSON形式で保管
定義すべきJSONデータ	<p>(必須)</p> <ul style="list-style-type: none"> @conformsTo(使用しているProject Open Data Schemaのバージョン) dataset(データセットの定義) <ul style="list-style-type: none"> datasetにおいて定義される内容(必須のものを抜粋) <ul style="list-style-type: none"> title(データセット名) description(データセットの説明) keyword(データセットに関連するキーワード) modified(データセットの最終改訂日) publisher(データ発行団体(と、任意でその親団体)) contactPoint(データセットの問い合わせ先) identifier(データセット/APIのデータカタログ/DB上のID) accessLevel(データセットの公開先(公開、一部公開など)) bureauCode(連邦政府通達に則った政府内でのコード) programCode(関連する政府プログラムのコード) <p>(※任意)</p> <ul style="list-style-type: none"> @context(使用するJSON-LDコンテキスト) @id(使用するJSON-LDノードID) @type(使用するJSON-LDデータタイプ) @describedBy(使用しているJSONスキーマファイルのURL)

(補足) data.govと他国の関連性(抜粋)

data.govには、data.gov.sgやdata.gov.uk等があり、大目的は同じであるものの、データ標準等に細かな差異が存在

	data.gov	data.gov.uk	data.gov.sg
組織目的	<ul style="list-style-type: none"> 米政府Project Open dataの一環として、データを資産として管理し、アクセス・発見・活用できるように、データをオープン化することで、民主主義を強化し、政府の効率と有効性を促進するだけでなく、経済的機会を創出し、市民の生活品質向上を目標とする 	<ul style="list-style-type: none"> 英国のデータ戦略であるNDSの一環としてより簡易に、分かりやすく、早くデータを見つけるようなプラットフォームを目指す *NDS (National Data Strategy) の目的としては、国民に信頼性のあるデータを提供しつつ、世界のデータ経済をリードするようなポジションの実現を目指す 	<ul style="list-style-type: none"> 政府の公開データに対して、ワンストップのアクセスの提供 政府のデータや分析の可視化や記事による伝達 アプリケーション開発の誘発による価値の創出 分析や研究の促進
データ来歴管理	<ul style="list-style-type: none"> DCAT-US Schema v1.1 (Project Open Data Metadata Schema)を活用 	<ul style="list-style-type: none"> ISO 19139 for non-spatial data GEMINI 2.3 for spatial dataに準拠 	<ul style="list-style-type: none"> メタデータにより管理
メタデータ形式	<ul style="list-style-type: none"> DCAT 	<ul style="list-style-type: none"> 不明 	<ul style="list-style-type: none"> GovTech定義
ID認証基盤	<ul style="list-style-type: none"> API Key 各ユーザにAPI Keyを付与し、以下いずれの方法で、API Keyをdata.govに提供すれば、認証が可能 	<ul style="list-style-type: none"> OS OAuth 2 API利用 ※ SSOの認証方式を利用 data.gov.ukに関連記載はないが、CKANのツール群として提供 	<ul style="list-style-type: none"> 政府が信頼できるデータ提供者をSSOTとして認証

技術 データの連携

データの連携にはローカルドライブからアップロード、クラウド、ライブデータ接続、URLからファイルを追加する方法があり、その内ローカルドライブ以外の手段では連携が可能

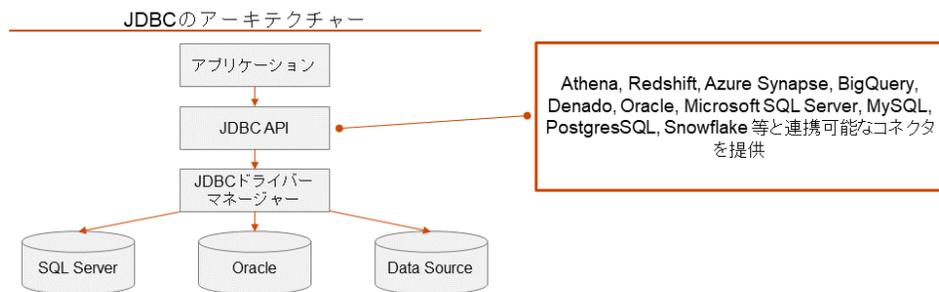
ローカルドライブ	クラウド	ライブデータ接続	URL
<ul style="list-style-type: none"> ローカルドライブからファイルを追加ボックスにアップロードすることで、データセットに複数のデータを追加可能 データとのリンクはないため、データソースとの同期は不可能 	<ul style="list-style-type: none"> data.worldによるアクセスを許可するようにクラウドサービスアカウントを構成することで、クラウド上のデータを追加可能 クラウドから供給されるファイルの定期的な同期が設定可能 	<ul style="list-style-type: none"> ライブテーブルとデータの抽出によるデータソースの追加が可能 ライブテーブルでは、データは引き続きデータソース上で存続し、data.worldに取り込まれず、クエリはデータソース上に変換され実行 データの抽出はデータソースから特定の区間で更新 	<ul style="list-style-type: none"> URLとアクセスする権限があれば、インポートしたい別のサイトのデータを追加可能 データソースが変更された場合、data.worldで自動で更新が可能

data.world

data.world

技術 認証・認可関連

data.worldはJDBC互換のデータベース・データウェアハウスの場合APIを使用して直接同期が可能



技術 データ加工関連

データの信頼性はデータ提供者に任せられているものの、連携したデータに対しての検証サービスを提供し、データ提供者の管理を支援

データの空白、重複、外れ値等を検出しアラートする機能を提供

413 file warnings
census_county.xlsx

To fix these issues, download and edit your file, then upload it again with the same file name. [Download file](#)

census_county

Structural (66)

66 blank cells detected

- Cell at row 310, column colonies appears blank.
- Cell at row 366, column colonies appears blank.
- Cell at row 381, column colonies appears blank.
- Cell at row 422, column colonies appears blank.
- + 62 similar issues

Dismiss

Numeric (291)

291 numeric values outside standard deviation detected

Dismiss

[Upload new file](#) [Done](#)

siyeh/usda-bee-colony-inventories-and-loss

review Contributors Discussion Activity Settings

+ Add to project [Launch workspace](#)

bookmark

contributors (1)

Brenda Griffith @siyeh

RECENT COMMENTS

Have something to say?

[Leave a comment](#)

census all.xlsx
99+ Add a description

Explore [Download](#) [Share](#)

census_year	colonies	ops_with_prod	ops_with_inv	ops_with_sales
1997	No data	No data	No data	1137
2002	2354005	12029	17357	937
2007	2902732	16905	27900	1046
2012	3202578	22027	30201	17300

Showing 1-4 of 4 rows, 5 columns [See all](#) [Switch to column overview](#)

figure 1.png
Add a description

[Download](#) [Share](#)

技術 メタデータ関連

data.worldではdata.world catalog collector (DWCC)というプログラムによりデータソースにより特定のメタデータを収集
収集するメタデータの項目はデータソースにより異なる

収集するデータソースの一例

JDBC データソース

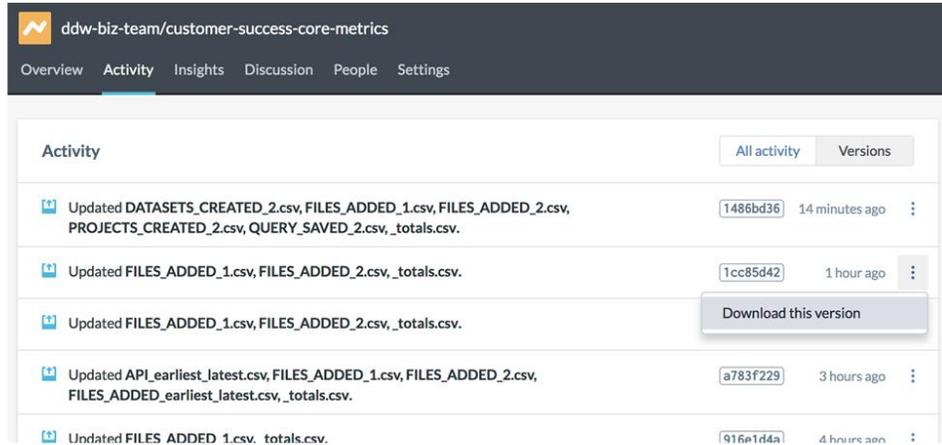
- ✓ データベース名
- ✓ 接続情報
- ✓ スキーマ名
- ✓ スキーマ別のテーブル名とビュー名
- ✓ 列名
- ✓ 列のデータ型
- ✓ 列の長さ
- ✓ 列の精度 (任意)
- ✓ テーブルと列の説明 (任意)

Tableauサーバー

- ✓ ブック名
- ✓ ダッシュボード名
- ✓ ダッシュボードのタイトル
- ✓ ダッシュボードが入っているプロジェクト
- ✓ ダッシュボード以外のビュー
- ✓ ダッシュボードビューの数
- ✓ ダッシュボードを持っているオブジェクトのタグ
- ✓ ビュー/ダッシュボードとワークブックの関係
- ✓ ダッシュボードのお気に入りの数

データ 来歴技術

すべてのdata.worldデータセットとプロジェクトはログに記録され、バージョン管理されているため、アクティビティの確認や以前のリビジョンに戻ることが可能



技術 コンプライアンス・安全性

データ提供者のデータの接続方法には3つの方法があり、どの手法で接続するかにより提供者はデータの安全性を管理可能

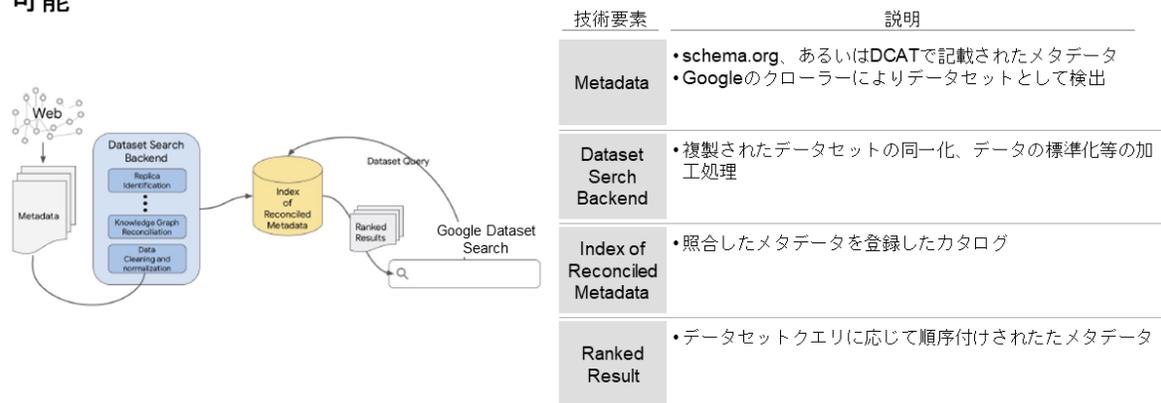
データの接続方法		データの権限対象	
メタデータの収集のみ	<ul style="list-style-type: none"> data.world側ではメタデータの収集のみ データをプラットフォーム上に保存しないため、安全性は高い 	非公開	<ul style="list-style-type: none"> 自分のみが閲覧可能
仮想クエリ機能	<ul style="list-style-type: none"> データとの接続により、ユーザーがクエリを実行すると、クエリ結果のみがフェッチ クエリ結果は5分間でタイムアウト データをプラットフォーム上に保存しないため、安全性は高い 	組織内の全員	<ul style="list-style-type: none"> 自分の所属する組織内に権限を付与 コミュニティには非公開
データの抽出・インポート	<ul style="list-style-type: none"> 仮想的な接続により、データをdata.worldに抽出・インポートが可能 小中規模のデータセットに対しては高速なクエリが実行可能 data.worldのデータは暗号化されたデータストアに保管 	コミュニティに公開	<ul style="list-style-type: none"> コミュニティにアカウントを持つ全員へ公開
		権限の種類と実行可能例	
		表示のみ	<ul style="list-style-type: none"> データのダウンロード クエリ実行およびデータのエクスポート ディスカッションでのコメント
		表示と編集	<ul style="list-style-type: none"> メタデータの編集 データの追加・削除 クエリの公開
		表示・編集および管理	<ul style="list-style-type: none"> データセットの追加・削除 管理者の追加・削除

Google

Google

Google Dataset Search : 技術 アーキテクチャー

Google Dataset SearchはGoogle検索のサブ事業に位置付けられており、Web上にあるデータからメタデータを検出、カタログ化することによりデータセット検索サービスを提供しているため、データ提供者はメタデータをWeb上に公開するのみで提供可能



技術 データフォーマット・メタデータ対応基準

Google dataset searchは一般的な構造化データ基準のschema.orgに加え、政府などで用いられているDCATの語彙を独自技術によりマッピングすることにより両者に対応し、幅広いカバレッジを実現

データフォーマット		データ・メタデータ基準	
JSON-LD (Google推奨)	<ul style="list-style-type: none"> ページの見出しまたは本文の <script> タグ内に埋め込まれる JavaScript 表記 ネストされたデータアイテムの表現が容易 	schema.org	<ul style="list-style-type: none"> Google, Yahoo!, Microsoftによって立ち上げられたマークアップ標準 schema.orgのdataset語彙はDCATを基に、Dublin CoreやFOAFの用語を使用
RDFa	<ul style="list-style-type: none"> 検索エンジンに伝えたいユーザー表示コンテンツに対応するHTMLタグ属性を追加することによってリンクデータをサポートするHTML5の拡張機能 HTMLタグ属性を使用して、構造化データとして公開するプロパティに名前を付与して使用 	DCAT*	<ul style="list-style-type: none"> DCATは、Web上で公開されているデータカタログ間の相互運用性を促進するように設計されたRDF語彙 DCATを使用すると、発行者は、複数のカタログからのメタデータの消費と集約が容易 標準モデルと語彙を使用して、カタログ内のデータセットとデータサービスを記述することが可能 データセットとデータサービスの発見可能性を向上
microdata	<ul style="list-style-type: none"> HTMLコンテンツ内に構造化データをネストするために使用される、オープンコミュニティのHTML仕様 HTMLタグ属性を使用して、構造化データとして公開するプロパティに名前を付与して使用 		

技術 データ来歴、加工記録技術

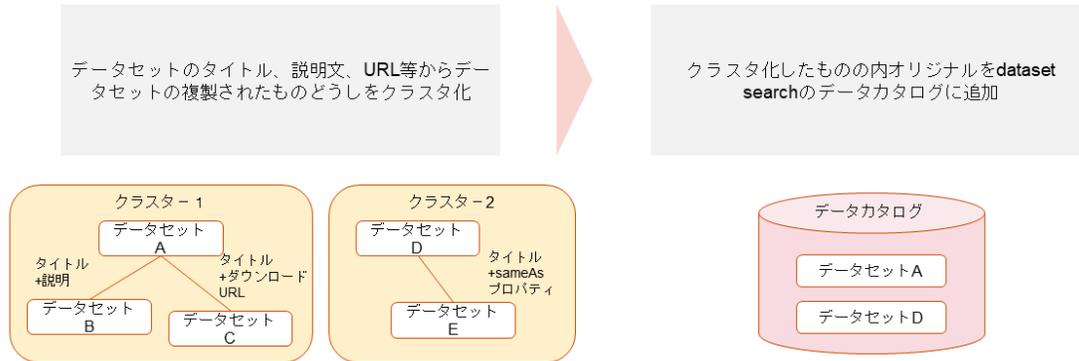
Google検索と同様にschema.orgのプロパティによりデータ提供者が任意で記述したものをGoogleは取得してるのみであるため、Googleによる来歴管理は行っていない

プロパティ例	プロパティの説明	
datePublished	データセットが公開された日付	
dateModified	データセットが公開された日付	
publisher	データセットの公開元	
creator	データセットの作成者	
temporalCoverage	データセットの対象期間	

(参考) 技術 認証関連技術

Google Dataset Searchでは団体ごとの認証を行っていないが、Web上にある対応データセット全てをクローリングにより取得することで同一のデータを取得してしまう問題があるため、複製されたデータセットを識別し同一化を行う独自技術を採用

データセットの同一化プロセス



Google Cloud Platform : 提供プロダクト概要

Googleデベロッパーツール、金融サービス等のクラウドソリューションを提供する場としてGoogle Cloud Marketplaceを提供しており、データセットも商品として提供

提供データセットの例

データ連携の仕方	カテゴリ	データセット	説明
1 直接連携	Research データセット	データセット検索結果となるデータ	ウェブ全体から約2,500万のデータセットをインデックス登録し、1か所でデータセットを検索して、そのデータへのリンクが掲載
	Kaggle データセット	Kaggle データセット	データサイエンス作業を行うために必要なコードとデータ(80,000以上の一般公開データセット)
	Earth Engine データセット	Earth Engine	40年以上の履歴画像と科学データセット(毎日更新)
2 Google Cloud Platform (GCP)を介して連携	Googleデータセット	Google トレンド	過去30日間のGoogleトレンドでの上位25件の検索トレンドと上位25件の急上昇ワード
		コミュニティモビリティレポート	さまざまなカテゴリの場所について、地理的に長期にわたる移動傾向を報告
		Google アナリティクス (サンプル)	Google ブランドの商品を販売する実在のeコマースサイトで表示される一般的なデータであり、トラフィックソースデータ、コンテンツデータ、トランザクションデータも包含
		Google Patents Research	Google 翻訳によるタイトルや抄録の機械翻訳、埋め込みベクター、抽出されたトピック、類似文書、前方参照など、Google 特許検索(patents.google.com)で使用されているデータ
	公開データセット	Severe Storm Event Details	1950年から今年に至るまでの米国における気象イベントの大規模な統合型データベース
		Census Bureau US Boundaries	米国国勢調査局が提供する地理情報プロダクトであるTIGER/Line Shapefileから派生した完全解像度の境界ファイル
		American Community Survey	全米の350万以上の世帯に対して毎年行われている、米国とそこに住む人々についての重要な情報を提供するアンケート
		すべての一般公開データセット	Google Cloud Marketplaceにある200以上のデータセット
	商用データセット	Crux Informatics	Crux Deliverが提供する検証され、クリーンで実用的なデータ
		HouseCanary	1億を超える住宅の信頼できる不動産情報、ローン情報、評価情報
Synthetic データセット	Cymbal Investments	架空のCymbal Investments グループによって管理されている自動取引ボットによる取引を表すデータ	

技術 データ・メタデータ対応基準

Google Cloud MarketplaceはデータセットのみでなくGCPで利用可能な様々なプロダクトを提供しているため、メタデータ定義は独自のものと想定

プロダクト例	プロダクト説明	メタデータイメージ
データセット	<ul style="list-style-type: none"> GCPで分析するためのデータセットを提供 データセットのソースをメタデータに記載 来歴(はないものの、予定される更新頻度をメタデータに記載可能) 	<p>概要</p> <p>In an effort to help combat COVID-19, we created a COVID-19 Public Datasets program to make data more accessible to researchers, data scientists and analysts. The program will host a repository of public datasets that relate to the COVID-19 crisis and make them free to access and analyze. These include datasets from the New York Times, European Centre for Disease Prevention and Control, Google, Global Health Data from the World Bank, and OpenStreetMap.</p> <p>詳細</p> <p>タイプ: Datasets カテゴリ: Datasets for COVID-19 research, Science & research データセットのソース: Various クラウドサービス: BigQuery 予定されている更新頻度: Daily</p>
API	<ul style="list-style-type: none"> GCP上で実行可能なAPIを提供 サービス名をメタデータに記載 来歴として最終更新日がメタデータに記載 	<p>概要</p> <p>Integrates text translation into your website or application.</p> <p>詳細</p> <p>タイプ: Cloud & APIs 最終更新日: 2021/07/23 カテゴリ: Machine learning Google Enterprise APIs サービス名: translate.googleapis.com</p>
仮想マシン	<ul style="list-style-type: none"> GCP上で実行可能な仮想マシンを提供 来歴として最終更新日・バージョンがメタデータに記載 	<p>概要</p> <p>Canonical is the commercial sponsor of Ubuntu, the leading open-source platform for cloud, personal computing and next-generation devices. Ubuntu delivers reliability, performance and interoperability to cloud and scale-out environments. Together with our partners, we ensure that Ubuntu runs reliably on every platform from the PC and the smartphone to the server and, crucially, the cloud.</p> <p>詳細</p> <p>実行対象: Google Compute Engine タイプ: Virtual machines, ARM VM 最終更新日: 2021/10/07 カテゴリ: Operating systems バージョン: 20.04 (Focal)</p>

技術 データ来歴、加工記録技術

メタデータ上では予定されている更新頻度程度の記録のみであり、来歴管理はデータ提供者に任されている。Google Cloud パートナー認証の基準が厳しいため、データの信頼性はそこで担保していると想定

The image shows a comparison between the metadata page for 'Broad References' in Google Cloud Marketplace and its corresponding data in Google Cloud Storage. The marketplace page lists the dataset with a 'Broad References' title and a description. The Cloud Storage page shows the actual data files with columns for name, size, and creation/modification dates. Annotations highlight that the marketplace metadata only shows the update frequency, while the storage page shows the actual creation and update dates of the data files.

• カタログへの記載は「予定されている更新頻度」のみで最終更新日などの記載はなし

• データセットのソースは任意で記述可能

• Google Cloud Platformのストレージには作成日時、更新日時、バージョン履歴等を記録可能

技術 認証・ID関連技術

Google Cloud にはパートナー制度があり、Marketplaceにプロダクトを提供するためには一定の要件を満たす必要があり、データの品質は原則提供者が担保。公開データセットに関しては別途審査を実施する想定

Google Cloud Marketplaceにプロダクトを提供するための要件

組織の要件	<ul style="list-style-type: none"> Google Cloud のサポートするリージョン内に存在 年間収益が100万ドル以上 Google Cloud Build パートナーとして認証
プロダクトの要件	<ul style="list-style-type: none"> プロダクトが製品版（アルファ版またはベータ版でない） プロダクト自体にMarketplaceで販売されているサービスを含まない
運用上の要件	<ul style="list-style-type: none"> Google Cloud Marketplaceのプロダクトを作成および維持、顧客へのタイムリーなサポートを提供するためにリソースを投入する意思

Google Cloud パートナーのランクとエンゲージメントモデル

	Sell	Service	Build
メンバー	<ul style="list-style-type: none"> 1社までの企業が無条件で認定 	<ul style="list-style-type: none"> 1社までの企業が無条件で認定 	<ul style="list-style-type: none"> 1社までの企業が無条件で認定
パートナー	<ul style="list-style-type: none"> 年間ビジネスプラン 2つの技術クルデンシヤル 2つの販売クルデンシヤル 	<ul style="list-style-type: none"> 年間ビジネスプラン 2億のプロフェッショナルレベル認定資格 	<ul style="list-style-type: none"> 年間ビジネスプラン 1億のプロダクトインテグレーション
プレミア	<ul style="list-style-type: none"> 年間ビジネスプラン 合計 20 個のクルデンシヤル 最低対象請求額 3つのお客様成功事例 	<ul style="list-style-type: none"> 年間ビジネスプラン 12億のプロフェッショナルレベル認定資格 最低対象請求額 3つのお客様成功事例 	<ul style="list-style-type: none"> 年間ビジネスプラン 2億以上のプロダクトインテグレーション 最低対象請求額 3つのお客様成功事例

ビジネスモデル:Googleのロックイン戦略に対する動向

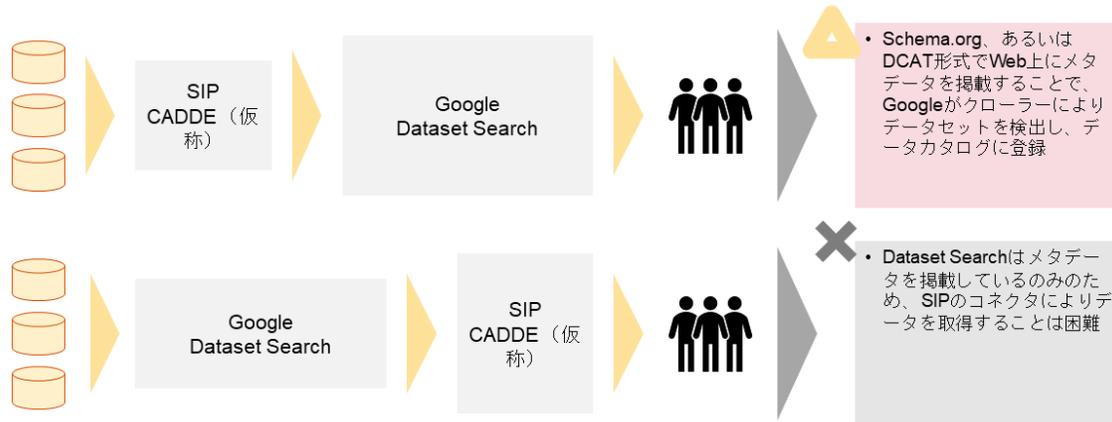
独占禁止への圧力を受け、Googleは Cloud Marketplaceの収益配分比率を20%から3%に大幅引き下げを行う意向を示しており、今後手数料による収益化はますます困難と認識

直近1年間のGAFAM困り込み戦略に対する動向



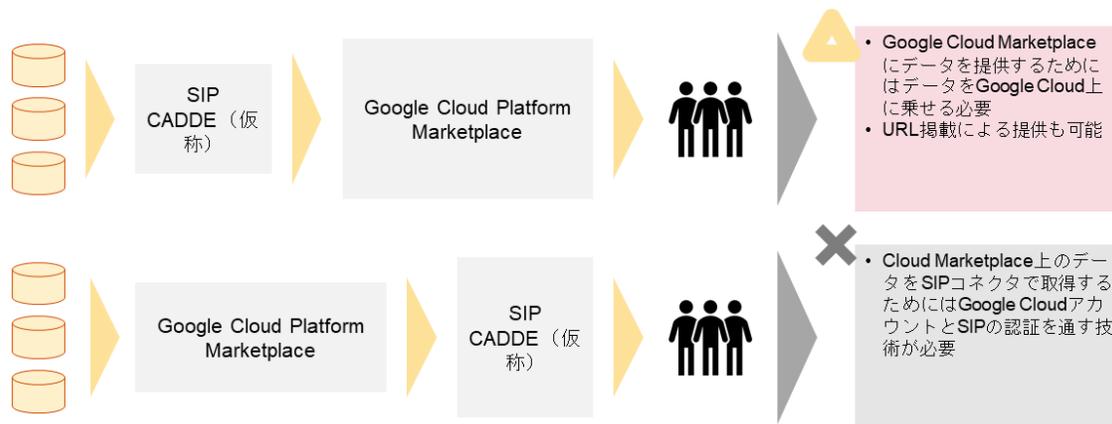
Google Dataset Searchとのデータ連携における実現可能性

SIPのメタデータを指定の形式でWeb上に記載することでDataset Search上から検索は可能であるが、Dataset Search自体が他のデータ基盤と連携しているとはいえないため、Dataset Search上のデータをコネクタで取得することは現状困難



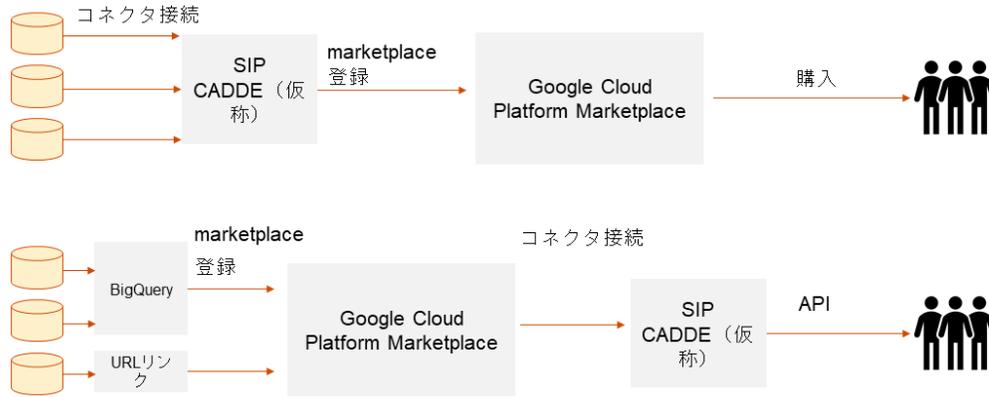
Google Cloudとデータ連携するにあたり、必要な技術要素

Marketplaceにデータを提供するためにはデータをGoogle Cloud上に乗せる必要があり、Marketplace上のデータをSIPコネクタで取得するためにはGoogle CloudアカウントとSIPの認証を通す必要があるとみられる



GCP想定アーキテクチャ

Google Cloud Marketplaceとの連携については、下記のようなアーキテクチャを想定しており、実現可能性についてヒアリングにて確認予定



Googleとデータ連携するにあたり、必要な法規則整理

Googleは日米デジタル貿易協定については課題や制約は少ないとみているが、拠点がアメリカにあるため、CLOUD Actについては日本にあるデータを含め懸念

日米デジタル貿易協定

- 2020年1月1日、発効した日米間の協定
- 日米貿易協定とともに発効された協定であり、「デジタル貿易に関する日本国とアメリカ合衆国との間の協定」、の略称
- 主に下記の点を制定
 - ✓ 関税賦課の禁止
 - ✓ 内国民待遇
 - ✓ 組織データ取引自由化
 - ✓ データのローカリゼーション制限
 - ✓ 消費者保護
 - ✓ 個人情報保護
 - ✓ ソースコード等の開示要求禁止
 - ✓ 暗号の開示要求禁止

CLOUD Act

- 2018年3月23日、米国トランプ大統領の署名を経て成立した新法
- 1986年連邦通信保存法 (SCA)を修正したもので「Clarifying Lawful Overseas Use of Data Act」の略称
- 主に下記の2点を制定
 - ✓ 米国政府は、米国所在の各種サービスプロバイダーに対し、米国内で保存されているデータのみならず、米国外で保存されているデータについても、所定の法的手続きを踏むことで提出要請が可能
 - ✓ 米政府に、「行政協定」を外国政府と締結する権限を付与。この「行政協定」があれば、外国政府が米国所在の各種サービスプロバイダーに対し、直接のデータの提出要請が可能

Amazon Web Service

Amazon Web Service

Amazon 概要

Amazonはデータ流通サービスとしてRegistry of Open Data on AWS(RODA)とAWS Data Exchangeを展開

Registry of Open Data on AWS (RODA)

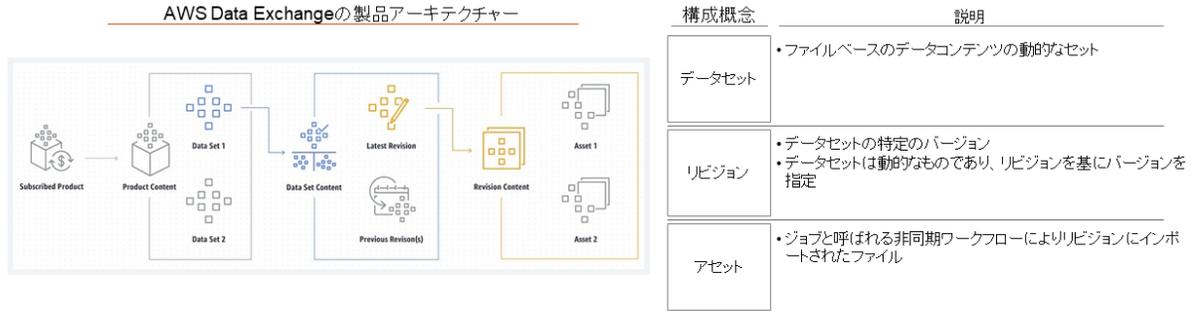
- 2018年に研究、分析のためのデータセット(Public Data Sets on AWS)をアップデートする形で発足
- GitHubのレポジトリに登録することで誰でもオープンデータをアップロード可能
- 現在約300のオープンデータセットを提供

AWS Data Exchange

- 2019年、AWS Marketplaceに追加される形で発足
- データ提供の為にAWS Marketplaceセラーとしての認証が必要
- 現在約3700のデータ製品を提供

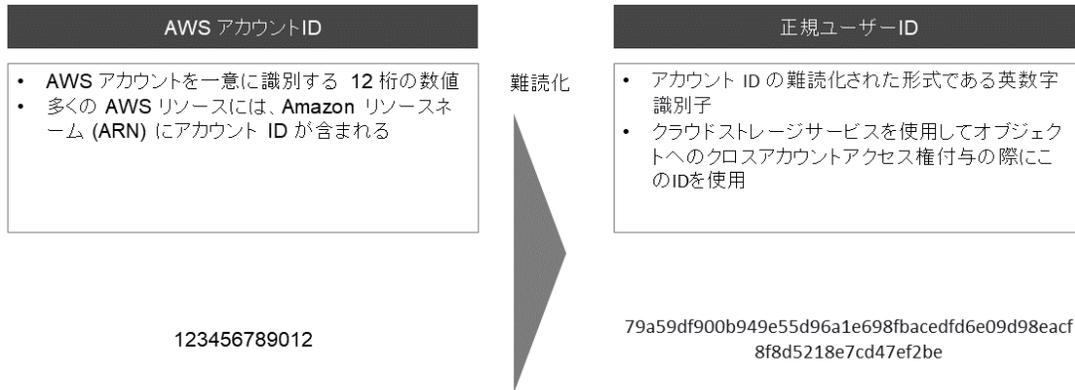
AWS Data Exchange : 技術 製品の概要

AWS Data Exchangeの製品にはデータセット、リビジョン、アセットという3つのコンポーネントが存在し、それぞれが階層構造を構成することで、動的なデータセットに対してバージョン管理が可能



AWS データ提供者・利用者ID技術

AWSではアカウントIDを12桁の数値として付与し、AWSリソースのリソースネームに使用。アクセス権の付与などにはアカウントIDを難読化させた正規ユーザーIDを使用することで、可読性と安全性を担保



(参考) 技術 ID関連技術

AWSリソースを一意に識別するためにARN (Amazon Resources Name) を採用
 これにより権限を付与された利用者に対してAWSリソースを提供者ID、リソースID
 から指定しアクセス可能

ARN形式

ARNの一般的な形式を次に示します。具体的な形式は、リソースによって異なります。ARNを使用するには、[イタリック体のテキスト](#)を、リソース固有の情報に置き換えます。一部のリソースのARNでは、リージョン、アカウントID、または、リージョンとアカウントIDの両方が省略されていることに注意してください。

```
arn:partition:service:region:account-id:resource-id
arn:partition:service:region:account-id:resource-type:resource-id
arn:partition:service:region:account-id:resource-type:resource-id
```

partition

リソースが属しているパーティション。パーティションはAWSリージョンのグループです。各AWSアカウントのスコープは1つのパーティションです。サポートされているパーティションは以下のとおりです。

- aws - AWS リージョン
- aws-cn - 中国リージョン
- aws-us-gov - AWS GovCloud (US) リージョン

service

AWS製品を識別するサービス名前空間。例えば、Amazon S3のs3です。サービスの名前空間を検索するには、Service Authorization Reference を開き、サービスのページを開いて、最初の文字で「service prefix」(サービスプレフィックス) という語句を見つけます。例えば、Amazon S3のページの最初の文字は、次のテキストが表示されます。

```
{
  "service prefix": "s3"
}
```

region

リージョンコード。たとえば、米国東部 (オハイオ) の場合は、us-east-2を使用します。リージョンコードのリストについては、リージョンエンドポイントを参照してください。

account-id

リソースを所有しておりハイパフがないAWSアカウントのID。たとえば、123456789012と指定します。

resource-id

リソース識別子。ARNのこの部分は、リソースの名前かID、またはリソースIDです。たとえば、IAM ユーザーの user/Bob、または EC2 インスタンスの instance/i-1234567890abcde9fです。一部のリソース識別子には、親リソース (sub-resource-type)/parent-resource/sub-resource または/リージョン (resource-type/resource-name/qualifier) などの修飾子が含まれます。

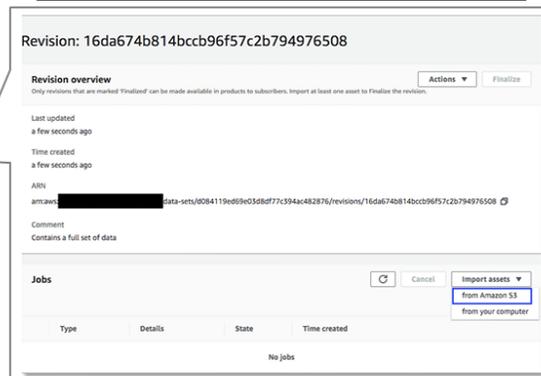
技術 データ来歴、加工記録技術

Data Exchangeでは3つのコンポーネントにより構成されており、その内リビジョン
 というコンポートメントによりバージョンを管理
 AWS特有の管理方法とみられる

データフォーマット

構成概念	説明
データセット	• ファイルベースのデータコンテンツの動的なセット
リビジョン	• データセットの特定のバージョン • データセットは動的なものであり、リビジョンを基にバージョンを指定
アセット	• ジョブと呼ばれる非同期ワークフローによりリビジョンにインポートされたファイル

リビジョンのイメージ



AWS Data Exchange ビジネスモデル

AWS Data Exchangeは各実行アクションに対して料金が設定されており、主にストレージ料金、手数料、およびデータ転送料金により収益化

料金	支払者	支払先	説明
サブスクリプション料金	データ利用者	データ提供者	購読の際には、適用される売上税とともに購読に関連する料金を前払い
ストレージ料金	データ利用者/データ提供者	AWS	AWS Data Exchange は、利用者がサービスにロードしたデータを保存するために請求
履行手数料	データ提供者	AWS	データ製品へのすべての新規サブスクリプションに対してAWS が収集した収益を、データ提供者の階層化された履行手数料 (TFF) として請求
データ転送料金	データ利用者/データ提供者	AWS	地域間でアセットをインポートまたはエクスポートする際には、標準的な S3 レートが適用

AWS 政府関与度

AWSはGAIA-Xを含む様々なEUのプライバシー、移植、およびデジタル主権プロジェクトに参画

AWS主体のデータ流通に関する政府連携プロジェクトについてはヒアリングにて確認予定

What's GAIA-X?

GAIA-X is a data economy-driven initiative that aims to bring together representatives from business, science, and politics to help define standards for the next generation of data infrastructure. This includes an open, transparent, and secure digital ecosystem, where data and services can be made available, collated, and shared in an environment of trust. AWS has participated in multiple GAIA-X technical working groups and supported the initiative from its beginning.

[Learn more about GAIA-X >](#)

Learn more about GDPR

The EU's General Data Protection Regulation (GDPR) protects EU individuals' fundamental right to privacy and the protection of personal data. It includes robust requirements that raise and harmonise standards for data protection, security, and compliance. Our customers can use all our services to process personal data (as defined in the GDPR) that is uploaded to the AWS services under their AWS accounts or generated within the AWS account (customer data) in compliance with the GDPR. We offer services and resources to help you comply with GDPR requirements that may apply to your activities, and we have over 500 features and services focused specifically on security and compliance with new features launching regularly.

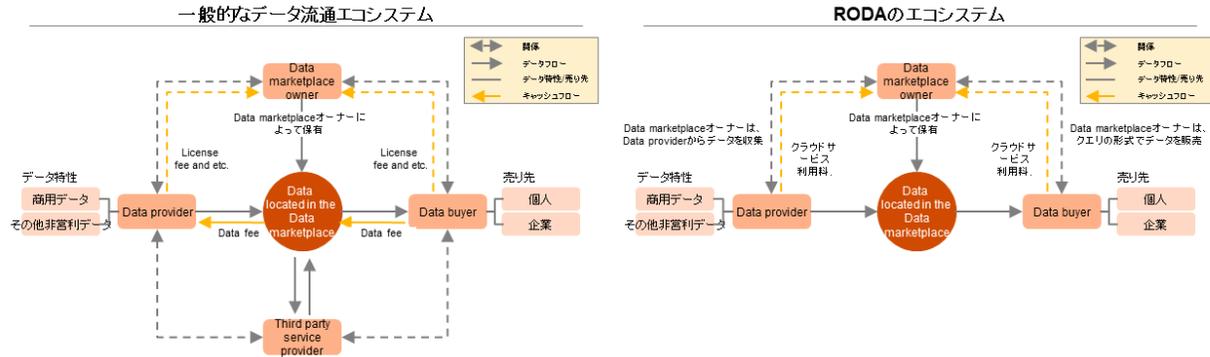
[Visit the GDPR Centre >](#)

Data Sovereignty

You can choose to store and process your customer data in any one or more of our [European Regions](#). Today, these include France, Germany, Ireland, Italy, Sweden, and beginning in 2022, Spain. You can also use AWS services with the confidence that customer data stays in the AWS Region you select. A small number of these services involve the [transfer of data to help](#), for example, develop and improve those services. You can opt-out of this transfer, unless the transfer is an essential part of the service (e.g. it's a content delivery service). Our systems prohibit – and are designed to prevent – remote access by AWS personnel to customer data for any purpose, including service maintenance, unless that access is either requested by you or if it is required to prevent fraud and abuse, or to comply with law. We are committed to important EU privacy, portability, and digital sovereignty programmes – including [Cloud Infrastructure Services Providers in Europe \(CISPE\)](#) Code of Conduct, the [European Commission Standard Contractual Clauses \(SCC\)](#), the [SWIPO Code of Conduct](#), and [GAIA-X](#).

Registry of Open Data on AWS ビジネスモデル

RODAはそれ自体の収益化を行っていないものの、RODAのデータ利用によるAWS使用手数料により収益化



Microsoft

背景 Microsoft データ連携基盤関連取組み

Microsoftにおいて、データ連携基盤関連の取組みとして、Azureの利用有無及び、取り扱うデータの種別(オープンデータ/商用データ)によって主に4つ事業が存在
グローバル未対応の為、AIDC DataCloudへの調査優先度を下げて、現在進んでいる
他の事業内容を中心に調査を推進

事業名	データ性質	Azureアカウント	対応地域	詳細
1 Microsoft Research Open Data https://msropendata.com/	オープンデータ	必須ではない*	グローバル	• Microsoft Research Open Dataは、自然言語処理、コンピュータビジョン、ドメインスペシフィックサイエンスなどの分野で最先端の研究を進めるために、Microsoft Researchが提供する無料のデータセットのレポジトリ
2 Microsoft U.S. Broadband Usage Percentages Data https://github.com/microsoft/USBroadbandUsagePercentages	オープンデータ	必須ではない*	グローバル	• マイクロソフトのエアバンド・イニシアチブとの取組みの一環として、地方でのブロードバンド・ギャップを解消するために開発したデータセットを公開
3 Azure Open Datasets https://azure.microsoft.com/open-datasets/	オープンデータ	必須	グローバル	• 機械学習のワークフローですぐに利用可能で、Azureサービスから簡単にアクセスできる厳選した一般公開のデータセットのレポジトリ
4 Azure Data Marketplace(PoC) https://github.com/Azure/commercial-marketplace-data-offers	商用データ	必須	グローバル	• Azure Marketplace内でローデータの販売・購入を実現し、データ取引市場を構築する為の事業開発活動
5 AIDC DataCloud https://www.aidcdatacloud.com/ (デモサイト)	商用データ	必須ではない*	日本国内	• AIDC Data Cloudは円滑なデータ流通を実現するための契約モデルに基づく、多様なデータ流通を可能とするクラウド基盤 • 一般的なECサイトと同様に利用開始時にアカウントを作成し、後は企業が必要なデータソースやクレンジング済みデータ、データラベルなどをカートに入れて購入する仕組みを提供

*事業はAzureで構築しているもの、ユーザはAzureを利用せず、事業のサービスの一部、又は全部利用可能の為、Azureの利用は必須ではない

技術 データ来歴、加工記録技術

データソースの関連情報はメタデータに記載し、Microsoft側の関係者より事実確認の想定データ登録後のバージョンアップは可能だが、変更詳細の記録について現時点不明

データ来歴

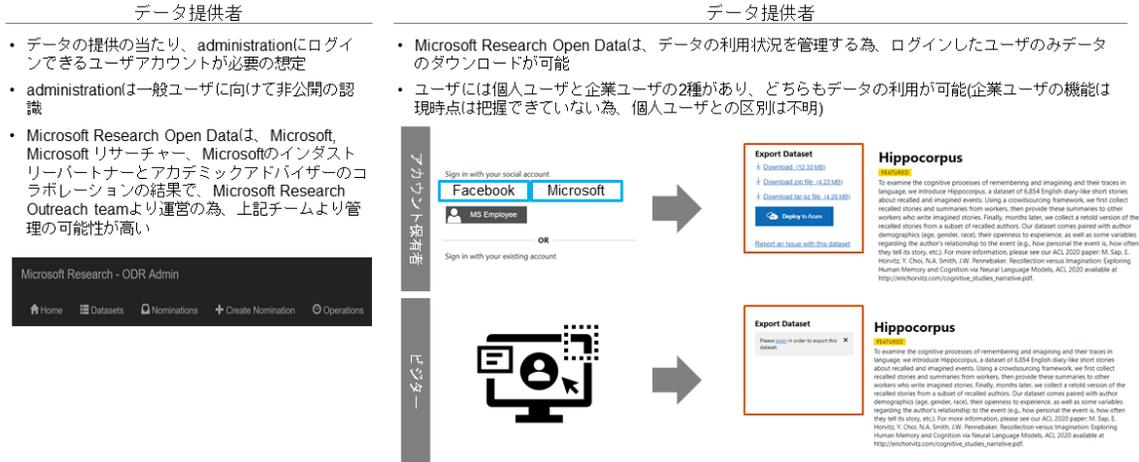
- データ指名に記載されたURL to Project Page、URL to Datasetなどの項目にて必要な情報を記載

メタデータ

- データ登録後のアップデートは可能だが、アップデートに関する詳細はユーザ向けに公開されていない認識

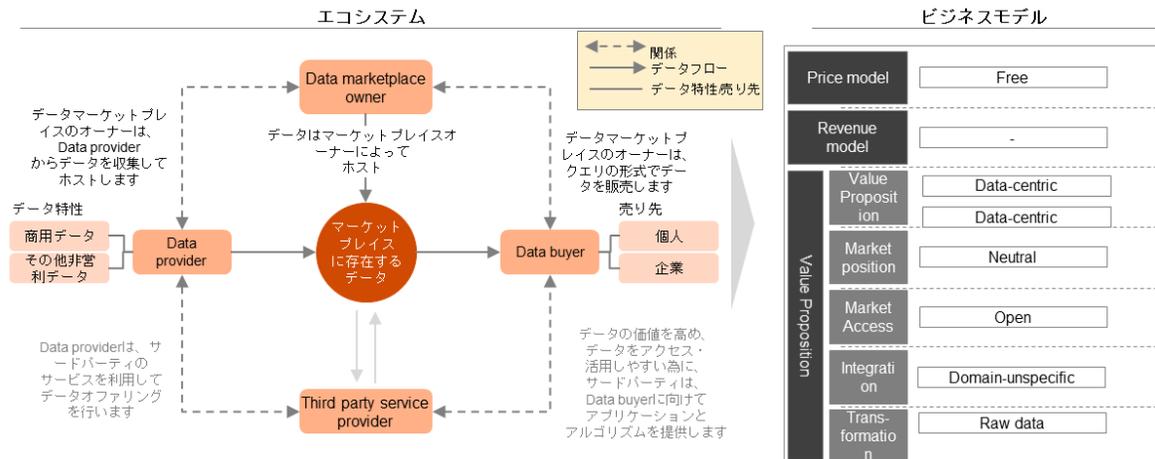
技術 ID関連技術

データ提供者にはadministrationの利用が必要で、データの利用にあたり、個人ユーザアカウント(Facebook/Microsoft)又は企業ユーザとしての登録が必要



ビジネスモデル 全体図

Microsoft Research Open Dataは基本無料のサービスとして提供される認識



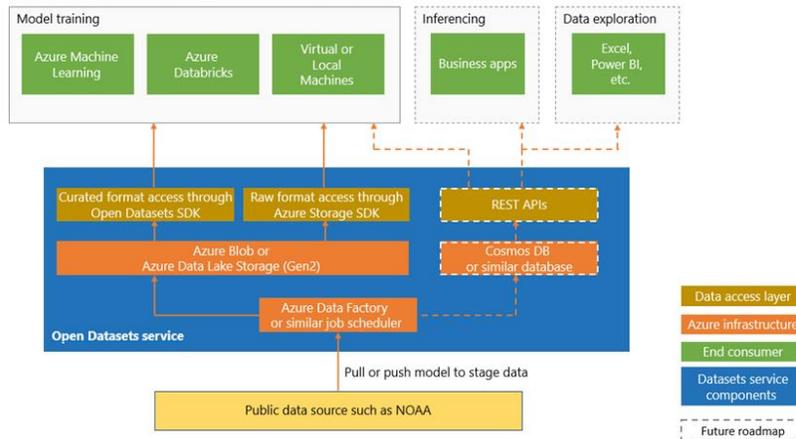
エグゼクティブサマリー Microsoft Azure Open Data

Microsoft Research Open DataはMicrosoftのオープンデータ共有の環境を推進する為の一環として、科学研究に必要なデータを無料サービスとして提供しつつ、オープンデータの共有にあたり必要な利用規約を策定する為のトライアル・模索と認識

深堀調査を通じて把握したファクト		ファクトから示唆	
知能目的 <ul style="list-style-type: none"> Microsoft Research Open Dataは、マイクロソフトの研究者が研究に関連して作成・公開したデータセットを公開するデータリポジトリ マイクロソフトの研究者や共同研究者がデータセットや関連する研究技術やツールを共有するためのシンプルなプラットフォームを提供 	データ連携基盤のありべき姿	<ul style="list-style-type: none"> データのディバインドを解消し、オープンデータの環境を育成する為のMicrosoftが主導する活動の一環として、Microsoft Research Open Dataは収益より、Microsoftがオープンデータ共有に関わるベストプラクティス(共有規則の策定など)へのトライアルとして戦略的価値を期待されている認識 オープンデータの共有環境を構築にあたり、データ共有のアライアンス関係の育成を目指すことと理解 	データ連携基盤の現勢
技術課題 <ul style="list-style-type: none"> データフォーマット・メタデータが対応するグローバル基準に関して記載不明(MS独自のルールで加工で、グローバル基準の対応はしていない可能性高) データの利用にあたり、MicrosoftまたはFacebookのユーザーアカウントが必要で、Microsoft側での認証技術を利用 APIは現在対応されていない理解 			
コンプライアンス <ul style="list-style-type: none"> データの登録にあたり、データの指名・審査、登録という一連のプロセスが必要 			
手続秘匿・秘開 <ul style="list-style-type: none"> サービス自体は無料で提供 			
運用条件 <ul style="list-style-type: none"> データの収集プロセスは基本Microsoftのチームより対応で、データオーナー(提供者)側の対応内容は不明；また、データの利用にあたり、Microsoftの個人ユーザーアカウントまたは企業アカウントでのログインが必要 法規制の対応について要確認 			

技術 全体図

高度な分析のためのデータのクリーニングと準備時間を削減する為、Azure Open Dataは選別されたデータソースからのデータに前処理を対応した上で、随時利用可能な状態を整える為データをAzure クラウドにも保存



技術 データ来歴、加工記録技術

データソースはMicrosoftより選定したため、科学研究に活用できるデータの為、来歴情報の確認はできるが、データ自体の可用性などSLAの担保は対応していないと認識データの前処理はAzure Data Factoryで対応する想定で、加工プロセスの管理が可能

データ来歴

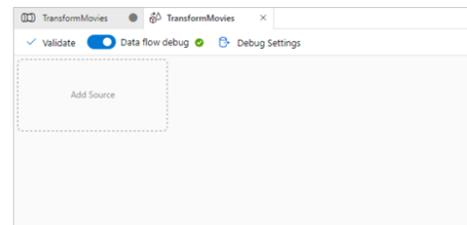
- Azure Open Dataには主に5つの種別で27データセットを有し、各データへの紹介(項目など)は公開ドキュメンテーションにて記載

種別	データセット例
Transportation	TartanAir, AirSim Simulation Dataset, NYC Taxi & Limousine Commission - yellow taxi trip records and etc.
Health and genomics	COVID-19 Data Lake, COVID-19 Open Research Dataset, Genomics Data Lake
Labor and economics	US Labor Force Statistics, US National Employment Hours and Earnings, US State Employment Hours and Earnings and etc.
Population and safety	US Population by County, US Population by ZIP Code and etc.
Supplemental and common datasets	Diabetes, OJ Sales Simulated Data

- マイクロソフトは、元のデータソースの品質や可用性を管理することができない為、データセットのSLAや可用性の保証もできないと明記

加工記録技術

- Azure Data Factoryにて実行したデータの処理は、データフローの形で管理可能



技術 認証

Azure Open Dataの利用にあたり、Azure環境の利用(アカウント)が必要の為、Azure Active Directory (AD)よりユーザーを認証；また、現時点はREST APIの開発はまだ対応できていない為、API経由の際の認証について、ヒアリングの中に確認予定※Azure ADで関連認証方法記載あり

Azure AD

- Azure ADで対応可能な認証手法は以下8つと定義
Password、SMS、音声、Microsoft Authenticator アプリ、OATH ソフトウェアトークン、OATH ハードウェアトークン(プレビュー)、Windows Hello for Business、FIDO2 セキュリティキー

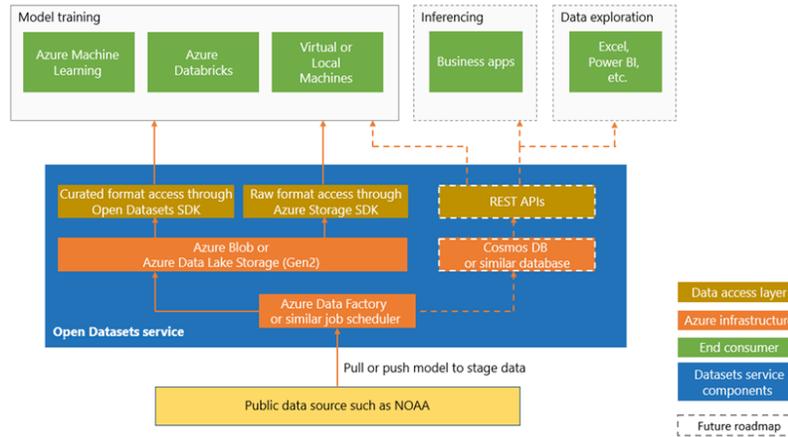
Bad: Password	Good: Password and...	Better: Password and...	Best: Passwordless
123456 qwerty password iloveyou Password1	SMS Voice	Authenticator (Push Notifications) Software Tokens OTP Hardware Tokens OTP	Windows Hello Authenticator (Phone Sign-in) FIDO2 security key

API認証(Microsoft Graph)



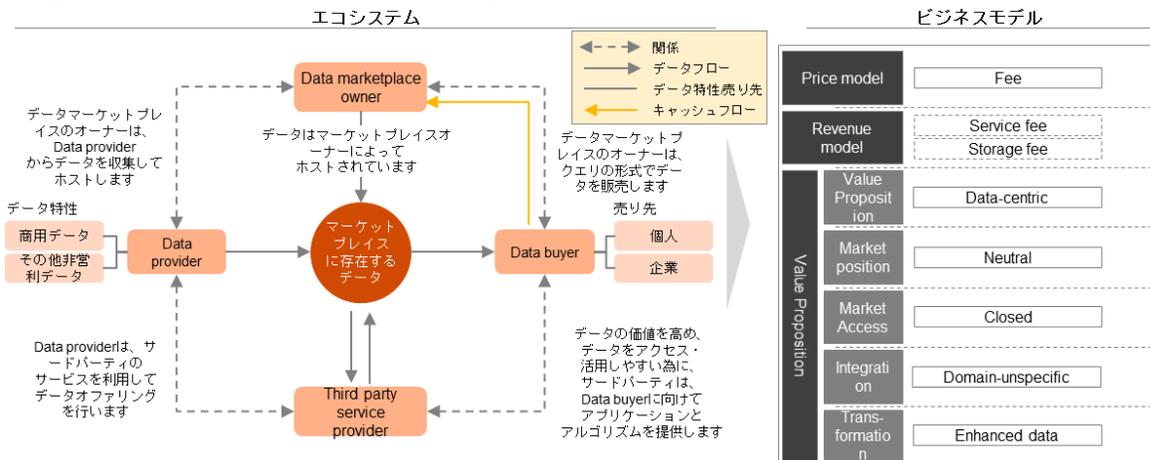
技術 ロードマップ

公開されたロードマップを踏まえ、今後Azureクラウドでの保存方式(パフォーマンス向上の為と推測)及びAPI連携を中心に開発活動を推進



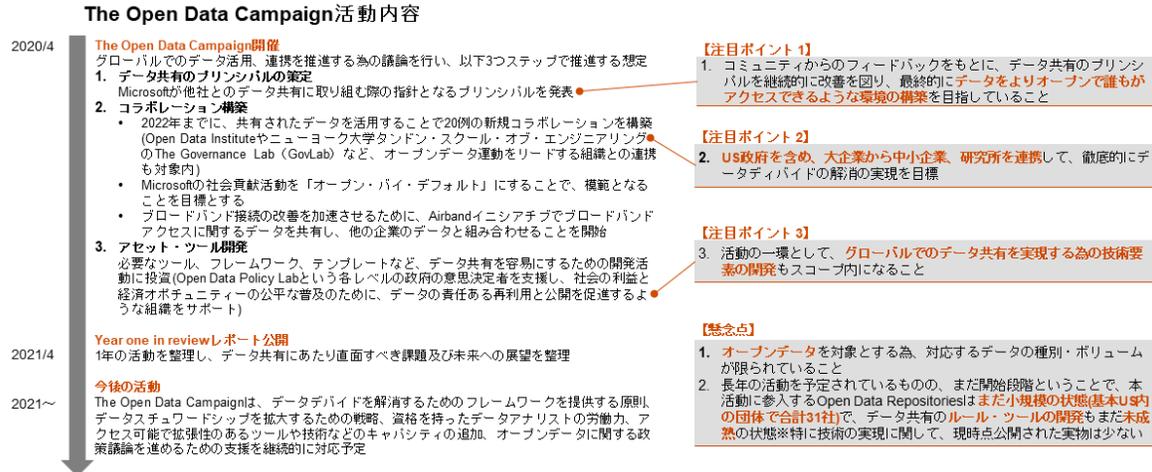
ビジネスモデル 全体図

Azure Open Datasetsにあるデータセットは無料でアクセスできますが、データにアクセスする Azure サブスクリプションには、大規模なデータセットの読み取りに関連するコストが課金される場合もある



The Open Data Campaign

2020年4月、Microsoftは、迫り来る「データデバイド」に対処し、あらゆる組織がアクセスできるオープンデータと、データ活用に必要な技術を整える為The Open Data Campaignを開催



(参考)Licensing, Governance, and Legal Tools

現在、オープンデータの共有に適切なルール・サポートツールが未成熟(または存在しない)であることを課題として見受けられる。この課題を解決する為The Open Data Campaignはデータ共有のシナリオに沿って利用規約の策定をし始め、現時点4種類の規約を公開

現時点公開されたデータ利用規約

O-UDA	Open Use of Data Agreement Open Use of Data Agreement (O-UDA) は、データを共有したいと考える個人や組織が、ユーザーに対する最小限の要求かつ、使用制限なしの前提で、簡単なデータ共有の実現を目的とする	C-UDA	Computational Use of Data Agreement Computational Use of Data Agreement (C-UDA)はAIのトレーニングでの利用を含め、特定のデータ使用シナリオに関して、法規則に準拠する利用方法を定義する
DUA-OAI	Data Use Agreement for Open AI Model Development DUA-OAIの目的は、人工知能 (AI) モデルを訓練するためのデータを共有し、訓練されたAIモデルをオープンソースライセンス構造で公開する際に、当事者が使用できる契約書のテンプレートを定義する	DUA-DC	Data Use Agreement for Data Commons DUA-DCは特定の主題に関する大規模なデータセットを所有している複数の関係者が、共通かつAPI対応可能なデータベースを通じてデータを共有したい場合に使えるテンプレートを定義する

(参考)Open Data Repositories

現在、The Open Data Campaignに参入した31種のOpen Data Repositoriesにおいて、29種がUS内のデータセットとなり、まだUS内のデータ共有を中心に取り組んでいる状態

#	Open Data Repositories	Country	#	Open Data Repositories	Country
1	Microsoft Azure Open Datasets	US	17	Open Data on GitHub	US
2	Microsoft Research Open Data	US	18	Open Images	US
3	Microsoft U.S. Broadband Usage Percentages Data	US	19	Open ML	- ※オンライン コミュニティ 2年に1回会議開催予定
4	Awesome Public Datasets	US	20	Public Data Sets for Testing and Prototyping on Azure	US
5	BotSpeak Database	Singapore	21	Quora list of datasets	US
6	BroadbandNow Open Data	US	22	Schlesinger-metoooproject	US
7	Bureau of Labor Statistics	US	23	Stanford Large Network Dataset Collection	US
8	California Open Justice Data	US	24	Stanford Open Policing Data	US
9	CodeSearchNet Corpus	US	25	Store-in-a-Box Solution Accelerator with product image collection	US
10	data.gov	US	26	Towards AI's Public Datasets for Machine Learning	US
11	Data.lacity.org	US	27	Udacity Self-Driving Car Dataset	US
12	Data World	US	28	University of California-Irvine Machine Learning Repository	US
13	FiscalData.Treasury.gov	US	29	USAFacts	US
14	Harvard Dataverse	US	30	USASpending.gov	US
15	HealthData.gov	US	31	Wide World Importers Sample Database	US
16	LA Neighborhood Data for Social Change	US			

IndiaStack

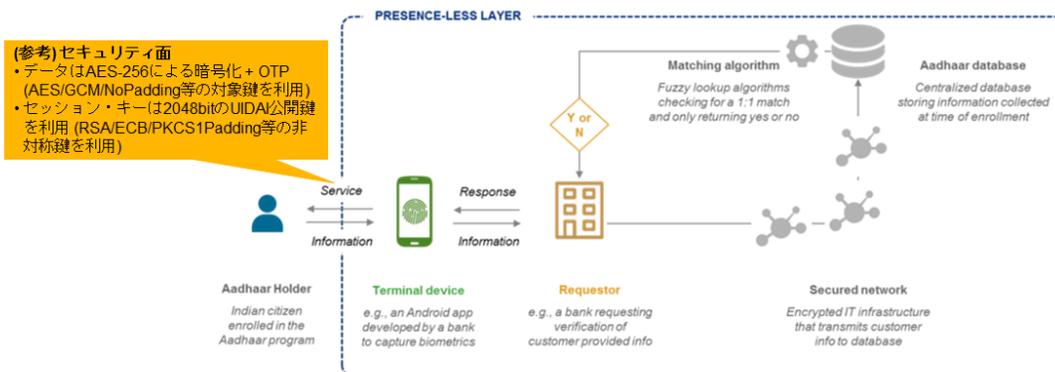
(前提) India Stackの4つのレイヤ

India Stackは本人確認、ペーパーレス、キャッシュレスの3つのレイヤが実装されており、4つ目となる同意レイヤが検討中。対象領域拡大と併せて、レイヤが追加される

用途レイヤ	目的	主なAPI	API概要	提供機関	提供開始年	補足
Consent Layer (同意レイヤ)	プライバシーを担保しながら、個人データを共有	未定	—	未定	未定	API開発に先行してDEPA(Data Empowerment and Protection Architecture) というフレームワークが考案されている
Cashless Layer (キャッシュレスレイヤ)	オンライン決済の実行	AEPS (Aadhaar Enabled Payment System) APB (Aadhaar Payment Bridge) UPI (Unified Payment Interface)	零細商店等での銀行取引サービス提供 政府の給付金/補助金の、銀行口座への振り込み スマートフォンを通じた、銀行口座間の送金	NPCI (インド決済公社)	2011 2016	—
Paperless Layer (ペーパーレスレイヤ)	KYC (Know Your Customer) 情報の提供や電子署名、文書のオンライン共有	Aadhaar eKYC (electronic Know Your Customer) eSign DigiLocker (Digital Lockerとも)	本人同意のもと、氏名・生年月日・性別・電話番号等の情報を第三者に提供 電子署名 運転免許証、卒業証明書等の文書を電子的に共有するサービス	UIDAI (固有識別番号庁) CCA (Controller of Certifying Authorities; 認証局の規制監督機関) MeitY (電子工学・通信技術省)	2012 2015	2015年~2020年4月でKYC情報が8兆回提供されている • DigiLockerへのアップロードは、認証機関・本人とも可能 • 保存された文書は、個人同意があれば、第三者への共有が可能(運転免許証など) • 2020年4月時点でDigiLockerのユーザー数は3700万人を超え、約37.5億件の文書が発行された • 文書発行機関には150以上の組織が、文書共有先機関には50近い組織が登録
Presence-less Layer (本人確認レイヤ)	リモートでの本人確認	Aadhaar Authentication	生体情報やデモグラフィック情報(氏名、生年月日、住所等)、OTP(ワンタイムPIN)を基にした本人確認機能	UIDAI	2010	提供開始から2020年4月までに、約396億回の本人確認が行われた

本人確認レイヤのアーキテクチャー

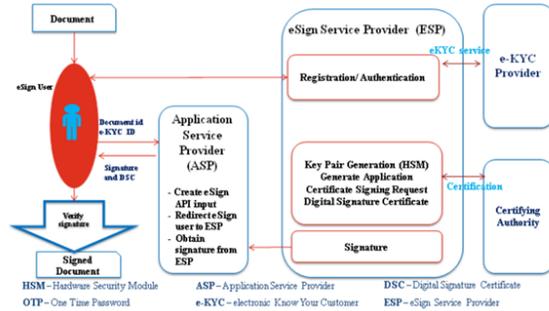
本人確認レイヤのAPIは、引数に個人情報を入力し、DB上の合致する情報の有無を返す。セキュリティは汎用的な手法で実装されており、導入の容易さを重視しているとみられる。



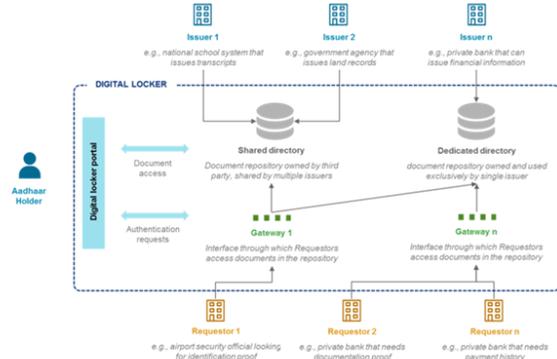
ペーパーレスレイヤのアーキテクチャー

ペーパーレスレイヤも、本人確認レイヤと同様にAPIの提供を行うが、Digital Lockerはデータ提供側/要求側双方にAPIが設定されるなど、詳細な仕様が異なる。

eSign (電子署名) のサービス概念図



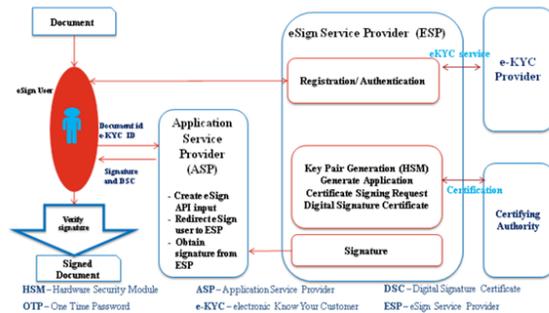
Digital Locker (文書の電子共有) のサービス概念図



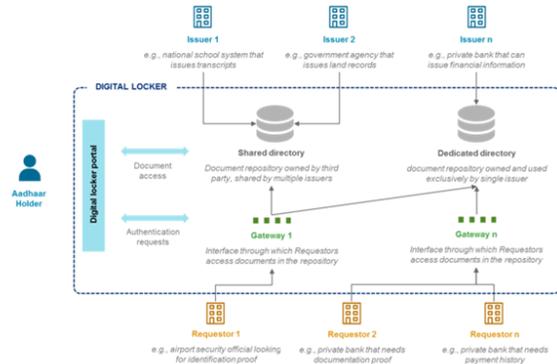
ペーパーレスレイヤにおける認証技術

eSignにおける認証はワンタイムパスワードに加え、FIDO2を採用。Digital LockerにおいてはOAuth2.0に加え、PKCE (Proof Key for Code Exchange) を採用するなど、オープンテクノロジーを活用しながら、ニーズに応じた使い分けを実現している。

eSign (電子署名) のサービス概念図



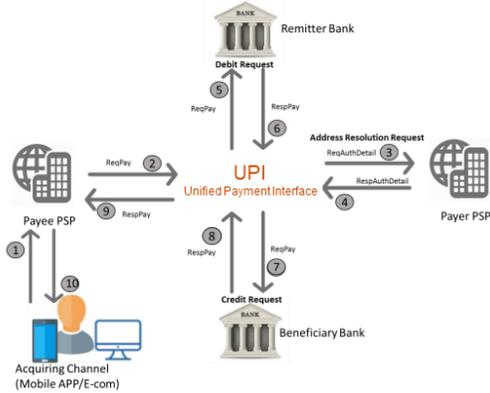
Digital Locker (文書の電子共有) のサービス概念図



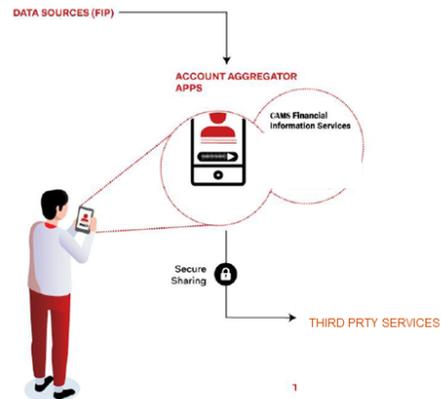
キャッシュレスレイヤ・同意レイヤのアーキテクチャ

キャッシュレスレイヤ・同意レイヤとも、アーキテクチャは定義されているものの、技術的な詳細は不明

キャッシュレスレイヤのアーキテクチャ



同意レイヤのアーキテクチャ



技術 データ・メタデータ対応基準

データ・メタデータ基準については不明であるものの、データはXML形式で出力される。APIの通信プロトコルについても、汎用プロトコルを採用しており、汎用性重視と考えられる。

データ
フォーマット

XML

- 本人確認レイヤ: APIにおける引数はXML形式で引き渡し
 - パーパレスレイヤ: データの保管形式については不明であるものの、APIにおける引数の引き渡し、APIコール結果となる出力結果は双方XML形式
- ※ APIの通信はHTTPSを使用

データ
メタデータ
基準

独自基準

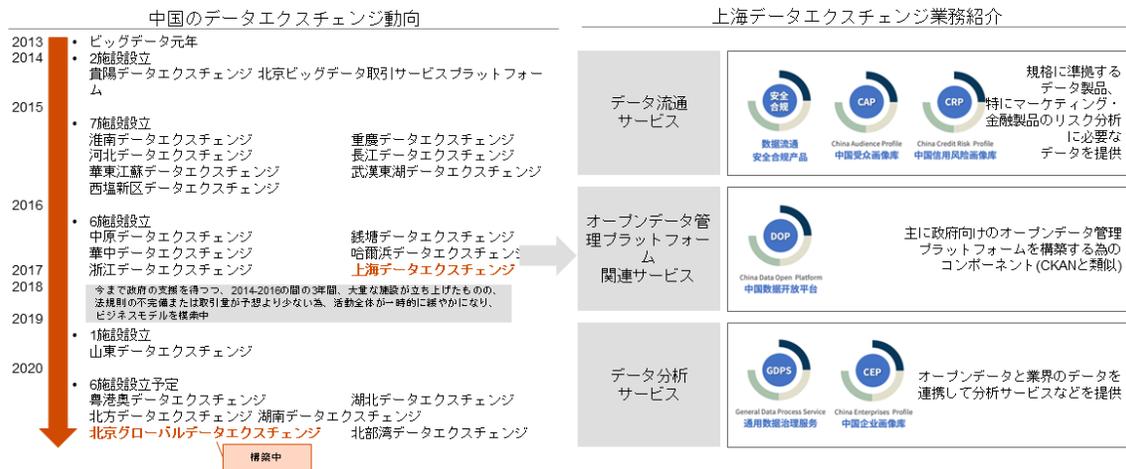
- 各データ領域ごとに責任省庁が任命されており、省庁ごとに基準を設定

上海データエクステンション

上海データエクステンション

9. 上海データエクステンション 概要

中国でビッグデータ元年と呼ばれる2013年以降、各地政府は積極的にデータ取引関連の動きを開始。上海データエクステンションは2016年に上海市人民政府等の承認を得て成立し、データ流通・オープンデータ管理プラットフォーム・データ分析サービスを提供



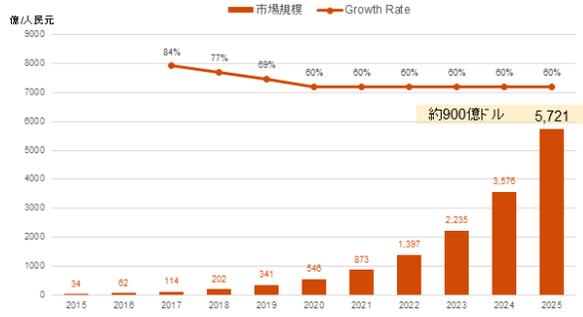
(参考)中国におけるデータ取引市場規模・事業形態

5Gの普及を踏まえ、2025年中国国内におけるデータ取引市場規模は約900億ドル
データ取引市場としてもデータの付加価値を高める為の加工関連のサービス(36%)
を提供する団体も多くみられる状況

データ取引の市場規模予測

- 5Gの普及を踏まえ、データの増長スピードが継続的に成長し、2025年中国国内におけるデータ取引市場規模は5,721億人民元(約900億ドル)になる見込み

中国データ取引市場 市場規模推移(2015~2025)

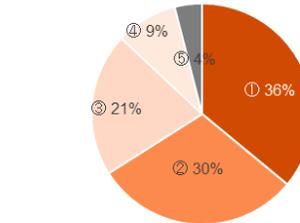


出典:PwC China, 国内ビッグデータ取引市場調査(2017年)

データ取引市場の収益モデル

- 主な収益モデルはデータの付加価値を高める為の加工関連のサービス(36%)となり、その次は自社データ販売(30%)、サードパーティーとしてのデータ取引市場(21%)の分布状況

中国データ取引市場 収益モデル分布状況(回答者割合)



- ① データの(価値を高める為の)サービス(データ収集・処理)による収益
- ② 自社データを提供することによる収益
- ③ データ取引市場市場を提供することによる収益(取引手数料)
- ④ ①~③におけるあらゆるサービスを提供し、会員ライセンス数による収益
- ⑤ その他

FHIR / HL7®

(参考) FHIR HL7® とは

FHIRは、HL7が公開している最新の医療データ交換のための標準規格であり、医療プロセス全般で使用される一連の機能を定義するプラットフォーム仕様を提供

HL7® とは

- Health Level Seven International (HL7)は、非営利団体であり、臨床現場や医療サービスの管理、提供、評価をサポートする電子医療情報の交換、統合、共有、検索のための包括的なフレームワークと関連標準を提供
- HL7とそのメンバーは、電子医療情報の交換、統合、共有、検索のためのフレームワーク（および関連規格）を提供し、これらの標準は、情報がどのようにパッケージ化され、ある当事者から別の当事者へと伝達されるかを定義し、システム間のシームレスな統合に必要な言語、構造、データタイプを設定
- HL7の標準規格は以下のように分類

カテゴリ名	説明
Primary Standards	システムの統合、相互運用性、コンプライアンスに不可欠な最も一般的な規格
Clinical Document Architecture (CDA®)	V3の進化版； CDA® (Clinical Document Architecture) 製品
EHR - Electronic Health Records	電子カルテを管理するための構成を可能にする機能モデルとプロファイルを提供
FHIR®	Fast Healthcare Interoperability Resources
Version 2 (V2)	テキスト情報交換； 医療現場における電子データ交換のためのアプリケーションプロトコル
Version 3 (V3)	XML情報交換； HL7の参照情報モデル (RIM) に基づく仕様 コンピュータ化された健康知識ベースを、担当者、情報システム、機関の間で共有することを容易にするために、手続き的な臨床知識を表現するための形式論
CCOW	HL7が従来から重視してきたデータ交換と企業のワークフローを補充する、内部アプリケーションのプログラミングと実行環境のインフラストラクチャの両方の標準として、アプリケーションの統合をポイントオブユースで促進
Cross-paradigm/ Domain Analysis Models	例：ドメイン解析モデル

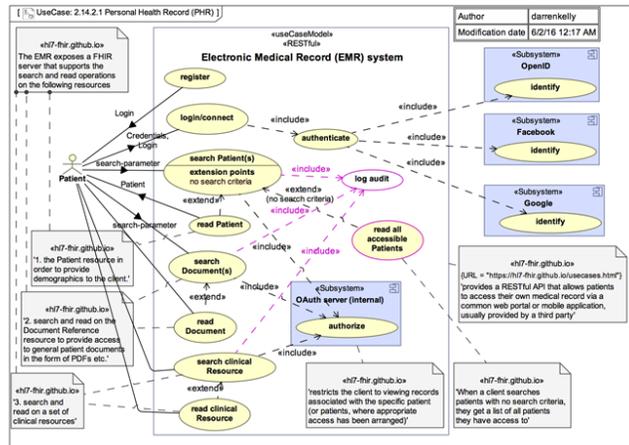
FHIR とは

- FHIR (Fast Healthcare Interoperability Resources : ファイヤー) は、HL7で定義された「リソース」のセットを利用して文書、メッセージ、サービス、RESTfulインターフェースなど、様々な手段で情報共有をサポートするヘルスケア情報交換規格
- Fast: 簡潔でわかりやすい仕様、実装者が使いやすい人間可読な形式のデータ、早くして簡単に実装可能
- Healthcare: 電子医療情報
- Interoperability: 既存の医療情報システムの情報を活用した相互運用性を確保 (米国等でのデジタルヘルス政策の中核は、「相互運用性の推進」)
- Resources: 「リソース」と呼ばれる「データ交換の小さな論理的に独立した単位」及びそのAPI仕様等を定義し、Restインターフェースによる交換、Documentsの交換、Messageの送受信、Serviceを公開したり呼び出したりしたものをサポート



技術 全体図

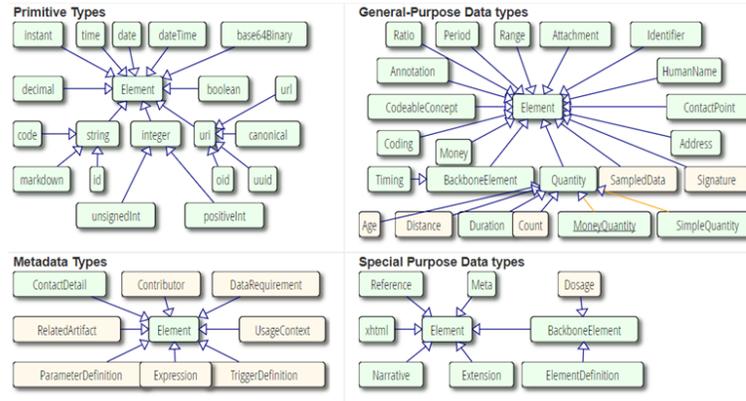
FHIR HL7の基本のユースケースとして、「患者が登録⇒ログイン⇒検索条件を入力して関心のある患者または自分と類似症状のある患者を特定⇒該当患者の臨床文書を検索⇒患者の臨床リソースを検索」という流れを想定



技術 データフォーマット・メタデータ対応基準

FHIR HL7により定義されたデータタイプは、プリミティブまたは複雑な性質を持つデータ構造、または他のリソースへの参照によって構成

各データタイプに定義されたコンテンツの内容



参考文献

1. NEDO「戦略的イノベーション創造プログラム(SIP)第2期/ビッグデータ・AIを活用したサイバー空間基盤技術/分野間データ連携基盤技術に関する国際動向調査」(2021年) (<https://www.nedo.go.jp/content/100931755.pdf>)
2. 内閣府「戦略的イノベーション創造プログラム(SIP)概要」(2021年) (<https://www8.cao.go.jp/cstp/gaiyo/sip/sipgaiyou.pdf>)
3. DATA-EX「データ基盤システム DATA-EXと海外の取組」(2021年) (https://www.kantei.go.jp/jp/singi/it2/dgov/data_strategy_tf/dai6/siryou4.pdf)
4. 流通流動化検討委員会「資金調達手法～証券化を中心に」(2007年) (https://system.jpaa.or.jp/patents_files_old/200704/jpaapatent200704_083-089.pdf)
5. NEDO「戦略的イノベーション創造プログラム(SIP)第2期/ビッグデータ・AIを活用したサイバー空間基盤技術」(2021年) (https://www.nedo.go.jp/activities/ZZJP_100158.html)
6. ミュンヘン連邦軍大学 サイバー防御研究機関 (CODE)「IMC: A Classification of Identity Management Approaches」(2020年) (<https://detips2020.github.io/PreProceedings/DETIPS2020-paper-01.pdf>)
7. NEDO「分野間データ連携基盤技術:採択テーマおよび委託先一覧」(2021年) (<https://www.nedo.go.jp/content/100899989.pdf>)
8. GAIA-X「Data Interoperability and Data Sovereignty for Food Production」(2021年) (<https://www.bmwi.de/Redaktion/EN/Artikel/Digital-World/GAIA-X-Use-Cases/fast.html>)
9. GAIA-X「AIQNET – Medical Data Ecosystem (previously KIKS)」(2021年) (<https://www.bmwi.de/Redaktion/EN/Artikel/Digital-World/GAIA-X-Use-Cases/aiqnet.html>)
10. GAIA-X「GAIA-X: The European project kicks off the next phase」(2021年) (https://www.bmwi.de/Redaktion/EN/Publikationen/gaia-x-the-european-project-kicks-of-the-next-phase.pdf?_blob=publicationFile&v=13)
11. Microsoft「The Future Of WindowsDirectory Services in Windows Server "Longhorn"」(2016年) ([https://docs.microsoft.com/en-us/previous-versions/technet-magazine/cc160894\(v=msdn.10\)?redirectedfrom=MSDN](https://docs.microsoft.com/en-us/previous-versions/technet-magazine/cc160894(v=msdn.10)?redirectedfrom=MSDN))
12. InfoWorld「Gaia-X: Will the European data cloud fail?」(2021年) (<https://www.infoworld.com/article/3623298/gaia-x-will-the-european-data-cloud-fail.html>)
13. International Data Spaces Association「Catena-X With GAIA-X: Will Data Space Be the Word of 2021?」(2021年)

- [\(https://internationaldataspaces.org/catena-x-with-gaia-x-will-data-space-be-the-word-of-2021/\)](https://internationaldataspaces.org/catena-x-with-gaia-x-will-data-space-be-the-word-of-2021/)
14. Ricardo N et al.「An Interledger Blockchain Platform for Cross-Border Management of Cybersecurity Information」(2020 年)
https://www.researchgate.net/publication/342255539_An_Interledger_Blockchain_Platform_for_Cross-Border_Management_of_Cybersecurity_Information/fulltext/5eeac2dc458515814a674d2e/An-Interledger-Blockchain-Platform-for-Cross-Border-Management-of-Cybersecurity-Information.pdf
 15. European Commission「The EU Cybersecurity Act」(2021 年) (<https://digital-strategy.ec.europa.eu/en/policies/cybersecurity-act>)
 16. International Data Spaces Association「More Funding for Data Spaces Promised」(2021 年) (<https://internationaldataspaces.org/more-funding-for-data-spaces-promised/>)
 17. PRESSEBOX「IDSA Accelerates Gaia-X with the Data Spaces Concept」(2020 年) (<https://www.pressebox.com/pressrelease/industrial-data-space-e-v/IDSA-Accelerates-Gaia-X-with-the-Data-Spaces-Concept/boxid/1032715>)
 18. International Data Spaces Association「IDSA Position Paper GAIA-X and IDS Version 1.0」(2021 年) (https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Position-Paper-GAIA-X-and-IDS.pdf)
 19. tributech「HOW GAIA-X, IDSA AND CATENA-X ARE CONNECTED AND COLLABORATE TO DEVELOP ONE DATA ECOSYSTEM STANDARD」(2021 年) (<https://tributech.io/blog/gaia-x-idsa-catena-x>)
 20. International Data Spaces Association「IDSA brochure International Data Spaces Enabling Data Economy」(2021 年)
https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-brochure-International-Data-Spaces-Enabling-Data-Economy.pdf
 21. International Data Spaces Association「IDS Reference Architecture Model 3.0」(2021 年) (<https://internationaldataspaces.org/wp-content/uploads/IDS-RAM-3.0-2019.pdf>)
 22. GitHub「International-Data-Spaces-Association/IDS-G」(2021 年)
<https://github.com/International-Data-Spaces-Association/IDS-G>
 23. Fraunhofer「Identity management with low organizational effort」(2021 年)
https://www.dataspaces.fraunhofer.de/en/software/identity_provider.html
 24. International Data Spaces Association「Membership Fee Regulations」(2021 年) (<https://internationaldataspaces.org/wp-content/uploads/IDSA-MembershipFeeRegulations-2021-01-01-EN.pdf>)

25. Fraunhofer「Sovereign and secure data exchange」(2021 年)
(<https://www.dataspaces.fraunhofer.de/en/software.html>)
26. Fraunhofer「General questions about the International Data Spaces」(2021 年)
(<https://www.dataspaces.fraunhofer.de/en/faq.html>)
27. International Data Spaces Association「General Data Privacy Policy」(2021 年)
(<https://internationaldataspaces.org/privacypolicy/>)
28. resources.data.gov「DCAT-US Schema v1.1 (Project Open Data Metadata Schema)」(2021 年)
(<https://resources.data.gov/resources/dcat-us/>)
29. data.world「Adding files to a dataset」(2021 年)
(<https://docs.data.world/en/64499-64813-0--Adding-files-to-a-dataset.html>)
30. data.world「Verify your data」(2021 年)
(<https://docs.data.world/en/64499-64520-3--Verify-your-data.html>)
31. data.world「WHAT IS CATALOGED」(2021 年)
(<https://docs.data.world/en/59261-61213-0--What-is-cataloged.html>)
32. data.world「Governance and auditability」(2021 年)
(<https://docs.data.world/en/64499-64528-11--Governance-and-auditability.html>)
33. data.world「How our data connections interact with your data」(2021 年)
(<https://docs.data.world/en/64499-64965-1--How-our-data-connections-interact-with-your-data.html>)
34. D.Brickley et al.「Google Dataset Search: Building a search engine for datasets in an open Web ecosystem」(2019 年)
(<https://www.semanticscholar.org/paper/Google-Dataset-Search%3A-Building-a-search-engine-for-Brickley-Burgess/bcbf86a3653270d5fe826ea9c5266e9c3a01613d>)
35. Schema.org「Dataset - Schema.org Type」(2021 年)
(<https://schema.org/Dataset>)
36. FourWeekMBA「The Power of Google Business Model in a Nutshell」(2021 年)
(<https://fourweekmba.com/google-business-model/>)
37. Google Cloud「データセットとビルド済みソリューション」(2021 年)
(<https://cloud.google.com/solutions/datasets>)
38. Google Cloud「About COVID-19 Public Datasets」(2021 年)
(<https://console.cloud.google.com/marketplace/product/bigquery-public-datasets/covid19-public-data-program>)
39. Google Cloud「Cloud Translation API」(2021 年)
(<https://console.cloud.google.com/marketplace/product/google/translate.google-apis.com>)

40. Google Cloud「Ubuntu 20.03 LTS(Focal)」(2021 年)
(<https://console.cloud.google.com/marketplace/product/ubuntu-os-cloud/ubuntu-focal>)
41. Google Cloud「Broad References」(2021 年)
(<https://console.cloud.google.com/marketplace/product/broad-institute/references>)
42. Google Cloud「gcp-public-data-broad-references」(2021 年)
(<https://console.cloud.google.com/storage/browser/gcp-public-data-broad-references?tab=objects?prefix=&forceOnObjectsSortingFiltering=false>)
43. Google Cloud「パートナーになる | Google Cloud」(2021 年)
(<https://cloud.google.com/partners/become-a-partner?hl=ja>)
44. CNBC「Google is slashing the amount it keeps from sales on its cloud marketplace as pressure mounts on app stores」(2021 年)
(<https://www.cnbc.com/2021/09/26/google-lowers-its-cloud-marketplace-revenue-share-to-3percent-from-20percent.html>)
45. KEN PAXTON ATTORNEY GENERAL of TEXAS「AG Pax-ton Leads Mul-ti-state Coali-tion in Law-suit Against Google for Anti-com-pet-i-tive Prac-tices and Decep-tive Misrepresentations」(2020 年)
(<https://www.texasattorneygeneral.gov/news/releases/ag-paxton-leads-multistate-coalition-lawsuit-against-google-anticompetitive-practices-and-deceptive>)
46. 日本経済新聞「日米貿易交渉どうみるか(下)デジタル協定、他交渉を先導」(2019 年)
(<https://www.nikkei.com/article/DGKKZO51346880U9A021C1KE8000/>)
47. RIETI「日米貿易交渉 どうみるか デジタル協定、他交渉を先導」(2019 年)
(<https://www.rieti.go.jp/jp/papers/contribution/ishikawa/08.html>)
48. AWS「AWS Data Exchange – データ製品の検索、登録、および使用について」(2019 年) (<https://aws.amazon.com/jp/blogs/news/aws-data-exchange-find-subscribe-to-and-use-data-products/>)
49. AWS「Paging Researchers, Analysts, and Developers」(2008 年)
(<https://aws.amazon.com/jp/blogs/aws/paging-researchers-analysts-and-developers/>)
50. AWS「AWS Data Exchange ユーザーガイド」(2021 年)
(https://docs.aws.amazon.com/ja_jp/data-exchange/latest/userguide/data-exchange-ug.pdf)
51. AWS「AWS Data Exchange を使用してデータ製品を動的に公開し、更新する」(2019 年) (<https://aws.amazon.com/jp/blogs/news/publish-and-update-data-products-dynamically-with-aws-data-exchange/>)

52. AWS「AWS アカウント ID」(2021 年)
(https://docs.aws.amazon.com/ja_jp/general/latest/gr/acct-identifiers.html)
53. AWS「AWS 認証情報の理解と取得」(2021 年)
(https://docs.aws.amazon.com/ja_jp/general/latest/gr/aws-sec-cred-types.html)
54. AWS「Amazon リソースネーム (ARN)」(2021 年)
(https://docs.aws.amazon.com/ja_jp/general/latest/gr/aws-arns-and-namespaces.html)
55. AWS「Data in AWS Data Exchange」(2021 年)
(<https://docs.aws.amazon.com/data-exchange/latest/userguide/data-sets.html>)
56. AWS「AWS Data Exchange の料金」(2021 年)
(<https://aws.amazon.com/jp/data-exchange/pricing/>)
57. AWS「海外のデータの合法的使用を明確化する (CLOUD) 法律」(2021 年)
(<https://aws.amazon.com/jp/compliance/cloud-act/>)
58. ZDNet Japan「データの流通体制をクラウド化、「AIDC Data Cloud」発表」(2021 年) (<https://japan.zdnet.com/article/35166363/>)
59. GitHub「microsoft/opendatacloud」(2020 年)
(<https://github.com/microsoft/opendatacloud/blob/master/docs/import-dataset.md>)
60. Microsoft Azure「Azure Open Datasets の価格」(2021 年)
(<https://azure.microsoft.com/ja-jp/pricing/details/open-datasets/#faq>)
61. InfoWorld「Gaia-X: Will the European data cloud fail?」(2021 年)
(<https://www.infoworld.com/article/3623298/gaia-x-will-the-european-data-cloud-fail.html>)
62. Microsoft「Closing the data divide: the need for open data」(2020 年)
(<https://blogs.microsoft.com/on-the-issues/2020/04/21/open-data-campaign-divide/>)
63. GeekWire「Microsoft launches Open Data Campaign to narrow ‘data divide’ as information is consolidated」(2020 年)
(<https://www.geekwire.com/2020/microsoft-seeks-narrow-data-divide-information-consolidated-among-companies-countries/>)
64. Support Centre for Data Sharing「Open Data Campaign – closing the data divide」(2021 年) (<https://eudatasharing.eu/data-sharing-practice-examples/open-data-campaign-closing-data-divide>)
65. Microsoft「REMOVING BARRIERS TO DATA INNOVATION」(2021 年)
(https://news.microsoft.com/wp-content/uploads/prod/sites/560/2021/03/Backgrounder-FAQ-Sheet_FINAL.pdf)

66. Microsoft「Closing the Data Divide: The Need for Open Data」(2021 年)
(<https://news.microsoft.com/opendata/#tools-and-capabilities>)
67. CCA「eSign API Specifications」(2020 年)
(<https://cca.gov.in/sites/files/pdf/esign/eSign-APIv3.3.pdf>)
68. DigiLocker「Issuer API Specification Version 1.11」(2021 年)
(https://img1.digitallocker.gov.in/assets/img/issuer_api/Digital%20Locker%20Issuer%20API%20Specification%20v1.12.pdf)
69. DigiLocker「Authorized Partner API Specification Version 1.11」(2021 年)
(<https://img1.digitallocker.gov.in/assets/img/Digital%20Locker%20Authorized%20Partner%20API%20Specification%20v1.11.pdf>)
70. DigiLocker「Dedicated Repository (Push) API Specification Version 1.7」(2016 年)
([https://cdntest.digitallocker.gov.in/assets/img/digital_locker_dedicated_repository_\(push\)_API_specification_v1_7_2.pdf](https://cdntest.digitallocker.gov.in/assets/img/digital_locker_dedicated_repository_(push)_API_specification_v1_7_2.pdf))
71. NPCI「IndiaStack–Can it drives Indian Fintech Ecosystem」(2021 年)
(<https://www.npci.org.in/newsletter/innovation-newsletter/topic-of-the-month.html>)
72. Global Times「Shanghai Data Exchange kicks off trading, with 100 traders making deals」(2021 年)
(<https://www.globaltimes.cn/page/202111/1239899.shtml>)
73. HL7 International「About HL7」(2021 年)
(<https://www.hl7.org/about/index.cfm?ref=nav>)
74. TechTarget Japan「医療分野におけるメッセージ交換の標準化規格「HL7」」(2012 年)
(<https://techtarget.itmedia.co.jp/tt/news/1201/20/news05.html>)
75. National Resource Centre for EHR Standards「HL7 Fast Healthcare Interoperability Resources Specification (FHIR)」(2021 年)
(<https://www.nrces.in/standards/hl7-international/hl7-fhir>)
76. HDX「公式ホームページ」(2021 年)
(<https://data.humdata.org/faq>)
77. HDX「Humanitarian Data Exchange Quality Assurance Framework」(2021 年)
(https://centre.humdata.org/wp-content/uploads/HDX_Quality_Assurance_Framework_Draft.pdf)
78. Australia Government「Australia’s National Digital Health Strategy: Framework For Action」(2018 年)
(<https://apo.org.au/sites/default/files/resource-files/2018-07/apo-nid182341.pdf>)
79. Blue Button「公式ホームページ」(2021 年)
(<https://www.va.gov/bluebutton/>)
80. data.gov「公式ホームページ」(2021 年)
(<https://www.data.gov/about>)

81. U.S. Government「PROJECT OPEN DATA」(2021 年) (<https://project-open-data.cio.gov/>)
82. UK Government「公式ホームページ」(2021 年) (<https://data.gov.uk/about>)
83. UK Government「National Data Strategy」(2019 年) (<https://www.gov.uk/guidance/national-data-strategy>)
84. UK Government「Technical documentation」(2021 年) (https://guidance.data.gov.uk/publish_and_manage_data/harvest_or_add_data/add_data/#add-an-organogram-to-a-dataset)
85. European Commission「Open data portals」(2021 年) (<https://digital-strategy.ec.europa.eu/en/policies/open-data-portals>)
86. EU Data Portal「公式ホームページ」(2021 年) (<https://data.europa.eu/en/about/about-dataeuropa.eu>)
87. EU Data Portal「Understanding the European Data Portal」(2021 年) (https://data.europa.eu/sites/default/files/edp_factsheet_portal_architecture_online.pdf)
88. Luxembourgish data platform「API Document」(2021 年) (<https://data.public.lu/en/docapi/>)
89. Dataportal.Asia「公式ホームページ」(2021 年) (<https://dataportal.asia/about/dpa>)
90. Daniel Cunningham「Asia Launches First Official Open Data Portal at Open Data Day 2021」(2021 年) (<https://www.businesswire.com/news/home/20210310006082/en/Asia-Launches-First-Official-Open-Data-Portal-at-Open-Data-Day-2021>)
91. Dataportal.Asia「Specifications」(2021 年) (<https://dataportal.asia/base/lib/Specifications.pdf>)
92. Dataportal.Asia「公式ホームページ Membership FAQs」(2021 年) (<https://dataportal.asia/about/faq>)
93. NDAP「What Is India's National Data and Analytics Platform (NDAP)?」(2020 年) (<https://analyticsindiamag.com/what-is-indias-national-data-and-analytics-platform-ndap/>)
94. NITI Aayog「NITI Aayog Releases Its Vision for the National Data and Analytics Platform」(2020 年) (<https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1600370>)
95. CHRONICLE「National Data And Analytics Platform」(2020 年) (<https://www.chronicleindia.in/current-affairs/352-national-data-and-analytics-platform>)
96. data.gov.sg「公式ホームページ」(2021 年) (<https://data.gov.sg/about>)

97. Wendell Santos「How Singapore Will Run the Country Using APIs」(2018 年)
(<https://www.programmableweb.com/news/how-singapore-will-run-country-using-apis/else-where-web-case-study/2018/06/24>)
98. Ross Thompson「Growing by Shrinking - Consolidating Data on the Open Government Portal」(2021 年) (<https://open.canada.ca/en/blog/growing-shrinking-consolidating-data-open-government-portal>)
99. CKAN「CKAN for Government」(2021 年) (<https://ckan.org/government>)
100. CKAN「CKAN for Government」(2021 年)
(<https://ckan.org/government>)
101. Government of Canada「Open Standards and Open Source Software in the Government of Canada」(2018 年)
(<https://events19.linuxfoundation.org/wp-content/uploads/2017/11/Open-Standards-and-Open-Source-Software-in-the-Government-of-Canada-S%C3%A9bastien-Lemay-Treasury-Board-of-Canada-Secretariat.pdf>)
102. Facebook「公式ホームページ」(2021 年) (<https://ja-jp.facebook.com/business/help/1910428752389787?id=2042840805783715>)
103. Facebook「Facebook for developers」(2021 年)
(<https://developers.facebook.com/docs/graph-api/#---api>)
104. goodjobsfirst「Tracking Subsidies, Promoting Accountability in Economic Development」(2021 年)
(<https://subsidytracker.goodjobsfirst.org/parent/facebook>)
105. workplace for Facebook「技術リソース:シングルサインオン(SSO)」(2021 年)
(<https://www.workplace.com/resources/tech/authentication/sso>)
106. Datos「公式ホームページ」(2021 年) (<https://datos-health.com/about-us/>)
107. GrowJo「Datos Health Competitors, Revenue, Alternatives and Pricing」(2021 年) (https://growjo.com/company/Datos_Health)
108. esri「ObjectID フィールドの基礎」(2021 年)
(<https://desktop.arcgis.com/ja/arcmap/10.3/manage-data/tables/fundamentals-of-objectid-fields.html>)
109. All Things D「Nokia Maps a Course for Its Location Business, Unveils “Here” Cloud Service」(2012 年) (<https://allthingsd.com/20121113/nokia-maps-a-course-for-its-location-business-unveils-here-cloud-service>)
110. HERE「公式ホームページ」(2021 年) (<https://www.here.com/jp/platform>)
111. Egil Juliussen「Monetizing Connected Car Data」(2021 年)
(<https://www.eetasia.com/monetizing-connected-car-data/>)
112. Quandl「公式ホームページ」(2021 年) (<https://www.quandl.com/about>)

113. GrowJo「Quandl Competitors, Revenue, Alternatives and Pricing」(2021 年) (<https://growjo.com/company/Quandl>)
114. Quandl「ログイン説明」(2021 年) (<https://data.nasdaq.com/sign-up>)
115. Nasdaq Data Link API Documentation「FIREWALL ERRORS」(2021 年) (<https://docs.data.nasdaq.com/docs>)
116. Microsoft「データ提供者及び利用者が安心して利用できるデータ取引サービス「AIDC Data Cloud」を発表」(2021 年) (<https://news.microsoft.com/ja-jp/2021/02/10/210210-aidc-data-cloud/>)
117. Microsoft「Facts about Microsoft」(2021 年) (<https://news.microsoft.com/facts-about-microsoft/>)
118. crunchbase「Microsoft introduction」(2021 年) (<https://www.crunchbase.com/organization/microsoft>)
119. Microsoft「About Microsoft」(2021 年) (<https://www.microsoft.com/en-us/about>)
120. Microsoft「About Microsoft Research Open Data」(2021 年) (<https://msropendata.com/about>)
121. Microsoft「Azure Open Datasets FAQ」(2021 年) (<https://azure.microsoft.com/ja-jp/pricing/details/open-datasets/#faq>)
122. Alibaba「公式ホームページ」(2021 年) (<https://www.alibabagroup.com/cn/about/overview>)
123. Dawex「公式ホームページ」(2021 年) (<https://www.dawex.com/en/news/data-exchange-platform-10-essentials/>)
124. RAY SHARMA「Deutsche Telekom Launches Data Intelligence Hub - A Virtual Marketplace for Trading Data」(2018 年) (<https://www.thefastmode.com/services-and-innovations/13434-deutsche-telekom-launches-data-intelligence-hub-a-virtual-marketplace-for-trading-data>)
125. Data Intelligence Hub「Implementing IDS」(2021 年) (<https://internationaldataspaces.org/implementing-ids/>)
126. Data Intelligence Hub「公式ホームページ」(2021 年) (<https://dih.telekom.net/en/faqs/#qualityensurance>)
127. Data Intelligence Hub「Getting Started Guide」(2021 年) (<https://dih.telekom.net/en/getting-started-guide/>)
128. Adobe「公式ホームページ」(2021 年) (<https://www.adobe.com/jp/about-adobe.html>)
129. britannica「Adobe Inc.」(2021 年) (<https://www.britannica.com/topic/Adobe-Systems-Incorporated>)

130. Eleanor Dickinson「IXUP acquires collapsed Data Republic IP for \$3M」 (2021 年) (<https://www.arnnet.com.au/article/688913/ixup-acquires-collapsed-data-republic-ip-3m/>)
131. Data Republic「公式ホームページ」 (2021 年) (<https://www.datarepublic.com/about>)
132. GrowJo「Data Republic Competitors, Revenue, Alternatives and Pricing」 (2021 年) (https://growjo.com/company/Data_Republic)
133. Data analytics seminar「Beyond the Data Science Bubble」 (2017 年) (<https://actuaries.asn.au/Library/Events/DataAnalytics/2017/DataAnalyticsSeminar2017SteveMillwardPresentation.pdf>)
134. Data Republic「Data Republic and Personally Identifiable Information (PII)」 (2021 年) (<https://help.datarepublic.com/en/articles/1454846-data-republic-and-personally-identifiable-information-pii>)
135. Data Republic「Data Republic Architecture」 (2021 年) (<https://help.datarepublic.com/en/articles/2820042-data-republic-architecture>)
136. Data Republic「Data Republic's Legal Framework」 (2021 年) (<https://help.datarepublic.com/en/articles/2442106-data-republic-s-legal-framework>)
137. McGrathNicol「Administrators' report to creditors pursuant to section 75-225 of the Insolvency Practice Rules (Corporations) 2016」 (2016 年) (<https://www.mcgrathnicol.com/app/uploads/Data-Republic-Administrators-report-to-creditors.pdf>)
138. Data Republic「Data Sandbox」 (2021 年) (<https://www.datarepublic.com/data-sandbox>)
139. Data Republic「Configuring your Contributor Node for Single Sign-On」 (2021 年) (<https://help.datarepublic.com/en/articles/4995634-configuring-your-contributor-node-for-single-sign-on>)
140. Data Republic「What is Privacy-Preserving Matching?」 (2021 年) (<https://help.datarepublic.com/en/articles/2219705-what-is-privacy-preserving-matching>)
141. COLIN HARPER「What is Streamr DATAcoin」 (2018 年) (<https://coincentral.com/streamr-datacoin-beginners-guide/>)
142. fundinguniverse「Axiom Corporation History」 (2021 年) (<http://www.fundinguniverse.com/company-histories/axiom-corporation-history/>)

143. Acxiom「ACXIOM ANNOUNCES NEW PARTNERSHIP WITH DEMYST」(2020 年) (<https://www.acxiom.com/news/acxiom-announces-new-partnership-with-demyst/>)
144. Precisely「White Paper」(2021 年) (<https://www.precisely.com/about-us/syncsort-is-now-part-of-precisely>)
145. Precisely「Connect」(2021 年) (<https://www.precisely.com/product/precisely-connect/connect>)
146. Jeremy Cox「The Knowledge Fabric from Pitney Bowes: A Platform for Customer Context Management」(2021 年) (<https://www.pitneybowes.com/content/dam/pitneybowes/us/en/campaign-pages/the-knowledge-fabric-from-pitney-bowes-by-ovum-white-paper-us-english.pdf>)
147. Precisely「Precisely APIs」(2021 年) (<https://developer.precisely.com/pricing>)
148. ownerIQ「公式ホームページ」(2021 年) (<https://www.owneriq.com/about-us>)
149. ownerIQ「CoEx Data Onboarding Policy」(2021 年) (<https://support.owneriq.com/hc/en-us/articles/360024257571-CoEx-Data-Onboarding-Policy>)
150. crunchbase「Harbr introduction」(2021 年) (<https://www.crunchbase.com/organization/harbr>)
151. Ingrid Lunden「Harbr raises \$38.5M to help enterprises exchange and share big data troves securely」(2020 年) (<https://techcrunch.com/2020/11/16/harbr-raises-38-5m-to-help-enterprises-exchange-and-share-big-data-troves-securely/>)
152. Harbr「公式ホームページ」(2021 年) (<https://www.harbrdata.com/about/>)
153. crunchbase「Data.world introduction」(2021 年) (<https://www.crunchbase.com/organization/data-world>)
154. Ron Miller「Data.world introduces enterprise data collaboration platform」(2018 年) (<https://techcrunch.com/2018/03/06/data-world-introduces-enterprise-data-collaboration-platform/>)
155. The Economic Times「OPENRISE TECHNOLOGIES PRIVATE LIMITED」(2019 年) (<https://economictimes.indiatimes.com/company/openrise-technologies-private-limited/U72200AP2013PTC090657>)
156. Openprise「公式ホームページ」(2021 年) (<https://www.openprisetech.com/>)

157. Openprise「Getting Started Data Sources」(2021 年)
(<https://help.openprisetech.com/hc/en-us/articles/360020263094-Data-Sources>)
158. Crux Informatics「公式ホームページ」(2021 年)
(<https://www.cruxinformatics.com/company>)
159. GrowJo「Crux Informatics introduction」(2021 年)
(https://growjo.com/company/Crux_Informatics#:~:text=Estimated%20Revenue%20%26%20Financials,currentl%20%247.4M%20per%20year)
160. Crux Informatics「Crux Informatics Closes \$36 Mln Funding Round」(2021 年)
(https://growjo.com/company/Crux_Informatics#:~:text=Estimated%20Revenue%20%26%20Financials,currentl%20%247.4M%20per%20year)
161. Nick Jordan「In Search of Efficient Data: Narrative's Origin Story」(2019 年) (<https://blog.narrative.io/in-search-of-efficient-data-narratives-origin-story>)
162. Narrative.io「公式ホームページ」(2021 年)
(<https://www.narrative.io/about>)
163. owler「Narrative.io introduction」(2021 年)
(<https://www.owler.com/company/narrativeio>)
164. Narrative.io「Knowledge Base What is the process for new Distribute Partners?」(2021 年) (<https://kb.narrative.io/what-is-the-process-for-new-distribute-partners>)
165. Narrative.io「Pricing plan」(2021 年) (<https://www.narrative.io/pricing>)
166. Narrative.io「API Document」(2021 年)
(<https://api.narrative.dev/#endpoint-datasets>)
167. NIEM「公式ホームページ」(2021 年) (<https://www.niem.gov/about-niem/history>)
168. NIEM「Value of NIEM」(2021 年) (<https://www.niem.gov/about-niem/value-niem>)

契約管理番号:	21500696-0
---------	------------