

データ未来会議（IPA, DSA共催）

# DATA-EXの展開にむけて

IDSA Japan Hub及びデータスペース技術国際テストベッド

2024.01.31

越塚 登

データ社会推進協議会（DSA）・会長

IDSA Japan Hub Coordinator

東京大学 大学院情報学環

# Dataspaces

データスペース

# Dataspace : 欧州がプッシュする重要なコンセプト

## ■ データスペースとは、相互にデータ交換／共有できる、個々のデータ基盤の集合体

▶ 以下の4つの性質を同時に満たす

### 1. 連邦型アーキテクチャ／Federated Architecture

- ▶ 物理的なデータ統合を行わ**ない**（データを集めない）
- ▶ データは管理している現状のまま

### 2. データの相互運用性／Interoperability

- ▶ データのサイロを打ち破り
- ▶ ベンダー依存性を排除する

### 3. データ主権（Data Sovereignty）とトレーサビリティ

### 4. 参加主体とデータのトラストの実現

# **PART 1**

## **DATA-EX (分野間データ連携基盤)**



# World of Data-Driven Innovation

データ利活用によりイノベーションが持続的に起こる世界

一般社団法人データ社会推進協議会

DSAについて >

**DSA**  
**VISION**  
ビジョン

データ利活用によりイノベーションが持続的に起こる世界へ。

一般社団法人データ社会推進協議会（DSA）では、産官学の連携により分野を超えた公正、自由なデータ流通と利活用による豊かな社会（「データ社会」という）を実現し、国内はもとより世界と連携し貢献を図ることを目的としています。その活動の一つとして、データ連携に係る既存の取組が協調した、連邦型の分野を超えたデータ連携をめざすプラットフォーム DATA-EXを推進してまいります。

# DATA-EX/DSA（データ社会推進協議会）

## https://data-society-alliance.org/

一般社団法人データ社会推進協議会

Language JP   [会員ログイン](#) [入会案内](#)

Amount of Data Created Daily (2024)  
exploringtopics.com/...\_data-generated-per-d...

「DATA-EX」とは [トピックス](#) [DSAについて](#) [委員会活動](#) [活動ライブラリー](#) [お問合せ](#) [検索](#)

HOME > 「DATA-EX」とは

## 「DATA-EX」とは

DATA-EXとは、データ連携に係る既存の取組が協調した「連邦型の分野を超えたデータ連携」を目指すプラットフォームです。

この取組みでは、SIP分野間データ連携基盤事業で開発したデータカタログ検索機能など分野間データ連携基盤技術（コネクタ）に加え、原本性保証・品質評価などの共通機能、データ管理機能、統計・解析、可視化などのデータ利用機能などの機能開発を行います。

DATA-EXは、国内のデータ連携のハブとなるとともに、GAIA-X等の国際的なデータ連携基盤との相互運用を見据え、海外の主要団体とも議論を重ね、社会実装を進めるものです。



なぜ「DATA-EX」が必要か？	「DATA-EX」取り組みマップ	DATA-EX分野間データ連携基盤の将来展望と開発環境	活動内容
------------------	------------------	-----------------------------	------

## 「DATA-EX」関連プロジェクト

エリア・データ連携基盤に関する取り組み

## なぜ「DATA-EX」が必要か？

最近では、データの活用がさまざまな分野で進み、人々の生活はより豊かになっています。しかし、個々のアプリケーションやサービスが独立して存在しているため、企業や業種等それぞれの分野の壁を越えたデータ流通ができないことが課題となっています。分野ごとにデータが分散しているため、必要なデータを取得するには複数のデータベースにアクセスする必要があります。そこで、DSAでは連邦型の分野を超えたデータ連携を目指すプラットフォームである「DATA-EX」の構築を推進しています。「DATA-EX」は、データを各分野ごとのデータベースに収集し、継続的に保持しながら、必要なデータのみを必要な時に抽出して活用する、連邦型のシステムです。「DATA-EX」によって多種多様なデータが統合されることで、例えば以下のような課題の解決に貢献することが期待できます。

## DATA-EX

DATA-EXとは、分野を超えたデータ連携を実現するために、DSA（データ社会推進協議会）が行う取組みの総称

その取組みの中核となる、DATA-EX分野間データ連携基盤（以下、「DATA-EX」）は、分野を超えてデータの発見と利用を可能とする技術的・社会的プラットフォーム

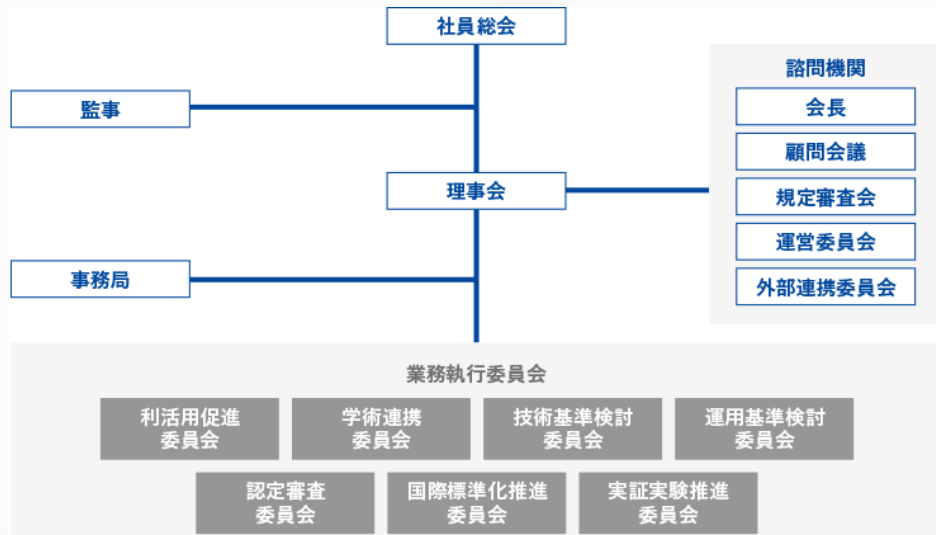
## ビジョン/Vision

“データ利活用によりイノベーションが持続的に起こる世界”  
“World of Data-Driven Innovation”

## ミッション/Mission

1. データ駆動型社会を構築し、イノベーションの民主化を実現  
Establish Data Driven Society with democracy of innovation
2. 世界規模で活用可能なデータ流通基盤を整備  
Develop data-distribution infrastructure for the world
3. 技術、サービス開発により社会実装を推進  
Accelerate social implementation with Technology and Service development
4. 世界と連携し、世界に貢献  
Collaboration and Contribution to the World

# Organization of Data Society Alliance (DSA)



## 理事・監事

### 会長

**越塚 登**  
 東京大学大学院  
 情報学環・教授  
[詳細を見る](#)

### 代表理事/理事長

**奥井 規品**  
 株式会社インターフュージョン・コンサルティング  
 代表取締役会長  
[詳細を見る](#)

### 代表理事/事務局長

**眞野 浩**  
 エプシロンジャパン株式会社  
 代表取締役  
[詳細を見る](#)

### 理事

- |   |  |
|---|--|
| <p><b>天野 雅典</b><br/>         富士通株式会社<br/>         JAPANビジネスグループ ビジネスクリエーター<br/>         シニアディレクター<br/> <a href="#">詳細を見る</a></p>      | <p><b>竹林 一</b><br/>         オムロン株式会社<br/>         イノベーション推進本部<br/>         インキュベーションセンター長<br/> <a href="#">詳細を見る</a></p>               |
| <p><b>高野 高志</b><br/>         株式会社ネクストウェイ<br/>         専務取締役<br/> <a href="#">詳細を見る</a></p>  | <p><b>寺澤 和幸</b><br/>         日本電気株式会社<br/>         クロスインダストリー企画本部<br/>         上席技術主幹<br/> <a href="#">詳細を見る</a></p>                   |
| <p><b>伊藤 直之</b><br/>         株式会社インターテック<br/>         事業開発本部 エンジニアリスト<br/> <a href="#">詳細を見る</a></p>                                  | <p><b>東條 真己</b><br/>         日鉄ソリューションズ株式会社<br/>         流通・サービスソリューション事業本部長<br/> <a href="#">詳細を見る</a></p>                            |
| <p><b>榎田 健治</b><br/>         大日本印刷株式会社<br/>         マーケティング本部 市場調査部<br/>         副部長<br/> <a href="#">詳細を見る</a></p>                   | <p><b>野口 誠</b><br/>         日本電気株式会社<br/>         デジタルトラスト推進本部長<br/> <a href="#">詳細を見る</a></p>   |
| <p><b>伊藤 雅典</b><br/>         日立製作所<br/>         社会インバケーション事業推進本部 事業戦略部<br/>         公共企画本部 本部長<br/> <a href="#">詳細を見る</a></p>         | <p><b>油谷 真紀</b><br/>         TIS株式会社<br/>         エグゼクティブフェロー デジタル社会サービス企画ユニット<br/>         ジェネラルマネージャー<br/> <a href="#">詳細を見る</a></p> |
| <p><b>木塚 伸介</b><br/>         富士ソフト株式会社<br/>         ソリューション事業本部<br/>         インフォメーションビジネス事業部 R&amp;D部<br/> <a href="#">詳細を見る</a></p> | <p><b>若目田 光史</b><br/>         株式会社日本総合研究所<br/>         リサーチ・コンサルティング部門 東 創発戦略センター<br/>         上席主任研究員<br/> <a href="#">詳細を見る</a></p>  |
| <p><b>真田 敦志</b><br/>         NTTコミュニケーションズ<br/>         ビジネスソリューション本部<br/>         第二ビジネスソリューション部 担当部長<br/> <a href="#">詳細を見る</a></p> | <p><b>監事</b></p>   |
| <p><b>杉山 恒司</b><br/>         株式会社ワフル<br/>         CDO (Chief Data Trading Officer)<br/> <a href="#">詳細を見る</a></p>                     | <p><b>飯倉 輝一郎</b><br/>         ひかり総合法律事務所<br/>         パートナー弁護士<br/> <a href="#">詳細を見る</a></p>  |
|   | <p><b>藤倉 孝文</b><br/>         三菱UFJ法律事務所・外国法共同事業<br/>         パートナー弁護士<br/> <a href="#">詳細を見る</a></p>                                   |

## A Members



## B Members





# 分野を超えたデータ基盤がmissing parts



ゼロ・カーボン社会



効率的な重要社会  
インフラの運用・管理



防災・災害対応  
迅速な救難活動、適切な情報提供



食の安心・安全  
日本の農産物の国際競争力強化



医療／ヘルスケア／福祉  
迅速な救急患者の病院搬送



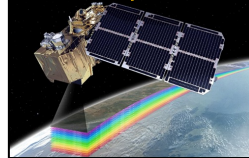
橋梁の歪み監視



被災状況モニタ



気象観測



リモートセンシング



食品の生産環境モニタ



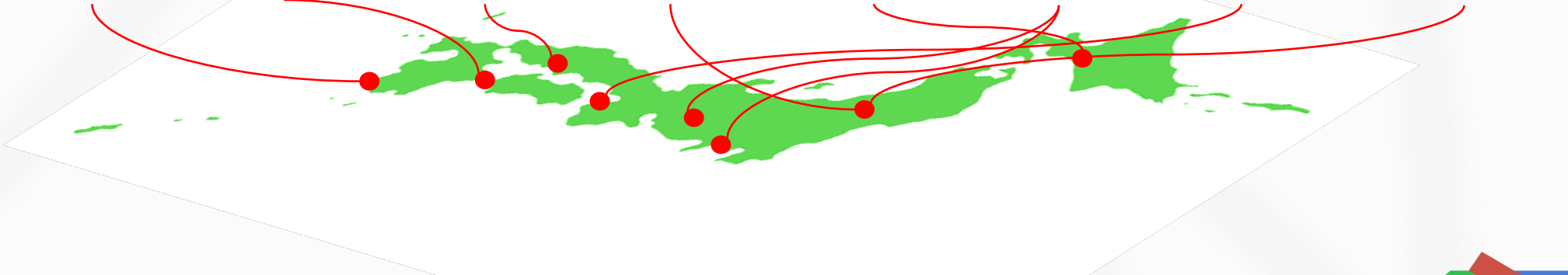
物流状況モニタ



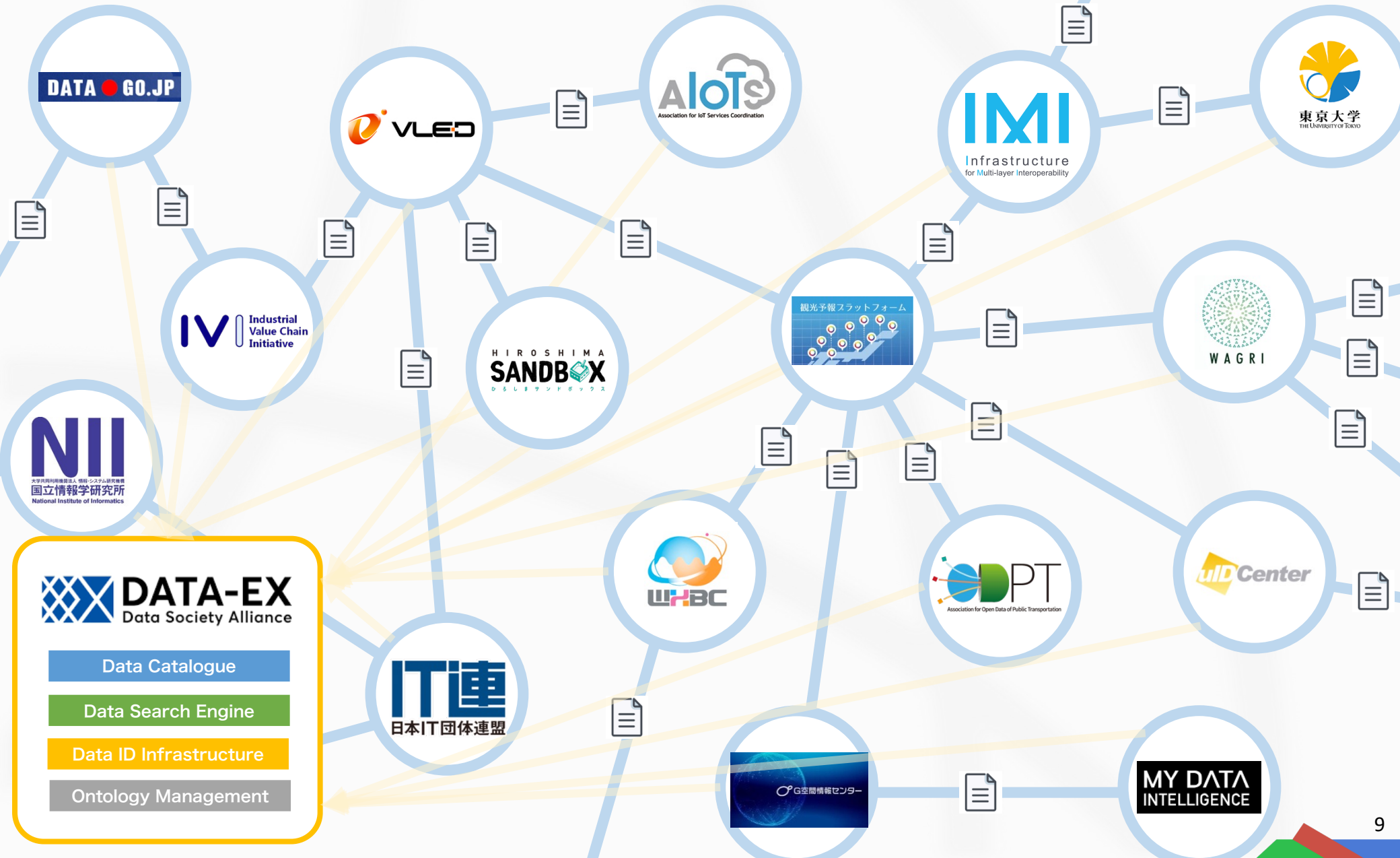
病院の稼働状況モニタ



患者の状況モニタ



# 分野間連携をするために、連邦型アーキテクチャ

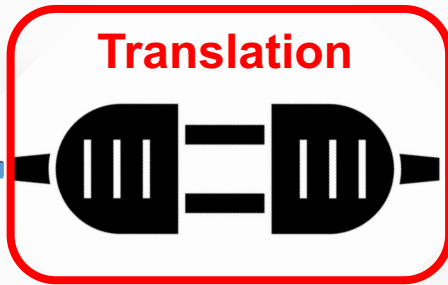


# “Connector”-based Federation Architecture

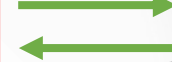
Your Data Platform



Your APIs and data formats



Connector



Standard APIs and data formats

Cross-Domain Global Dataspace

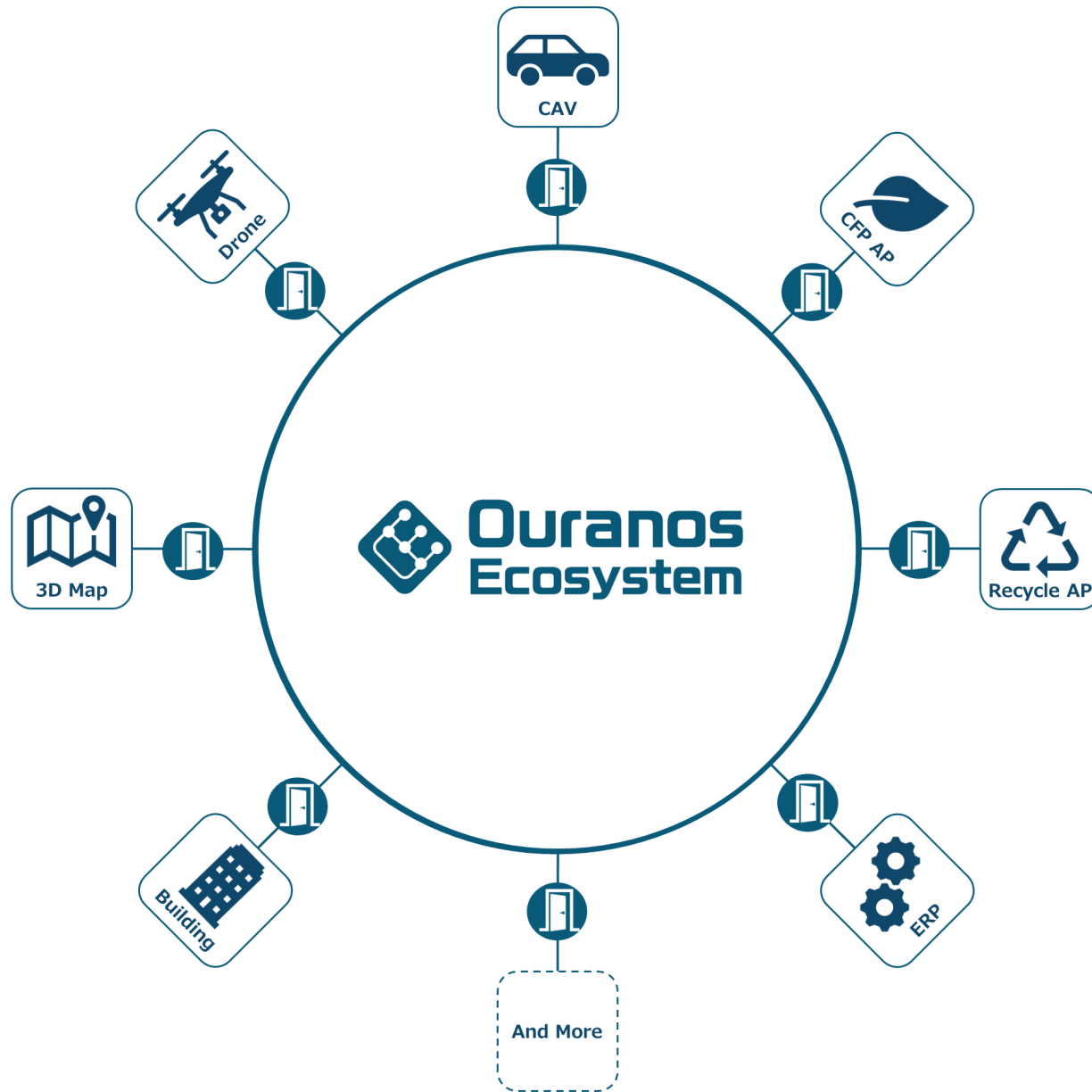
Domain-Specific Existing Local Data Platform



# 最近の進展 1

## Ouranos (ウラノス)

# Ouranos Ecosystem (METI)



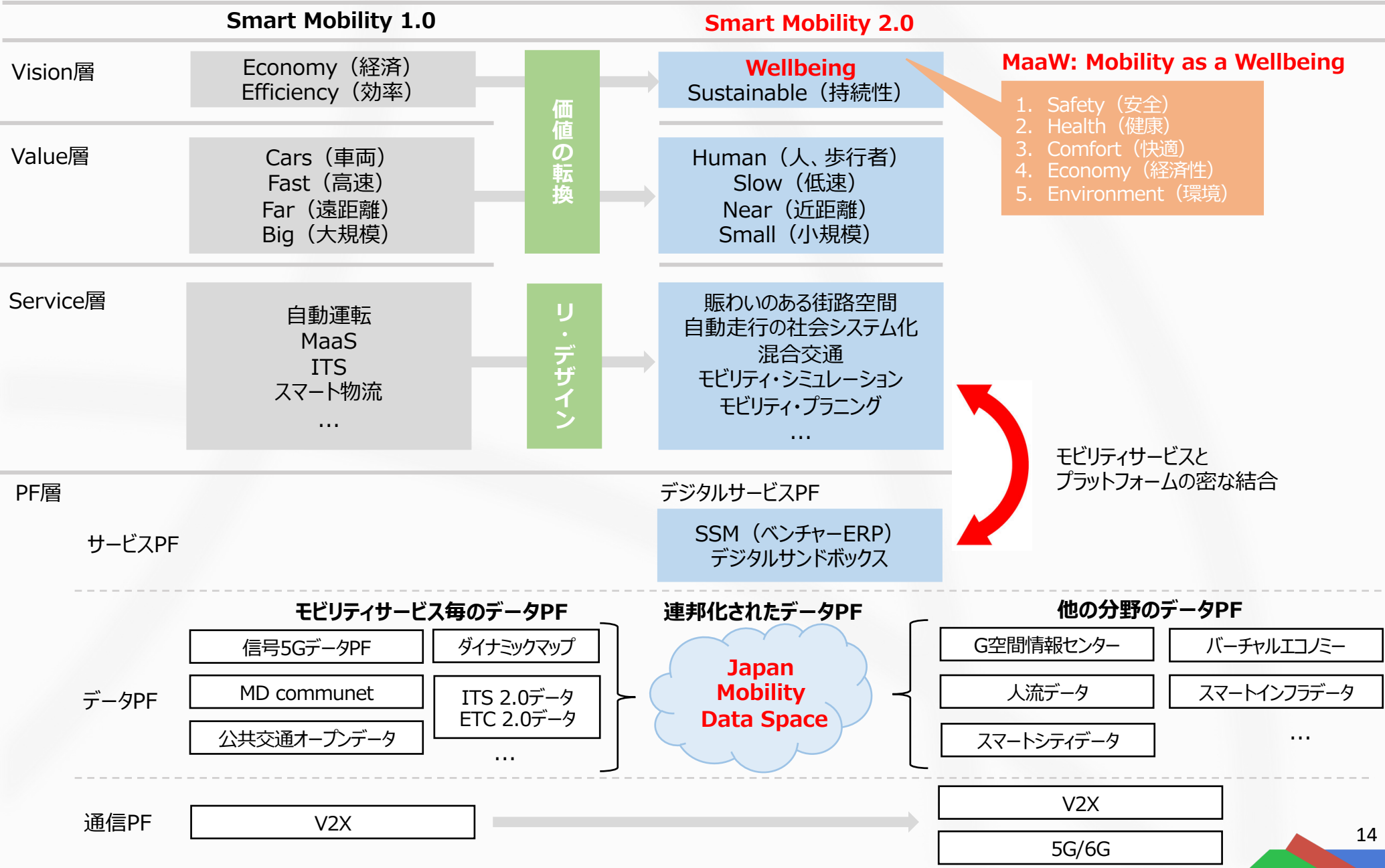




# 最近の進展 2

## Japan Mobility Data Spaces

# Smart Mobility 2.0の概念



# Japan Mobility Data Space (JMDS) : ビジョン・コンセプトと取組内容

<http://mobility-data-space.jp/>

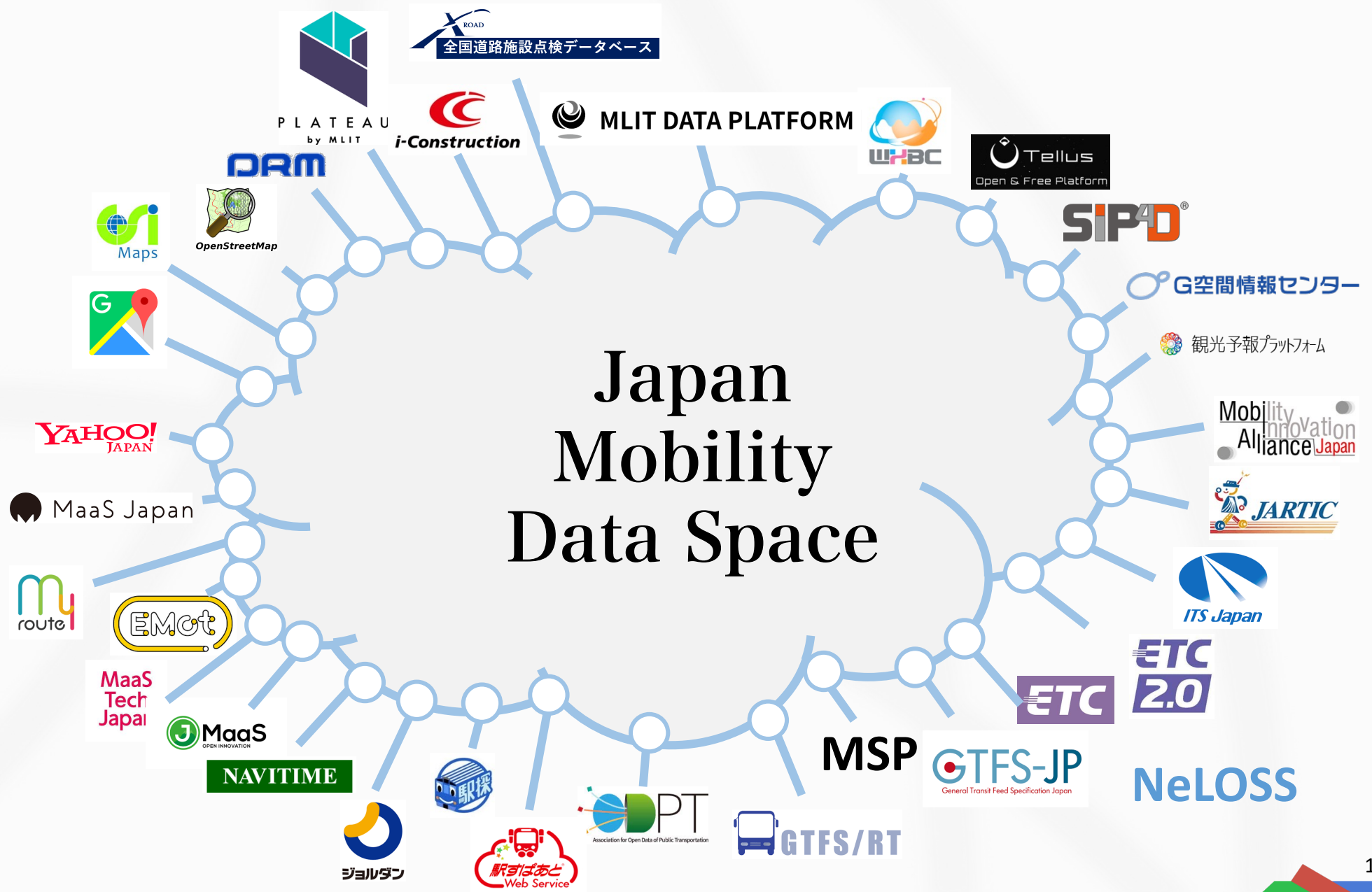
## ■ ビジョン・コンセプト

- ▶ モビリティに関連するデータの流通・共有環境を提供し、産官学の様々な組織間のデータエコノミーやデータエコシステムを創出する。
- ▶ 我が国におけるモビリティ事業の効率化やサービス向上、安全性の向上等に貢献し、モビリティのリ・デザイン及びスマートモビリティ2.0を実現する。

## ■ 取組内容

- ▶ 有用なデータを保有するデータホルダーと、革新的なモビリティソリューションを提供したいサービスを結びつけ、双方がWin-Winの関係になる。
- ▶ モビリティデータ流通のための技術仕様の策定
- ▶ 政府・自治体、民間企業、大学等が運用する様々なモビリティデータのプラットフォームを連携し、我が国のモビリティデータ取得のワンストップサービスを提供
- ▶ モビリティデータの利活用やそれによるイノベーションの創出を支援
- ▶ 国際連携を推進し、グローバルに貢献

# Japan Mobility Data Spaceのイメージ

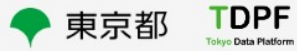




# 最近の進展 3

## TDPF（東京都）サービス開始

# TDPF: Tokyo Data PlatForm (2023~)



[トップ](#) [入会について](#) [会員一覧](#) [サービス紹介・申込み](#) [よくある質問・お問い合わせ](#)

[ログインはこちら](#)  
データ連携基盤サイト



# TOKYO DATA PLATFORM

東京データプラットフォーム

お知らせ

NEWS

2024.01.09 システムメンテナンスを実施します (1/15 22:00~23:00)

2023.12.22 年末年始休業のお知らせ

[一覧を見る](#)

# **PART 2**

## **国際連携活動の進展**

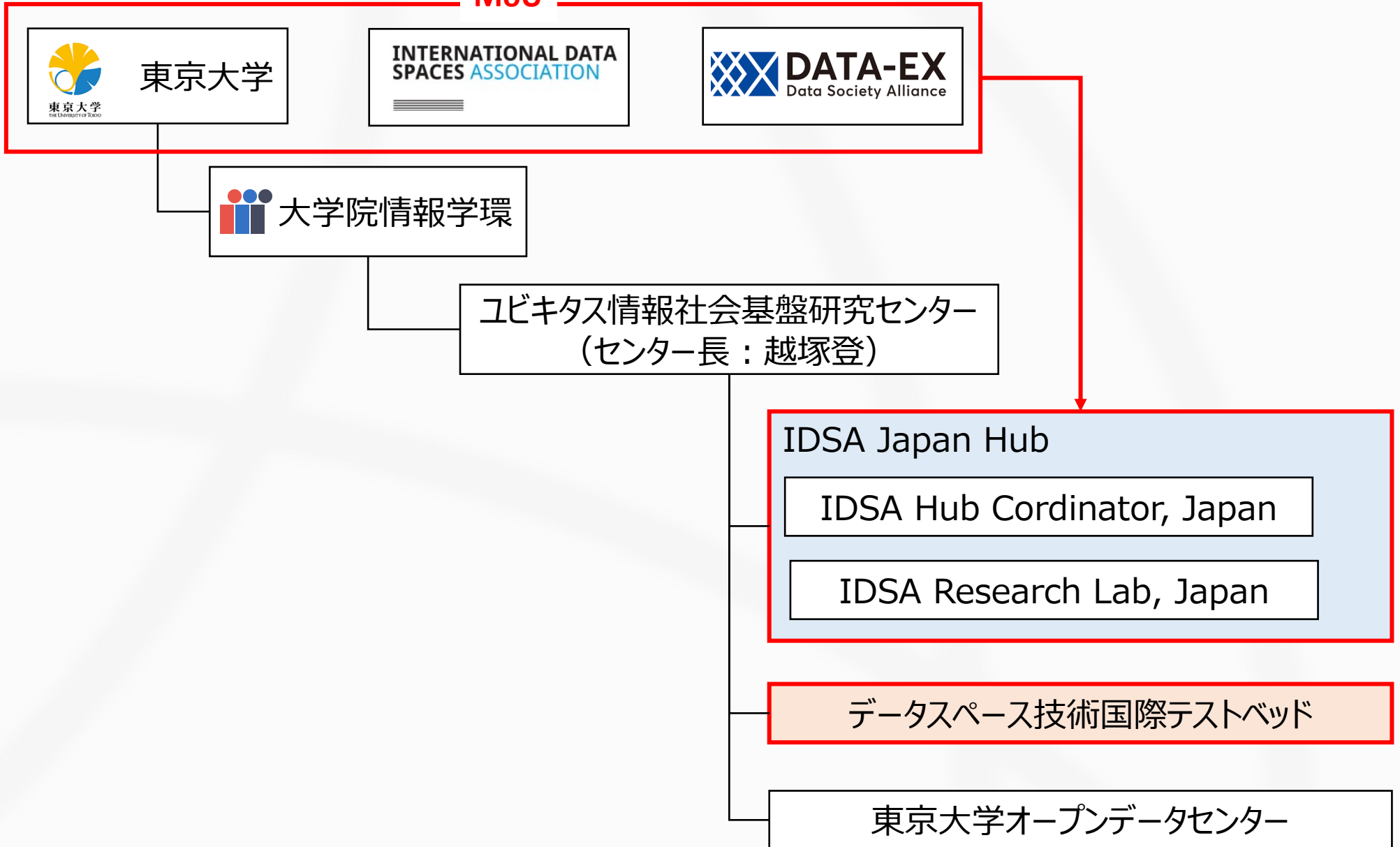


## 2-1 IDSA Japan Hub



# IDSA Japan Hub

MoU



# International Data Spaces Association

**INTERNATIONAL DATA  
SPACES ASSOCIATION**



News, Blog, Events



Newsletter



Contact



Calendar



WHY

WE

MAKE

OFFERS

ADOPT

PUBLICATIONS

KNOWLEDGE BASE

CONNECT



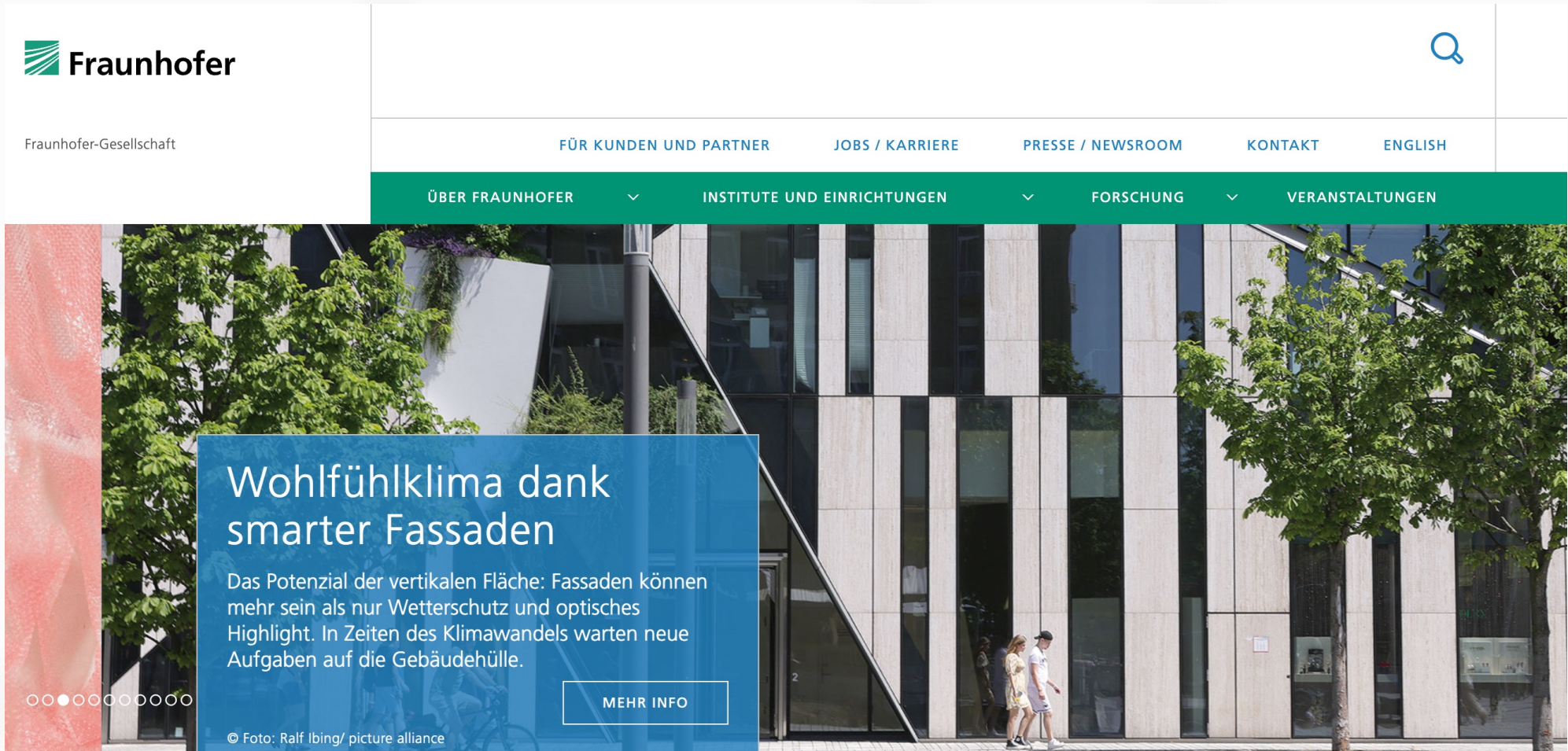
**INTERNATIONAL DATA SPACES**

The future of the data economy is here



# Fraunhofer 研究機構

- Fraunhofer 研究機構（独: Fraunhofer-Gesellschaft）は、ドイツ全土に75の研究施設・研究ユニットを持つ欧州最大の応用研究機関。



# IDSA Japan Hub

## IDSA Hub in Japan facilitated by University of Tokyo (UTokyo) in collaboration with Data Society Alliance (DSA)

The IDSA Japan Hub is facilitated by the University of Tokyo (UTokyo) in close collaboration with DSA (Data Society Alliance). UTokyo, as the oldest and largest national university in Japan, conducts research in a wide range of fields, including computer science, data science, digital infrastructure, economics, law, and other areas of expertise. At UTokyo, the Research Center for Ubiquitous Information Society Infrastructure, Interfaculty Initiatives in Information Study, plays a central role in this facilitation.

MORE 



Your contact person:

**Prof. Noboru Koshizuka**

Interfaculty Initiative in Information Studies

UTOKYO.JP



## **2-2**

# **Data Spaces Discovery Day Tokyo**

**November 22, 2023**

**日本橋ホール**





## International Data Spaces Association expands global reach with new hub in Japan

NOV 24, 2023 | NEWS



# LinkedIn IDSA Page

**International Data Spaces Association (IDSA)** 5,885 followers  
2d · 🌐

**IDSA breaks new ground: Inaugurates hub in Japan for global data space collaboration** <https://lnkd.in/ewyzvB98>

The International Data Spaces Association (IDSA) has just inaugurated a new IDSA Hub in Japan, expanding the horizons of data spaces beyond European borders.

The [Universität Tokio](#) and the Data Society Alliance have joined hands to steer the Japanese IDSA Hub. The formalization of this collaboration was marked by the signing of the Memorandum of Understanding (MoU) at the prestigious Data Spaces Discovery Day in Tokyo.

The event, held at the Nihonbashi Hall in Tokyo and virtually, drew a crowd of 350 participants. It served as a dynamic platform for impactful discussions, unveiling a shared commitment to unlocking the potential of data spaces across various domains and international borders.

The event, held at the Nihonbashi Hall in Tokyo and virtually, drew a crowd of 350 participants. It unleashed lively discussions on how the economic potential of data spaces can be exploited across different domains and international borders.

The establishment of the IDSA Hub in Japan marks a significant milestone in our journey, reinforcing our commitment to international collaboration and emphasizing the pivotal role of standards and interoperable data spaces on a global scale.



🌐 Anna Derevianko and 23 others

3 reposts

### Reactions



👍 Like    💬 Comment    🔄 Repost    ➦ Send

 Add a comment...  



## **2-3 IDSA Malaysia Hub**



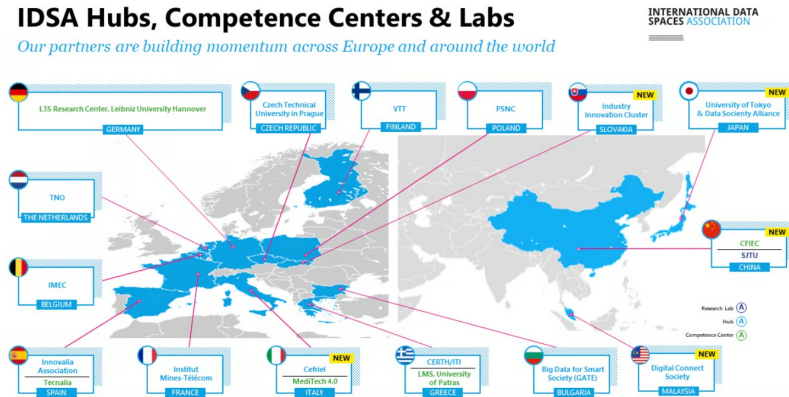
# IDSA Malaysia Hub (Since Jan. 29, 2024)



## Unlocking the Future: International Data Spaces Association (IDSA) Hub Malaysia Inauguration

### IDSA Hubs, Competence Centers & Labs

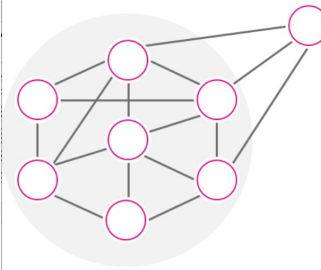
*Our partners are building momentum across Europe and around the world*



## What are Data Spaces?

Data spaces are expected to play a major role in supporting digital economic strategies and speeding up digital transformation both within and across domains. Data spaces approach a decentralized and dynamic data ecosystem with many-to-many interactions. In short, a data space is the sum of all end points that are able to share data with each other.

According to the [Commission Staff Working Document on Common European Data Spaces](#), "a common European data space brings together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing." This document also defines **key features** of a common European data space, namely:



- **Federated data architecture:** no physical data integration, leave data where it is, which the participation comprises an open number of organizations/ individuals.
- **Interoperability:** no silos, no vendor-dependency. Data that is made available can be reused against compensation, including remuneration, or for free.
- **Data sovereignty and traceability:** a clear and practical structure for access to and use of data in a fair, transparent, proportionate and/non-discriminatory manner and clear and trustworthy data governance mechanisms. Data owners will have the possibility, in the data space, to grant access to or to share certain personal or non-personal data under their control.
- **Trusted participants:** a secure and privacy-preserving infrastructure to pool, access, share, process and use data, in particular personal data protection, consumer protection legislation and competition law, are fully respected

# ホスト組織 = Digital Connect Society (DCS)



Dr. Lee Wah Pheng  
President of IDSa Malaysia Hub  
President of Digital Connect Society



Dr. Lim Yee Mei  
Secretary of IDSA Malaysia Hub  
Secretary of Digital Connect Society



## 2-4 GAIA-X Luxemburg Hub

# **PART 3**

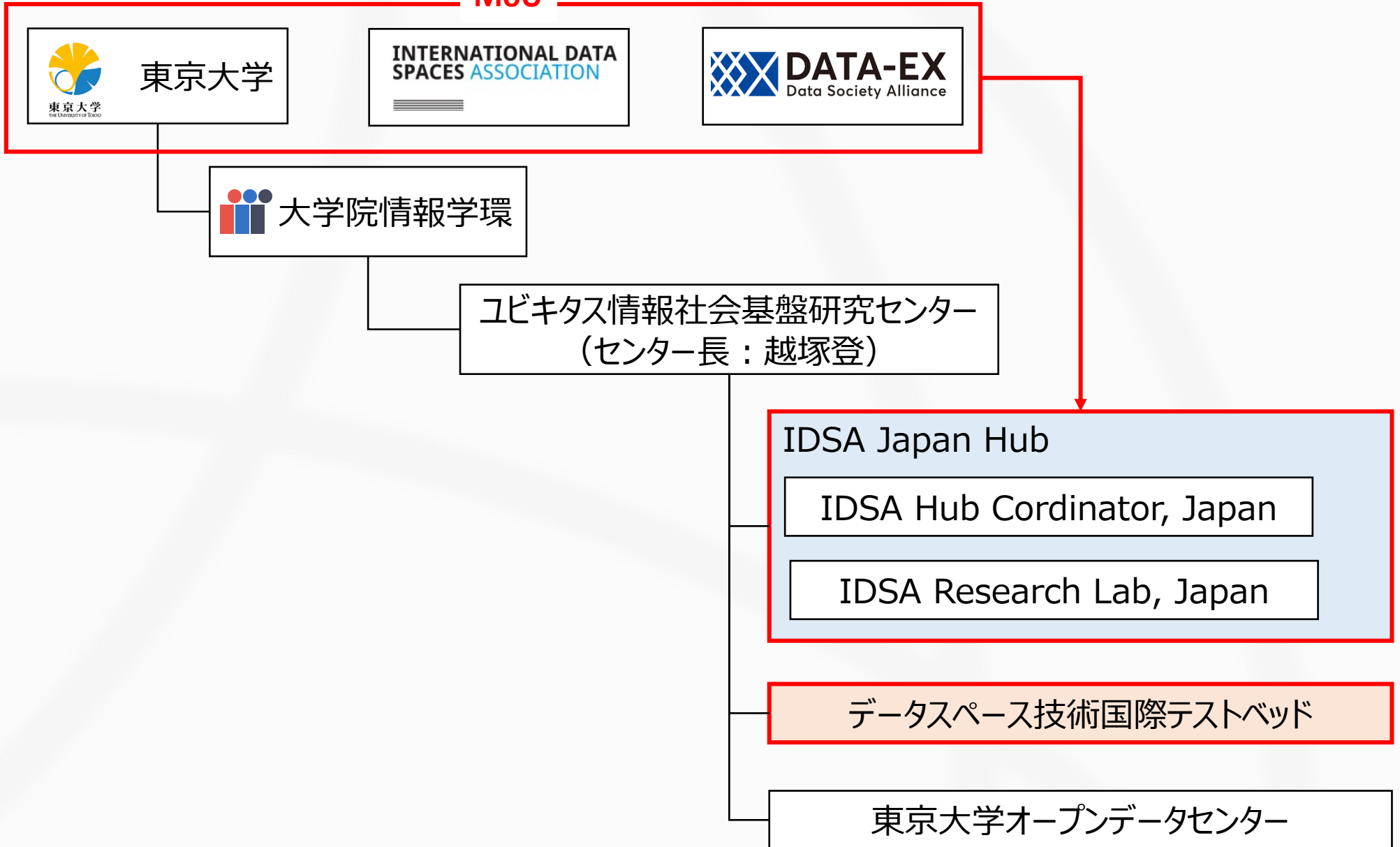
## **データスペース技術国際テストベッド**

### **International Dataspace Technology Testbed**

**東京大学**

# IDSA Japan Hub

MoU



# データスペース技術国際テストベッド：概要

## ■ 背景

- ▶ データスペースを構築するために必要な分散システムは、巨大システムとなっている。
- ▶ 個社・個人（企業、大学、研究所、学生、技術者...）がこれらをテストすることは、大きな困難が生じている。
- ▶ 「今すぐ、データスペースを触りたい」、と言われても触れない。

## ■ 活動内容

- ▶ データスペース技術のオープンなテストベッドを構築する。
  - ◆ テスト環境は、東京大学の計算機システム環境内に構築
  - ◆ テスト環境は、自由に使える実験環境という位置付け
  - ◆ テスト環境を使って基盤システムの使い方、実装の仕方、管理・運用の方法の学習
  - ◆ データスペース基盤／データ連携基盤関連技術の研究開発
  - ◆ 様々なデータスペース基盤／データ連携基盤関連技術間の相互運用性の研究開発
  - ◆ DATA-EXの開発が完了すれば（現在進行中）、DATA-EXの新しいコンポーネントも稼働させたい
- ▶ データスペース技術の「技術コミュニティ」の形成
  - ◆ オンラインマガジンの発行
  - ◆ トレーニングコースや、ハンズオンワークショップの開催
  - ◆ 海外のデータスペース技術コミュニティとの共創活動
  - ◆ 最終的には、「データ学会」を設立

# 3-1 CADDE 4.0 OSSとAPI仕様

githubに

# CADDE: OSS for Dataspaces in Federated Architecture

<https://github.com/CADDE-sip>



Product Solutions Open Source Pricing

Search or jump to...

Sign in

Sign up



## CADDE

Overview Repositories 5 Projects Packages People 1

### Popular repositories

**connector**

Public

2023年3月版分野間連携基盤コネクタ(CADDE)

Python ☆ 7 🍴 4

**cdl-front-server**

Public

2023年3月版来歴管理機能

Java ☆ 2 🍴 1

**cdl-chaincode-go**

Public

2023年3月版来歴管理機能(ブロックチェーンプログラム)

Go ☆ 1 🍴 1

**documents**

Public

SIP分野間データ連携基盤のドキュメント

🍴 1

**catalog\_tool**

Public

2023年3月版カタログ作成ツール

HTML 🍴 1

### People



### Top languages

Python Java Go HTML



# External Specifications of CADDE 4.0

<https://github.com/CADDE-sip/documents/tree/master/doc/4>

Product Solutions Open Source Pricing

Sign in Sign up

CADDE-sip / documents Public
Notifications Fork 1 Star 0

Code Issues Pull requests Actions Projects Security Insights

**Files**

master

Go to file

- > contacts
- > doc
  - > 1
  - > 2
  - > 3
  - > 4
    - > 50\_V4\_外部仕様
      - readme.md
    - > 5
    - > 6
    - > 7
    - > 8
      - CADDE\_v4\_document\_list\_20...
      - readme.md
      - LICENSE
      - README.md

documents / doc / 4 /

CADDE-sip2 Hotfix 20230331 (#1) ab39183 · 10 months ago History

Name	Last commit message	Last commit date
..		
50_V4_外部仕様	Hotfix 20230331 (#1)	10 months ago
readme.md	Update readme.md	10 months ago

readme.md

## 4. External Specification(APIs)

CADDEの機能とサービスの外部仕様であるAPI仕様（機能やサービスのネットワークを経由して利用するインタフェース定義）を定義した資料です。

### 著作者及びライセンスについて

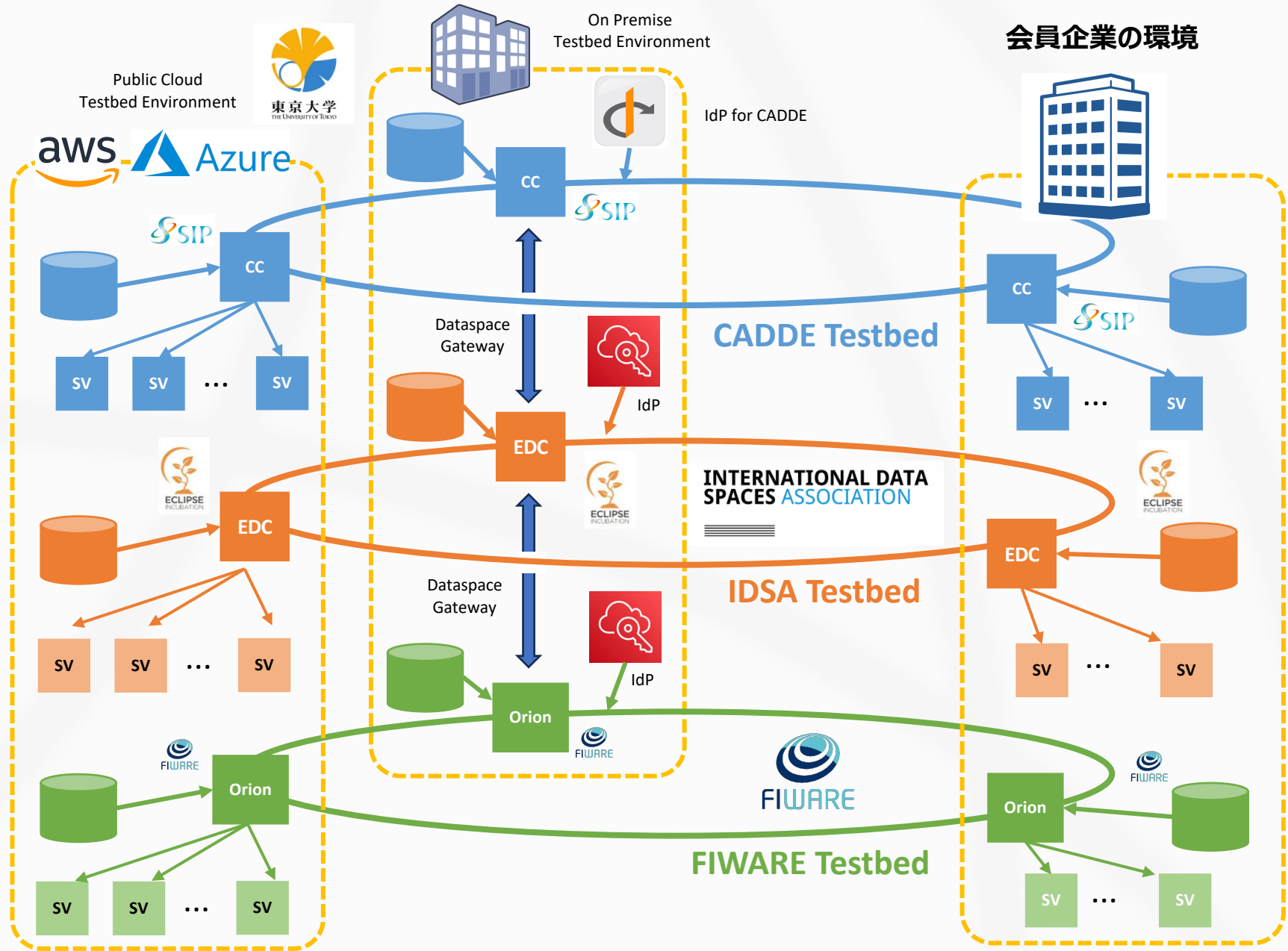
本ディレクトリ(documents/doc/4/)に格納されている資料の著作者は、SIP分野間データ連携基盤技術社会実装コンソーシアム（日本電気株式会社、エブリセンスジャパン株式会社、株式会社日立製作所、大学共同利用機関法人 情報・システム研究機構 国立情報学研究所、株式会社ザイナス）です。

[クリエイティブ・コモンズ 表示 4.0 国際 ライセンス](#)(CC BY 4.0)で提供します。



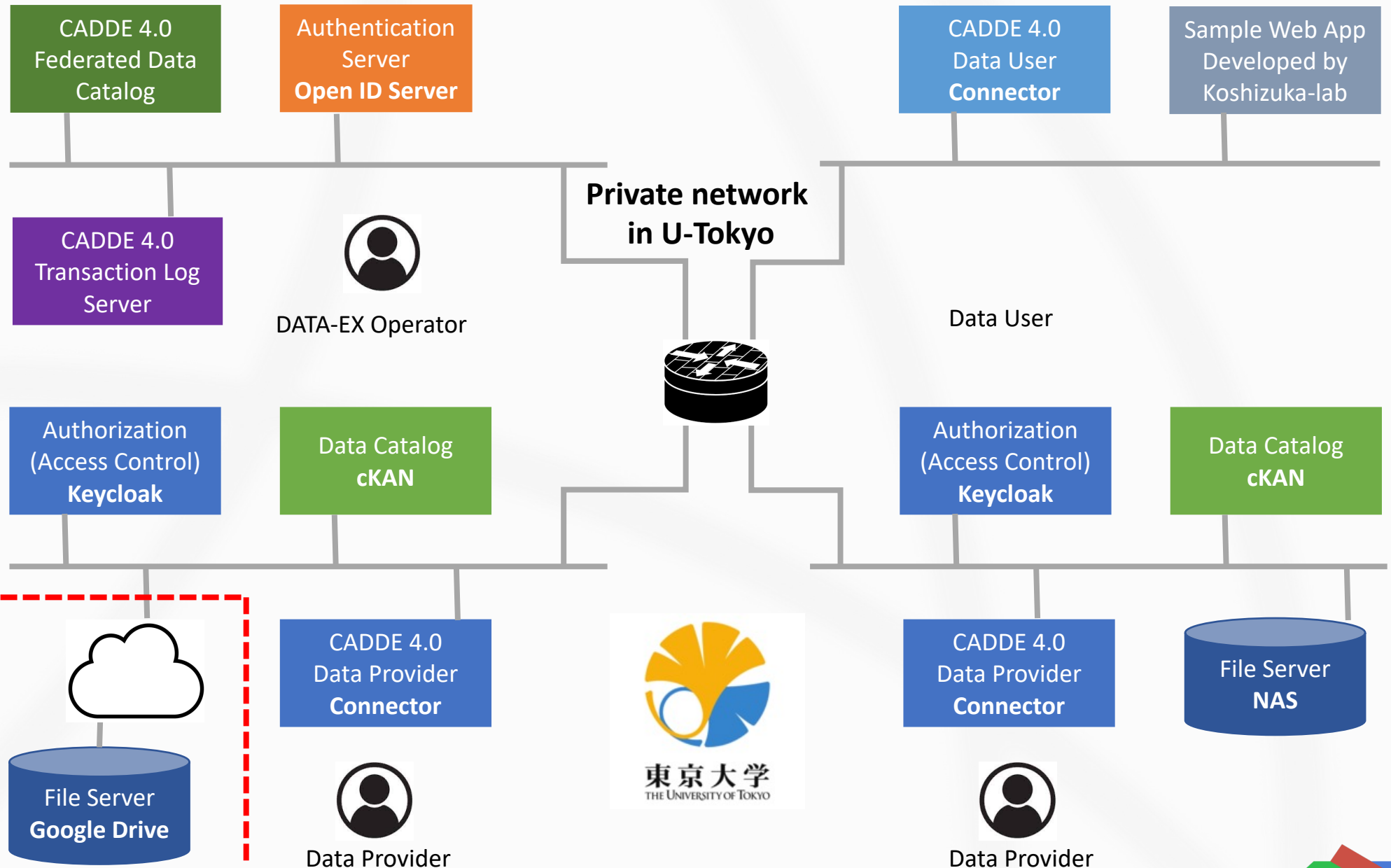
## 3-2 テストベッドのシステム構成

# データスペース技術国際テストベッドの構成図





# CADDE 4.0 Testbed in ITDT at U-Tokyo



A vertical bar on the left side of the slide, divided into a blue upper section and a dark blue lower section.

# 3-3 UT-CADDE Deployment

# CADDE 4.0 in International Dataspace Technology Testbed

# UT-CADDE

# UT-CADDE

### Search data

# Data Download

# UT-CADDE

### Search data

Distribution Title	Provider	Data Type	Last Updated Time	Description	
hello dataspace	taniguchi	TXT	2023-02-28	hello data spaceと書いてあるtxtへのリンク	<input type="button" value="Download"/>
shakespeare	taniguchi	JSON	2023-03-02	https://www.umayadia.com/Note/Note028WebAPISample.htm	<input type="button" value="Download"/>



## 3-4 テストベッドから見えてくる世界観

## 使ってみて感じるメリット

1. データの適正な送受信が、標準化される（ユーザインタフェース上も）
  - ▶ CADDE/DATA-EX特有の機能がなかったとしても、
2. データを配信するときに、アカウント作成やパスワードの作成、パスワードの送付などが不要
  - ▶ 他社のデータ交換用サーバーに、データを送る必要もない。



3. 1:1（相対型）、1:N（ブロードキャスト型）の両方の配信パターンが可能
4. （来歴管理実装すれば）データの送受信の記録が残る
5. （DDP IDをきちんと設計すれば）DDP IDから、データの物理的ロケーションが隠蔽される
  - ▶ セキュリティー上の効果
  - ▶ オープンデータであれば、他のノードにキャッシュしてアクセス効率高めることも可能



A vertical bar on the left side of the slide, consisting of a blue upper section and a dark blue lower section.

# 3-5 グローバル企業が社内で使うFederation

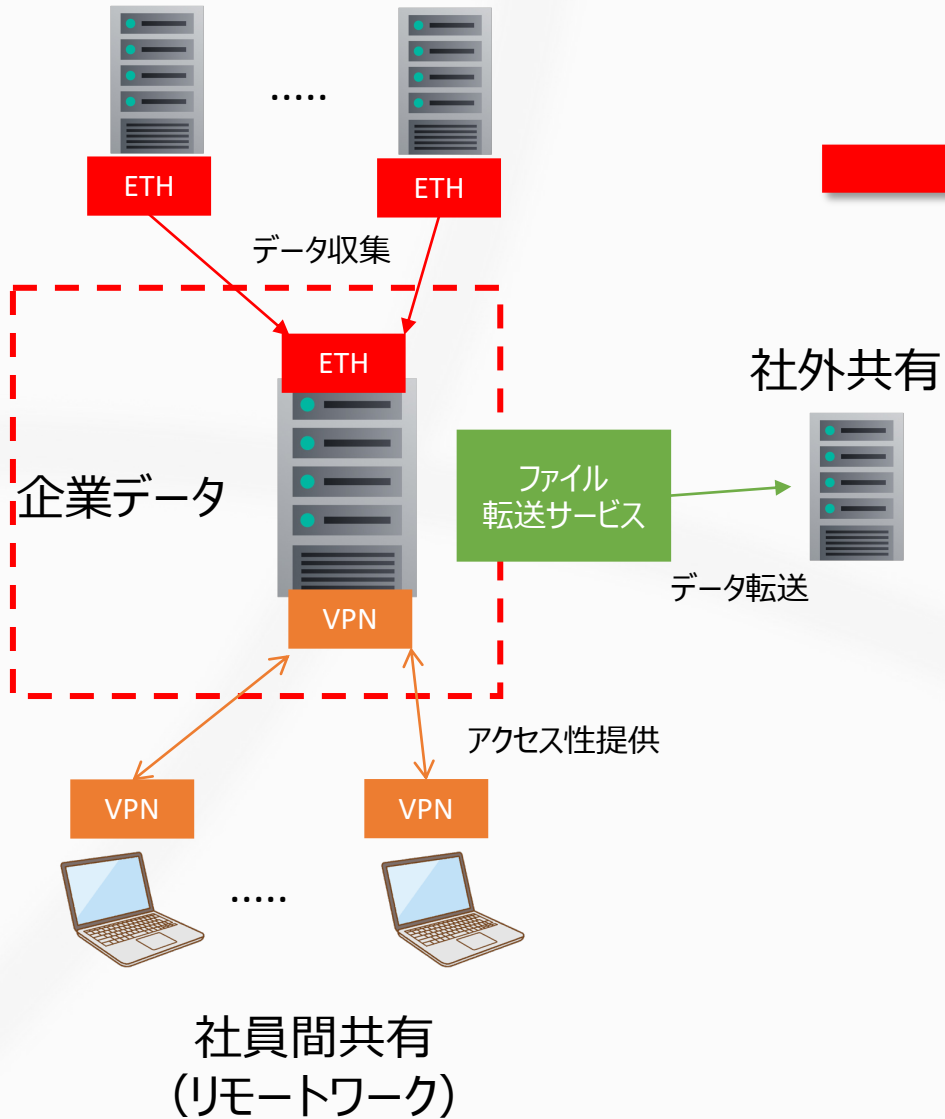
## 社外データ共有方式のそれぞれと課題

方式	サービス例	課題
Public Cloud File Servers	Google Drive, One Drive, Box,	当該ドライブ上のユーザアカウント作成が必要 (ユーザアカウント管理とコストの問題)
ファイル転送サービス 利用	クリプト便、...	(共通) 他社転送サービスにデータを置く必要 (パスワード認証方式、いわゆるPPAP) パスワード 送付が必要 (メール認証方式) メールアカウントの運用の信頼 性に基づく
ファイル転送サービス 導入		(パスワード認証方式、いわゆるPPAP) パスワード 送付が必要 (メール認証方式) メールアカウントの運用の信頼 性に基づく

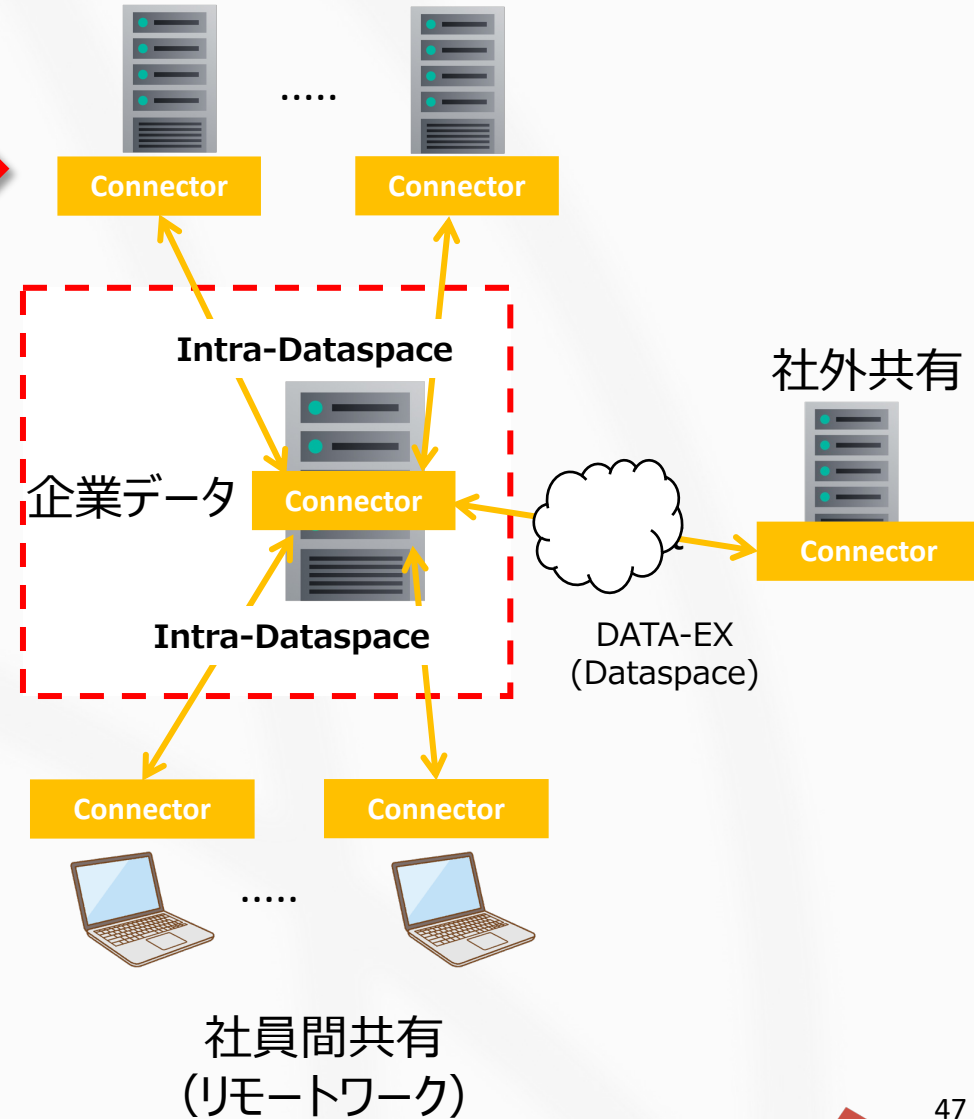
※ 標準化されていないため、これらが、連携社毎に多種多量に発生する

# データスペース後のデータ共有アーキテクチャの世界観

グループ企業間・事業所間共有



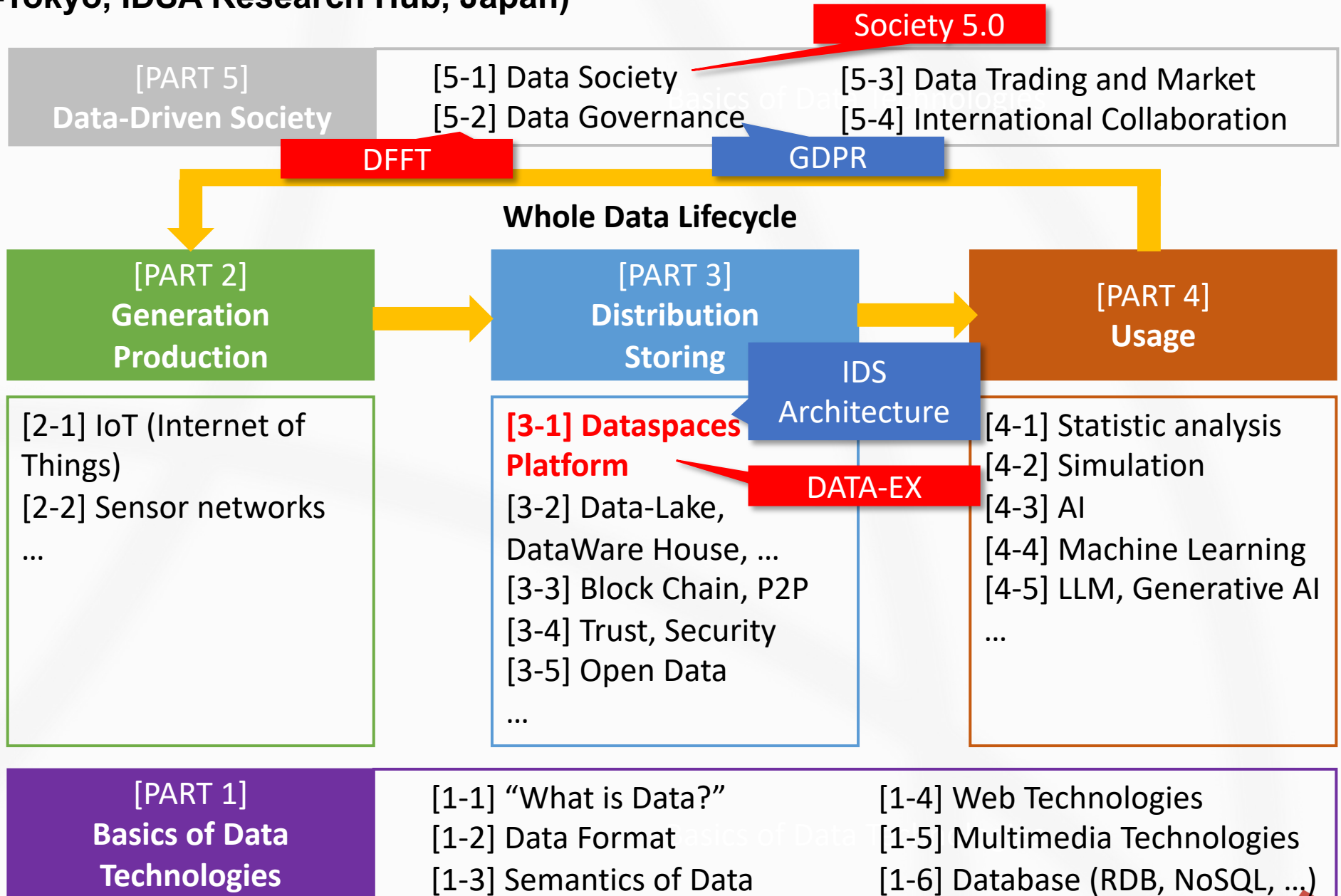
グループ企業間・事業所間共有





# 3-5 人材育成教材プロジェクト

# Data Science Training Program (U-Tokyo, IDSA Research Hub, Japan)



# Example of Web-based “Data Science Training Program”

ホーム » EnPiT Pro 2022 IoT技術者のためのデータ活用論 (東京大学)

## EnPiT Pro 2022 IoT技術者のためのデータ活用論 (東京大学)

### ■科目名

IoT技術者のためのデータ活用論

### ■担当教員・連絡先

越塚登 (noboru@koshizuka-lab.org)  
葛杭麗 (hangli.ge@koshizuka-lab.org)

### ■講義の目的

IoTシステムでは、センサー等から得られたデータを分析・解析を行い、その結果をアクチュエーターなどを通して機器や設備の自動制御に役立てたりする。本講義では、IoTでデータを扱うために必要な基礎知識を習得することを目的とする。

### ■講義の実施方法

オンライン教育システム上で、講義資料と課題を週2回のペースで公開します。受講生の方には、それぞれのペースで講義資料を閲覧しながら課題を進めて頂きます。質問等がある場合には、Slackのチャンネル上に書き込んで頂ければ、担当教員が回答いたします。

### ■教材について

各自のPC等を利用し、講義を受講して下さい。

### ■講義計画 (9月~11月)

- 第1講 (9/17) ガイダンス+データ活用論イントロダクション
- 第2講 (9/21) 実習環境Jupyter
- 第3講 (9/24) 様々な情報 (1) 数値データと文字データ
- 第4講 (9/28) 様々な情報 (2) 画像、音声、動画などのマルチメディアデータ
- 第5講 (10/01) Python入門
- 第6講 (10/05) データの圧縮と暗号化とPython演習
- 第7講 (10/08) Pythonの重要なライブラリ : Numpy, Pandas, matplotlib
- 第8講 (10/12) Pythonで時系列データの扱い
- 第9講 (10/15) Pythonで試すマルチメディアデータ処理
- 第10講 (10/19) Pythonで画像データの取り扱い
- 第11講 (10/22) unicodeとunicode実習
- 第12講 (10/26) Web時代のデータ形式とデータベース論
- 第13講 (10/29) Web形式のデータ形式+データベースの扱い\_Python練習
- 第14講 (11/02) データの統計分析と可視化
- 第15講 (11/05) Open Data 概論と技術
- 第16講 (11/09) IoT、データと法制度

※ 教材は、Google Colaboratoryの講義資料配布ページに、notebook形式で掲載されます。

## Lecture 12 データベース

### 本講の目的

本講では、データベースの概要、特に関係データベースについて学習します。

### 想定履修時間

90分

### 授業アンケートと練習問題の提出

以下のURLにあるGoogle Formにアンケートと練習問題を回答して送ってください。

<https://goo.gl/forms/MJNCE0ddBJE.i0S2>  
<https://goo.gl/forms/MJNCE0ddBJE.i0S2>

### 1. はじめに

データを格納するためには、まずはプログラム言語の変数を用います。変数は、主に限定的に実装されるので、そのプログラム言語で書いたプログラムが動いている間はデータを保持しますが、プログラムが終了すればデータは保持されません(揮発性記憶)。そこで、プログラムが終了してもデータを保持するために、一般的にはファイルがよく使われます(不揮発性記憶)。ただ、このファイルも小さなプログラムならこれで充分ですが、データが大規模で複雑なものになると、様々な方法で、データを検索でき、データの読み書きが可能で、データが正しい形式になっていることを保証する仕組みが必要になります。これがデータベース(database)です。

データベースには、様々な種類のものがあります。現在、一般的に使われており、しっかりとした理論的背景をもっているものが、関係データベース(relational database)です。

関係データベースのコンセプトは、1970年6月にIBMサン・研究所のコード博士が発表した「大規模な共有データバンクのためのリレーショナル・モデル」という論文で提唱されました。当初は、学術的な世界でのみ取り扱われましたが、80年代に入ると、ハードウェアの進化に伴って商用に使える関係データベース用のソフトウェアが徐々に登場してきました。そして1990年代以降になると、UNIX/Windowsといったオープン系テクノロジーが、また最近ではLinux系

のサーバーで動くデータベースが増えてきました。関係データベースを用いて大規模な商用システムや科学ソフトウェアの心臓部が構成されています。現在の商用データベースの約80%はリレーショナル・データベースが占めています。IoTで扱う、大量のセンサーデータなどを扱う基本的な仕組みにすることもできます。

そこで、今回は、この関係データベースの基本概念を紹介します。

### 2. 関係データベース (Relational Database) とは?

関係データベース (Relational Database) は、データを行と列から構成される2次元の表形式で表します。列は各項目を表し、行はデータのエントリー (レコード) を表します。データ同士は複数の表と表の関係によって関連付けられ、SQL (問い合わせ言語) によりユーザーの目的に応じて自由な形式で簡単に操作できます。

#### 2.1 利点

- プログラムとデータの分離
- プログラムとデータの独立性が高いため、データ構造に修正が入ったとしてもプログラムへの影響は極めて小さい
- 柔軟かつ容易なデータの取り出しが可能
- データベース操作の容易化
  - SQLにより、データベースの構築や問合せが簡単になりました。

#### 2.2 関係モデル

関係データベースにおけるデータは表に似た構造で管理されるが、関係 (Relation) と呼ぶ概念でモデル化される。

関係は以下などの要素が必要である。

- 組 (タプル, Tuple) 表における行に相当する
- 属性 (アトリビュート, Attribute) 表における列に相当する
- 定義域 (インスタンス, Instance) またはドメイン (Domain) : データの型 (Type) に相当する
- 候補キー (主キー, Main key)
- 外部キー

具体的に関係 (Relation) は、以下のように構成される。

- 関係は、組の集まりで、組には、見出しの組 (一つ) と本体の組 (一つ以上) から構成される。
- 組には、いくつかの属性から構成される。
- 属性は、属性名と定義域から構成される。(例: 属性A1 = (住所、文字列)、属性A2 = (電話番号、数字列)、など)

The image shows a video lecture interface. The top part features a video player with a man speaking, overlaid with the text "Data is so IMPORTANT that data should be enclosed in a computer" and "Silos of Data". Below the video is a slide titled "System Architecture of DATA-EX". The diagram illustrates the flow of data between three stakeholders (A, B, and C) and various data services. It includes components like Data-EX Platform, Open Vocabulary Registry, Data Catalog (Metadata), and Data Catalog (Mark). It also shows connectors for each stakeholder and services like Resource Discovery, History Management, and Authorization (Access Control List). At the bottom, it lists DATA-EX Co-operated Services: Data Trading Market, Information Bank, and Personal Data Store.

# **PART 4**

## **今後の活動計画**

### **IDSA Japan Hub / 東大テストベッド**

# IDSA Japan Hub／データスペース技術国際テストベッド活動計画（案）

## IDSA Japan Hub

- IDSA技術の勉強会
  - データ及びデータスペース技術に関する人材育成・トレーニング活動
    - ▶ IDSA重要文書の邦訳、教材作成（日本語）、講習会・ハンズオンの実施
    - ▶ 成果は、DSA、IPAはじめ、国内外に提供
  - 共同開発プロジェクトの実施
  - 日本国内のコネクタ実装のIDSAコネクタ登録
  - Data Space Discovery Day Tokyo 2024の開催
  - 国際標準化活動や、海外でのIDSAやデータ関連のイベント参加
- ※ DSA DATA-EX開発TF 国際連携SWGとの連携

## データスペース技術国際テストベッド

- テストベッドの開発（2023年度内にテストベッドの正式提供）
  - ▶ CADDE環境
  - ▶ EDC（Eclipse Data Connector）環境
  - ▶ FIWARE環境
  - ...
- テストベッドを用いた研究プロジェクト（他のテストベッドとの接続等）
- テストベッドを用いたハンズオンセミナーやワークショップの実施



# **PART 5**

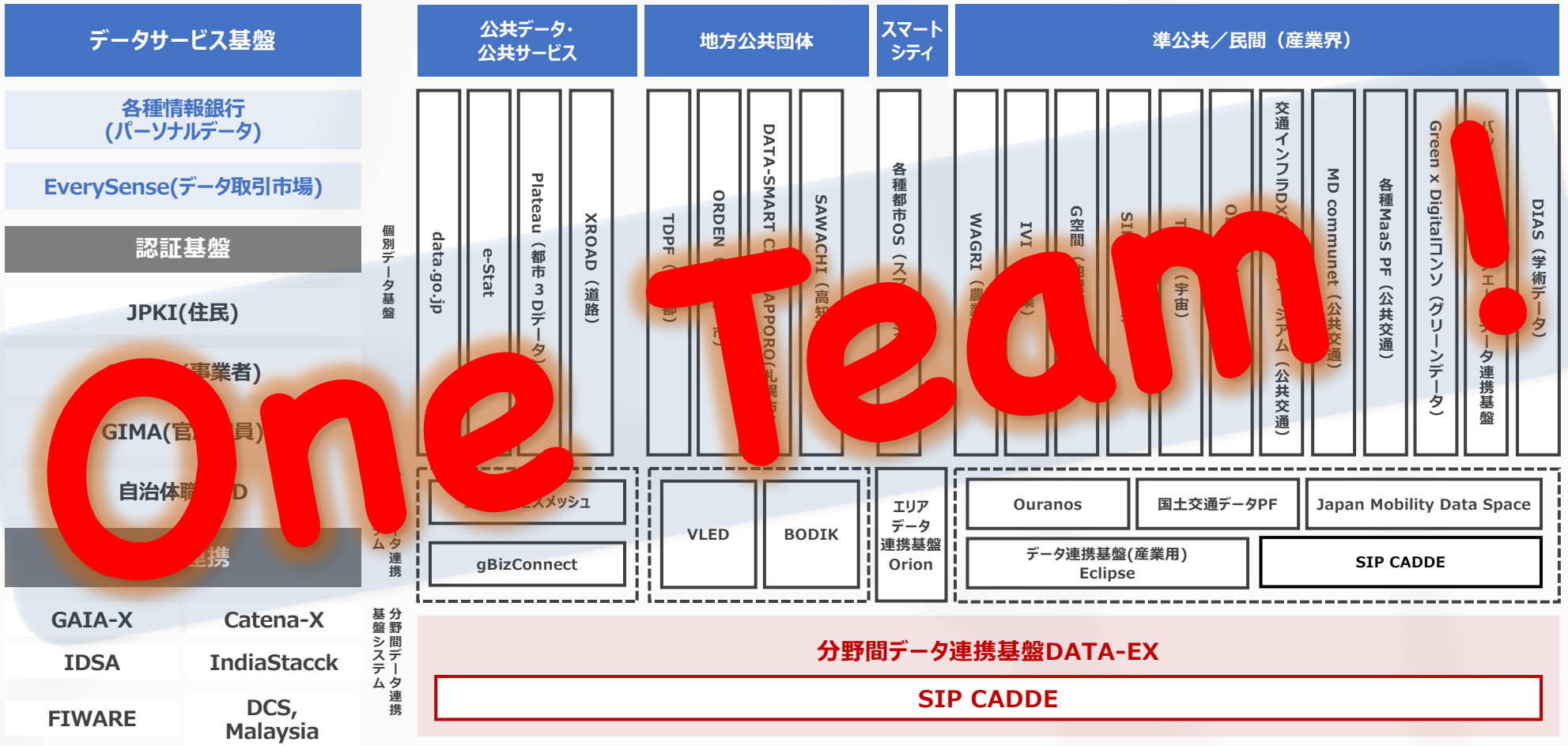
## **Final Message**

我が国のデータ基盤の取り組みは**結構**盛ん

# 過小評価 + データを使っている自覚が無い

ただ、コミュニティが育っていない

# 日本のデータプラットフォーム全体の状況（越塚の個人的なビュー）



***(C) 2023 Koshizuka-Lab.,  
All Rights Reserved.***