

資料-2.

データ連携基盤の整備にむけて 想定される課題について

越塚 登

東京大学 大学院 情報学環・教授

自己紹介

■ 越塚 登 (越塚登)

■ 学位

- ▶ 博士 (理学)
 - ◆ 東京大学大学院理学系研究科情報科学専攻 (1994)

■ 役職

- ▶ 東京大学大学院情報学環・学際情報学府
 - ◆ 学環長・学府長・教授
 - ◆ ユビキタス情報社会基盤研究センター長
 - ◆ 東京大学オープンデータセンター (UTODC) センター長
- ▶ 東京大学教養学部 学際科学科 (兼務)
- ▶ YRPユビキタス・ネットワーキング研究所・副所長

■ 研究分野

- ▶ Computer Science、特に、Embedded Computing、Operating Systems、Network Systems、IoT (Internet of Things)、Smart Environments (House/Building/City)

■ 連絡先

- ▶ E-mail: noboru@koshizuka-lab.org
- ▶ Web: <http://www.koshizuka-lab.org/>





**[Issue 1]
Smart City**

**日本の都市サービスは
世界的にも高品質・高レベル**

スマートシティサービス（1）

- バスロケーションサービス
- オンデマンドバス
- 自動運転バス（実験）
- 経路検索など、交通情報提供
- 乗り合いバス、カーシェア
- 人と荷物共載
- 健康ポイント
- ヘルスケア
- お年寄りの見守りサービス
- 見守りロボット
- スマート救急車
- スマート病院
- 電子カルテ共有
- 地域ポイント／地域通貨
- デジタル化市役所
- プログラミング教育
- デジタルアート
- スマート農業
- スマート漁業
- データ教育
- アイデアソン、ハッカソン、コンテスト
- 自販機のDigital Signage化：災害時の飲料水vending
- インフラ管理

スマートシティサービス（２）

- 観光支援ソフト
- Digital Signage
- 人流解析→マーケティング
- オープンデータカタログ
- 気象情報
- 鳥獣被害防止
- 子供の見守り
- 防災：水位観測、ライブカメラ
- 市の専用アプリ
- チャットボット
- 母子手帳
- お薬手帳
- 駐車場満空情報サービス
- レポーティング（ちばレポ）
- 運転記録、ドライブレコード
- 町工場のIoT化
- 道路情報の取得（レポーティング）
- 防災（ハザードマップ、避難所地図、避難経路...）
- レンタル自転車
- 除雪車
- 無料Wifi
- Smart Lighting
- 施設情報：店舗、混雑、コインロッカー、トイレ、.....

更なるスマート化

その上で更に「スマート」な街とは？

1 都市機能のデジタル化

- ▶ すべての都市サービスがネット上で得られる

2 全体に対する最適化

- ▶ 目の前の課題解決だけでなく、、、
- ▶ 街全体を見た最適化

3 自律的に発展するエコシステム

- ▶ サービスの提供者とサービスの消費者の、両極構造ではなく、、、
- ▶ 皆で一緒になって街をよくしていく
 - ◆ Citizens' Involvement (市民参画)、Civic Tech., Citizens' Science, Prosumer,...

4 ダイバーシティに対する最適化

- ▶ 全員に同じサービスではなく、、、一人一人に違うサービスが提供される
- ▶ 人が街にあわせるのではなく、街が人にあわせる

5 Digital Twins/Mirror World

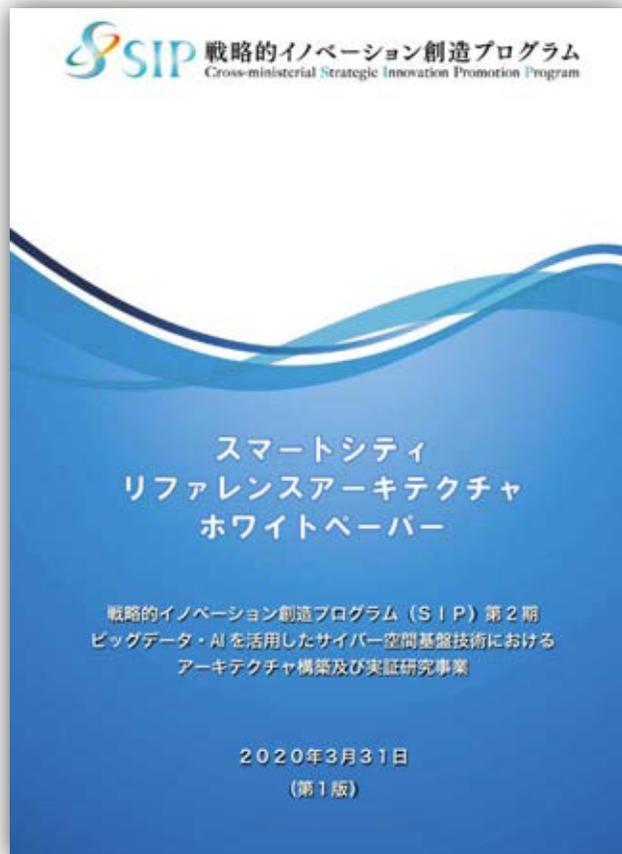
- ▶ 都市サービスは、実空間のサービスとサイバー空間のサービスとを複合して提供

Smart Cityは第2フェーズへ

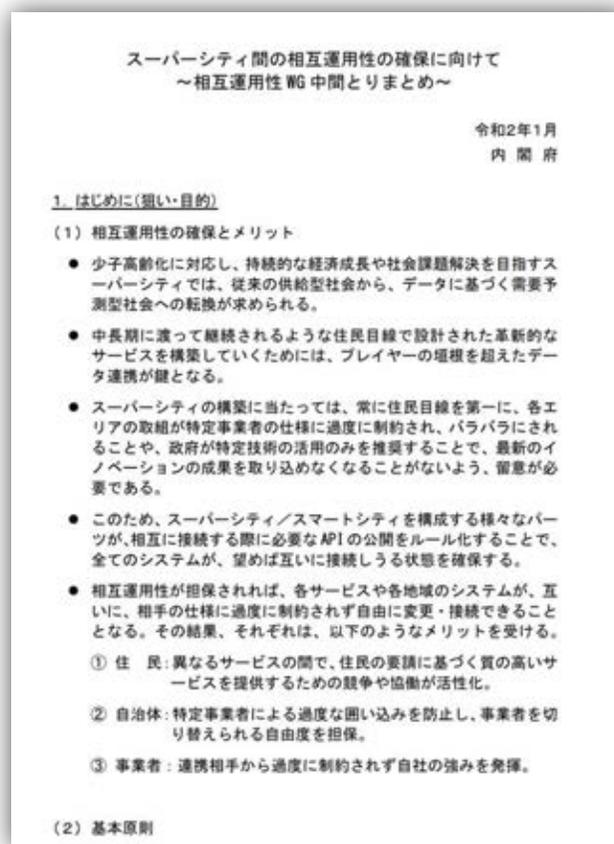
課題は**何**を作るかではなく、**どう**作るか

“What” → “How”

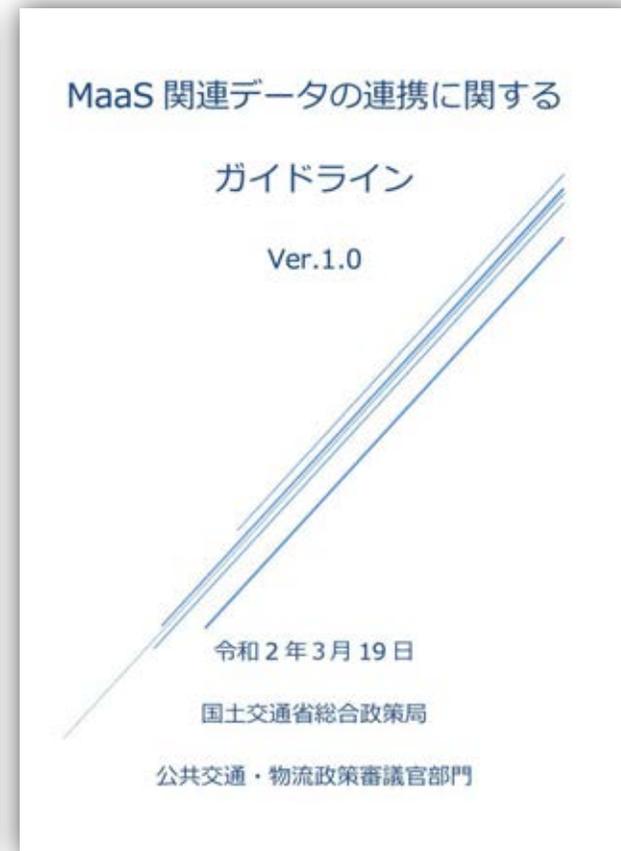
各施策が連携して同じ方向に進むことが大切：共通アーキテクチャ



内閣府SIP Smart City Architecture, 2020 [1]



内閣府Super City Architecture, 2020 [2]



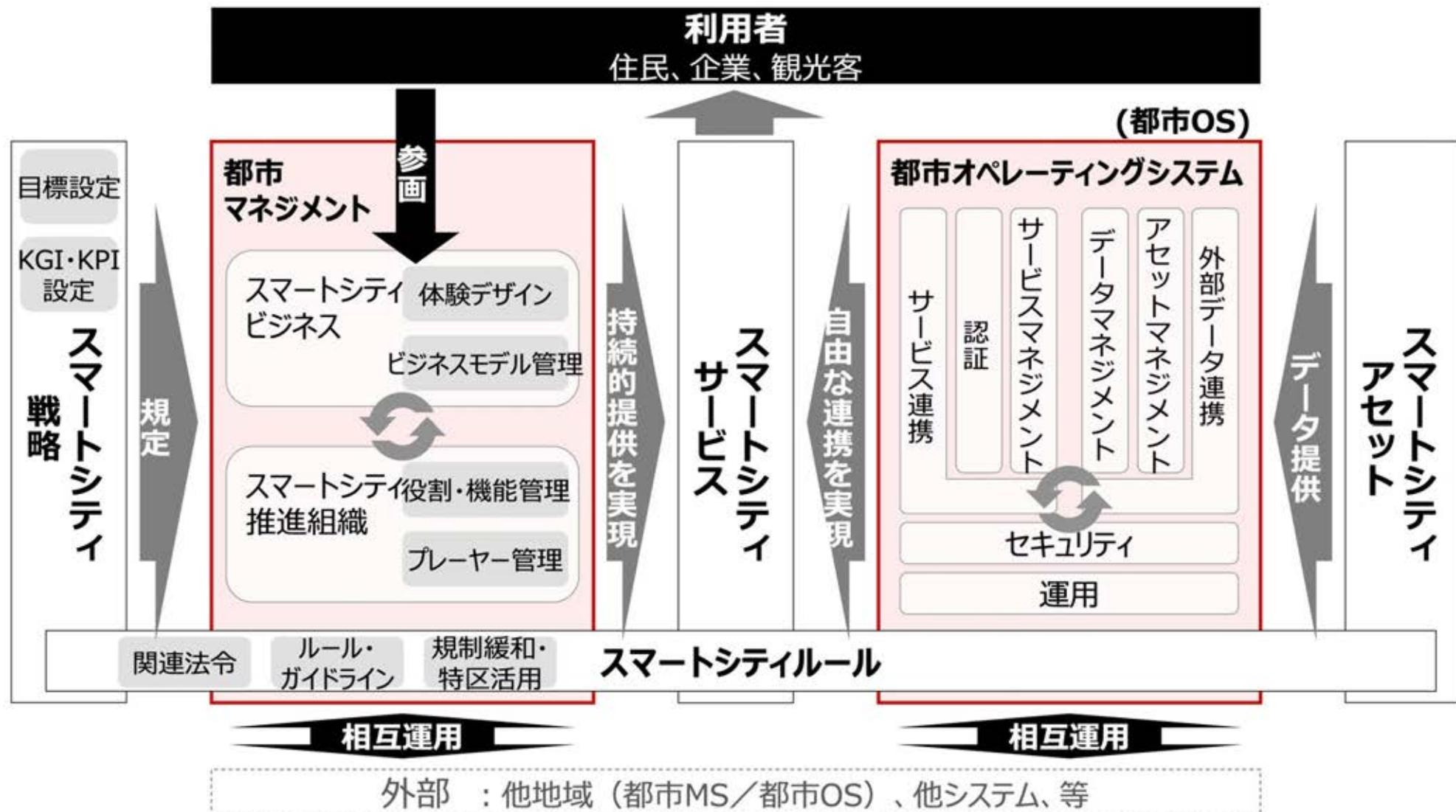
国交省 MaaS Architecture, 2020 [3]

[1] <https://www8.cao.go.jp/cstp/stmain/20200318siparchitecture.html>

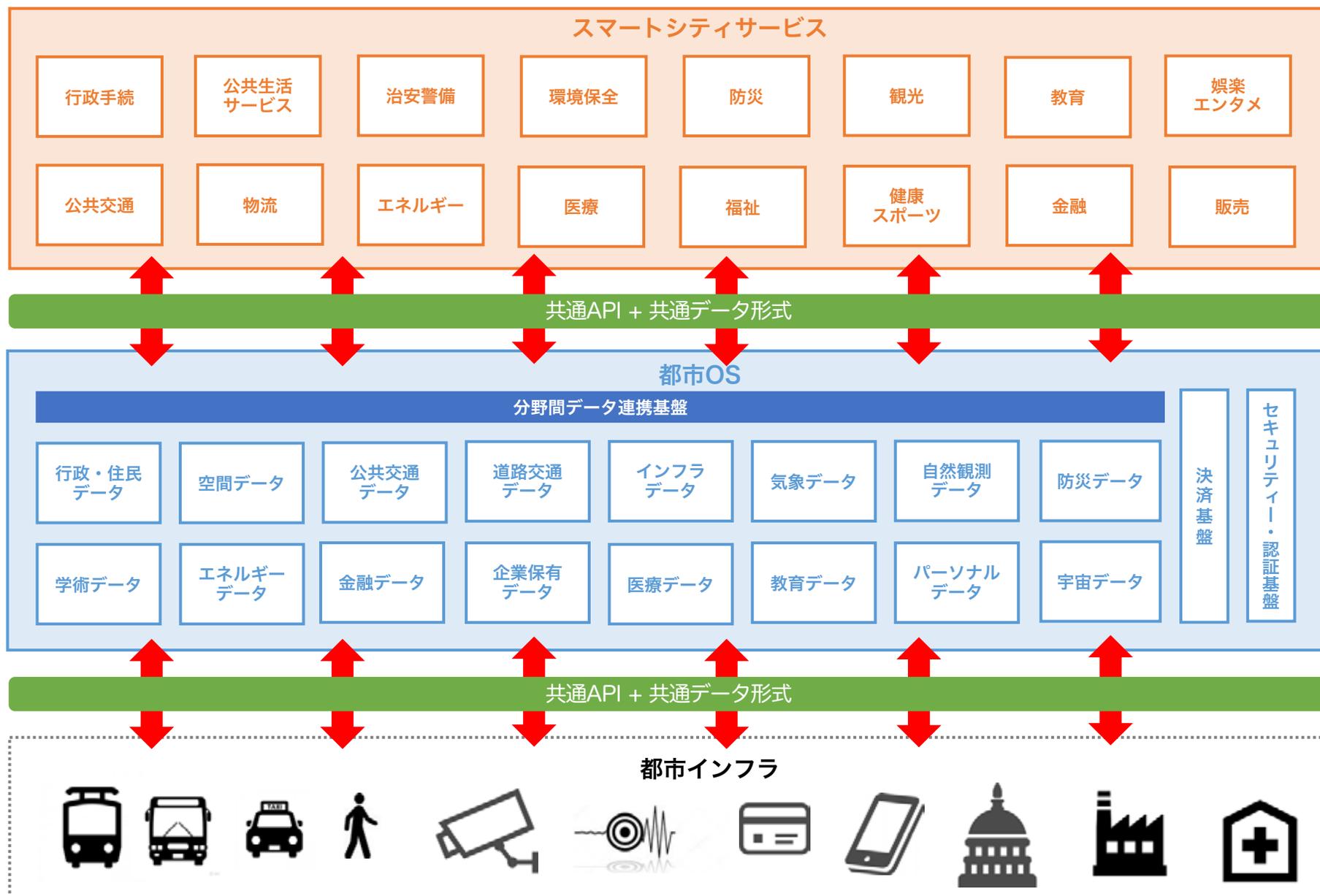
[2] <https://www.kantei.go.jp/jp/singi/tiiki/kokusentoc/supercity/pdf/sogowgchukantorimatome.pdf>

[3] http://www.mlit.go.jp/report/press/sogo12_hh_000181.html

SIP事業によるスマートシティアーキテクチャ



都市OSと分野間データ連携基盤





[Issue 2] データ基盤とは何か？

データ基盤とは何か？

1 データが探せる

国内外、どこにどのようなデータがあるかわかる
(カタログ機能、検索機能)

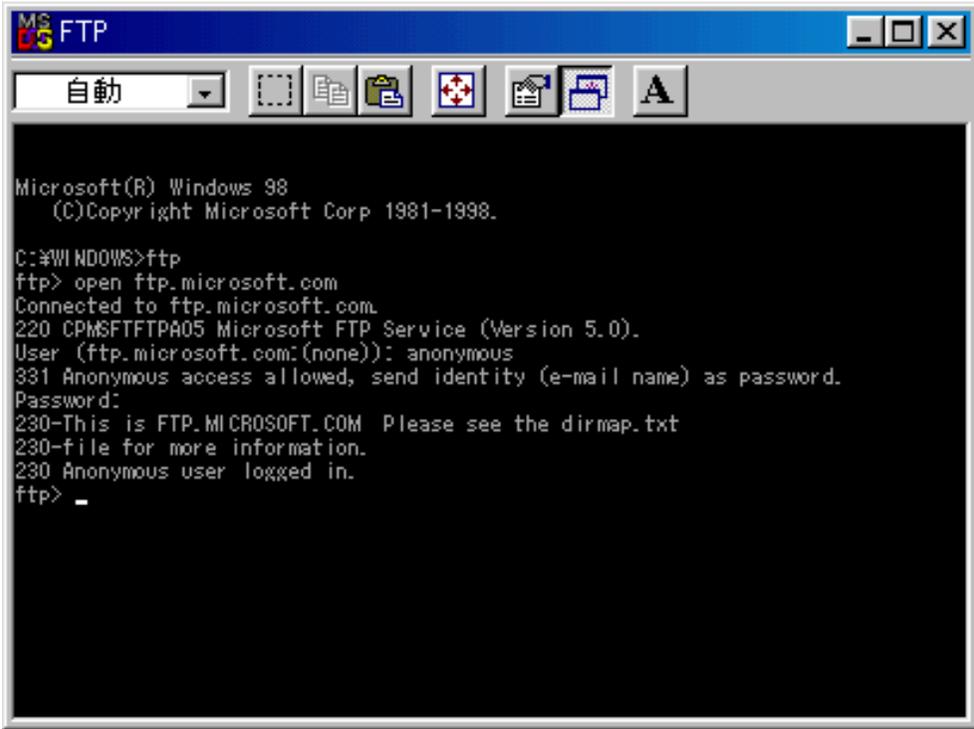
2 探したデータ取得できる

ダウンロード、API取得 (技術的課題)
契約、取引、売買 (社会的課題)

3 取得したデータがつながる

データ連携、インターオペラビリティ、データクレンジング

データ基盤システム（歴史）



```

Microsoft(R) Windows 98
(C) Copyright Microsoft Corp 1981-1998.

C:\WINDOWS>ftp
ftp> open ftp.microsoft.com
Connected to ftp.microsoft.com
220 CPMSFTFTP05 Microsoft FTP Service (Version 5.0).
User (ftp.microsoft.com:(none)): anonymous
331 Anonymous access allowed, send identity (e-mail name) as password.
Password:
230-This is FTP.MICROSOFT.COM Please see the dirmap.txt
230-file for more information.
230 Anonymous user logged in.
ftp> _
  
```

Anonymous FTP



```

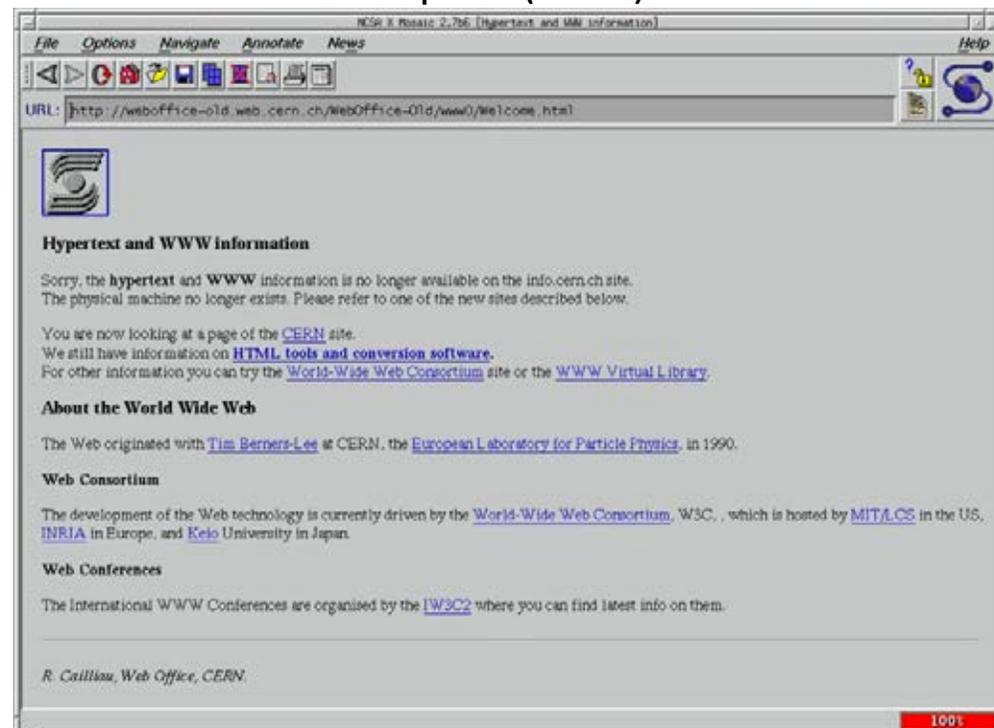
Internet Gopher Information Client v2.1.1-1

Suomen gopherit (finnish)

1. ----- Suomen Gopherit, Finnish Gophers -----
2. Finnish University and Research Network, FUNET/
3. Svenska Handelshgskolan (Hanken)/
4. Teknillinen Korkeakoulu, TKY/
5. Teknillinen Korkeakoulu/
6. Helsingin Kauppakorkeakoulu/
--> 7. Helsingin Yliopisto/
8. Joensuun Yliopisto/
9. Jyvaskylän Yliopisto/
10. Kuopion Yliopisto/
11. Lapin Yliopisto/
12. Lappeenrannan Teknillinen Korkeakoulu/
13. Oulun Teknillinen Oppilaitos/
14. Oulun Yliopisto/
15. Tampereen Teknillinen Korkeakoulu/
16. Tampereen Yliopisto/
17. Turun yliopisto/
18. Valtion Teknillinen Tutkimuskeskus (VTT)/

Press ? for Help, Q to Quit, M to go up a Menu      Using Directory...
  
```

Gopher (1991)

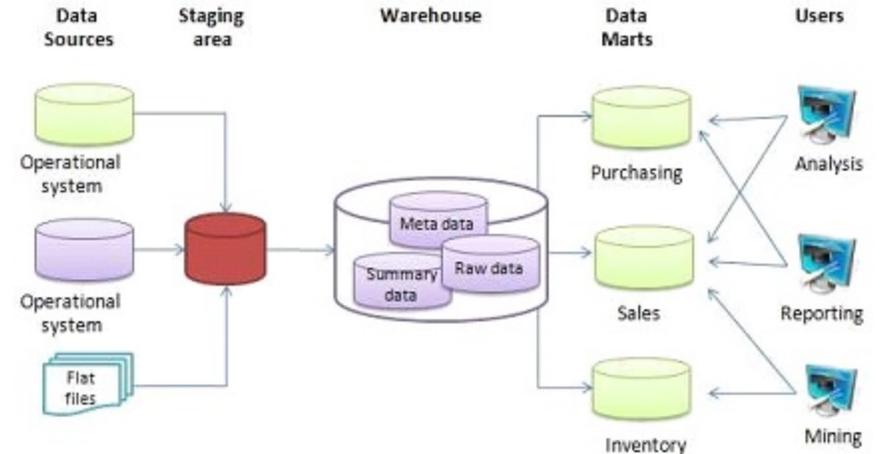


WWW (NCSA Mosaic) (1993)

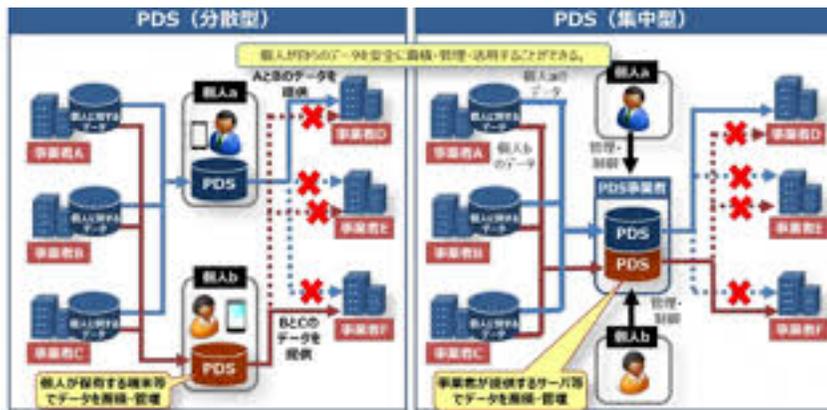
データ基盤システム (近年の例)



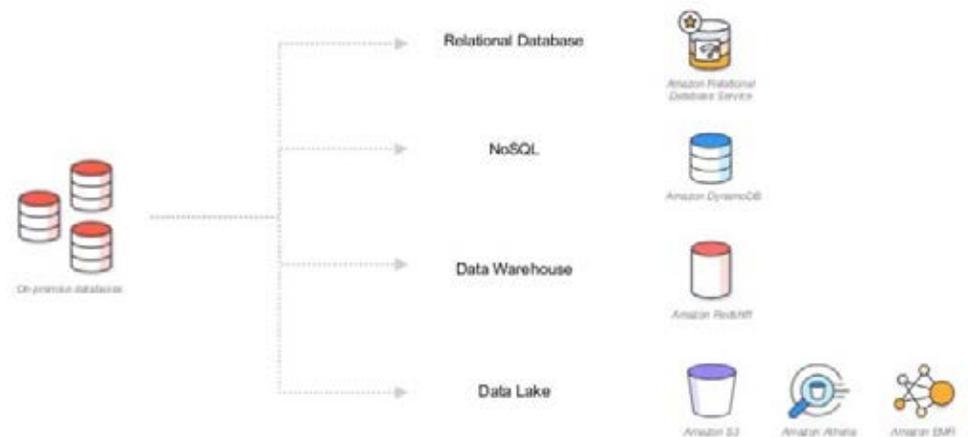
(Open) Data Catalog (cKAN)



Data Warehouse



Personal Data Store(PDS), 情報銀行



Data Lake



[Issue 3]
分野毎データ基盤事例

データ基盤整備の事例（1）



data.go.jp



札幌オープンデータ



odpt.org (公共交通オープンデータセンター)



Personal Data Store (おもてなしクラウド)

データ基盤整備の事例（2）



観光データ基盤 (kokosil)



観光データ基盤 (kokosil)



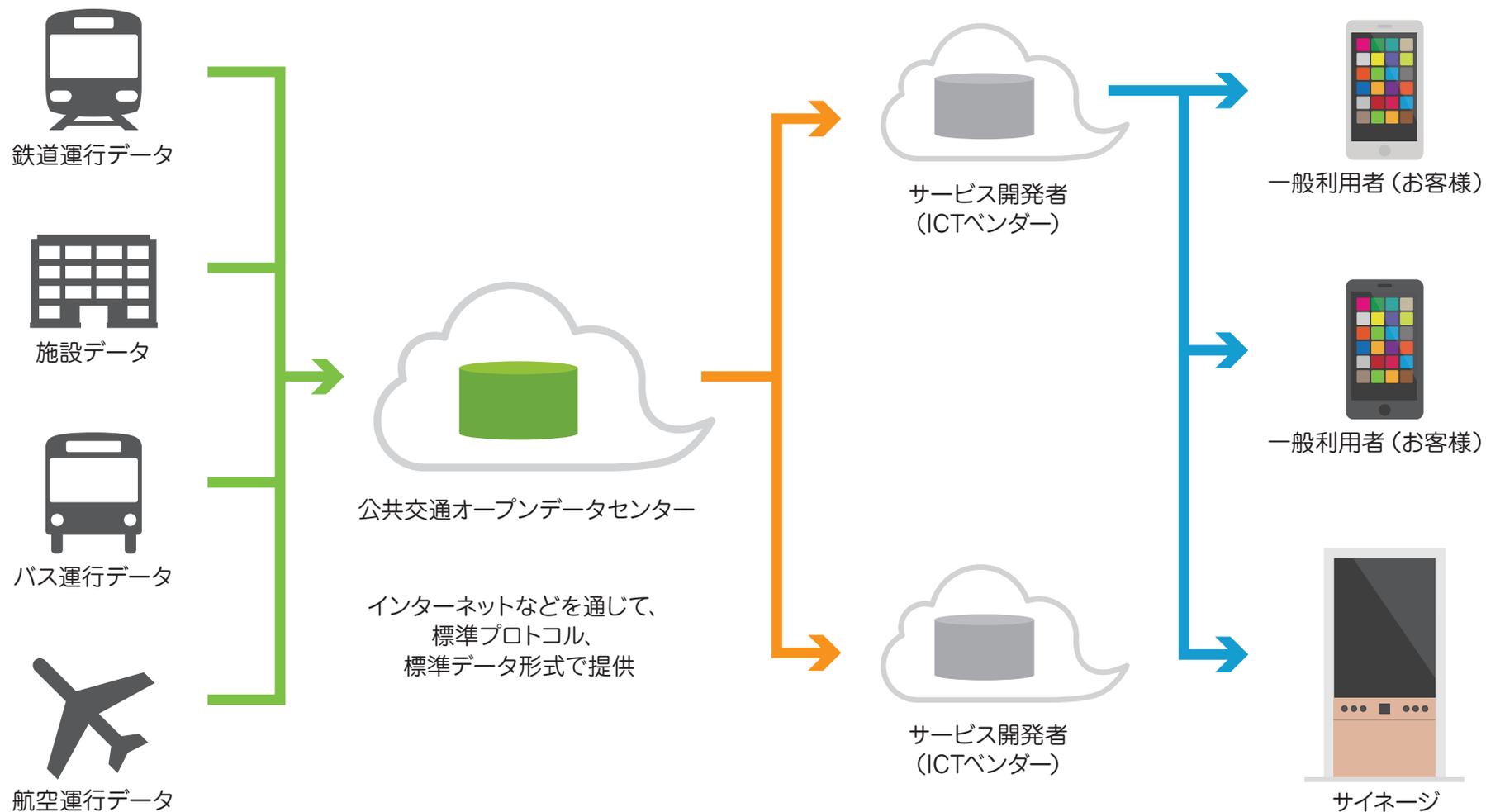
電子母子手帳・電子お薬手帳



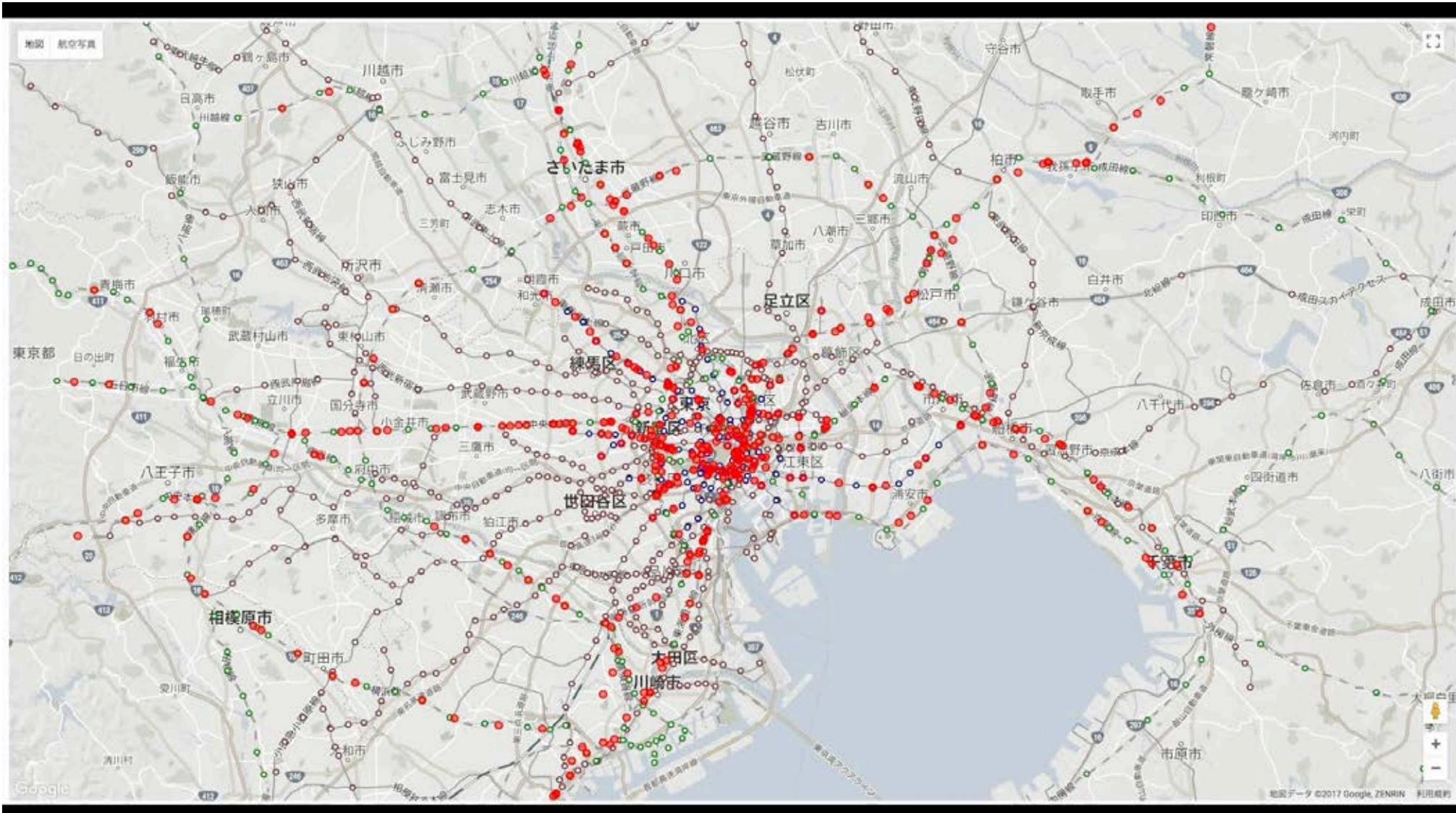
バリアフリーデータ基盤 (自律移動支援プロジェクト)

【交通分野】公共交通オープンデータセンター

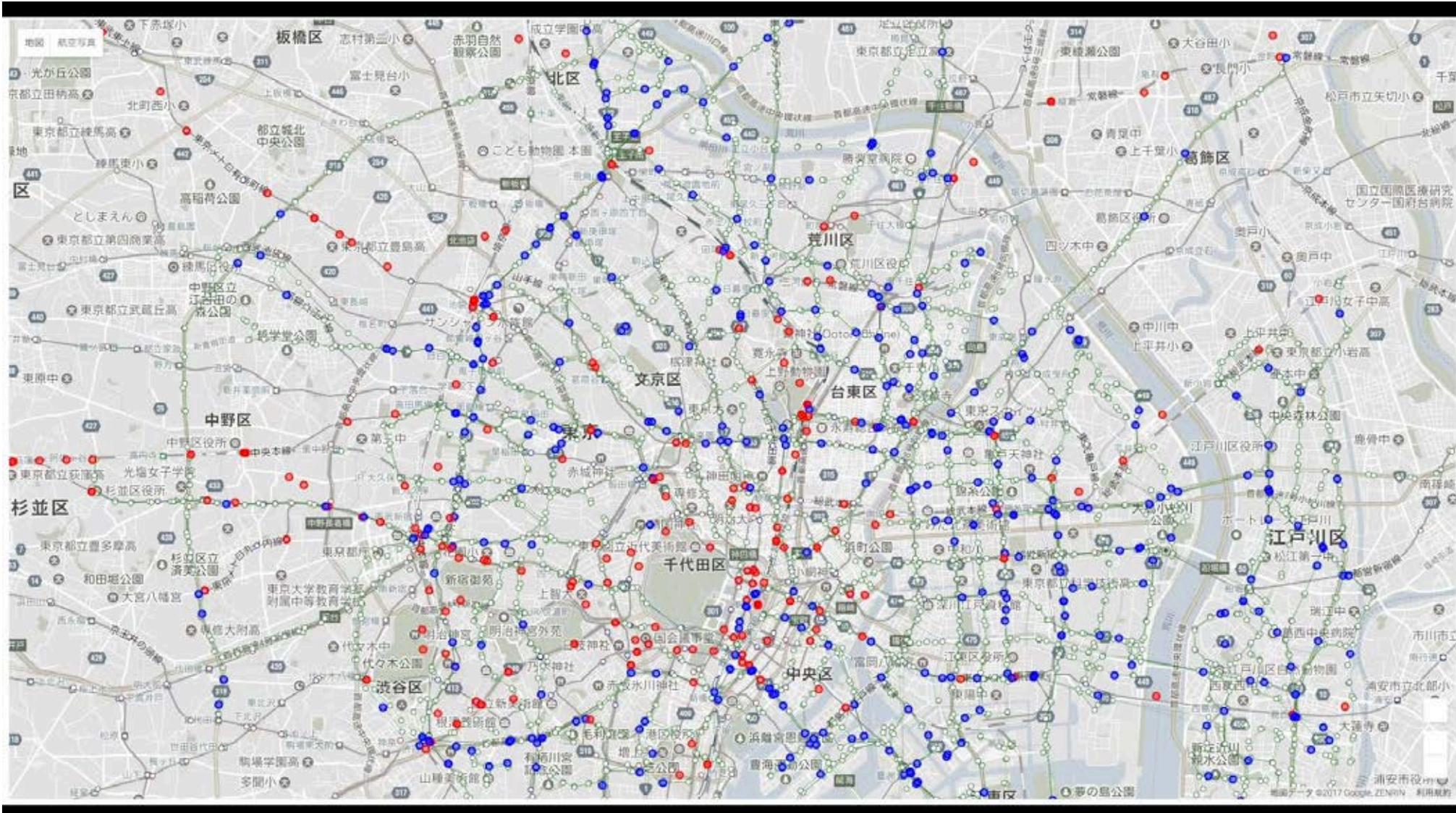
- 鉄道、バス、航空などの交通事業者の提供する公共交通データを集約し、標準化されたデータ形式でサービス開発者であるICTベンダーに提供
 - ▶ ICTベンダーは、公共交通オープンデータセンターから、各交通事業者の指定した条件（有償のケースも想定）でデータを入手し、様々なサービス開発に活用できる



【交通分野】 JR東日本と東京メトロの車両位置の可視化



【交通分野】 JR東日本と東京メトロ、都バスの車両位置の可視化



【インフラ分野】NEDO SIP事業「高度なインフラ・マネジメントを実現する多種多様なデータの処理・蓄積・解析・応用技術の開発」(NEXCO東日本、YRP UNL、東大、他)

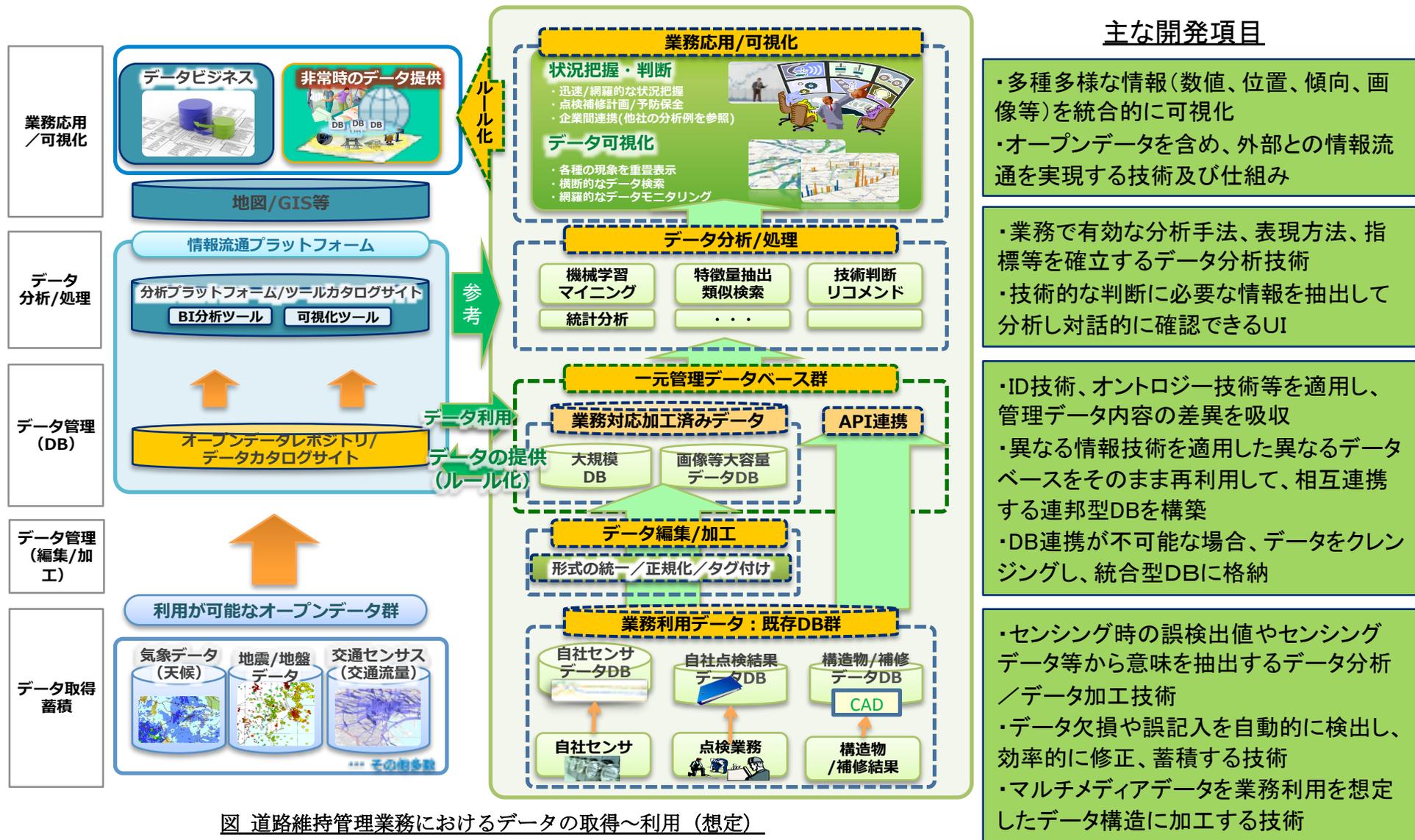


図 道路維持管理業務におけるデータの取得～利用(想定)

【インフラ分野】点検・評価・診断業務におけるマルチスケールモニタリングUI

The screenshot displays a comprehensive multi-scale monitoring interface for infrastructure inspection. It is organized into several key sections:

- Top Panel (Route Overview):** Shows a route map for '上信越自動車道 軽井沢～佐久' with various interchanges (IC) and kilometers (km) marked. Below it is a detailed view of the road with '上り' (up) and '下り' (down) lanes, and specific bridge locations like '下田IC', '下田橋', '神原西橋', '神原東橋', '大島橋', '野上川橋', '富岡TN', and '富岡IC'.
- Navigation & Settings:** Includes buttons for 'KP 指定', '構造物 指定', '表示要素 指定', '縮尺 100%', and '全路線図'.
- Bridge Information (橋梁情報):** A list of bridges with their respective KP values and lengths:
 - PC橋 赤羽根橋 100.0 KP 橋長 52.140 m
 - PC橋 上之原橋 100.0 KP 橋長 45.540 m
 - PC橋 西丸山橋 100.0 KP 橋長 72.500 m
 - PC橋 弥津町屋橋 100.0 KP 橋長 72.000 m
 - PC橋 野行田一号橋 100.0 KP 橋長 31.100 m
 - PC橋 裏山橋 100.0 KP 橋長 31.100 m
 - PC橋 赤羽根橋 100.0 KP 橋長 52.140 m
- Bridge Details (橋梁概要):** A detailed view of the '上之原橋' (Ueno-harabashi) bridge, including a 3D perspective image and technical specifications:
 - 橋名: 上之原橋
 - 構造種別: PC 連続ラーメン2跨
 - 構造階数: 上り線 3, 下り線 2
 - 構造形式: 上り線 275m, 下り線 203m
 - 設計水平面: 上り線 0.1%, 下り線 0.1%
 - 構造形式: PC 橋
 - 構造種別: PC 連続ラーメン2跨
 - 構造階数: 上り線 3, 下り線 2
 - 構造形式: 上り線 275m, 下り線 203m
 - 設計水平面: 上り線 0.1%, 下り線 0.1%
 - 構造形式: PC 橋
 - 構造種別: PC 連続ラーメン2跨
 - 構造階数: 上り線 3, 下り線 2
 - 構造形式: 上り線 275m, 下り線 203m
 - 設計水平面: 上り線 0.1%, 下り線 0.1%
- Diagram Information (図面情報):** A grid of thumbnail images representing various technical drawings such as '数量集結表', '配置図', '下部工・構造一般図', '縦断面図', '各部詳細', '構造一般図', '全体一般図', '橋台補修図', '平面図', '橋台補修図 2', and '橋台補修図 3'.
- Inspection Information (点検情報):** A grid of inspection photos with associated KP values and grades:
 - 84.49 KP 左端昇駐 本体の損傷 (Grade B)
 - 74.665 KP 車道部 横打ち材の損傷 (Grade A1)
 - 69.974 KP 中分帯 付属物の損傷 (Grade AA)
 - 84.49 KP 左端昇駐 本体の損傷 (Grade B)
 - 74.665 KP 車道部 横打ち材の損傷 (Grade A1)
 - 69.974 KP 中分帯 付属物の損傷 (Grade A)
 - 84.49 KP 左端昇駐 本体の損傷 (Grade B)
 - 74.665 KP 車道部 横打ち材の損傷 (Grade AA)
 - 69.974 KP 中分帯 付属物の損傷 (Grade AA)

【防災分野】 災害時避難ナビの画面



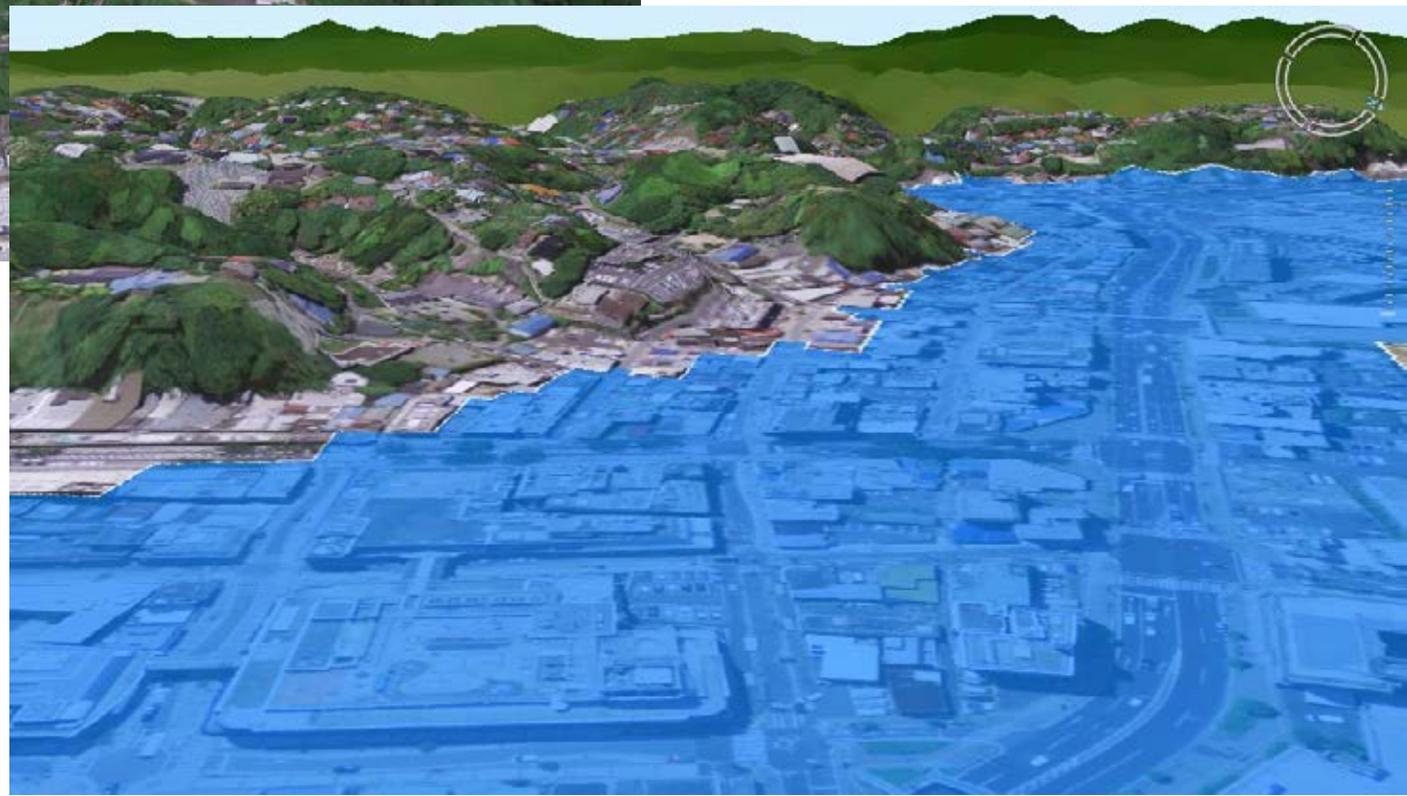
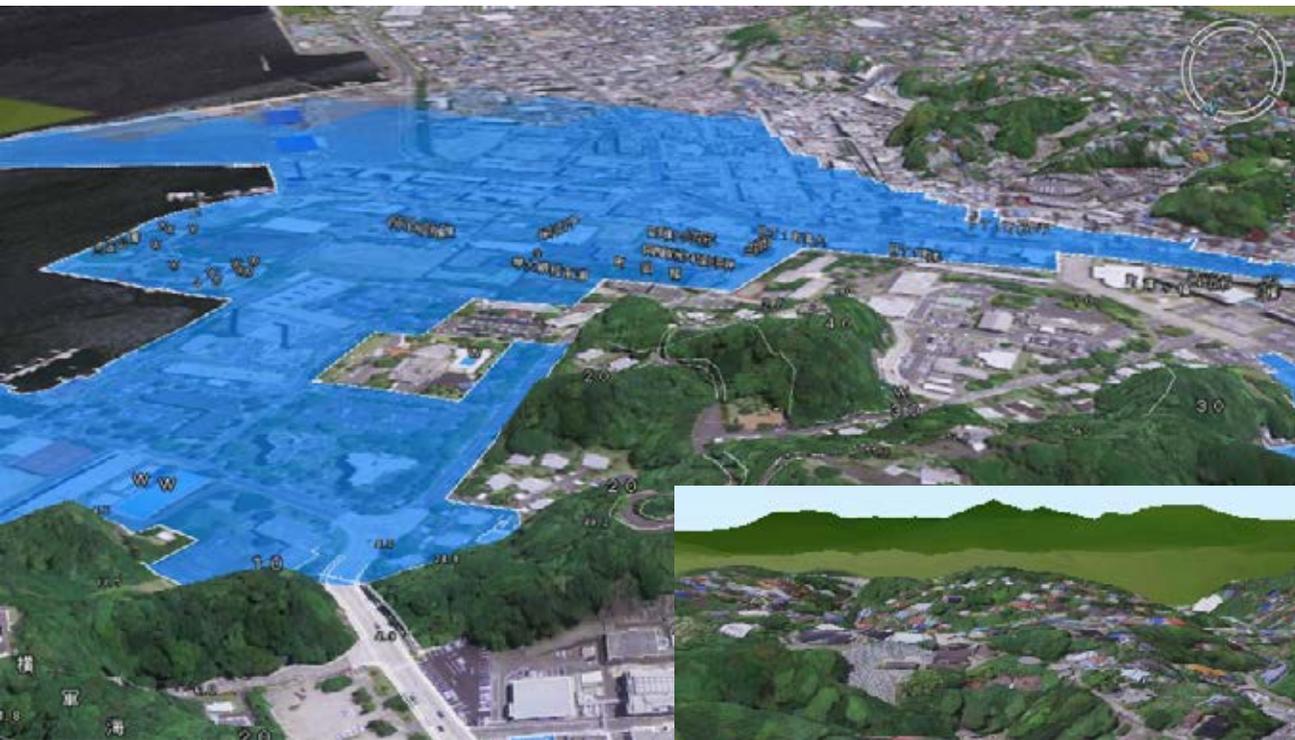
時間がある場合は高台へ誘導

高台の方向



時間がない場合は避難ビルへ誘導

【防災分野】 3次元地図による津波シミュレーション（横須賀）



【福祉分野】 オープンバリアフリーマップ（バリアフリー施設）

オープンバリアフリーマップ

バリアフリー

- AED
- 多目的トイレ
- トイレ
- エレベーター
- バリアフリー設備

注意

- 階段
- 段差
- 狭い道
- 勾配
- 傾斜

絞り込みキーワード

空白で区切って複数指定できます

検索オプション

- 表示中の画面内でのみ検索する

ページあたりの表示件数: 500

検索結果

2 pages

1 2

- 立野公園便所
- 鴨居3丁目公園便所
- 岩戸3丁目公園便所
- 旗山崎公園便所

猿島公園便所

横浜旗山猿島1番
だれでもトイレです。

Copyright © 2014 YIN Ubiquitous Networking Laboratory. All rights reserved.

【防災分野】 災害時対応機能を備えた電子母子手帳（東京大学、日本医科大学）

平時

被災時



電子母子手帳

電子母子手帳

電子母子手帳

データ入力

データ入力

データ入力



手帳情報の確認



担当医師



行政の防災部署

災害でスマホ損壊・通信不具合あっても
災害直前までのスマホの通信ログで
被災時の状態を推定

【福祉分野】 オープンバリアフリーマップ（段差情報）

オープンバリアフリーマップ

検索条件

バリアフリー

- AED
- 多目的トイレ
- トイレ
- エレベーター
- バリアフリー設備

注意

- 階段
- 段差
- 狭い道
- 勾配
- 傾斜

絞り込みキーワード

空白で区切って複数指定できます

検索オプション

- 表示中の画面内でのみ検索する

ページあたりの表示件数 500

検索 リセット

検索結果

2 pages

< 1 2 >

• 狭い道/勾配

狭い道/傾斜

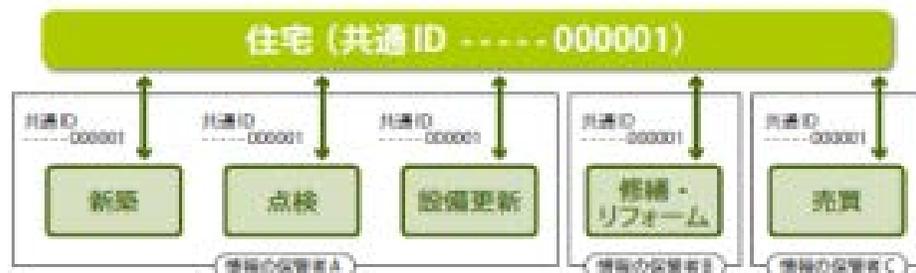
- 道幅 1~1.5m
- 道の横き 3%以上

地図データ ©2015 Google, ZENRIN 利用規約
利用規約 | プライバシーポリシー
Copyright © 2014 YRP Ubiquitous Networking Laboratory. All rights reserved.

【福祉分野】 バリアフリートイレの情報提供：トイレの状況がわかる映像



【住宅分野】 住宅履歴管理「いえがるて」



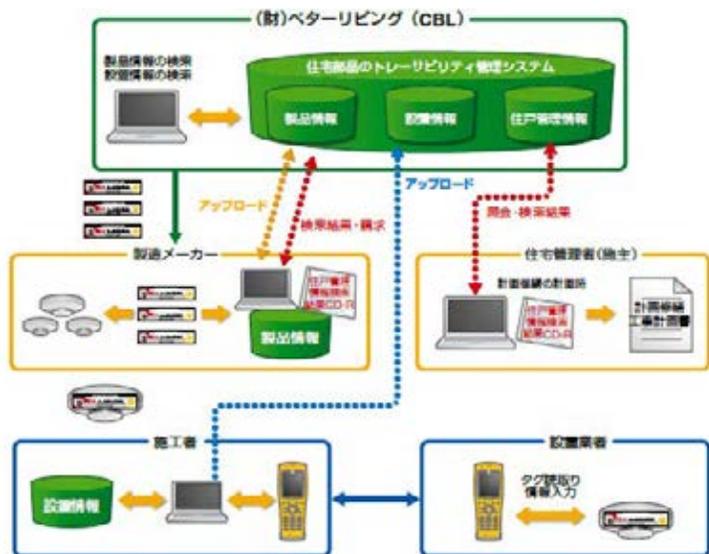
■ 概要

- ▶ 住宅の長期にわたる設備更新、修繕・リフォーム、売買などの局面で、その住宅の過去の諸情報（住宅履歴情報）を記録・活用する
- ▶ 個々の住宅を唯一識別すると共に個人情報やプライバシーを保護する目的で「共通ID」と呼ばれる個体識別番号（ucode）の配布
- ▶ 135 万戸分のucodeを発行済

■ 事業主体

- ▶ 一般社団法人住宅履歴情報蓄積・活用推進協議会

【住宅分野】 住宅備品のトレーサビリティ／ベターリビング



■ 概要

- ▶ メーカーが認定部品を販売する際、住宅部品ひとつひとつにucodeタグ（電子タグ）付きの証紙が貼られて出荷。
- ▶ 各認定部品の個品がどの家屋に設置されたかの管理を少ない手間で実施可能
 - ◆ 既に、トラブル時の回収を効率化する実績
- ▶ 設置業者は、RFID上のucodeと設置・保全・廃棄等の作業情報を自動的に紐付けし、ベターリビングが運営するサーバに送信・登録し、トレーサビリティ情報を管理。

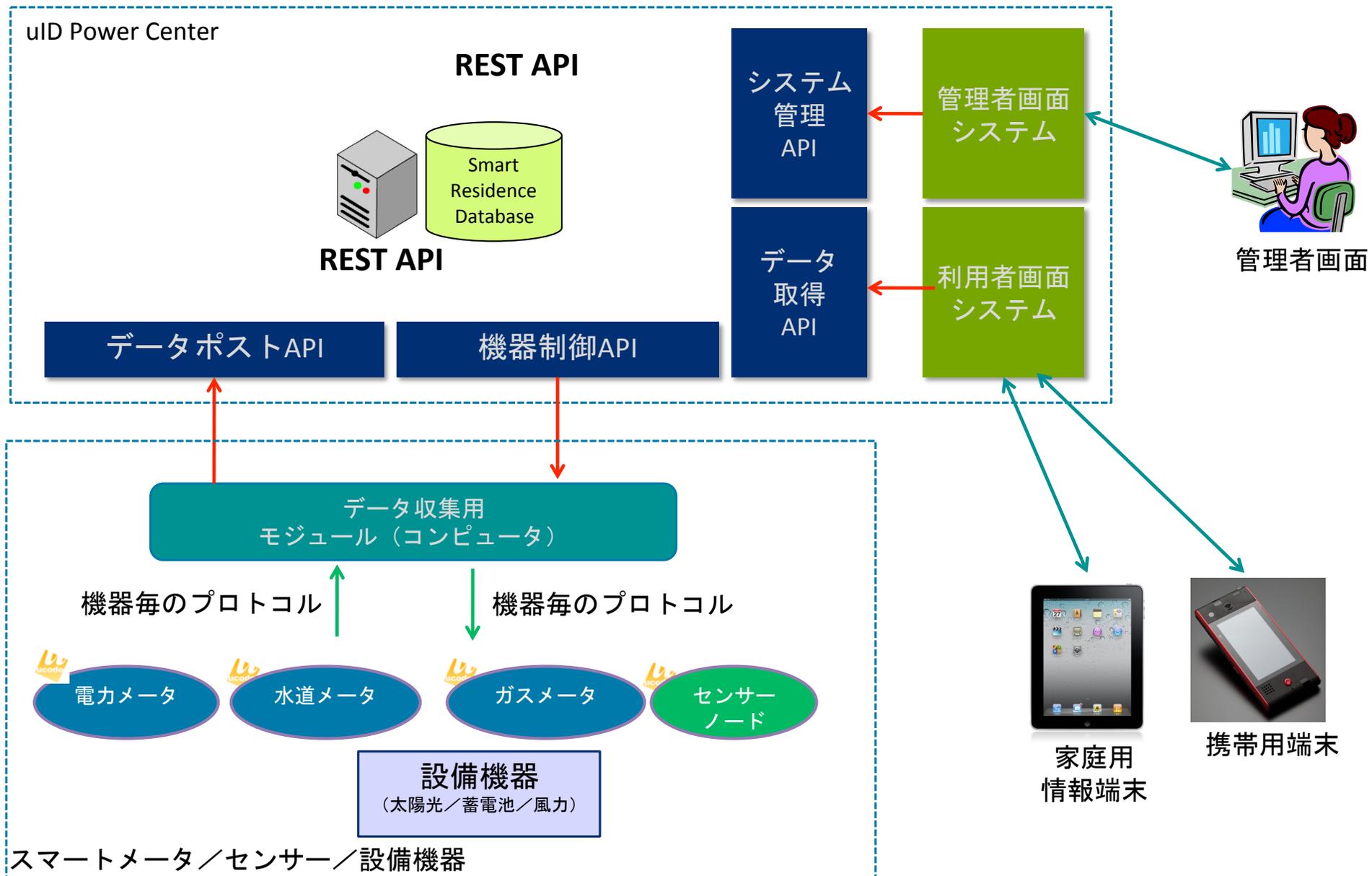
■ 事業主体

- ▶ 一般財団法人ベターリビング

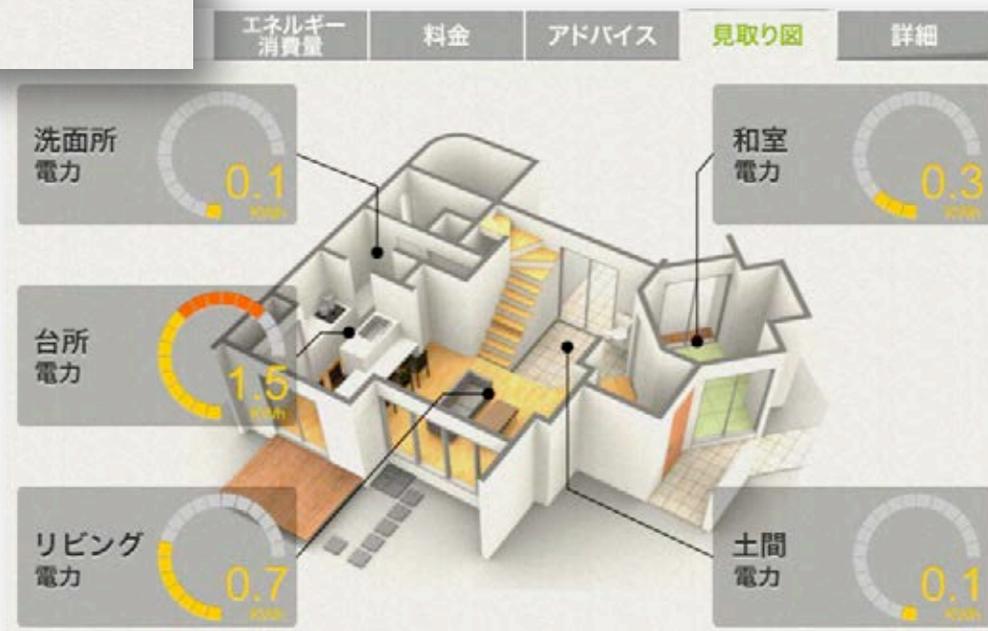
【エネルギー分野】スマート住宅（Low Energy House）



【エネルギー分野】 システム構成

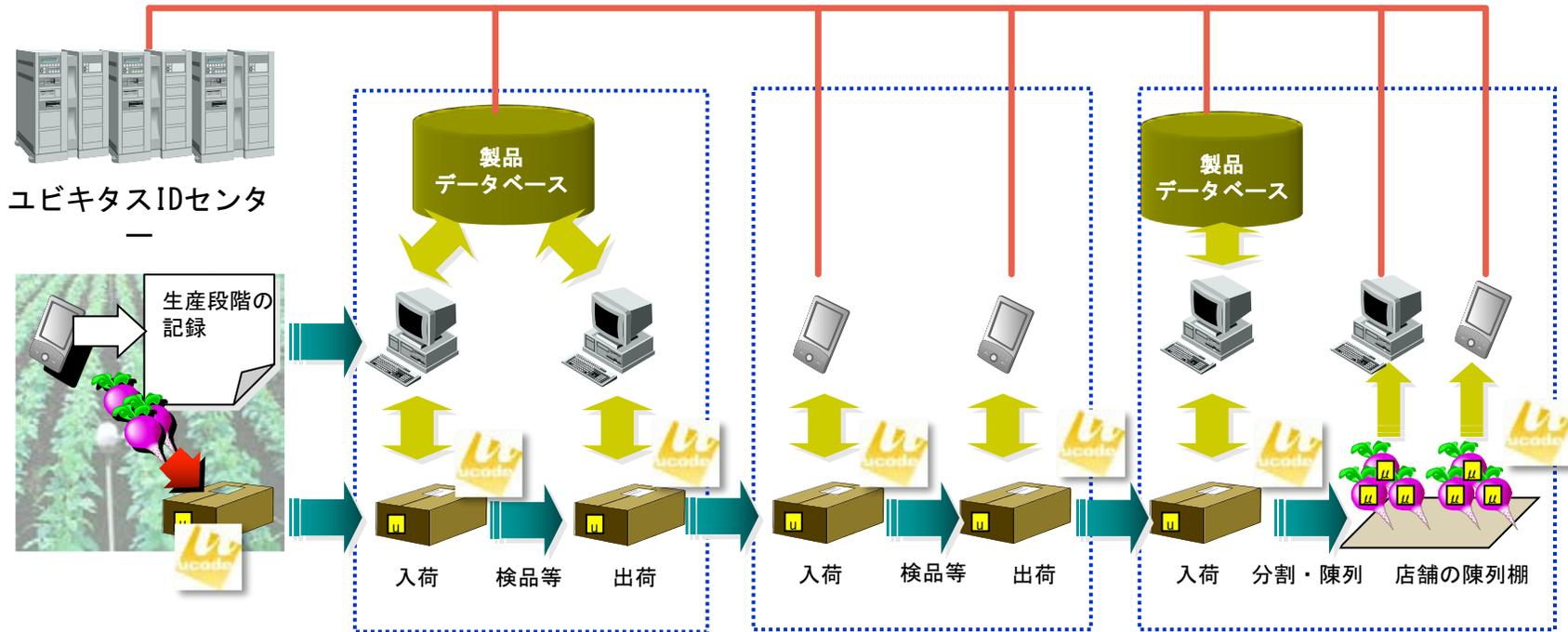


【エネルギー分野】 電力、水道の使用量のビュー



【食品分野】 食品トレーサビリティデータ基盤

インターネットによる接続



生産段階



準備段階



卸売り段階



小売り段階



【食品分野】食品トレーサビリティデータ基盤

戻る
青森りんご
日本語

商品をさがす
商品をさがす

飛馬ふじ

品種 飛馬ふじ



特徴

JA相馬村では、りんごの中で一番のシェアを誇る「サンふじ」の栽培方法を究め、味にコクのある高品質開発を続け、出来上がったサンふじをかわり商品「飛馬ふじ」として販売しています。「飛馬ふじ」は世界自然遺産白神山の恩恵を受けて生まれたサンふじの中でも日本一のサンふじを目指したりんごです。土壌診断・摘果時期・葉摘み時期・収穫時期・食などの各種条件をクリアし、さらに透過光方式センサー選果システムにより高品質なものだけを「飛馬ふじ」として厳選しています。

収穫日 2014-11-12

生産者氏名 大場 勉



生産者のコメント

ごだわりりんご「飛馬ふじ」の企画に参加して、会員全員が栽培方法を統一してりんご作りを行うことのむずかしさを感じました。しかしながらいざ出来あがったりんごを試食してみると、うま味や重量感など、いままで作ってきたりんごと明らかな違いを感じ、販売するためだけではなく消費者に「美味しい本物のりんご」を食べていただくかということを実感しました。「飛馬ふじ」は消費者へ自信を持ってお届けできるサンふじですので、ぜひ手に取って食べていただければと思います。

生産地 弘前(青森県弘前市人字紙漣沢字山越185-1)



通販関連サイト



リンク: [サンふじ](#)



リンク: [りんごジュース](#)



【物流分野】 電力メーターデータで不在配送解消（横須賀市、佐川急便）

■ 背景と課題

- ▶ 宅配における再配送は全小口配送の20%
- ▶ そのコストは全国で数千億円にのぼる。
- ▶ スマートメーターの導入が進み、2020年に東京電力管内、2024年に全国で導入が完了予定

■ 解決策

- ▶ スマートメーターから取得されるデータを用いて、各家庭の将来の在不在を人工知能技術（機械学習）で予測し不在先を回避
- ▶ 配送成功率は98%、不在配送は88%減少、総移動距離は6%減少（シミュレーション結果）
- ▶ 現状では、訪問により「不在」が配達者に特定された。本システムにより、在宅先を回るルートだけを「配達者」に提示することで「不在」は特定されず、よりプライバシーが守られる



AIが在宅かどうか予測 再配達負担軽減で実証実験へ

2020年7月9日 17時35分

集められた家庭の電力メーターのデータからAI=人工知能が、住民が在宅かどうかを予測し、宅配便のドライバーの再配達の負担を軽減するシステムの実証実験が、神奈川県横須賀市で、行われることになりました。

この実証実験は東京大学と佐川急便、それにデータ分析などを手がける「日本データサイエンス研究所」などが行い、9日、都内で開かれた記者会見で仕組みが紹介されました。

【福祉分野】 電力メーターデータでフレイル自動検知（三重県東員町）

- 介護予防におけるフレイルの重要性
 - ▶ 介護給付額の拡大、独居高齢者拡大の中で、要介護前に至る過程（筋力・活動低下等）である「フレイル」をいかに早期特定し、介護予防の手を打つかが課題
 - ▶ 現状の特定方法は対面式・能動的な検診に依存しているため、自治体活動からの「漏れ」が問題化。「住んでいるだけで受動的にフレイル判定ができる」技術が必要
- 電力データとAIを用いたフレイル検知の実現可能性
 - ▶ センサを用いた“フレイル判定”が研究開発・実証が活発化
 - ▶ スマートメータから取得可能な電力データ等とAIを組み合わせることで、フレイル判定の実現可能性が高い
 - ▶ 実証できた場合、世界的に先駆的かつ地域課題の有効な手法
- 東京大学・三重県の連携下での実現
 - ▶ 実施主体：東京大学・JDSC社・ネコリコ社
 - ▶ フィールド：三重県東員町
 - ▶ 東京大学・三重県間での連携協定を活用

要介護状態になる前にAIが検知 三重で実証実験へ

© 2020年01月16日 07時00分 公開

[ITmedia]



▶ 全国内企業が対象に「IFRS類似の新リース会計基準とは

▶ 新しい日常に合わせた「オンライン活用術」テレワークも学習も徹底指南！

合同会社ネコリコと日本データサイエンス研究所（以下、「JDSC」）、ならびに東京大学大学院情報学環 越塚登研究室は、3者共同で、AIと電力データを用いたフレイル検知に関する実証実験を2020年に三重県東員町において実施すると発表した。

20年1月設立予定の「東大・三重連携 介護予防に向けたAI・データ活用研究会」と連携し、東京大学地域未来社会連携研究機構、東京大学高齢社会総合研究機構、三重大学、三重県、東員町が参画して、東員町における20年中的の実施に向けて共同実験を進める。

「フレイル」とは、健康な状態と要介護状態の中間に位置し、身体的機能や認知機能の低下が見られる状態を指す。適切な治療や予防を行うことで要介護状態に進まらずにすむ可能性があることから、フレイルの早期発見、早期対応が重要なものとなっている。近年は、厚生労働省も自治体におけるフレイル予防対策を推進している状況だ。

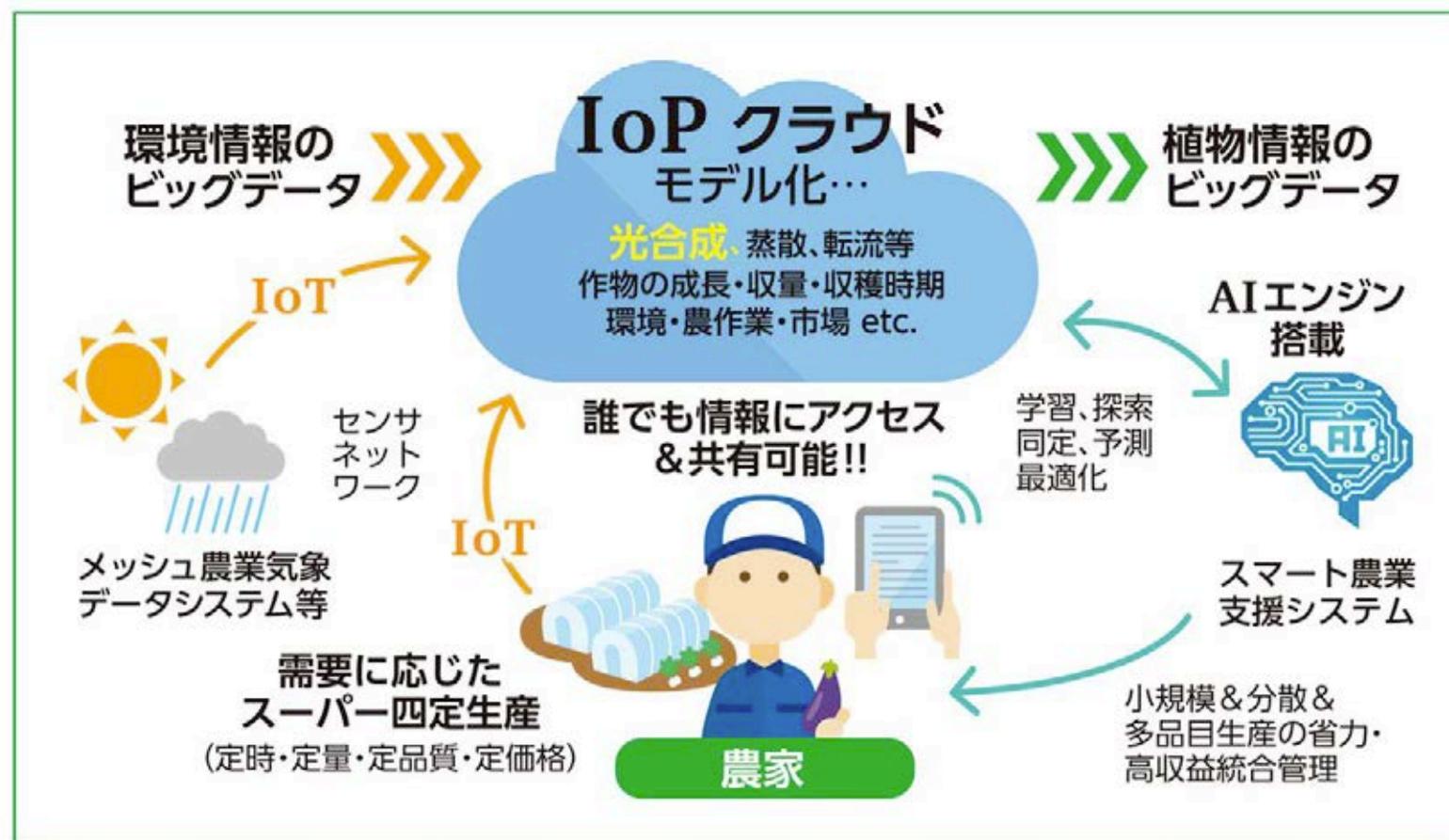


単身高齢世帯の増加にどう対応していくのか（写真提供：Gettyイメージズ）

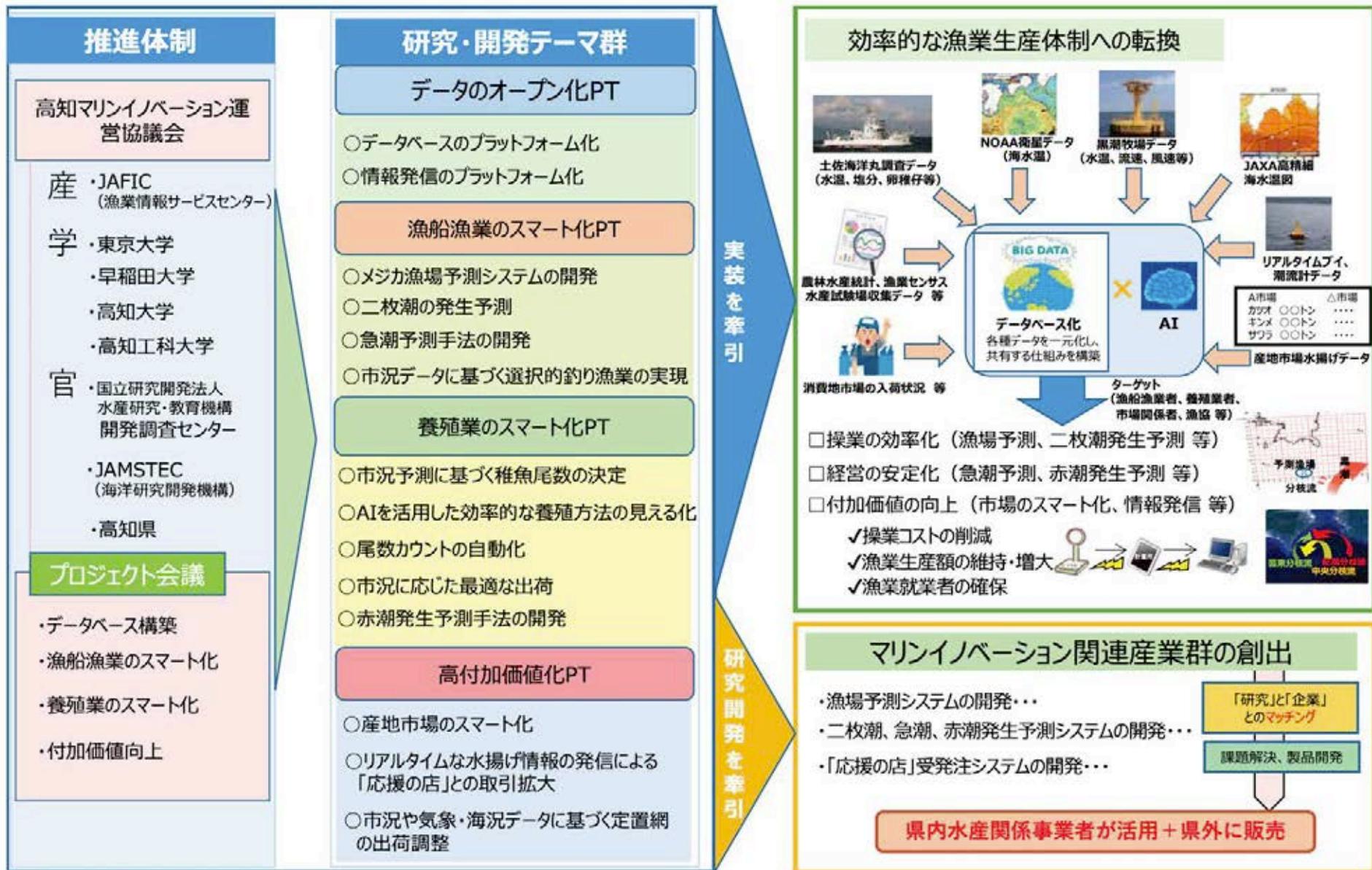
【農業分野】IoP (Internet of Plants: 高知県)

IoPクラウドで変わる施設園芸

データに基づく農業を行うためにIoPクラウドを構築します。最先端の研究からビックデータ（作物情報・農作業・環境情報）のデータベースにAIを組み合わせ機械学習などにより栽培・生産管理の最適化や出荷時期の予測などを行います。いつでも、誰でもハウスの状況や生育状態に合わせて活用できる環境が、農家の生産向上に繋がります。



【漁業分野】高知マリンイノベーションプロジェクト（高知県）



分野毎データ基盤の整備（国内事例）。

- COVID-19
 - ▶ 感染防止、混雑データ
- 公共オープンデータ
 - ▶ 総務省統計(e-stat)
 - ▶ 政府オープンデータ (data.go.jp)
 - ▶ 地方自治体オープンデータ (VLED)
- 交通分野
 - ▶ 公共交通オープンデータ (ODPT)
 - ▶ GTFS (日本ではバス共通フォーマット)
 - ▶ 道路情報 (渋滞情報、通行止め情報、...)
 - ▶ MaaS (MaaSアーキテクチャ、国交省)
 - ▶ 航空離発着データ
- 地理空間分野
 - ▶ G空間 (G空間センター)
 - ▶ 地図 (国土地理院、ゼンリン地図、Google Map、...)
 - ▶ 場所情報コード (国土地理院)
 - ▶ パブリックタグ (国土地理院)
 - ▶ 自律移動支援のバリアフリーマップ (国交省)
- 産業分野
 - ▶ ロボット革命イニシアティブ (RRI)
 - ▶ インダストリアル・バリュー・チェーン・イニシアチブ (IVI)
 - ▶ 建築 (BIM、国交省データ)
- 農業データ (WAGRI)
 - ▶ IoP (高知県)
 - ▶ 食品トレーサビリティ (農水省 食品トレーサビリティガイドライン) : 牛、米、...
- 金融
 - ▶ FINTECH
 - ▶ 仮想通貨、電子マネー
- 経済
 - ▶ EDI (Electronic Data Interchange)
 - ▶ デジタル領収書
- 自然環境
 - ▶ 宇宙データ (Tellus)
 - ▶ 気象データ (気象庁、WXBC)
 - ▶ 海洋データ
- 医療
 - ▶ 医療 (#2 SIP)
 - ▶ HL7 FHIR (Fast, Health, Interoperable, Resources) ¥
 - ▶ 医薬品 (添付文書、電子お薬手帳、)
 - ▶ 医療データ (電子カルテ、レセプトデータ、電子母子手帳) 福祉・介護
- ウェブページ
 - ▶ アスコエ社「行政メニューの統一」
- 防災
 - ▶ SIP4D
- 都市
 - ▶ スマートシティ、スーパーシティ
- 観光
 - ▶ TravelXML 1.0
 - ▶ 観光予報プラットフォーム
- 教育
 - ▶ eポートフォリオ (教育のデータ化)
- 住宅
 - ▶ 超長期優良住宅 (家カルテ)
- 学術データ
 - ▶ NII
 - ▶ mdx (東大、NII、他)
- デジタルコンテンツ
 - ▶ ジャパンサー https://jpsearch.go.jp
- 図書館
 - ▶ 国会図書館図書DB
 - ▶ カーリル (民間サービス)

COVID-19のデータ基盤事例（アカデミック分野を中心に）

- IEEE DataPort
 - ▶ <https://ieee-dataport.org/topic-tags/covid-19>
- google data set search
 - ▶ <https://datasetsearch.research.google.com/search?query=covid-19&docid=Gsh%2ByOoeNPtHtq5%2BAAAAAA%3D%3D>
- Microsoft (COVID-19 Open Research Data)
 - ▶ <https://azure.microsoft.com/ja-jp/services/open-datasets/catalog/covid-19-open-research/>
- AWS (AWS COVID-19 データレイク)
 - ▶ <https://aws.amazon.com/jp/covid-19-data-lake/>
- Kaggle (COVID-19 Open Research Dataset Challenge (CORD-19))
 - ▶ <https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge>
- EU-data portal (COVID-19 Coronavirus data)
 - ▶ <https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data>
- Github (COVID-19 Open-Data)
 - ▶ <https://github.com/GoogleCloudPlatform/covid-19-open-data>
- COVID-19 open source data sets: a comprehensive survey (*1)
 - ▶ <https://link.springer.com/article/10.1007/s10489-020-01862-6>
- US The COVID Trakcing Project
 - ▶ <https://covidtracking.com>
- John Hopkins Univ. (CORONAVIRUS Resource Center)
 - ▶ <https://coronavirus.jhu.edu>
- dataworld
 - ▶ <https://data.world/datasets/covid-19>
- UK Gov. (All data related to coronavirus (covid-19))
 - ▶ <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datalist?filter=datasets>
- COVID 19 Resources (Academic Data Science Alliance)
 - ▶ <https://academicdatascience.org/covid>
- WHO
 - ▶ <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov>



[Issue 4]
分野間データ連携基盤
dataex.jp

デジタルサービスの質・量の向上には
インフラの整備が重要
(情報通信の歴史観からくる基本戦略)



データ基盤は最重要課題

他にも、通信基盤（有線、無線）、制度基盤、教育基盤、...

Society 5.0に向けて
日本が考える**自由**・**健全**・**適正**な
データ流通（DFFT）の姿を確立する必要



データ流通連携基盤の確立

米国型でも中国型でもない形とは？

むしろ世界の大勢は、米国型でも中国型ではない。

日本国内で機は熟している

各分野内におけるデータ流通連携基盤（例）：機は熟している



政府オープンデータ



自治体オープンデータ



官民連携データ



おもてなしクラウド



情報銀行



データ取引市場



公共交通



製造業



気象



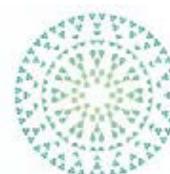
学術



AI



宇宙



農業

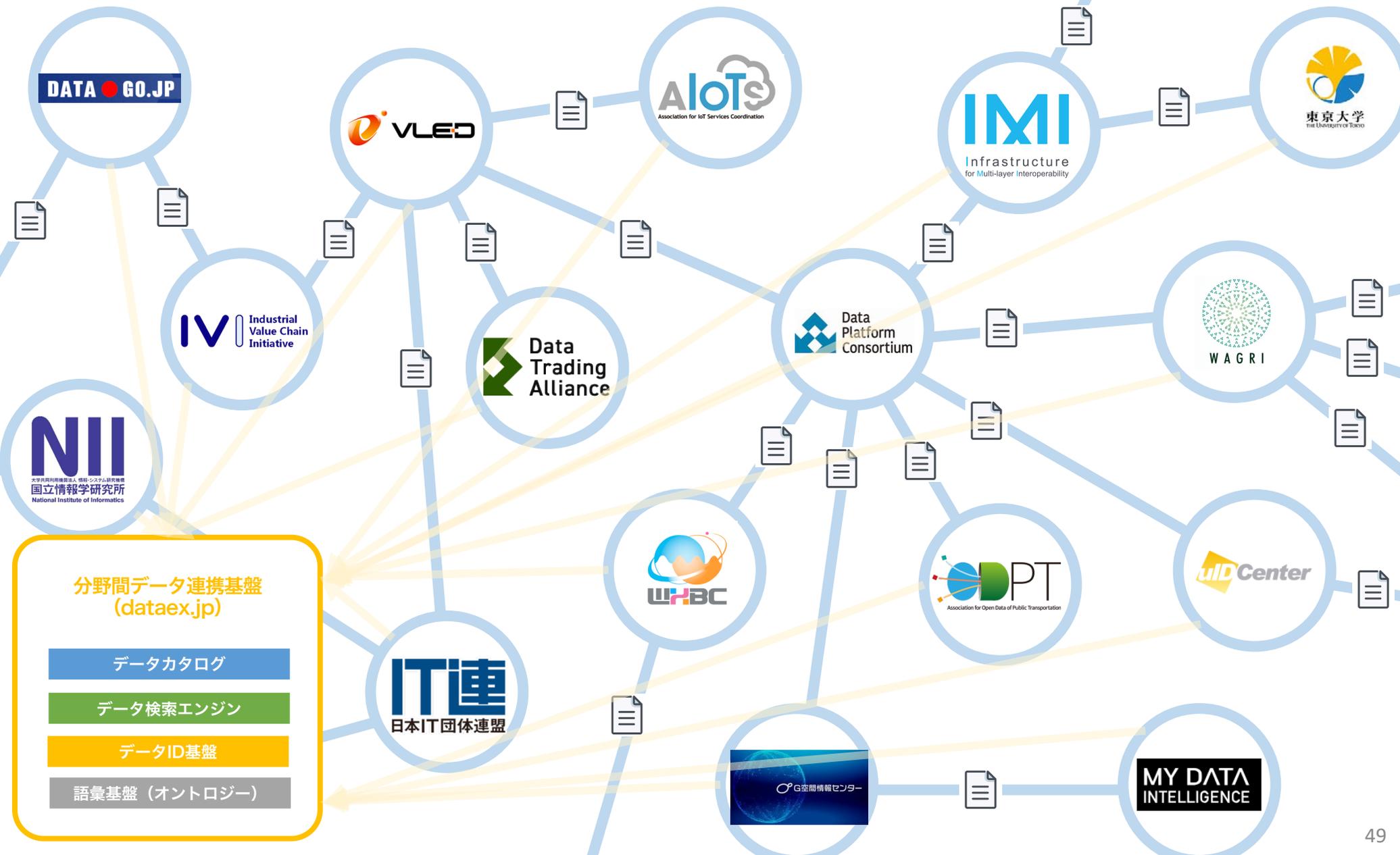


地理空間

データに関する国内の取組一覧

| データに関する取組 | | キーパーソン |
|-----------|---------------------------------|---|
| 1 | 内閣府SIP分野間データ連携基盤 | 安西祐一郎（慶應義塾大学名誉教授・元塾長）PD 越塚登（東京大学教授）SPD |
| 2 | dataex.jp（準備中） | 越塚登（東京大学教授）準備会会長 |
| 3 | 内閣府スーパーシティ（アーキテクチャ） | 越塚登（東京大学教授）座長 |
| 4 | 内閣府SIPスマートシティアーキテクチャ | 越塚登（東京大学教授）座長, SPD |
| 5 | 国土交通省 MaaS関連データ連携、標準バス形式GTFS | 越塚登（東京大学教授）座長 |
| 6 | 一般社団法人オープン＆ビッグデータ地方創生推進機構（VLED） | 坂村健（東京大学名誉教授、東洋大学教授）理事長 越塚登（東京大学教授）理事 |
| 7 | 公共交通オープンデータ協議会 | 坂村健（東京大学名誉教授、東洋大学教授）会長 |
| 8 | 東京大学オープンデータセンター | 越塚登（東京大学教授）センター長 |
| 9 | 一般社団法人インダストリアル・バリューチェーン・イニシアティブ | 西岡靖之（法政大学教授）理事長 |
| 10 | 一般社団法人 データ流通推進協議会（DTA） | 村井純（慶應義塾大学教授）理事長 |
| 11 | 一般社団法人官民データ活用共通プラットフォーム協議会（DPC） | 須藤修（東京大学名誉教授、中央大学教授）理事長 |
| 12 | 気象ビジネス推進コンソーシアム（WXBC） | 越塚登（東京大学教授）会長 |
| 13 | 農業データ連携基盤協議会（WAGRI） | 神成淳司（慶應義塾大学教授、政府CIO補佐官）会長 |
| 14 | 東京都官民連携データプラットフォーム | 宮坂 副知事以下、越塚、柴崎先生、吉村先生、等 |
| 15 | ヨコスカモビリティチャレンジ | 中村 文彦（横浜国大副学長）会長 越塚登（東京大学教授）展開戦略TF座長 |
| 16 | 札幌オープンデータ協議会 | 越塚登（東京大学教授）会長 |
| 17 | G空間情報センター／一般社団法人社会基盤情報流通推進協議会 | 関本義秀（東京大学准教授）代表理事 |
| 18 | 一般社団法人 日本IT団体連盟・情報銀行推進委員会 | 井上貴雄 委員長 |
| 19 | 一般社団法人 IoTサービス連携協議会（AIoTS） | 坂村健（東京大学名誉教授、東洋大学教授）理事長 |
| 20 | 国交省・パブリックタグ情報共有プラットフォーム | 越塚登（東京大学教授）座長 |

分野間データ連携基盤：既存の分野毎データ基盤をつなぐ



重要なことは、**既存**の取り組みを生かして
分散連邦型で基盤とすること

Distributed and Federated

dataex.jp

分野間データ連携基盤

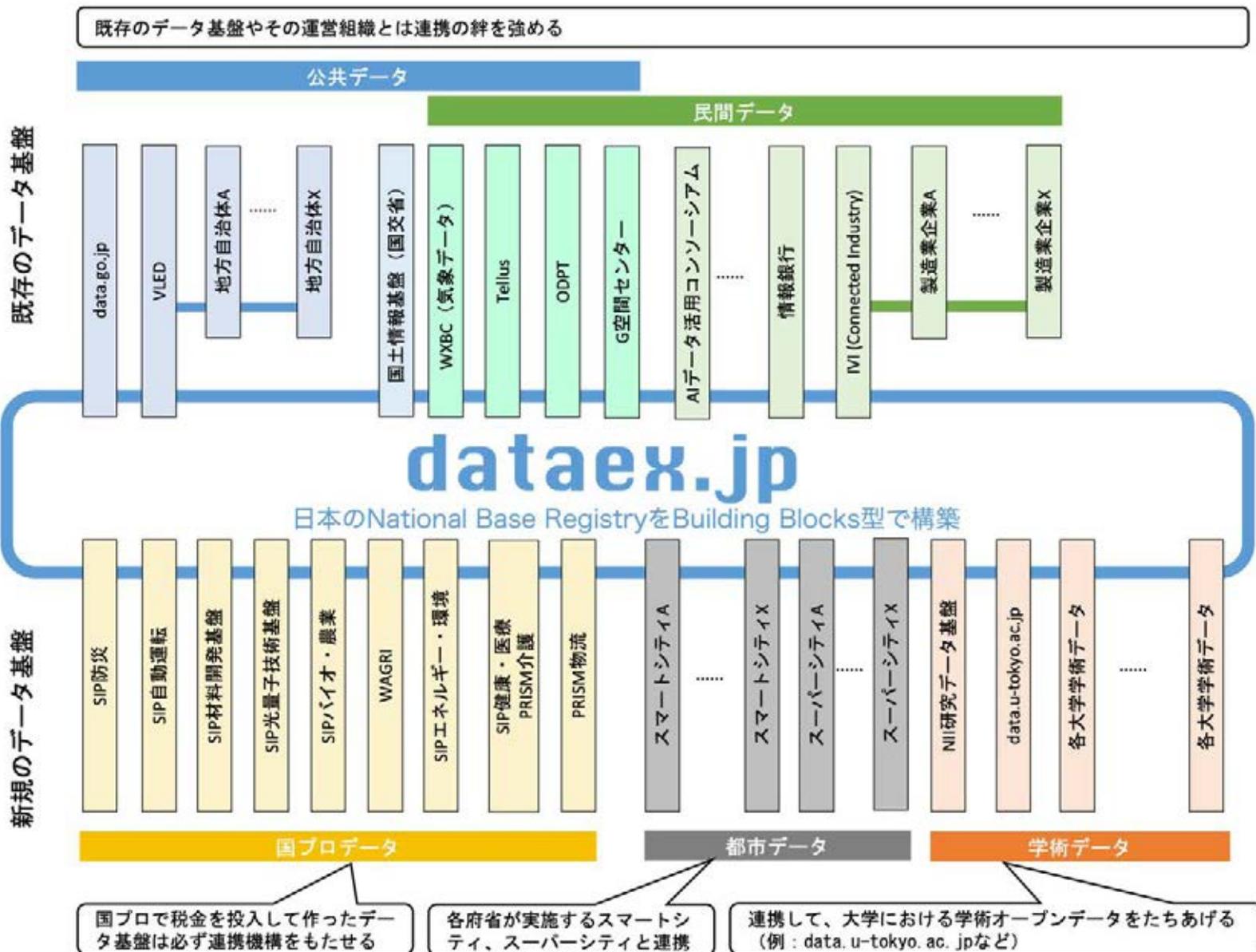
2020年度中に設立予定（準備中）

分散連邦型による分野間データ基盤（データは持たない）

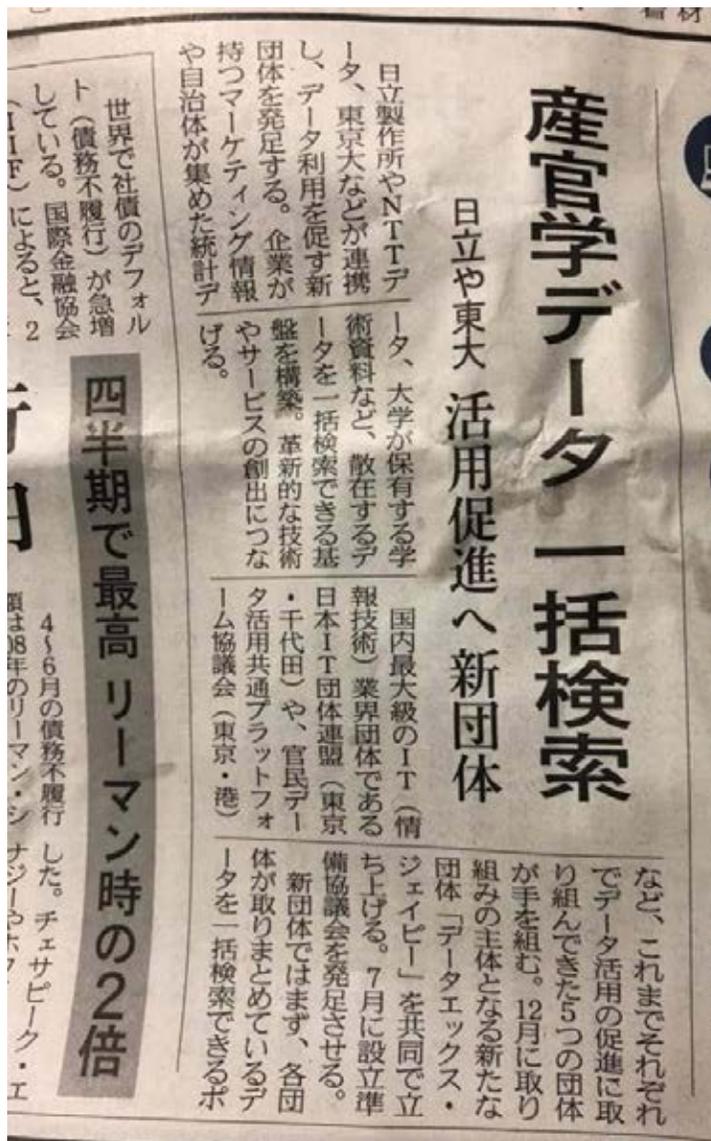
Building Blocks型で作る日本のNational Base Registry

最小限の軽いシステム（日本型アーキテクチャ）

dataex.jp: 分野間データ連携基盤



dataex.jpのアウトリーチ状況



2020年7月18日(土) 日本経済新聞朝刊 1面



出典：内閣府データ連携基盤サブワーキンググループ（第3回）資料1より抜粋

図 分野間データ基盤連携イメージ

また、分散型の分野間データ連携を促進するルールや仕組みを整備するため、民間の促進団体として「dataex.jp」の本年内の設立が計画されている（本年7月に設立準備協議会を発足済み）。

今後、dataex.jpでは、コネクタ提供を行うポータルサイトの運用や、技術標準等の国際標準化を推進することとしており、必要に応じて、求められる機能の拡充が必要である。

こうしたdataex.jpの活動に加えて、IPA デジタルアーキテクチャ・デザインセンター（DADC）において、重要分野ごとのルール設定の支援等を行うこと、およびその知見を蓄積することを必要に応じて検討していく。

第2回データ戦略TF資料

【アーキテクチャ】国等の取組状況（2/2）

現在内閣府SIP※にて検討が進んでいる「分野間データ連携基盤」や、設立準備がなされている「dataex.jp」の取組も踏まえ、他分野とデータ連携が可能なアーキテクチャ等を検討する

概要

| | | |
|--------------------------|--|--|
| <p>分野間データ連携基盤</p> | <ul style="list-style-type: none"> 関係府省庁で整備が進められている分野ごとのデータ連携基盤やその他の様々なデータを相互に連携させる分野横断のプラットフォーム 内閣府のSIPの取組※にて研究され、2022年までに実用版の構築と社会実装を目指す | <ul style="list-style-type: none"> 官民連携DPFにおいても、DPFが分野間データ連携基盤の連携先の1つとなることも見据え、SIPの研究状況を踏まえたシステムを構築していく |
| <p>dataex.jp</p> | <ul style="list-style-type: none"> 上記研究成果も活用して、産官学一体でデータ流通・利活用を促進する連携組織 データ流通・促進に係る5団体にて準備協議会を本年7月に設立 | <ul style="list-style-type: none"> dataex.jpの設立準備協議会の検討・取組と適宜連携を図る |

※内閣府 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
 出典：内閣府 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」（2020年7月9日）、
 「dataex.jp」設立準備協議会「プレスリリース」「dataex.jp」設立準備協議会発足のお知らせ」（2020年7月17日）より



[Issue 5]
データ連携基盤の整備にむけて
想定される課題

データ基盤の構成

■ 集約手法

- ▶ 集中型 vs 分散型（連邦型）

■ 連携技術

- ▶ システム連携（同じシステムをつかう）
- ▶ データ連携（共通データベース化）
- ▶ API連携（分散データベース化）
- ▶ 個別対応（何も仕切らない、Connectivity only）

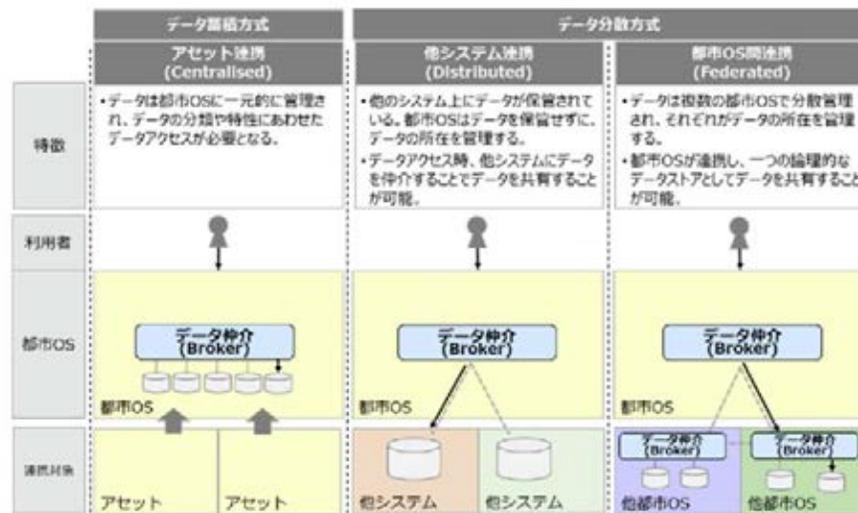
■ 経営方法

- ▶ 収益性 vs 共益性（ニュートラル）
 - ◆ 収益性：ちゃんとお金がまわる、お金が回る規模感
 - ◆ 共益性：説明責任、特定の人だけでなく広く平等に使われる
- ▶ （サービス＋基盤）統合 or 分割

■ 外部連携

- ▶ 国との連携
- ▶ 他地域との連携
- ▶ 所掌外の分野との連携

他



費用負担

- ・ 公的負担
- ・ 業界負担
- ・ 収益事業化（自走）
- ・ それらのバランス

データの種類

■ データの種類

- ▶ オープンデータ
- ▶ 制限データ（有料、Licensed、...）
- ▶ パーソナルデータ（含、個人データ）
- ▶ その他（オルタナティブデータ、SNS、掲示板書込み、CGM、ログデータ...）

■ 変更頻度による分類

- ▶ アーカイブ（記録、文書、写真、...）：read only, no update
- ▶ 静的データ（統計等）
- ▶ 動的データ（リアルタイム統計、今日のデータ、...）
- ▶ 動的データ（機械的計測、センサーデータ）

■ オリジナリティによる分類

- ▶ 一次データ（オリジナルソースのデータ）
- ▶ 二次データ（加工・編集済データ、組み合わせデータ、等）

データの種類毎に、必要な技術や、制度・ルールが異なる

データの分野・用途

- 基礎統計
- 公共交通
- 人流
- パーソナルデータ
- 購買データ
- 地理データ（地図、等...）
 - ▶ 屋外よりは屋内・地下街
- 防災
- バリアフリー
- 図面・台帳類
- 写真、コンテンツ、ライブラリ
- 駐車場
- 気象、災害（水位）
- エネルギー
- ライブカメラ
- コミュニティデータ（SNS等）
- ...

**データの分野・用途毎に、
必要な技術や、制度・ルールが異なる**

- 犯罪やテロ、非合法行為にデータが利用された場合の考え方
- 反社会的活動への利用
- ↓
- データ提供側としてケアするのか？

データコミュニティ

- ステイクホルダー間のコミュニケーション
 - ▶ ラウンドテーブル
 - ▶ アイデアソン・ハッカソン・ワークショップ・コンテスト
 - ▶ アセスメント
 - ▶ 開発者サイト (Developers' Site)
 - ▶ 学校連携 (地元の大学、高専、..等)

- 人材育成、スキル訓練
 - ▶ 講習会
 - ▶ オンライン教材
 - ▶ 等

- 事例共有
 - ▶ データ利用事例の公開 (例: OpenData 500、...)
 - ▶ アプリの公開
 - ▶ 等

データに関するルール・規則・規制（国レベルの事例）

■ データ取引契約

- ▶ 「AI・データの利用に関する契約ガイドライン」（平成30年、経済産業省）
- ▶ 「データの利用に関する契約ガイドライン産業保安版」（第2版。平成31年、経済産業省）
- ▶ 「農業分野におけるAI・データに関する契約ガイドライン」（令和2年、農林水産省）

■ データ取引市場

- ▶ データ取引市場運営事業者認定基準_D2.0（DTA）
- ▶ データ取引市場運営事業者認定基準_説明_REV1.1（DTA）

■ データ提供主体の真正性

- ▶ 「行政手続におけるオンラインによる本人確認の手法に関するガイドライン」（平成31年2月、各府省情報化統括責任者（CIO）連絡会議決定）
- ▶ 米国NIST「デジタルアイデンティティガイドライン（SP800-63-3）」

■ パーソナルデータの取り扱い

- ▶ 個人情報保護法（日本）－2000個ルール問題
- ▶ GDPR（EU）
- ▶ FIPPS（米国）
- ▶ 「情報信託機能の認定に係る指針ver2.0」
- ▶ 「情報信託機能の認定スキームの在り方に関する検討会とりまとめ」（令和元年10月、総務省・経済産業省 情報信託機能の認定スキームの在り方に関する検討会）
- ▶ 消費者向けオンラインサービスにおける通知と同意・選択ガイドライン、ISO/IEC 29100 (ISO/IEC 29100:2011 (Information technology –Security techniques – Privacy framework))
- ▶ 「スマートライフ個人情報都立扱い方針」（経産省）

■ データの信頼性

- ▶ 「DMBOK」：データ品質マネジメントの指針
- ▶ 「ISO8000シリーズ」：提供元のデータガバナンスの評価
- ▶ 「品質の要求、評価及び報告のための規則」（国土地理院、令和元年7月）
 - ◆ 観測データの品質の評価の取組み・・・地理空間情報の品質要求、評価、報告、
- ▶ 「2030年の科学技術を見据えた気象業務のあり方（提言）」（2020.8.20 交通政策審議会気象分科会）
 - ◆ 観測データの品質の評価に向けた取組み・・・気象データの「品質の見える化」
- ▶ 一般社団法人データ流通推進協議会のデータカタログ作成ガイドライン
- ▶ メタデータによりデータの属性を説明する取組み

■ データガバナンス

- ▶ データの所有権、コントロール権の設定
- ▶ データ漏えいへの対応
- ▶ データ独占
- ▶ 不正競争防止
- ▶ レピュテーション対応（炎上しない）
 - ◆ サービスアセスメント

■ データ標準ライセンス

- ▶ Open Source License
 - ◆ GPL (General Public License)
- ▶ Open Data License
 - ◆ Creative Commons
 - ◆ 政府標準規約

パーソナルデータの課題

■ 価値観のトレードオフ

- ▶ ダイバーシティ対応の実現 ⇔ プライバシーの保護



解決のためには、住民との間の議論が不可欠（アセスメント）

■ パーソナル化されたサービス...スマートシティの重要なコンセプトの一つ

- ▶ 街に人があわせるのではなく、人に街があわせる
- ▶ 全員に同じサービスではなく、個々人のダイバーシティに応じたサービス
- ▶ 反面、おせっかいとも



- ▶ パーソナルデータが不可欠

■ プライバシーの保護

- ▶ パーソナルデータの流通のコントロール権は個人に
 - ◆ 個人は自分のデータの制御を理性的に実施可能か？
- ▶ 個人情報関係の法制度の2000個問題

パーソナルデータ、個人情報の扱いは難しく、且つ漏洩時のリスクも大きい



データ基盤で管理してもらえると事業者側は嬉しい面もある

(技術課題) データ連携を含む、データ利活用のV字型プロセス (仮説)

企画立案フェーズ 実装フェーズ

サービス層

実現したいサービス要求

サービスを実現する

(データ)
連携基盤層

使えるデータを探す

1. Data Finding Tech.
2. Meta Data
3. Rating

データをつなげる (連携する)

1. Social Interoperability
2. Tool Interoperability
3. Semantic Interoperability
4. Format Interoperability
5. Access Interoperability

(個別データ)
利活用層

(個別) データを検討

データを使えるようにする

(技術課題) データをつなげる (連携) ためには以下の相互運用性が不可欠

相互運用の内容

1 Social Interoperability
社会・契約面での連携

ライセンスの違い、売買、取引の違い

2 Tool Interoperability
ツールの連携

データを扱うツールの相互運用性
・プログラミング言語、データ利活用ライブラリ、などの相互運用性

3 Sematic Interoperability
意味の連携

データの意味の相互運用性
・パラメータ表現・語彙表現や、単位行現、などの相互運用性

4 Format Interoperability
フォーマットの連携

データ形式の相互運用性
・ファイル形式、パラメータデータ形式、文字コード、などの相互運用性

5 Access Interoperability
アクセスの連携

データ取得方法の相互運用性
・通信プロトコル、データベース言語、などの相互運用性

データの品質・信頼性

■ データの技術的品質

- ▶ 5 Star Open Data
- ▶ 「機械学習の技術的負債」：膨大なデータクレンジングコスト
- ▶ 「ネ申エクセル」問題
- ▶ 文字コード、外字、...
- ▶ ベストエフォート品質？

■ データの信頼性

- ▶ データの出所、責任は誰に？
- ▶ 改ざんはされていないか？ → ブロックチェーンで担保
- ▶ データのバイアス
- ▶ データの説明責任

■ データの誤りが原因による損害

- ▶ 法的課題
- ▶ 風評課題

データ品質は課題山積

For more info...

**<http://www.koshizuka-lab.org/>
noboru@koshizuka-lab.org**



Copyright © 2020

Koshizuka-Laboratory
The University of Tokyo
All Rights Reserved