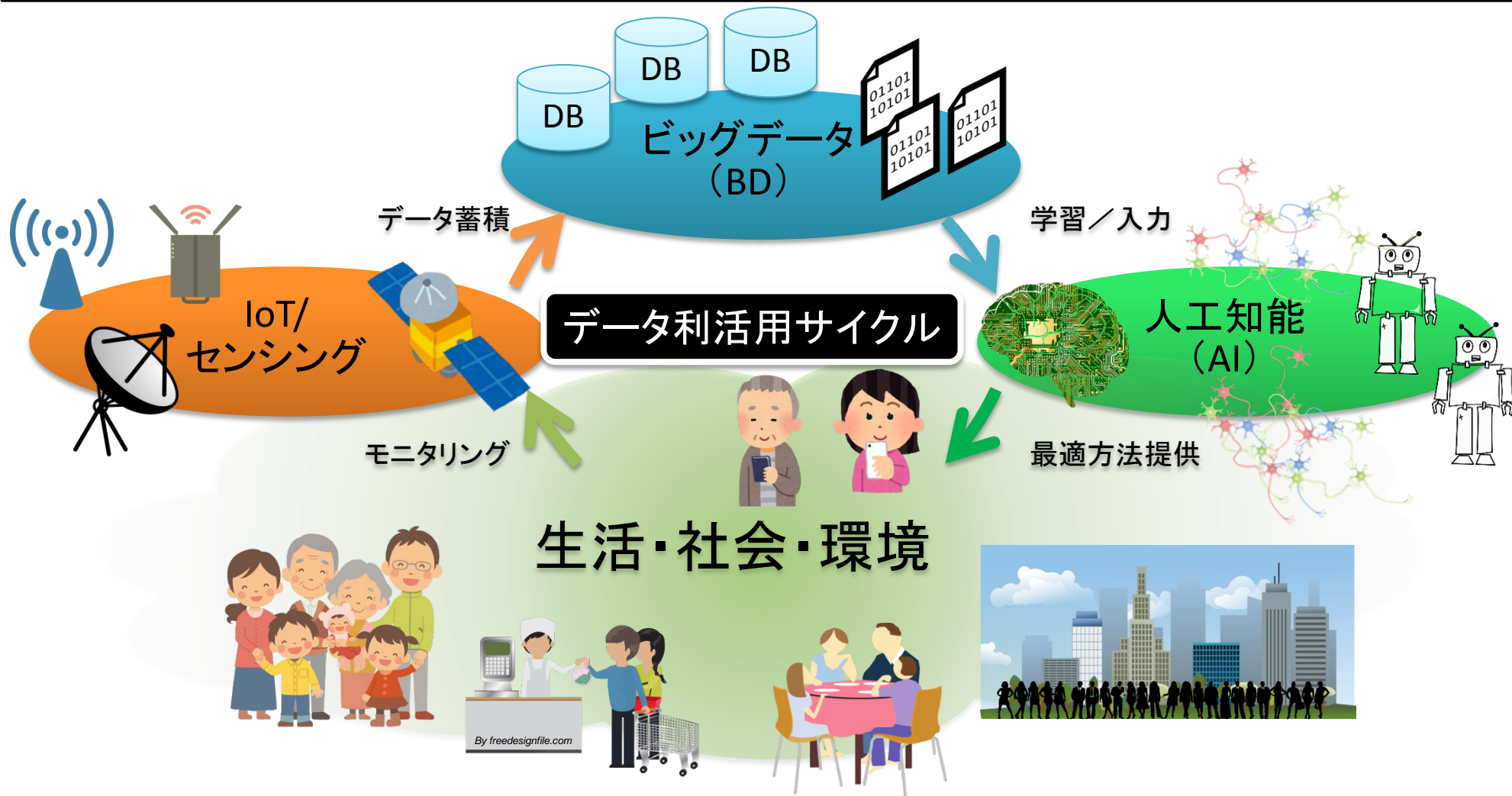


# データ利活用を中心とした 研究開発の現状及び課題



# 社会におけるデータ利活用の重要性

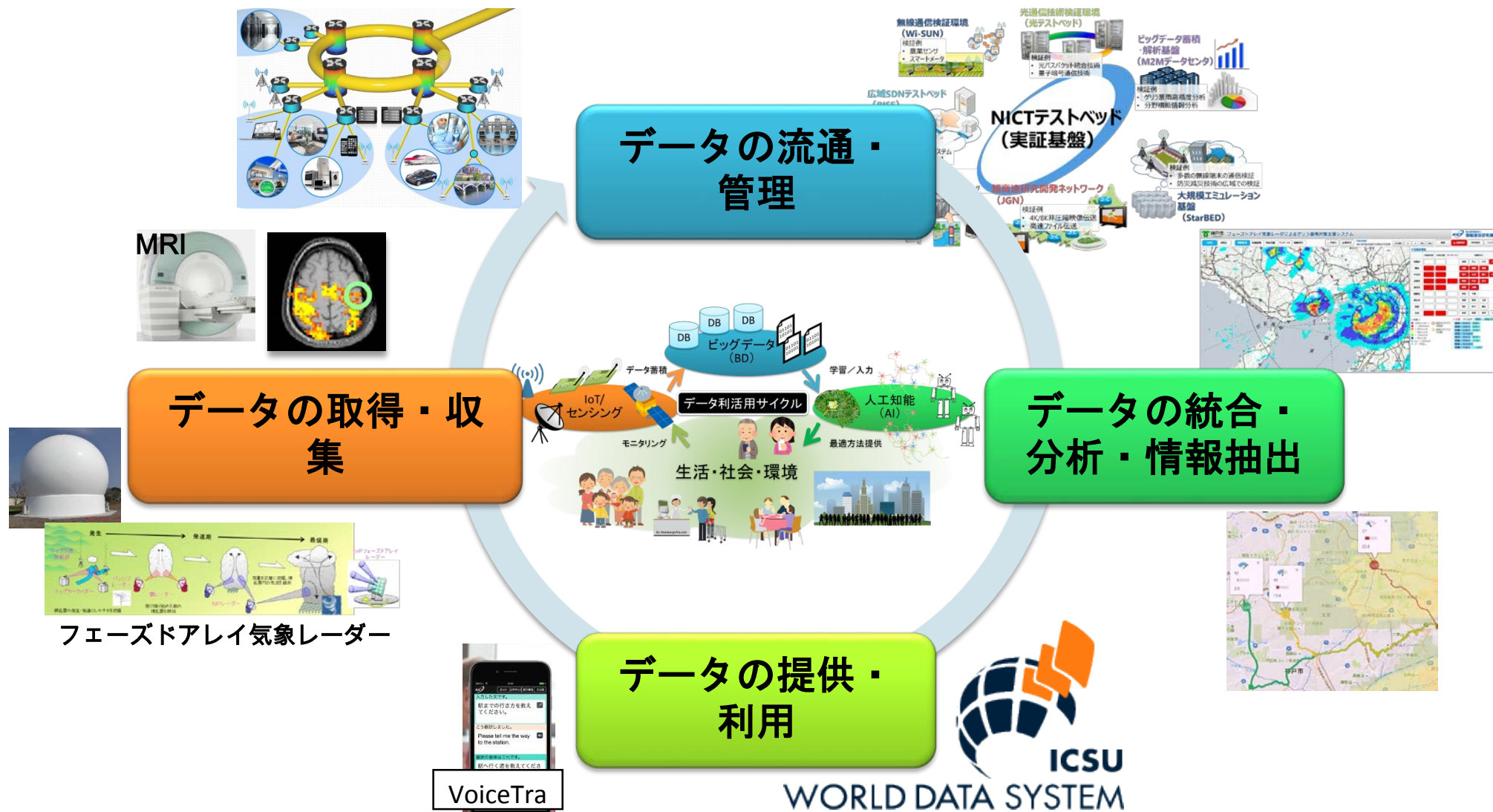
IoT、ビッグデータ(BD)、人工知能(AI)技術の加速により、多種多様なデータ利活用が可能となり、社会における課題の解決や新たなサービスの創出などを通じて、生活の質の向上につながる事が期待される。



# NICTのデータサイエンスの概要



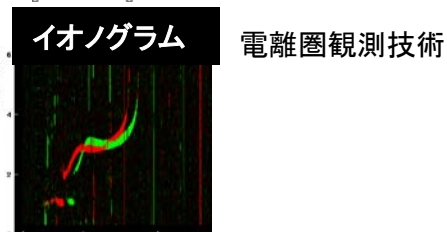
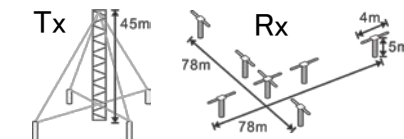
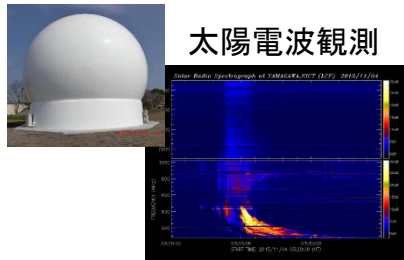
NICTでは、データの「取得・収集」から「流通・管理」、「統合・分析・情報抽出」、「提供・利用」までの各フェーズにおいて研究開発を行い、それらの連携にも取り組んでいる。



- 宇宙から微生物までの科学データやソーシャルデータまで幅広い分野にわたるデータを取得・収集。
- 貴重で高精度なデータを取得するための研究開発を行いつつ、ソーシャルメディアを活用したデータの収集も実施。

## 宇宙環境

- 太陽電波観測データ
- 電離圏観測データ



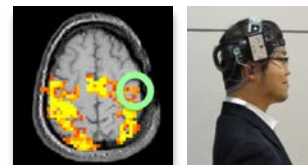
## 気象・地表

- 衛星搭載機器による気象・大気観測データ
- 地上3D気象レーダー観測データ
- 航空機搭載合成開口レーダー(SAR)測定データ



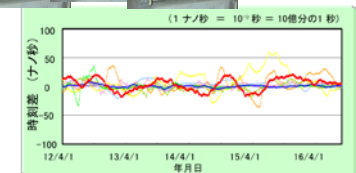
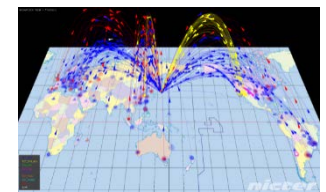
## 生体・生物化学

- 脳機能計測データ
- 音声コーパスデータ
- バクテリアによる化学物質計測データ



## ソーシャル・その他

- ウェブアーカイブ
- Twitterデータ
- サイバー攻撃情報
- 原子時刻データ

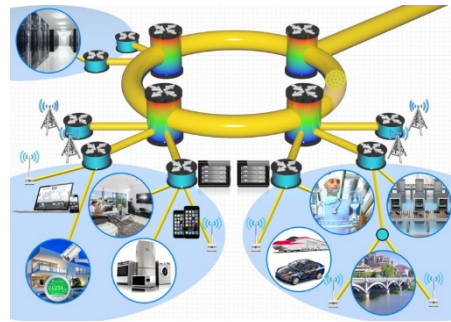


# データの流通・管理技術への取組・研究開発の例

- 大容量・多種多様なデータの流通のためのネットワーク技術の研究開発
- データの流通・利用における安全・安心の確保につながる研究開発
- データの利活用に関する技術を実証するための環境構築

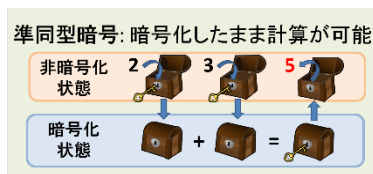
## ネットワーク技術の研究開発

- IoTやクラウドなど多様化する利用環境に柔軟に対応できるネットワーク技術
- 急増する通信量に対応する大容量マルチコアネットワーク技術



## 暗号技術・プライバシー保護技術の研究開発

- 暗号化したまま処理を可能とする技術
- プライバシーを保護したままデータ分析を可能とする技術



## NICTテストベッド



様々な技術の研究開発・実証を行うための基盤として、データの取得・流通・蓄積・解析のプラットフォーム(テストベッド)を構築

# データの統合・分析・情報抽出技術への取組み・研究開発の例

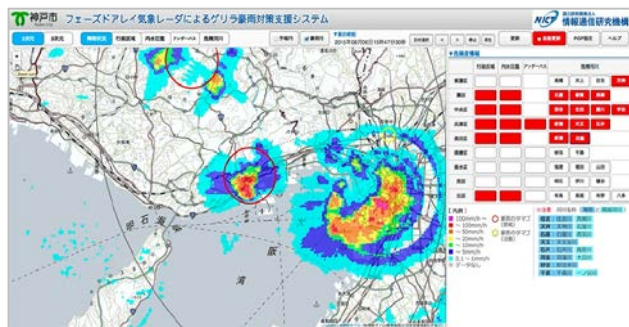
- 大量の科学・ソーシャルビッグデータから情報の抽出や分析を行う研究開発
- 科学データとソーシャルデータを融合して有益な情報を生成する研究開発

## ソーシャルデータからの情報抽出

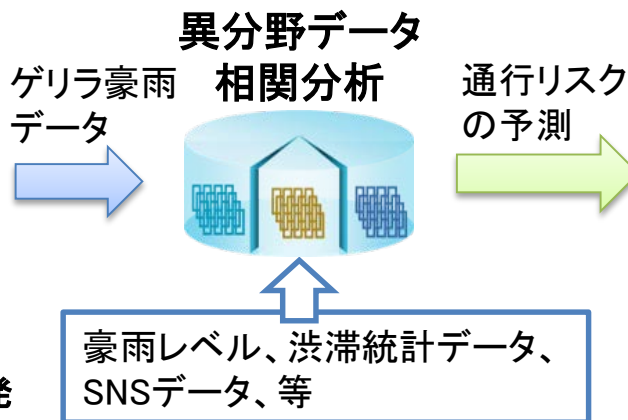


- 大量のウェブ情報から様々な質問に対する回答を抽出したり、質問自体を生成する技術 (WISDOM X)
- Twitter等から災害に関する情報の抽出する技術 (DISAANA)

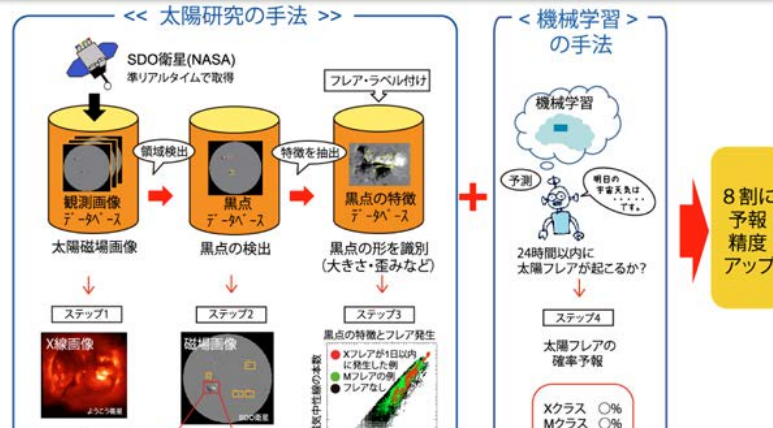
## 科学データとソーシャルデータの連携



フェーズドアレイ気象レーダーデータと自治体ハザードマップを利用したゲリラ豪雨対策支援システムの研究開発



## AIによる科学ビッグデータの分析



大量の太陽面画像からAI解析による太陽フレア発生予測精度の向上

# NICTのデータ提供・利用の例

取得した科学データを学界や一般に提供しているほか、アプリケーションを通じてデータの分析結果を提供している。

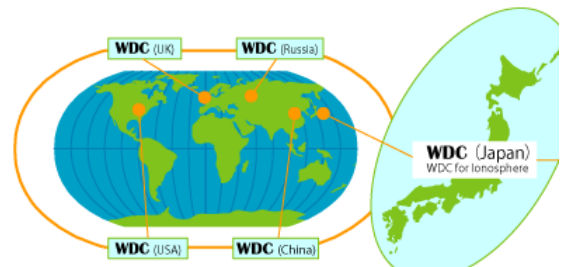
## ホームページによるデータの公開

- 宇宙環境情報(太陽電波、放射線データ等)
- SMILES衛星大気観測データ
- Pi-SAR2(航空機搭載高性能合成開口レーダ)観測データ
- 長波標準電波の電界強度
- サイバー攻撃統計情報
- テラヘルツデータベース

これ以外に有償または共同研究ベースでデータを提供

## オープンデータの取組

世界科学データシステム(国際科学会議(ICSU)/WDS)の事務局を支援

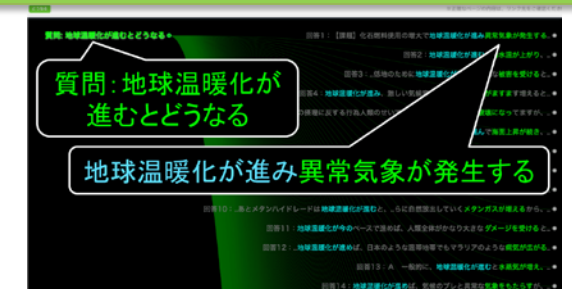


World Data Center for Ionosphere and Space Weather

## データ分析結果を提供するアプリケーション



対災害SNS情報分析システム  
DISAANA



大規模Web情報分析システム  
WISDOM X



多言語音声翻訳  
アプリVoiceTra

宇宙・地球から脳・微生物、実環境、社会活動といった多種多様な対象を観測・測定し、データ化する技術の研究開発と、これらデータを分析して利活用を推進する研究開発を連携して行っているが、さらなる社会的課題の解決や新たな価値創造につなげるためには、例えば次のような課題が挙げられる。

- 科学データの取得には、機器、装置の特性等の専門的な知見が必要であり、データの前処理・解析においても各分野の専門的な知見の活用が必要。
- 専門性の高いデータは分野に限定的な利用に留まっている事例が多いため、データを最大限利用するためには、誰でも利用できるデータ体系や分野を超えたデータ連携が必要。
- ニーズとシーズを結び付ける有用なデータベースの構築には、社会実装の方向性の明確化が必要。
- 取得したデータの学术界、産業界・公共への提供のあり方(制度、システム)の検討が必要。
- 大学、研究機関の有するデータとの連携による科学技術研究の推進と、産業界が有するデータの連携によるNICT研究開発成果の最大化に向けた検討が必要。
- データの利活用にあたっては、セキュリティ、プライバシー、権利関係等についても制度的、技術的な検討が必要。



# 異分野データ連携の課題検討

～ スマートIoT推進フォーラム 異分野データ連携プロジェクトの報告(検討中) ～

# スマートIoT推進フォーラム／異分野データ連携プロジェクト

**IoT推進コンソーシアム**  
 会長：村井純（慶應義塾大学環境情報学部長兼教授）  
 IoT・ビッグデータ・人工知能時代に対応し、企業・業種の枠を超えた産官学での利活用の促進

- 平成27年10月23日、民主導の組織として設立。
  - 技術開発、利活用、政策課題の解決に向けた提言等を実施。
- ← 総務省、経済産業省 等の協力

会員数3,007社（平成29年1月31日現在）

**技術開発WG**  
 （スマートIoT推進フォーラム）  
 座長：徳田英幸（慶應義塾大学教授）  
 ネットワーク等のIoT関連技術の開発・実証、標準化等

**先進的モデル事業推進WG**  
 （IoT推進ラボ）  
 先進的なモデル事業の創出、規制改革等の環境整備

**IoTセキュリティWG**  
 IoT機器のネット接続に関するガイドラインの検討等

**データ流通促進WG**  
 データ流通のニーズの高い分野の課題検討等

会員数2,206者（平成29年1月31日現在）

**技術戦略検討部会**  
 部会長：森川博之（東京大学教授）

- 産学官の今後の戦略の策定や具体的なプロジェクト組成、テストベッド活用ノウハウの共有、国際標準化活動の推進を実施

- 技術・標準化分科会
- テストベッド分科会
- IoT人材育成分科会

- 国内外の動向把握と技術・標準化戦略、普及展開戦略の検討 等
- 技術実証・社会実証を促進するテストベッドの要件とその利活用促進策の検討 等
- IoTの活用等に必要な専門知識の要件に関する検討、技術開発人材等の育成の推進 等

**研究開発・社会実証プロジェクト部会**  
 部会長：下條 真司（大阪大学教授）

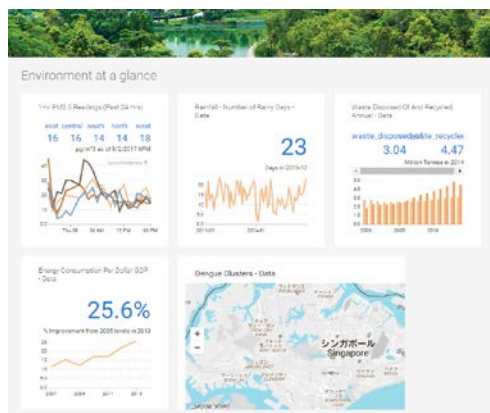
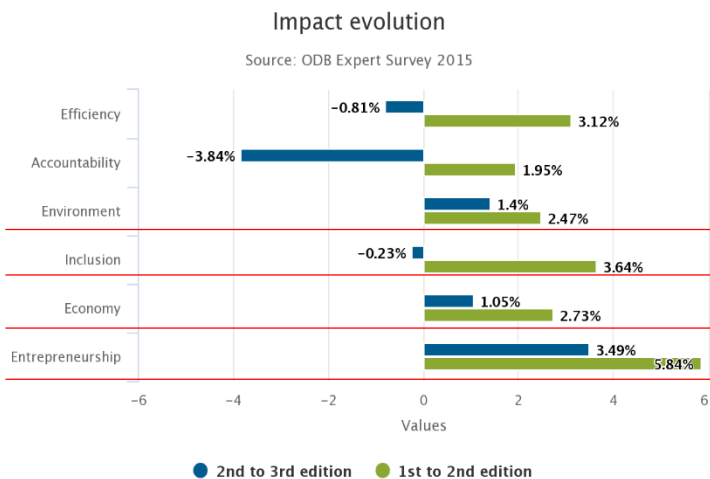
- 各プロジェクト成果の情報共有、对外発表。また、具体的な検討結果を技術戦略検討部会を通じ国際標準化へ向けて議論を展開

- 自律型モビリティプロジェクト
- スマートシティプロジェクト
- 身近なIoTプロジェクト
- 異分野データ連携プロジェクト**

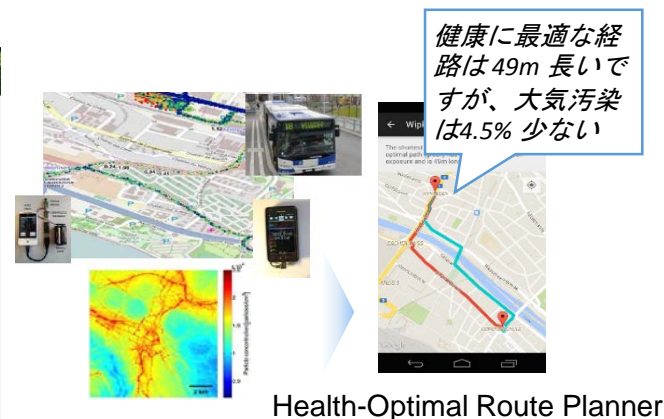
- 自律型モビリティシステムの早期実現に向けた技術開発、実証 等
- スマートシティの社会実証に向けた技術、課題の検討 等
- IoTサービス普及の課題や、生活に身近なIoTの社会実証によるリファレンスモデルの構築 等
- 異分野ソーシャルビッグデータの横断的な流通・統合を行うための課題の検討 等**

# 異分野データ連携の背景

- G8オープンデータ憲章(2013)や各国のオープンデータ政策を受け、**政府のデータポータル上で多種多様なオープンデータを積極的に整備**  
 グローバル・オープンデータ・インデックスの調査(2015)で、122カ国中1位は台湾、日本は31位
- オープンデータの世界的な潮流の中で、**様々なオープンデータを分野横断的に利活用し、国際的な環境問題の解決や住民参加による社会的課題の解決に向けた活動が活発化**
- **オープン・スマートシティに代表されるように、IoTデータを利活用しエネルギー・交通など都市リソースの最適化や災害時の安心・安全の保障、住民の生活品質(QoL)の向上に資する付加価値の高いサービス開発に期待**



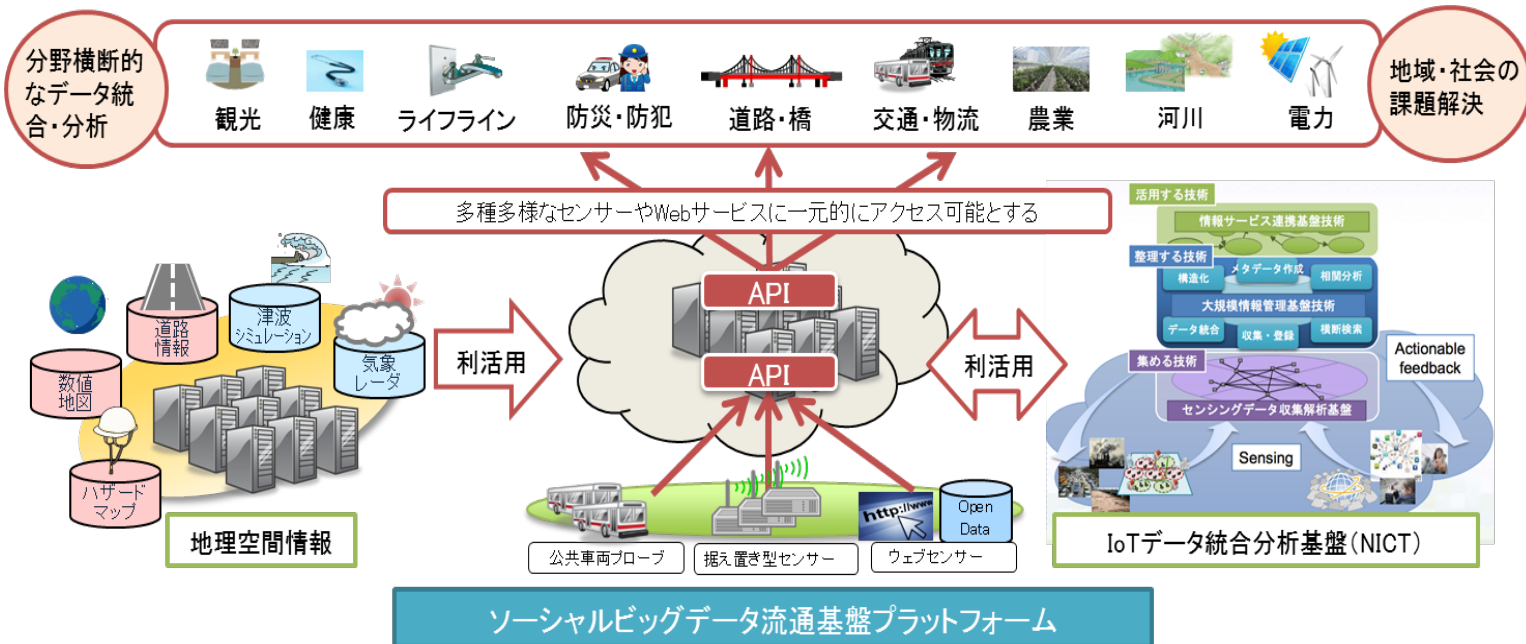
**Data.gov.sg:** シンガポールの気象や大気汚染、エネルギー、交通など様々な動的オープンデータをAPIで提供



**OpenSense/OpenSense II (スイス):** 公共機関やクラウドソーシングによりチューリッヒ市内の大気汚染データを収集・共有し、健康対策等に応用

# スマートIoT推進フォーラム 異分野データ連携プロジェクト

- IoTデバイスやウェブ等からのデータ、国・地方自治体が公開するオープンデータや利活用範囲が広まりつつあるG空間データ、さらには参加者自らが保有するデータとの新しいデータ連携を目指して、異分野ソーシャルビッグデータの横断的な流通・統合を行うための課題を検討
- 異分野データ連携の在り方について、基盤技術、社会実装の両面から課題を整理・体系化し、技術報告書の公開等を通じた提言を行う
- 研究開発・社会実証プロジェクト部会に属する3プロジェクトの1つ。産学官から21機関37名のメンバーが参加（平成28年12月時点）



# ケーススタディ(1)

## ● 実空間データの分野横断的利活用による環境問題対策支援 (NICT)

- ✓ センシングデータ、科学データ、ソーシャルメディアデータ等を対象に、**時間・空間・概念的な相関性を分析し分野横断的な検索や予測を行う技術を開発**
- ✓ **ゲリラ豪雨や大気汚染等の環境問題**による様々な被害発生のリスク分析に利用

## ● スマートシティを実現するソーシャルビッグデータ流通基盤 (慶應大学SFC)

- ✓ IoTセンサからウェブまで、**多種多様なセンシングデータを透過的に取得・流通させる仮想センシング基盤を開発**
- ✓ オートモーティブセンシング (ごみ収集車による環境データ収集) や、市職員による参加型センシングによる**ごみ資源情報等の収集・整理に活用** (みなレポ)

## ● G空間(地理空間)情報アーカイブ (東大CSIS)

- ✓ 官民が保有する**G空間関連データを共有・提供**し、データを円滑に組合せて利活用する**オープンデータ・プラットフォームを構築**
- ✓ **G空間情報センターによる利活用サービスの提供**:
  - ① G空間情報の流通支援、
  - ② 政府・自治体向け「情報信託銀行」、
  - ③ G空間情報の研究開発、
  - ④ 災害対応情報ハブ、
  - ⑤ G空間オープンソースハブ

総務省情報通信審議会 諮問第22号「新たな情報通信技術戦略の在り方」中間答申 (H27.7)

交通・気象・人の活動状況等をリアルタイムに収集・解析

衛星から地球環境全体をミクロに観測

ビッグデータ観測

環境負担(PM2.5やCO<sub>2</sub>等の排出)が最小となるように最適な配送経路を選択

次世代レーダー、住民のソーシャル情報等を組み合わせた確かなゲリラ豪雨予測等

鉄道に応じて最適経路を選択

ゲリラ豪雨早期探知と連動した交通リスク予測に基づく地図ナビゲーション

仮想センシング基盤SOXFire

IoT Sensor Publish

Crowd Sensor Retrieve

Web/SNS Sensor Publish

Virtual IoT Sensor Subscribe

Virtual Crowd Sensor Subscribe

Virtual city context sensor Subscribe

Virtual WEB Sensor Subscribe

App. App. App.

Agent for analyzing data

● meta node ● data node

運用検証

防災 減災 農業 交通 地域活性化

避難誘導 流通 観光

データの配信

情報のダウンロード、API

機能開発

キーワード、地図による検索 情報の重ね合わせ 座標、ファイル形式の変換 有償データの代理販売 データの評価活用事例の紹介

G空間プラットフォーム

データの登録

情報の登録 登録支援

データの収集

救護地図 道路情報 避難所地図 ハザードマップ タンク監視 航空写真 津波シミュレーション 気象レーダー カーブプロブ 携帯電話

基礎データ 静的データ 動的データ

# ケーススタディ(2)

## ● 大規模ドライブレコーダデータによる運転者指向サービス (東大生産研)

- ✓ 交通・物流事業者から数千台、数千運転者規模のドライブレコーダデータを収集。ドライバーの運転操作の特徴を抽出し運転経歴や事故履歴との相関を分析することで、**運転者管理**や**運転者教育**に活用
- ✓ さらに、特定の運転操作が頻発する道路の箇所を発見・可視化することで**道路の潜在的なリスク**を発見し、**運転者への注意喚起**や**道路改善**につなげる



## ● 「はたらく車」走行データによる自治体業務の高度化

 (ゼンリンデータコム)

- ✓ 公用車数百台から**走行データ**(OBD2等)を収集。公用車のカーシェアリングに向けた**稼働率**や**走行範囲**の分析・管理や、道路保全業務に必要な**路面性状の経時変化**の取得簡易、**防犯パトロール経路の最適化**等に活用
- ✓ 走行データと活用モデルをオープン化し、活用モデル開発の促進や、他自治体でのモデル活用により**都市間での比較・評価**を行い自治体経営に還元



## ● 実空間データのデータ形式やスキーマの共通化

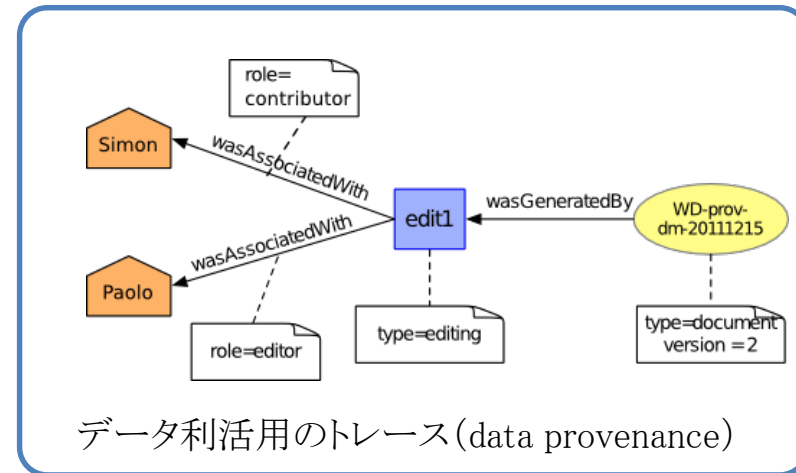
- ✓ 実世界の様々な事象・現象(イベント)を表すデータモデルと、それらの時空間的・意味的な関係に基づくデータ操作に基づく共通スキーマの定義
- ✓ データプラットフォームによる時空間情報の表現形式(測地系等)やデータ構造(JSON, XML等)の相互変換、画像・PDF等への時空間情報の付与
- ✓ 巨大な実空間情報を手軽に扱えるようにすべく、データの時空間的な内挿化や分割を高速処理する技術の開発

## ● データ利活用におけるスケーラビリティの向上

- ✓ 多様化・豊富化するデータを横断的かつリアルタイムに処理できるよう、エッジ、フォグ、クラウドコンピューティングを効果的に連携させ処理全体を最適化したり、アプリケーションとデータ流通基盤の間でデータ量や更新頻度等を動的にマッチングする技術の開発
- ✓ 従来のデータ利活用技術(収集、統合、検索、分析、可視化等)を、マルチスケール、マルチモーダル、マルチメディアなデータソースに対応できるよう拡張・強化
- ✓ データプラットフォームに、データ本体だけでなく、データへのリンクを登録しリンク経由でアクセスできるようにすることで、ビッグデータの大規模なリアルタイム流通を実現

## ● 安心・安全なデータ利活用のための技術

- ✓ センサーの種類や計測方法、キャリブレーション方法、検出限界値など、**データ取得・生成に関する品質をメタデータに追加**
- ✓ 第三者から入手したデータの生成・加工・流通経路 (provenance) を明らかにする**トレーサビリティ技術**や、データが本物であり一切の改竄を含まないことを証明する**真贋性保証技術の開発**
- ✓ **プライバシー保護**に加え、データのリアルタイム性に配慮した**秘匿化手法の開発**
- ✓ データの縮尺や鮮度、用途等によって異なる**情報の有効範囲**について、データの提供者、利用者間で**コンセンサスを形成する仕組みが必要**
- ✓ データプラットフォームにデータへのリンクのみを登録し、検索や閲覧は可能にするがデータ本体の譲渡は選択可能にすることで、**データの利活用拡大と知財保護を両立**





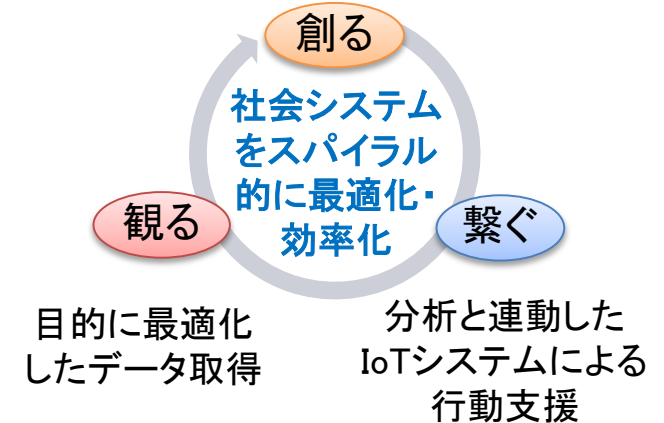
## ● 社会システムとデータ連携基盤の統融合

- ✓ 都市や地域全体でデータを有機的に流通させ、データ利活用を通じた社会システムの最適化・効率化を実現
- ✓ エネルギーや交通など都市リソースの最適化だけでなく、市民生活の品質向上(QoL)や安心・安全の保障など付加価値の高いスマートシティサービスの開発
- ✓ 様々な組織や団体から提供される多種多様なデータをワンストップで検索・入手できるデータ・マーケット・プレイスや、クラウド上でデータ・マッシュアップを行える環境を整備。災害時だけでなく平時も利用できるようにすることが重要

## ● データ流通から価値流通へのシフト

- ✓ 国、自治体、企業、個人など様々なステークホルダーの間で、データの提供とその利活用による対価の関係を構築し相互にインセンティブを持たせることで、**実空間に関するデジタル情報を「財」として課金を伴って流通させるプラットフォームを実現**

異分野IoTデータの横断的分析



データ流通

価値流通

情報第一次産業

自然界に働きかけて直接に富(情報)を取得する産業。実空間からのセンシング。需要が明らかであればより多くのセンシングが行われる。

情報第二次産業

第一次産業が採取・生産した原材料(情報)を加工して富を作り出す産業。知の生産を行うソフトウェア、情報を可視化するアプリ。需要が明らかであればより多くの加工が行われる。

情報第三次産業

左記以外の産業で、商品やサービスを分配することで富を創造する。小売業やサービス業。IoT的な情報を扱う第三次産業は現存しない。

# 課題解決指向なデータ利活用に向けて

## ● データ駆動型の課題解決

- ✓ 自治体に散在している**地域情報**は地域政策立案の基礎資料としての価値を有しており、**官民で蓄積されている実空間情報と組み合わせ**て有効活用すべき(オープンデータ2.0、官民データ活用推進基本法など)
- ✓ **データ利活用の目標設定や評価・改善**においては、業務全体の最適化・効率化など、**課題解決による価値を重視することが重要**
- ✓ データ保有者が抱える課題から始めて信頼関係を構築し、然る後に**データ保有者が気づいていなかった利活用へと発展させることが重要**。そのためには、データ保有者の課題に興味を持ち、データ解析を行うことができる**研究機関とのマッチング**をサポートする**組織の介在が必要**。研究機関、メーカーとその顧客企業の3者が密に連携を取り、相互にメリットのある連携を組むことが**成功の鍵**

## ● データ利活用を介した協働の促進

- ✓ コミュニティ全体で積極的にデータを共有し課題解決に取り組むことが**重要**(ガバメント2.0等)。データの収集、共有、分析を共同で行うことで**課題に対する集合知を形成し、課題解決の効率化や未知の課題への迅速な対応を促進**



# 今後の検討に向けた論点の一例について

産学官の様々な機関が取得・蓄積するデータは、現在、加速度的に増加している。これらを融合的に活用できれば、学術、産業、公共といった分野で、新たな価値の創造につながることを期待される。これを踏まえると、以下の論点が考えられるのではないか。

- 出来るだけ多くのデータを共有し、社会全体としてイノベーションを創発するためには、データの共有化を進める方策を検討し、産学官連携のあり方を含め、データ連携の形を検討すべきではないか。
- その際、独自にデータを利活用する方式や、外部とも連携する方式など、分野・データ毎の特性を踏まえた検討が必要ではないか。
- データをオープンにして外部機関によるデータの利活用を促進するために、データ提供のあり方やテストベッド環境のあり方を検討すべきではないか。
- データの利活用にあたっては、セキュリティ、プライバシー、権利関係等についても制度的、技術的な検討が必要ではないか。
- これら高度なデータの利活用に対応する人材の育成が必要ではないか。