

4-3 xData プラットフォームによるデータ連携スマートサービス開発

4-3 *Developing Cross-data Smart Services using xData Platform*

伊藤 禎宣

ITO Sadanori

統合ビッグデータ研究センターでは、ユニバーサルコミュニケーション分野のスマートデータ利活用基盤技術の研究開発として、データ連携分析によるスマートサービスの開発基盤を構築する取組を続けている。xData プラットフォームは、実空間の多様なセンサデータを収集して、分野横断的に相関を発見・学習・予測するデータ連携分析の基盤技術を提供し、分析結果を用いたスマートサービス開発を行うための基盤である。分析やサービス開発を効率化するため、API や情報資産の整備を進めている。本稿ではこの取組状況について述べる。

The NICT Big Data Integration Research Center continues to research and develop the platform for smart service development through cross-data analysis as a smart data utilization platform technology in the field of universal communication. The xData platform is being built to discover correlated data and predict risks by use of data mining and machine learning with collecting various real-space sensing data. We are also implementing xData APIs and xData information assets to promote the use of NICT's analysis technology, development of smart services, and technology transfer. This paper describes the status of these efforts.

1 はじめに

近年、気候変動による気象災害の頻発化、交通課題の増加、地域サービスの衰退、インフラ維持管理の負担増など、ICT の活用によって解決が期待される様々な社会課題の増加が指摘 [1] されている。総務省では、防災や都市計画などの具体的な課題を抱える地方公共団体などによる、データを活用した効果的な政策立案や住民サービスの向上を推進 [2] している。

xData プラットフォームは、実空間の多様なセンサデータを収集し、分野横断的に相関を発見・学習・予測するデータ連携分析の基盤技術を提供する。プラットフォームの利用者は、同基盤技術やユーザー開発環境を利用することでスマートサービスの開発を効率的に進めることができる。

ICT による課題解決のための参加型開発基盤としては、データサイエンティストが様々な分析課題に取り組むデータ分析コンペティションプラットフォームである Kaggle [3]、プログラマがアプリケーション開発に取り組むソースコードホスティングプラットフォームである GitHub [4]、プログラマが情報技術の課題解決を共有するプログラミング技術のナレッジプラットフォームである Stack Overflow [5] などが知られている。

xData プラットフォームは、データ連携分析を中心とした NICT の情報資産と、利用者の独自データやノウハウを活用し、共創による課題解決を推進することを目指している。本稿では、xData プラットフォームの取組とその活用事例について紹介する。

2 xData プラットフォーム

2.1 プラットフォームの概要

xData プラットフォームは、リモートセンシングを始めとする様々な分野のセンシングデータやソーシャルビッグデータから、実空間のイベント情報を抽出し、横断的に利用可能な形式で集積するイベントデータウェアハウス (EvWH) と、それらの相関を発見・予測するデータマイニングや機械学習などのデータ連携分析機能を実装した API (xData API)、これらを活用したアプリケーション開発を支援するツールによって構成されている (図 1)。イベントデータウェアハウスには、NICT リモートセンシングデータを始め、環境、交通、健康等の様々な分野のセンシングデータから抽出した数百 TB のイベントデータがアーカイブされている。xData プラットフォームには、これらのイベントデータを収集し、データ連携分析を行い、その結果を

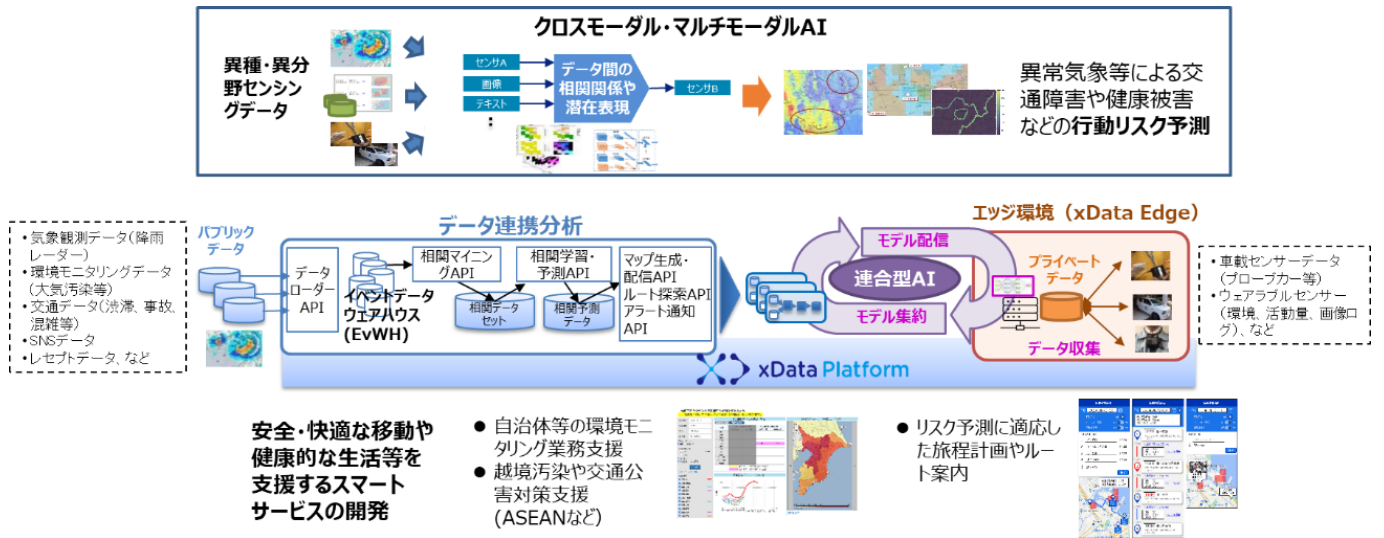


図1 xData プラットフォームの概要

マップ形式などでアプリケーション向けに配信するためのAPIなどが実装されている。プラットフォーム利用者は、これらのAPIを用いて、様々なデータを分野横断的に収集して組み合わせたり、関連データの発見や予測を行ったり、予測結果を使ってアラート通知やルート案内などの行動支援を行うスマートサービスの開発を行うことができる。

また、xData プラットフォームの利用者が、独自にデータを収集したり、データ連携分析モデルの調整や予測結果データの加工・配信をカスタマイズしたりできるようAPIや開発ツールを仮想化コンテナ (Docker) にパッケージ化した開発環境 xData Edge も提供している。xData Edge をプラットフォーム利用者のサーバ等に配備し、利用者のデータやノウハウを活用した共創による課題解決を可能にしている。

本節では、xData プラットフォームの構成要素の実装について説明する。

2.2 xData API

xData API は、xData プラットフォームによるデータ連携分析のフレームワークに沿ってアプリケーション開発を行うための共通機能を提供する。現在提供している主要なAPIを表1に示す。主要なAPIとして、イベントデータウェアハウス (EvWH) にイベントデータの収集・抽出・変換を行うデータローダーAPIや、イベントデータの時空間統合や相関ルール発見を行う相関マイニングAPI (FP-growth手法、SHUIM手法など)、時空間相関パターンの機会学習を行う相関学習・予測API (CRNN手法、LGBM手法など)などの分析APIがある。このほか、予測結果からGISデータ (リスクマップ) を配信するマップ生成・配信API、リスクマップに基づいてルート探索を行うAPIやアラ

表1 主な xData API

名称	概要
データローダーAPI	多様なイベントデータをEvWHへ格納・取得
相関マイニングAPI	FP-growth手法 FP-growthアルゴリズムによる頻出アイテム集合の発見
	SHUIM手法 Spatial High Utility Itemset Mining (SHUIM)手法[9]による局所的に高効用な頻出アイテム集合の発見
相関学習・予測API	CRNN手法 Convolutional Recurrent Neural Network (CRNN)手法[7]による時空間相関パターンの学習・予測
	LGBM手法 Light Gradient Boosting Machine (LightGBM)手法による時系列変化パターンの学習・予測
マップ生成・配信API	イベントデータ配信・マップ配信などのデータ配信
ルート探索API	リスクマップに基づき、指定されたリスク許容値を満たすルートを探検
アラート通知API	リスクマップを監視し、指定された条件を満たした時に通知を送信

ート通知を行うAPIなどが実装されている。分析APIが実装している基盤技術については、本特集号 **4-1**「スマートデータ利活用基盤技術の概要」及び **4-2**「Multi-modal and Crossmodal AI for Smart Data Analysis」で解説している。また、APIのリファレンスやサンプルの整備 [6] も進めている。

これらのAPIはWeb APIとして実装され、JSON-RPC v2.0のデータ形式で、xData Web API endpointのURLにHTTPS POSTすることで呼び出すことができる。利用申請により発行されるAPIKeyと秘密鍵ペアによる認証のほか、リクエスト元のIPアドレスによるアクセス制限も行っている。

2.3 システム構成

xData プラットフォームのシステム構成を、図2に示す。主な構成要素としては、イベントデータウェアハウス (EvWH)、Web経由でのAPIアクセスを提供するWeb API endpointとgateway、各種APIの実行モジュール群、外部のデータ提供元からイベントデータを収集 (Extraction-Transform-Load: ETL) するデータローダー、Pythonプログラム統合開発環境 Jupyter-

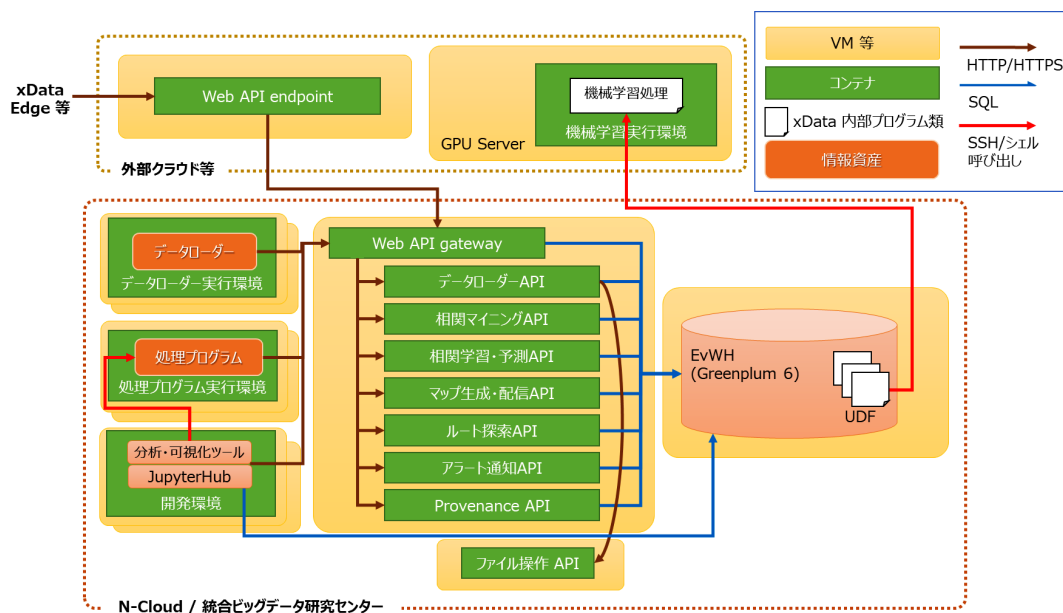


図2 xData プラットフォーム システム構成図

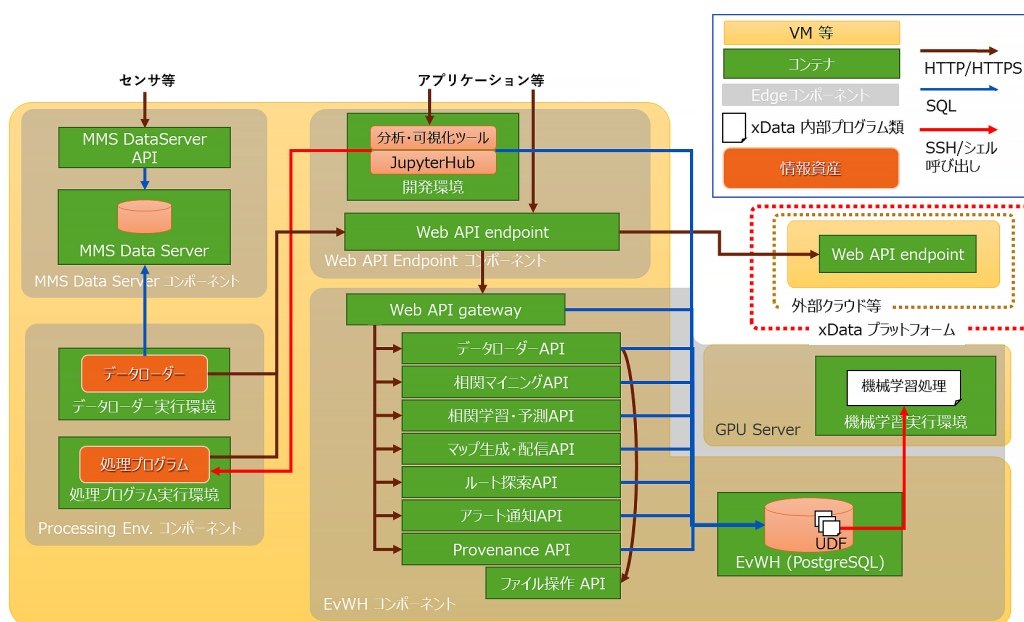


図3 xData Edge システム構成図

Hub 等の開発ツールなどが挙げられる。EvWH では、In-Database 分析技術 [9] により、トランザクション生成や地理空間計算の高速処理を行う。分析 API の処理は、EvWH のユーザー定義関数 (User-Defined Function: UDF) として定義され、大量のデータを扱う処理を DB 内で効率的に実行できるようにしている。また、外部からのアクセスを受け付ける Web API endpoint は、WAF (Web Application Firewall) サービスを利用しセキュアな外部接続を実現している。

また、利用者の開発環境である xData Edge の構成を図 3 に示す。xData Edge は、xData プラットフォームと同じく、EvWH、API 実行モジュール群、データ

ローダーや開発ツール (JupyterHub 等) で構成される。利用者は、Ubuntu OS 用のインストーラパッケージを使って、自分の計算機やクラウド環境に xData Edge を導入することができる。xData Edge からは、Web API endpoint を通して、データ収集や分析処理を xData プラットフォーム側と連携して実行することができる。例えば、プラットフォーム側から学習済みモデルを取得したり、プラットフォーム側の計算機資源を使って、モデル学習や予測を実行したりすることができる。また、個人情報などを含むプライベートデータの処理を xData Edge 内で実行することも可能である。

4 スマートデータ利活用基盤技術

xData プラットフォームと xData Edge を構成する各機能モジュールの実行に必要なバイナリや設定ファイル等は、コンテナ型仮想化技術 (Docker) を用いてコンテナにカプセル化している。これにより、多様なユーザー環境に対し、機能モジュールのデプロイや

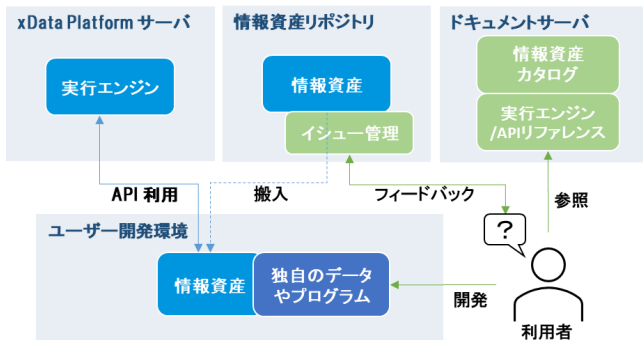


図4 アプリケーション開発モデル

ロールバックを容易に行うことができる。また、xData Edge では、関連するコンテナをグループ化したコンポーネントの単位(図3の灰色部参照)でデプロイを行うことができる。例えば、利用者がマルチメディアセンシングの情報資産(後述)を利用する場合、MMS Data Server コンポーネントをデプロイすることで、関連するコンテナ(MMS Data Server API や MMS Data Server など)をまとめて xData Edge に組み込むことができる。これにより、利用目的に応じた xData Edge の構成変更を簡単に行うことができる。さらに、機能モジュールのビルド・テスト・リリースを自動化する、CI/CD (Continuous Integration / Continuous Delivery) パイプラインにより、xData Edge に組み込まれた機能モジュールを常に最新版に保つことができる。

2.4 アプリケーション開発モデル

xData プラットフォームのアプリケーション開発モデルを図4に示す。

「情報資産」は、大気汚染予測、交通渋滞予測といった幅広い応用分野ごとに、処理プログラムや予測モデル、データセットなどをパッケージしたものである。情報資産カタログ(図5)では、情報資産の概要、分析手法の技術的特徴、動作環境、応用例などを紹介しており、利用者は目的に合う情報資産を選択することができる。現在、環境品質短期予測、煙霧越境汚染被害予測、マルチメディアセンシング、移動環境リスク予測、環境睡眠品質予測などの情報資産が掲載されている。図6に、環境品質短期予測情報資産による処理手順の例を示す。この情報資産では、大気環境のオープンデータやウェアラブルセンサで取得したデータなどを、「データローダー」で収集、変換して EvWH へ格納



図5 情報資産カタログ

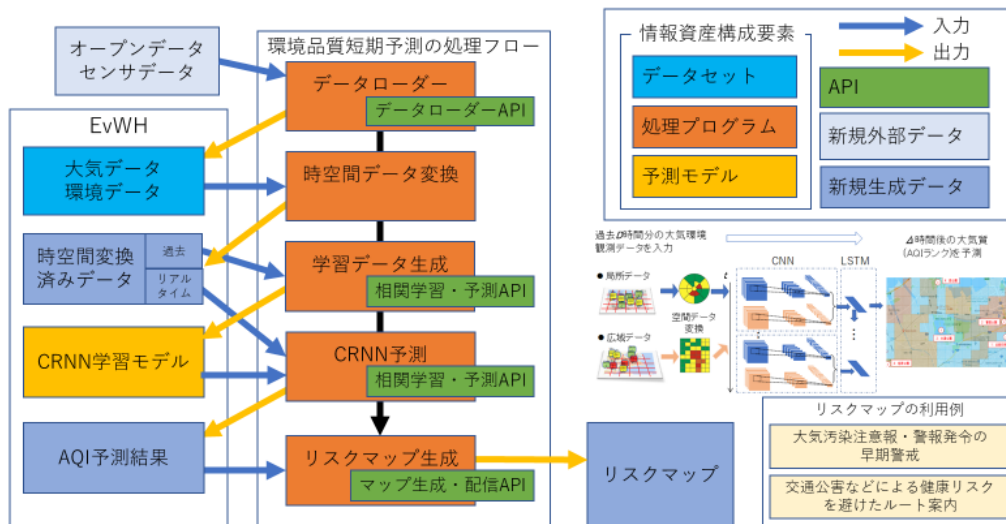


図6 環境品質短期予測情報資産による処理手順の例

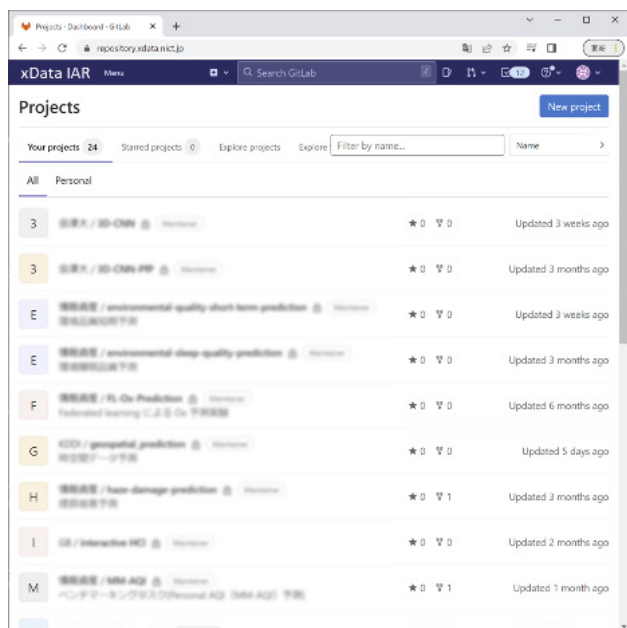


図7 情報資産リポジトリ

し、「学習データ生成」と「CRNN 予測」によりデータの時空間的な相関パターンを学習・予測して、生成された「AQI 予測結果」データから「リスクマップ生成」で GeoJSON 形式のマップデータを出力する。利用者は、例えば、情報資産に含まれるデータローダーや予測モデルを編集することで、独自に収集したデータを利用した予測のカスタマイズが可能となる。

「情報資産リポジトリ」(図7)には、情報資産を構成する処理プログラムや予測モデル、サンプルデータ、実行手順書などのドキュメントが整備されている。情報資産リポジトリでは、情報資産の保存や閲覧、バージョン管理のほか、課題管理も行われている。

図4の「実行エンジン」は、情報資産の実行基盤(ランタイム)である。実行エンジンには、情報資産の実行に必要な EvWH と API 実行モジュール、その他の必須ライブラリなどがパッケージ化されている。

利用者は、情報資産カタログからベースとなる情報資産を選択し、「ユーザー開発環境」に情報資産を搬入し、データセットや予測モデル、処理プログラム等のカスタマイズを行うとともに、独自のデータセットやサービス等と組み合わせた技術検証や実証実験を実施することができる。編集した情報資産を利用者が継続的に使用したい場合、技術移転契約を締結し、実行エンジンとあわせて利用者自身の環境へ搬出することで、xData プラットフォームとは独立して実行することが可能となる。

3 xData プラットフォームの活用事例

3.1 プラットフォーム利用の流れ

利用者は、以下の手順に沿って xData プラットフォームを利用する。

- 利用開始手続き
 - 共同研究契約を締結し、情報資産リポジトリのアカウント発行、WebAPI キーの発行などの利用開始手続きを実施する。
- xData Edge ユーザー開発環境のインストール
 - xData Edge インストーラパッケージをダウンロードして、利用者のサーバ環境やクラウド環境にインストールし、アカウント等の初期設定を行う。
- 情報資産の搬入
 - 情報資産リポジトリにアクセスして、該当の情報資産をダウンロード(クローン)し、xData Edge へ搬入する。
- カスタマイズ
 - 利用者の開発環境において、独自データのロードや予測モデルの変更など、情報資産の改修作業を行う。必要に応じて、変更部分を情報資産リポジトリにアップロード(プッシュ)する。
- カスタマイズ済み情報資産の搬出と独自サービスへの組み込み
 - 技術移転契約を締結した上で、編集した情報資産と実行エンジンを利用者環境に搬出・導入し、独自のサービスやアプリケーションに組み込む。

3.2 環境品質短期予測に基づく環境基準監視業務支援

国内の各自治体では、大気汚染防止法に基づく大気環境の監視業務を実施しており、光化学オキシダント濃度が基準値を超過した場合に、光化学オキシダント注意報・警報を発令する。そこで我々は、環境モニタリング事業者と共同で、環境品質短期予測情報資産から派生した光化学オキシダント注意報予測情報資産を活用した自治体の環境基準監視業務支援の取組を実施している。図8に、環境基準監視業務支援のアプリケーション例を示す。地域や測定局を指定することで、濃度推移グラフに過去の観測値と予測値を重畳表示するほか、注意報の発令確率を地図上に表示することができる。

事業者は、ユーザー開発環境を利用し、自治体ごとの収集データや監視業務フローに応じて、前駆物質の観測データを追加で入力したり、大気測定局のグルーピングや予測時間の変更など、情報資産のカスタマイズを実施する。このカスタマイズされた予測モデルを

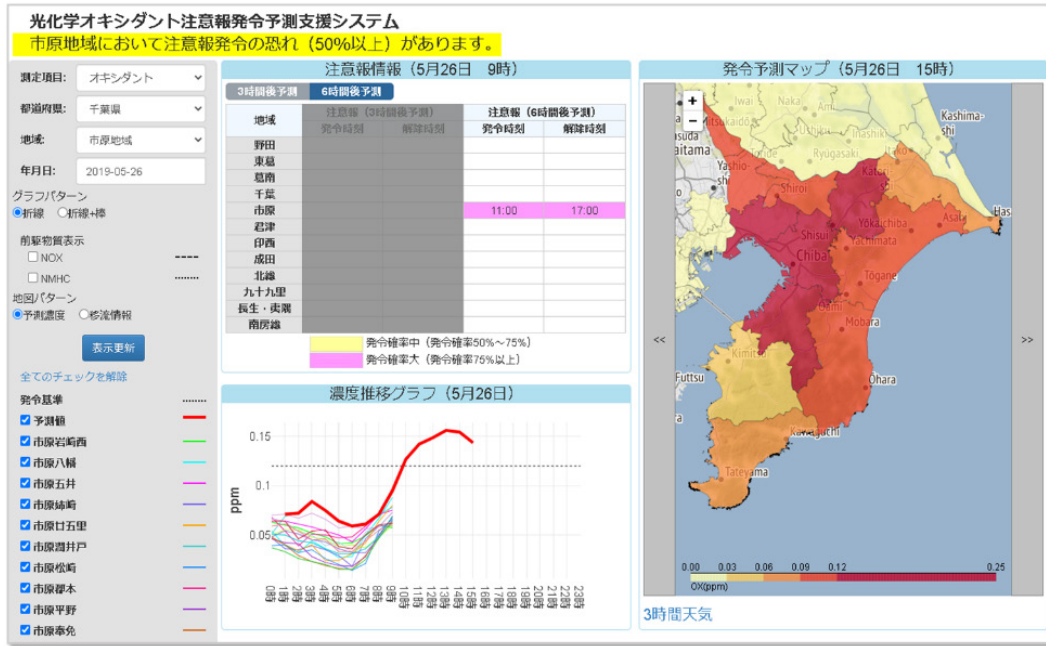


図8 大気環境監視業務支援アプリケーション

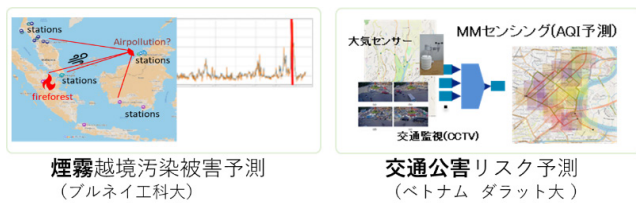


図9 ASEAN 地域の環境問題対策への応用例 (ASEAN IVO プロジェクト)

アプリケーションに組み込み評価実験を行ったところ、千葉県対象の注意報発令レベルの光化学オキシダント予測で92%(8時間以内)の予測精度を実現した。また、自治体の環境基準常時監視業務支援への受容性に関する評価を行ったところ、注意報発令の早期警戒(警戒解除)への利用可能性がアンケート調査により確認された。さらに対象となる自治体を拡大し、カスタマイズと性能評価を実施している。

事業者がxDataプラットフォームの情報資産を活用することで、対象期間や自治体の変更はマニフェストファイル(観測局情報などの設定ファイル)から設定でき、検証結果の参照も容易になるため、過去の大気環境データを利用した予測精度検証の作業を効率的に進めることができた。また、事業者が持つ地域ごとの気象特性に関わるノウハウを取り入れ、予測精度を向上させるためのスクリーニング処理を行う独自プログラムの追加や、アプリケーションと連携した総合的な判断(風向による移流の評価など)を効果的に行うことができた。

3.3 ASEAN 地域の環境問題対策への応用

環境問題が深刻な ASEAN 地域の4研究機関(ブルネイ工科大、ベトナム ダラット大、シンガポール国立大、フィリピン イザバラ州立大学)とNICTが共同で行った ASEAN IVO プロジェクト“Reusable, Shareable, and Transferable Smart Data Platform for Collaborative Development of Data-Driven Smart Cities”(令和元~3年度)では、現地で収集したデータを用いて情報資産をカスタマイズし、煙霧越境汚染被害対策(ブルネイ工科大)やマルチメディアセンシング[10]による交通公害対策(ベトナム ダラット大)などに応用した(図9)。例えば、煙霧越境汚染被害予測では、xDataプラットフォームの環境品質短期情報資産を派生させ、ブルネイ及び周辺国の環境観測データを様々な情報源から収集するようデータローダーを改修するとともに、収集されたデータセットに合わせた予測モデル(CRNN)のチューニングを行い、ASEAN 地域の大気汚染移流を予測できるようにした[11]。さらに煙霧越境汚染被害予測の情報資産の一部を MediaEval 国際ベンチマーキングタスク[12]に公開し10チームの研究者らが性能改善を競うなど、研究者を巻き込んだ情報資産の性能改善を効果的に実現した。これらの結果、ASEAN IVO プロジェクト終了評価では、全8プロジェクト中第2位の高い評価を得た。

ブルネイ工科大での煙霧越境汚染被害予測の研究成果は、同大学での越境汚染災害の研究や地元政府の公衆衛生危機管理への応用検討に発展した。ベトナムダラット大の交通公害リスク予測の研究成果は、同大学とダラット市インテリジェントオペレーションセン



図 10 Smart Sustainable Mobility ハッカソン

ターでの安全な住環境や観光支援、交通対策の研究開発に引き継がれた。現地の研究機関が新たな技術を開発し、開発環境問題対策に向けた研究開発の継続的な実施につながっている。

3.4 移動環境リスク予測を用いたカーナビアプリのハッカソン

移動環境リスク予測情報資産を活用したスマートサービスを具現化するための Smart Sustainable Mobility ハッカソン [13] を、平成 31 年 2 月に東京都で実施した(図 10)。異常気象による交通障害リスクの予測データを用いてルート案内アプリを開発するもので、ITS (Intelligent Transport Systems) 事業者、技術者、大学教員など約 20 名が参加し、豪雨や豪雪に伴う交通障害の移動環境リスク予測に基づいて、アラートや運転リスクを通知し、それらを回避するルート案内や運転支援を行うカーナビアプリケーションの試作と運転シミュレーターを使ったデモを行った。

xData プラットフォーム API の機能(リスクマップ配信、ルート探索、アラート通知など)を利用し、交通リスクデータを参照しながら、状況に応じたカーナビ画面表示や経路案内を容易にカスタマイズできるナビサービス試作環境を開発した。例えば、最優秀賞を受賞した「S-KNOW」では、移動環境リスク予測の結果に応じて、豪雪時、路上カメラやドライブレコーダーの動画をリアルタイムに表示する機能や、運転スキルにも合わせた運転ガイダンスを行う機能を持つアプリが試作された。参加者へのアンケート調査では、アイデアを迅速に形にできたことが高く評価され、新聞報道(日刊自動車新聞 平成 31 年 2 月 27 日)でも、交通リスクデータ活用をテーマにソフト開発を競うハッカソンを通じ企業の垣根を超えカーナビを考える活動と

して紹介された。

このハッカソン等を通じて開発された移動環境リスク予測に基づくルート案内を応用し、異常気象時の交通障害リスクを回避するカーナビの実証実験を委託研究(課題番号 201)等と連携し実施した。現在、これらの成果の社会実装に向けた検討を進めている。

4 おわりに

NICT 統合テストベッド研究開発推進センターへの xData プラットフォームの技術提供を進めており、サービスレイヤテストベッド(Data Centric Cloud Service: DCCS)として応用開発や実証実験を加速する環境構築を行っている。既に環境モニタリング事業者による光化学オキシダント注意報予測情報資産の応用開発で連携を実施しているほか、環境品質短期予測などの情報資産の公開へ向けた整備を実施している。また、xData プラットフォームと DCCS の間で情報資産をシームレスに移行できるよう、機能モジュールの共通化や情報資産利用規約の策定を実施している。これらにより、今後の応用開発や実証実験を DCCS と協力しながら推進していく。

また、幅広く利用者を受け入れるための環境整備を実施している。xData プラットフォームに関心を持った潜在的な利用者が、事前評価や自己学習をしやすいするため、お試し利用環境の整備や、ローコード開発環境、ビジュアル分析環境などの整備を進めている。

今後、Beyond 5G (B5G) / 6G 時代におけるサイバーフィジカルシステムの構築・活用へ向けて、大容量・低遅延・超多様な IoT データの流通・蓄積・分析・予測を行う B5G / 6G サイバー空間としての機能を拡充していく。

【参考文献】

- 総務省, “令和 4 年版情報通信白書,” <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r04/pdf/01honpen.pdf>
- 総務省, “地方公共団体におけるデータ利活用ガイドブック Ver.2.0,” https://www.soumu.go.jp/menu_news/s-news/01ryutsu06_02000212.html
- “Kaggle,” <https://www.kaggle.com/>
- “GitHub,” <https://github.co.jp/>
- “StackOverflow,” <https://stackoverflow.com/>
- “xData プラットフォーム,” <https://www.xdata.nict.jp/>
- Peijiang Zhao and Koji Zettsu, “Decoder Transfer Learning for Predicting Personal Exposure to Air Pollution,” BigData 2019, pp.5620–5629, 2019.
- Kiran, R. U., Saideep, C., Zettsu, K., Toyoda, M., and Reddy, P. K., “Discovering Partial Periodic Spatial Patterns in Spatiotemporal Databases,” 2019 IEEE International Conference on Big Data (IEEE BigData 2019), Los Angeles, CA, USA, 2019.
- Joseph M. Hellerstein, Christopher Ré, Florian Schoppmann, Daisy Zhe Wang, Eugene Fratkin, Aleksander Gorajek, Kee Siong Ng, Caleb Welton, Xixuan Feng, Kun Li, and Arun Kumar, “The MADlib analytics library: or MAD skills, the SQL.” Proc. VLDB Endow. 5, 12, pp.1700–1711, 2012.

4 スマートデータ活用基盤技術

- 10 Minh-Son Dao, Kazuki Tejima, Tuan-Vinh La, Rage Uday Kiran, and Koji Zettsu, "Improving the Awareness of Sustainable Smart Cities by Analyzing Lifelog Images and IoT Air Pollution Data," 2021 IEEE International Conference on Big Data (Big Data), pp.3589-3594, 2021.
- 11 E. N. Aziz, A. Kasem, W. S. H. Suhaili, and P. Zhao, "Convolution Recurrent Neural Network for Daily Forecast of PM10 Concentrations in Brunei Darussalam," The 6th International Conference on Low Carbon Asia & Beyond, 2020.
- 12 MediaEval 2021, "Insight for Wellbeing: Cross-Data Analytics for (transboundary) Haze Prediction," <https://multimediaeval.github.io/editions/2021/tasks/wellbeing/>
- 13 Ito, S. and Zettsu, K., "Report on a Hackathon for Car Navigation Using Traffic Risk Data," 3rd International Conference on Intelligent Traffic and Transportation (ICITT 2019), Amsterdam, The Netherlands, 2019.



伊藤 禎宣 (いとう さだのり)

ユニバーサルコミュニケーション研究所
統合ビッグデータ研究センター
主任研究員
博士 (知識科学)
ヒューマンコンピュータインタラクション、
ユビキタスコンピューティング

【受賞歴】

- 2016年 建設ロボット研究連絡協議会
第16回建設ロボットシンポジウム
優秀論文賞
- 2015年 建設ロボット研究連絡協議会
第15回建設ロボットシンポジウム
優秀論文賞
- 2008年 マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO2008)
優秀論文賞