プレスリリース

2021 年 4 月 27 日

日立造船、データブリックスの「レイクハウス・プラットフォーム」を採用、

メソドロジックの支援により IoT 基盤と連動するデータ分析・機械学習システムの運用を効率化

※本プレスリリースは、データブリックス・ジャパン株式会社および株式会社メソドロジックの両社がそれぞれ配信しています。重複のご案内となりましたら、ご容赦いただけますようお願い申し上げます。

データ分析基盤コンサルティングを実施している株式会社メソドロジック(本社：東京都港区麻布台、代表取締役社長：大西俊幸、以下「メソドロジック」)は、日立造船株式会社（社長兼 COO：三野　禎男、以下「日立造船」）の、IoT 基盤のデータ分析の課題に対して、米 Databricks 社の日本法人であるデータブリックス・ジャパン株式会社（本社：東京都港区、社長：竹内 賢佑、以下「データブリックス」）が提供する統合データ分析基盤『レイクハウス・プラットフォーム』の導入支援をしたことを発表いたします。
これにより日立造船は、同社の IoT 基盤と連動するデータ分析・機械学習システムの運用を効率化するとともに、データブリックスを用いたデータの利活用を実施していきます。

2018 年より日立造船は、遠隔監視および IoT、ビッグデータ、AI などの ICT 活用の拠点として Hitz 先端情報技術センター「A.I/TEC（エイアイテック）」の運用を開始しました。主力事業となるゴミ焼却発電施設では、遠隔監視・運転支援サービスの先駆けとして先進的な取り組みを展開しており、全社的にデジタル化を加速させるため、事業製品や工場を横断した IoT 基盤の構築を開始しました。日立造船では、IoT と連動したデータ分析、機械学習のシステムの導入運用を検討・実施していくにあたり、以下の課題が挙がりました。

１．大量データの分析に関する不安

従来は、IoT で収集したデータによる特徴量抽出などは、個人のローカル環境で実施してきました。しかし、データ分析者がハンドリングするデータ量の増加に伴い、ローカル環境では対処が困難なケースが発生するようになったため、大規模データを効率的に処理するための仕組みを構築する必要性が高まっていました。

２．データ分析者の育成

データ分析者は案件に応じて様々な手法を選択して分析を行います。部門全体として分析力を向上するためには、分析者がそれぞれの案件で得たナレッジを展開することが不可欠です。しかし、そのためのインフラ整備が

十分になされておらず、効率的なナレッジを共有することが困難な状況でした。そのため、データ分析者は、Python、SQL での分析内容について、ナレッジを効率的に共有できる方法を模索していました。

３．機械学習でのモデル管理・スムーズなデプロイ環境

適切かつ意味のあるデータを分析し、恒常的に価値を創出し続けるためには、分析イテレーションを効率的に繰り返し、デプロイしたモデルについても適切に管理・監視し続ける必要があります。そのため、蓄積したデータを十分に活用することができる、機械学習の CI/CD（継続的インテグレーション/継続的デリバリー）環境が求められておりました。

日立造船様の課題や検討事項	メソドロジックからのご提案
✓ AWS、マルチクラウドサービスの活用 ✓ 大量データ処理時の対応	✓ **Databricks統合データ分析プラットフォーム** マネージドSparkクラスタを軸にして、機械学習を中心としたデータ分析基盤にて、長時間クエリやデータマートの作成処理を実行する
✓ デプロイしたモデルの精度監視・再学習を容易にする ✓ 推論結果と推論実施モデルの紐付け	✓ **ML flowの利用** 機械学習のCI/CD環境を用意、モデルデプロイの様々なデータフローに対応可能 AutoMLにも対応し、拡張性あり
✓ モデルの世代管理・モデル実行結果の再現性確保 ✓ モデルの評価などで過去のある時点のデータを使って分析を実行するためにデータ保存設計が必要	✓ **DeltaLakeのタイムトラベル機能** 過去のスナップショットから任意の時点のデータをクエリ対象として分析を実行可能
✓ データガバナンスの確保 ✓ データアクセス管理、利用者ごとのアクセス制御	✓ **DeltaLakeによるデータ変更監査** Delta Lake のトランザクションログにはデータ変更の詳細が記録され、監査用に変更履歴を記録
✓ 利用者に応じた複数ツールの活用ツールを提供したい。	✓ **Databricks NoteBookで協同作業** R、Python、SQL等で共通で利用できるWebインターフェイスを利用可能
✓ 全体状況が一目で把握できる可視化を行う	✓ **ReDashによるダッシュボード** その他、BIツールとも幅広く連携が可能

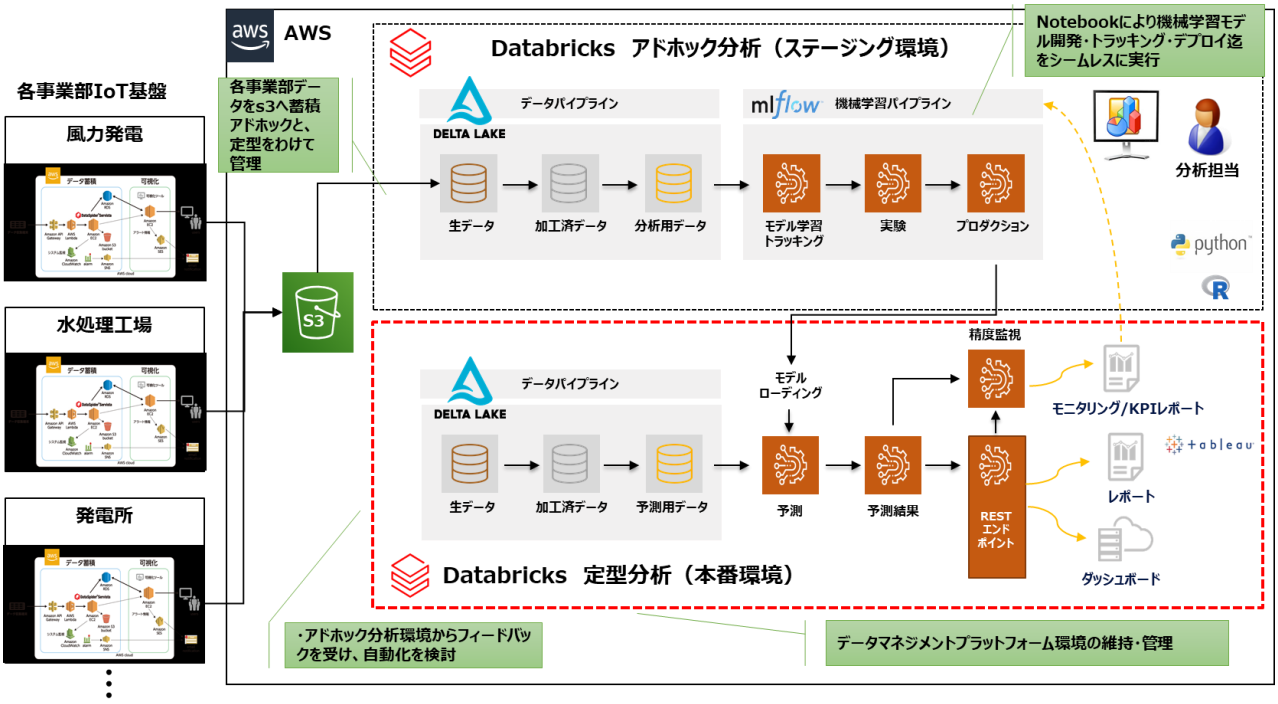
これらの課題を解決するために、データブリックスのパートナーであるメソドロジックは、日立造船の要件詳細をヒアリングした上で、データブリックスの「レイクハウス・プラットフォーム」による課題解決を提案しました。

1．　今後のデータ量増加に対する分散処理環境としての Apache Spark
2．　データ分析者の協同開発環境に、Databricks Notebook
3．　機械学習のモデル管理、柔軟なデプロイに MLflow

データブリックスは、大量データ処理の中心的な分散処理エンジン Apache Spark を中心としたプラットフォームです。ストリーミング処理とバッチ処理の両方に対応できることが強みであり、IoT データ取集に強みを発揮します。また、Spark 向けに特化されたストレージソフトである Delta Lake（データレイク）により、高速かつ高信頼性のデータパイプラインを構築することを可能にします。

データサイエンティスト向けには MLflow を使用しての協同開発によるノウハウ共有、データの調査やモデルの構築を一箇所に集約して実行可能な環境になっています。
日立造船では、データブリックスの各種機能を 5 カ月間にわたり検証、評価し、導入を決定しました。今後、日立造船では、様々なデータに対して データブリックスを用いたデータの利活用を実施していきます。

■導入イメージ図



株式会社メソドロジック　代表取締役社長である、大西 俊幸は、次のように述べています。
「日立造船様の IoT 基盤で収集したデータに対して、弊社がご提案したデータブリックスによる統合データ分析基盤を採用いただき、厚く御礼申し上げます。機械学習のシステムは、増大するデータ量の管理と、データエンジニアや、データ分析者の連携が非常に重要となります。データブリックスは、分散処理の Apache Spark™ と、MLflow により、大量データの高速処理と、コラボレーション開発により IoT 基盤に大きな効果を生み出す

ことができると確信をしております。今後も、データブリックス社とのパートナーシップにて、お客様のデータ＆AI　の民主化を推進していければと思っております」

データブリックス・ジャパンの社長である、竹内 賢佑は、次のように述べています。
「この度、日立造船様に当社のレイクハウス・プラットフォームが採用されことを大変光栄に存じます。製造業は日本を支える重要な産業の一つであり、日立造船様の IoT 基盤への当社プラットフォームの適応は、他の製造業様へ多くの示唆を提供するものと確信しております。当社はこれからも国内企業へのデータと AI の可能性を最大限に引出させるよう、引き続き"データと AI の民主化"に全力で取り組んで参ります」

株式会社メソドロジックについて

メソドロジックは、データモデリングとアーキテクチャー設計を軸とした IT コンサルティング会社です。
ビジネスの本質を捉え、業務と、システムの構造を可視化するモデリングと、最適な構造に変革するアーキテクチャー設計を実施しております。最先端の技術を活用した IT アーキテクトを通じてエンタープライズ向けシステムの構築、継続性のあるシステム運用、改善を支援しております。

名称	株式会社メソドロジック
代表	代表取締役社長　大西　俊幸
住所	東京都港区麻布台 2-4-5 メソニック 39MT ビル
事業内容	データモデリング、アーキテクチャー設計
URL	http://www.methodologic.co.jp

データブリックス・ジャパン株式会社について

データブリックスは、グローバルで唯一のデータ＆AI 企業です。Comcast、Condé Nast、Nationwide 、そして H&M といった 5,000 以上の世界のリーディング企業が、データブリックスが提供するオープン、且つ統合されたプラットフォームを活用し、データエンジニアリング、機械学習、アナリティクスなどのデータプロジェクトに取り組んでいます。データブリックスは多くのベンチャーキャピタルから支援を受け、米国サンフランシスコに本社を構え、世界中に営業拠点を有しているグローバル企業です。Apache Spark™、Delta Lake、MLflow のクリエーターによって創立され、企業のデータチームが抱える、世界の最も困難な課題を解決するサポートをしています。
その他のデータブリックスの詳細に関しては、Twitter、LinkedIn、Facebook をフォローしてください。

名称	データブリックス・ジャパン株式会社
代表	社長 竹内 賢佑
住所	東京都港区六本木 1 丁目 4-5 アークヒルズサウスタワー16 階
事業内容	データ&AI 向け統合データ分析プラットフォームの提供
URL	https://databricks.com/jp/

本プレスリリースに関するお問い合わせ先

株式会社メソドロジック

執行役員　COO 兼 コンサルティング事業部長：白石 章

電話　03-6380-4334　Email：info@methodologic.co.jp

データブリックス・ジャパン 広報担当（株式会社井之上パブリックリレーションズ）

担当：関口、中村、塚田

03-5269-2301　databricks@inoue-pr.com