

全社共通のデータ分析基盤としてデータブリックス製品を導入 データサイエンティストの育成を軸にDXを推進、 全社横断でのデータ連携を想定したRAGのパイプラインを構築



導入前の課題

- 分析環境の構築に手間がかかり、データサイエンティストの迅速な分析業務の障害になっていた
- データストアが異なり、データのソース/種類/量の増加に比例して、データのサイロ化も進行していた
- 社内向け対話型AIのリリース後、部門や用途ごとにカスタマイズされた生成AIへのニーズが高まっていた

データブリックスによる解決

- 分析基盤の構築工数が少なく、データサイエンティストが分析業務に集中できるようになった
- 非構造化データと構造化データを同一プラットフォーム上で処理できるようになり、データのサイロ化を解消することができた
- 部門単位だけでなく、部門間の情報障壁を超えた生成AIの利用までを視野に入れることが可能になった

DX推進を担う「二刀流人財」の育成プログラムを確立

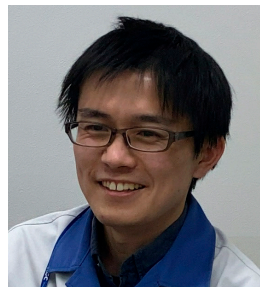
AGC株式会社は、建築用ガラスや自動車用ガラスなど多くの製品で世界トップクラスのシェアを持つ総合素材メーカーだ。グループビジョンとして「Look Beyond」(=先を見据える)を掲げ、2024年2月には、同年から3か年の中期経営計画「AGC plus-2026」を策定した。主要戦略としてコア事業と戦略事業の強化を共に加速させる「両利きの経営」の進化などを打ち出し、「2030年のありたい姿」の実現をめざしている。

AGCでは2018年から、業務知識を持ち、かつデータ解析のできる「二刀流人財」の育成に取り組んでおり、2019年10月には独自のデータサイエンティスト育成プログラム「Data Science Plus」を確立した。この育成プログラムの運営を担うのが、同社DX推進の中心となるデジタル・イノベーション推進部だ。

同部では2025年までに、データサイエンスで自部門の課題を解決することのできる上級データサイエンティストを100名まで育成することをめざしている。これまでの成果について、同部 デジタルソ



デジタル・イノベーション推進部
デジタルソリューション部
ビジネスプロダクト開発チーム
デジタル開発ユニットリーダー
吉中 泰輝 氏



デジタル・イノベーション推進部
デジタルソリューション部
ビジネスプロダクト開発チーム
石川 祐之助 氏



デジタル・イノベーション推進部
デジタルソリューション部
ビジネスプロダクト開発チーム
石原 俊也 氏

リューション部 ビジネスプロダクト開発チーム デジタル開発ユニットリーダーの吉中泰輝氏は、次のように説明する。

「既に育成されたデータサイエンティストは各事業部に分散配置され、現場でさまざまなデータサイエンスプロジェクトを担当しています。具体的なユースケースとしては、AIを用いた自動検査システムで、リアルタイムに取得する画像データから欠陥を抽出、判別することにより、製品の高品質化に貢献した事例などが挙げられます」(吉中氏)

一方で3つの課題を抱えていた データサイエンスプロジェクト

その反面、デジタル・イノベーション推進部は、さまざまなデータサイエンスプロジェクトを部門横断的にサポートするうえで、いくつかの課題を抱えていた。第一に、開発スタイルや計算資源が統一されおらず、分析環境が乱立していたことだ。これにより機械学習のモデル

ノウハウの共有を効率的に行うことができず、モデルのガバナンスや精度の維持、再現性にも問題があった。

第二に、クラウド環境上に個別に分析環境を構築するためには、ネットワーク/コンピューティングリソース/ストレージなどのインフラの設定が必要で、データサイエンティストがこの最初のステップでつまづくケースも散見された。第三に、データスタアが異なるため、データのソース/種類/量の増加に比例して、データのサイロ化も進行していた。

「これらの課題を解決するためには、分析基盤の構築工数が少なく、データサイエンティストが分析業務に集中できること、基盤が多種多様なデータソースにも対応できること、データの品質やガバナンスを担保できること、MLOpsを容易に実践できることが必要だと考えました。そこでこれら全ての要件を満たす統合データ分析基盤の導入を計画しました」(吉中氏)

個々のレベルに合わせて利用できるデータ分析基盤を獲得

AGCでは、複数の製品を比較検討し、最終的にデータブリックスの提供するデータ分析基盤「データインテリジェンスプラットフォーム」を採用した。

データインテリジェンスプラットフォームは、データブリックス独自の「レイクハウスアーキテクチャ」をベースに、未加工の非構造化データを格納するためのデータレイクと、処理済みの構造化データを格納するためのデータウェアハウスを単一のプラットフォームに統合したデータ分析基盤だ。ユーザー企業は非構造化データと構造化データを同一のプラットフォーム上で処理できるようになり、大規模データのリアルタイム分析などをより迅速に行うことが可能となる。AGCの抱えるデータのサイロ化という課題にも直接的な効果をもたらす特長だ。

同製品を採用した理由について、デジタル・イノベーション推進部 デジタルソリューション部 ビジネスプロダクト開発チームの石川祐之助氏は、次のように説明する。

「異なるスキルレベルを持つデータサイエンティストが各々のレベルに合わせて利用できること、多様なデータソースから生成されるデータの一元的な管理・活用が可能になり、データのサイロ化を解消できること、さまざまなユースケースを有益な資産として会社レベルで共有し、活用できること、そして一連のデータパイプラインやMLOpsをスケラブルに内製できることが挙げられます」(石川氏)

さらにデータインテリジェンスプラットフォームでは、分析環境がSaaSとして提供されているため、すぐに機械学習モデルの開発を開始できること、また社内にPythonユーザーが多いため、ノートブック形式でモデルの開発や運用ができることも決め手になった。

こうしてAGCは2022年末より、データインテリジェンスプラットフォームを採用したデータ分析基盤の運用を開始した。

ガラス寸法最適化パイプラインや社内向け対話型AIなどを構築

DX推進を担うデジタル・イノベーション推進部では、新たに構築したデータ分析基盤を活用して、さまざまなデータサイエンスプロジェクトを完成させている。例えばその1つが、ガラス寸法最適化パイプラインだ。

ガラス製品を作る場合、最初は1枚の非常に大きい板ガラスから、用途別にさまざまなサイズのガラスを切り出していくが、最後にどうしても余部分が出てくる。ガラス寸法最適化パイプラインは、これを最小化するための仕組みで、需要データや型式マスタ、価格データなどを取り込み、データインテリジェンスプラットフォーム上で機械学習を利用して複雑なパイプラインを自動で最適化する。

また2023年6月には、社内向け対話型AI「ChatAGC」を構築し、運用を開始した。ChatAGCは、AGCの従業員が安心・安全にチャットAIを業務に活用することを目的に開発された生成AIで、OpenAIの技術をMicrosoft社が企業向けに提供するAzure OpenAI Serviceが利用されている。また、ChatAGCの運用開始にあわせて、ITインフラや基幹システムを担当する情報システム部とデジタル・イノベーション推進部が主軸となり、生成AIの活用方法の探索を行う全社プロジェクトを立ち上げた。

プロジェクトの背景について、デジタル・イノベーション推進部 デジタルソリューション部 ビジネスプロダクト開発チームの石原俊也氏は、次のように説明する。

「生成AIを活用することで、従業員の業務効率の向上と、さらにはお客さまへのより良い製品やサービスの提供が可能になります。使わないという選択肢はありませんでした。しかし、社内データをインプットすることになるので、セキュリティ上、通常のChatGPTをそのまま利用することはできません。ChatAGCは、端的に言えば、ChatGPTのセキュリティを強化して社内向けにリリースした対話型AIで、従業員は入力情報を社外の組織に送信することなく、ChatGPTと同等の対話機能を利用できます」(石原氏)

そして現在、デジタル・イノベーション推進部が注力する取り組みが、全事業のデータを利用可能にするRAGのパイプライン構築だ。

部門単位さらには部門を超えたコラボレーションを促進する仕組みとして、RAGのパイプラインを構築

RAG(Retrieval Augmented Generation: 検索拡張生成)は、カスタムデータを活用することで大規模言語モデル(LLM)アプリケーションの有効性を向上させるアーキテクチャアプローチだ。質問やタスクに関連するデータや文書を検索し、LLMのコンテキストとして提供するもので、現在ドメイン固有の知識にアクセスする必要があるチャットボットやQ&Aシステムのサポートなどに活用さ

れている。

AGCにおけるRAGパイプライン構築の背景について、石川氏は次のように説明する。

「社内にChatAGCをリリースしたことで、ディスプレイ用ガラスや電子部材を製造する電子カンパニーから“対話型の生成AIを用いて、さらなる情報の利活用をしたい”という要望が上がってきました。それというのも、ChatAGCの裏側はあくまでChatGPTであるため、質問を投げて得られる回答の精度には限界があります。そこで電子カンパニーは、自部署にある営業情報と開発情報を取り込んで生成AIに質問を投げることで、事業活動の成果により直結する回答を得られるのではないかと考えたのです」(石川氏)

例えば営業部門の商談情報から顧客ニーズを把握できれば、研究部門は新たな製品開発のアイデアを得ることができる。一方、開発部門が現在開発中の製品やサービスについての知識を得ることで、営業部門はいち早く見込み客にアプローチできる。

「部門単位で、さらには部門を超えたコラボレーションを促進するための仕組みがRAGでした。実はChatAGCのリリース以降、早い段階で、他の部署からも、社内のさまざまな情報資産に紐付けて生成AIを活用したいというニーズは出てきていました。具体的な要望が上がってきたのは電子カンパニーからでしたが、RAGのパイプラインを開発するに当たっては、今後の展開までを見据えておく必要があったのです。この場面でも非常に有効だったのが、データインテリジェンスプラットフォームとデータブリックスの知見でした」(石川氏)

RAGは1つのステップであり、 今後は生成AIによる 意思決定支援までを実現したい

AGCは、2023年11月から1か月をかけて、電子カンパニーをファーストユーザーとしてRAGパイプラインを開発した。この時の開発基盤として利用したのも、データインテリジェンスプラットフォームだ。

「さまざまな形式のデータを同一プラットフォーム上で処理できるデータインテリジェンスプラットフォームを基盤にすることで、その上に部門や用途ごとのベクトルストアを構築し、UIから利用したいベクトルストアを組み合わせることで、目的に沿った回答の生成が可能になります。また、データはそのオーナー自身が管理しつつも、同一プラットフォーム上でベクトルストアとして整備することで、部門間の情報障壁を超えた生成AIでのデータ利用も可能になります。現在電子カンパニー以外のカンパニーや研究所、本社のバックオフィス系部門など社内の半分以上の部署で試用が始まっています」(石原氏)

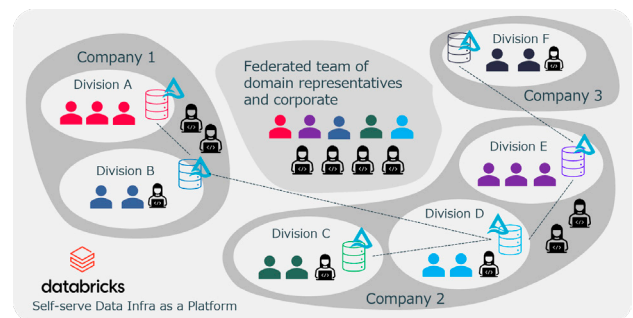
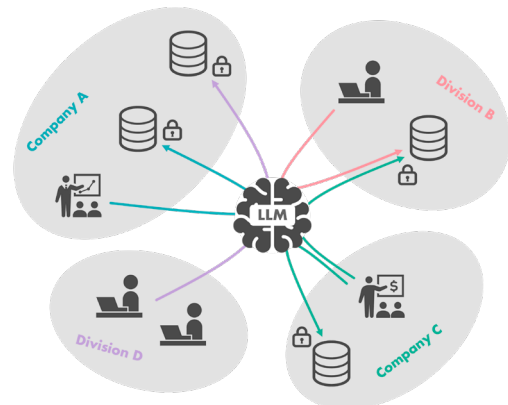
また今回のRAGパイプライン開発プロジェクトには、データブ

リックスのプロフェッショナルサービスチームも参画した。データブリックスの知見が役に立った場面について、石川氏は次のように振り返る。

「例えばデータブリックスの根幹技術となっているオープンソースの分析エンジンであるApache Sparkの詳細について、さらには生成AI自体の開発方法などについて教えていただきました。またプロジェクトの過程では、“将来こんな機能がリリースされる予定なので、この部分はしばらくは着手しなくてよいでしょう”というフォローをいただきました。スピード面、コスト面を含めて、非常に効率的な開発ができたと思います」(石川氏)

最後に今後の取り組みについて、吉中氏は次のように展望する。

「RAGは生成AIを活用するための1つの技術に過ぎません。2023年からは既に、意思決定支援まで一気に通貫に行う“Agent AI構想”に着手しています。生成AIの活用を判断の自動化にまで拡張するもので、こちらもできる限り早く実現したいですね」(吉中氏)



User's Profile

AGC株式会社

東京都千代田区丸の内一丁目5番1号

導入時期: 2022年~

<https://www.agc.com/>

フロート板ガラスや自動車用ガラスなど多くの製品で世界トップクラスのシェアを持つ総合素材メーカー。2024年2月には、新たな中期経営計画「AGC plus-2026」を策定、主要戦略としてコア事業と戦略事業の強化と共に加速させる「両利きの経営」進化などを掲げ、「2030年のありたい姿」の実現をめざしている。