

Serendie データ分析基盤

Serendie Digital Infrastructure - Data Analysis System -

*DXイノベーションセンター

要 旨

三菱電機のデジタル基盤“Serendie”(セレンディイ)⁽¹⁾が提供する技術基盤の一つ“データ分析基盤”へ、当社の事業データを集約している。データを生み出すコンポーネントやデータを蓄積するシステムは多岐にわたっているため、事業ドメインごとの多様な方式に対応してデータを集約する、及びデータごとに異なる利用条件に細かく対応して利活用を進める、という課題がある。今回、これに対応するため、データを管理する連邦型組織を構築するとともに、データメッシュ型アーキテクチャを採用してデータ分析基盤を実現した。今後は、この取組みに基づいてデータの集約と利活用を加速させるとともに、データ分析基盤のグローバル対応も進めていく。

1. ま え が き

当社グループは、各種事業で当社のコンポーネントやシステムが稼働することで生まれるデータをデジタル空間へ集約し、分析することによって、顧客の潜在課題・ニーズを把握し、新たな価値を提供する“循環型 デジタル・エンジニアリング企業”を目指して変革を進めている。その実現のためには、データの事業横断的な分析・活用から価値創出を可能にするデジタル基盤が不可欠である。当社グループが培ってきた様々な分野でのナレッジを新たな価値創出に広く活用するため、組織や事業を越えた活動を促進するデジタル基盤Serendieを構築した。Serendieの活用によって、社内だけでなく、顧客やパートナーとの共創を積極的に推進していく。

Serendieでは四つの技術基盤“データ分析基盤”“WebAPI連携基盤”“サブスクリプション管理基盤”及び“顧客情報管理基盤”を提供する。このうちデータ分析基盤は技術基盤の中核であり、各事業が個別に収集してきたデータを集約し、事業横断的に分析・活用する共通システムである。例えば、電力システム事業の“BLEnDer”(ブレンダー)、ビルシステム事業の“Ville-feuille”(ヴィルフィーユ)、空調・家電事業の“Linova”(リノバ)など、これまでそれぞれのシステムの中で、異なる形式で別々に保管されていたデータを一つのデータプールに集約し、分析することによって、新たなソリューションの開発に活用できるようにしている⁽²⁾(図1)。

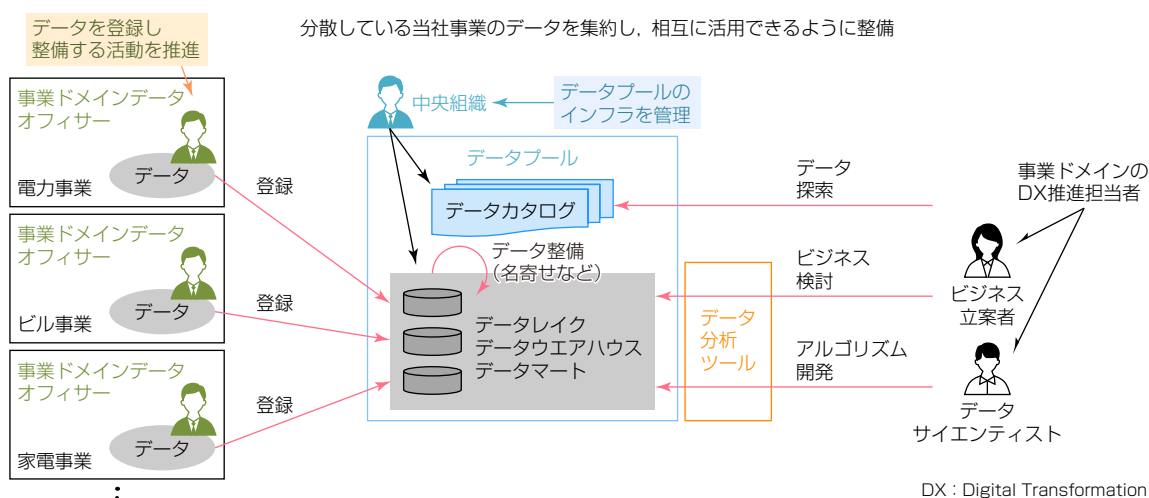


図1-データ分析基盤の全体構成

顧客やパートナーとの共創を積極的に推進していくためにデータ分析基盤が備えておくべき機能としては、全社にわたって広くデータを集約できており、データは常に最新の状態で保たれて容易に理解できる形で格納されていること、及びビジネス立案者やデータサイエンティストがいつでもデータ分析を行える環境が整っていることである。この実現に向けて、構築開始以来、データ分析基盤の変革を行ってきた。

本稿では、2章でデータ管理組織の変革について述べて、3章ではそれに伴って実施したデータ基盤のアーキテクチャの変革について概要を述べる。また4章ではデータ分析ツールと分析環境について述べる。

2. 事業データ管理組織

2023年度末に、データ分析基盤の構築を開始し、その管理組織については中央集権型でトライアルを開始した。

この章では、データプール構築以前、データプール初期段階、及び現時点までのデータプール管理組織の形態の変革について述べる(図2)。

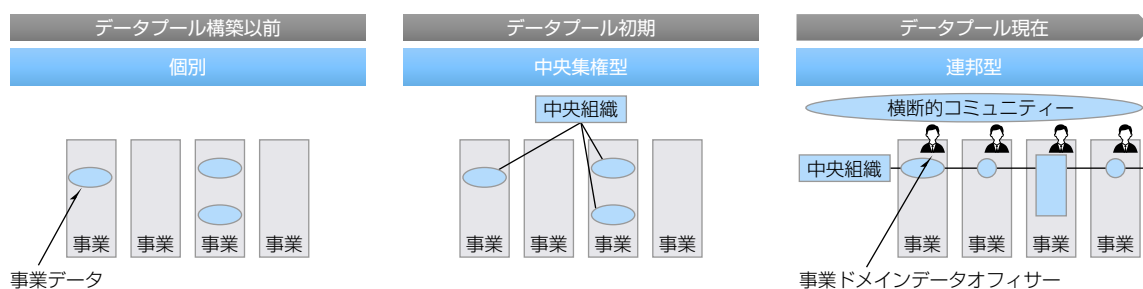


図2-データプール管理組織の変革

2.1 データプール構築以前の状態

2023年度初期の段階では、事業ドメインによってはデータ集約の取組みが行われていないか、又はドメイン独自で取組みが進行しているかの状態であった。データを蓄積するシステムは事業ドメインごとに分かれて構築されており、ドメイン内で統合して構築されているものもあれば、より細かな事業ごとに個別に構築されているものもあった。すなわち、事業ドメインによってデータ集約の状況は様々であり、データがサイロ化されている状態であった。

また、事業間を跨(また)いでデータを連携し統合ソリューションを創出する取組みも存在したが、システム間で個々に連携を行っている状況であり、事業を跨いだ分析やソリューション構築を活発化するにはまだ効率的ではなかった。

2.2 中央集権型管理組織の発足

データプールを管理する組織は、中央集権型でトライアルを開始した。すなわち、中央組織でデータ基盤を整備するとともに、データ管理面でも中央組織が主導することにした。具体的には、データの登録、各データサーバーとデータプールをつなぐデータパイプラインの構築、メタデータの収集などに加えて、各事業ドメインでのデータの利用条件を調査し、それに基づいて利用者に対してデータの共有などを実施してきた。中央組織にエンジニアを集めて、集中的にこれらの作業を行うことで、短期間でデータプールの立ち上げができた。

しかしながら、データの収集範囲が事業ドメインの広範に及ぶに当たり、次第にアジリティ面で次の課題が見えてきた。

- (1) 個々のデータは、それぞれの利用条件に基づいて管理され、その利用条件は個々に更新されていく。中央集権的組織だけでこれらを把握してデータ利活用を行うには、各事業ドメインに対する問合せが頻繁に発生し、アジリティを損なう要因になった。
- (2) 事業ドメインでのデータの蓄積状況やデータの構造は、コンポーネントやシステムの実装に依存して様々である。個々データの状況を理解した上で収集していくのは手間がかかる。結果的に、各事業ドメインでのコンポーネントのエンジニアに対して頻回にヒアリングする等が生じて、アジリティを損なう要因になった。

2.3 連邦型管理組織化と中央組織の再構築

今回、これまで中央集権的に実施してきたデータの利用条件の管理やデータ登録などのデータ管理の責務を、全社横断的な役割に注力する中央組織と各事業ドメインとに分散させることで、データ集約でのアジリティーを確保することを目的として、データ管理組織の再構築を行った。併せて、事業ドメインが作成したデータや分析ノウハウを、他の事業ドメインでも積極的に共有していく横断的コミュニティを構築することで、データ利活用を促進する(図3)。

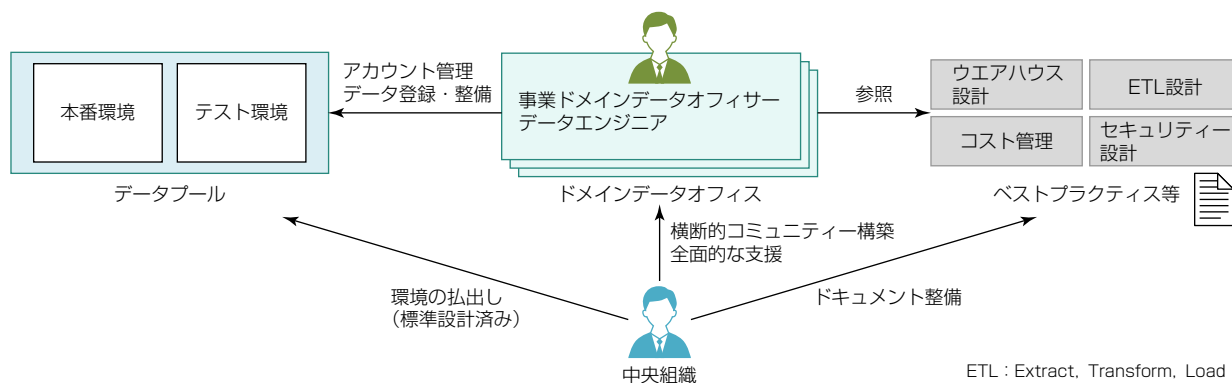


図3-中央組織と事業ドメインデータオフィサーの役割分担

これを実現するため、事業ドメインで、各ドメインでのデータやデータカタログの登録・管理、データ利活用を推進する役割を持つ“事業ドメインデータオフィサー”を任命した。事業ドメインデータオフィサーは、実業務を担当するデータエンジニア等も含めてドメインデータオフィスを構成する。

また、中央組織では、これまでどおりデータ基盤を継続的に整備していくことに加えて、データプール活用のためのドキュメントを整備したり、エンジニアによる支援を行ったりすることによって、ドメインデータオフィスでの業務を全面的に支援する。

3. データプールを実現するデータメッシュアーキテクチャー⁽³⁾

2章で述べたデータプール管理組織の変革に伴って、各組織の役割分担をシステムに落とし込むために、データ基盤のアーキテクチャー変革も実施した。事業ドメインがそれぞれ自律分散的にデータを集約可能にするために、これまでは一つの領域を共有する形であったデータプールを分割し、事業ドメインが所有するデータ格納領域である“ドメインアカウント”と、事業横断でデータを利活用するために使う領域である“プロジェクトアカウント”とに分離し、お互いをメッシュ的に接続することにした。また、利用者が自由にデータを探索するために“データカタログ”も設置した(図4)。

この章では、これらの構成要素について述べる。



図4-データ基盤アーキテクチャーの変革

3.1 ドメインアカウント

ドメインアカウントは、事業ドメインごとに設置するデータを格納する領域である。事業ドメインデータオフィサーは、自身の事業ドメインに存在するデータを収集してドメインアカウントに登録したり、利用者を登録したりする作業を行う。また、利用者がいつでもアクセスできるようデータを維持管理するとともに、利用者が容易に利用できるようデータを整備するといった、データの民主化に向けた活動を行う。

ドメインアカウントに登録された利用者は、同アカウントに登録されているデータを閲覧でき、事業ドメイン内でのデータ共有・利活用を進めることができる。

またドメインデータオフィサーは、データごとに利用条件を定めて、その条件に合致した他アカウントからの利用申請に対して、当該データへのアクセス許可を与えることができる。これによって他アカウントに所属する利用者にもデータを閲覧させることが可能になり、事業を跨るデータの共有・利活用を推進できる。

3.2 プロジェクトアカウント

プロジェクトアカウントは、データを利活用する組織のための一時的なビューや、一時的な格納場所を提供する領域である。プロジェクトアカウントは、新たな事業創出を企図するスクラムプロジェクトのプロダクトオーナーが管理することを想定している。

スクラムプロジェクトでは、事業横断的にデータを探し、データを取得し、分析等を行ってソリューションの検討を行うために、データカタログとドメインアカウントに接続する。そして、データカタログを参照してデータを探索し、所望のデータを所有するドメインアカウントに対して当該データの利用申請を行い、当該データへのアクセス権を取得する。なお、複数のドメインアカウントに対してデータの利用申請を行うことで、事業ドメインを跨いでデータを参照することも可能になる。

3.3 データカタログ

各事業ドメインが持つデータが、他アカウントから容易に見え、理解され、安全に利用されるように、中央組織がデータカタログを整備する。事業ドメインデータオフィサーはデータカタログにデータの説明を登録できる。データカタログには、データの内容や構造を述べるメタデータだけでなく、当該データの利用条件や機密区分(個人情報か否かなど)も掲載する。これによって利用者は、データを受け入れる領域の適切なカスタマイズや、利用条件に合わせた共創テーマの設定などを事前に行うことができるようになる。

今後、顧客との広範な共創を進めるために、データカタログは可能なかぎり当社グループ内に広く公開できるよう運用の改善を図っていく。

4. 事業ドメインを跨るセルフサービスのデータ分析環境

プロジェクトアカウントからは、事業ドメインを跨いでデータを参照できる。これを利用し、プロジェクトアカウントの領域と、各データ分析ツールとの間にコネクタを構築することで、データ分析ツールから事業ドメインを跨いだ分析が可能になる。

分析ツールとしては、データサイエンスの事前知識なくノーコードで容易にデータを視覚化するBI(Business Intelligence)ツールや、AIを活用したツール、プロフェッショナルな分析を行うことができるツールなど、複数の分析ツールを用意している。3章に述べたデータメッシュの機能と併せることで、ビジネス立案者やデータサイエンティストが、いつでも事業ドメインを跨いでデータを表示、評価、分析できる、セルフサービスの分析環境を提供する。

5. む す び

Serendieの技術基盤の一つであるデータ分析基盤の概要を述べるとともに、そのコアになるデータプールについて、各事業のデータを集約して継続的に管理していくための組織面及び技術面のアーキテクチャーについて概要を述べた。

今後はSerendieの海外展開に伴い、データプールのグローバル対応を行っていく予定である。

参考文献

- (1) 三菱電機：Serendie セレンディ
<https://www.mitsubishielectric.co.jp/serendie/>
- (2) 三菱電機：IR DAY 2024：三菱電機のDX戦略（2024）
- (3) Dehghani, Z. : Data Mesh Principles and Logical Architecture（2020）
<https://martinfowler.com/articles/data-mesh-principles.html>

