

re:Invent で発表されたライフサイエンス関連事例

アマゾン ウェブ サービス ジャパン合同会社
ヘルスケア・ライフサイエンス事業開発部
シニア事業開発マネージャー
片岡 勇人



自己紹介

名前: 片岡 勇人 (かたおか ゆうと)

所属: アマゾン ウェブ サービス ジャパン合同会社
ヘルスケア・ライフサイエンス事業開発部
シニア事業開発マネージャー

経歴: 国内半導体メーカー
→外資系ヘルスケア企業
→現職

役割: ✓ヘルスケア・ライフサイエンス領域のお客様の取組みご支援
✓クラウドに対する日本のお客様固有の要件にお応えするため
アマゾン ウェブ サービス (AWS) グローバルチームと連携



Our Mission

Amazonは、
地球上で最もお客様を
大切にする企業であること



AWS は生活者・患者をとりまくステークホルダーのインフラをご支援

Provider (医療機関)

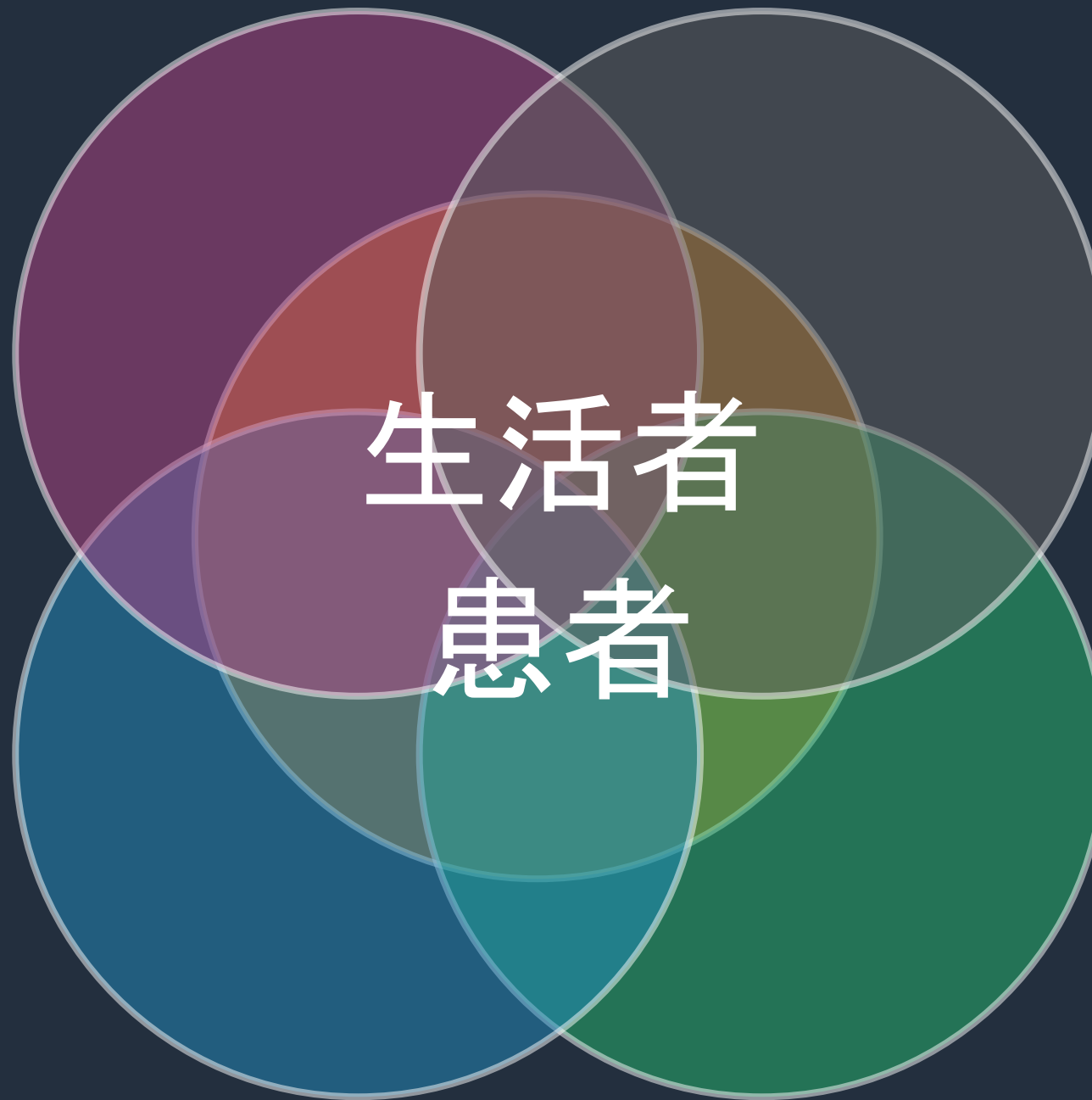
病院
クリニック
歯科
調剤薬局

介護施設・サービス
在宅支援

Vendor (技術提供)

医薬品
健康医療機器

ISV・ソフトウェア
SI・サービス



Payer (保険者)

健康保険組合
協会けんぽ
国保

企業 (総務担当)
～福利厚生～

(保険会社)

Government (政府)

法規制
公衆衛生
研究者

国内における製薬関連のお客様

(一部抜粋)



国内における公的研究・ゲノミクス・予防・介護関連のお客様

(一部抜粋)



京都大学
KYOTO UNIVERSITY



国立循環器病研究センター
National Cerebral and Cardiovascular Center



TAKARA BIO INC.



XCOO [ténku:]



Genesis Healthcare



FiNC
Technologies



国内における医療関連のお客様

(一部抜粋)



京都大学
KYOTO UNIVERSITY



藤田医科大学
FUJITA HEALTH UNIVERSITY



国立循環器病研究センター
National Cerebral and Cardiovascular Center



特定機能病院 / 地方独立行政法人 大阪府立病院機構
大阪国際がんセンター

AOI 国際病院 医療法人社団 葵会
AOI UNIVERSAL HOSPITAL



東京都済生会中央病院
TOKYO SAISEIKAI CENTRAL HOSPITAL



平成医療福祉グループ
HEISEI MEDICAL WELFARE GROUP



日本医師会 ORCA 管理機構



クラウド利用における製薬業界のトレンド



データの民主化



バリューチェーンのDX



リアルワールドデータの
活用

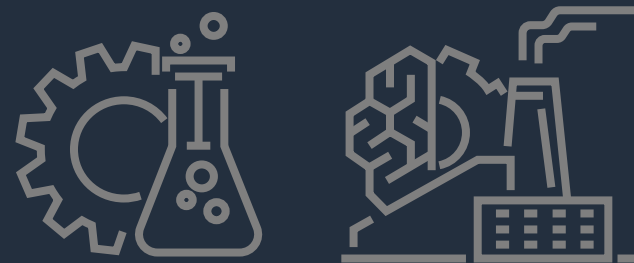
本セッションでお話しする内容

1. 膨大な社内データの活用で広がるデータの民主化
2. 機械学習やIoTで進むバリューチェーンのDX
3. マーケットプレイースで簡素化されるリアルワールドデータ利用

クラウド利用における製薬業界のトレンド



データの民主化



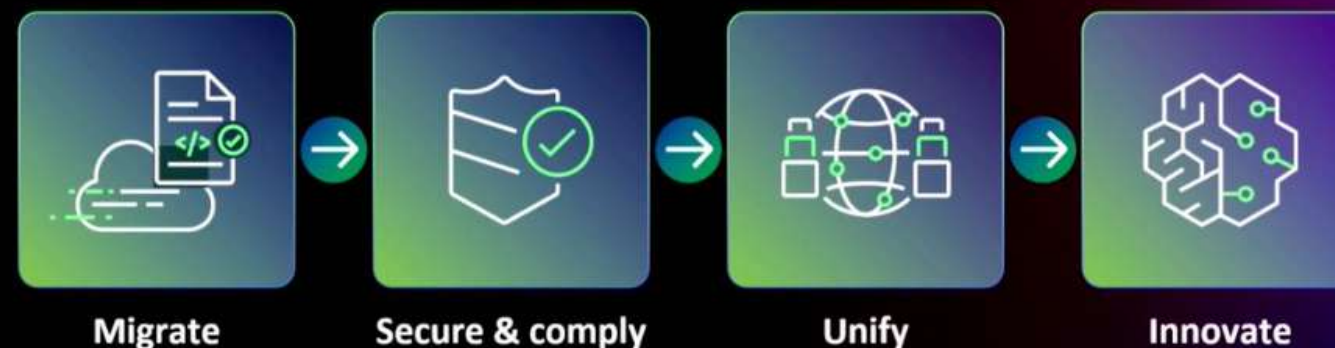
バリューチェーンのDX



リアルワールドデータの
活用

ギリアド: 爆発的に増えるデータを支える データメッシュアーキテクチャ

Migration to innovation journey



マイグレーションからセキュリティコンプライアンス対応

- 数千のVMや数十のエンタープライズアプリケーション（GxP含む）、数万のユーザーが対象
- GxP対応に関わる「神話」vs 様々な検証済みアプローチやフレームワークが活用できる「事実」
- データセンターの70%がAWSへ移行済み

データ統合からイノベーション

- データ統合により今後2.5PB相当のデータがクラウド上で扱えるように
- 機械学習を活用し、非構造化データを含む膨大な社内データの検索時間を50%削減
- データメッシュを活用したエンドトゥーエンドの部門横断的なAIプラットフォームを構築

ギリアド: 爆発的に増えるデータを支える データメッシュアーキテクチャ

従来のモノリシックな管理から脱却し、データメッシュアプローチでエンタープライズレベルでデータ利活用のスケーラビリティを確保

Why data mesh?

1. モノリシックなアプローチはスケールしない
2. トップダウンのデータガバナンスが機能しない
3. データの信頼性が低下
4. データパイプラインの複雑性が増加

データメッシュ・プリンシプルス (原則)

1. ドメインオーナーシップ
2. Data as a product
3. セルフサービス
4. フェデレーションガバナンス

成功要因

1. ドメインの定義づけとデータカルチャーの醸成
2. CxOからのスポンサー
3. ユースケースと目標の明確化
4. ステークホルダーへの周知

データメッシュアーキテクチャに基づくAIデータプラットフォームを7つの異なるビジネスユニットに構築し、そのうち3つのユニットで既に成果が出ている。

ノボノルディスク: NNEDH (ノボノルディスク エンタープライズデータハブ) が実現する分散型データレイク

データの民主化

中央集権的なデータレイクが抱える課題

First generation data lake

NOVO NORDISK ENTERPRISE DATA LAKE (NNEDL)

- Centralized data lake
- 2,000+ users
- 2.3 PB of data
- Supports multiple data domains
- Data ownership delegation using AWS S3 access points
- Data access based on user persona
- Integrated data warehouse with Amazon Redshift
- GxP compliant
- GDPR compliant

The diagram illustrates the data flow from various sources (FSX, Oracle, etc.) through ingestion and consumer accounts into a centralized data lake. The lake is organized into layers: Source, Raw, Transient, and Curated. It is integrated with Amazon Redshift and an API Gateway, leading to data visualization tools.

デザインパターン

- シングルのアカウントで扱えるリソースの制限
- 特定のユースケースごとの構築
- ポリシードキュメント管理の煩雑さ
- アカウントごとに複製されるデータ

オペレーション

- データセットへのユーザーベースのアクセス
- データの再利用の難しさ
- エキスパートや技術への依存
- データパイプラインごとに必要となるバリデーション



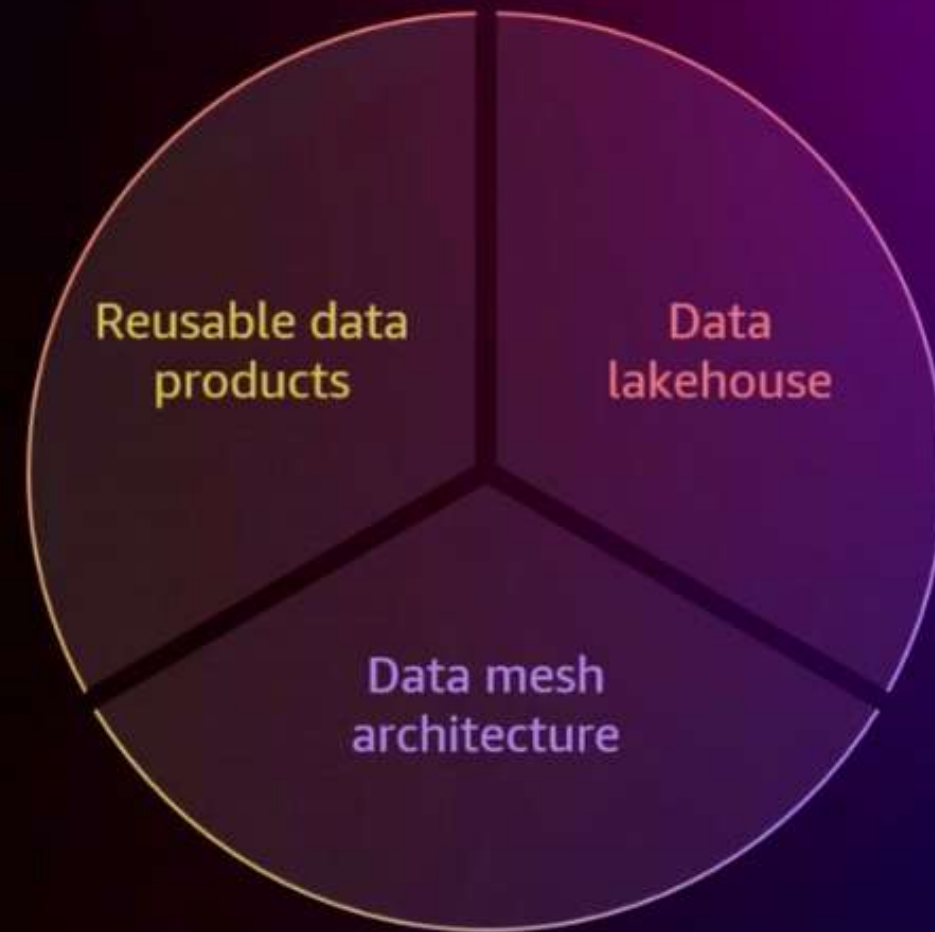
ノボノルディスク: NNEDH (ノボノルディスク エンタープライズデータハブ) が実現する分散型データレイク

データの民主化

Next generation data lake

NOVO NORDISK ENTERPRISE DATA HUB (NNEDH)

- Distributed data lake
- Data discovery
- Role based access to dataset
- Integrated data sharing using AWS Lake Formation
- Pipeline blueprints
- Automated and distributed infrastructure
- GxP compliant
- GDPR compliant



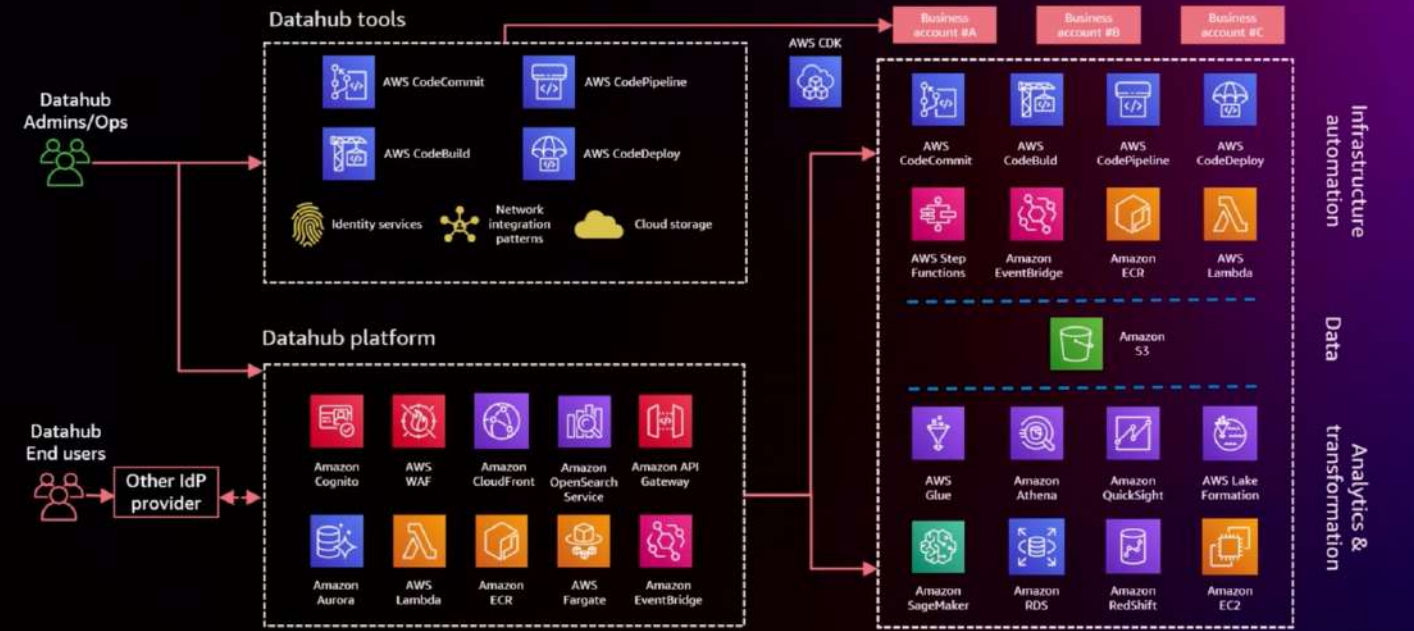
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

<https://youtu.be/YCNVdK5kPWk>

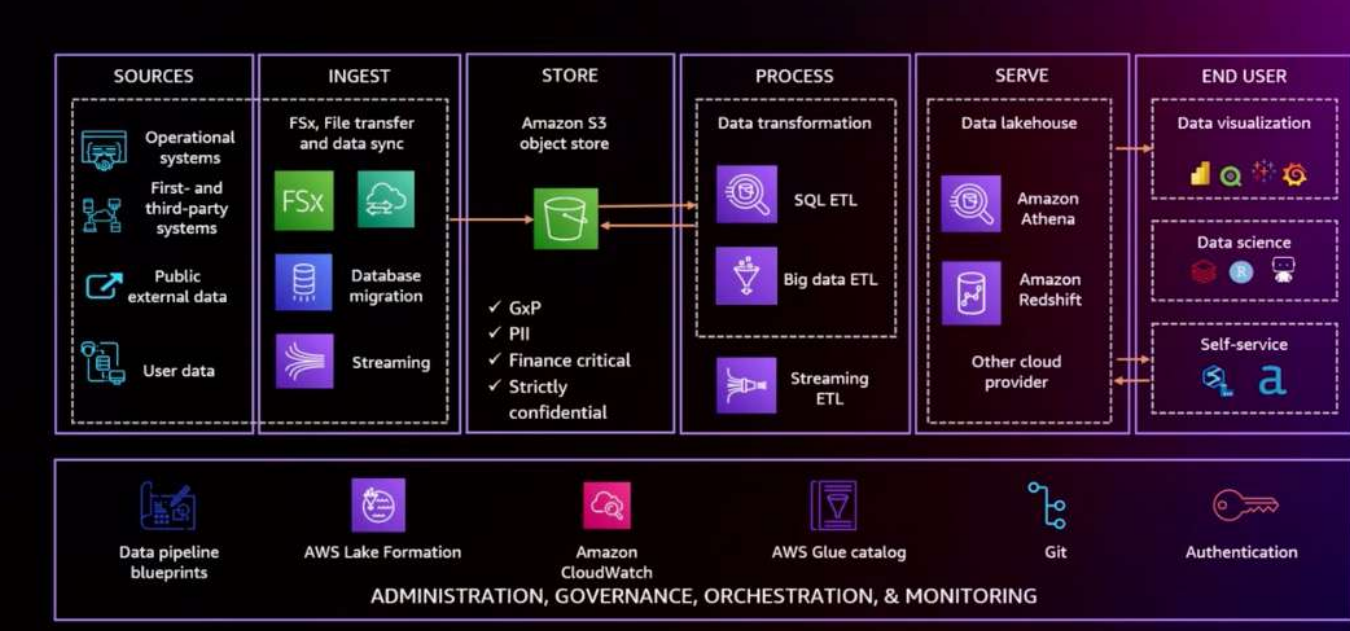
ノボノルディスク: NNEDH (ノボノルディスク エンタープライズデータハブ) が実現する分散型データレイク

データの民主化

AWS accounts & services used in NNEDH



Data architecture in NNEDH business account



クラウドネイティブなDXに向けたベストプラクティス イーライリリー、アッヴィ、ギリアド、ファイザー4社座談会

データの
民主化

イーライリリー

クラウドネイティブな環境開発に注力し、自動化と高度な分析を実現。数年後には完全にクラウド以降が完了予定。パデュー大学との提携による学位取得やe-learning基盤整備など人材育成にも注力。



Tim Coleman
Vice President & Information
Officer at Eli Lilly

アッヴィ

多種多様な治験データをFDAに提出するために必要な膨大な前処理のマンユアルワークをAIで自動化。買収したアラガンがさまざまな場所に保管していたデータを統合することでコスト削減・臨床試験の迅速化。



Brooks Fowler
Vice President of Clinical
Data Strategy at AbbVie

ギリアド

クラウド上でデータ統合基盤を構築し、AWSマーケットプレースを通じて利用するSaaSサービスと連携し幅広いデータ活用。社員のAWSの習熟度に合わせたトレーニングコースを開設計し、レベル別人材育成。



Nicole Quash
Executive Director & Head of
Development Systems, Gilead

ファイザー

コロナワクチンの開発以降、イノベーションのペースが加速。デジタルヘルスで患者データを収集し、新しい治療法の開発に注力。データセンターの75%をクラウドに移行したことでモダナイゼーションが加速。



Wes Barnes
Sr. Director of Digital Hosting
Solutions Pfizer

クラウド利用における製薬業界のトレンド



データの民主化



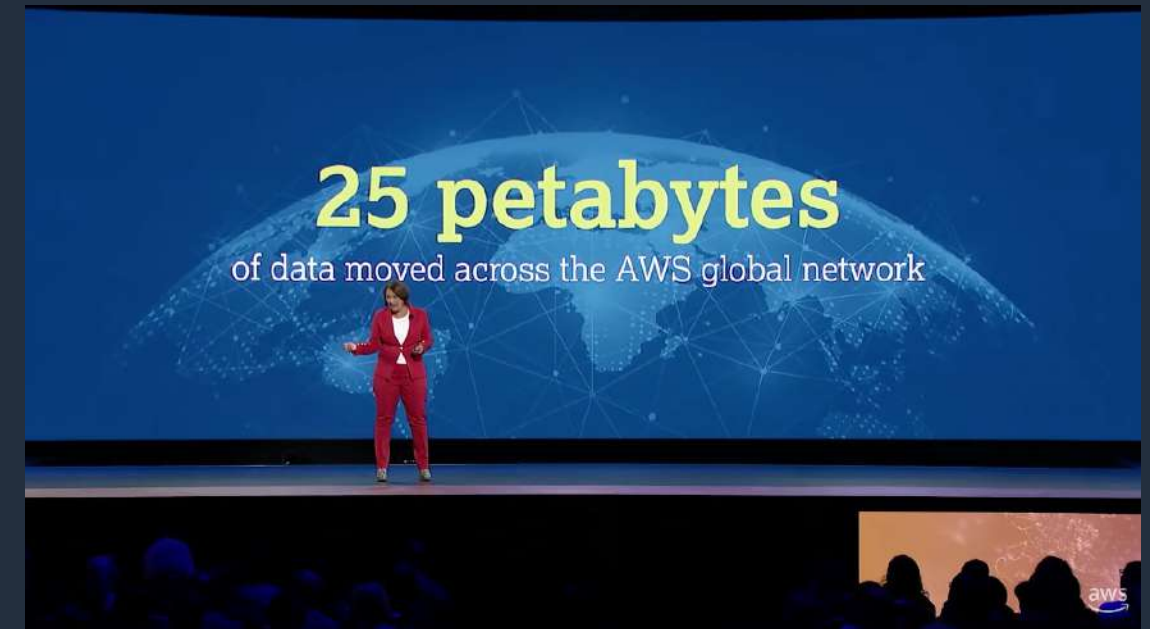
バリューチェーンのDX



リアルワールドデータの
活用

AstraZeneca: Omicsデータの解析とAI活用

- 2026年までに社内で200万の全ゲノム解析を計画しており、すでに25PBのデータが活用される。
- 解析プロセスにおいては、AWS Step Functions, AWS Lambda, AWS Batch, Amazon S3などにより効率化を実現。
- 30時間で1,100億件の統計処理を実行
- マルチオミックス、そして画像などを含めたマルチモーダルデータを分析している
- 機械学習の基盤としてAmazon SageMakerを活用しており、数百のプロジェクトが同時に利用している
- さらに患者様の自宅からヘルスデータを収集する技術に取り組んでおり、早期診断や疾病予測に活用



KEK: クライオ電子顕微鏡が生み出すテラバイト級のデータ解析にかかる計算時間・コスト最適化の検証

KEK: High Energy Accelerator Research Organization



KEK Cryo-EM facilities

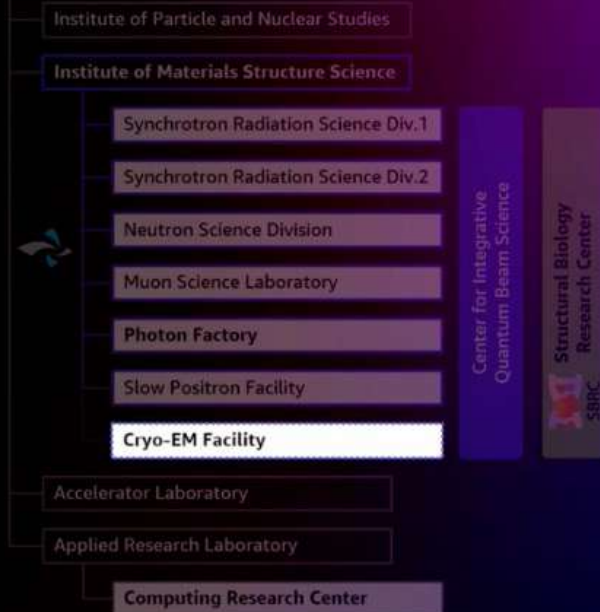


Talos Arctica
200 kV
Cryo-EM
Falcon 4

Coming soon



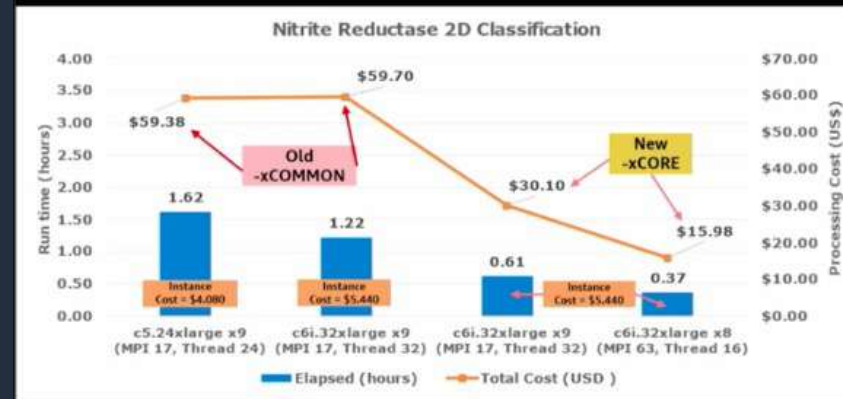
Titan Krios G4
300 kV Cryo-EM
Falcon 4i
Selectris X Filter



aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cryo-EM benchmarks – Life science – RELION on AWS

NITRITE REDUCTASE: 2D CLASSIFICATION | VARYING INSTANCE COUNTS AND COMPILER SETTINGS



- KEY POINT:**
Using the right compiler flags can shorten the runtime and reduce the processing costs
- RESULTS:**
- Using `-xCORE` instead of `-xCOMMON` provides a **~3.4x shorter** runtime and **3.7x lower cost**
 - Using 8 instances vs. 9 instances with `-xCORE` is **~1.6x faster** and a **~1.9x lower cost**
 - Using more MPI threads and **all** virtual cores vs. less MPI threads and **mostly** physical cores yields the shortest time and lowest cost

Notices and Disclaimers: Performance varies by use, configuration, and other factors. Learn more at www.intel.com/PerformanceIndex. Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. No product or component can be absolutely secure. Your costs and results may vary. Intel technologies may require enabled hardware, software, or service activation. © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

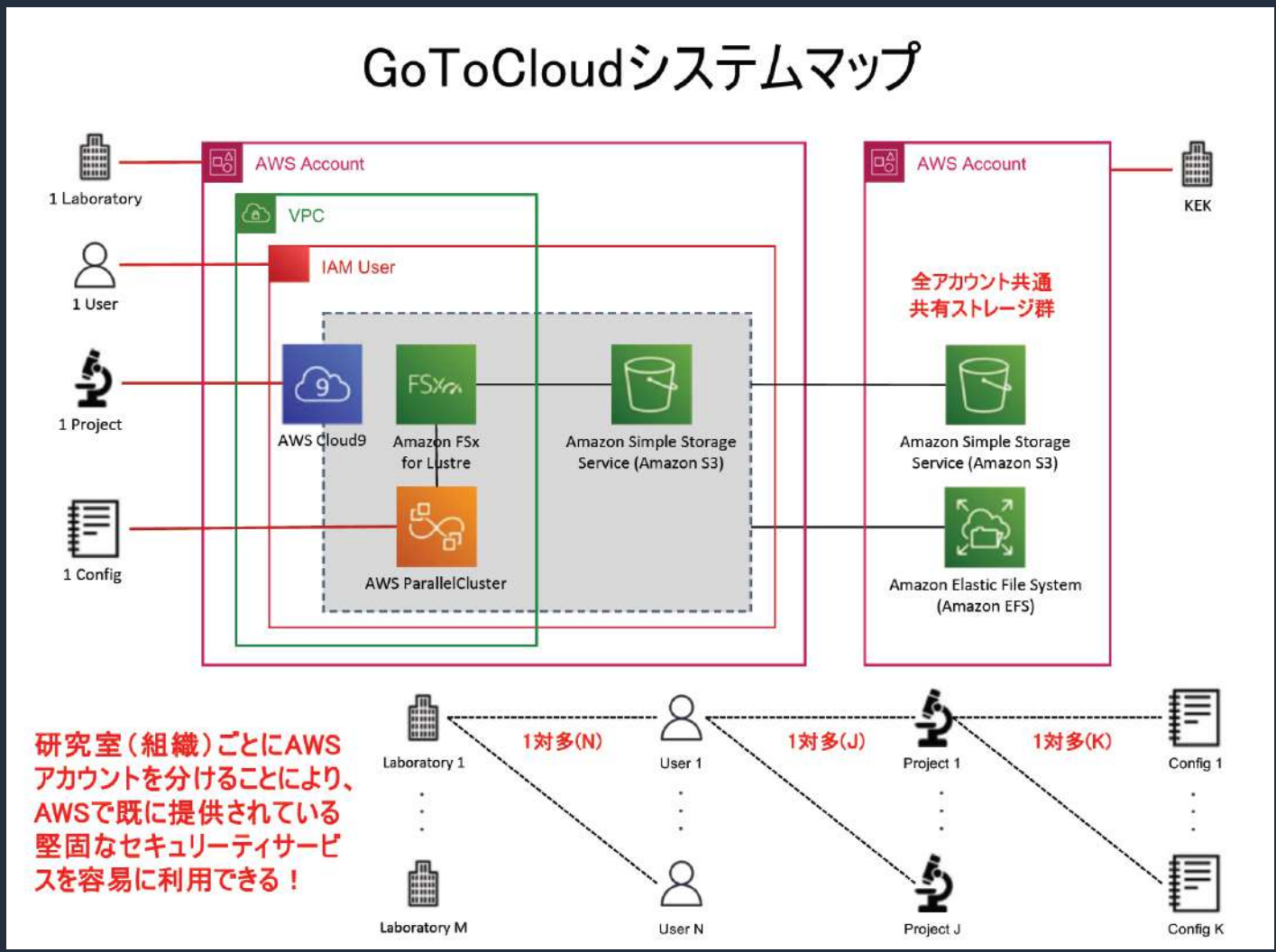
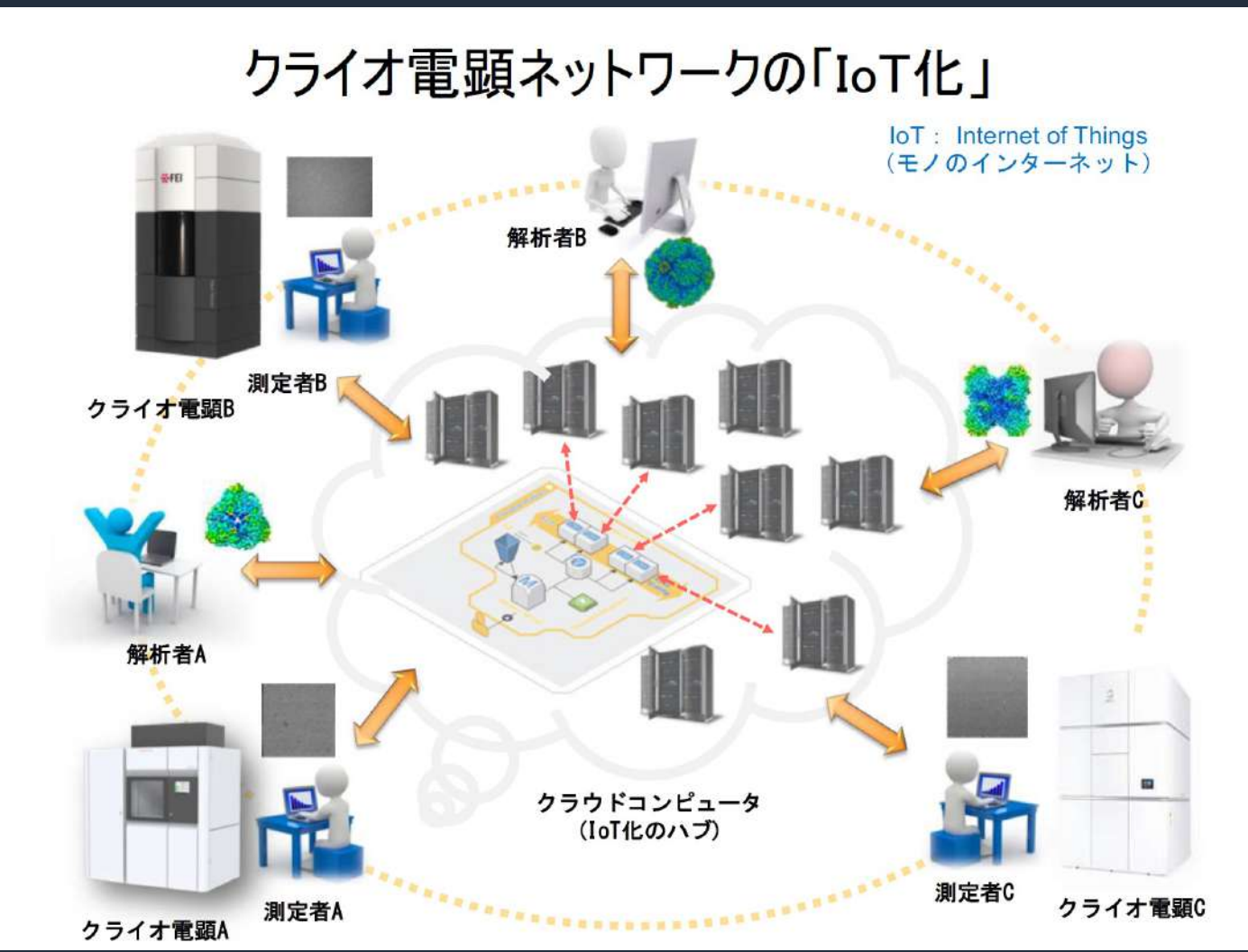


- クライオ電子顕微鏡を所有しており、日本のあらゆる研究者に利用が開かれている。
- クライオ電子顕微鏡から生まれる大規模データを処理するハイエンドな計算環境を多くの研究者が容易に低コストで利用できるためにAWSを活用。

- Amazon EC2 C6iシリーズのインスタンスを活用し、かつコンパイラを調整することで、ランタイムを抑えてトータルコストが50%に。

<https://youtu.be/14awm2bB1vo>

(参考) CBI学会2021年大会 AWSスポンサーセッション KEK : AWSを活用したクライオ電子顕微鏡ネットワークのIoT化



日経xTECH Active: 第一三共のAWSを活用した 創薬化学研究プラットフォームの取り組み



「創薬ワークフローハンズオン」で創薬領域のユースケースに合わせたAWSの利用方法を体系化

生命・疾患
の理解

ターゲット
選定

医薬品候補の
探索や改良

前臨床試験

臨床試験

新薬申請
市販後評価

PubMed論文
Abstract分析

DeepChem on
Sagemaker

Cryo-EM on AWS
ParallelCluster
(CryoSPARC)

動物行動検出

MDV社
RWD分析

GenomicsCLI

Cryo-EM on AWS
ParallelCluster
(Relion)

白血球画像分類

SNS医療情報
抽出・分析

SageMaker DGL-LifeSci

Protein Folding
(OpenFold/AlphaFold2/RoseTTAFold/OmegaFold)
on AWS Batch

GROMACS on
ParallelCluster

VirtualFlow on
ParallelCluster

ハンズオン

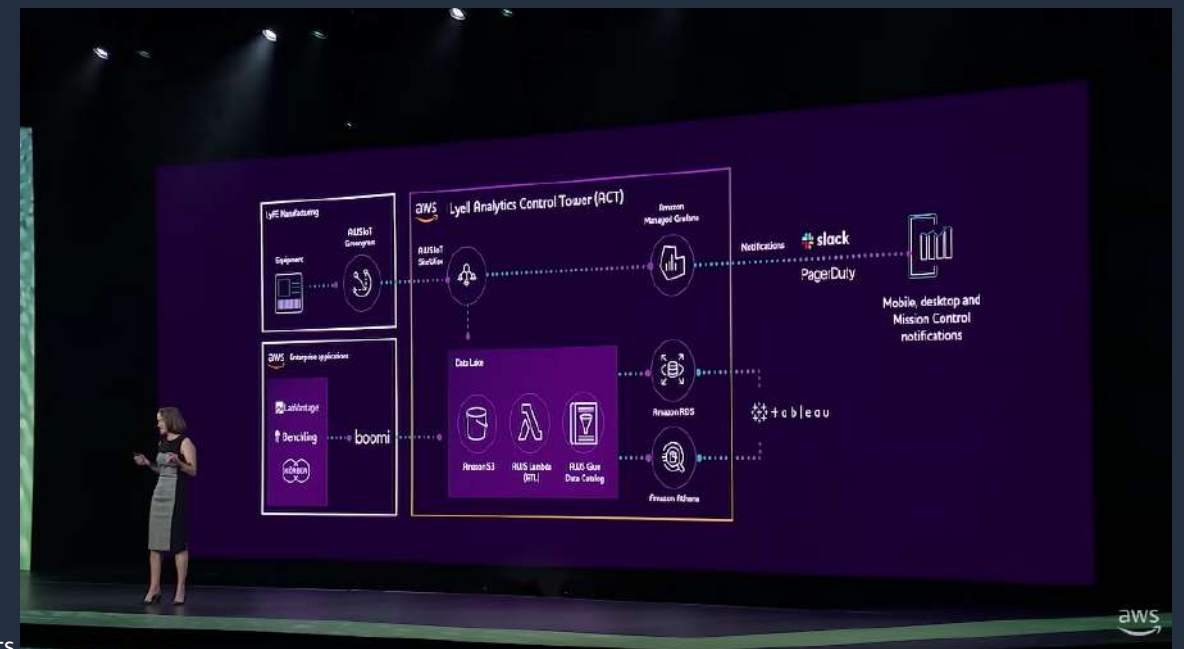
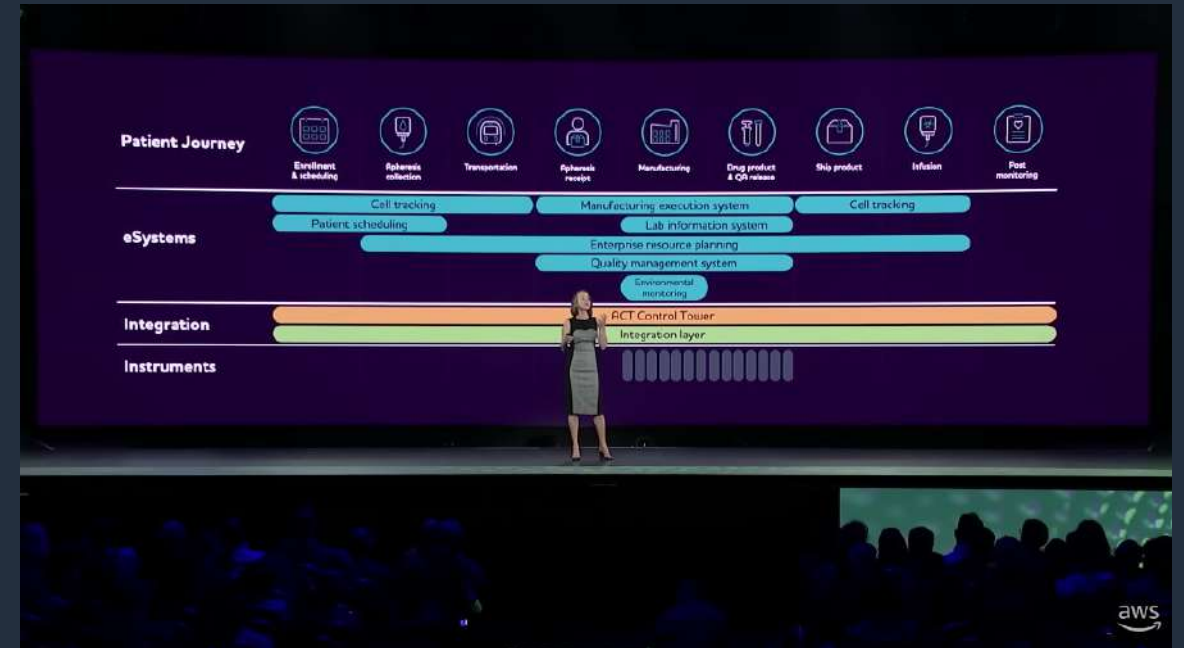
サービスの評価や使い方を体
験いただける手順書

ソリューション

すぐに実際にご利用いただけ
るテンプレートや手順書

Lyell: ペイシエントジャーニーに渡る膨大なデータをリアルタイムに活用する細胞療法プラットフォーム

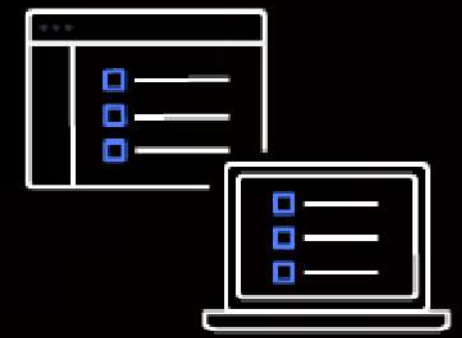
- Lyell社は固形がん領域への細胞療法を提供。細胞療法とは患者様の免疫細胞を採取・改変後、投入することでがん細胞の死滅を目指す治療。
- 細胞の採取、輸送、施設でのエンジニアリングプロセス、再度の輸送と投入というペイシエントジャーニーのあらゆるプロセスでの品質管理が求められる。
- LyFE Manufacturing Centerで製造・サプライチェーンに関わる膨大なデータを活用する基盤としてAWSを採用。
- IoTの技術を活用することで、リアルタイムに施設の機器の稼働状況を可視化したり、異常があればアラートが通知される。
- 現在4つのがん疾患で治療法を開発中で、今後の大規模治験、商業化に向けたスケーリングにも対応。



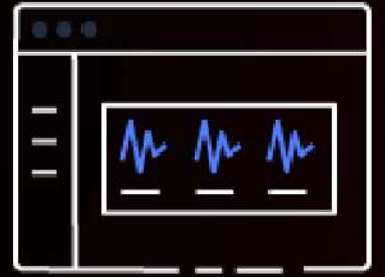
Johnson & Johnson: サプライチェーンの可視化を統合的に実現するフルスタックのAWS Observability

バリューチェーンDX
製造・サプライチェーン

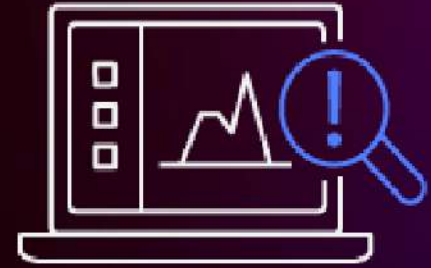
Challenges – No longer an afterthought...



Observability



Reliability



Reporting



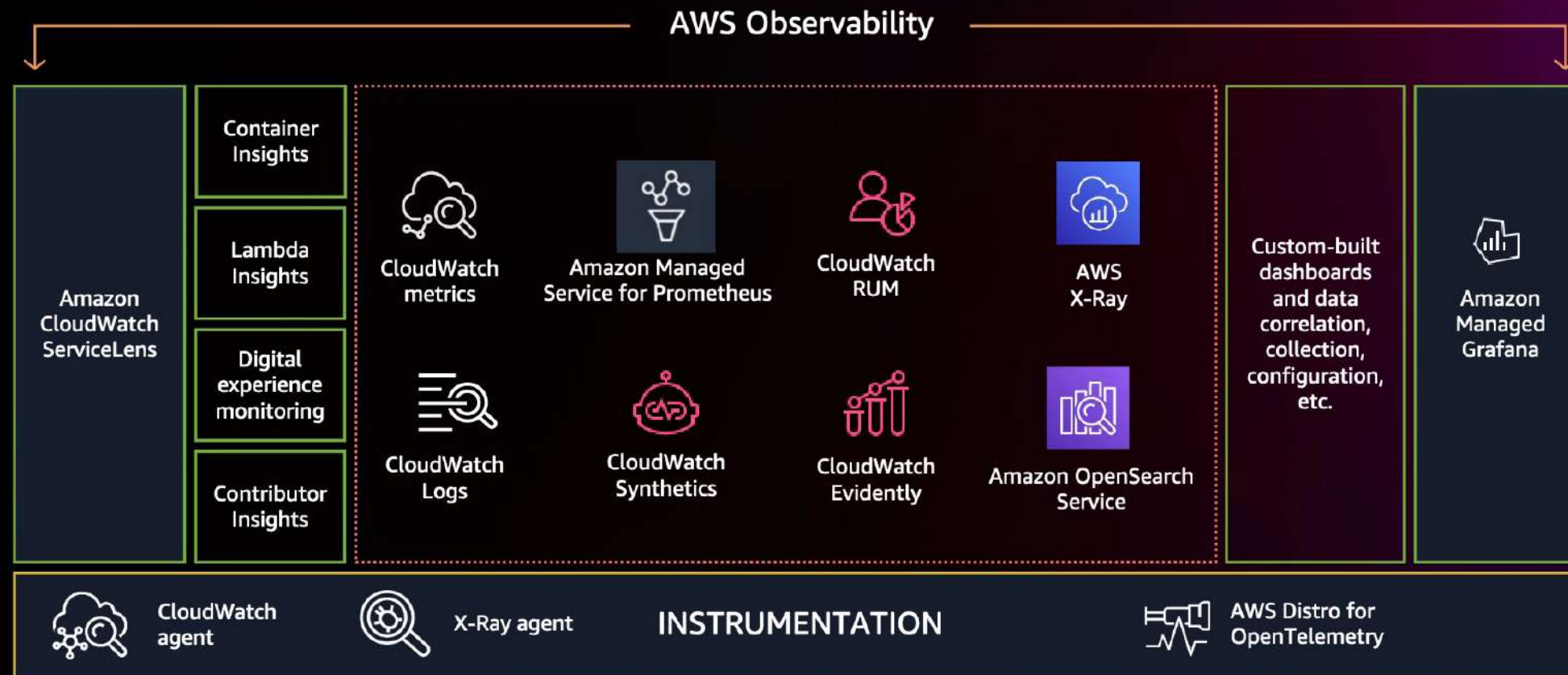
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

<https://youtu.be/EUVRbApIJS8>

Johnson & Johnson: サプライチェーンの可視化を統合的に実現するフルスタックのAWS Observability

バリューチェーンDX
製造・サプライチェーン

Full-stack observability at AWS



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

<https://youtu.be/EUVRbApIJS8>

モデルナ: Amazon Connect と AI を用いて オムニチャネルでのシームレスな顧客体験とデータ活用

バリューチェーンDX
コマーシャル

What is OC3? Moderna's Omnichannel Global Cloud Contact Center platform

THE OC3 PLATFORM DEFINE AND LAY THE FOUNDATION FOR A CONSISTENT CUSTOMER EXPERIENCE IN ALL THEIR INTERACTIONS WITH MODERNA

What is the
OC3 program?



OC3 is a technology transformation that will help Moderna provide a streamlined customer interaction experience across all lines of business and markets

Why is it
important?



- Provide a consistent customers interaction experience
- Transfers control and ownership of customer interaction data to Moderna
- Furthers Moderna's vision of being a data-driven organization
- Takes us closer to AI-driven capture of adverse events and product quality complaints

How
does it
work?



- An integrated Salesforce and Amazon Connect ecosystem for a seamless call flow
- Introduces additional channels such as voice bots, chat, and SMS
- Systems owned by Moderna; resources provided by partners—vendor agnostic



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

- 患者のニーズやその他の環境変化にダイナミックに適応しながら、個別化された体験を提供できる能力の獲得
- 関連するすべての業務部門や市場データなどのあらゆるチャネルにおけるインタラクションの整流化
- データに基づく意思決定できる組織への変革

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

<https://youtu.be/A5HMy9P6xbI>



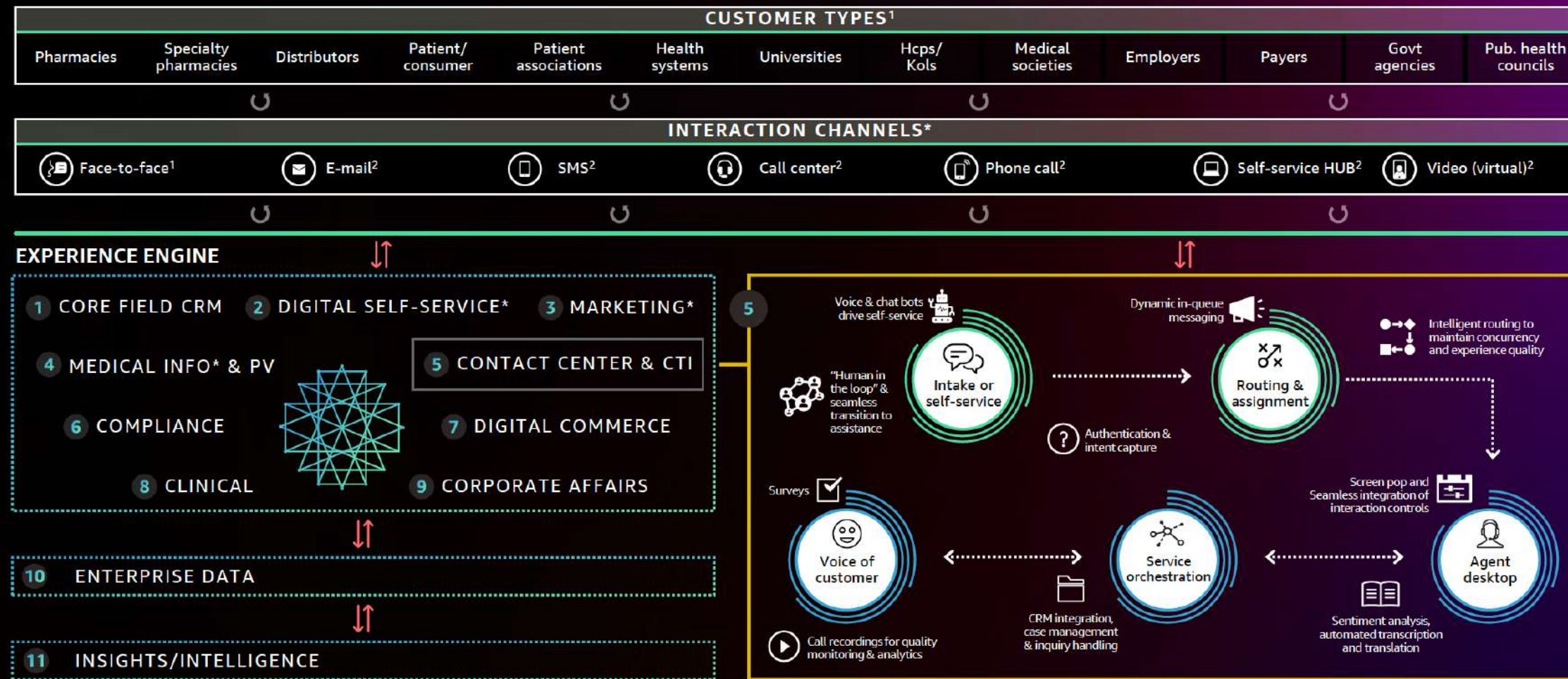
モデルナ: Amazon Connect と AI を用いて オムニチャネルでのシームレスな顧客体験とデータ活用

バリューチェーンDX
コマーシャル

AWS and Salesforce

AN INTEGRATED EXPERIENCE ENGINE, AN ECOSYSTEM APPROACH

¹IN PERSON ²DIGITAL



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



<https://youtu.be/A5HMy9P6xbI>

モデルナ: Amazon Connect と AI を用いて オムニチャネルでのシームレスな顧客体験とデータ活用

バリューチェーンDX
コマース

OC3を活用することで得られたもの

上質な顧客体験

- 顧客の満足度向上
- 従業員（エージェント）の体験向上
- オムニチャネルに対応したセルフサービスの増加

プロセスの標準化

- 標準化されたプロセスをリージョンの状況に合わせて迅速に修正
- コンプライアンスの向上
- メトリクスのトラッキングが可能

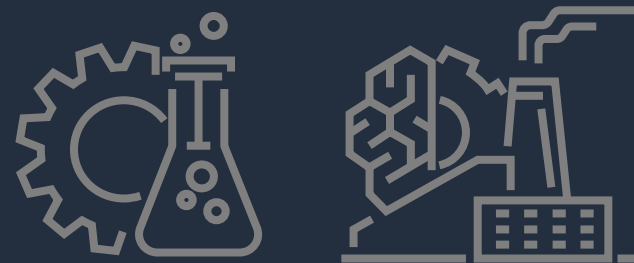
拡張性・柔軟性

- 市場への展開スピード
- 事業展開に応じた拡張性
- モジュール化されたプラットフォームで変化に柔軟に対応

クラウド利用における製薬業界のトレンド



データの民主化



バリューチェーンのDX



リアルワールドデータの活用

AWS Data Exchange (ADX) によりデータの提供・購入が簡便に

aws お問い合わせ サポート

製品 ソリューション 料金 ドキュメント 学ぶ パートナーネットワーク AWS Marketplace カスタマーサポート

AWS Data Exchange 概要 最新の更新 プロバイダー向け 機能 料金 よくある質問 ユースケース

AWS Data Exchange

クラウド内サードパーティのデータを簡単にサブスクリプションする

利用可能なデータ製品を詳しく見る データプロバイダーになるには

延べ 35,000 人が参加する日本最大の AWS イベント「AWS Summit Online」登録開始！

AWS Data Exchange を利用すると、クラウド内のサードパーティーのデータを簡単に検索、サブスクリプション、および利用できるようになります。認定データプロバイダーには、次のような業界のリーダーが含まれます。ロイターは、年間 220 万件を超える独自のニュース記事のデータを複数の言語で収集しています。Change Healthcare は、ヘルスケア事業で年間 140 億以上の取引件数と年間 1 兆 USD の売上を持つ企業です。Dun & Bradstreet は、3 億 3 千万件以上のグローバルな業務記録のデータベースを管理しています。Foursquare は、2 億 2000 万人の固有の消費者から位置データを取得し、それらのデータには 6,000 万か所を超えるグローバルな商業施設が含まれています。

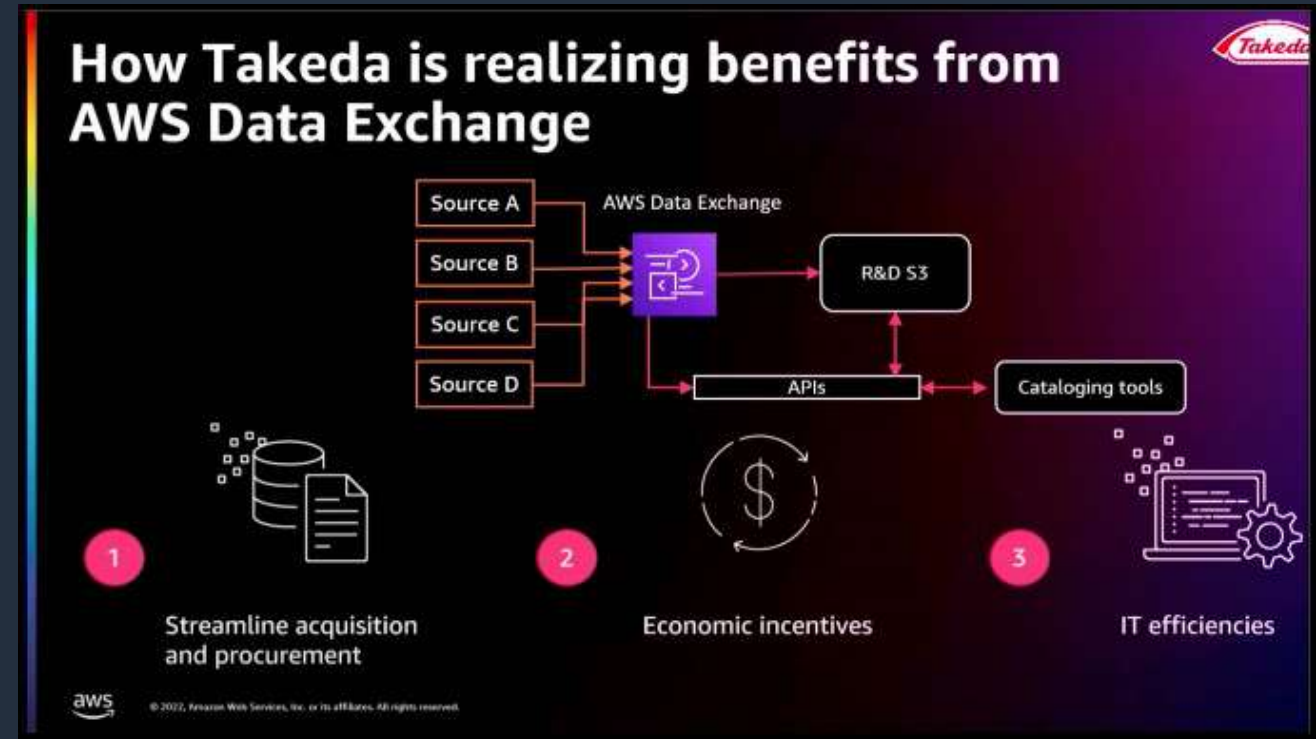
データ製品にサブスクリプションした後、AWS Data Exchange API を使用して Amazon Simple Storage Service (S3) に直接データをロードし、さまざまな AWS 分析および機械学習 (ML) サービスを使用して分析することができます。たとえば不動産保険会社の場合、過去の気象データをサブスクリプションして、いろんな地域の保険適用要件を調整できます。レストランであれば、人口と場所のデータをサブスクリプションして、拡張に最適な地域を特定できます。学術研究者であれば、二酸化炭素排出量データをサブスクリプションして気候変動に関する研究を行うことができます。医療従事者だと、過去の臨床試験の集計データをサブスクリプションし、研究活動を推し進めることができます。

- 数百万のAWSユーザーが、企業が提供する3,600以上のデータを利用・購入できるプラットフォーム
- 利用者はサブスクリプションするだけで、煩雑な契約や環境構築が不要
- テキスト、動画、画像、音声などフォーマットは様々

<https://aws.amazon.com/jp/data-exchange/>

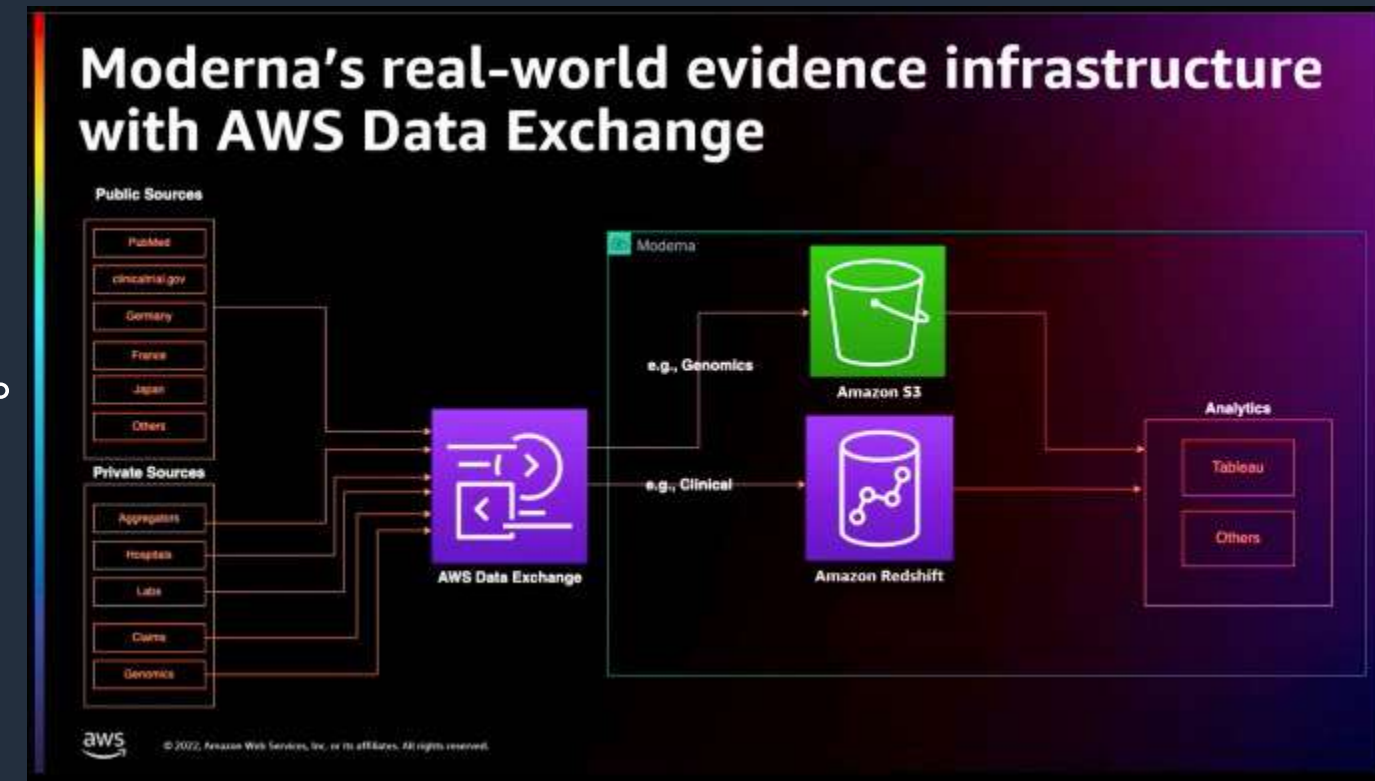
タケダ: ADXを利用したリアルワールドデータ戦略

- リアルワールドデータを活用することで、アンメットニーズを明らかにしたり、治験デザインを最適化したり、経済的価値を明らかにできたりする。
- 通常、ベンダーと会話を始めて、実際にデータが扱えるようになるまで2、3ヶ月かかる。データサンプルにアクセスするだけで2、3週間かかり、そのデータを貯めて、解析する環境を構築するのにさらに時間がかかり、さらにETL処理も煩雑。
- AWS Data Exchangeを活用することで、データ評価から購入、活用までのプラットフォームが統一され、スピードが向上した。
- 従来、複数のデータを管理するために、複数のソフトウェアやスクリプト言語をメンテナンスする必要があったが、ADXにより統合されたデータセットをデータプロバイダーから直接環境に流し込むことができる。



モデルナ: ADXとAmazon Redshiftによる リアルワールドデータ活用

- Public データやプライベートデータ（電カル、ラボデータ）など活用しているがプライベートデータは特に標準化がされておらず、ETL処理に手間がかかる。新たなデータを利用し始めるのに10日ほどかかっていた。
- データへのアクセス方法（FTP、API、など）が無数にあり標準化されていなかった。ETLパイプラインも煩雑。カタログ化もされず、管理工数がかかる。
- ADXにより、Single point of entryが実現。データの種類によりAmazon S3とAmazon Redshiftを使い分け。RedshiftのData Sharingを活用してリアルタイムにデータを利用可能に。
- 3日程度でパートナーを探してデータ取得までが完了する。Evaluateフェーズでは、ADX上でベンダーがサンプルデータを提供しているため、事前に検証が可能。ETL処理も不要となった。



本セッションでお話しした内容

1. 膨大な社内データの活用で広がるデータの民主化

ギリアド、ノボノルディスク、イーライリリー、アッヴィ、ファイザー

2. 機械学習やIoTで進むバリューチェーンのDX

アストラゼネカ、KEK、Lyell、Johnson & Johnson、モデルナ

3. マーケットプレイスで簡素化されるリアルワールドデータ利用

タケダ、モデルナ

Thank you !