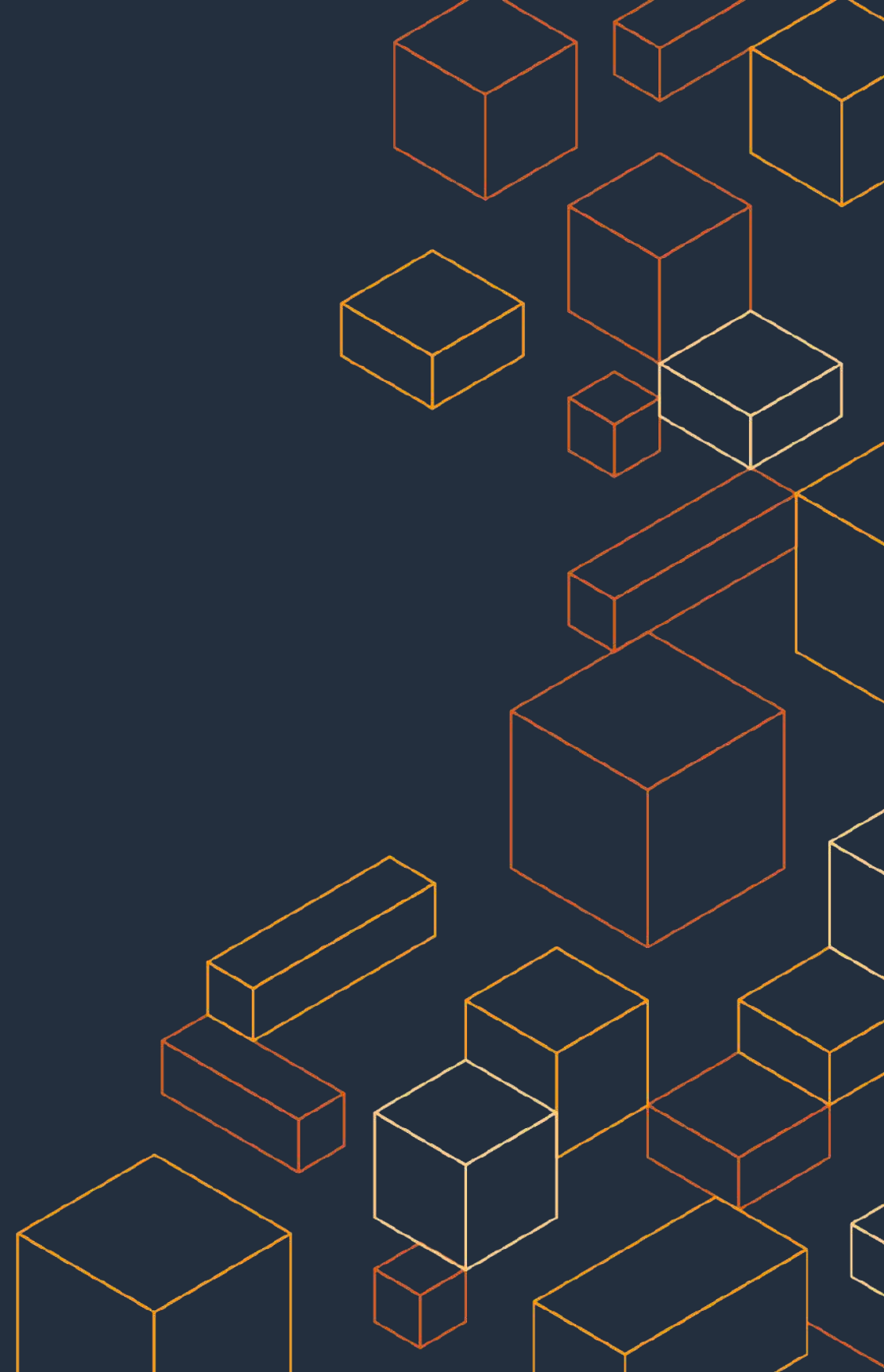




創薬研究領域における 国内外のAWS活用事例ご紹介

2022年4月21日

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 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
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



本日のセッションの対象と目的

対象：創薬研究部門および関連するIT部門、パートナーの皆様
※このセッションは特に「ビジネス課題」にフォーカスしてご説明します。

目的：創薬研究領域のクラウド活用事例を通じて、研究プロセスの効率化、データ活用など創薬研究領域のDX推進のヒントを得て頂くこと。

Our Mission

地球上でもっともお客様を
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)

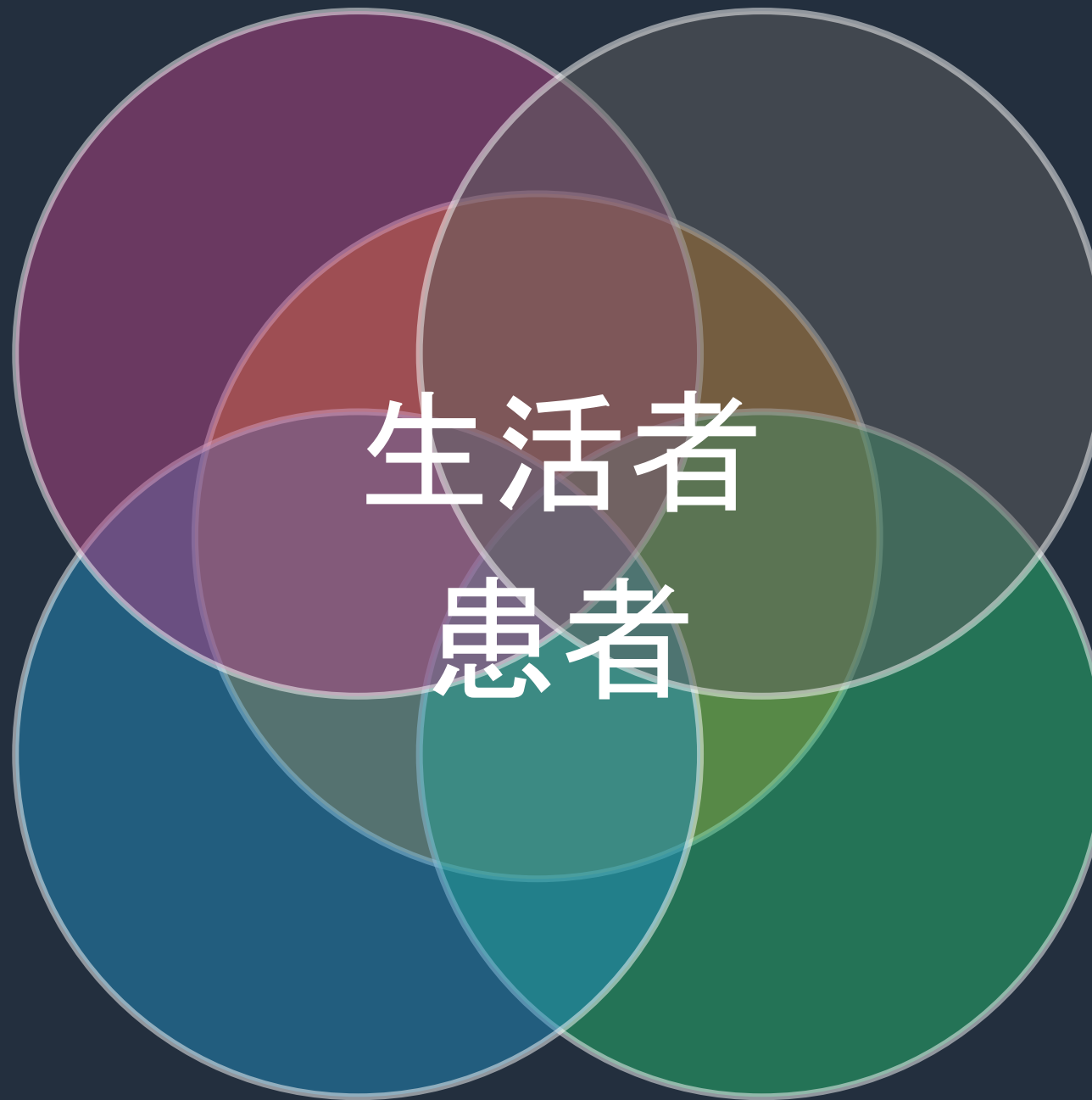
病院
クリニック
歯科
調剤薬局

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～

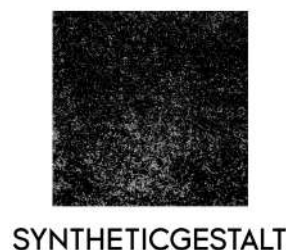
(保険会社)

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:]

JMDC
● + x ◀

Genesis Healthcare



R:RUNNET

FiNC
Technologies



Moff

wiseman

国内における医療関連のお客様

(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



日本医師会ORCA管理機構

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



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



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



Abbott

OMRON



sysmex

Lighting the way with diagnostics

TERUMO

CureApp

SUSMED

Sustainable Medicine

welby



LPIXEL



DeepEyeVision



Dental Systems



Allm SHAPING HEALTHCARE

京都ProMed



Media Contents Factory

MEDLEY

MICIN

Integrity Healthcare

JMDC

● + x ◀

MDV
medical.data.vision



Antaa

スギ薬局

MedPeer



KAKEHASHI

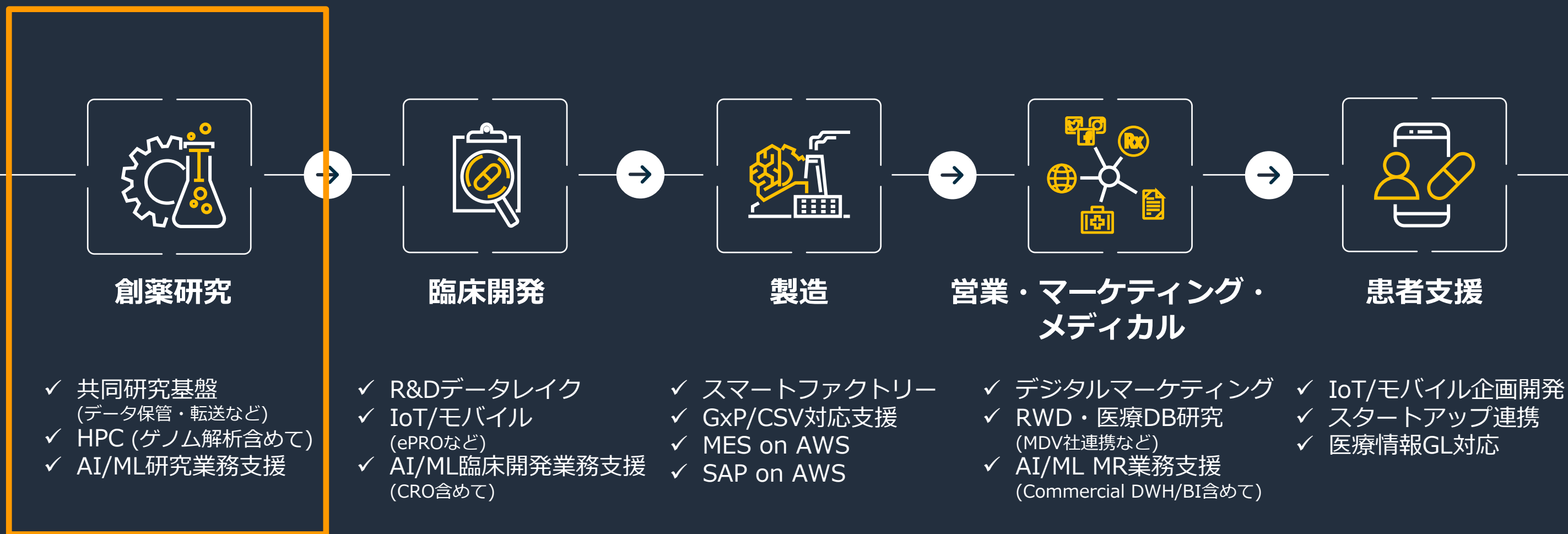
Solamichi System



MG-DX

Medication Guidance
Digital Transformation

製薬バリューチェーンの各段階でご支援しております。



インフラから見た創薬研究領域の課題



データの大容量化に伴う計算リソースの制限



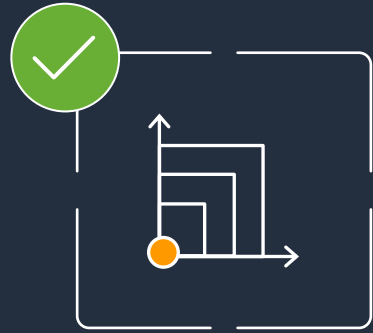
データのサイロ化によるデータ利活用の停滞



属人的な作業による生産性の低下

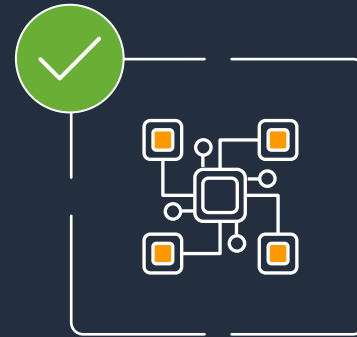
創薬研究領域でAWSを活用するメリット

スケーラブルな 計算環境



利用に応じて
スケーラブルに
規模を拡大しつつ
効率よく高速な
計算環境を提供

ラボデータの連携



研究機器データ等を
蓄積し、知財は守り
つつ社内外に対して
セキュアに
データを共有・活用

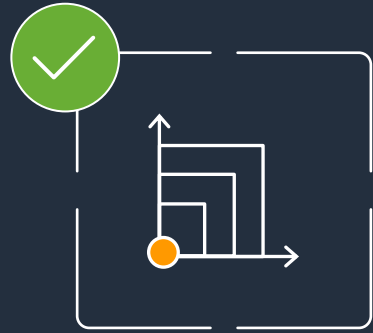
作業の自動化



AI/機械学習で
繰り返しの業務を
効率化、研究者が
より付加価値の
高い業務を行えるように

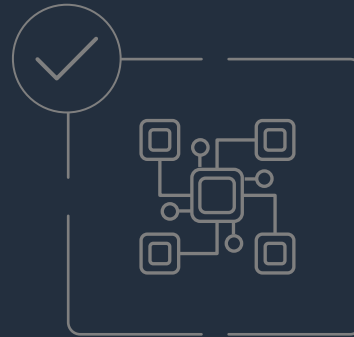
創薬研究領域でAWSを活用するメリット

スケーラブルな 計算環境



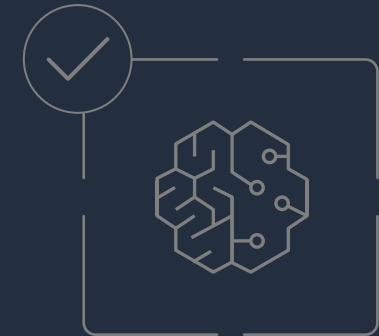
利用に応じて
スケーラブルに
規模を拡大しつつ
効率よく高速な
計算環境を提供

ラボデータの連携



研究機器データ等を
蓄積し、知財は守り
つつ社内外に対して
セキュアに
データを共有・活用

作業の自動化



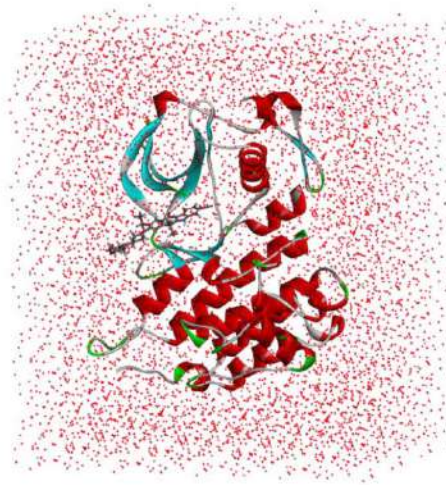
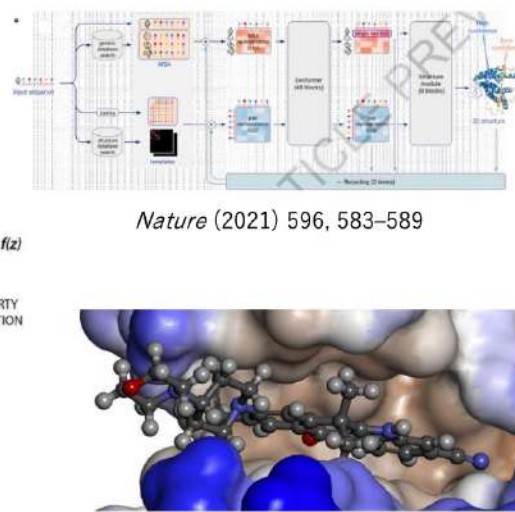
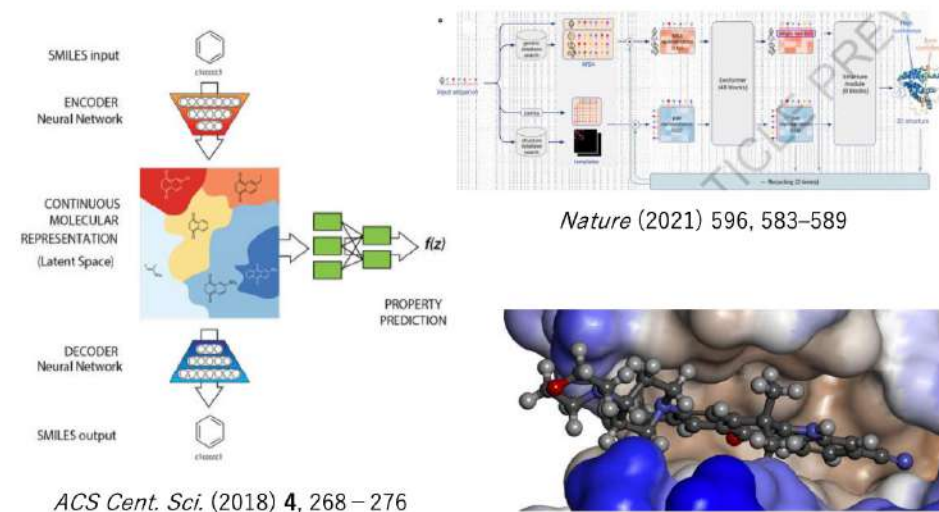
AI/機械学習で
繰り返しの業務を
効率化、研究者が
より付加価値の
高い業務を行えるように

「中外製薬の創薬研究におけるAWS活用事例紹介 –創薬計算化学での活用事例紹介–」

創薬計算化学での活用事例

社内で計算化学を活用した創薬を実施するためには？

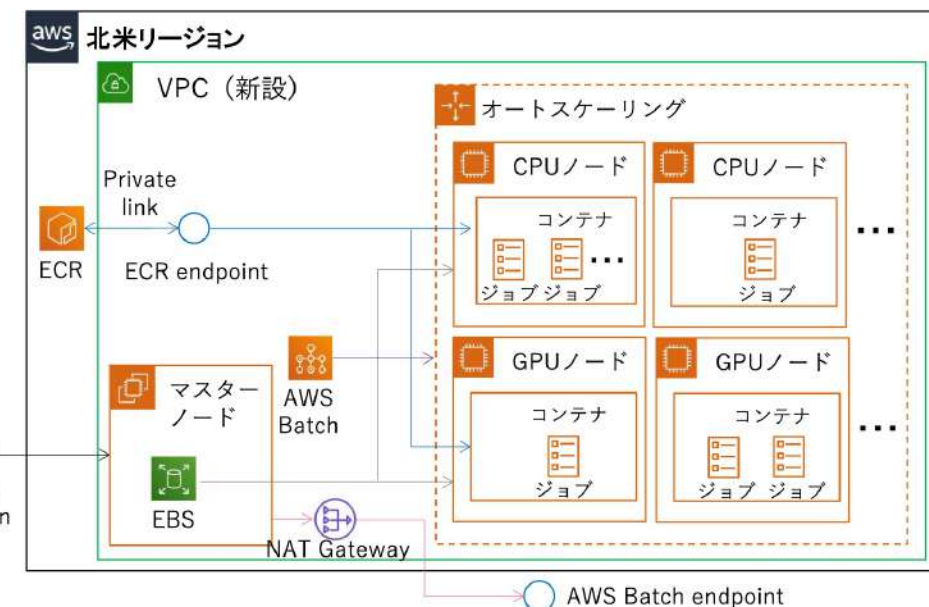
AIや分子シミュレーション関連の高負荷計算を大規模に実行できる環境が必要である。



創薬計算化学での活用事例

創薬計算化学用クラウド環境

AWS Batchを利用した計算化学用クラウド環境を構築した。



「社内外で十分な利用実績があるAWSが第一選択肢になった」 2021年10月26日CBI学会登壇

- ・ 社内既存システムとの連携、構想段階からのAWS担当者との定期的な意見交換
- ・ 豊富な外部情報ソース・ベンダー候補、計算機リソース・サービスの種類が潤沢。創薬関連ソフトウェアの稼働実績が充実。



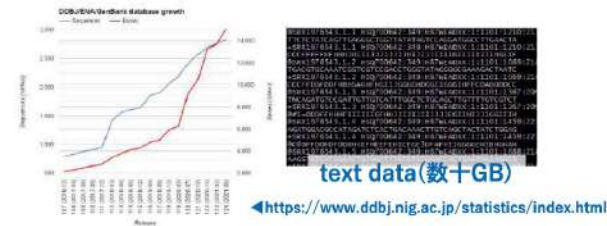
「中外製薬の創薬研究におけるAWS活用事例紹介 –Bioinformaticsでの活用事例紹介–」

NGS解析での活用事例

NGSデータ解析の課題

- NGSデータ解析は「データ量が大きく」、「解析ToolがOSSで開発されている」。

○ データ量が大きく、データ数も右肩上がり



○ ツールの種類が多く、OSSで開発



○ 日々の悩み

- NGSデータのダウンロード時のネットワーク負荷・ディスク容量にリソースが必要。
- データ取得時期に波があり、オンプレの計算リソースが余る。
- 新しいサーバーの導入の度に、解析環境の構築が必要。
- ツールインストールに時間がかかる、バージョンも不定期に更新され、統一しにくい。

本問題を解決するために「コンテナ仮想化・クラウド・Workflow言語」を導入した。



23

NGS解析での活用事例

NGS解析プラットフォーム

- AWS上にNGS解析プラットフォームを構築中、**本格的な運用は年明けを予定**。

○ AWS上に構築した解析プラットフォームの概略図



- ① 実行パラメータの登録を検知し、Cromwellに実行ジョブを投げる。
- ② CromwellにCWLのパイプラインを登録しておく。
- ③ 実行ジョブは、AWS Batchでスケーリングされる。
- ④ 出力結果はCromwell上でID管理され、S3に保存される。
- ⑤ S3を参照して、データ解析を行える。

本プラットフォームで実験・解析・解釈がシームレスにつながる。



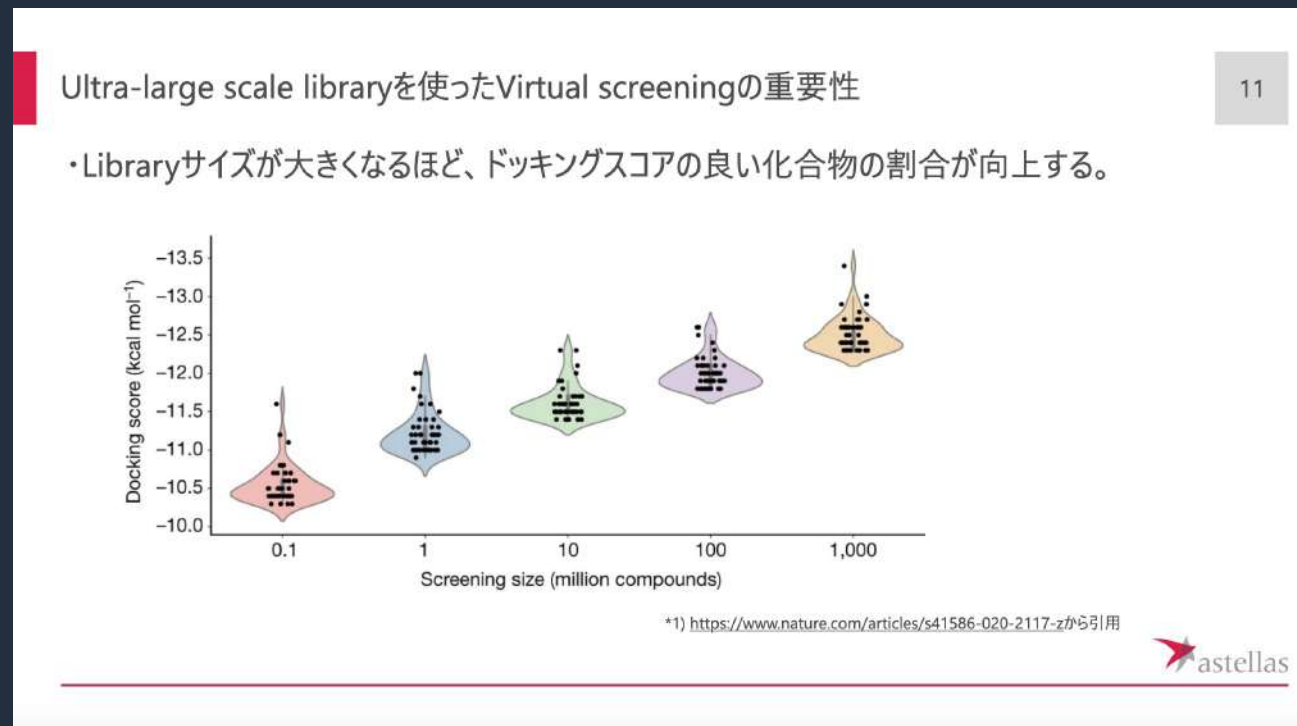
27

クラウド（AWS）に構築するメリット

- スケーラブルな計算リソースとストレージ
- クラウド上の解析結果を用いて、新規サービスの展開が容易
- サーバー自体の保守/運用の軽減



「アステラス製薬の創薬研究におけるAWS活用事例について～クラウドによる機械学習ベースのUltra-large scale virtual screening～」



1. 実現したいこと

数億単位のVirtual screeningをコストを抑えつつ現実的な時間内で実行する。

2. 課題

Enamine REALなど大規模Virtual libraryが利用可能になり、既存の社内サーバーで処理ができない

3. AWSに決めた理由

- ・セキュリティ対応の豊富な実績
- ・AWSの手厚い技術サポートで研究者が環境を構築可能
- ・オンプレミスよりコストが低減

「アステラス製薬の創薬研究におけるAWS活用事例について～クラウドによる機械学習ベースのUltra-large scale virtual screening～」

Ultra-large scale virtual screeningを実施するための二つの方法

15

Approach	Brute Force(BF) based approach	Machine Learning(ML) based approach
Application	VirtualFlow ^{*1} , Giga docking ^{*2}	Active learning glide(AL-Glide) ^{*3}
Method	全化合物をドッキング	1. 少数のドッキング結果から予測モデルを構築し、全化合物のドッキングスコア予測 2. 上位化合物に対してドッキングによるリスコアリングを実施
Pros	Full Dockingのため取りこぼしが無い	機械学習を活用するため、必要な計算リソースはBF-basedに比べて少ない
Cons	必要な計算リソースが膨大(数千CPU以上)	Full Dockingで上位に来る化合物の7割程度の回収率との報告

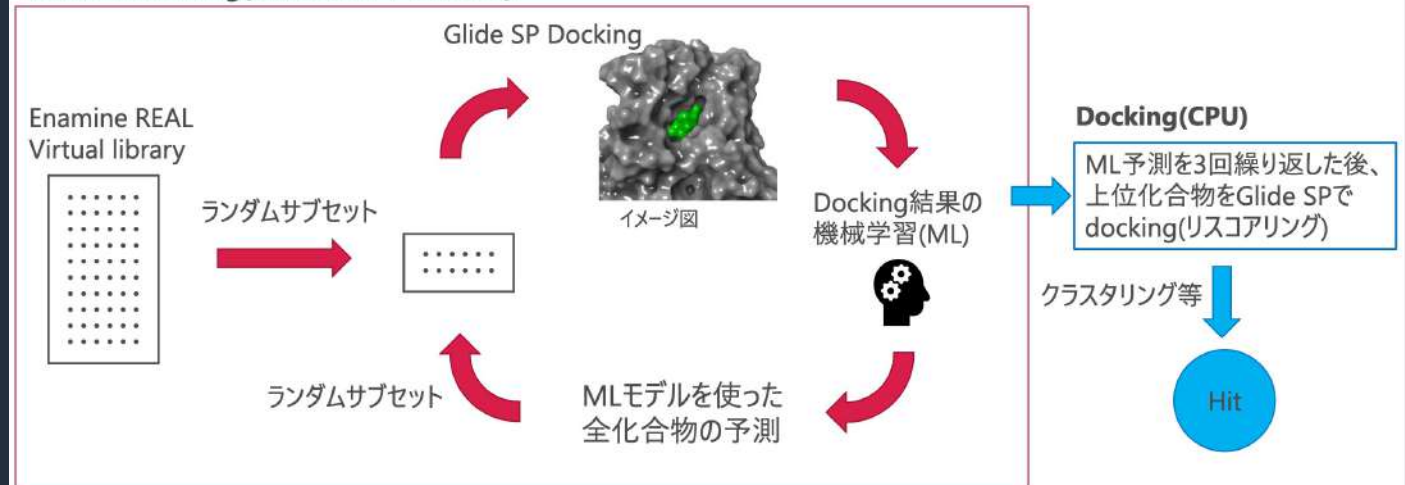


*1) <https://www.nature.com/articles/s41586-020-2117-z> *3) <https://www.schrodinger.com/>
*2) <https://www.eyesopen.com/>

AL-Glideによるvirtual screening概要

23

Machine Learning(予測: CPU, 学習: GPU)



ParallelClusterでスケラブルにCPU, GPUインスタンスを自動管理。



「AWS ParallelClusterのインスタンス変更などもConfigに書き込むだけなので簡単」

- CPU/GPUインスタンスの併用も問題なく活用できた
- コスト最適化のためにSpot instanceを活用し、最後まで計算を完走



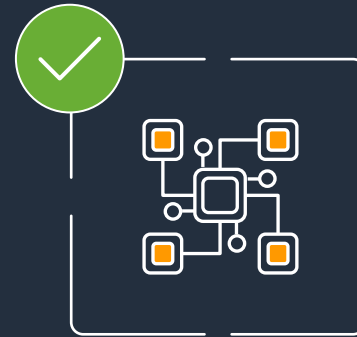
創薬研究領域でAWSを活用するメリット

スケーラブルな 計算環境



利用に応じて
スケーラブルに
規模を拡大しつつ
効率よく高速な
計算環境を提供

ラボデータの連携



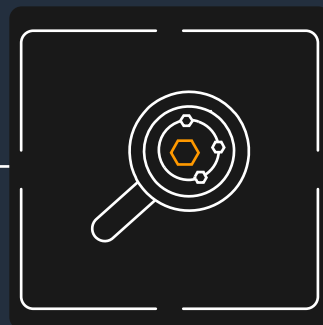
研究機器データ等を
蓄積し、知財は守り
つつ社内外に対して
セキュアに
データを共有・活用

作業の自動化



AI/機械学習で
繰り返しの業務を
効率化、研究者が
より付加価値の
高い業務を行えるように

(1) 創薬研究AI/ML – セルジーン



Research &
Discovery

Lab of
the future

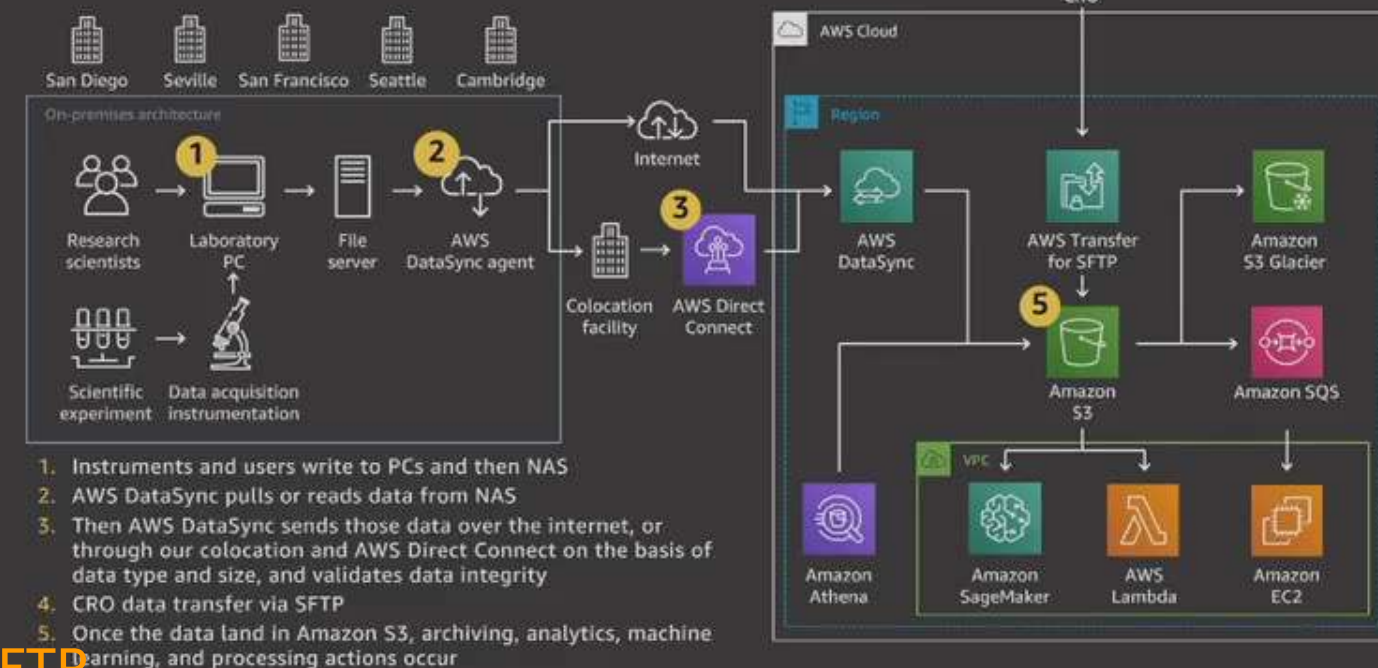


～ハイブリッド・アーキテクチャー～



1. PCからNASにデータ書き込み
2. AWS DataSync でNAS連携
3. データ (種類やサイズ) に
応じて、閉域網 or VPN転送
4. CRO からはAWS Transfer for SFTP
5. S3 に保管されたら、他のアクションが起動

Our hybrid architecture



AWS
re:Invent

(1) 創薬研究AI/ML – セルジーン



Research &
Discovery

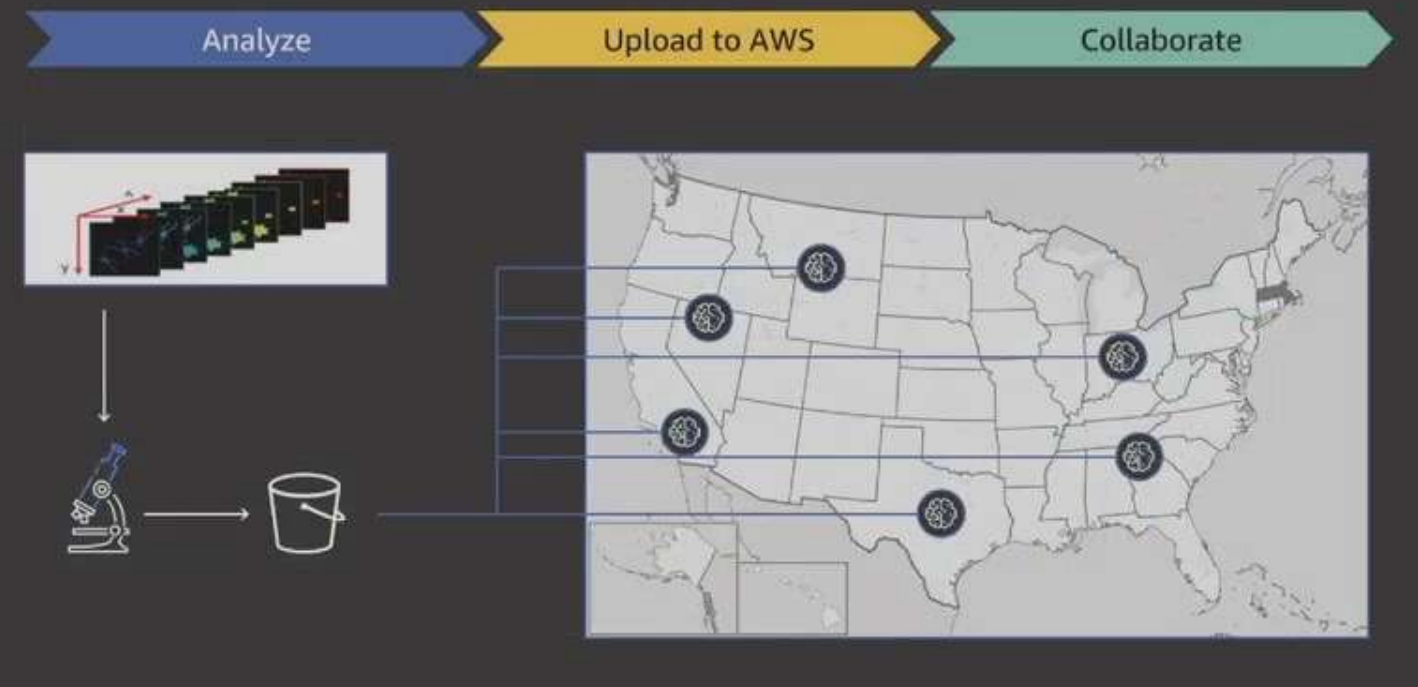
Lab of
the future



～病理画像診断: 病理医とコラボレーション～



Pathology collaboration



AWS
re:Invent

(1) 創薬研究AI/ML – セルジーン



Research &
Discovery

Lab of
the future



～病理画像診断: 解析例～



Cell phenotyping



+



Feature
extraction

Nucleus area
Nucleus compactness
Nucleus DAPI mean
Nucleus DAPI std
Nucleus CD3 mean
Nucleus CD3 std
Cytoplasm area
Cytoplasm compactness
Cytoplasm DAPI mean
Membrane area
Membrane compactness
Membrane DAPI mean
And others

ML
algorithm

Phenotype
prediction

AWS
re:Invent

(1) 創薬研究AI/ML – セルジーン



Research &
Discovery

Lab of
the future

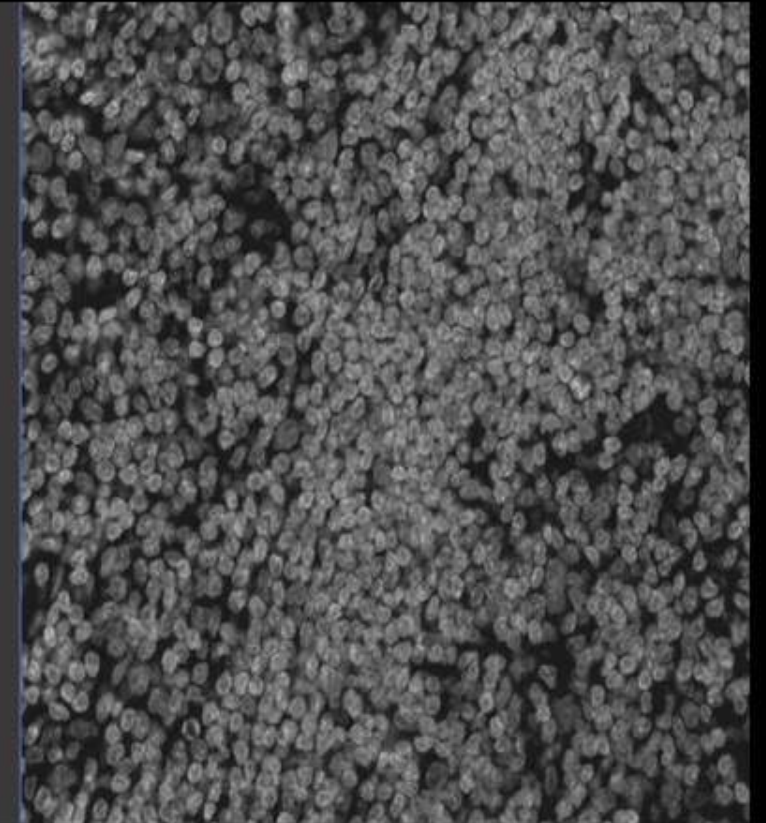


～病理画像診断: チャレンジ～



Challenges

- Lack of training data
- Pixel value variation
- Cells clustering
- Number of cells



AWS
re:Invent

(1) 創薬研究AI/ML – セルジーン

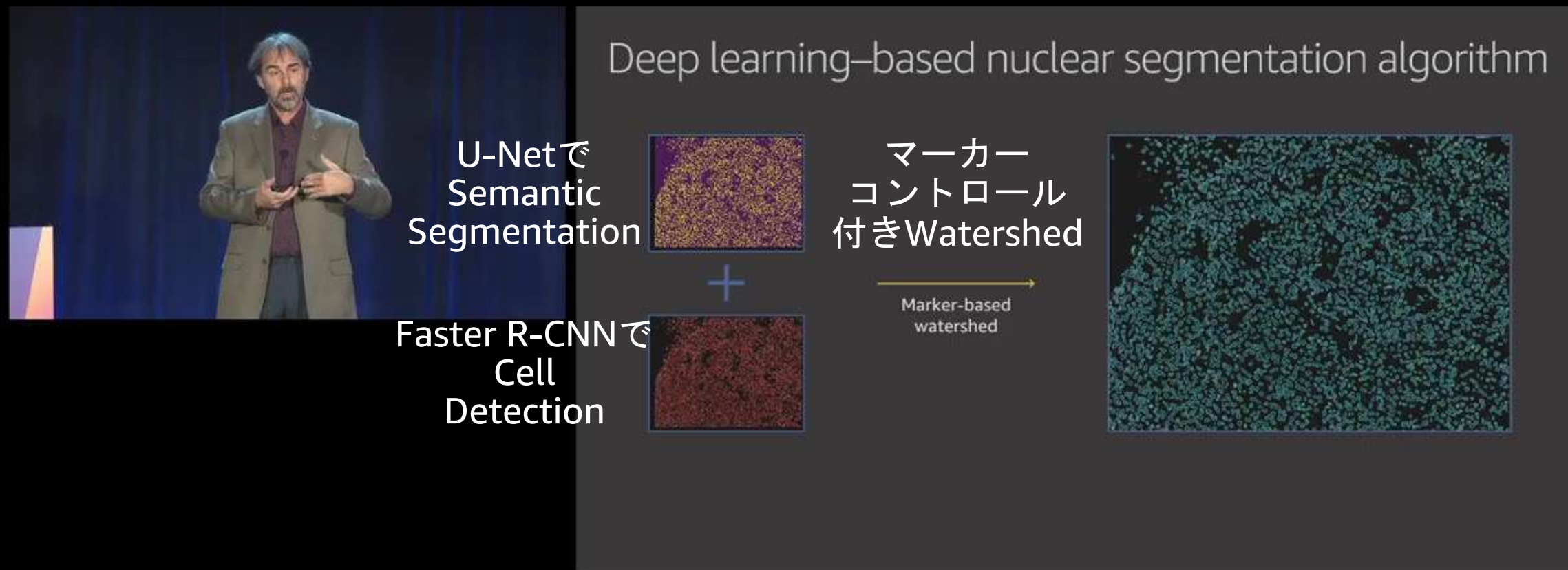


Research &
Discovery

Lab of
the future



～病理画像診断: セグメンテーション結果～



AWS
re:Invent

(1) 創薬研究AI/ML – セルジーン



Research &
Discovery

Lab of
the future



～病理画像診断: 細胞表現型予測の結果～



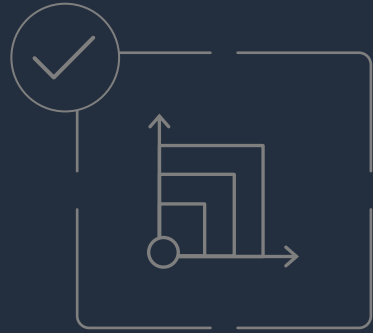
Machine learning model selection for phenotyping

Model	Accuracy (mean)	Accuracy (std)
LASSO regression (L1 penalty)	96.04%	0.83%
Ridge regression (L2 penalty)	95.72%	0.18%
ElasticNet (L1 + L2 penalty)	95.84%	0.55%
Support vector machine	95.84%	0.32%
Decision tree	93.21%	0.62%
Random forest	94.28%	0.71%
Gaussian Naïve Bayes	87.45%	0.86%
Linear discriminant analysis	93.77%	1.08%

AWS
re:Invent

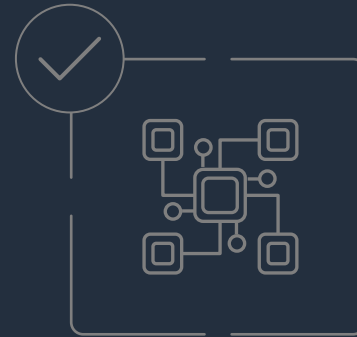
創薬研究領域でAWSを活用するメリット

スケーラブルな 計算環境



利用に応じて
スケーラブルに
規模を拡大しつつ
効率よく高速な
計算環境を提供

ラボデータの連携



研究機器データ等を
蓄積し、知財は守り
つつ社内外に対して
セキュアに
データを共有・活用

作業の自動化



AI/機械学習で
繰り返しの業務を
効率化、研究者が
より付加価値の
高い業務を行えるように

アストラゼネカ： R&D 領域で直面しているチャレンジ

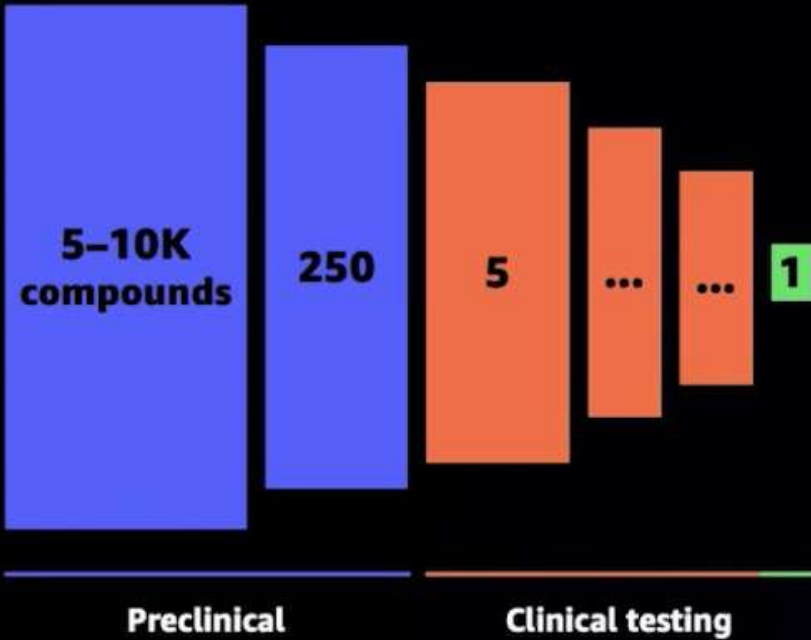


Anna Berg Åsberg
VP R&D IT
AstraZeneca

Biopharma R&D productivity and success rates are an industry-wide challenge

10万の新薬ターゲット/アイデアから新薬にたどり着くのはたったの1個

新薬が患者に届くまでに10年以上かかる



10+ years

アストラゼネカ：R&D ITで創薬プロセスのスピードと精度を向上する



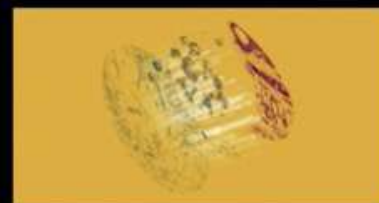
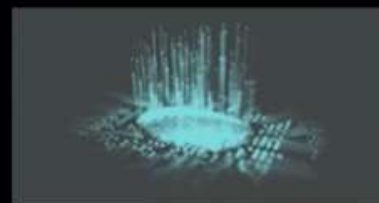
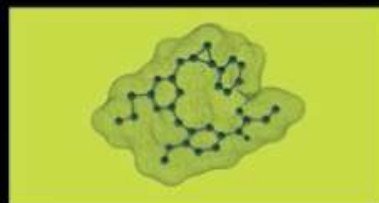
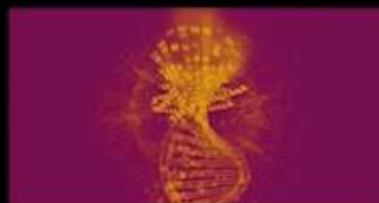
Anna Berg Åsberg
VP R&D IT
AstraZeneca

Accelerating delivery of life-changing medicines by improving speed and accuracy with technology

ターゲット探索

医薬品設計

デジタルラボ



画像分析

臨床試験

自然言語処理

アストラゼネカ：技術的課題・Why AWS・アーキテクチャのTenets

技術的課題

1. サイロ化されたデータ
(Data silos)
2. パフォーマンスやキャパシティ不足
(Limited performance and capacity)
3. AI/ML活用不足
(Limited use of AI/ML)

Why AWS

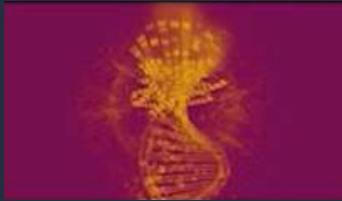
1. スケール
(Scale)
2. 潤沢な計算機資源
(Computer power)
3. 豊富なサービス群
(Service offerings)
4. アーキテクト支援
(Architecture guidance)

アーキテクチャのTenets (教義・信条)

1. マネージドサービス
(Managed services)
2. マイクロサービス
(Microservices-based architecture)
3. オープンソース
(Open source)
4. IaC
(Infrastructure as code)

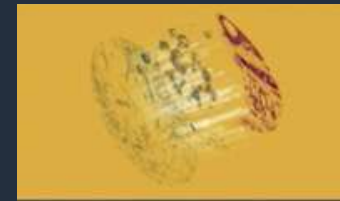


アストラゼネカ：AWSサービスを活用して創薬プロセスの各フェーズで引き起こすイノベーション



ターゲット探索

- ✓ ペタバイトクラスのゲノムデータ解析で、1000億もの統計的な評価を30時間以内に実行



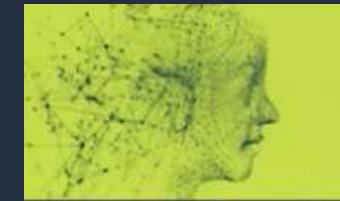
画像分析

- ✓ 機械学習を用いて、病理画像のアノテーションを自動化
- ✓ 研究者はより付加価値の高い作業に従事可能に



医薬品設計

- ✓ AIを用いて分子の設計を簡略化
- ✓ 開発したAIツールを70%以上のプロジェクトで利用



臨床試験

- ✓ 過去の臨床試験データの活用
- ✓ 被験者が家にいながら、データ収集が可能



デジタルラボ

- ✓ 研究員のコロナ禍のリモートワークを可能に
- ✓ 研究機器をクラウドにつなげて、機器データを活用

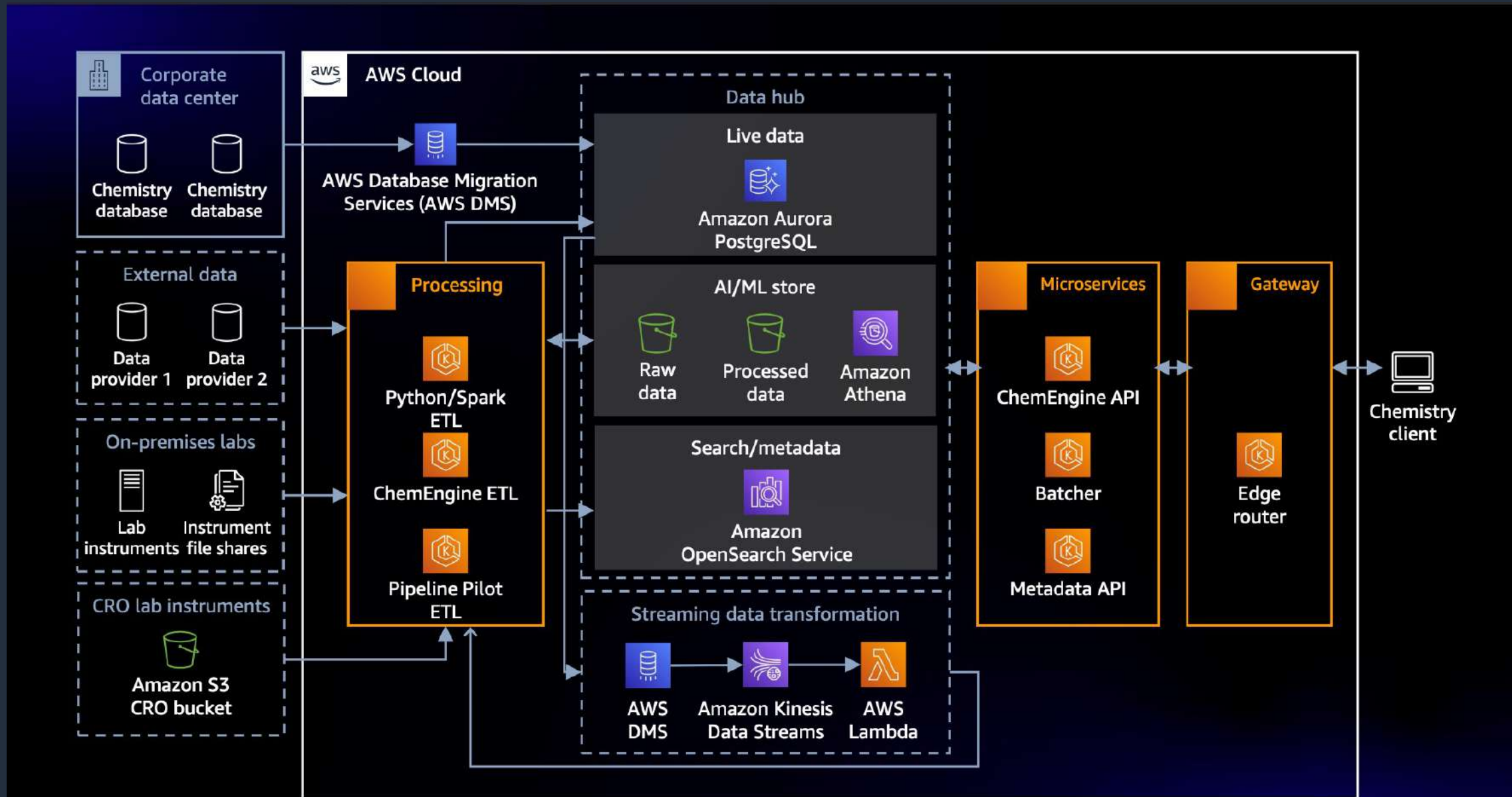


自然言語処理


- ✓ 機械学習で自然言語処理を行うことで、医薬品のフィードバックを迅速に収集

アストラゼネカ：分子フィンガープリント検索


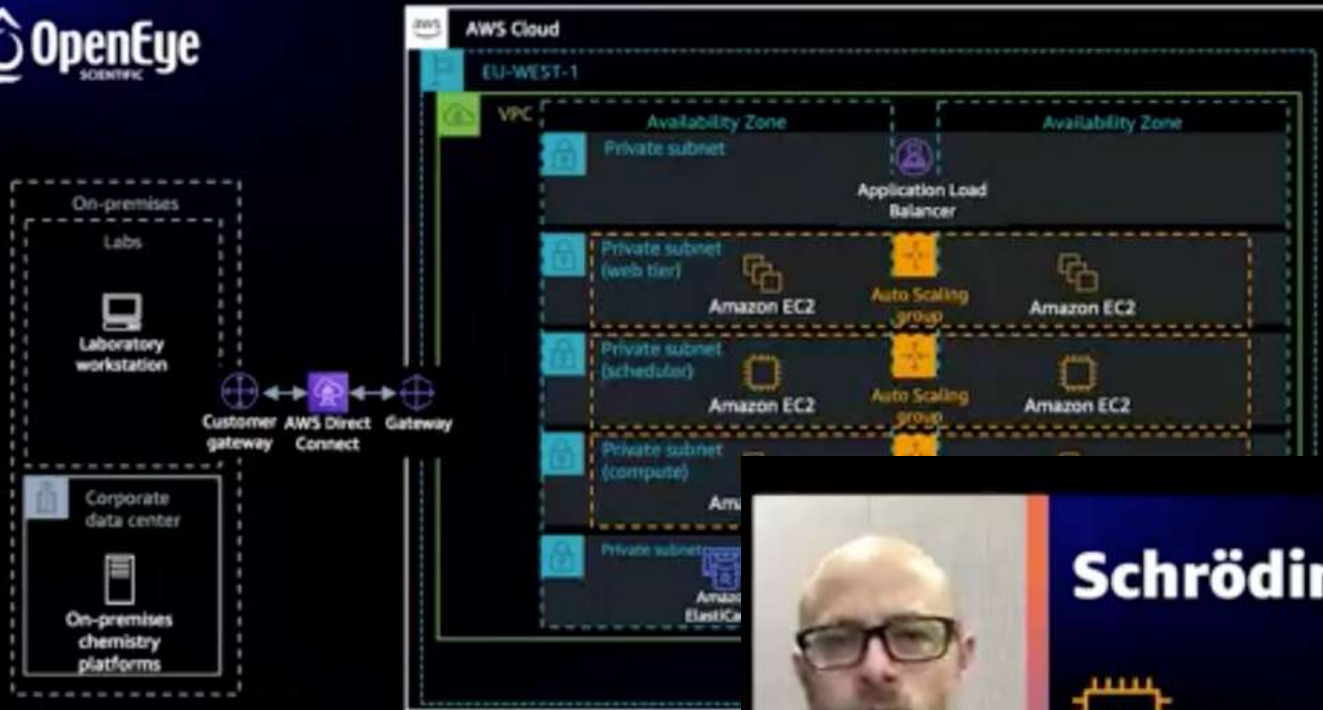

Amazon OpenSearch Service で高速な検索、Amazon EKS と AWS Lambda を用いた ETL のワークロードを実装



アストラゼネカ：パートナーソリューションとの連携



Chris Scott
Global Solutions Architect
AWS



Chris Scott
Global Solutions Architect
AWS

Schrödinger FEP+: Enter the GPU

Amazon EC2

- Up to 8 NVIDIA Tesla V100 GPUs, each pairing 5,120 CUDA Cores and 640 Tensor Cores
- High-frequency Intel Xeon E5-2686 v4 (Broadwell) processors for p3.2xlarge, p3.8xlarge, and p3.16xlarge, and 2.5 GHz (base) Intel Xeon 8175M processors for p3dn.24xlarge
- Provides up to 100 Gbps of aggregate network bandwidth

Potential drug molecules

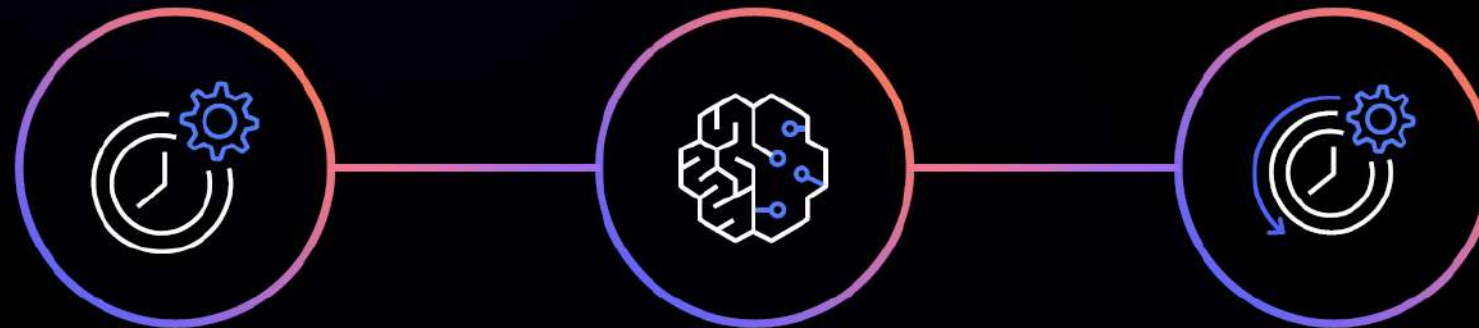
FEP+ active learning

Circa 300,000 FEP+ calculations per annum

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

アストラゼネカ：AI/MLプラットフォームの成果

Outcomes



創薬に要する
時間を
大幅に短縮

200以上の
ホストデータと
AIエンドポイント

モデルの迅速
なデプロイ

AWS で実現するクラウド HPC 2022 春

～最大 65% のコスト削減 HPC 向け Hpc6a インスタンスに Deep Dive～

10:00 - 10:05	オープニング
10:05 - 10:20	「HPC専用で作られたAmazon EC2 Hpc6aインスタンス」 アマゾン ウェブ サービス ジャパン合同会社
10:20 - 10:35	「HPC インスタンスで活用されるプロセッサ技術概要」 日本 AMD 株式会社
10:35 - 10:55	「解析ソフトウェア性能を最大化するにはクラウドが必要」 シーメンス株式会社
10:55 - 11:15	「Cradle CFD のHPCチャレンジとAWS Hpc6aによる性能評価」 ヘキサゴンマニュファクチャリングインテリジェンス
11:15 - 11:35	「AWS HPCを用いた高精細気象予報モデルの構築と検証」 メトロウェザー株式会社
11:35 - 11:55	「並列陽的時間領域FEMによる波動音響シミュレーションのクラウドHPC環境における計算性能」 株式会社 安藤・間 技術研究所
11:55 - 12:00	Q&A

従来のコンピューティング最適化インスタンスに比べて**最大 65% 優れたコストパフォーマンス**を提供するHPCワークロード向けAmazon EC2 インスタンス、**Hpc6a の提供開始**。実際の活用事例を中心にご紹介します。



<https://pages.awscloud.com/JAPAN-event-OE-jp-EIB22-WWSO-CMP-HPC-20220428-reg-event.html>





Thank you!