

# 生成AIデータ基盤としての AWSマネージドアナリティクスサービス

平井 健治

SOLUTIONS ARCHITECT

2024/05/16



# 自己紹介

名前：平井 健治（ひらい けんじ）

所属：アマゾン ウェブ サービス ジャパン  
ソリューションアーキテクト

流通小売業のお客様の技術支援を担当

経歴：データベース、データ活用に関するコンサルタント

好きなAWSサービス：Amazon S3、AWS Lambda、  
Amazon Redshift、Amazon DataZone



# はじめに：生成AIを有効に活用するためは？

Amazon Bedrock <

Amazon Bedrock > Chat playground

Chat playground Info Load examples Compare mode

**AI Claude 3 Sonnet v1** | ODT Change

自社のデータを用いて生成AIを有効に活用するために必要なことについて教えてください。

自社のデータを用いて生成AIを有効に活用するために必要なことは以下の5点です。

1. 高品質でラベル付きの大量のデータセットの確保
2. データの前処理とクレンジングによる、ノイズの除去と整形
3. 適切な生成AIモデルの選定と、ハイパーパラメータの最適化
4. 大規模な計算リソース(GPU、高速ストレージ等)の確保
5. 継続的なモデル改善のためのモニタリングと再学習プロセスの構築

これらの点に注力することで、自社のデータから最大限の価値を引き出し、生成AIを事業に効果的に活用できるようになります。

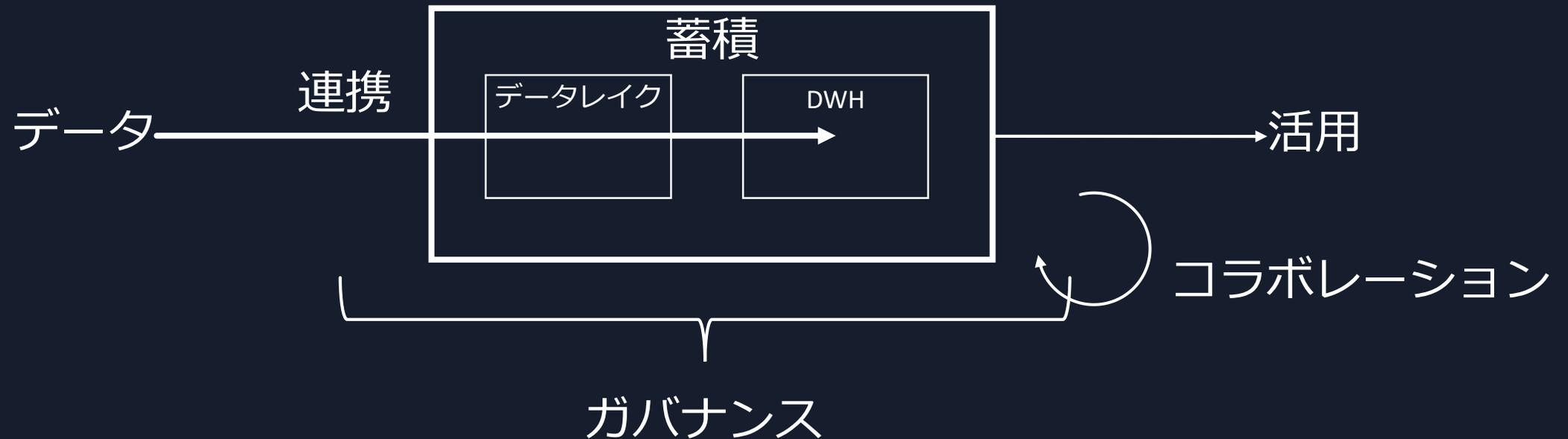
Write a prompt... (Shift + ENTER to start a new line, and ENTER to generate a response) Image Run

**アナリティクス環境の整備が重要**

# 本セッションの内容

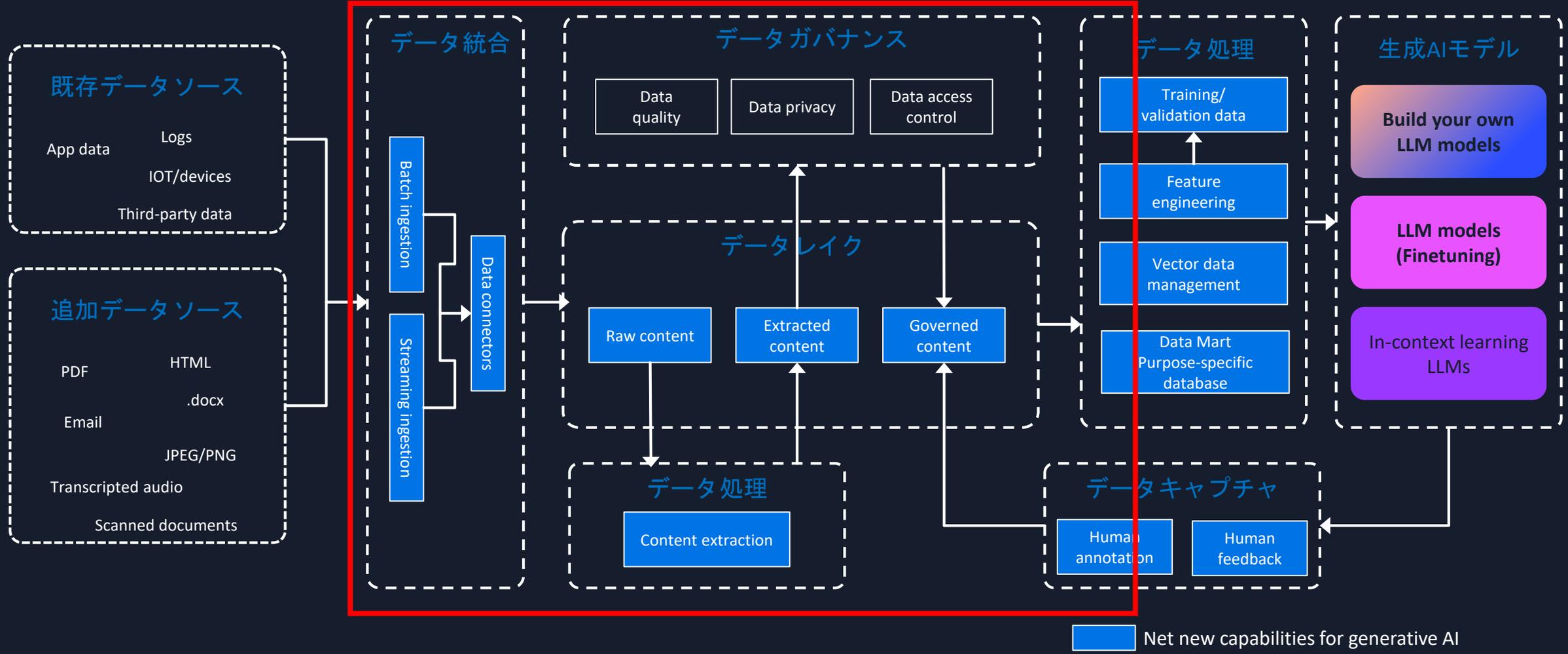
生成AIの活用に重要となるアナリティクス環境について、必要な要件と、要件を実現するAWSマネージドアナリティクスサービスを、以下の順に紹介します。

蓄積、連携、ガバナンス、コラボレーション



# 本セッションの範囲

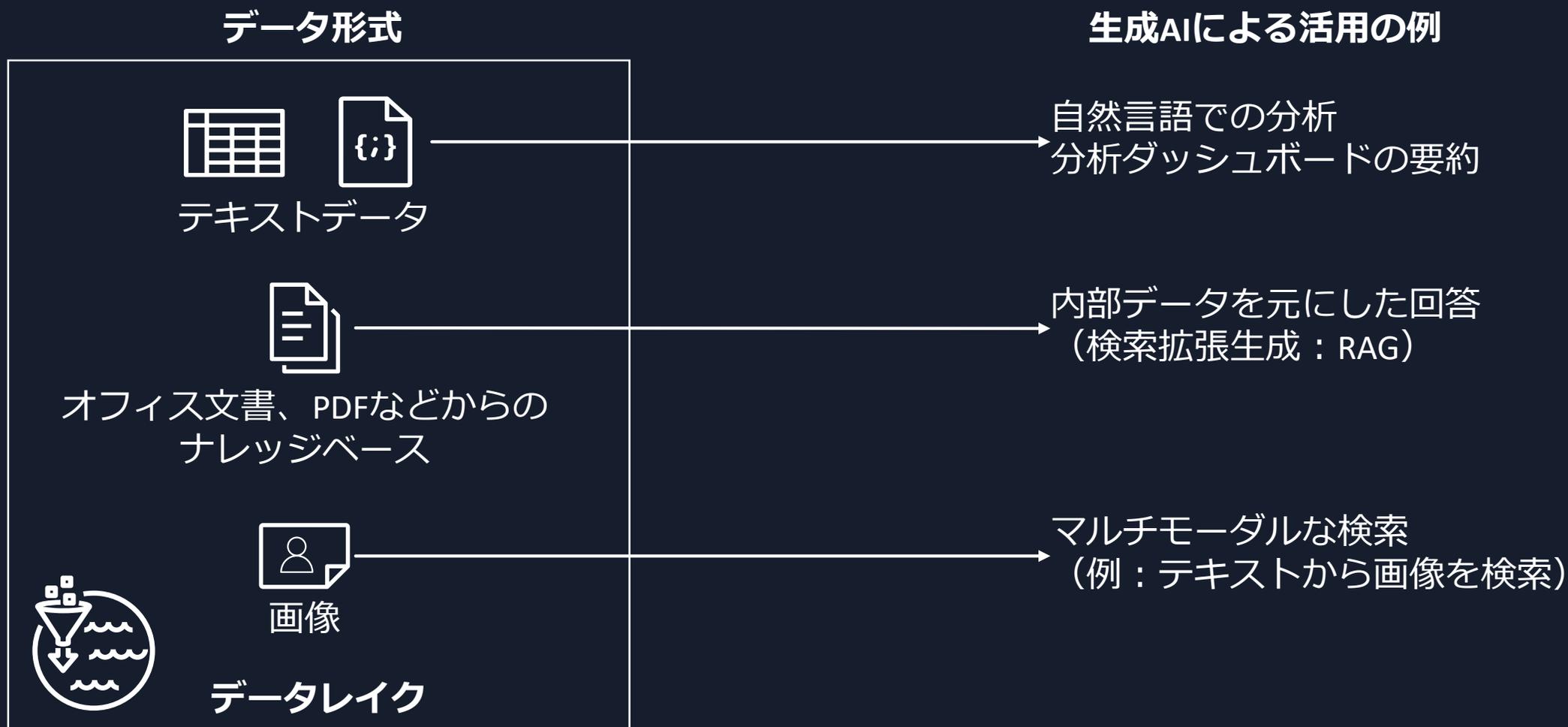
生成 AI アプリケーションにおける AWS データ サービスの役割、との位置付け



# データの蓄積



# 生成AIにより多様な形式のデータ活用が容易に



多様、大量のデータを保存することがデータレイクにますます期待される

# データレイクの基盤となる Amazon S3

拡張性と耐久性を兼ね備えたクラウドストレージ

上限無し：サイジング不要

高い耐久性：99.999999999%

安価：

\$0.025/GB/月\*（スタンダード）

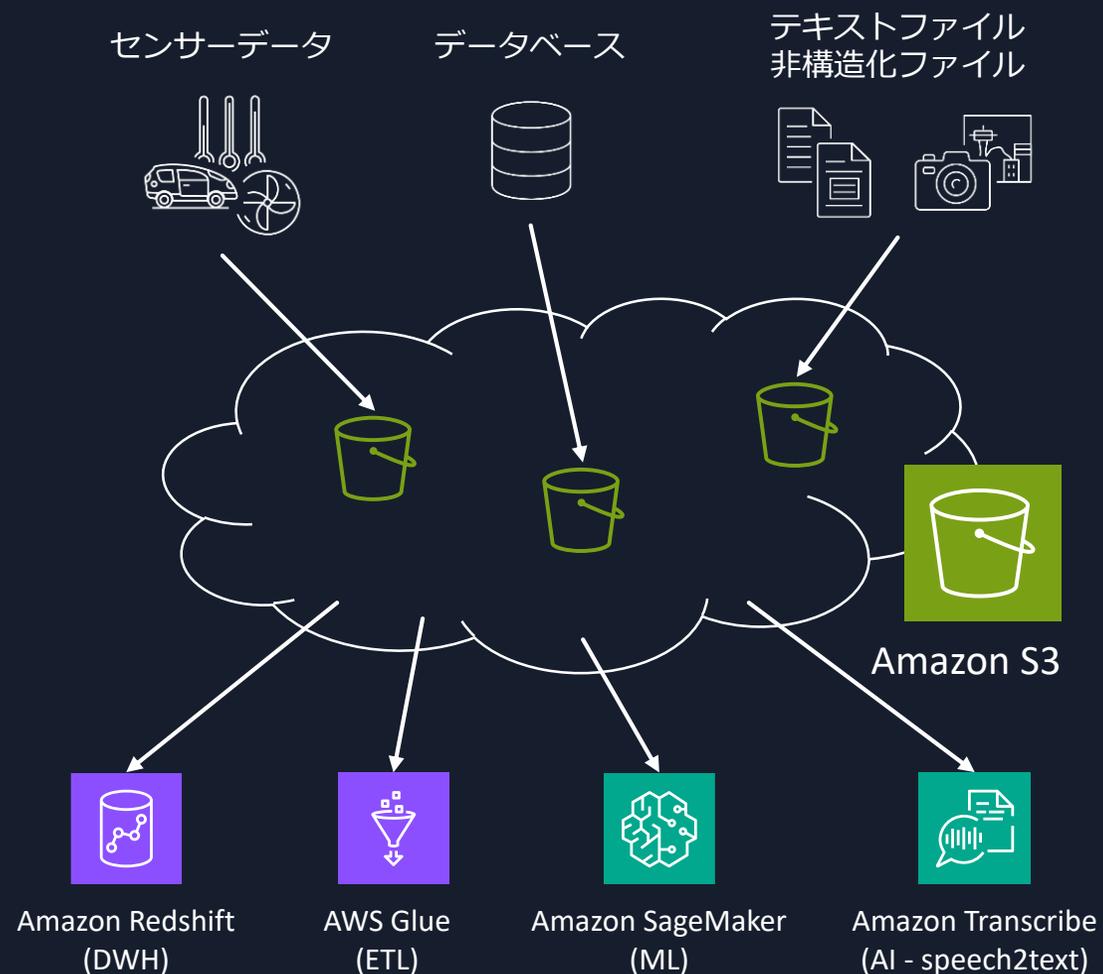
\$0.0138/GB/月\*（標準-低頻度アクセス）

APIアクセス：

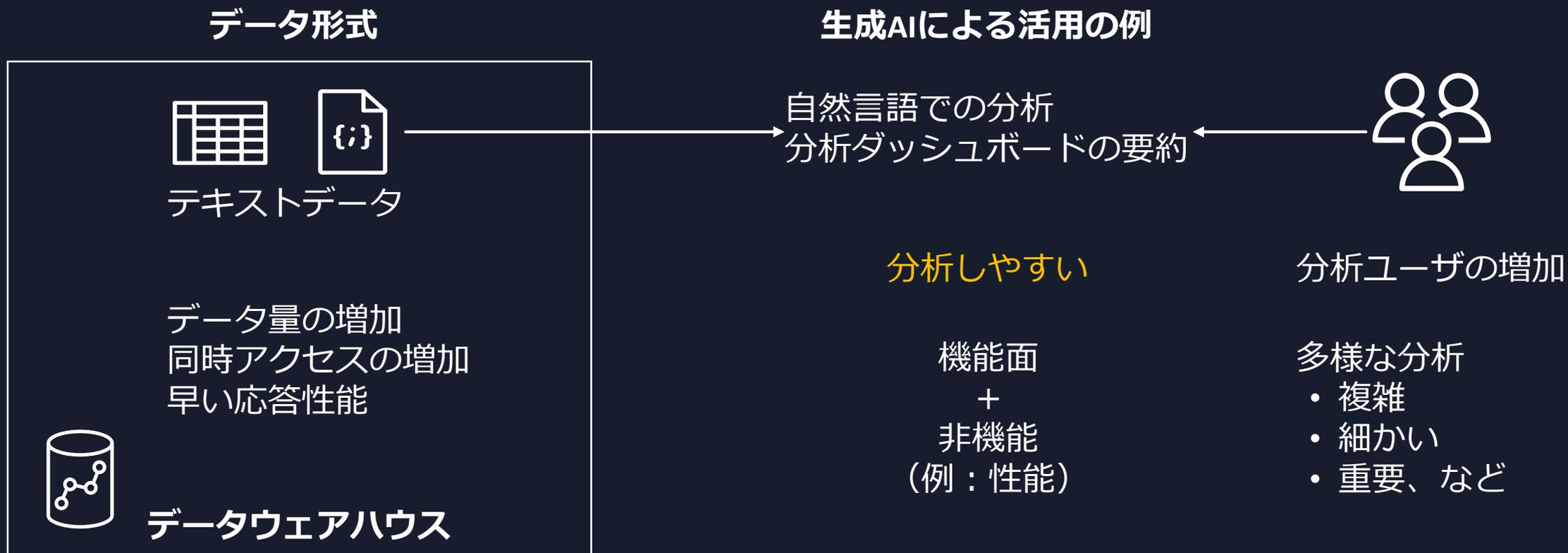
多様な言語にライブラリを提供

AWS各種サービスと連携

\* 2024年5月時点の東京リージョン価格



# 簡単に分析できるようになるとどうなるか？



複雑、あるいは細かい分析を可能とするデータウェアハウスが用意される

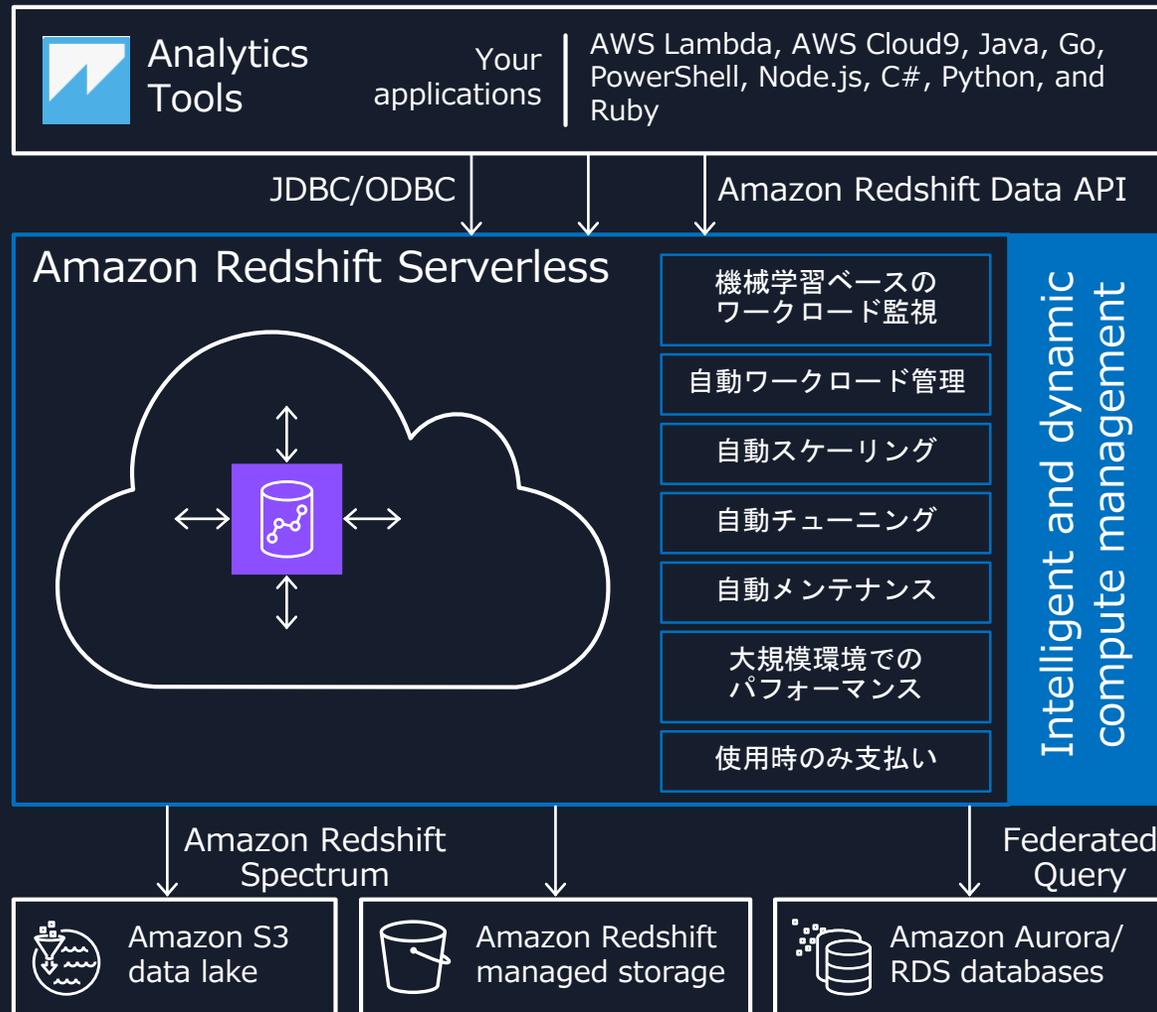
# クラウドデータウェアハウス Amazon Redshift

コストパフォーマンスが高く、数万の顧客を抱え最も広く使用されているクラウドデータウェアハウス

誰でも簡単に分析できる：  
インフラストラクチャを気にせず、数秒でデータからインサイトを得ることに集中

全てのデータを分析する：  
DWH、データレイク、オペレーショナルDBに跨るあらゆるデータに対し、データの移動なしにリアルタイムな分析や予測分析が可能

高いコストパフォーマンス：  
他のクラウドデータウェアハウスに比べて最大 3 倍のコストパフォーマンスを実現



# Amazon Q generative SQL in Amazon Redshift

## 自然言語で入力すると推奨 SQL クエリを出力

Amazon Redshift クエリエディター v2 に搭載

テーブル名やカラム名など Amazon Redshift 内のデータ構造を理解した上でクエリを生成

「地域別の総収益を調べる」といった意思を示すと、それを実現するクエリが自動生成される

クエリ、データ、スキーマなどの情報は安全かつプライベートに管理され、生成 AI 基盤モデルの学習には使用されない

バージニア北部とオレゴンにてパブリックプレビュー中  
プレビュー版の提供期間中は、Amazon Q の生成 SQL を無料で利用可能



Which state has the most venues?

```
1 SELECT
2   venuestate,
3   COUNT(*) AS num_venues
4 FROM
5   tickit.venue
6 GROUP BY
7   venuestate
8 ORDER BY
9   num_venues DESC
10 LIMIT
11  1
```

Regenerate SQL [Add to notebook](#)

Ask me a question to generate a SQL suggestion for your connected database. [➔](#)

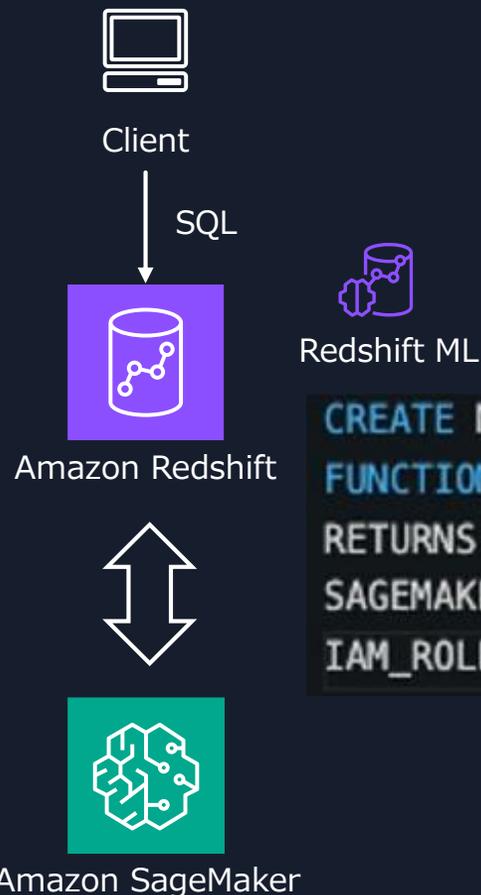
# Amazon Redshift ML で大規模言語モデル (LLM) 対応

Amazon Redshift ML で大規模言語モデル (LLM) がサポートされ、より高度な処理を実現

Amazon SageMaker JumpStart で利用できる事前定義済みのモデルや、カスタムモデルをデプロイして利用可能

## ユースケース例

製品のフィードバックデータに対してフィードバックの要約、エンティティの抽出、感情分析、製品フィードバックの分類



```
CREATE MODEL my_llm_model
FUNCTION my_llm_model(super)
RETURNS super
SAGEMAKER '<endpoint_name>'
IAM_ROLE default/'<iam_role>';
```

東京を含む5つのリージョンでプレビュー中



# Amazon Redshift Serverless AI 駆動スケーリングと最適化

従来と比較して最大 10 倍の価格性能比

## インテリジェントな自動スケーリング

- データ量やクエリの複雑さなどの動的なワークロードの変化に合わせて最適化

## より優れた価格性能比

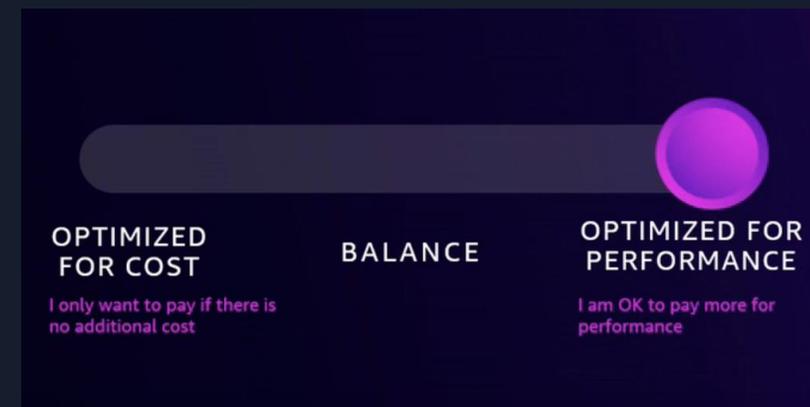
- コストを最適化しながら一貫したパフォーマンスを維持する、機械学習主導の総合的な最適化

## 希望する価格性能目標を設定

- コスト最適化、バランス重視、パフォーマンス重視のシンプルな3つの軸で柔軟に調整



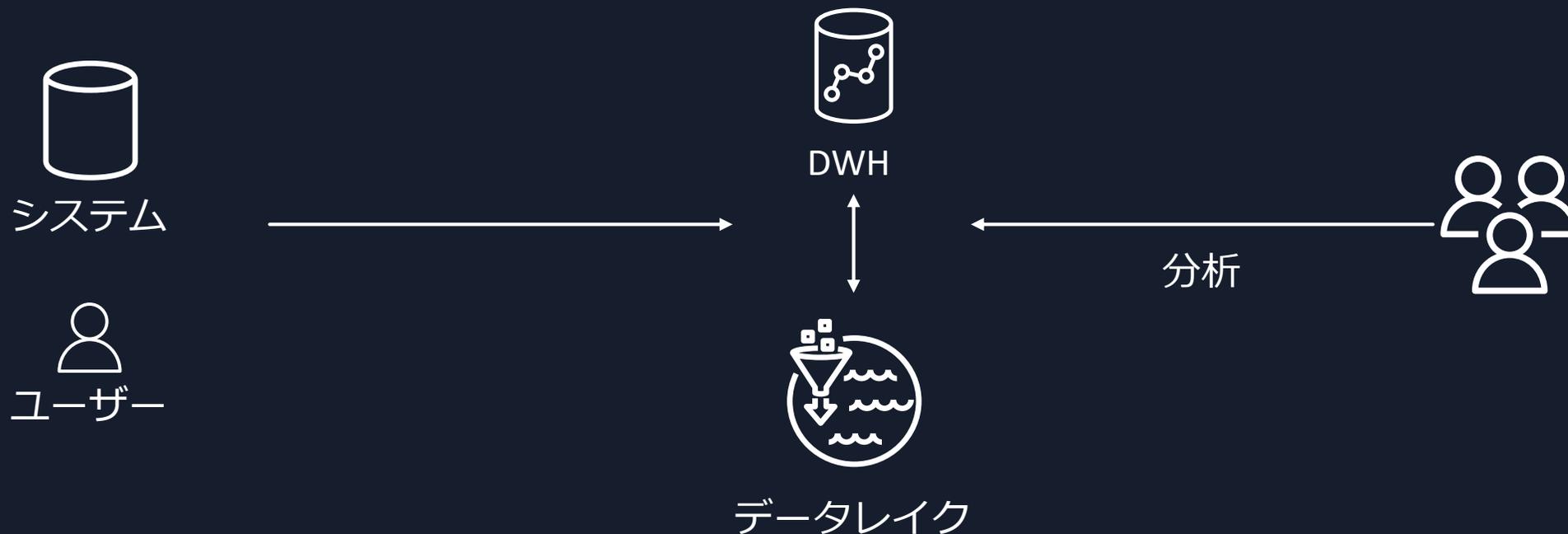
Serverless



# データの連携



# 分析するデータは分析環境に取り込む必要がある



手間をかけずに鮮度の高いデータを分析したい

# Zero-ETL でデータを統合する

Zero-ETL : パイプラインの管理無しに、ペタバイト級のトランザクションデータのニアリアルタイム分析を可能にする簡単でセキュアな方法

従来のデータ連携 (ETL)



Zero-ETLによるデータ連携



Amazon Redshift 以外も含む Zero-ETL の全体像 : <https://aws.amazon.com/jp/what-is/zero-etl/>

# Amazon Redshift Zero-ETL 統合の範囲拡大

Zero-ETL : パイプラインの管理無しに、ペタバイト級のトランザクションデータのニアリアルタイム分析を可能にする簡単でセキュアな方法



# ETLが必要なケースは存在する

## 従来のデータ連携（ETL）



例えば、

- Zero-ETLでは対応していないオンプレミスのシステムなどとデータを連携したい
- クレンジングなど、分析するための前処理を実施してデータレイクに保存したい

## Zero-ETLによるデータ連携



# データインテグレーションサービス : AWS Glue

あらゆる規模ですべてのデータを検出、準備、統合するシンプルかつサーバーレスなETLサービス

## スケーラブルなデータ インテグレーションエンジン



組み込みデータ変換



多様なジョブ実行エンジン  
Apache Spark, Python shell, Ray, Streaming



ジョブのモニタリング

## コネクタとデータ収集



コネクタ



コネクタのマーケットプレイス



多様なインターフェース

## 生産性の高い データオペレーション



多様なペルソナへの対応



生産性向上



DataOps

## 一元化された 統合データガバナンス



データカタログ



データ品質



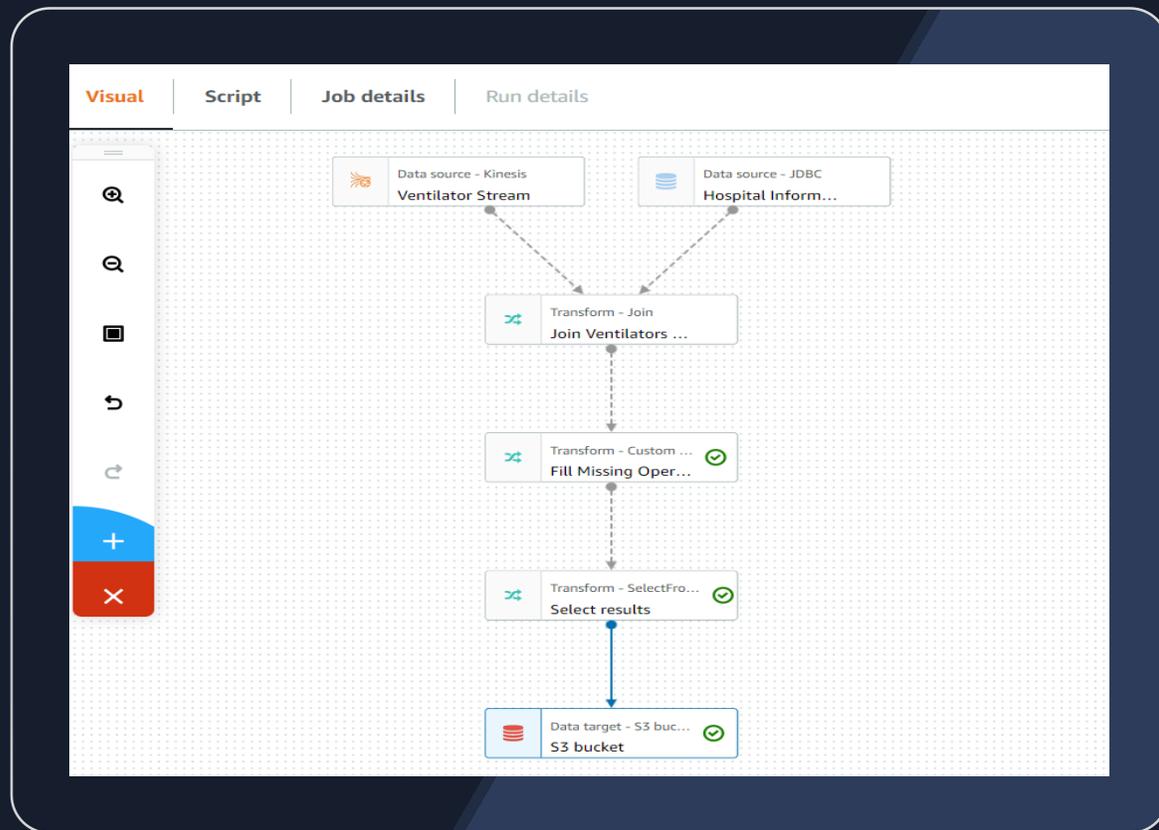
クローラー



AWS Lake Formation  
連携

# AWS Glue Studio

AWS GLUE の ETL ジョブを GUI で簡単に作成・編集



ETL ジョブの作成、実行、および  
モニタリングを簡単にできるようにする  
グラフィカルインターフェイス

GUI で作成したフローから  
ソースコードを生成

データ抽出、出力、ジョイン等の汎用的な  
処理はビルトインされノンコーディングで  
実現

Transform nodeを配置することで、**任意の  
コードやSQL**を挿入可能

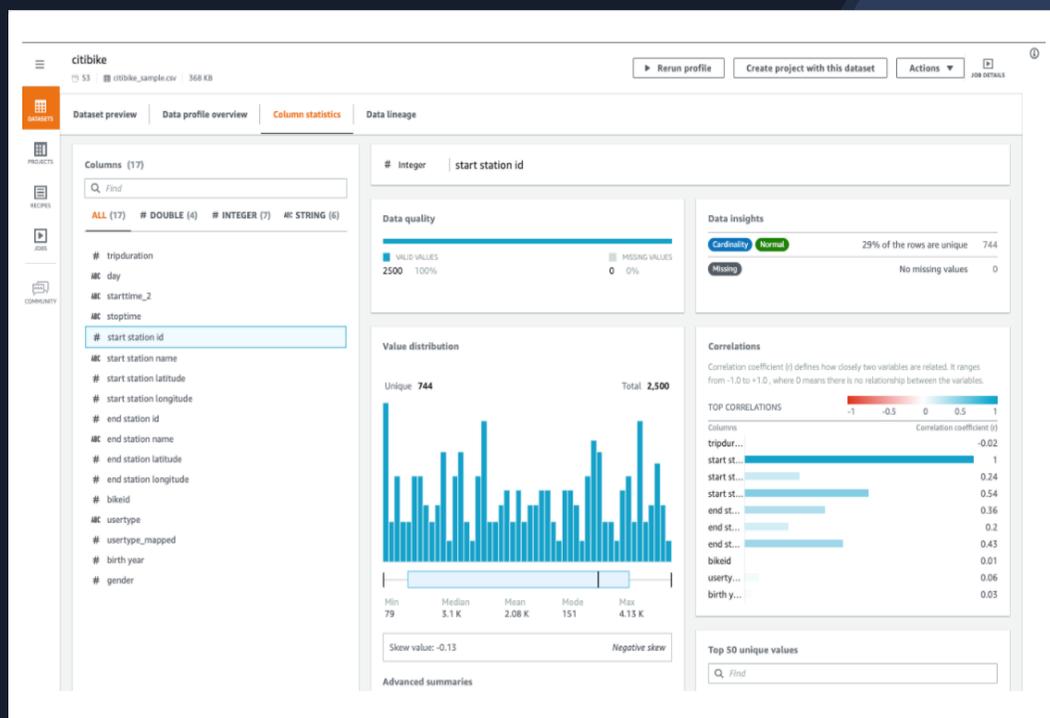
分散処理を意識せずに開発

# AWS Glue DataBrew

ノーコードデータプレパレーションサービス



**Business Analyst**  
**Data Scientist**



豊富なビジュアルインターフェイス

250 種類以上の組み込みトランスフォーメーション

データパターンと異常を把握するためのプロファイリングデータ

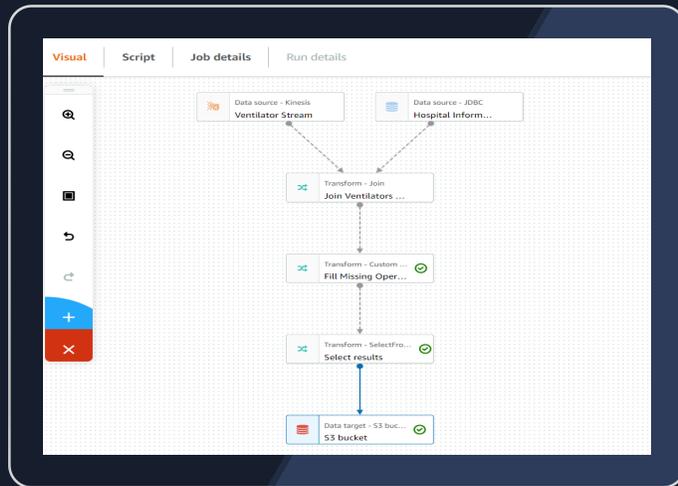
大規模なデータセットを大規模に処理

# 参考：全てのデータ利用者向けの選択肢



**ETL developer  
Data Engineer**

## AWS Glue Studio



ローコードで迅速に開発

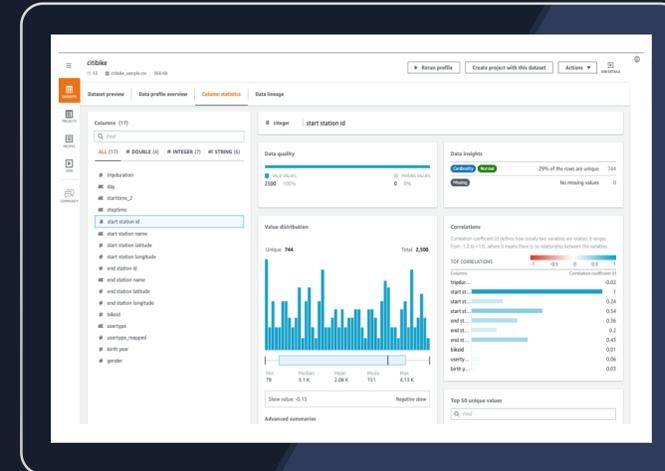
コーディングで柔軟にカスタマイズ

用途別の実行エンジン



**Business Analyst  
Data scientist**

## AWS Glue DataBrew



ノーコードでセルフサービス型データ準備

250以上のビルトイン変換機能

データプロファイル

# Amazon Q data integration in AWS Glue

AIエキスパートによりAWS GLUEジョブの学習、構築、および実行に必要な時間と労力を削減

## 自然言語によるAWS Glueジョブの作成

AWS管理コンソールから利用可能

例えば、

「Redshift から読み込み、  
null フィールドをドロップし、  
Parquet ファイルとして S3 に書き込む  
Glue ETL ジョブを作成して」  
と指示するとコードを作成する

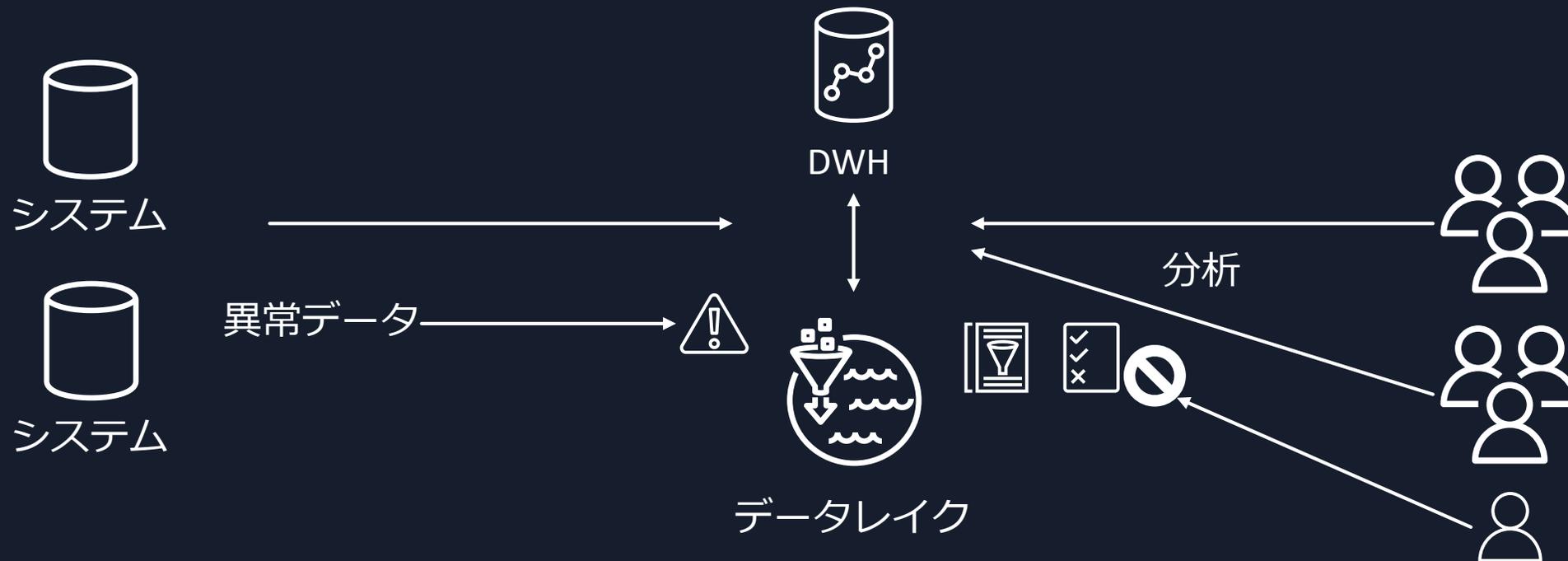
チャットインターフェースによる  
トラブルシューティングや  
Q&Aによる詳しいガイダンスの提供

The screenshot shows the AWS Glue console interface. On the left, the 'Script (Locked)' tab is active, displaying a Python script for a Glue ETL job. The script includes imports for sys, awsglue, pyspark, and GlueContext, and defines a function to find null fields in a schema. A red arrow points to the script with the text 'コピー & ペースト' (Copy & Paste). On the right, the Amazon Q chat interface is visible, showing a prompt: 'Write a Glue ETL job that reads from Redshift, drop null fields, and write to S3 as parquet files.' Below the prompt is a text input field with the placeholder 'Ask me anything about AWS' and a 'Max 1000 characters' limit.

# ガバナンス



# データが活用されるにはガバナンスが必要



データの品質チェック、データカタログ、権限制御により、  
データの沼地を避けて安全な状態を作り上げたい

# AWS Glueはデータガバナンス機能を提供している

## スケーラブルなデータ インテグレーションエンジン



組み込みデータ変換



多様なジョブ実行エンジン  
Apache Spark, Python shell, Ray, Streaming



ジョブのモニタリング

## コネクタとデータ収集



コネクタ



コネクタのマーケットプレイス



多様なインターフェース

## 生産性の高い データオペレーション



多様なペルソナへの対応



生産性向上



DataOps

## 一元化された 統合データガバナンス



データカタログ



データ品質



クローラー



AWS Lake Formation  
連携

# AWS Glue Data Quality

データレイクのデータの品質を自動的に測定・モニタリング・管理

ルールベースとMLベースのアプローチを  
組み合わせてデータ品質をモニタリング

- 組み込みのデータ品質ルールとアクション
- 自動ルール推奨機能：データの統計を自動的に計算して品質ルールを推奨
- MLベースの異常検出（プレビュー）

データパイプラインと保存データをサポート

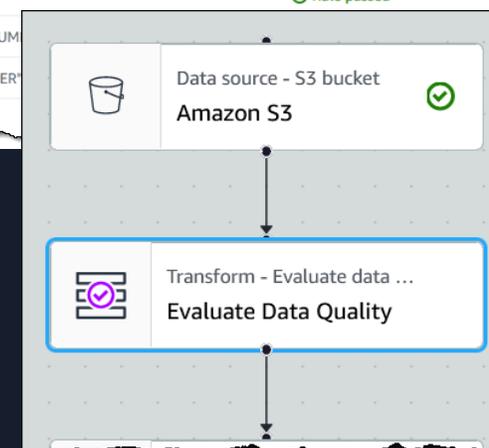
Rulesets | Data quality results

dqrun-89769d34ba974a7b9117d6cdf4a26b1108239b38

Name	DQ status	Status detail
RS1	DQ failed	Data quality score is 94.444
Rule_1 RowCount between 1021908 and 4087632	Rule passed	
Rule_2 IsComplete "TAG_NUMBER_MASKED"	Rule passed	

Observation	Related metrics	Rule recommendations	Observed data
RowCount of 74562.0 is higher than the detected upper bound of 65565.0.	Dataset: RowCount • ActualValue: 74562.0 • ExpectedValue: 60388.0 • LowerLimit: 55409.0 • UpperLimit: 65565.0	RowCount between 55408.0 and 65566.0	RowCount

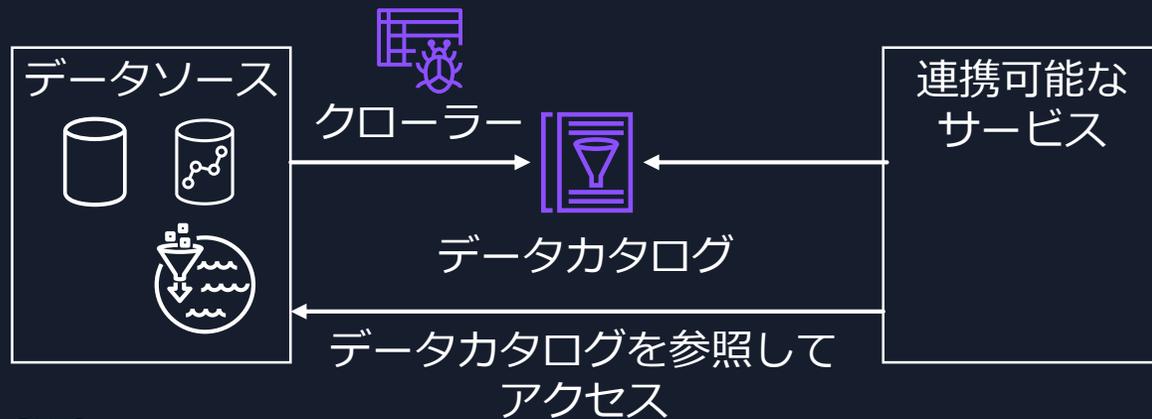
Rule_8 ColumnLength "TICKET_DATE" = 10	Rule passed
Rule_9 IsComplete "TICKET_NUMBER"	Rule passed
Rule_10 StandardDeviation "TICKET_NUM"	
Rule_11 ColumnValues "TICKET_NUMBER"	
Rule_12 IsComplete "OFFICER"	



# AWS Glue データカタログ

様々なデータソースのメタデータを一元管理

- クローラーでメタ情報を収集し、カタログを自動作成
- AWS GlueジョブやAmazon Athena、Amazon Redshift Spectrumなどからデータへのアクセスが便利に
- Apache Hive メタストア互換



### テーブル情報

Table details		Advanced properties	
Name	bank_additionalssampledata	Description	サンプルデータ
Database	sampledata	Location	s3://[REDACTED]
Connection	-	Output format	org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
Deprecated	-	Input format	org.apache.hadoop.mapred.TextInputFormat
Last updated	October 13, 2022 at 07:36:3	Output format	org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe
Serde serialization lib	org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe	Input format	org.apache.hadoop.mapred.TextInputFormat

### スキーマ情報

Schema (21)

View and manage the table schema.

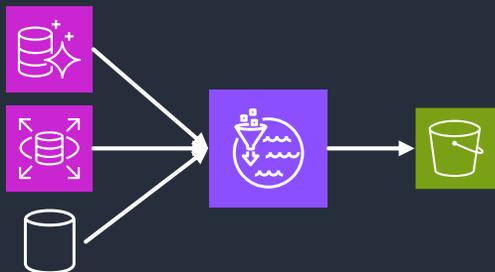
Filter schemas

#	Column name	Data type	Partition key	Comment
1	age	bigint	-	年齢
2	job	string	-	職業
3	marital	string	-	-
4	education	string	-	-

# 安全なデータレイクを簡単に構築できる

## AWS Lake Formation

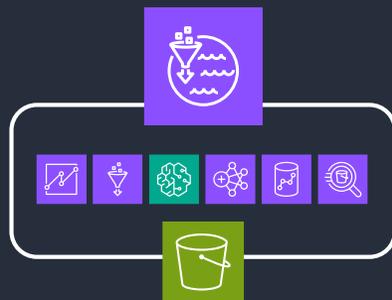
### ブループリント機能 によるデータの取り込み



AWS Glue の Connection を用いて  
様々なデータソースから  
データレイクにデータ取り込み

洗い替えだけでなく  
増分データ取り込みにも対応

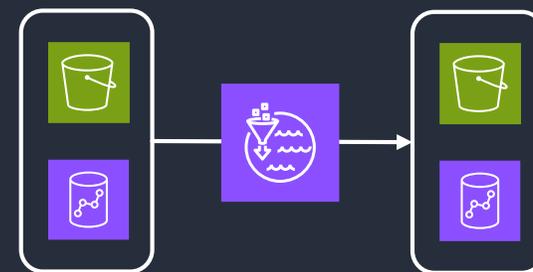
### 一元管理による きめ細かいアクセス制御



データレイク管理者が  
AWS サービスの Amazon S3 へのア  
クセスを一元的に制御

細かなアクセス制御に基づいて  
行レベル/列レベル/セルレベルのセ  
キュリティを実現

### 安全なデータ共有の制御

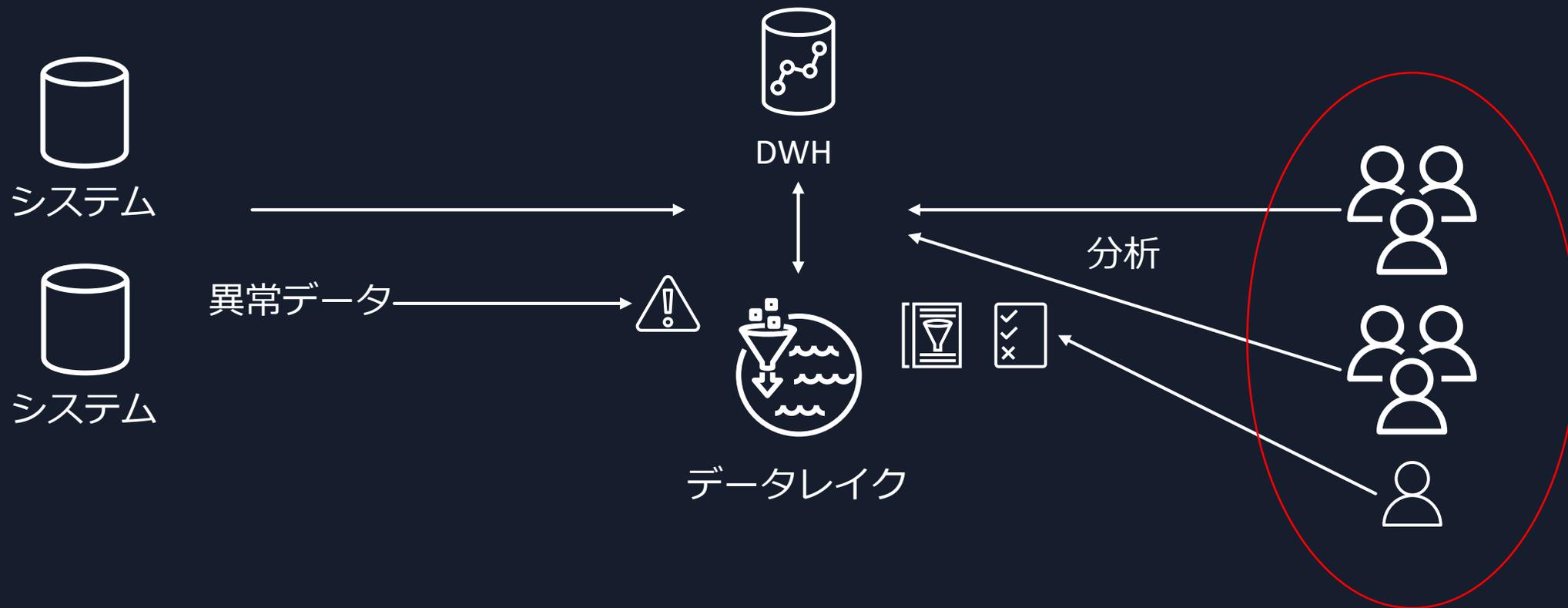


Amazon S3 データレイクや Amazon  
Redshift のデータを  
安全に  
別 AWS アカウントや  
別リージョンに共有し制御可能

# コラボレーション

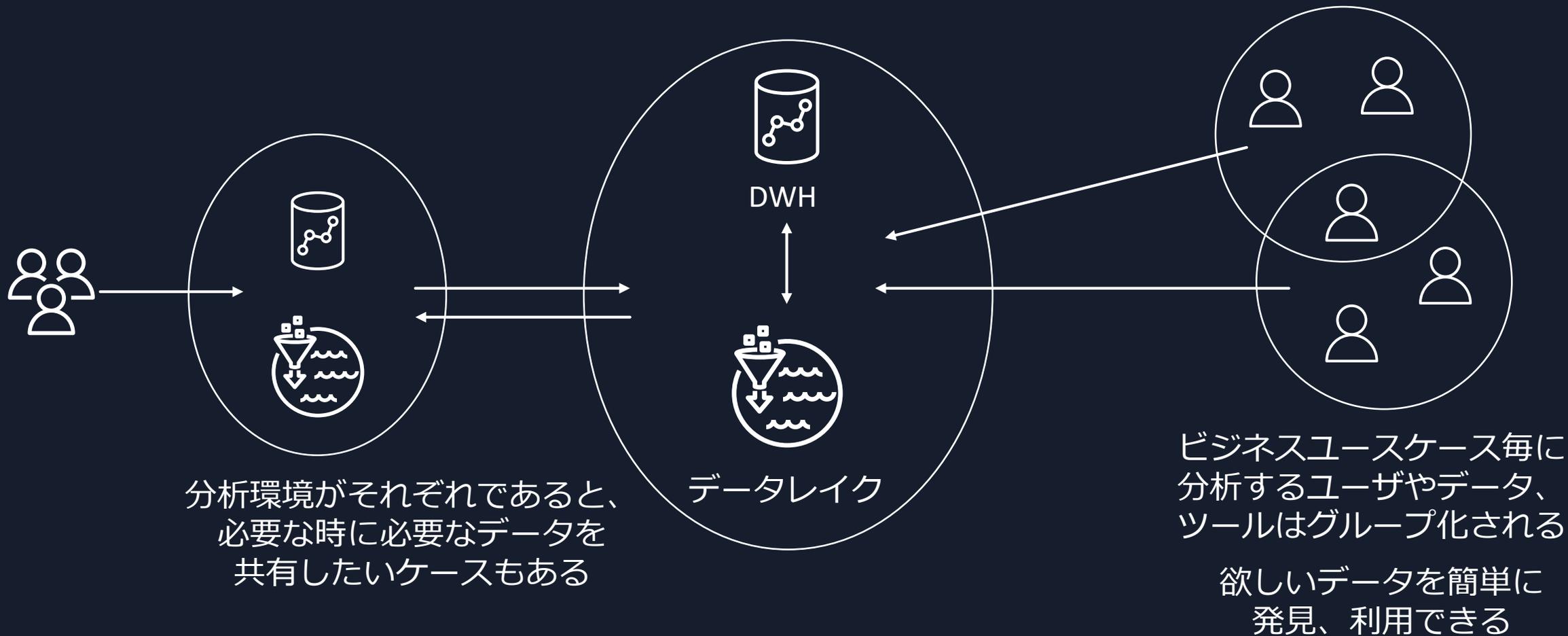


# データ活用におけるコラボレーションとは？



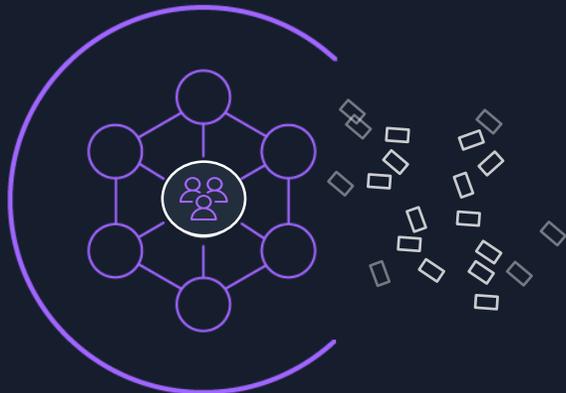
こういったコラボレーションが  
行われているか？

# データ活用におけるコラボレーションとは？



データガバナンスの一元化と分散化の適切なバランスを実現したい

# データの活用を促進する Amazon DataZone



## データプロデューサー

様々なソースや、アカウント、リージョンからデータを取り込む



## Amazon DataZone

データポータル

ビジネスデータ  
カタログ



プロジェクトと  
環境

ガバナンスとアクセスコントロール



## データコンシューマー

分析ツールを使用して、アクセス、コラボレーション、利用を簡単にする

組織の境界を越えたデータアクセスの管理  
データとツールを共有して関係者を結びつけ、ビジネス・インサイトを促進  
機械学習でデータの発見とカタログ作成を自動化

# Amazon DataZone データポータル

MARKETINGDATA... | Admin | 891377069679

Marketing

マーケット

マーケットセールステーブル  
技術名: mkt\_sls\_table | アセットタイプ: Glue Table

以下の結果をすべて表示: "マーケット"

### カタログを作成する

非表示

- ビジネス用語集を作成する**  
データアセットとリンクできるビジネス用語とその定義の標準語彙を作成します。単純なリストでも、各用語が複数の用語とつながる階層でもかまいません。  
[ビジネス用語集を作成](#)
- メタデータフォームを作成**  
メタデータを収集したり、公開されたデータアセットを管理したりするために添付できるメタデータフォームを作成します。文字列、日付、ブール値、整数、ビジネス用語集などのタイプのフィールドを使用してキーと値のペアを作成できます。  
[メタデータフォームを作成](#)
- カタログヘデータを公開する**  
データソースに接続し、メタデータを収集して生成する自動ジョブを実行し、簡単に見つけられるように豊富なコンテキストを追加することで、データアセットをカタログに公開します。データを公開するプロジェクトを選択してください。  
プロジェクト: MarketingDataAnaly... | [データ概要を表示](#)

### 最近閲覧したプロジェクト

- MarketingDataAnalysisProject  
環境: datalake02
- SalesDataPublishingProject  
環境: test\_env01, detalake01

### タスク

- サブスクリプションリンクエラストが作成されました | 2024年3月23日 午前12:37:11  
arn:aws:iam:891377069679:role/AdminはプロジェクトAWS Logoのためにアセット MarketingDataAnalysisProjectの利用を要請した
- サブスクリプションリンクエラストが作成されました | 2024年3月22日 午後08:14:46  
arn:aws:iam:891377069679:role/AdminはプロジェクトAWS Logoのためにアセット MarketingDataAnalysisProjectの利用を要請した

# ビジネスデータカタログ

メタデータにビジネス価値を付加し、ディスカバリーを強化する

SALESDATAPUBLI... アセットを検索

Admin 891377069679

戻る

## マーケットセールステーブル

説明なし

検索名: mkt\_sls\_table - アセットタイプ: Glue Table

インベントリアセットを表示

サブスクライブ

ビジネスメタデータ スキーマ サブスクリプションリクエスト

概要  
セールス部門が管理する販売データ。マーケティング施策の参考情報。

用語集の用語 [これは何ですか?](#)

日本  
機密

メタデータフォーム [これは何ですか?](#)

### AWS Glue テーブル

Glue データカタログ ID	891377069679
場所	s3://amazon-datazone-891377069679-us-east-1-461606003/dzd_56shjrdz7g7607/datazone/570jhlmrf2utnr/sys/athena/tables/910e6a89-4d4b-451d-9781-ab5255612be0
リージョン	us-east-1
テーブル ARN	arn:aws:glue:us-east-1:891377069679:table/detalake01_pub/mkt_sls_table
テーブル名	mkt_sls_table

アセットの詳細

公開されたリビジョン  
18

所有プロジェクト  
SalesDataPublishingProject  
サブスクリプションの承認  
必須

発行者  
sales01  
公開日  
2024年2月02日 午後07:23:37

SALESDATAPUBLI... アセットを検索

Admin 891377069679

戻る

## マーケットセールステーブル

説明なし

検索名: mkt\_sls\_table - アセットタイプ: Glue Table

インベントリアセットを表示

サブスクライブ

ビジネスメタデータ スキーマ サブスクリプションリクエスト

このアセットのスキーマを表示または変更するには、このセクションを使用します。

フィルター

名前	データタイプ	説明	用語	データ品質の結果
注文番号 ord_num	int	Order number for the sale		
販売量 sales_qty_sld	int	Quantity sold for the order		
卸売原価 wholesale_cost	double	Wholesale cost of the item		
定価 lst_pr	double	List price of the item		
販売価格 sell_pr	double	Selling price of the item for this order		
ディスカウント disnt	double	Discount percentage given on the item		
配送モード ship_mode	int	Shipping mode used to deliver the order		
倉庫ID warehouse_id	int	ID of the warehouse from where the item was shipped		
アイテムID item_id	int	ID of the item ordered		

# 生成AIによるメタデータ生成の自動化

テーブル名とカラム名に加えて、概要とカラムの説明を自動生成

Market Sales Table

説明なし

生成されたメタデータを確認

概要

It tracks information like the catalog page the item is listed on, order details such as order number, shipping mode and customer IDs, pricing, warehouse details for fulfillment, and sales quantity sold.

The catalog page number indicates where customers can find the item in the catalog. Order details capture who the order was shipped to, how it was shipped, and a reference order number. Pricing columns store the list price, selling price, discounts applied, and internal wholesale cost. Customer IDs link to the billed and shipped customer for the order. The sales quantity provides the volume sold for each line item. Warehouse ID indicates which warehouse fulfilled the order.

Together, these columns provide a comprehensive view of catalog sales transactions, enabling analysis of product pricing, customer order behaviors, sales volumes and fulfillment. The table supports insights into profitability, customer order patterns, item popularity, and logistics monitoring. **Use Case** Based on the detailed schema and description provided, the mkt\_sls\_table contains rich transactional data on catalog sales that could be valuable for several business use cases and end users:

- Marketing and sales teams could utilize this table to analyze product pricing, identify best selling items, monitor customer order patterns and item popularity to optimize catalog layout, pricing and promotions. The sales, pricing and customer data provides insights to boost revenues.
- Product managers could study sales volumes and trends to inform decisions on inventory planning, production levels, supplier orders and new product development. Granular order data aids demand forecasting and inventory optimization.
- Finance teams may leverage the cost and pricing data to perform profitability analysis by product, region or channel. Data on discounts and volumes helps modeling for pricing optimization.
- Customer service could lookup order details like shipping mode, warehouse and billed customer ID to address any order inquiries. Customer IDs link to demographics for segmentation.
- Logistics professionals can analyze warehouse fulfillment volumes and shipping modes to identify process bottlenecks and optimize distribution networks.
- Executive leadership can track overall sales KPIs and drill down into product-level trends to monitor growth, market share and strategic initiatives.

In summary, the transactional data in mkt\_sls\_table enables critical sales, marketing, product, finance, customer service and supply chain operations. It would provide business value across consumer goods, retail, manufacturing and distribution sectors. The table supports data-driven insights for pricing, inventory, customer patterns, logistics and overall growth.

少なく表示

ビジネスメタデータ

このアセットのスキーマを表示または変更するには、このセクションを使用します。

メタデータの自動生成

緑色のアイコンは、データアセットのメタデータ候補が自動的に生成されたことを示します。このアイコンをクリックすると、各提案を編集、承認、または拒否できます。また、アセットに関連して自動生成されたすべての提案に対して、「すべて承認」または「すべて拒否」を選択することもできます。詳細はこちら

名前	データタイプ	説明	用語
Order Number ord_num	int	Order number for the sale EMPTY	
Sales Quantity Sold sales_qty_sld	int	Quantity sold for the order EMPTY	
Wholesale Cost wholesale_cost	double	Wholesale cost of the item EMPTY	
List Price lst_pr	double	List price of the item EMPTY	
Selling Price sell_pr	double	Selling price of the item for this order EMPTY	
Discount disnt	double	Discount percentage given on the item EMPTY	
Shipping Mode ship_mode	int	Shipping mode used to deliver the order EMPTY	

素早く簡単にビジネスメタデータが充実化され、データ活用をより進めることができる

# AWS Glue Data Qualityの統合

## AWS Glue Data Quality (AWS管理コンソール)

The screenshot shows the AWS Glue Data Quality console. On the left is a navigation menu with options like 'Getting started', 'ETL jobs', 'Data Catalog tables', and 'Data Catalog'. The main area displays 'Data quality info' for a specific table. It includes a search bar and a table of rules. The table has columns for Name, Status, and Status detail.

Name	Status	Status detail
auto-gen-ruleset- 27 rule(s)	Completed	Data quality score is 96
Rule_1- IsComplete "id"	Rule passed	-
Rule_2- Uniqueness "id" > 0.95	Rule passed	-
Rule_3- ColumnLength "id" = 36	Rule passed	-
Rule_4- IsComplete "drivers"	Rule passed	-
Rule_5- IsComplete "passport"	Rule passed	-
Rule_6- IsComplete "prefix"	Rule passed	-
Rule_7- ColumnValues "ethnicity" in ["nonhispan...	Rule passed	-
Rule_8- Completeness "zip" >= 0.5	Rule passed	-
Rule_9- IsComplete "healthcare_expenses"	Rule passed	-



## Amazon DataZoneデータポータル

The screenshot shows the Amazon DataZone Data Quality portal. The top section is 'Data quality scores overview' for a table named 'Patients'. It displays an overall score of 96 and a breakdown of metrics: ColumnValues.Compliance (66.67 Failed), Completeness (100 Passed), MaximumLength (100 Passed), MinimumLength (100 Passed), and Uniqueness (100 Passed). Below this is a 'Data quality run results' table showing details for the 'auto-gen-ruleset'.

Rule set	Attribute	Applicable column	Result	Result details
auto-gen-ruleset (27 rules)			Failed	26/27 have passed. Data quality score is 96% (26/27)
Rule 1 IsComplete "id"	Completeness	id	Passed	-
Rule 2 Uniqueness "id" > 0.95	Uniqueness	id	Passed	-
Rule 3 ColumnLength "id" = 36	MaximumLength MinimumLength	id	Passed	-

The bottom section shows a 'Report snapshot' with a line chart of 'Data quality score (%)' over time. A tooltip for a snapshot on 22/09/2024 shows a score of 62.79 for 'Ruleset1'.

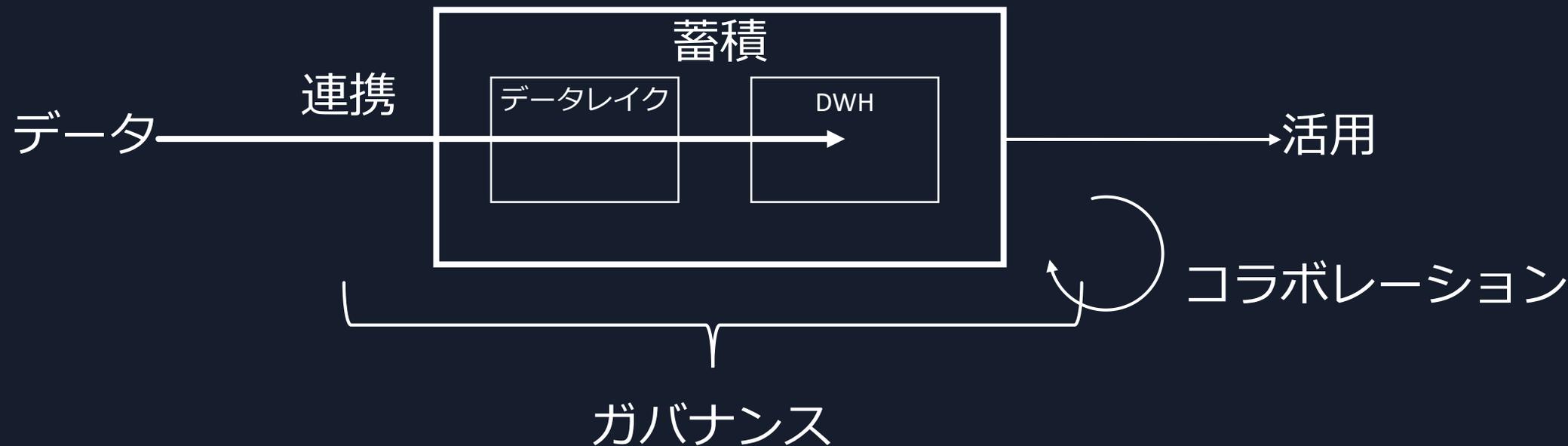
AWS Glue Data Qualityのデータ品質スコアや履歴がAmazon DataZoneのデータポータルより確認できる

# まとめ

# 本セッションでお話ししたこと

生成AIの活用に重要となるアナリティクス環境について、必要な要件と、要件を実現するAWSマネージドアナリティクスサービスについて、以下の順に紹介しました。

蓄積、連携、ガバナンス、コラボレーション



# 本セッションでお話ししたこと

お話しした求められること	ご紹介したAWSサービス
多様、大量なデータを保存したい	Amazon S3
大量データを高速に分析したい	Amazon Redshift
手間をかけず鮮度の高いデータを分析したい	Amazon RedshiftのZero-ETL統合
ETLが必要なケースも存在する	AWS Glue
データが活用されるにはガバナンスが必要	AWS Glue、AWS Lake Formation
コラボレーションしてデータ活用を促進したい	Amazon DataZone

マネージドサービス、サーバーレス技術に加えて  
生成AIをはじめとするMLによる機能強化により更なる利便性と省力化を実現

# AWSはお客様のデータニーズに応える比類のない機能を提供します



## 包括的

最も包括的なサービスセットであらゆるユースケースに最適な価格パフォーマンスを提供



## 統合型

Zero-ETLを含むデータ統合の選択肢が豊富なため、すべてのデータに簡単に接続可能



## ガバナンス

エンドツーエンドのデータガバナンス機能により、データをより迅速に活用可能

# Thank you!

