

2014.6.22

第14回 関西DB勉強会

Snowflake

Data Cloud Summit 24 参加報告

- Snowflake × AI で考える最新マーケティング基盤 -

2024/6/22




自己紹介 & 会社紹介

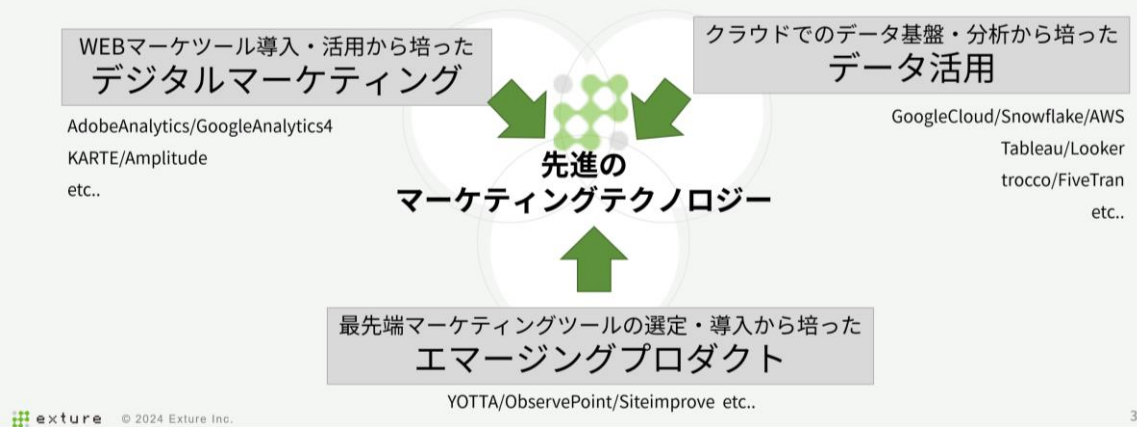
喜田 紘介 (@kkkida_twtr)

趣味： 🧑‍🦱 🧑‍🦱 🐕 🔍 🎤 🐘



 **exture** エクスチャ株式会社

デジタルマーケティング × データ活用 × 最先端プロダクト
先進のマーケティングテクノロジーで
貴社のマーケティングを次の次元へと導きます



- 🎯 マーケのためのデータ収集～活用コンサル
- 🎯 ビジネスユーザが「使える」データ×AI整備
- 🎯 目的特化ツールで即効性のある施策実行

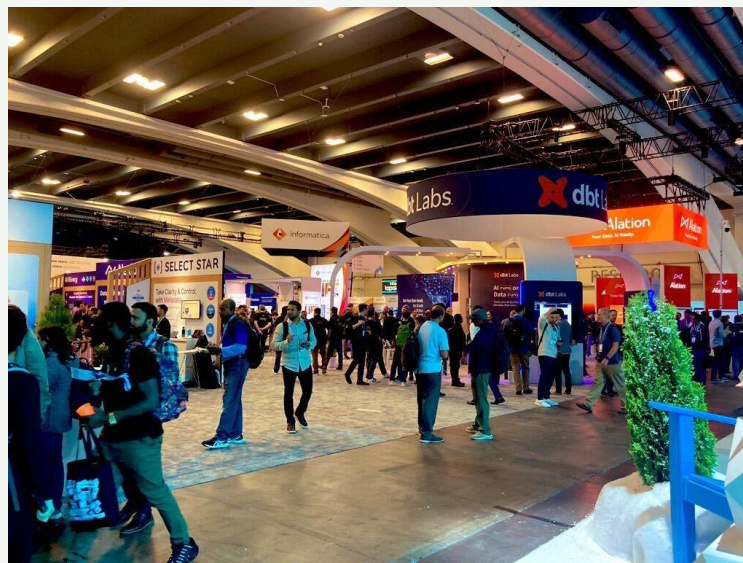
Snowflake主催のData Techイベント、開催規模は年々拡大しグローバルに存在感を示す

1.5万人の参加者（日本人参加者 250人）

- 昨年1.2万人から**25%up**
- 最終日のDev Dayは**5万人**の登録

400+セッション、180社のスポンサー

サンフランシスコの有名ホール Moscone Center



海外カンファレンス体験

🎥 深夜のリアルタイム現地配信

🎥 おしゃれ民泊から思い出配信

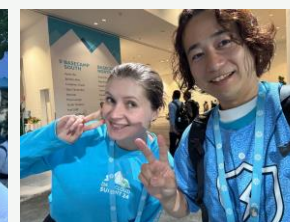
[エクスチュア株式会社 youtube](#)



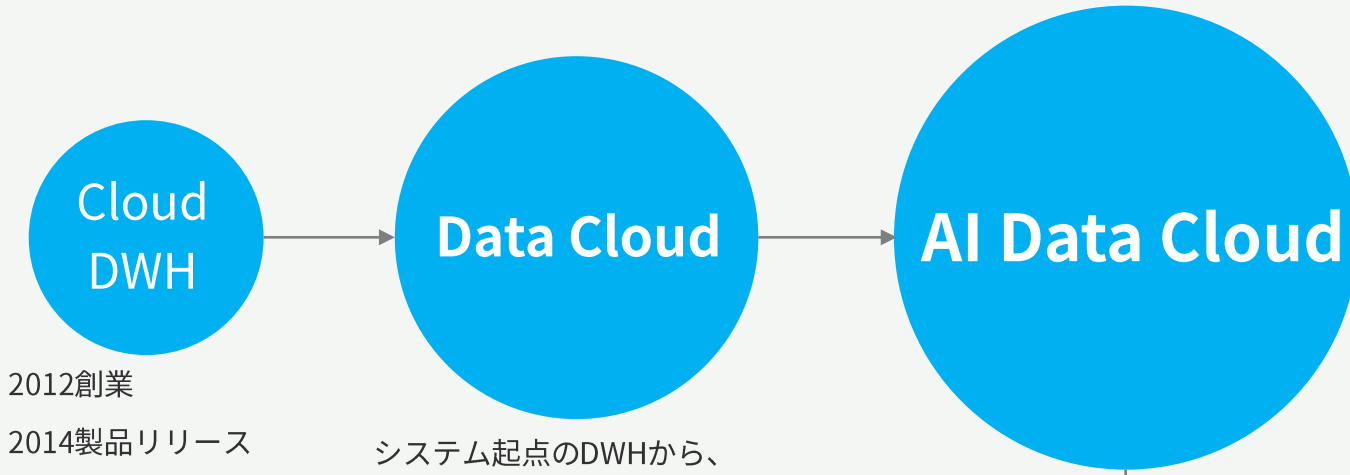
🏢 聖地巡礼



👋 つなぐりに感謝



キーノートより - Snowflakeが唱えるビジョン -



2012創業

2014製品リリース
最新の技術で実現した
DWHのあるべき姿

システム起点のDWHから、
人やビジネスを中心としたデータ
利活用にフォーカスしたSaaSへ

2023年 キーノートより



ENTERPRISE AI

簡単、効率的、信頼できることがビジネスの成功のためのAIに必須である
Snowflakeはそれらを提供する「AI Data Cloud」へとリブランディング

NATIVE APPS

データの移動なく、統制の取れた環境でAI/ロジックを動かしてビジネスに寄与
Native Appsのエコシステムは拡大し、業界特化やインハウス型も実現

DATA FOUNDATION

AI/Appの土台となるデータ基盤は既に「completed」と強調
あらゆるデータを最適なコンピュートエンジンで扱い、高品質なデータを揃え
セキュリティとアクセシビリティを両立

キートンより - AI Data Cloud のいま -

AI EXPERIENCES

Copilot | Document AI | Universal Search

DEVELOPER OPERATIONS

DEVELOPER EXPERIENCES

Notebook | Studio | Copilot | Worksheet | VSCode | APIs | Git Integration

OPERATIONALIZATION

Snowflake CLI | Database Change Management | Snowflake Trail | Serverless Tasks

DISCOVERY & DISTRIBUTION

MARKETPLACE - Data / Apps / AI

INGESTION

STREAMING

Snowpipe Streaming
Snowpipe

CONNECTORS

SaaS Connectors
Database Connectors

TRANSFORMATION

Task

Serverless | Cost Optimized
Snowpark DataFrames
Dynamic Tables

ANALYTICS

Geospatial
Search Optimization
Time Series
Query Acceleration Services

MACHINE LEARNING

SNOWFLAKE ML

Snowpark ML API
Feature Store
Model Registry
ML Functions

GENERATIVE AI

SNOWFLAKE CORTEX AI

Studio
Cortex Analyst
Cortex Search
Foundation Models

SNOWFLAKE ARCTIC

DATA PRODUCTS

Streamlit
Secure Data Sharing
Snowflake Native App
Snowflake Data CleanRoom

SNOWFLAKE HORIZON

Compliance | Security | Privacy | Interoperability | Access

DATA & COMPUTE

DATA TYPE

Structured | Semi-Structured | Unstructured

TABLE

Snowflake Table | Iceberg Table | Hybrid Table

RUNTIME

SQL | Python | Java | Scala | Snowpark Container Services

PROCESSOR

CPU | GPU

Snowflake Cortex - エンタープライズAIへのコミット -

すぐ使えるマネージドAI/MLサービス + あらゆる開発者体験をAI用途に統合

Snowflake Cortex AI

Snowflakeによる目的特化型AIサービス群

Cortex Functions 環境準備ゼロ、最小工数でAIアプリ開発
SQL内で実行できるAI/ML関数 → 例: [complete関数](#)

Cortex Analyst ビジネスユーザがデータから直接示唆を得る時代へ
構造化データに対するクエリ生成 + 結果の説明

Cortex Search 社内に散在する文書データを統合してRAGに利用
非構造化データに対するセマンティック検索

用途に応じて選択、ファインチューニング可能なLLMモデル
Snowflake Arctic / Mistral / llama / genna

AI/ML開発、ライフサイクル支援

高度なアプリ開発用のフレームワーク群

Cortex Playground ノーコードでGUIのAIアプリをリリース
数クリックで構築完了するCortex AI特化のインターフェース

Notebook / Git / VSCodeExtension モダンな開発ツールへの標準対応
弱点とされてきた開発者ツールを刷新

Snowpark ML 高度なMLワークフローのための開発機能セット
MLモデルの学習～管理、pandas構文へのネイティブ対応

DWH、Data Lakehouseとして進化を続けてきたパワフルな開発
インターフェース (SQL、Python、REST API)

Snowflakeそのもののユーザ利便性を向上するAIインターフェース群

Copilot

Universal Search

AI Object description

Auto Classification

Document AI

demo - AI with MyData -

SQL as cell13

```
1 -- complete関数の実行
2 -- レビューコメントの感情を評価
3
4 SELECT SNOWFLAKE.CORTEX.COMPLETE(
5   'mistral-Large',
6   CONCAT('determine the overall sentiment this review in bullet points: <review>', content, '</review>')
7 ) FROM reviews LIMIT 3;
```

NO	CONTENT
0	1 this was really good
1	2 this was really sad
2	3 It was a terribly boring time

SQL as cell14

```
1 -- 同じことをもっと端的に回答させるプロンプトをインプット
2
3 SELECT SNOWFLAKE.CORTEX.COMPLETE(
4   'Llama2-70b-chat',
5   [
6     { 'role': 'system', 'content': 'You are a helpful AI assistant. Analyze the movie review text and determine the overall sentiment. Answer with just \\'Positive\\', \\'Negative\\', or \\'Neutral\\' },
7     { 'role': 'user', 'content': CONCAT('reviews are in bullet points: <review>', content, '</review>') }
8   ], 0
9 ) as response
10 FROM reviews;
11
```

RESPONSE
0 { "choices": [{ "messages": "Positive" }], "created": 1718960417, "r
1 { "choices": [{ "messages": "Negative" }], "created": 1718960417, .
2 { "choices": [{ "messages": "Negative" }], "created": 1718960417, .

強力なNotebook機能がリリース

- SQLとPythonの実行環境を統合したNotebook
 - Pythonのライブラリをポチポチ選べて準備不要
 - セルごとにSQL、Python、Markdownを切り替えてどんどん書ける
- 前のセルの結果は以降のセルで変数として呼び出せる
 - SQLで取得した結果をdataframe化してPythonで加工、描画
- ipynb形式 / Git統合
 - Notebook用フォーマットでコードを保管、Git管理
 - Snowflake公式からも様々なハンズオンスクリプトがこの形で公開されておりあらゆるSnowflake機能の習得に

LLM関数のデモ

- 映画のレビューコメントのデモ（公式マニュアルにサンプルあり）
- SQLセルの中で準備ゼロで書ける汎用/マネージドLLM関数
 - SQLで取得したテーブルの各行に対して普通のSQL関数のように動作し、当然その場で与えた指示通りに結果を生成してくれる
 - ✗データを外部LLMに送信する
 - ◎データのある場所で動作するAI
- プロンプトエンジニアリングにも対応

←レビューコメントからお気持ちを説明する例

最後のセルでは同じコメントを文章で説明ではなく、Positive/Negative/Neutralの3択で答えてくださいを指示

スコア化して点数ごとに集計、悪いコメントを探すようなSQLとの組み合わせ技が可能（感情分析については特化したモデルも提供、サミットではそのデモもあったよ）

Snowflake Horizon - データへのアクセス性の追求 -

活用Readyなデータを整える = Data Productsと、そのデータが利用者に届くまでを重視

Snowflake Horizon

データへのアクセス性とセキュリティ、ガバナンスを統合した新たな機能ブランディング

データへのアクセス性

Universal SearchとInternal Marketplaceで、自身がアクセス可能なあらゆるデータを一発検索
AIが自動的にデータの説明を生成するため運用負担もなし

プライバシー、ガバナンス

既存のポリシー（列選択、マスキング、GROUP BY強制など）の統合と下流テーブルやIcebergへの伝播、ネイティブデータクリーンルーム

コンプライアンス&セキュリティ

継続的なデータ品質チェック、ネイティブのリネージ機能によるデータ源流の把握、環境全体に対するセキュリティ診断など

マーケットプレイスの活用

3rd Party データ/アプリ流通基盤

Data Productsの隆盛

データを「利用可能な状態」に整えるデータカタログ製品、DataOps製品が隆盛、データの主管部門の責務でデータを整え広く展開するData Productsの浸透

高度分析用エンジン

ツール独自のデータストアやロジックを組み込んだNative Appが登場
グラフDBなど専用ロジックをアプリケーションのエンジンに選択可能に

データメッシュ組織の実現

Data Products / Native Appsを自社で有効活用

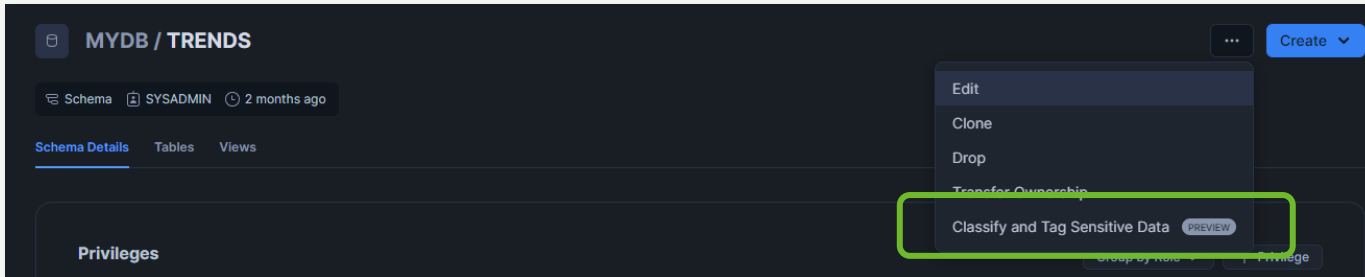
Internalマーケットプレイス

各部門で用意した分析Readyなデータセットを簡単に検索、業務利用可能に
業務ロジックを組み込んだNative Appを社内展開することも可能

Icebergテーブル / Polarisカタログ

Snowflake外部との相互運用性を実現するオープンフォーマットの
テーブルタイプ&カタログ

demo - Horizon / プライバシーデータを最上流で保護 -



Classify and Sensitive Data

スキーマ単位ですべてのオブジェクトを検査し、値のパターンや意味、センシティブかどうかを判定、自動タグ付け

ポリシーとの連動

センシティブデータのタグを持っていればマスキングするポリシーと自動的に対応付ける

タグの伝播、可視化 (2024/6時点では未リリース)

上流のテーブルで定義したタグは下流のビューにも伝播し、リネージ機能で可視化することでセンシティブな値がどこで使われているか漏れなくチェック可能

The screenshot shows the 'Classification Results' page. A table lists columns and their recommended semantic categories. The 'EMAIL' row is highlighted in green. A callout box explains the classification.

COLUMN ↑	RECOMMENDED SEMANTIC CATEGORY	SAMPLE VALUES
1 AGE_RANGE	—	40代, 20代, 10代, 30代, 60代以上
2 TREND_KEYWORD	—	3, 5, 6, 7
3 USER_EMAIL	EMAIL IDENTIFIER	user_00001@example.com +9

Other semantic tag values

GENDER	QUASI_IDENTIFIER
PHONE_NUMBER	IDENTIFIER
IP_ADDRESS	IDENTIFIER
URL	IDENTIFIER
PAYMENT_CARD	IDENTIFIER
AGE	QUASI_IDENTIFIER
LAT_LONG	QUASI_IDENTIFIER

↑メールアドレスであることが正しく判定されている
←他にも様々な値のsemanticパターン (=意味付け) が組み込みで用意されている他、カスタムでの定義も可能

Exture視点で見るこれからのマーケティング基盤のポイント

これまで確定的な手法がなかった「離反」の早期発見・対処しLTVを向上

既存のマーケ支援ツールでできることはやり尽くしている状態から、AIならではのアイデア×手数の実行にシフト



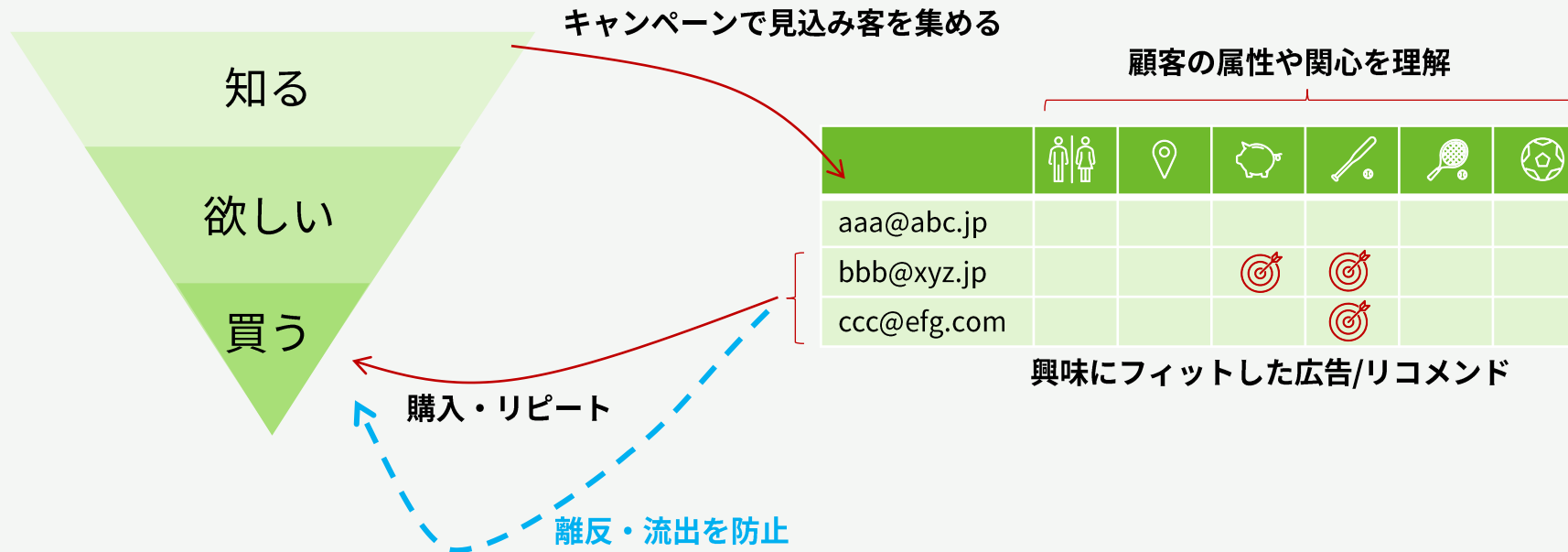
関心の高い人に広告を出してリピートしてもらう

→実施済み、広告効果をいかに高めるか

離反の可能性を早期発見、対策

→確定的な手法がなく、ツール機能で提供されない

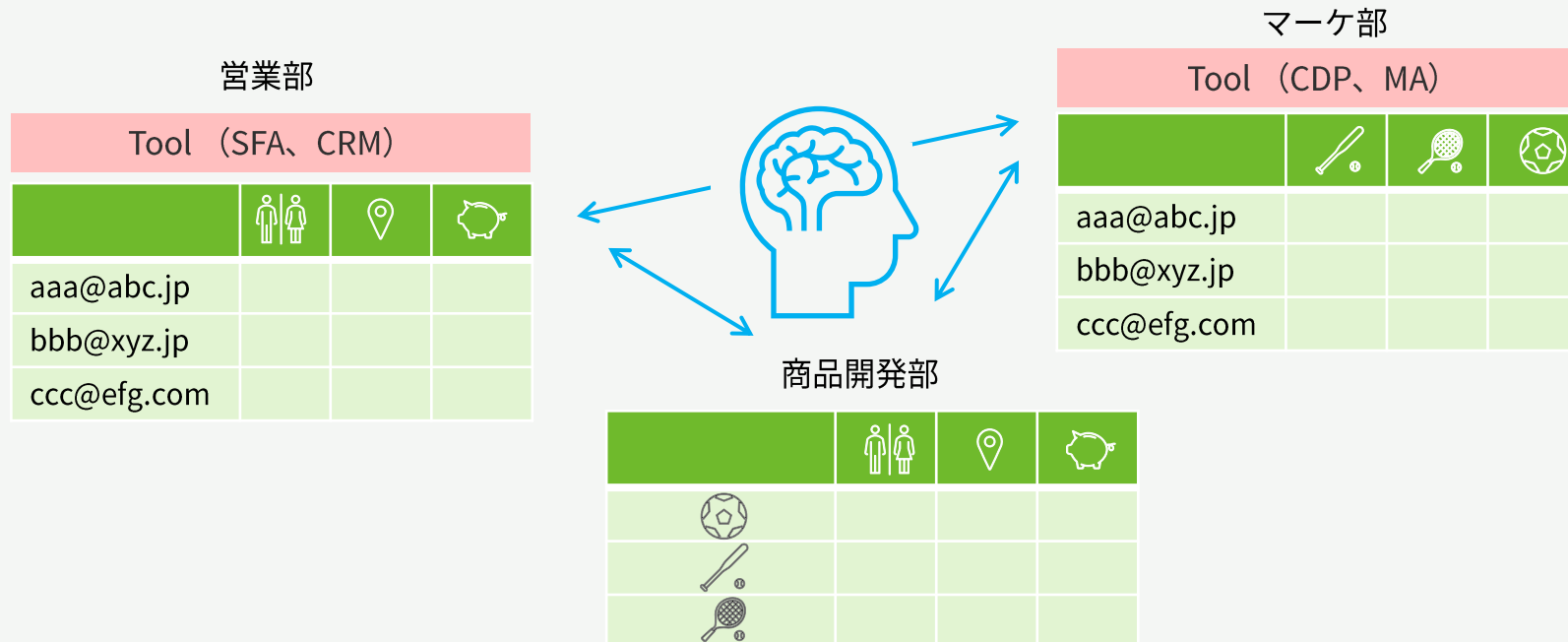
AI施策実行のスピード&手数はカギ ロコミ分析、ネガティブな発信者に対して早期に手を打つなど



「サイロ化解消」を実用段階に進める Data Products × Composable IT × AIミックス基盤

ツールに閉じた活用から、データを切り離し一元管理し、AIで活用できる状態へ

- Data Products** . . . データ所有部門が分析Readyなデータを整えて利用部門に公開
- Data Discovery** . . . ビジネスサイドから、社内外のData Productsを探して使える
- AI Activation** . . . 一元的にアクセス可能になったデータはAI・MLでの活用が可能
- Composable CDP** . . . 従来ツールが担った機能はコンポーザブル（組み合わせ式）で実現

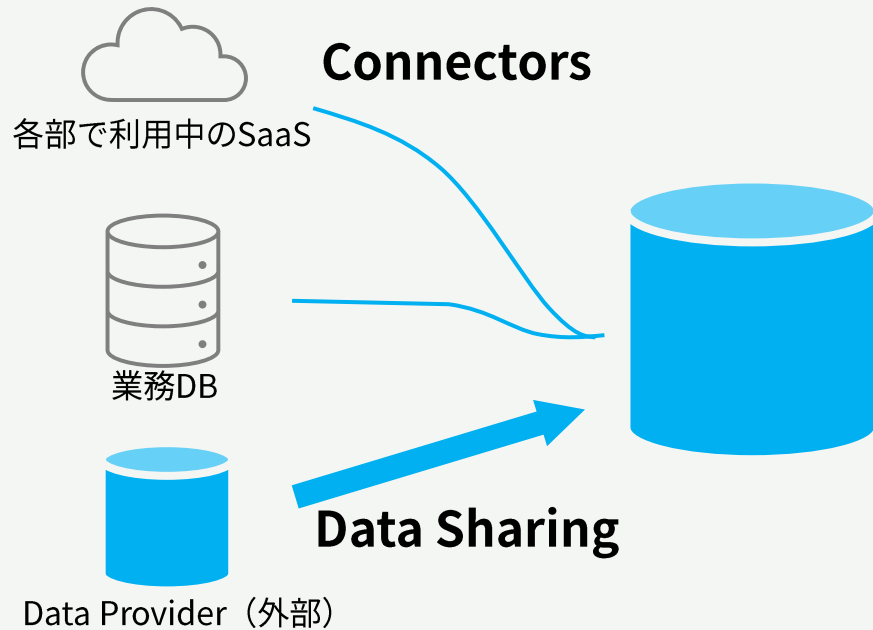


Exture視点で見るとこれからのマーケティング基盤のポイント

ZeroETLとライブデータで「モーメント」を重視した新たな顧客体験づくり

SaaSコネクタ / Databaseコネクタの充実でZeroETLへ、ライブデータを使ったアイデア創出

ライブデータで作られた高品質なData Products



データ加工

リアルタイム加工

Dynamic Table

定期実行ジョブ

Serverless

Cost Optimized

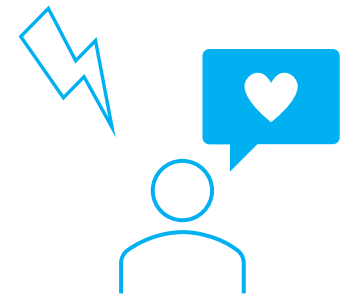
Defined by Notebook

Horizon

ガバナンス、クオリティ



瞬間のユーザ感情・関心に
アプローチする新たな顧客体験

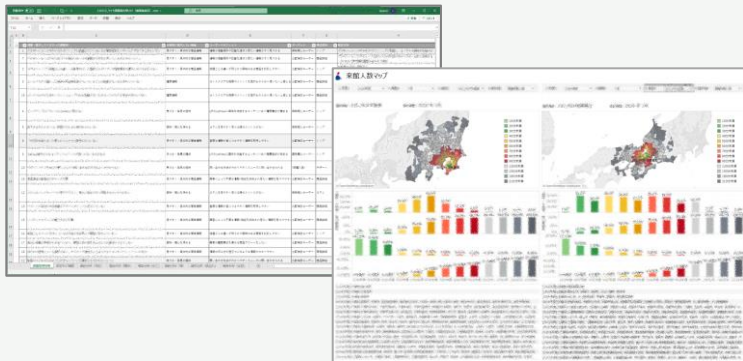


データ×AI活用のフェーズにあわせて、立案・プレ導入から構築・定着まで全面的にサポートします

マーケティングツールの導入やデータ基盤の構築に際して、“どう事業で活用できるか”の視点で立案から定着までを継続してサポートいたします

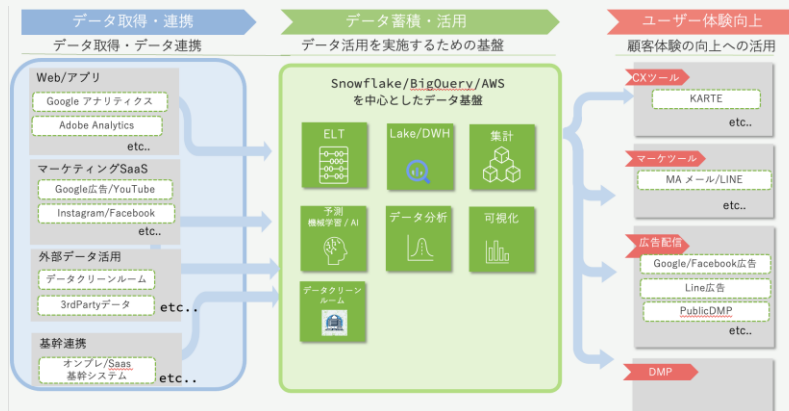
活用立案、PoC

現在使っているレポート、課題よりデータ活用方針を立案
ヒアリングシートや実データを使ったBIモック等より事業部との活用擦り合わせを実施



本格導入 ツール・基盤構築

貴社環境にマーケティングツールやデータ基盤を本格導入実施
各種連携やデータマート作成、ダッシュボードの作成などを実施



マーケティング/CX 高度活用/定着

CXデータ活用やBIを事業で利用し成果を出す為のサポート
データドリブンの事業への定着やモデルの作成を実施



Data to the People

すべての人にデータを