



最短・最速でアドホックな 分析環境を手に入れる方法

松本健治 杉山裕亮

クラウドエース株式会社

スピーカー自己紹介



松本 健治

クラウドエース株式会社執行役員
事業推進本部・本部長

呉服屋から一転して国産ソフトウェアメーカーの営業・営業企画に従事し、クラウドエースへ参画。営業責任者から事業本部長を経て 2021 年より執行役員に就任。Google プロダクトを利用したデータドリブン マーケティングを推進するため 2020 年 2 月から D.Table 株式会社の取締役副社長に就任。

スピーカー自己紹介



杉山 裕亮

クラウドエース株式会社
マーケティング部・部長

学生起業して以降、フリーランスのマーケターとして 10 業種以上の事業に携わる。2018 年にクラウドエースへ入社し、マーケティングを統括。スポーツや地域社会のデジタル活用を推進するため、新しいテクノロジーの活用や生活様式にチャレンジしている。

Vision

すべての人が先端技術を活用し
思い描いた夢を実現できる世界

Mission

正直を仕事にする



本日のキーワード

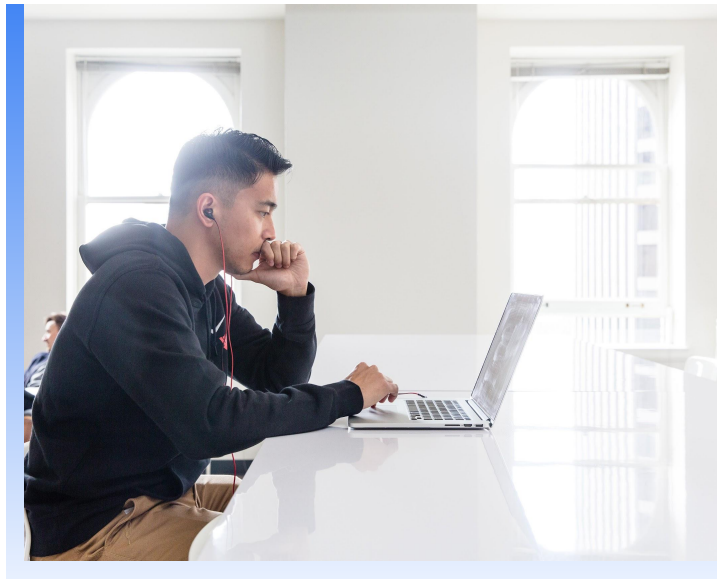
- セマンティックレイヤー
- コンポーザビリティ
- データの民主化
- ETL と ELT
- ローコードとノーコード



なぜアドホックな 分析環境をつくるのか

アドホックなデータ分析環境が欲しい理由

- すぐにデータが欲しい
- ビジネスサイドもデータを使いたい
- 意思決定スピードを上げたい
- 誰でも簡単に扱いたい
- 他に影響を与えたく無い



アドホックな分析ができるまで

データ収集

すでにあるデータ、これから欲しいデータ。複数に散らばったデータを集める。

データ加工

そのままでは使いづらいデータに意味を与えるため定義する。

分析

データと分析したいことがそろったので分析を開始する。



データ変換

形式がバラバラだったり、分析しづらい状態のデータを変換する。

データ探索

分析したいことを専門部署に依頼する。事業サイドにデータ分析の知識がなかったり、技術サイドの事業理解が低いと、手戻りが発生しやすい。

最短・最速のルートを探すには

目的を決めておく

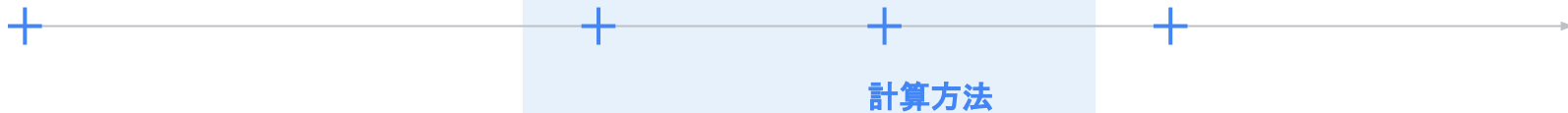
どのテーブル見ればいいのか後回し

セマンティックレイヤー

計算方法は後回しで先に指標を作る




データを集める

ここから始める



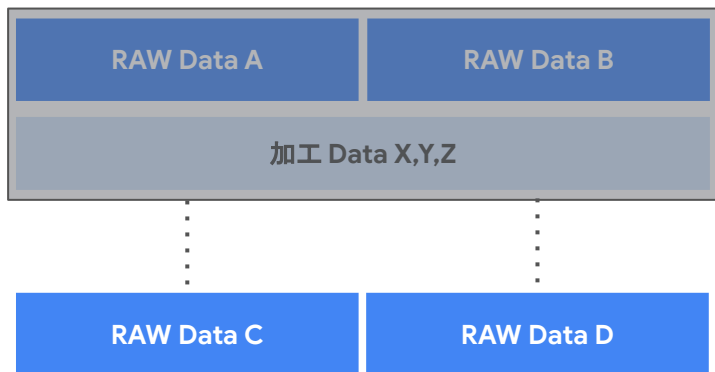
BI が民主化を加速

- データガバナンスの強化
- DataOps の推進
- API 経由でさまざまなアプリに
エンベッド

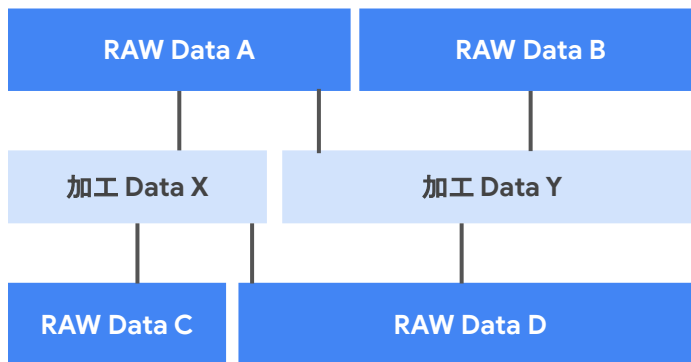
	第 1 世代	第 2 世代	第 3 世代
構造	 集中	 分散	 民主化
ユーザー	専門家	IT 人材	誰でも
高度な分析に かかる時間	数日～ひと月	1日～数時間	瞬時～数分
価値	データを一元管理して、専門の部署が分析を行って、レポートを作成できる	データソースをつなぐだけでグラフを作り、直感的に分析できる	分析したいデータに 自由にアクセス でき、 AI による新たなインサイト が提供される

コンポーザビリティ

構成・結合が容易でない



構成・結合が容易である



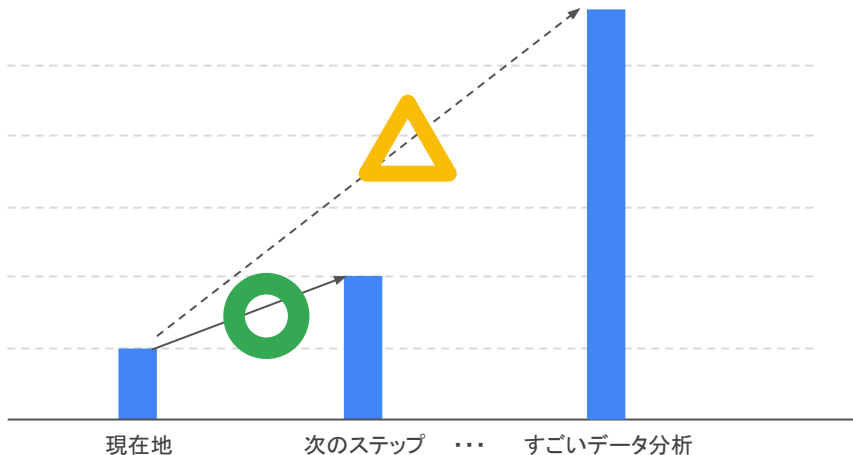


ビジネスに正解は無いが、真実（データ）はひとつ。データが民主化されていると、ファクトベースで意思決定ができる。

ステップ論：モチベーションとデータ量

現在地はどこで、次の目標地点はどこか

必ずしも SQL などを扱えるスキルを身に付けないといけないということではない。今持っているスキルセットで、現在保有しているデータを整理するだけで、できることはまだまだある。いきなり大量のデータ分析を始めるのは難易度が高い。

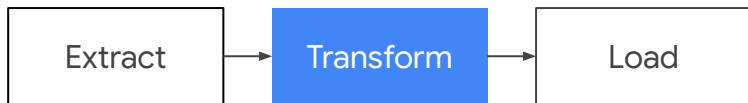




どうやって分析環境つくるか

データは活用は前処理がとにかく大変

活用方法が決まっていたら ETL



ETL で最も効率的な変換処理をおこない、活用まで一気に持っていく。

活用方法が決まっていなければ ELT

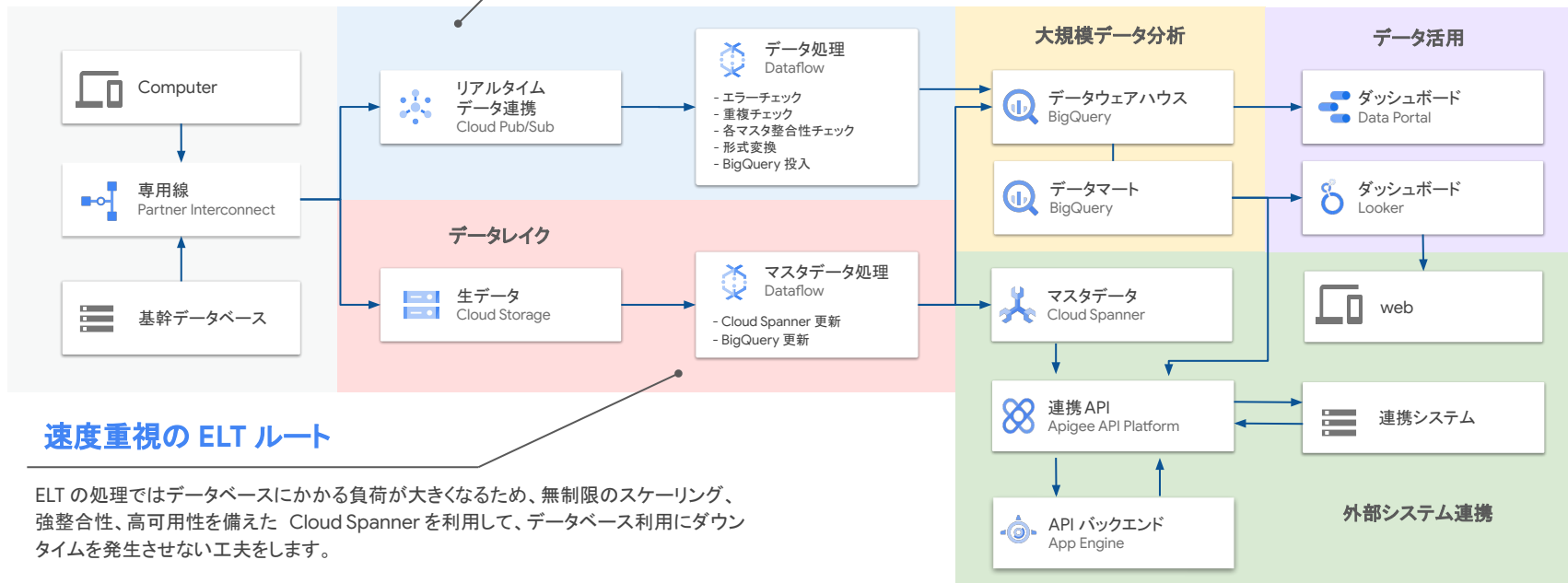


ELT でとりあえずクラウドに上げて活用する準備をしておく。

サンプル アーキテクチャ

加工重視の ETL ルート

データの量や形式、場所も増えていくと、スムーズに有効活用できなくなるため、ETL でデータを利用しやすく統一して DWH に保存することで効率化します。

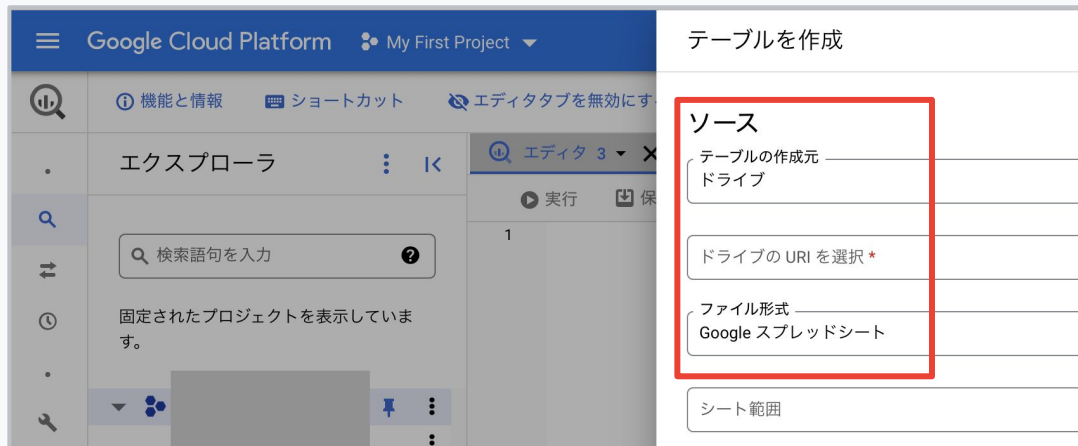
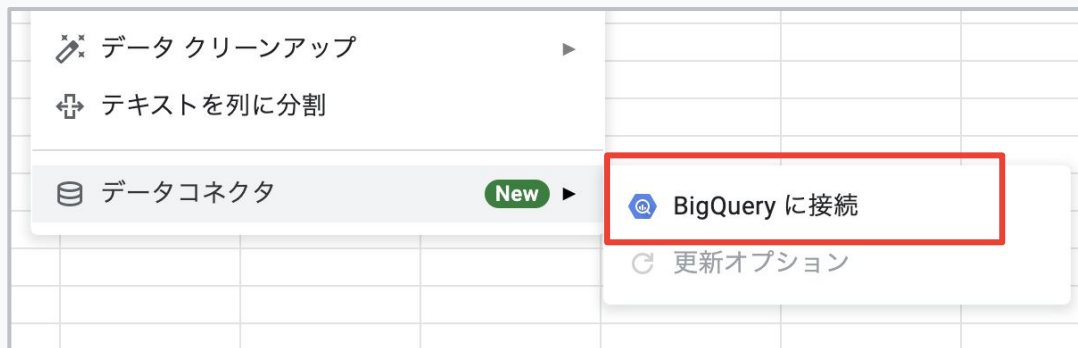




Connected Sheet

Spreadsheet を UI として BigQuery のデータを利用可能。

BigQuery は Drive の Spreadsheet をテーブルとして指定可能なため、相互参照可能な関係にある。



ノーコードとローコードに向いているケースとは

いい例

瞬間的に何らかの変化をしないといけない、システムが解決になる場合。

いつでも捨てられるもので、いつ捨ててもいいつもりで作り、レガシー化させない。

例: ワクチン接種の受付システムなど

よくない例

大掛かりで長期的な運用を前提としたシステムにした場合。レガシー化が早い。

自由で気軽に作りすぎて管理が行き届かず、シャドーIT などリスクの温床となる。



インハウスで何をやるか

パートナーから 得られるもの

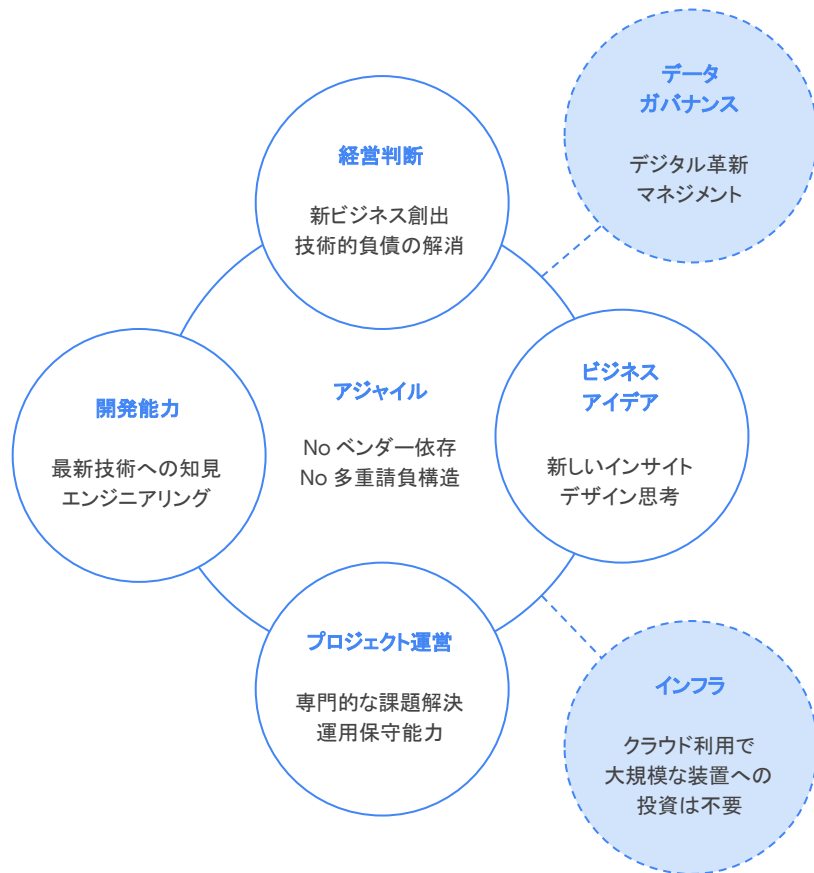
- ノウハウ
- 時間短縮



Cloud Ace の提言「SI 2.0」

SI 2.0 が目指すのは IT の民主化です。

アジリティの高い開発を繰り返し持続的に行える体制を作るためには、投資先や人材計画の見直し、内製化や開発手法の再検討、ユーザー企業とベンダーの関係性など、変化を求められます。



まとめ

- セマンティックレイヤー(目的から定める要になる)
- コンポーザビリティ(疎結合なデータ)
- データの民主化(専門家以外も使えるように)
- ETL と ELT(状況に応じて使い分ける)
- ローコードとノーコード(いつでも捨てるつもりで)
- ビジネスに正解は無いが真実(データ)はひとつ

Thank you.

