データサイエンスすいすい会

第8回
「データ活用を推進するマネージャが知っておくべき自然な摂理」

GRI

ForecastFlow

ナビゲータ

■ 古幡征史(PhD)
■ 株式会社GRI 取締役
■ 経歴
- 筑波大学　第三学群社会工学
- 筑波大学大学院　経営・政策科学研究科修了(MS)
- University of Western Sydney, PhD in Computer Science（コンピューターサイエンス博士号）
- Universite Toulouse 1 Capitole, Doctorat en Informatique（コンピューターサイエンス博士号）

- 丸文株式会社/KPMGコンサルティング/北陸先端科学技術大学院大学/University of Southern California/ドワンゴ

■ 分野
- データ分析＆シミュレーション
- データ分析基盤構築
- AI開発
- BI構築
- マーケティング、サプライチェーン、物流管理、配送管理、営業、経営、人事管理、財務管理、オークション、制度設計
- 東京ガス、読売新聞、ソニーネットワークコミュニケーションズ、ドワンゴ、アミューズ、ジブラルタ生命保険、野村総研、東京証券取引所、スパークス、米国運輸省、米国海軍研究所、NASAを含む50社以上

©GRI Inc.

すいすい会の紹介

■データサイエンスの領域に関して知見を共有しあい、みんなで実践的な理解度を高めたい

■自分の知見などを自由に発言してください
（呼び水のテーマ紹介をします）

■しばらくの期間は、自動機械学習を中心にお話をさせてください

■資料：　GRIホームページ: https://gri.jp/news/12924

■Slack　ForecastFlowチャンネル:
https://join.slack.com/t/forecastflowusers/shared_invite/enQtNTgyMjcxOTg0NzcxLTBkOWEzYWMwNDJmNTkyMDQzYmIxYWU0YWI4ZmU3ZDU0ZTMxNDUwODAxMWFmYmU1YjJiZGI0MjRhYWYyYTNlZTQ

■YouTube: GRIチャネル: https://www.youtube.com/channel/UCDVGqf-dgczYHMfPzpQ0jNQ

©GRI Inc.

過去のすいすい会の動画（YouTube）

- 第１回「ビジネスでAIを活用するための問題設定」
  - https://www.youtube.com/watch?v=ppkgYD69NDU
- 第２回「アンケートと機械学習」
  - https://www.youtube.com/watch?v=-s1PcLQUBNI
  - https://www.youtube.com/watch?v=ylVT3OyayEk
  - https://www.youtube.com/watch?v=sFVabtfNvIk
- 第３回「分類問題の精度評価手法と不均衡データ」
  - https://www.youtube.com/watch?v=q6WJDTOgotA
- 第４回「自動機械学習での特徴量の作り方」
  - https://www.youtube.com/watch?v=Ms52EnCRk8g
- 第６回「予測スコアを用いた効果的な施策実施」
  - https://youtu.be/8Sm8ex6taE8
- 第７回「時系列のセンサーデータを扱ったPredictive Modelでの予測について」
  - https://youtu.be/P09VdyyVtHs

©GRI Inc.

データサイエンスすいすい会

第8回
「データ活用を推進するマネージャが知っておくべき自然な摂理」

GRI

ForecastFlow

データは21世紀の新しい石油

https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

©GRI Inc.

石油は精製して付加価値、データも同様

原油探鉱 → 石油開発生産 → 精油 → 蓄積 → 輸送 → 販売 → 消費

水平統合＆垂直統合

©GRI Inc.

データの価値（生データから、いかにして加工して活用するか）

データ分析基盤

AI（人工知能）

BI（可視化）

MA（マーケティングオートメーション）

©GRI Inc.

小さな所から変化は始まっている（優秀な若手の流出）

- まだＯＯで疲弊してるの？
- まだＯＯしてないの？
- 人生で無駄してない？

転職を考える理由
- 上司／会社がデータ活用の理解がない
  - 必要な投資をしない（ツール／基盤）
  - 分析結果に向き合ってくれない
  - 分析に関する指示してくれない
- ずっと非効率な仕事の仕方をしたくない

©GRI Inc.

マネージャの難しい立場

- データ分析をやったことないから
  - 分析の指示が難しい
  - 費用対効果の説明が難しい
  - 何をすれば良いか分からない

では、どのように進めるべきか？

©GRI Inc.

マネージャが知っておくべきこと

- データ利活用の仕事のやり方
- 人材像と組織像
- 育成方法
- データ分析基盤
- 自動機械学習ツールの活用
- 落とし穴

©GRI Inc.

データ利活用の仕事のやり方

データ利活用の狙うべき仕事のやり方

今までのデータの使い方
（過去に起こったことを基準にデータを利用）

Descriptive Analytics
Diagnostic Analytics

決まったことをデータ確認
＋
何か起きたらデータ確認

- 受動的
- 硬直的
- 対処に時間を要する
- 人が作表する

AIやBIツールを利用したデータの使い方
（これから起こることを想定しデータ利活用）

Predictive Analytics
Prescriptive Analytics

これから起こることを想定した自動化
＋
基準から外れた状況の対策

- 能動的
- 柔軟
- 準備をしておく
- 人が自動化の設定する

※組織の中でデータに関する
共通認識レベルが上がる

©GRI Inc.

データ利活用の狙うべき仕事のやり方

今までのデータの使い方
（過去に起こったことを基準にデータを利用）

＜定型レポート確認＞

データ収集
↓
定型帳票作成
↓
課題点の洗い出し

＜課題対応＞

課題発生
↓
データ収集
↓
データ集計（作表）
↓
問題の数値化
↓
対応策

AIやBIツールを利用したデータの使い方
（過去＋これから起こることを想定してデータを利用）

＜自動化の設定＞

分析対象／モデルの選定
↓
データフロー作成
↓
ダッシュボード作成／
機械学習で訓練
↓
施策の立案／アラート基準

＜定型業務＞

データフロー実行（定時）
↓
ダッシュボード更新確認／
アラート確認／
機械学習で推論
↓
施策の実施／課題の検知
↓
効果検証

©GRI Inc.

データ利活用の狙うべき仕事のやり方

今までのデータの使い方
（過去に起こったことを基準にデータを利用）

＜定型レポート確認＞

データ収集
↓
定型帳票作成
↓
課題点の洗い出し

＜課題対応＞

課題発生
↓
データ収集
↓
データ集計（作表）
↓
問題の数値化
↓
対応策

80%の作業時間

AIやBIツールを利用したデータの使い方
（過去＋これから起こることを想定してデータを利用）

＜自動化の設定＞

分析対象／モデルの選定
↓
データフロー作成
↓
ダッシュボード作成／
機械学習で訓練
↓
施策の立案／アラート基準

＜定型業務＞

データフロー実行（定時）
↓
ダッシュボード更新確認／
アラート確認／
機械学習で推論
↓
施策の実施／課題の検知
↓
効果検証

40%の作業時間

©GRI Inc.

人材像と組織像

データ活用組織を作る際、真っ先に欲しい人物

■データエンジニア

–ビジネスアナリストが使いやすいデータ準備

■分析エバンジェリスト

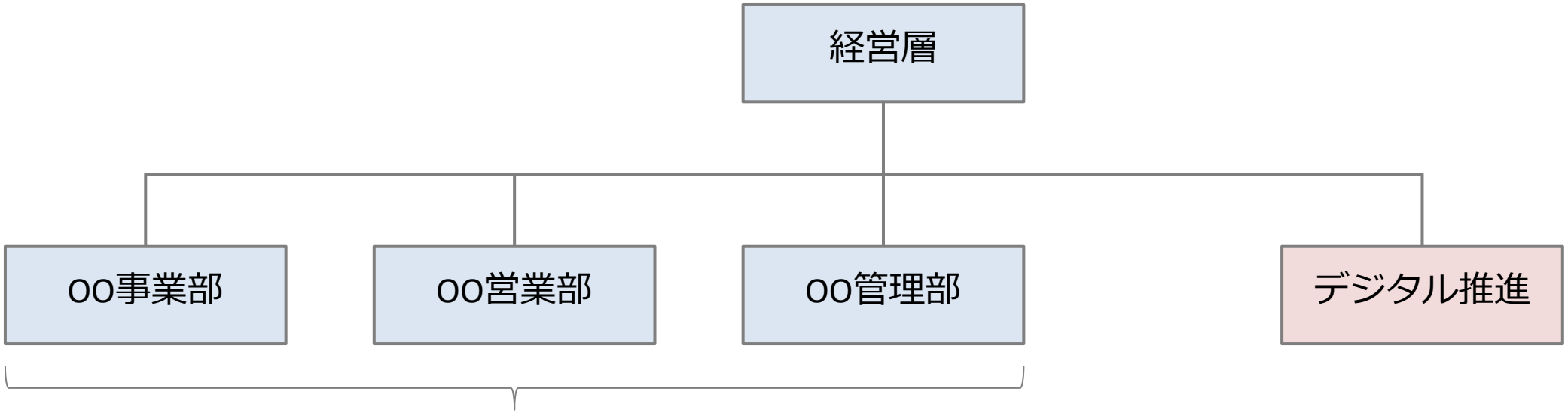
–分析テンプレート作り

–教育

※多くのビジネスアナリストを育成できる

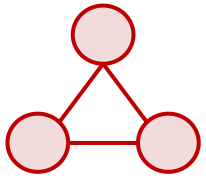
©GRI Inc.

DX組織



- 原局の各チームで分析PDCA（定期）
- 個別分析テーマ（アドホック）

マネージャ



施策　分析

ビジネスアナリスト
たまにデータエンジニア

- 全社横断分析
- 戦略的重点分析
- 分析ツール選定
- データ分析基盤の管理
- 分析テンプレート作成
- 教育

ビジネスアナリスト
データエンジニア
分析エバンジェリスト
データサイエンティスト

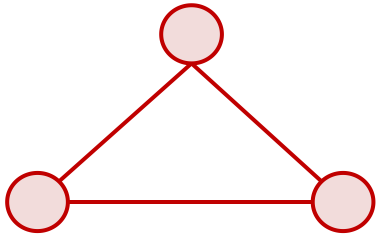
ワークフロー、人種、データの特徴から、このようになる
原局のビジネスアナリストの孤立を防ぐ心がけ

©GRI Inc.

DX組織

マネージャ

- 予算獲得
- KPI設定
- 優先順位
- 他部門折衝
- 教育



施策系 ビジネスアナリスト

- ビジネス理解（分析対象の理解）
- ダッシュボードで数値確認
- 施策立案
- 施策選定
- 他部門との交渉

分析系 ビジネスアナリスト

- データ理解
- データ収集
- データ整形
- ダッシュボード作成
- 機械学習活用
- データ分析
- 効果検証

©GRI Inc.

それぞれ全く異なる人種であること

©GRI Inc.

システムエンジニアとデータエンジニアの違い

システムエンジニア	データエンジニア
• 正しく処理できるシステムを構想（具体度が高い）	• 分析で使いやすい中間テーブルを構想（抽象度が高い）
• 処理を単一の処理で考えがち（ファンクション作り）	• 処理をフローで考えがち（パイプライン作り）
• データを正規化する方向性で考える（冗長で無駄なメモリ使用を防ぐ）	• 分析に使いやすくする方向性で考える（列が長いようなテーブルを考える）
• リレーショナルデータベースが中心	• 列指向データベースが中心

©GRI Inc.

データエンジニアとビジネスアナリストの違い

GRI BIGDATA SCIENCE & RESEARCH

データエンジニア	ビジネスアナリスト
• 分析で使いやすい中間テーブルやパイプライン作りをやりたい	• 前処理は興味がないがやらなければいけないこと
• ビジネスアナリストのやりたいことを理解できる	• 分析や施策、効果検証をしたい
• 人数が少なく、データのカバー領域が広い	• 人数が多く、データのカバー領域が狭い

データエンジニアには、より高価なETLツールの提供

©GRI Inc.

育成方法

会社としてビジネスアナリストを育てるために

■分析エバンジェリストから教育の展開

–分析テンプレートを作り、使い方をビジネスアナリストに広める

■ビジネスアナリスト同士の社内交流の場を用意

–分析例の共有＆フィードバック

■ビジネスアナリストを現場とＤＸ推進の兼務

–一定期間教育し、元の現場に戻る

■ビジネスアナリストの共有

–複数の現場を同時に担当

©GRI Inc.

現場でビジネスアナリストと仕事をするために

- １か月、３か月、６か月くらいの単位で、自動化したい内容を決める
  - ビジネスインパクトの想定が困難だが…

- 達成するために必要なタスクを確認する
  - データ収集、データの意味確認、データ前処理、分析、解釈、施策立案

- アジャイル開発の仕組みを取り入れる
  - １週間 or ２週間の単位に区切り、任せる範囲を確認
  - 担当者の悩みを積極的に確認する

©GRI Inc.

データ分析基盤

RDBと列指向DBの違い

RDB

MySQL ORACLE Microsoft SQL Server

- システムエンジニアがよく扱う
- 大規模データの扱いが遅い
- データ更新が速い

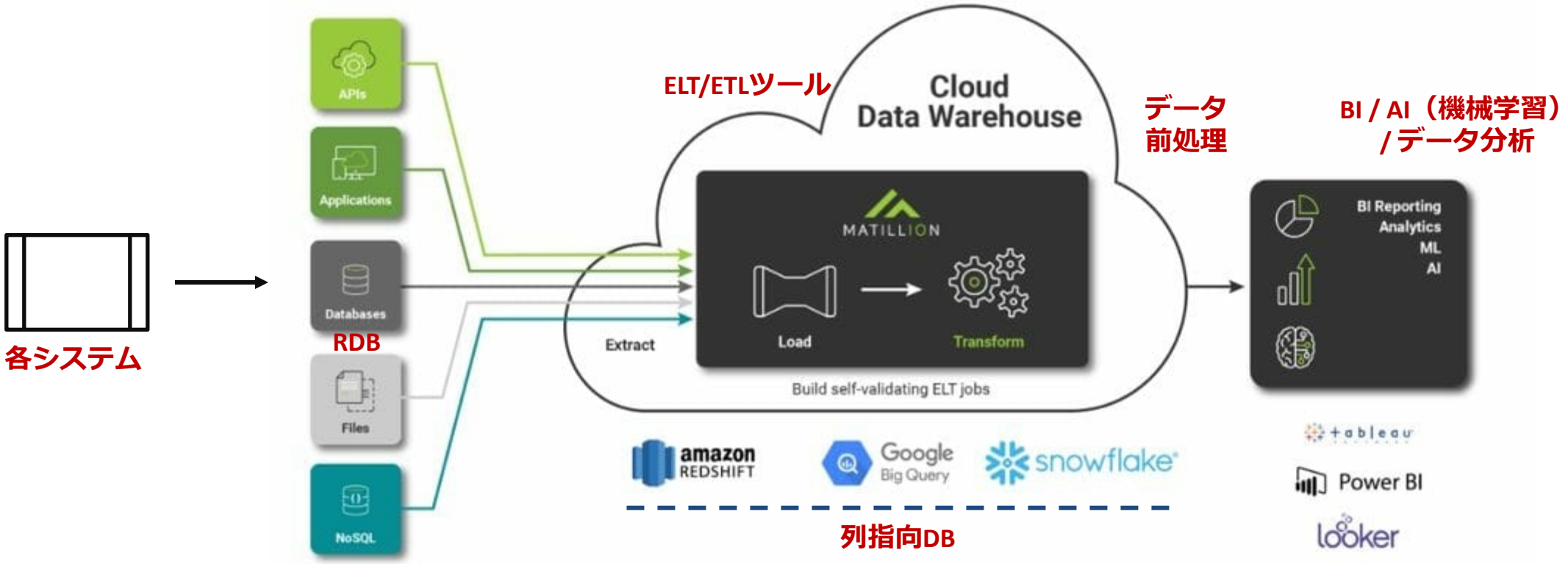
列指向DB

BigQuery amazon REDSHIFT snowflake VERTICA

- データエンジニアがよく扱う
- 大規模データの扱いが速い
- データ更新が遅い

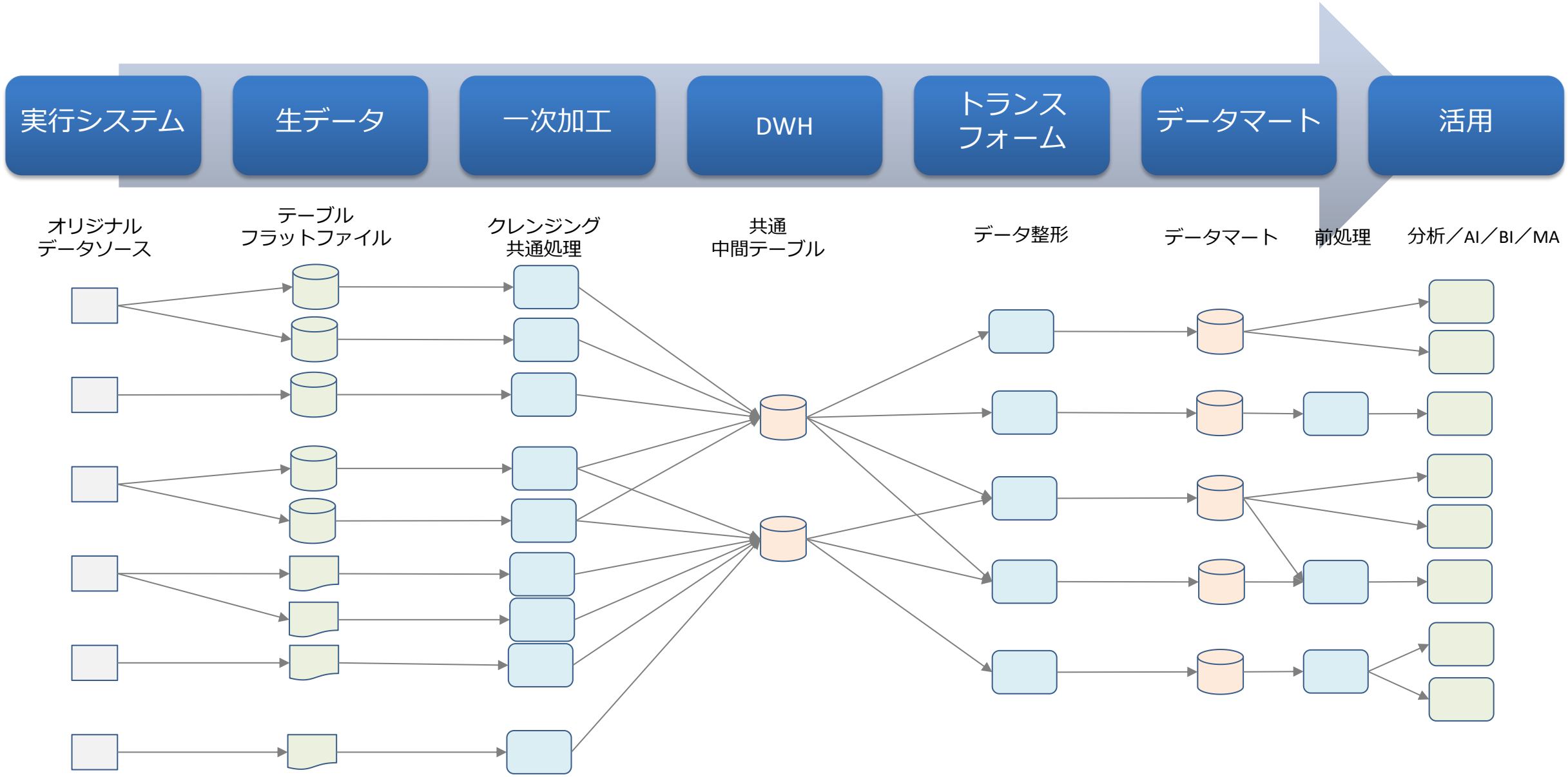
©GRI Inc.

システムの登場人物



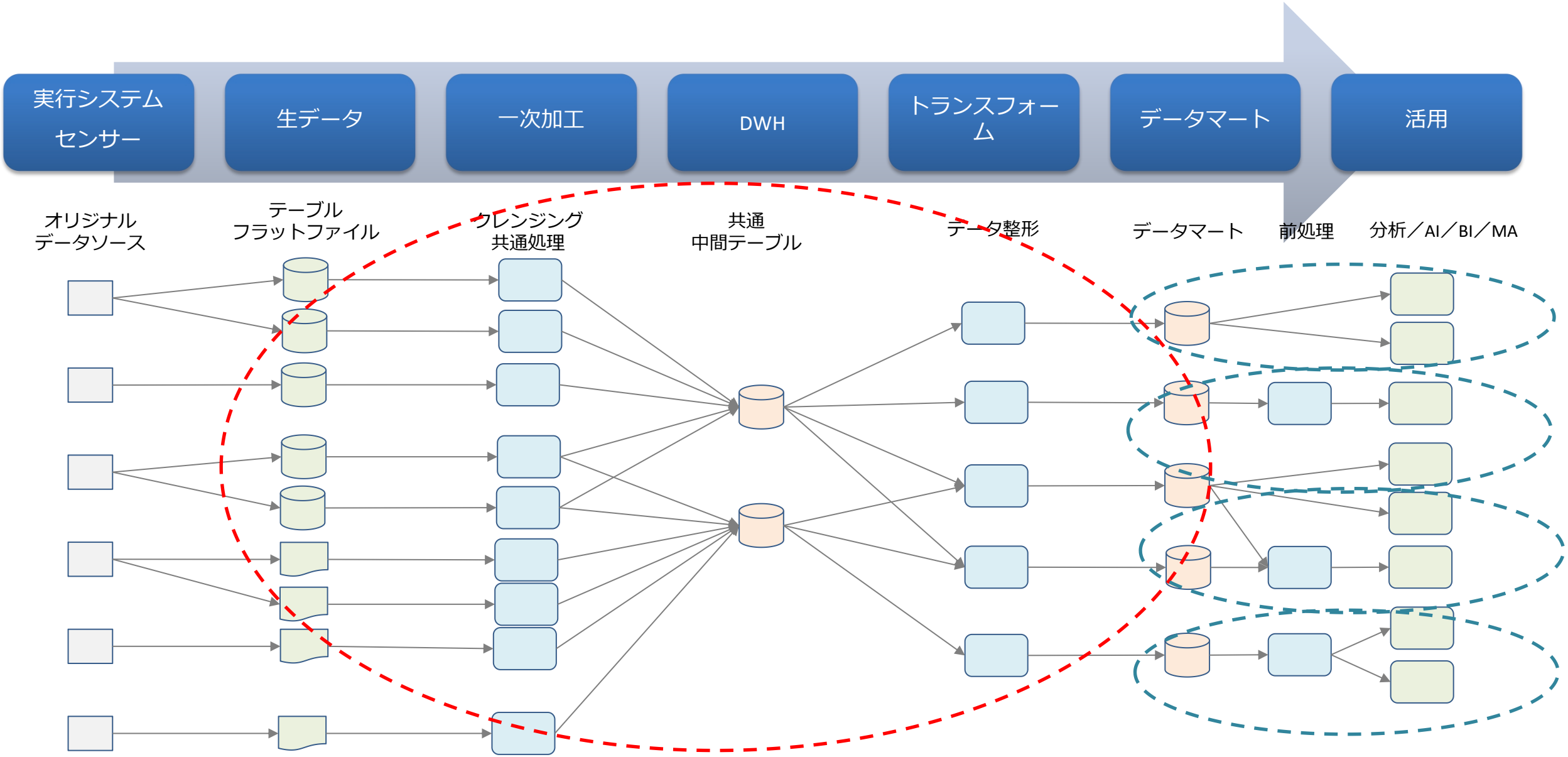
©GRI Inc.

データフロー



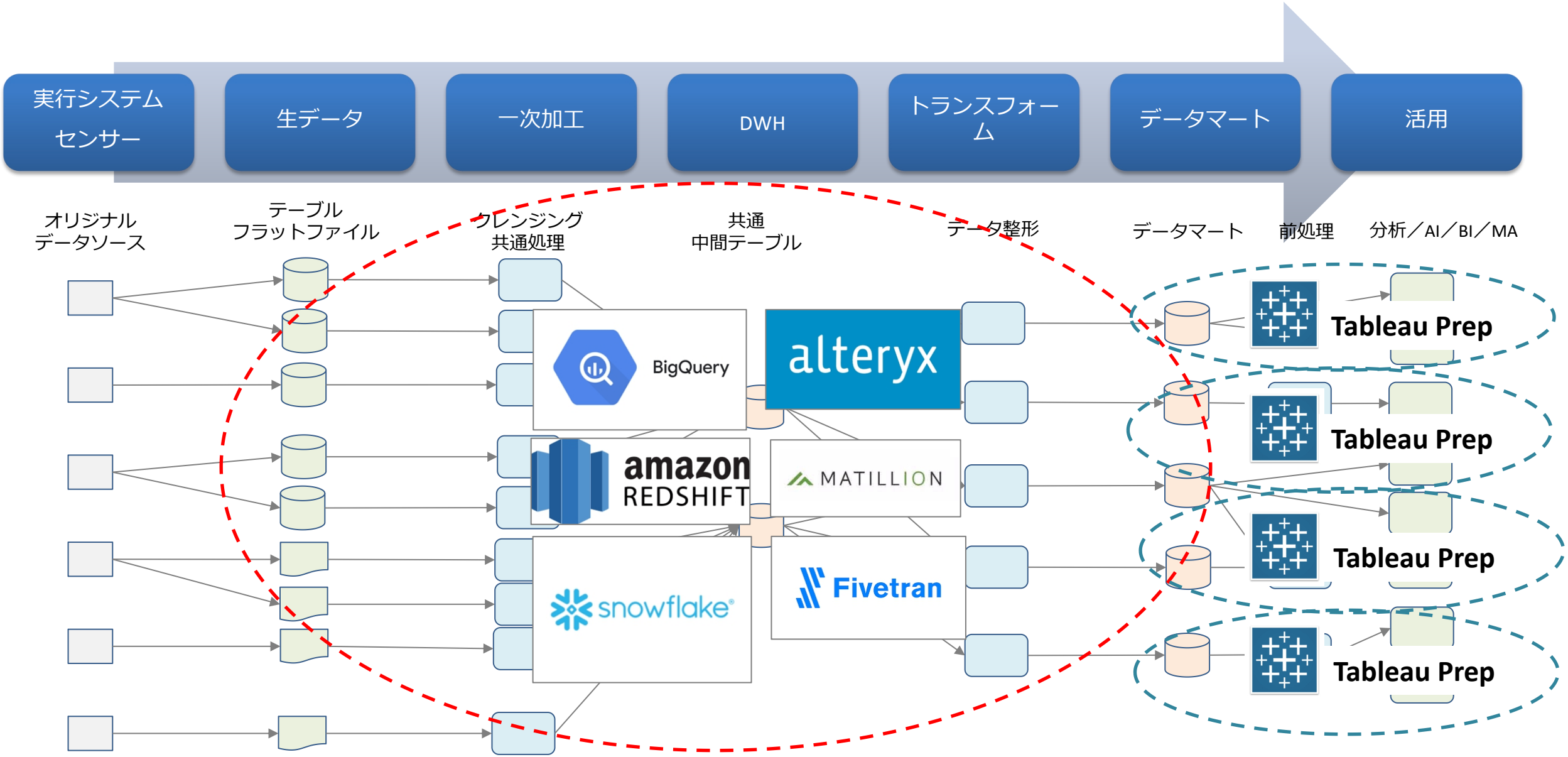
©GRI Inc.

データエンジニアとビジネスアナリストの担当領域



©GRI Inc.

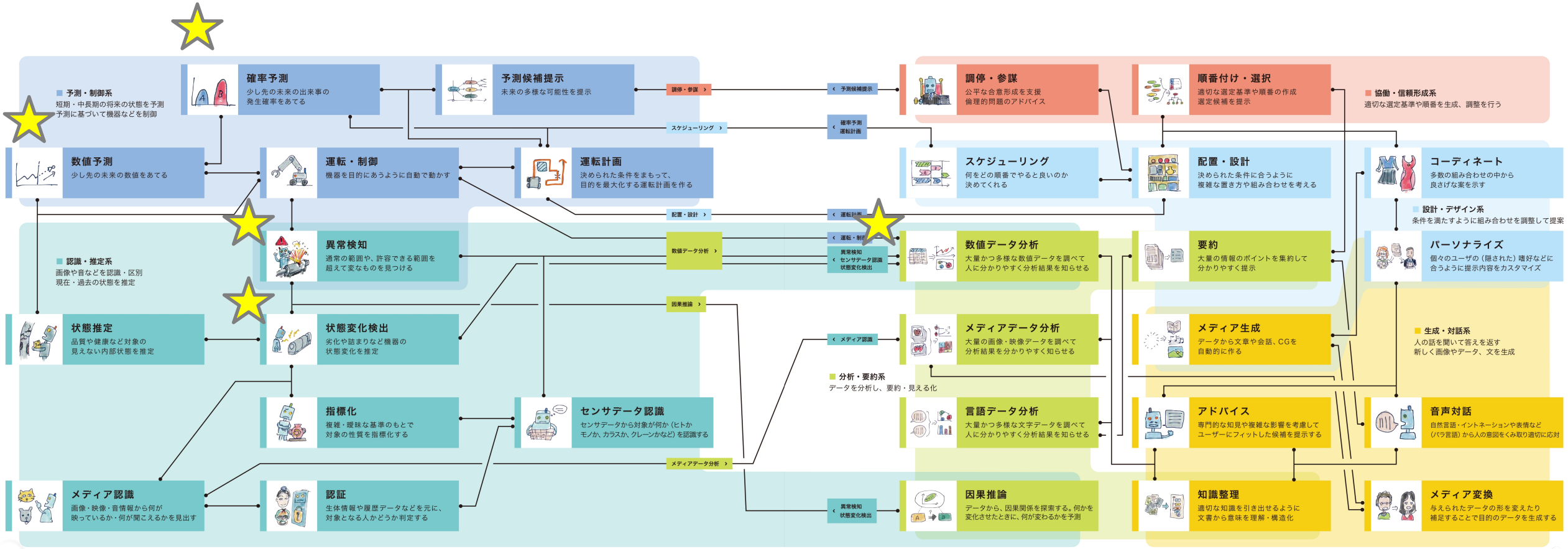
データエンジニアとビジネスアナリストのETL/ELTツールの例



©GRI Inc.

自動機械学習ツールの活用

AI & 機械学習の利用



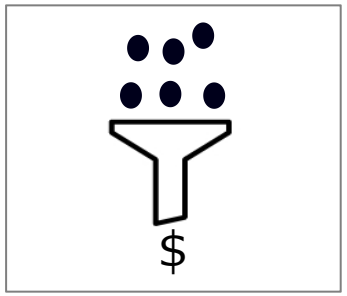
https://www.ai-gakkai.or.jp/resource/aimap/

©GRI Inc.

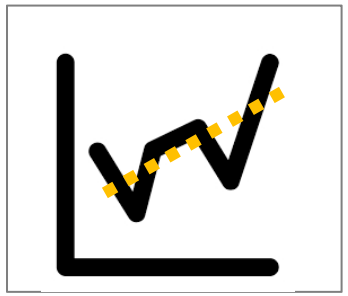
自動機械学習の応用例

■ サブスクリプション・ビジネス
- 解約防止施策
- Life-Time Value の予測

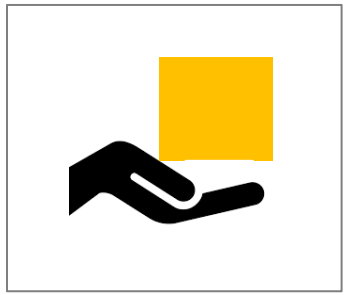
■ 優良顧客の育成
- 優良顧客の態度変容予測
- 新規顧客の定着パターン



■ リード・ナーチャリング
- リード顧客をコンバージョンへ導く施策選定



■ 売上予測／数値予測
- 店舗ごとの売上予測
- 視聴率予測



■ レコメンド
- 商品の購買予測確率からレコメンド

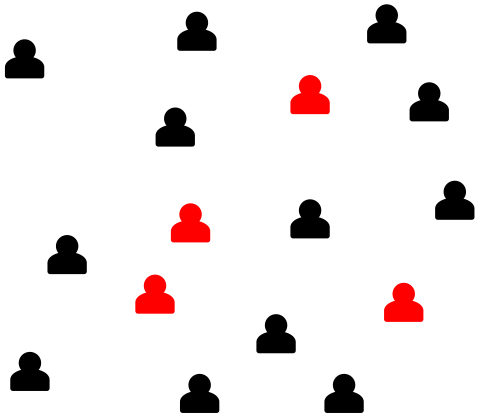
■ 工場での検品／IoT
- 印刷工場での入稿画像の検版
- 飼育牛の発情検知、分娩検知

©GRI Inc.

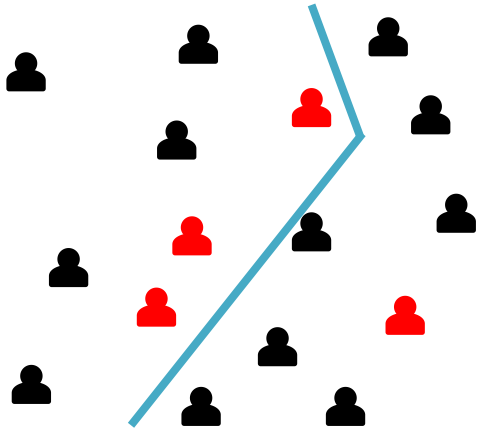
機械学習（ML: Machine Learning）の例

機械学習の問い：　継続と解約するユーザの見分け方を学習してください

ある一定期間後に継続／解約しているユーザ



継続状態以外のデータのみで、見分ける法則を
機械に学習させる（法則が分かると便利）



解約判定　　継続判定

©GRI Inc.

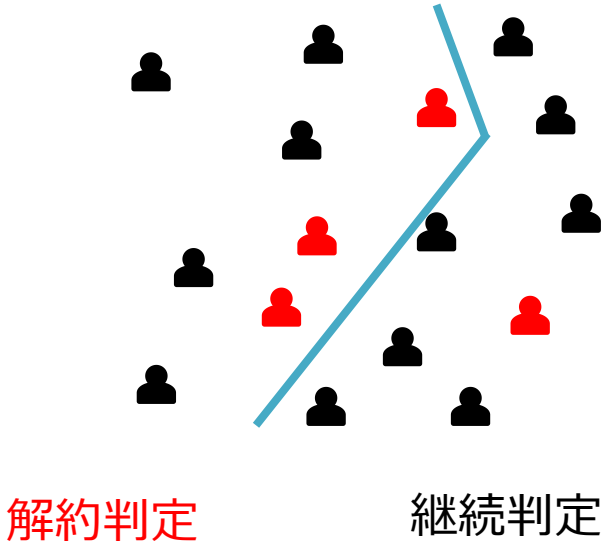
機械学習（ML: Machine Learning）の例

機械学習の問い：　一枚表の特徴量を使ってターゲット（継続と解約するユーザ）の見分け方を学習してください

一枚表

ターゲット　　特徴量

会員ID	会員状態	会員月数	居住地域	年齢	支払方法
1	継続	20	東京	32	クレジットカード
2	解約	4	埼玉	29	コンビニ
3	継続	15	千葉	63	自動引き落とし
4	継続	32	大阪	42	クレジットカード
5	解約	19	北海道	58	コンビニ
6	継続	20	鹿児島	23	自動引き落とし

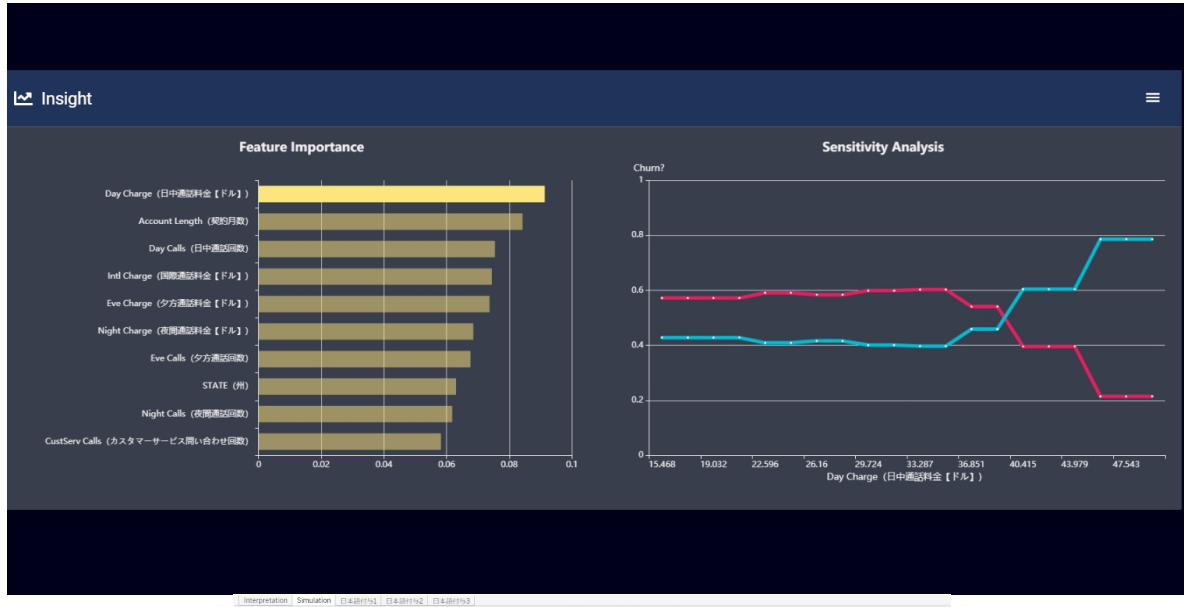


※大量の組み合わせから自動判定
例）100の特徴量を含む含まないの組み合わせだけで、1267650600228229401496703205376通り

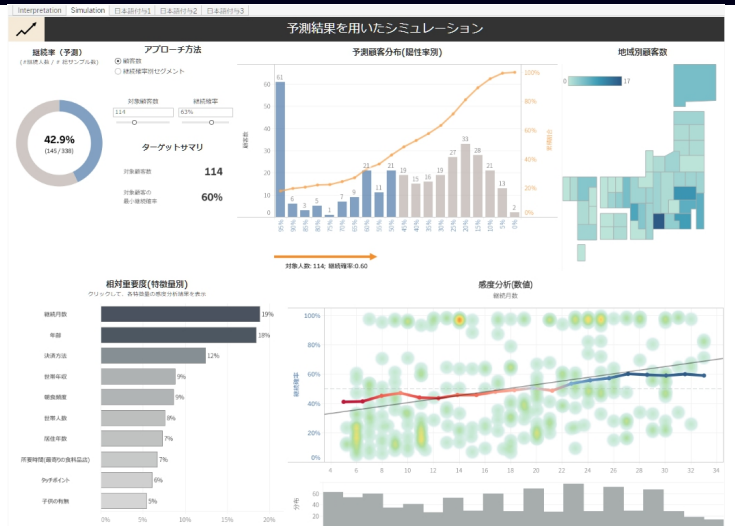
©GRI Inc.

自動機械学習でうれしいこと

- 一枚表さえあれば、自動で顧客の見分け方を学習
  - どのように見分けたかが分かる
  - つまり予期の仕方が分かり、事前的な施策に活かせる



- 学習した結果を使って、新しいデータを使い、ユーザごとに予測スコアを得られる
  - 予測スコアを用いて効果的な施策選定可能
  - より厳密な効果測定に使える
  - What-if分析で、仮にこうなったら、どのようなことが起こるかシミュレーションできる



https://public.tableau.com/profile/griinc6648#!/vizhome/MachineLearningInterpretation_JA/Interpretation

©GRI Inc.

自動機械学習ツールの活用

落とし穴（売上増加と費用削減の違い）

売上増加に繋がるテーマ

- 増えた売上で再投資
- 成功するSaaSはT2D3の成長軌道
  - triple, triple, double, double and double
  - 若手に大きなチャンス

費用削減に繋がるテーマ

- 人件費を削減するためだけ？
- 削減できたらマネージャも減らすことができる

（資本主義の前提は時間と共に成長）

- 削減できる人に対する育成方針をセットに
- 単に人員削減すると明るい未来を社員が描けない

- その技術を使った新しいサービスを生み出すことを考えるべき

©GRI Inc.

まとめ

本日のまとめ

- データ利活用した仕事のやり方（何かが起こる前に準備しておく）
- 人材像と組織像（似て非なる職種）
- 育成方法（エバンジェリストを作る。孤立させない、共有させる）
- データ分析基盤（適したツールを選定）
- 自動機械学習を使う例（予測数値ではなく施策に活用を追いかける）
- テーマ選定重要（売上系と費用系で違う）

©GRI Inc.

残っている質問（今後、説明していきたい内容）

- その他のビジネスの実践例を知りたい
- どのくらい正解データの数があるべきか
- 効果検証
- 様々な自動機械学習ツールの比較
- 様々なETL/ELTツールの比較
- Feature Store、分析基盤、CDPの説明

©GRI Inc.

ForecastFlowトライアル・アカウントの配布

■訓練する際、最大1,000行になります

■トライアル・アカウントの詳細ページ

- https://forecastflow.jp/free/

■申し込みメールを下記へ

- forecastflow@gri.jp

- ForecastFlow無料アカウント申請フォーム
  - お名前：
  - 所属名：
  - 部署名：
  - ForecastFlowで利用するメールアドレス：

©GRI Inc.