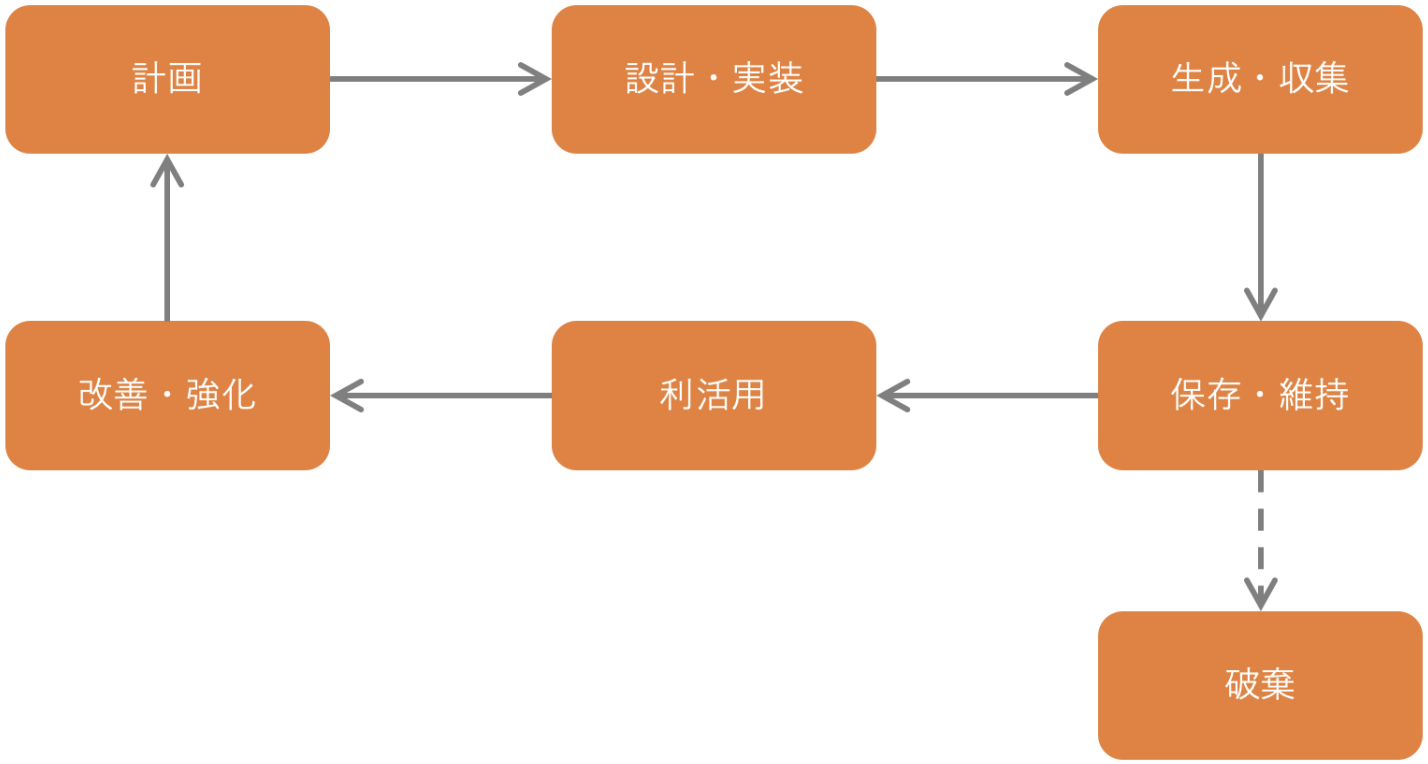
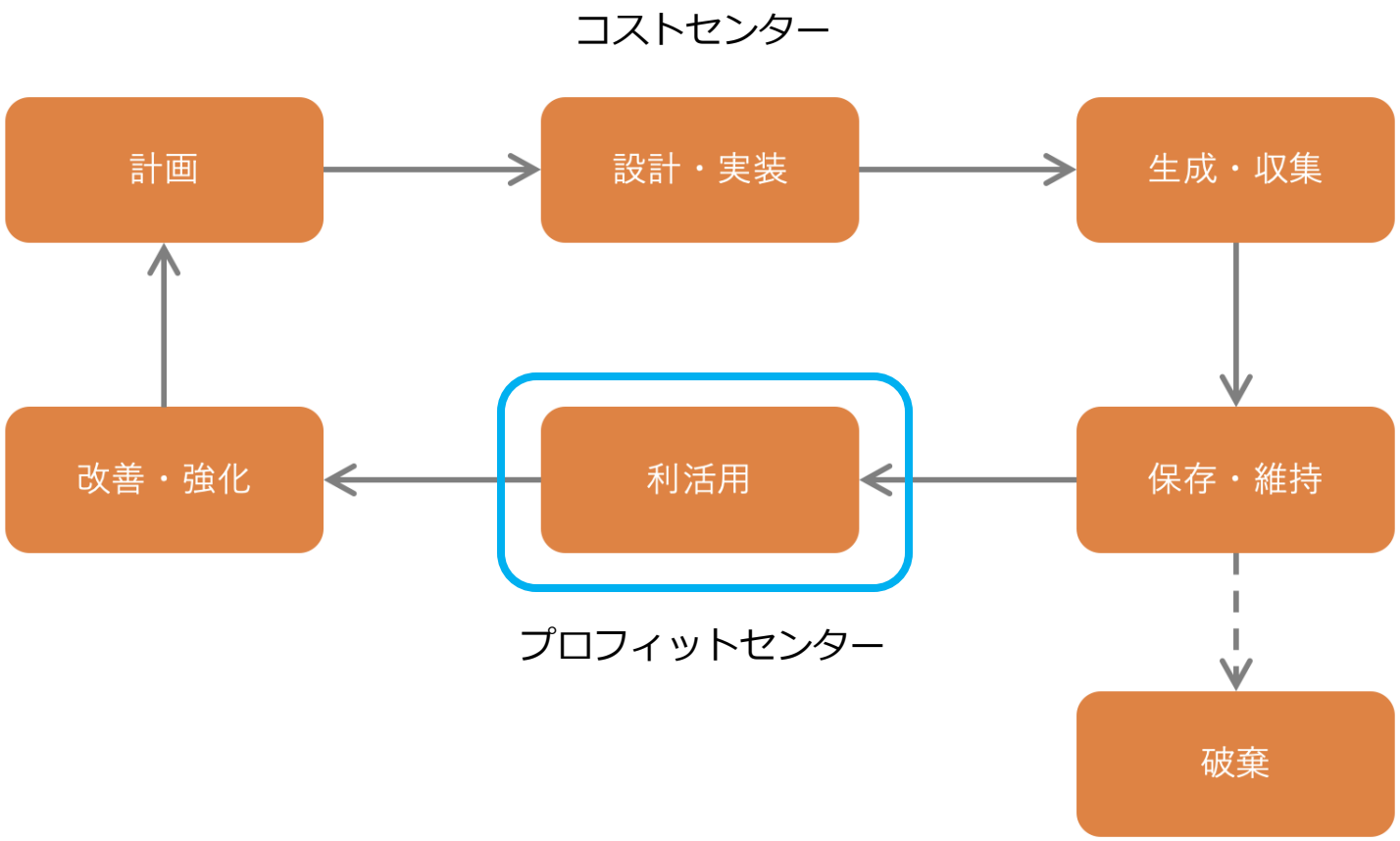
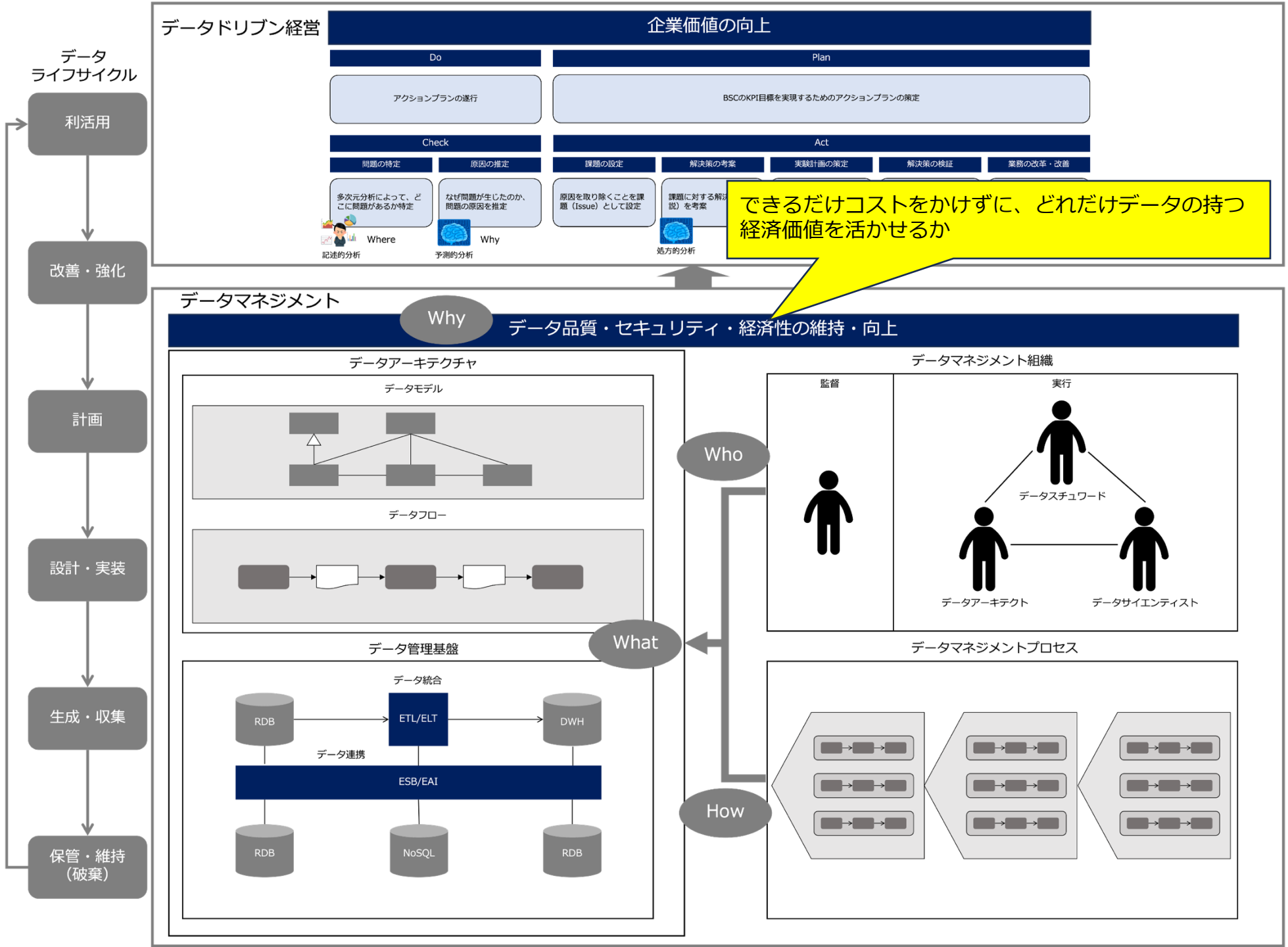
データのライフサイクル

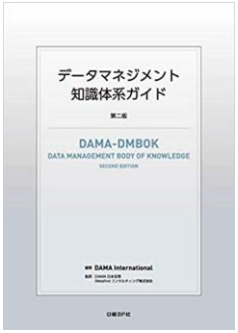


データのライフサイクル




データ
ライフサイクル
利活用
改善・強化
計画
設計・実装
生成・収集
保管・維持
（破棄）
データドリブン経営
企業価値の向上
Do
Plan
アクションプランの遂行
BSCのKPI目標を実現するためのアクションプランの策定
Check
Act
問題の特定
原因の推定
課題の設定
解決策の考案
実験計画の策定
解決策の検証
業務の改革・改善
多次元分析によって、どこに問題があるか特定
なぜ問題が生じたのか、問題の原因を推定
原因を取り除くことを課題（Issue）として設定
課題に対する解説）を考案
Where
Why
記述的分析
予測的分析
処方的分析
できるだけコストをかけずに、どれだけデータの持つ経済価値を活かせるか
データマネジメント
Why
データ品質・セキュリティ・経済性の維持・向上
データアーキテクチャ
データモデル
データフロー
データ管理基盤
データ統合
RDB
ETL/ELT
DWH
データ連携
ESB/EAI
RDB
NoSQL
RDB
Who
What
How
データマネジメント組織
監督
実行
データスチュワード
データアーキテクト
データサイエンティスト
データマネジメントプロセス

データマネジメント



データ資産

品質とセキュリティの高いデータを「データ利活用」のためにコストをかけずに迅速に準備することができる

データの準備ステップ

○○のために○○のようなデータ分析がしたい！

↓

必要なデータの特定

↓

データの構造化

↓

データの調査（有無、所在、セキュリティレベルなど）

↓

データの取得

↓

データの理解・クレンジング・加工

↓

データ分析

データ利活用に関するさまざまな問題

どこに何のデータがあるかわからない！

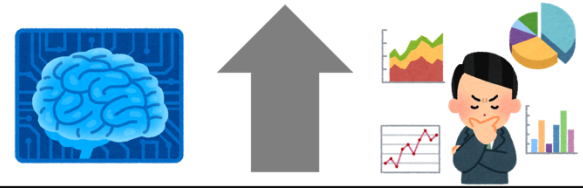
業務担当者

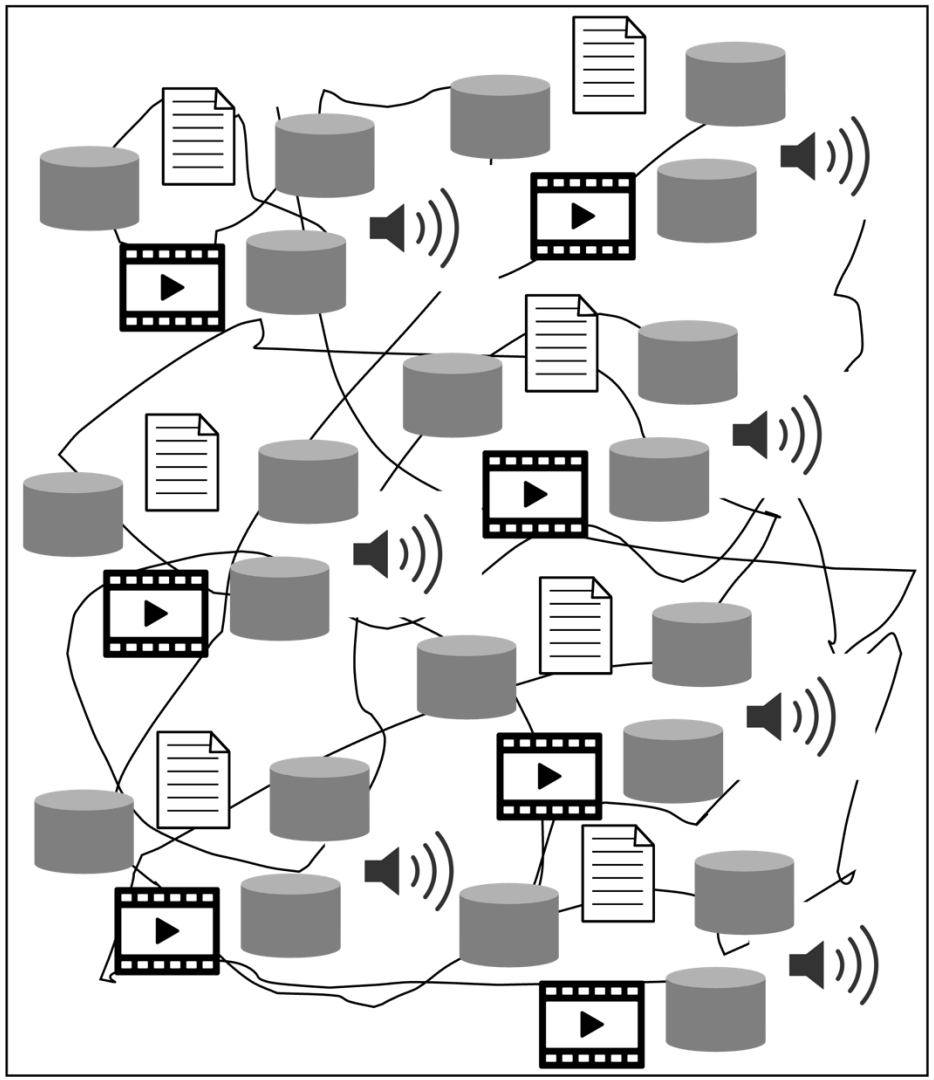
データ分析のための準備に時間がかかる！

あちこちに似たようなデータがあってどのデータを信用していいのかわからない！

機密情報に簡単にアクセスできてしまう！

企業価値





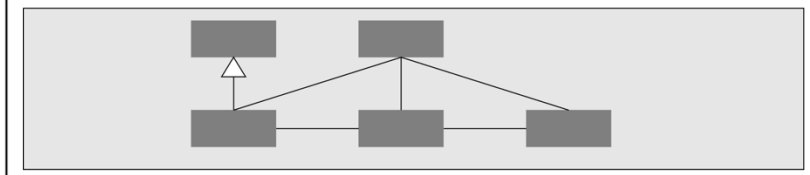
会社に眠るさまざまなデータ

資産化

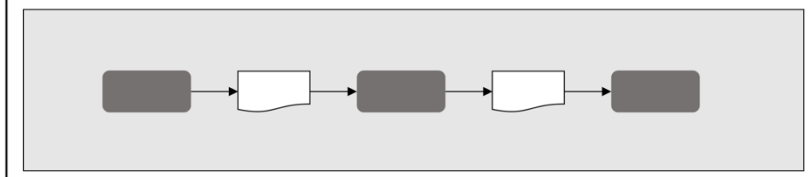
データマネジメント

データアーキテクチャ

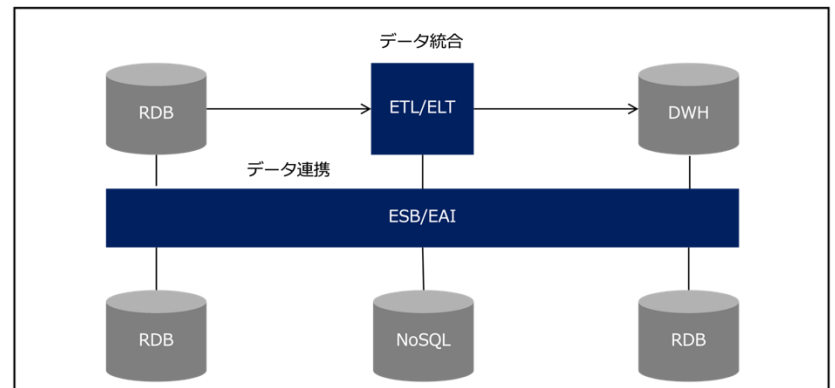
データモデル



データフロー



データ管理基盤



資産としての（稼ぐ力を持つ）データ

データ利活用に関するさまざまな問題

どこに何のデータがあるかわからない！

業務担当者

データカタログの整備

データ分析のための準備に時間がかかる！

データ品質の管理

あちこちに似たようなデータがあってどのデータを信用していいのかわからない！

データの統合・連携

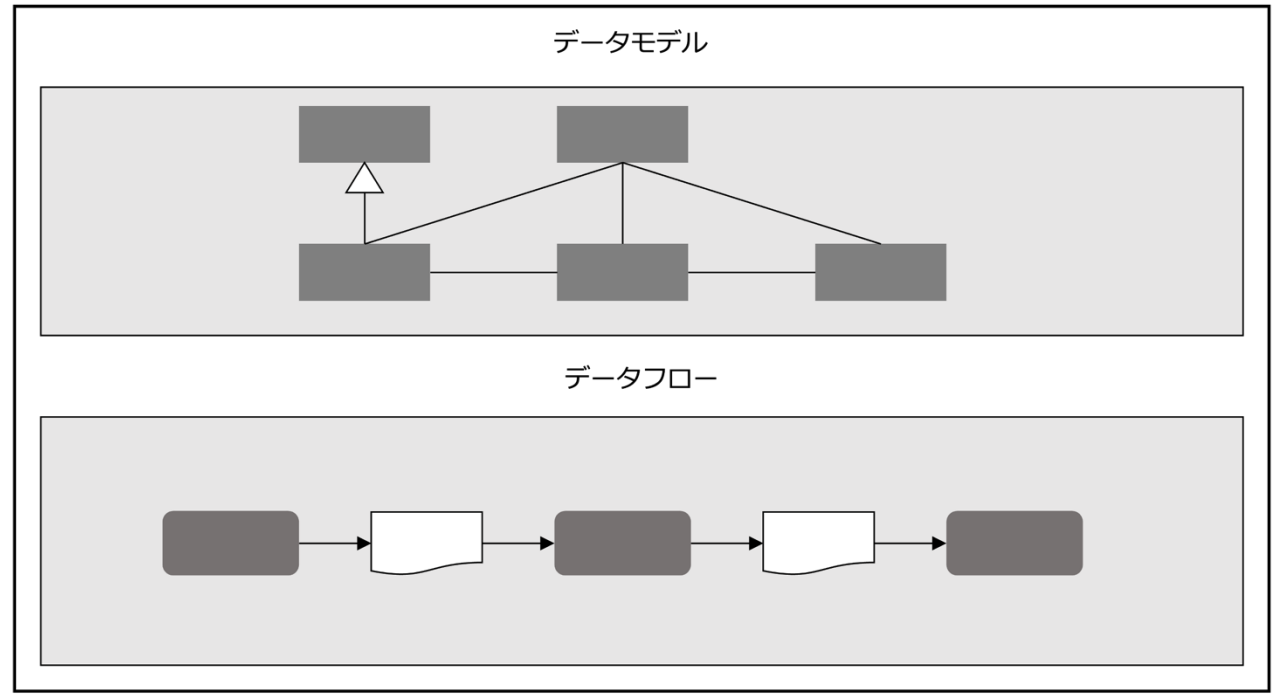
機密情報に簡単にアクセスできてしまう！

データセキュリティの管理

データの有無、所在、構造、来歴、品質の状態、セキュリティレベル、利用するときの連作先など。

メタデータ

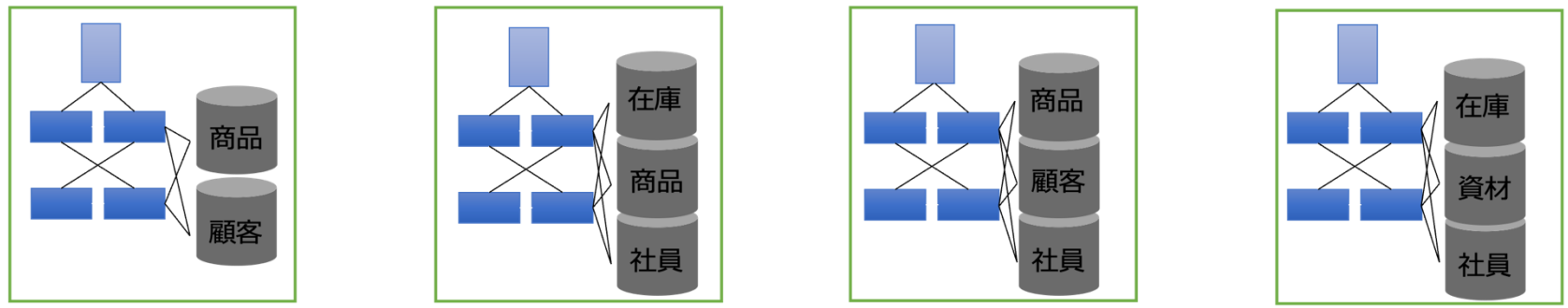
データアーキテクチャ



企業の現状

ビジネスの変化の加速＋ビジネスと情報システムの密接化

必要に応じて情報システムを開発



chaos

大規模で複雑なシステム（サイロ）が乱立

重複し整合していないデータの散在

全体を理解できる人や資料がない

雁字搦め

データ利活用に関するさまざまな問題

どこに何のデータがあるかわからない！

業務担当者

データカタログの整備

データ分析のための準備に時間がかかる！

データ品質の管理

あちこちに似たようなデータがあってどのデータを信用していいのかわからない！

データの統合・連携

機密情報に簡単にアクセスできてしまう！

データセキュリティの管理

品質の悪いデータの例

例えば、あるエリアの不動産物件の適正な家賃を推定する場合、広さや駅からの距離、築年数といった変数の影響が大きくなります。

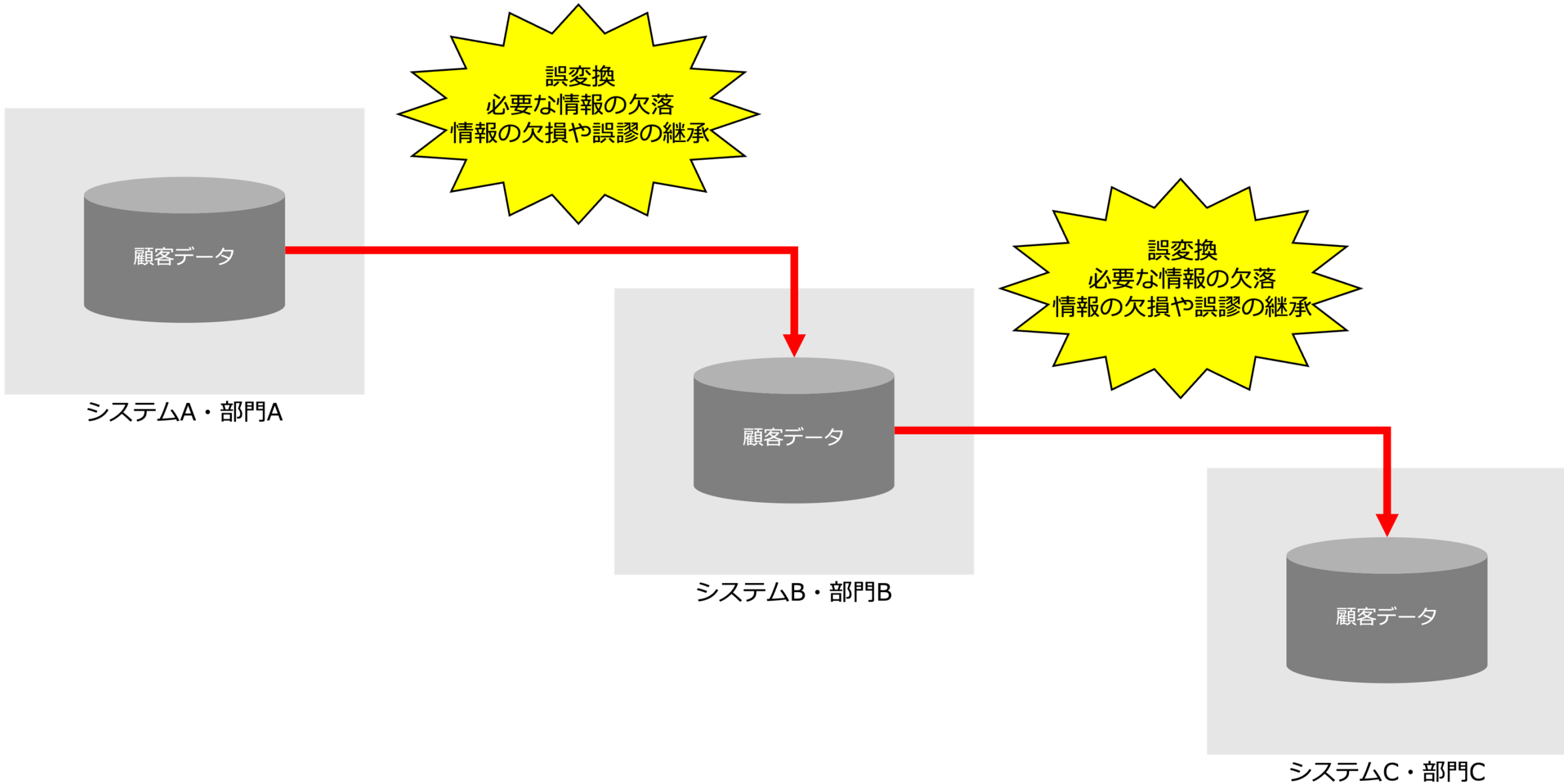
このとき、とある物件の建築年のデータが抜けたまま分析した場合、プログラムがどう判断するのかというと「建築年＝””」→「西暦00年に建てられた」→「2024年前に建てられた物件」と捉え、広くて駅近なのに、ものすごく古い物件なので「適正家賃は0.01円」という分析結果を出してきたりするのです。

出所：DATA is BOSS 収益が上がり続けるデータドリブン経営入門

データマネジメントの欠如がもたらすリスク

- 誤請求
- 顧客サービスコールの増加とそれを解決する能力の低下
- 事業機会の逸失による収益損失
- 合弁・買収の間に発生する業務統合の遅延
- 不正行為発覚の増加
- 不正なデータに起因する業務上の意思決定不備がもたらす損失
- 良好な信用力の欠如による事業の損失

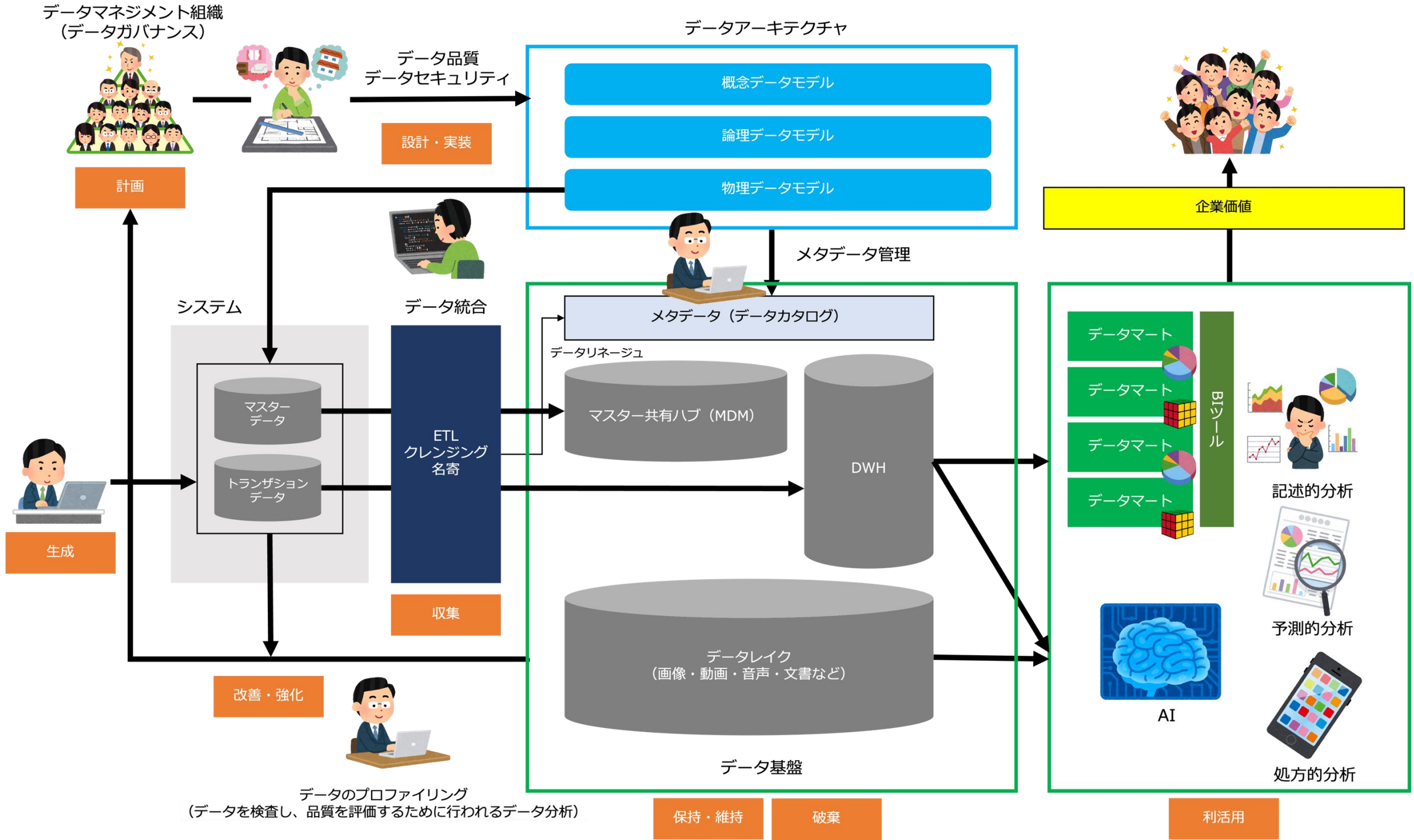
バケツリレーの問題


誤変換
必要な情報の欠落
情報の欠損や誤謬の継承
顧客データ
システムA・部門A
誤変換
必要な情報の欠落
情報の欠損や誤謬の継承
顧客データ
システムB・部門B
顧客データ
システムC・部門C

データ品質確保の重要性

- データクリーニングや、データクレンジングといった言葉を聞いたことがあると思います。データドリブンな意思決定をするには、扱うデータの質が高いことが前提です。データの質が高くないと、「人間なら絶対こんな返答しないよね」という結果を平気で返してくることがあるので、それを踏まえた活用が大事です。
- データの質の担保は、AIによってデータドリブンがしやすくなるだけに、今後さらに重要になるはずです。経営としては、社内でこれらを維持できるよう、努めなければなりません。

出所：DATA is BOSS 収益が上がり続けるデータドリブン経営入門


データマネジメント組織
（データガバナンス）
計画
データ品質
データセキュリティ
設計・実装
データアーキテクチャ
概念データモデル
論理データモデル
物理データモデル
企業価値
メタデータ管理
メタデータ（データカタログ）
データリネージュ
システム
データ統合
マスター
データ
トランザション
データ
ETL
クレンジング
名寄
マスター共有ハブ（MDM）
DWH
データレイク
（画像・動画・音声・文書など）
データ基盤
データマート
データマート
データマート
データマート
BIツール
記述的分析
予測的分析
AI
処方的分析
生成
収集
改善・強化
データのプロファイリング
（データを検査し、品質を評価するために行われるデータ分析）
保持・維持
破棄
利活用

データ利活用に関するさまざまな問題

どこに何のデータがあるかわからない！

業務担当者

データカタログの整備

データ分析のための準備に時間がかかる！

データ品質の管理

あちこちに似たようなデータがあってどのデータを信用していいのかわからない！

データの統合・連携

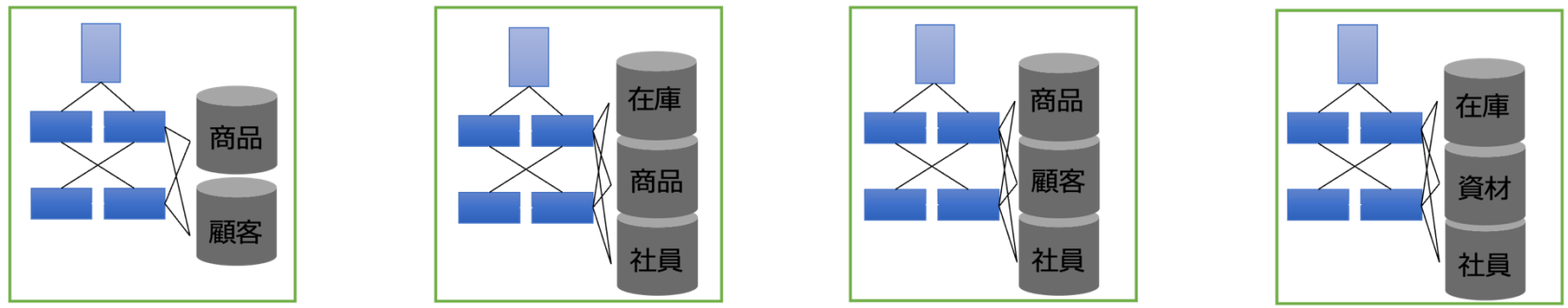
機密情報に簡単にアクセスできてしまう！

データセキュリティの管理

企業の現状

ビジネスの変化の加速＋ビジネスと情報システムの密接化

必要に応じて情報システムを開発



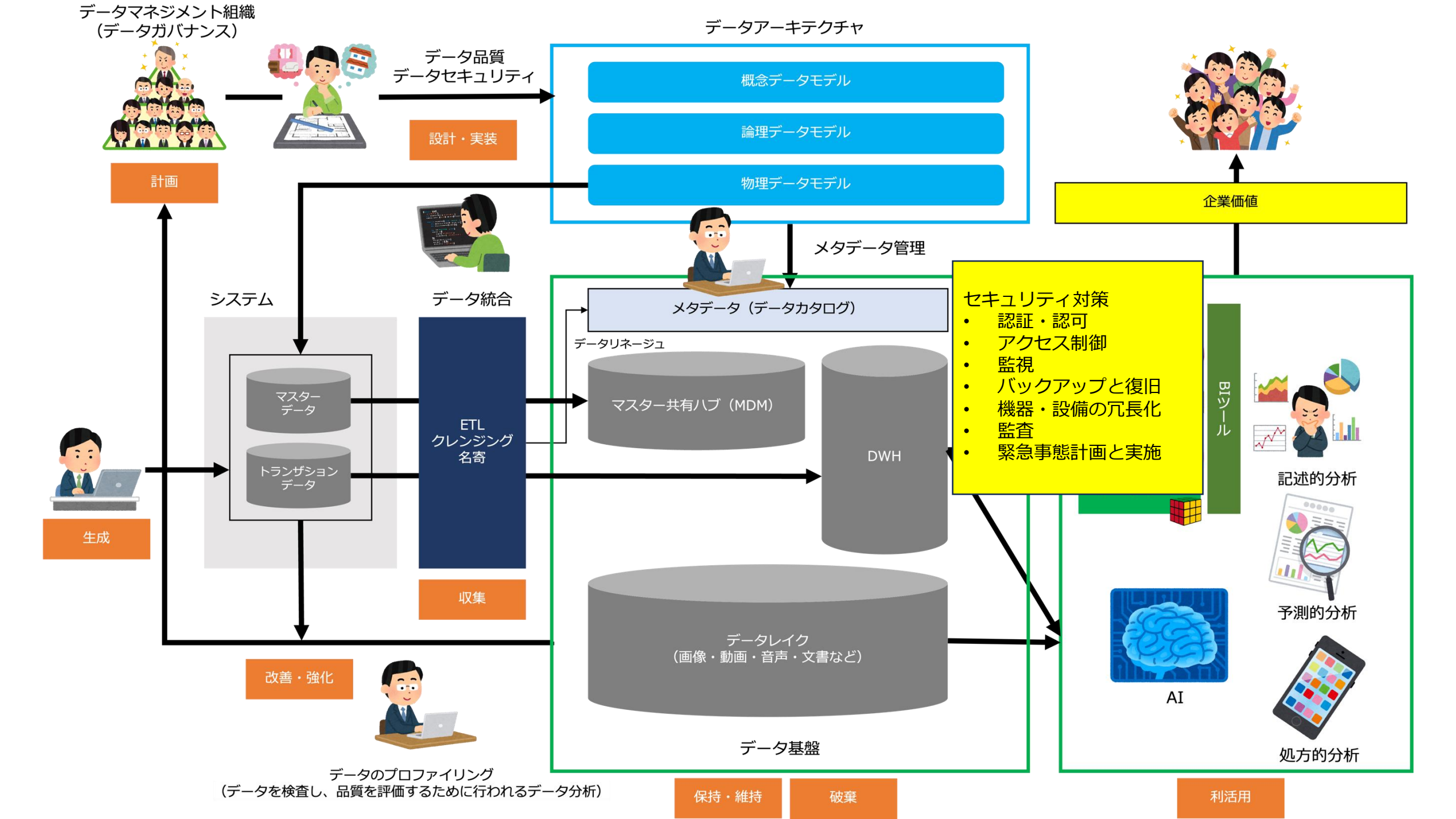
chaos

大規模で複雑なシステム（サイロ）が乱立

重複し整合していないデータの散在

全体を理解できる人や資料がない

雁字搦め


データマネジメント組織
（データガバナンス）
計画
データ品質
データセキュリティ
設計・実装
データアーキテクチャ
概念データモデル
論理データモデル
物理データモデル
企業価値
メタデータ管理
システム
データ統合
メタデータ（データカタログ）
データリネージュ
マスター
データ
トランザション
データ
ETL
クレンジング
名寄
マスター共有ハブ（MDM）
DWH
セキュリティ対策
• 認証・認可
• アクセス制御
• 監視
• バックアップと復旧
• 機器・設備の冗長化
• 監査
• 緊急事態計画と実施
BIツール
記述的分析
予測的分析
処方的分析
AI
生成
収集
データレイク
（画像・動画・音声・文書など）
データ基盤
改善・強化
データのプロファイリング
（データを検査し、品質を評価するために行われるデータ分析）
保持・維持
破棄
利活用

データ利活用に関するさまざまな問題

どこに何のデータがあるかわからない！

業務担当者

データカタログの整備

データ分析のための準備に時間がかかる！

データ品質の管理

あちこちに似たようなデータがあってどのデータを信用していいのかわからない！

データの統合・連携

機密情報に簡単にアクセスできてしまう！

データセキュリティの管理

データを整備するケイパビリティの不足

こうした場合は、そもそもデータを整備するケイパビリティが不足しています。

具体的には、どういったデータやどんな分析ができれば顧客の姿が見えてくるか描いた上で、収集するデータの定義、データインフラの整備、データの更新方法などを交通整理していく必要があります。

それは事業の特性や組織の状況によって変わってくるので、社内外のリソースをうまく使って、継続的にデータを整備していくための体制づくりから、しっかり取り組むのが得策です。

出所：DATA is BOSS 収益が上がり続けるデータドリブン経営入門

データマネジメントの方法

- ビジネスとして経済価値のるデータを特定する方法
- 企業全体のデータ構造やデーフローを可視化し、データの見取り図（データアーキテクチャ）を作成する方法
- データをカタログ化して必要なデータに関する情報（データの有無、所在、構造、来歴、品質の状態、セキュリティレベル、利用するときの連絡先など）をすぐに入手できるようにする方法
- データを統合・連携することでデータの品質やセキュリティを向上させるとともにデータ利活用の効率を上げる方法
- データのライフサイクルを通してデータの品質（一意性、完全性、一貫性など）やセキュリティを組織的に統制する方法

データの準備ステップ

○○のために○○のようなデータ分析がしたい！

必要なデータの特定

データの構造化

データの調査（有無、所在、セキュリティレベルなど）

データの取得

データの理解・クレンジング・加工

データ分析

効率化

データマネジメント

データアーキテクチャの設計

データのカタログ化

データ管理基盤の整備

データ品質の管理

データドリブン経営

Plan

BSCのKPI目標を実現するためのアクションプランのスケジュール設定

Do

アクションプランの遂行

Check

問題の特定	原因の推定
多次元分析によって、どこに問題があるか特定	なぜ問題が生じたのか、問題の原因を推定

Where

記述的分析

Why

予測的分析

解決策が組み込まれたビジネスプロセスをベースにアクションプランを策定

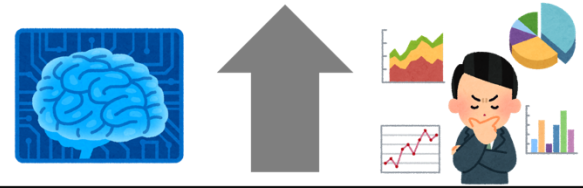
Act

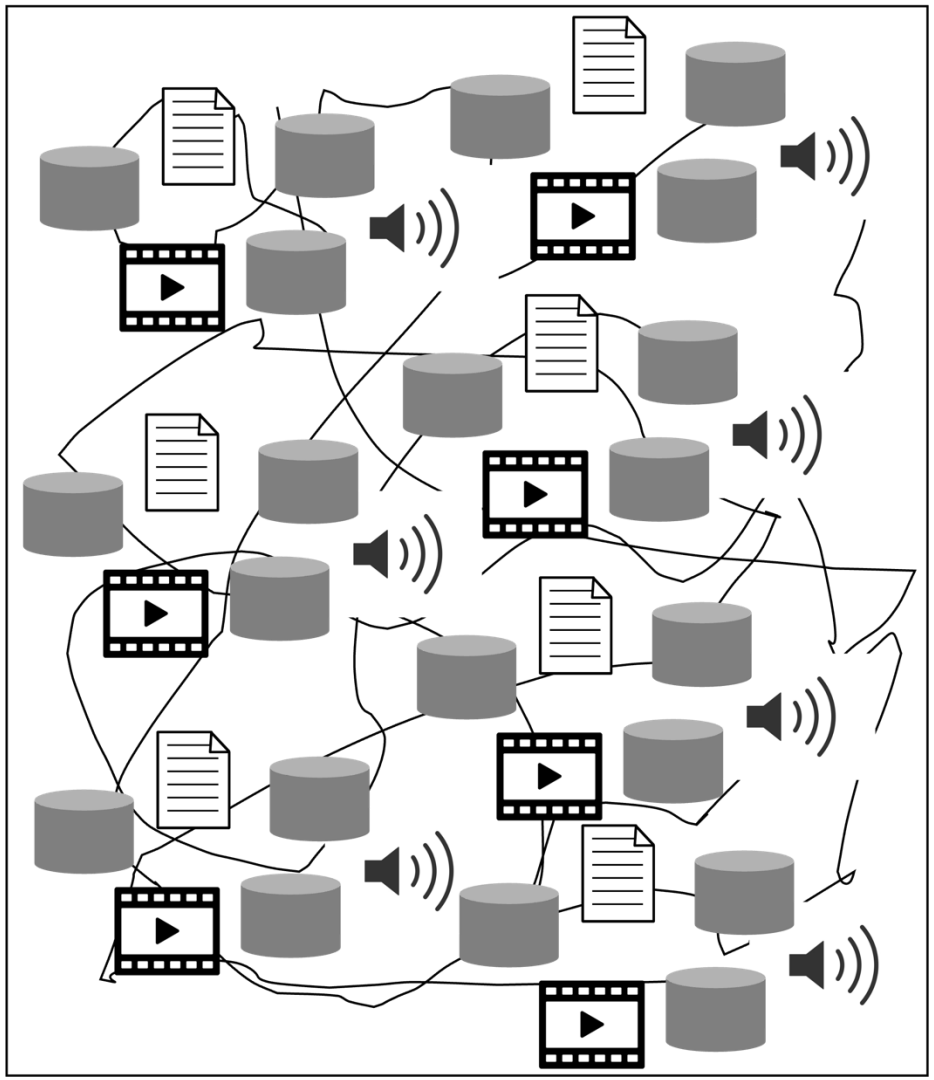
課題の設定	解決策の考案	実験計画の策定	解決策の検証	業務の改革・改善
原因を取り除くことを課題（Issue）として設定	課題に対する解決策（仮説）を考案	課題解策を具体的な実験計画に落とし込む	実験を行い課題と解決策に因果関係があるか検証	因果関係が実証された解決策をビジネスプロセスに組み込む

処方的分析

ランダム化比較実験
（因果推論）

企業価値





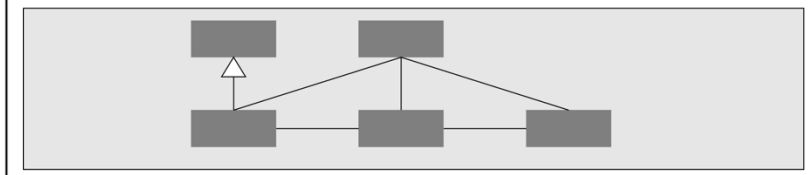
会社に眠るさまざまなデータ

資産化

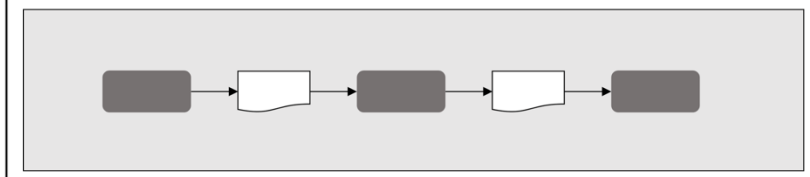
データマネジメント

データアーキテクチャ

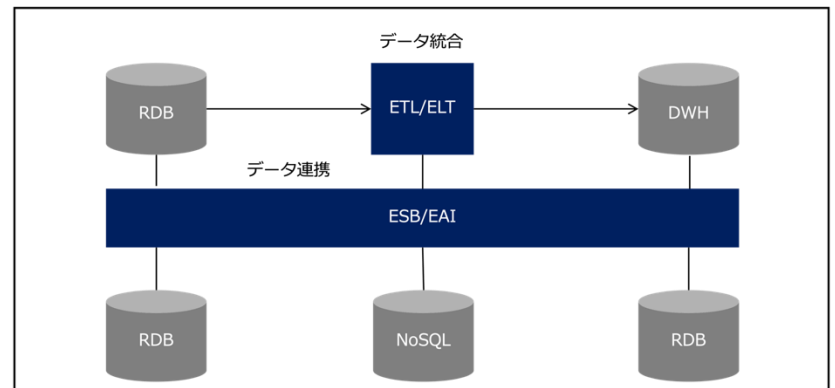
データモデル



データフロー



データ管理基盤



資産としての（稼ぐ力を持つ）データ

データの準備ステップ

○○のために○○のようなデータ分析がしたい！

必要なデータの特定

データの構造化

データの調査（有無、所在、セキュリティレベルなど）

データの取得

データの理解・クレンジング・加工

データ分析

効率化

データマネジメント

データアーキテクチャの設計

データのカタログ化

データ管理基盤の整備

データ品質の管理

データドリブン経営

Plan

BSCのKPI目標を実現するためのアクションプランのスケジュール設定

Do

アクションプランの遂行

Check

問題の特定	原因の推定
多次元分析によって、どこに問題があるか特定	なぜ問題が生じたのか、問題の原因を推定

Where

記述的分析

Why

予測的分析

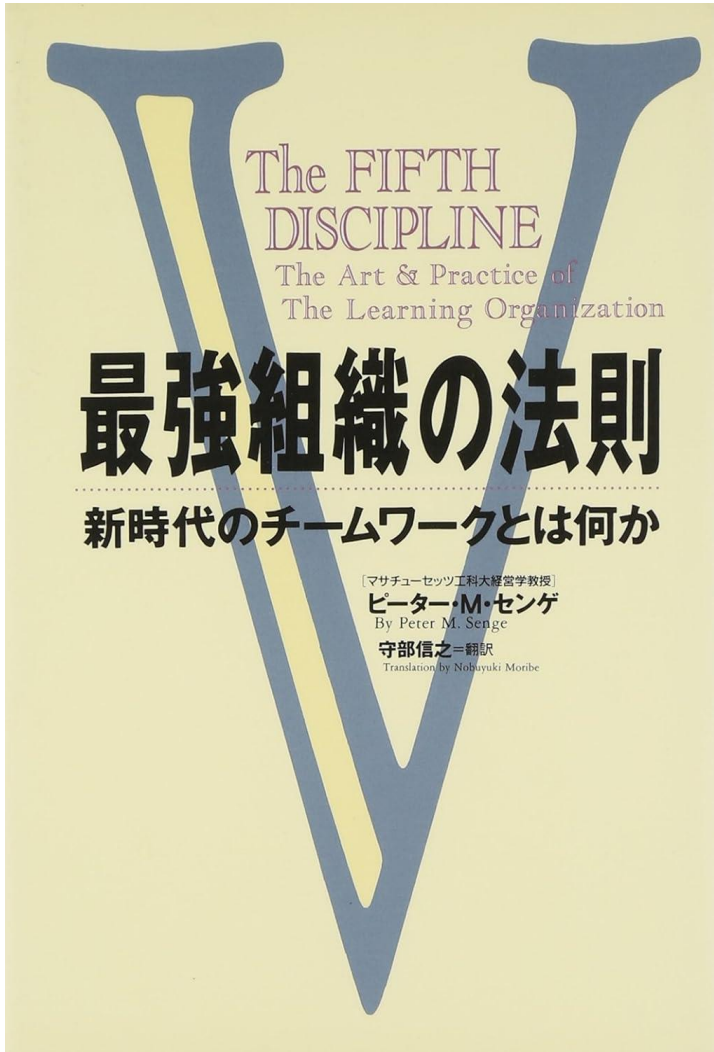
解決策が組み込まれたビジネスプロセスをベースにアクションプランを策定

Act

課題の設定	解決策の考案	実験計画の策定	解決策の検証	業務の改革・改善
原因を取り除くことを課題（Issue）として設定	課題に対する解決策（仮説）を考案	課題解策を具体的な実験計画に落とし込む	実験を行い課題と解決策に因果関係があるか検証	因果関係が実証された解決策をビジネスプロセスに組み込む

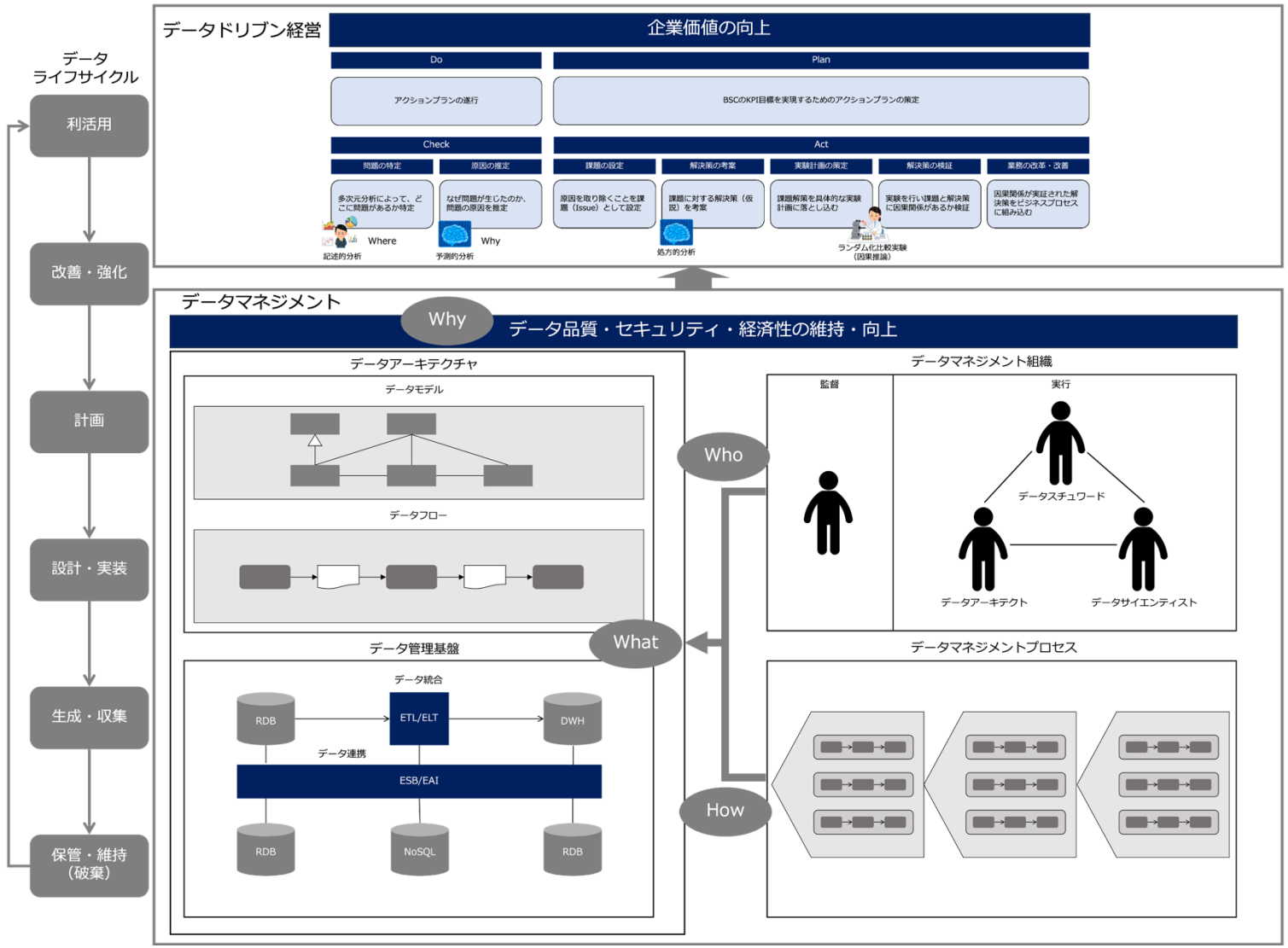
処方的分析

ランダム化比較実験
（因果推論）



競争相手より早く学べる能力、それが競争力を維持する唯一の鍵である

競争相手より早く学べる能力をつけることで企業の競争優位性を上げることができる



データマネジメントの重要性

データマネジメントの目的は、データの品質、セキュリティ、経済性を維持、向上すること

データは利活用されて初めて経済価値を提供する

データマネジメントを導入することで品質とセキュリティの高いデータを「データ利活用」のために
コストをかけずに迅速に準備することができるようになる

データドリブン経営（データを利活用して仮説検証のPDCAを繰り返す経営）の回転速度を上げることで
組織の学習スピードが上がり企業の競争優位性を上げる