

SHIONOGIにおける次世代データマネジメント

塩野義製薬株式会社
DX推進本部 データサイエンス部
雑賀 恵美

初版 2023.03.01
改訂 2023.03.23
改訂 2023.04.20



SHIONOGI

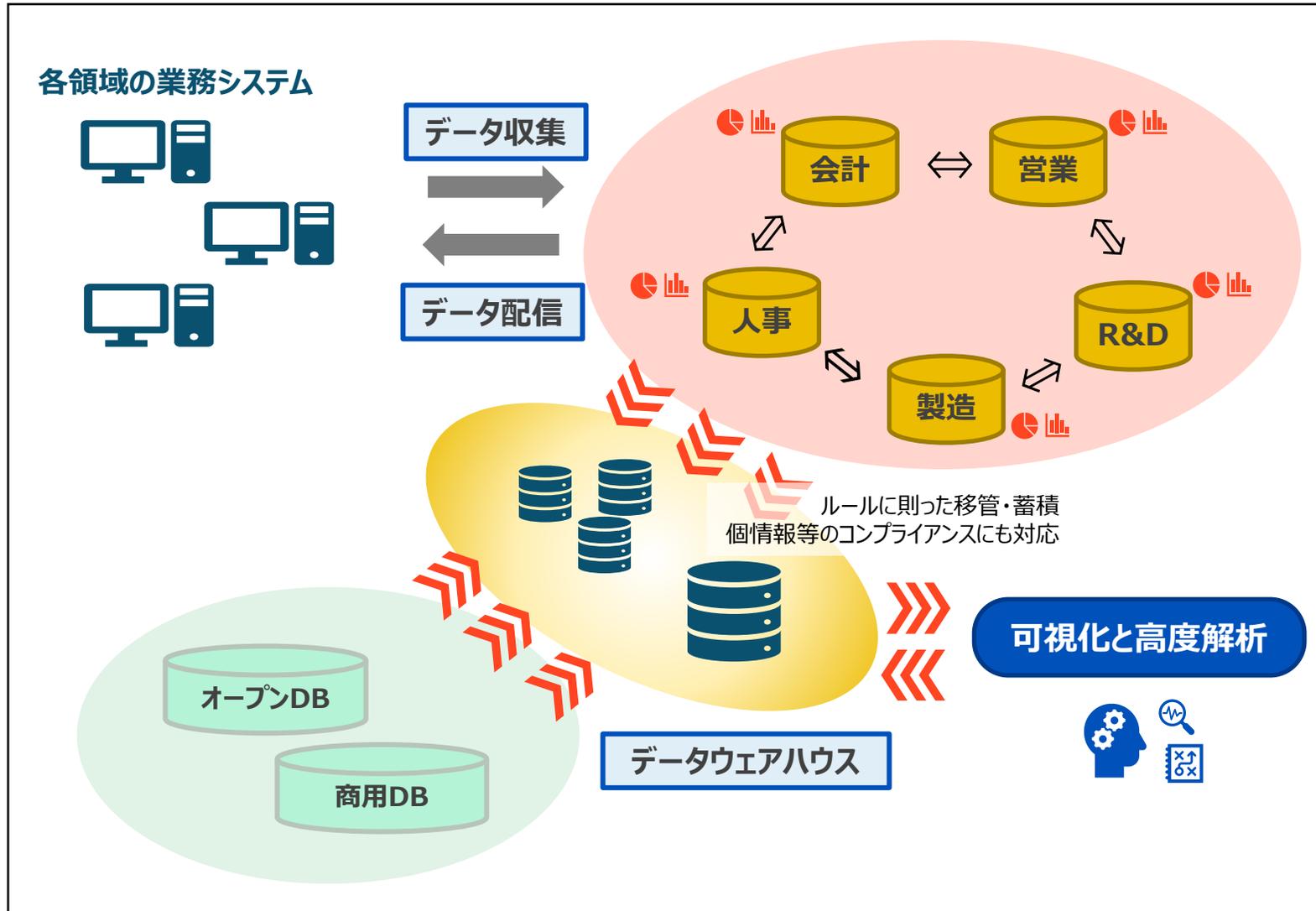
SHIONOGI DATA SCIENCE FES 2023

COI開示

雑賀 恵美

講演内容に関連し、発表者に開示すべきCOI関係にある企業などはありません

Central Data Management 構想



1. 業務システム間のデータ収集・配信のハブ

2. データを活用しやすい形で準備しておくデータウェアハウス

SHIONOGIが目指す次世代データマネジメント

ニーズを先取する、能動的なデータマネジメント

中央集権型を脱却した、分散型のデータコミュニティの形成

なぜ、能動的データマネジメントが必要なのか？

なぜ、能動的データマネジメントが必要なのか？

従来

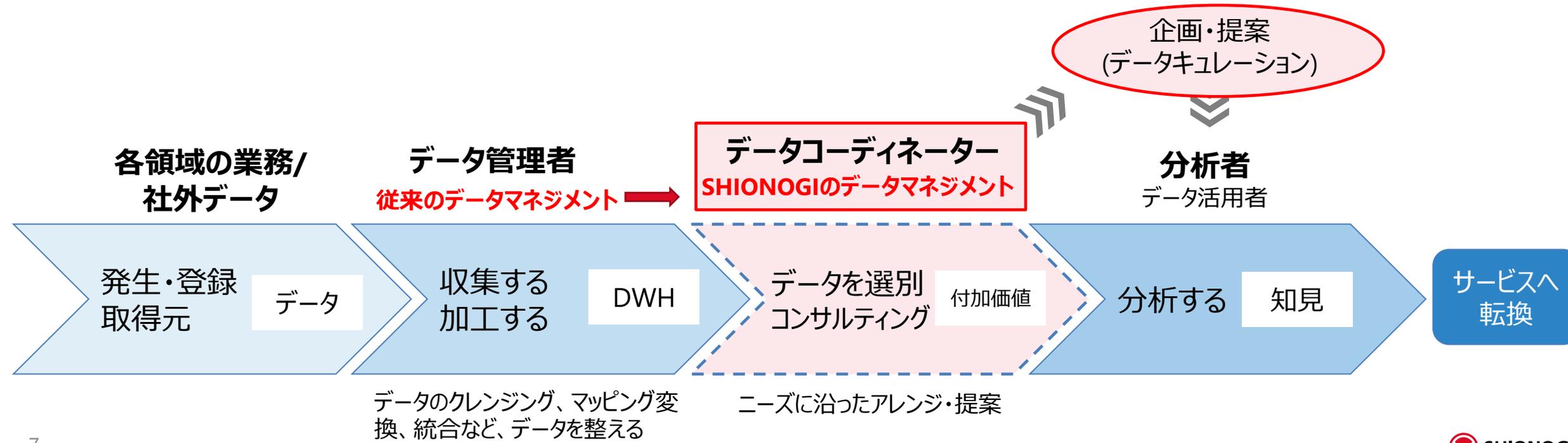
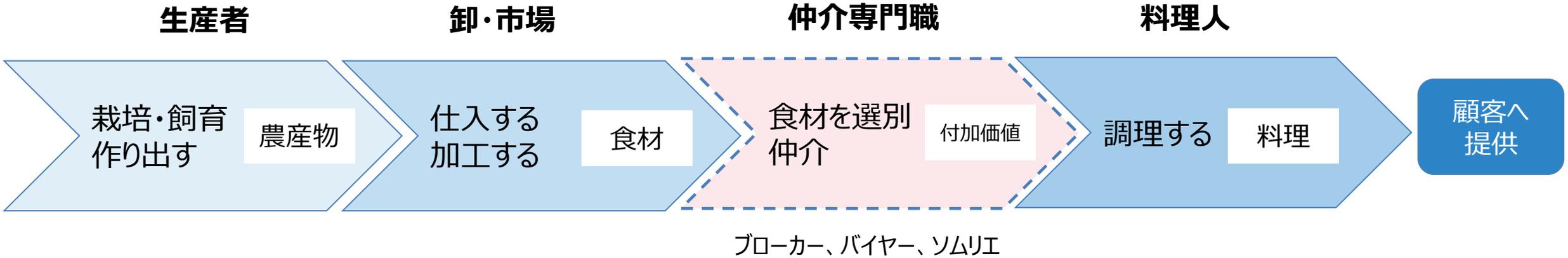
ビジネスサイドの業務目的が起点となって、データの登録が始まる

ビジネスサイドは、データの二次利用の有用性に
気付いていない・気付けない

SHIONOGI

データマネジメントサイドから、データ活用シーンと要件を提案する

データが価値を得るまで1



データ活用者の課題を解決する

直面する課題

- 新鮮で美味しい食材を入手したい
- この料理にはどの食材が適しているだろうか？調理方法も食材によって違うな
- 新しいメニューも考えないと・・・

料理人



- 最新の正確なデータを入手したい
- この仮説を立証するには、どのデータを利用すればよieldだろうか？
- 次は、どのようなテーマの分析をすればよieldだろうか・・・

分析者



これらの課題を解決できるデータマネジメントとは

データが集まる・正しく使える仕掛けを作る

- 鮮度、正確性が高いデータを常に利用できる状態に維持
 - 法令遵守、個人情報保護
- ⇒ タイムリーで信頼できるデータ。安心・期待を持ってデータを預けられる。

目的にマッチしたデータニーズを予測する

- 周辺データとの関連性などを考慮し、最も適したデータと方法をレコメンド
- ⇒ その為にデータを熟知し、データソムリエの役割を担う。

ビジネスサイドにデータから得られる仮説を提案する

- ビジネスサイドも気づいていないビジネス課題の探索
- ⇒ 場合によっては自ら実証する。

なぜ、分散型データコミュニティが必要なのか？

なぜ、分散型データコミュニティが必要なのか？

中央集権型

データマネージャのみがデータ管理者として、データウェアハウスを運営する



ビジネスサイドが一次目的に利用しやすいデータと、分析者が利用しやすいデータは異なる
中央管理のデータマネージャは、データハブとして前者の視点を取り入れているが、
データウェアハウスとしては、分析者の視点も取り入れる必要がある



分散型

データ活用者もデータ管理者として、データウェアハウスの運営に参加する

立場による興味の違い

生産者

- 種、苗、肥料の種類
- 植え付け、収穫の時期
- 土壌、気候の条件

市場

- 流通量、市場の価値
- 在庫

料理人

- 栄養、旨味、新鮮さ
- 産地
- 農薬有無

ビジネスサイド（一次目的）

- システム連携
- 自動化、効率化

データ管理者（データハブの視点）

- 正規化したデータ管理
- データ型、データサイズ、チューニング
- データ連携ルール、統制

分析者（データ活用者）

- 正規化されていないデータマート
- データ項目の意味・説明、関連データ
- 取得期間、鮮度

これらの興味（メタデータの提供）に素早く的確に応えることが重要
データ活用者自身がデータ管理者になることで解決

データ活用者がデータマネジメントに参加することで解決を図る

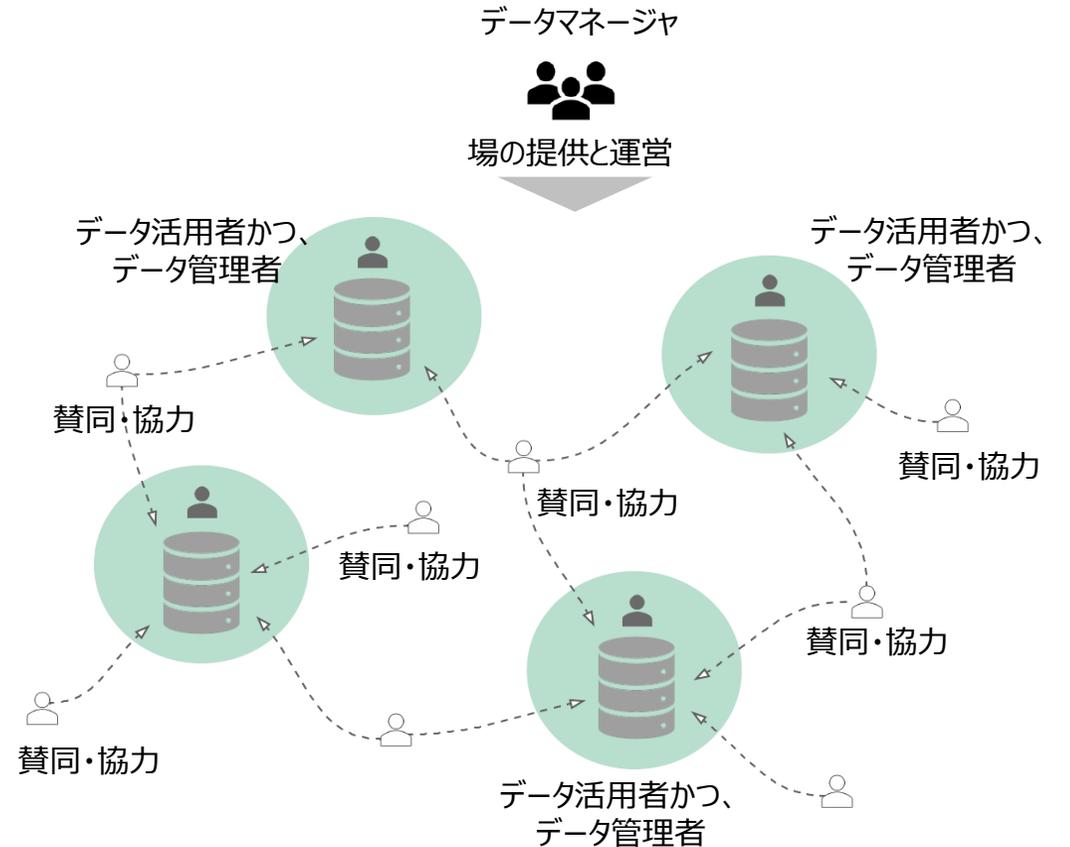
データ活用者自身が、DWHへ必要とするデータを
持ち込み、データ管理者となる

- 賛同者を募ることで、DWHへ持ち込める（参加者による意思決定）
 - 自ら必要なメタデータ（データの情報）を追加し共有できる
- ⇒重複タスクの防止。公平性、透明性が高い。

データマネージャは、分散型データコミュニティの場を
提供し自律性を引き出すことに注力

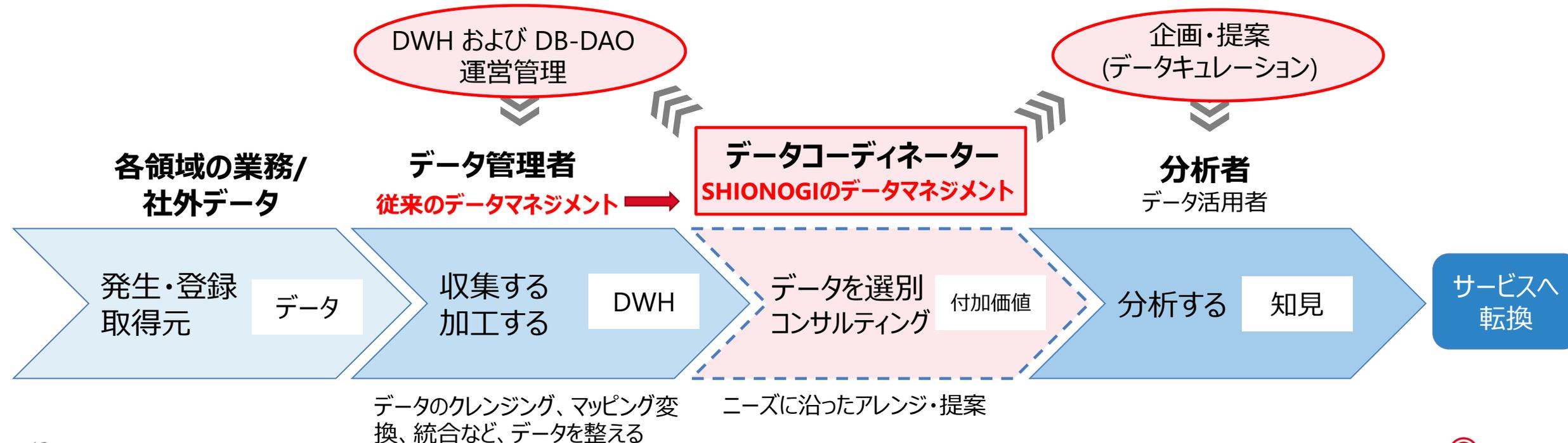
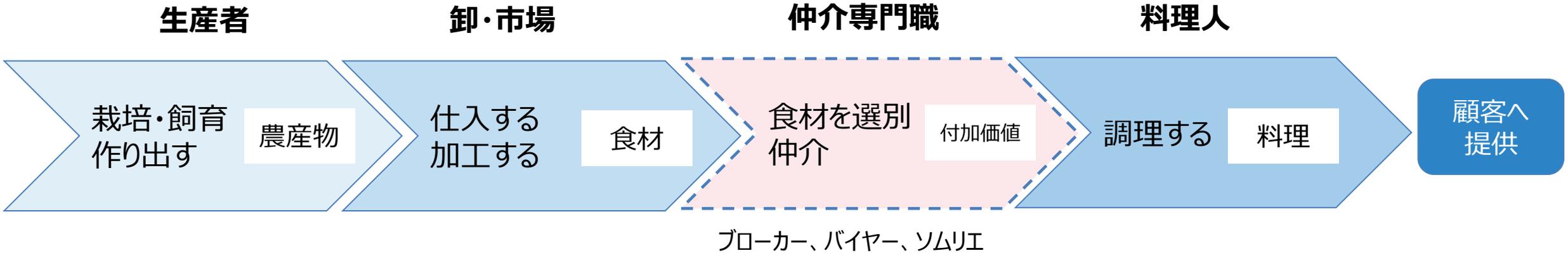
- 権限が集中しない自由な発想
- ⇒コラボレーション、クリエイティビティ、イノベーションに繋がる。

分散型データコミュニティ（DB-DAO*）

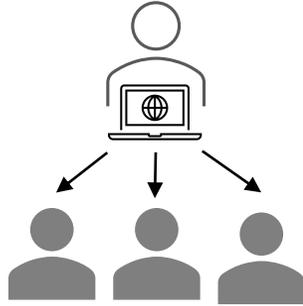


*DAO : Decentralized Autonomous Organization（分散型自律組織）

データが価値を得るまで2



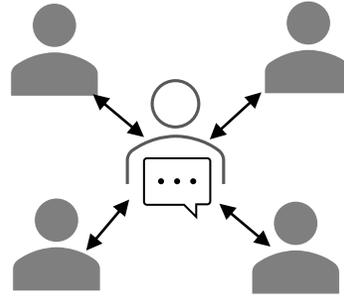
参考：インターネットにみる概念の変遷



WEB1.0

- 一方通行の情報発信
- 情報発信者が限定的

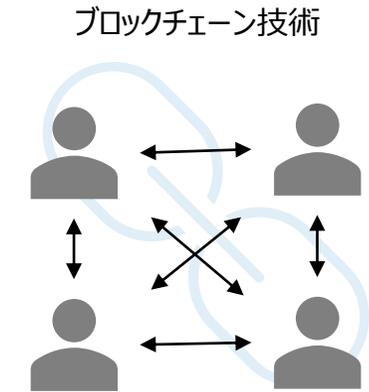
テキストや静止画が主なコンテンツである
WEBサイト



WEB2.0

- 双方向のやり取り
- 中央集権型：データは運営主体が管理
- 管理者によるコントロールやマージンの徴収が発生

Facebook、YouTube、Instagram、
Twitter、食べログ、Uber Eats

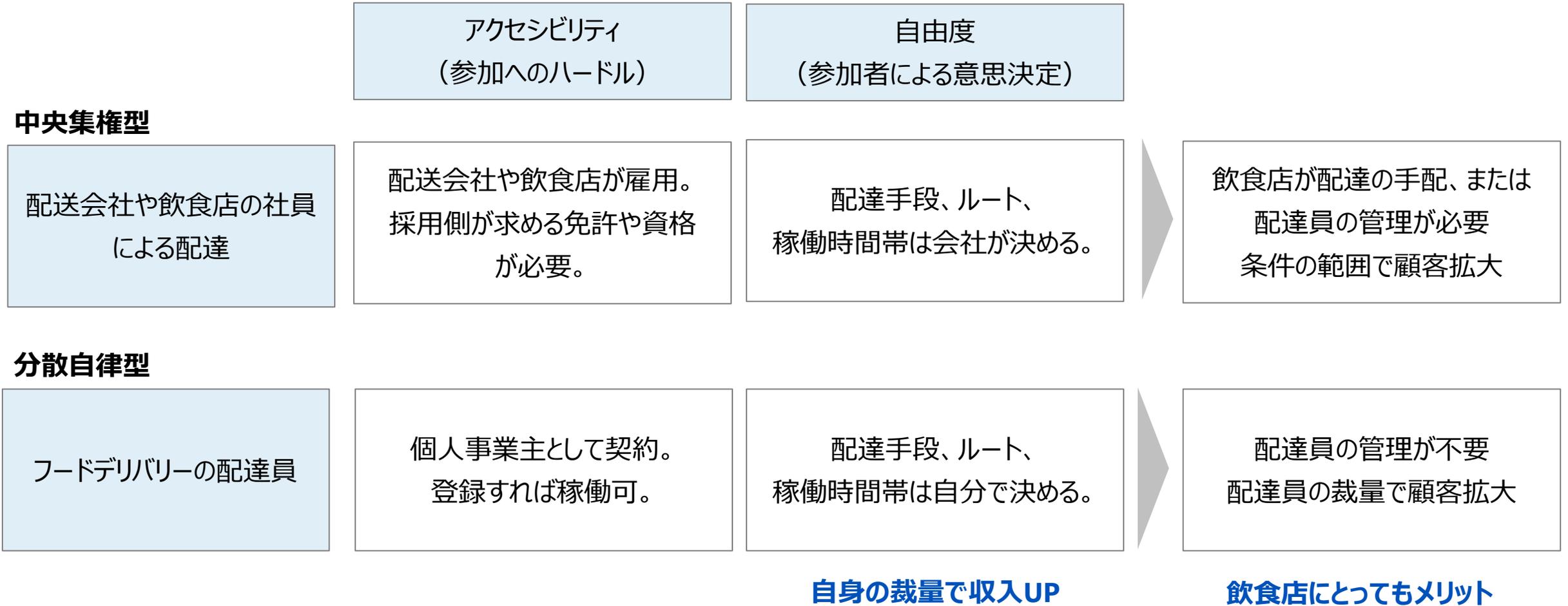


WEB3.0

- ユーザ同士が直接やり取り
- 分散自律型：データはユーザが管理
- 特定の管理者を持たない

DeFi（分散型金融）
NFT（非代替性トークン）
DAO（分散型自律組織）
Mastodon

参考：インターネットサービス以外の事例（フードデリバリーの配達員）



SHIONOGI データ活用業務イメージ

データ活用 – 業務イメージ

データベースの価値の向上

探しやすい

分かりやすい

使いやすい

利用目的と成果の透明化 (ガバナンス体制強化)

興味本位ではなく成果に責任を持つ

データサイエンスの生産性（スピードと質）の向上

データ活用による成果創出だけに留まらず、その**成果の二次利用**により生産性を向上

同じような検討をしない・させない

先人の失敗から学ぶ

既存の成果をさらに深掘りする

ご清聴、ありがとうございました。