



DataSpider × HULFT DataCatalog

シナジーで切り拓く新たなビジネスの可能性

ご利用中のDataSpiderをさらに活用する方法があります



データの活用・分析をスムーズに

今こんな課題を抱えていませんか？

DWH構築支援

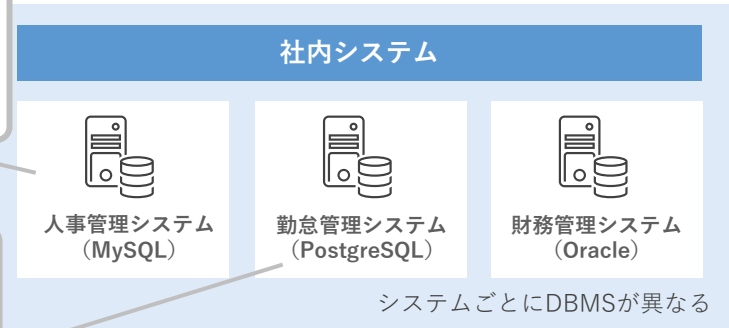
悩み

社内システムが複数のDBMSで構築されているため、DWHの構築に時間がかかる

システムごとに
項目名がバラバラ

社員ID	氏名
EMP0001	山田太郎
EMP0002	山田花子
EMP0003	鈴木一郎

社員番号	氏名
EMP0001	山田太郎
EMP0002	山田花子
EMP0003	鈴木一郎



現状の問題

- 同じ意味のデータでもシステムごとに表現がバラバラ
- 全体把握には各システム担当者への個別ヒアリングが必要

解決策

- DataSpiderで複数システムのデータ項目を紐づけ
- HULFT Datacatalogを参照して、**DWHに必要なキー項目を把握**し、スクリプトを実装

効果

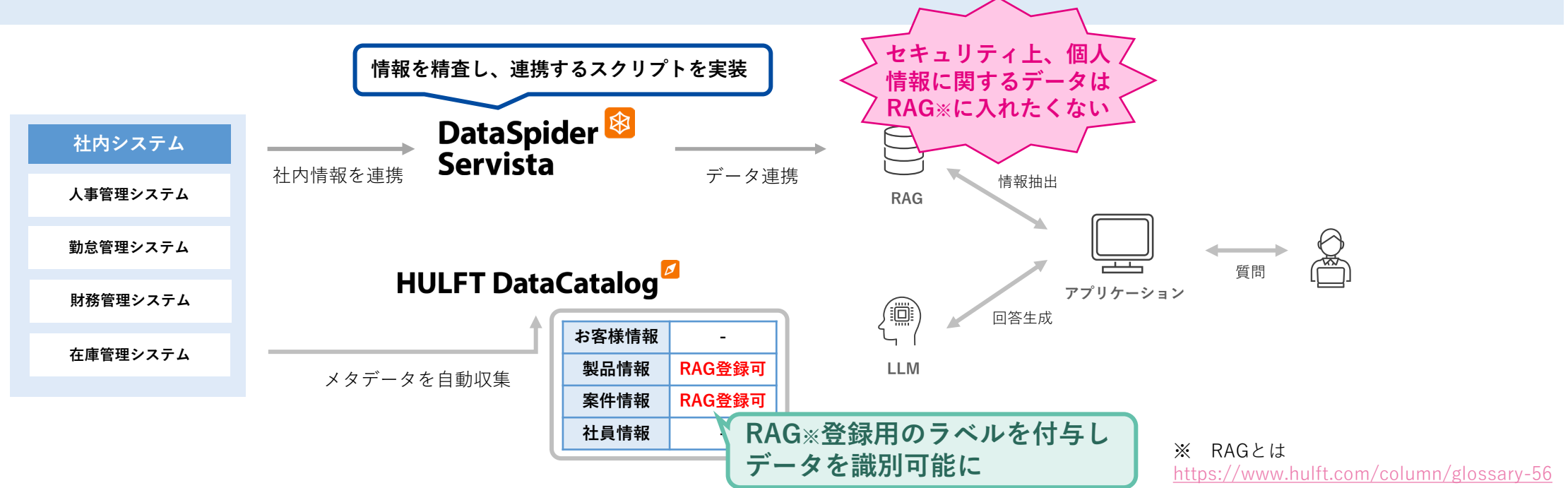
- システム担当者へのヒアリング時間の削減され、**DWH構築が迅速化**
- システム拡張にも柔軟に対応

データ品質の向上とDWH構築時間の短縮を実現

RAG※データの精査

悩み

セキュリティ規約により、RAG※に連携できるデータとできないデータを精査する必要がある



現状の問題

- RAGに連携するデータを精査するために、社内の担当者にヒアリングが必要
- 精査した内容を一元管理する場所がない

解決策

- 各データにメタデータとして、「**RAG登録可**」のラベルを付与
- メタデータをもとにデータ検索、登録可否の識別

効果

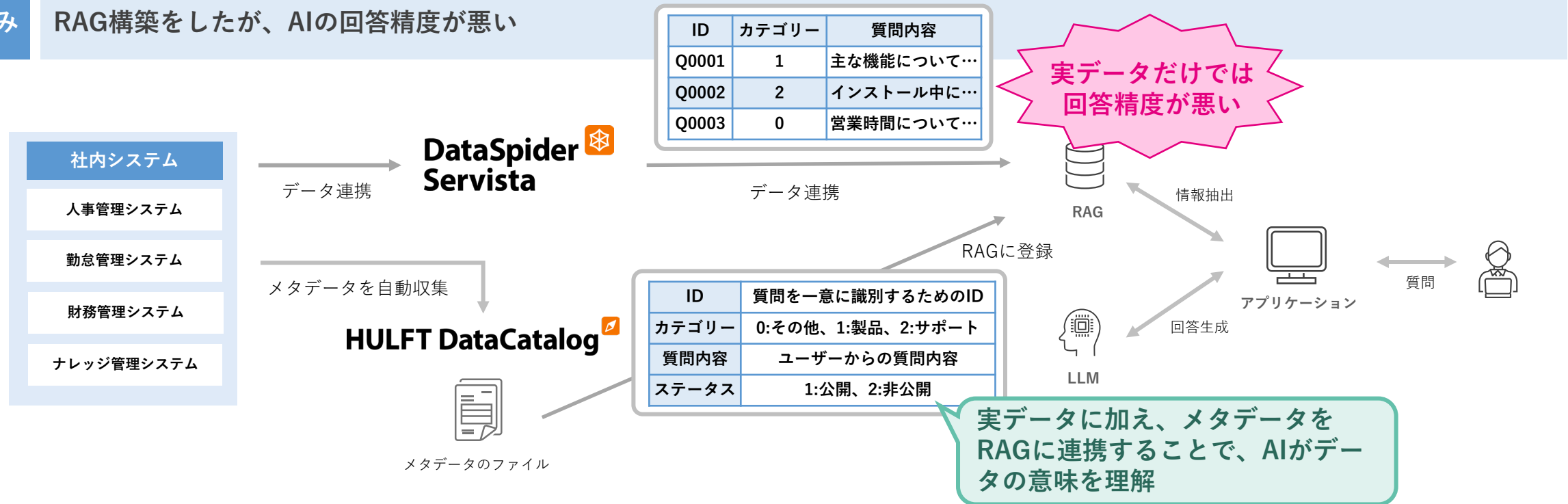
- データ識別作業が効率化され、人的コストを削減
- 社内のセキュリティ上、**RAG※に取り込めないデータも精査可能**

データを精査し、セキュリティ強化

AIの回答精度改善

悩み

RAG構築をしたが、AIの回答精度が悪い



現状の問題

- 項目名などのカラム情報と、実データでRAGを構築
- AIが「データの意味」を理解していないため、回答精度が低い

解決策

- 実データだけでなく、メタデータ（**データ意味やデータ項目の意味、タグでグループ化された情報**）をRAGに取り込む

効果

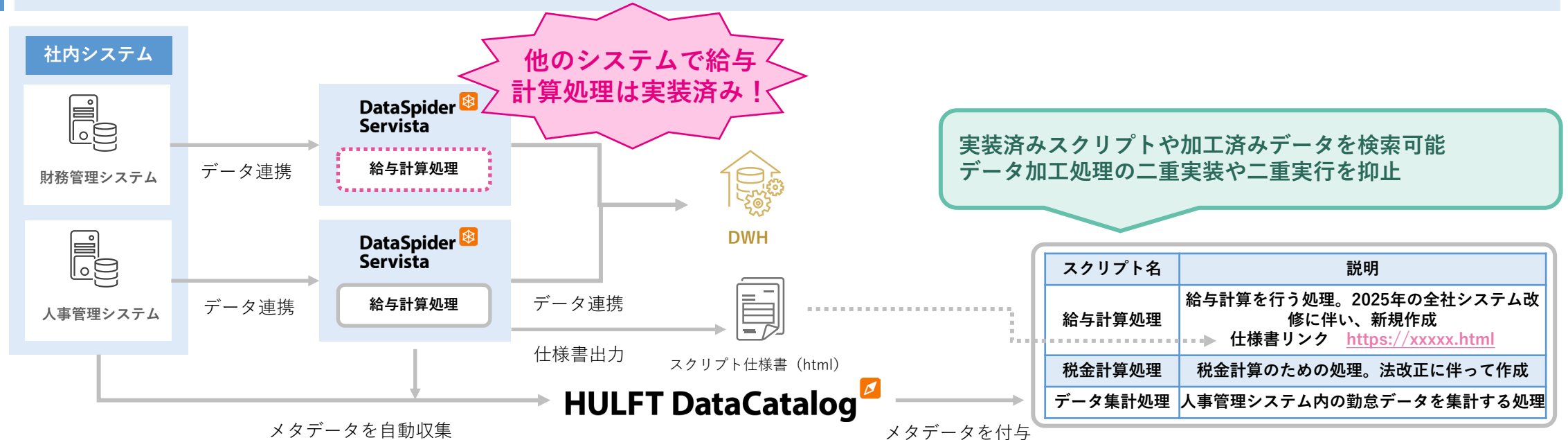
- 学習データが精査され、**AIの回答精度が向上**
- 「データの意味」は人だけでなく、AIにとっても有用な情報

メタデータを学習することで、AIの回答精度が飛躍的に向上

資産の有効活用

悩み

複数システムで類似のデータ加工処理が存在し、不必要な開発工数が発生



現状の問題

- 重複したプログラムの管理コストが発生
- どの処理がどのシステムで行われているかが把握できず、障害発生時に対応が遅れる

解決策

- DataSpiderと接続し、**スクリプトの一覧を取得**
- **データの来歴を参照**し、スクリプトの概要を把握
- 給与計算処理のスクリプト仕様書 (html) を生成し、メタデータにリンクを登録

効果

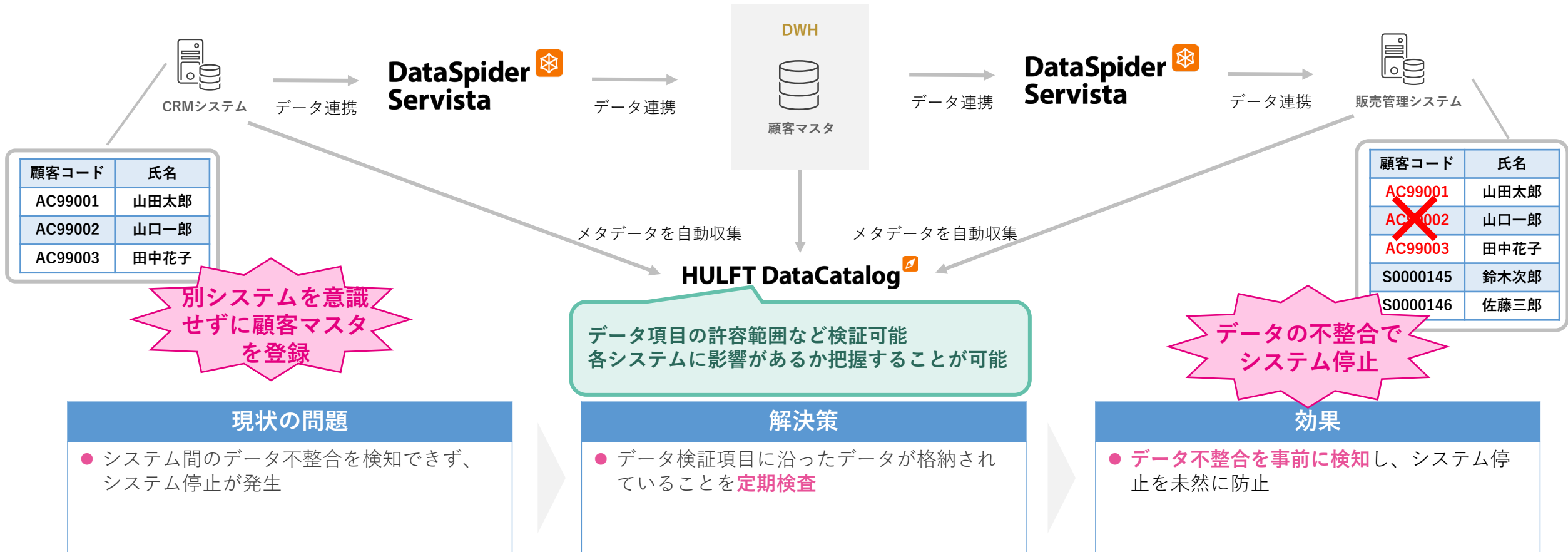
- 既存の**スクリプトを検索**でき、不要なプログラム作成工数の削減
- 処理の共通化による障害対応の迅速化
- **データ加工処理や加工済みデータを有効活用**

データの再利用で開発コスト削減

データの定期検査

悩み

他システムに影響を与えるデータの不整合を検知できない

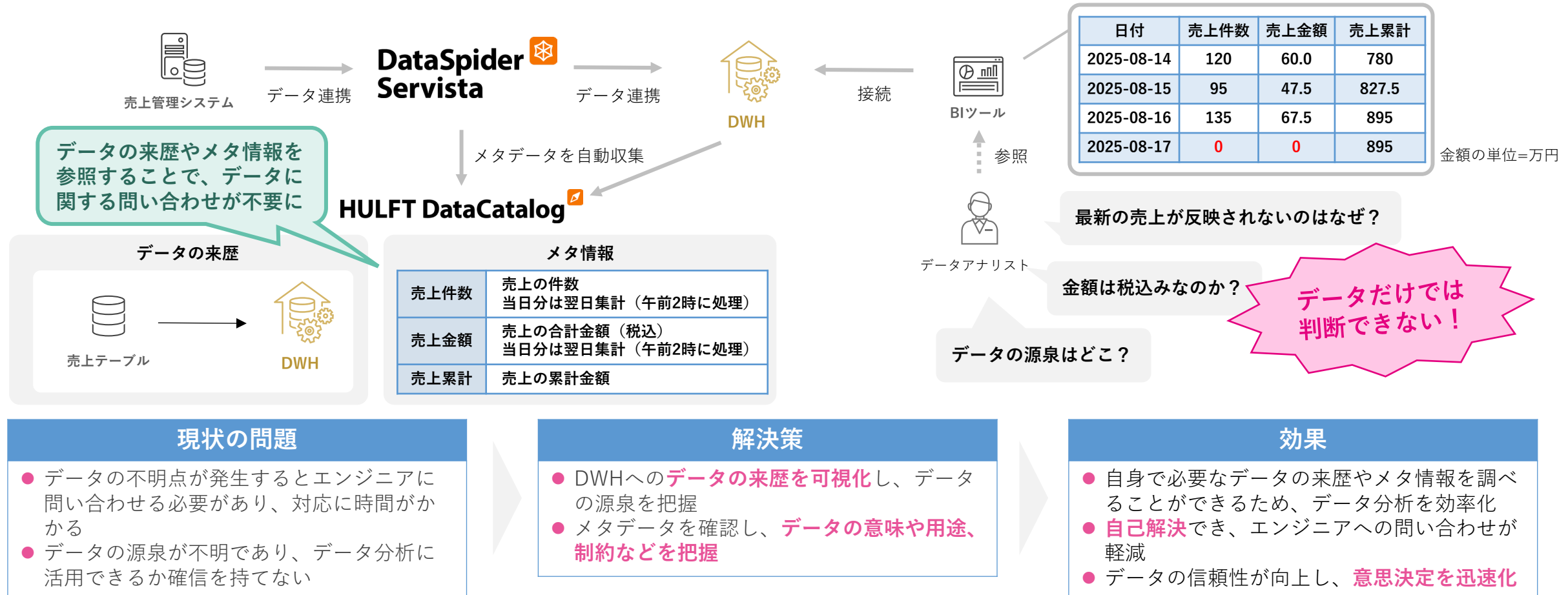


定期検査で安定したデータ連携を実現

データドリブンな意思決定

悩み

登録されているデータの説明や来歴がわからないため適切な意思決定ができない

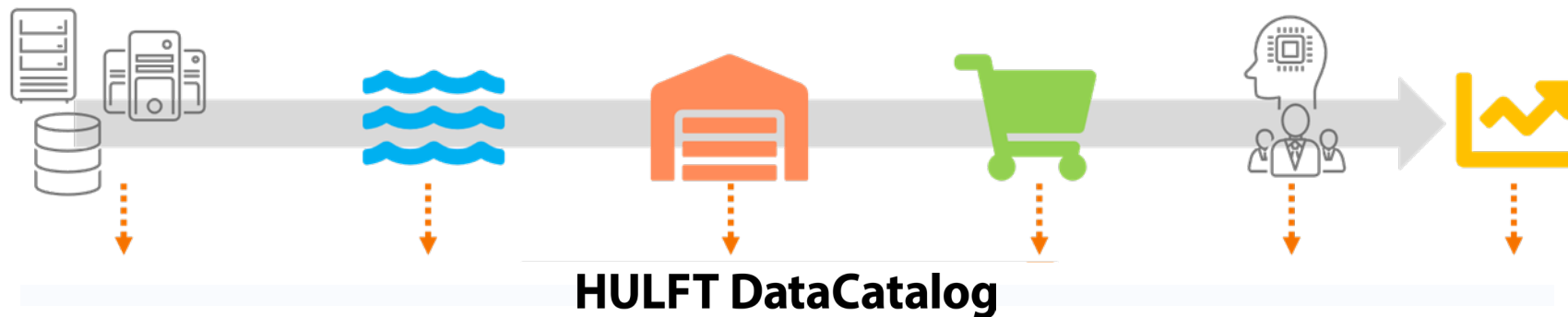


データの透明性を確保し、迅速な意思決定を実現

HULFT DataCatalogとは？

HULFT DataCatalogとは

散在したメタデータを収集・整理・カタログ化
データからひらめきを生み出すメタデータマネジメントプラットフォーム



メタデータの
自動収集管理



ビジネス用語で
探索発見理解



データの
信頼性を管理



セキュリティと
データ利活用の両立

期待される効果

データドリブン・デジタル経営へ変える。維持する。発展させる。

問題点	どこにデータがあるか 分からない、見つからない	信じられるデータなのか分からない	社内データ利活用状況が 分からない、見えない
アプローチ	データ所在の可視化 どこにデータが データ来歴の可視化 どこからどう変化した	用語／タグ どんなデータなのか データナレッジ共有 皆が平等にデータを	アクティビティレポート 誰がいつどんなデータを
短期的効果	✓ システム横断のデータ所在の可視化 ✓ 問い合わせする、対応する間接時間の削減	✓ 真のデータの意味を可視化 ✓ データに対する属人性を排除	✓ 誰が・いつ・どんなデータを参照している か可視化 ✓ データ人財の発掘とロールモデル人財展開 を支援
長期的効果	✓ まずはデータを探索する。社内の意識や行動に変革をもたらします。 ✓ データに触れる楽しさを醸成します。 ✓ データに基づく意思決定を加速させます。		

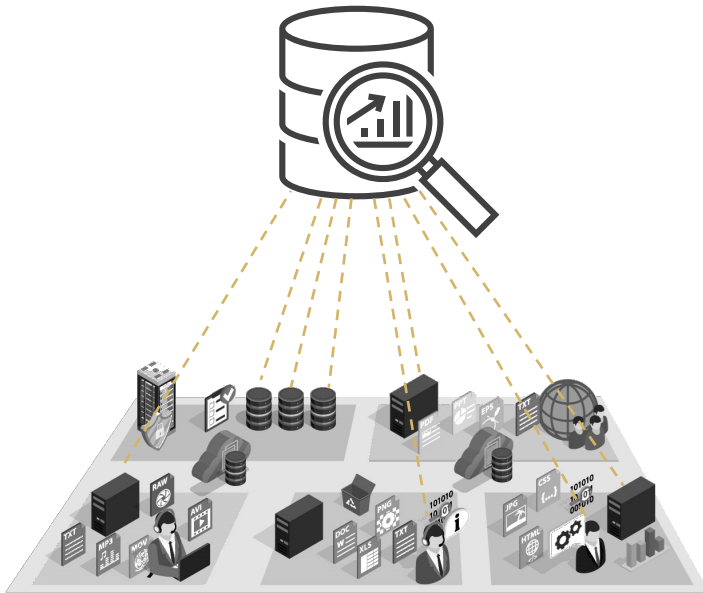
今回紹介したHULFT DataCatalogの機能

メタデータの検索と可視化

- カタログ化されたメタデータの検索と可視化ができます。
キーワードを基に横断的に探索ができ、どこにデータがあるのか一覧として可視化する事が可能です。

HULFT DataCatalog

メタデータリポジトリ



assets

localhost:8080/assets?q=有休&options=

HULFT DataCatalog

有休

検索結果: 5 件

絞り込み条件	検索結果: 5 件																		
アセット種別 ☐ Table (2) ☐ Script (2) ☐ Project (1)	<table border="1"><thead><tr><th>アセット名</th><th>接続先種別</th><th>説明</th></tr></thead><tbody><tr><td>有休未取得者マスタ HRDB > public > 有休未取得者マスタ</td><td>PostgreSQL</td><td>有休未取得者の一覧を記載しています。毎月25日に</td></tr><tr><td>有休テーブル HRDB > public > 有休テーブル</td><td>PostgreSQL</td><td>社員Noを主キーに有休日数に関する情報を記載する</td></tr><tr><td>有休未取得者 DataSpider4.4 > root@有休未取得者 > 有休未取得者</td><td>DataSpider</td><td></td></tr><tr><td>root@有休未取得者 DataSpider4.4 > root@有休未取得者</td><td>DataSpider</td><td></td></tr><tr><td>有休未取得者マスタデータ削除 DataSpider4.4 > root@有休未取得者 > 有休未取得者マスタデータ削除</td><td>DataSpider</td><td></td></tr></tbody></table>	アセット名	接続先種別	説明	有休未取得者マスタ HRDB > public > 有休未取得者マスタ	PostgreSQL	有休未取得者の一覧を記載しています。毎月25日に	有休テーブル HRDB > public > 有休テーブル	PostgreSQL	社員Noを主キーに有休日数に関する情報を記載する	有休未取得者 DataSpider4.4 > root@有休未取得者 > 有休未取得者	DataSpider		root@有休未取得者 DataSpider4.4 > root@有休未取得者	DataSpider		有休未取得者マスタデータ削除 DataSpider4.4 > root@有休未取得者 > 有休未取得者マスタデータ削除	DataSpider	
アセット名	接続先種別	説明																	
有休未取得者マスタ HRDB > public > 有休未取得者マスタ	PostgreSQL	有休未取得者の一覧を記載しています。毎月25日に																	
有休テーブル HRDB > public > 有休テーブル	PostgreSQL	社員Noを主キーに有休日数に関する情報を記載する																	
有休未取得者 DataSpider4.4 > root@有休未取得者 > 有休未取得者	DataSpider																		
root@有休未取得者 DataSpider4.4 > root@有休未取得者	DataSpider																		
有休未取得者マスタデータ削除 DataSpider4.4 > root@有休未取得者 > 有休未取得者マスタデータ削除	DataSpider																		
接続先種別 ☐ DataSpider (3) ☐ PostgreSQL (2)																			
接続先名 ☐ DataSpider4.4 (3) ☐ PostgreSQL 9.6_HRDB (2)																			
タグ																			
用語																			

HRDB > public > 有休未取得者マスタ

有休未取得者マスタ

説明

有休未取得者の一覧を記載しています。毎月25日に自動で更新されます。

URL

jdbc:postgresql://localhost:5432/HRDB

オーナー

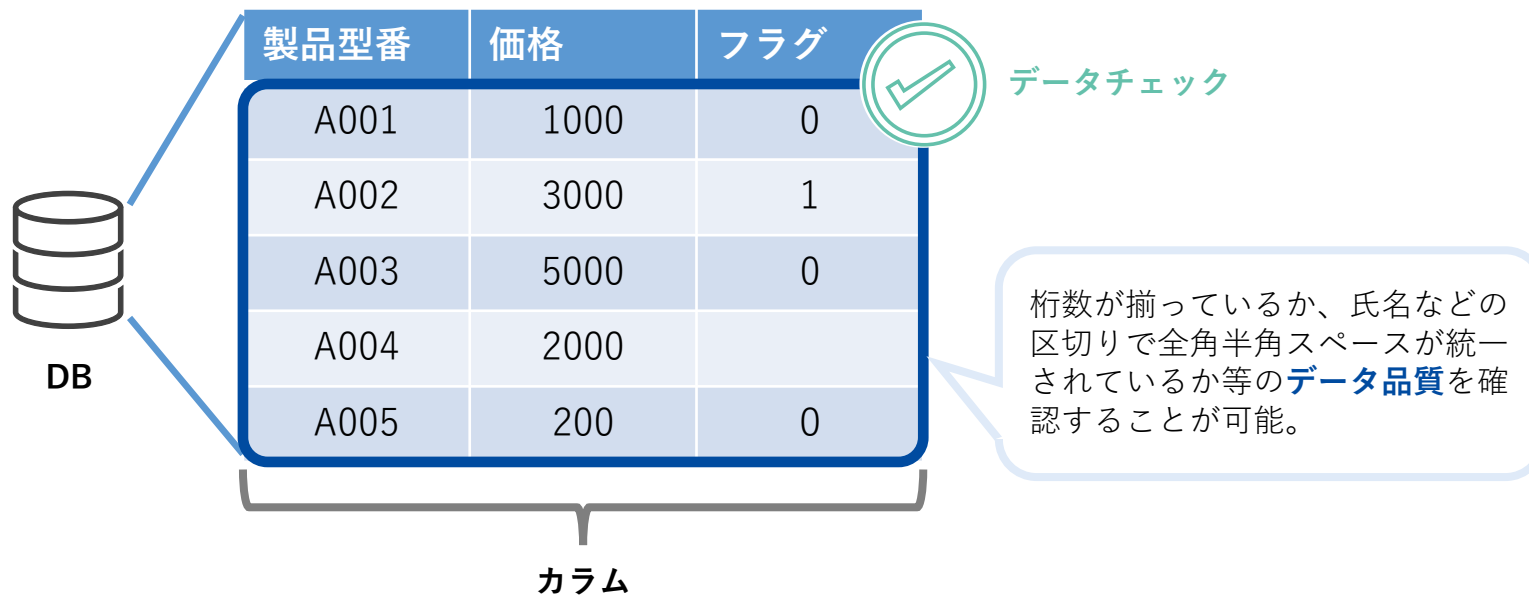
レコード件数

4 件

データテスト

●データの健康状態を検査

データがデータ運用者にとって意図した範囲や書式で格納されているかをテストします。
これによりデータの健康状態を評価し、いち早く必要な対処を行うことができます。



チェック項目

項目	内容
isComplete()	Nullの値があるかどうか
isUnique()	重複している値があるかどうか
hasMin(num)	指定した数値より下の数値はあるか
hasMax(num)	指定した数値より上の数値はあるか
hasPattern(string)	指定した正規表現にあったデータであるか

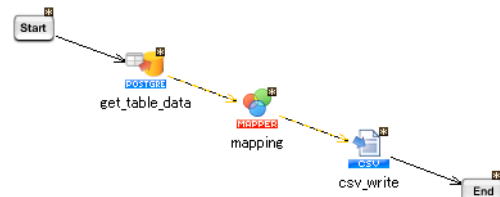
データリネージュ(来歴)

- DataSpider Servistaにて連携されたデータの来歴を可視化する事ができます。
「何が」「どこへ」「どのような形で」を可視化する事で目的のデータか確認します。

DataSpider Servista

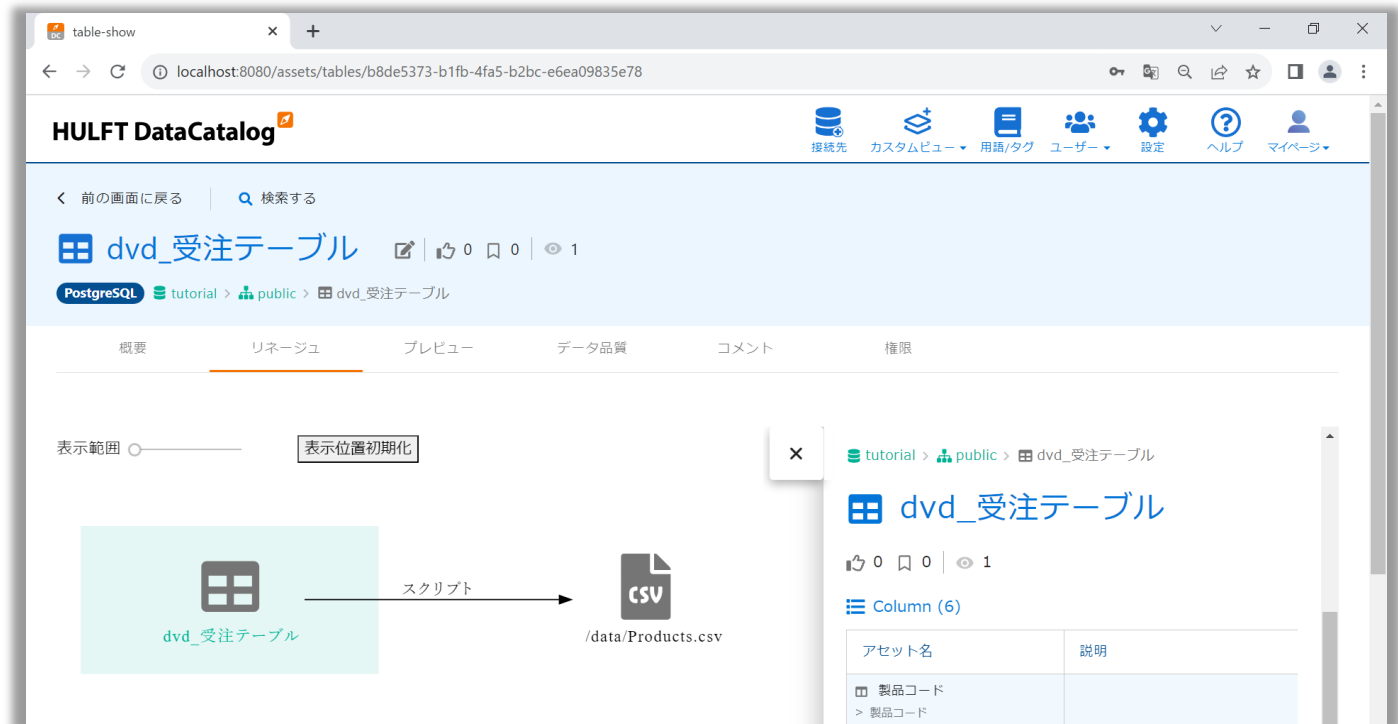
データ連携スクリプト

データ連携スクリプト



データ連携処理を収集解析

HULFT DataCatalog



Appendix

データのサイロ化によって様々な問題が発生

大企業での「データのサイロ化」に対して、データ連携の重要性が高まってきている

原因

組織内の部門間でのコミュニケーション不足や特有の企業文化、技術的な制約など

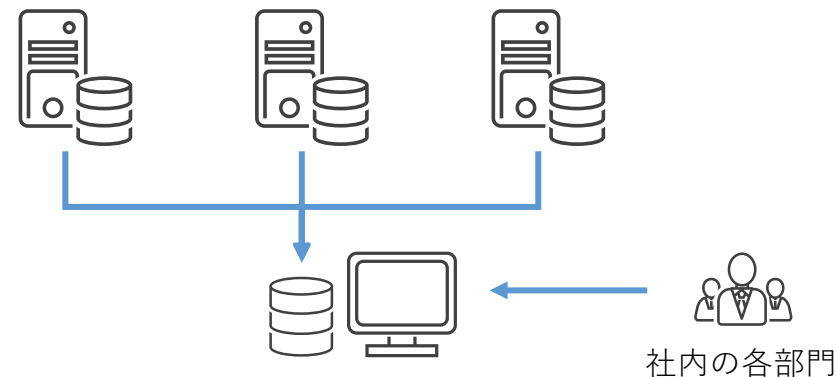
部門間のデータ共有に障壁がある状態（サイロ化）



問題

- ✓ 情報共有の障害
- ✓ 重複作業が発生
- ✓ 非効率な意思決定

データ共有が正常にできる状態

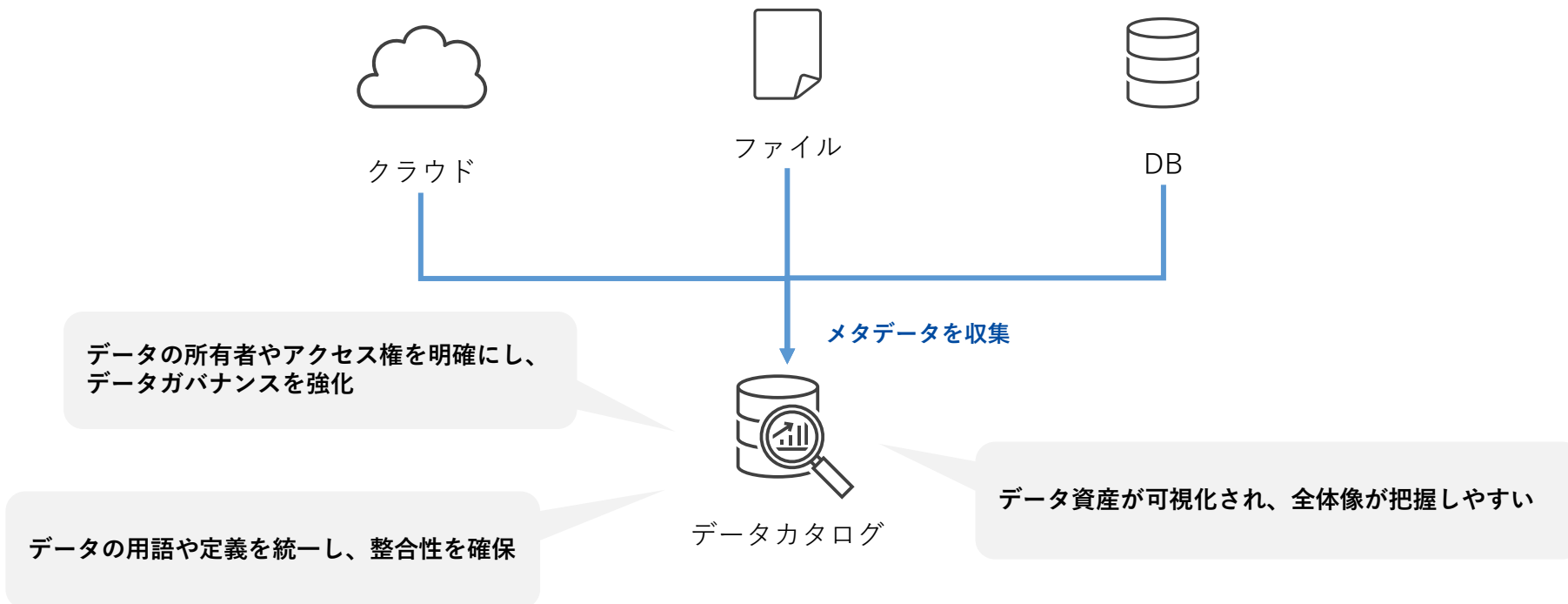


- ✓ 情報の一元管理が可能
- ✓ 重複作業の削減による業務の効率化
- ✓ データに基づいた効率的な意思決定

データ共有のためには、データ連携の仕組みが必要

データ連携だけでは不十分

各システムからデータを連携する際に、**信頼されたデータ**を扱うことが重要
データカタログを利用してデータを「**見える化**」することで、情報を精査



データ連携 × 見える化 → 効果的な資産の利活用

Appendix : その他のHULFT DataCatalog機能

メタデータの収集

リソースに対して接続し、物理的なメタデータを自動的に収集しカタログ化します。

● 収集できるメタデータ

DB系

スキーマ名、テーブル名、カラム名、
データタイプ、サイズ、コメント等



ファイル系

バケット情報、ファイル名、
フォルダ名、パス、サイズ等



ETL/EAI系

接続先一覧、プロジェクト名、スクリプト名
スクリプト内コンポーネント、コメント等



定期的に自動クロール

HULFT DataCatalog



自動的に最新状態の
カタログを構成

メタデータの収集

リソースに対して接続し、物理的なメタデータを自動的に収集しカタログ化します。

● サポートリソース一覧

■ DB系

■ Oracle Database

■ Microsoft SQL Server

■ IBM Db2

■ MySQL

■ PostgreSQL

■ JDBC

■ Amazon RDS

■ Azure SQL Database

■ Google BigQuery

■ Snowflake

■ ファイル系

■ Amazon S3

■ Azure Blob Storage

■ Google Cloud Storage

■ ETL/EAI系

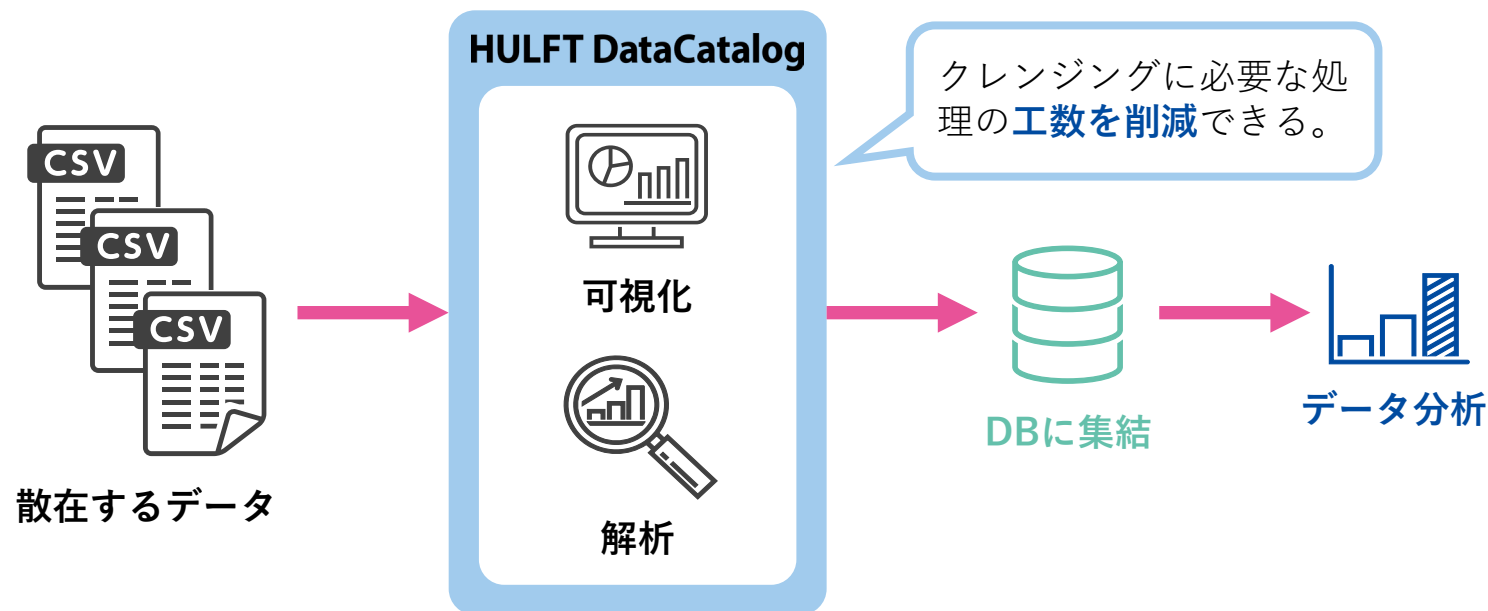
■ DataSpider Servista

データプロファイリング

● データ解析による状態の可視化

データの解析を行い、各列（フィールド）値の状態やデータ傾向・分布等の情報をメタデータとして取得します。

この機能により、データの品質を維持することに貢献し、誤った分析の予防やデータ利用者が個別に実施していたデータ補正コストを減少します。



ユーザー管理設定

● HULFT DataCatalog へのログイン認証方式を「ローカル認証」「SSO/SAML認証」「Active Directory認証」の3つから選択することが可能です。

- ユーザー認証方式の変更は、設定メニューのユーザー管理設定画面から「**ユーザー認証方式を変更する**」ボタンを押下し、行ってください。

ユーザー認証方式画面

現在選択のユーザー認証方式

☒ ローカル認証 ☐ SSO/SAML認証 ☐ Active Directory認証

ユーザー認証方式を変更する

ユーザー認証方式の変更について

認証元的环境設定も確認してください。

認証方式をローカル認証以外に変更する場合は、認証元的环境設定が必要になる場合があります。正常に動作することをご確認の上、変更手続きにお進みください。

選択できる認証方式

接続未設定の認証方式は選択できません。

設定画面にて接続情報を入力してから、認証方式の変更手続きを行ってください。

- ローカル認証
- SSO/SAML認証（SSO/SAML接続設定はこちらから）
- Active Directory認証（Active Directory接続設定はこちらから）

ユーザー認証方式の変更

ユーザー認証方式の選択

接続未設定の認証方式は選択できません。

設定画面にて接続情報を入力してから、認証方式の変更手続きを行ってください。

- ローカル認証
- SSO/SAML認証（SSO/SAML接続設定はこちらから）
- Active Directory認証（Active Directory接続設定はこちらから）

☒ ローカル認証 ☐ SSO/SAML認証 ☐ Active Directory認証

認証元的环境設定も確認してください。

認証方式をローカル認証以外に変更する場合は、認証元的环境設定が必要になる場合があります。正常に動作することをご確認の上、変更手続きにお進みください。

ユーザー認証方式を変更する

変更を中止して設定画面にもどる

テーブル設計書や定義書との統合自動化支援

- データの論理構造が定義されているテーブル設計書や定義書などの設計ドキュメント情報を物理的なメタデータと統合しカタログ化します。

PK	FK	論理名	物理名	型	サイズ	NULL	初期値	説明
1		所属ID	DEPARTMENT_ID	VARCHAR2	32			所属IDを指定する 10文字以内
2		所属	DEPARTMENT	VARCHAR2	192			所属の表示名
3		登録ユーザ数	USERS	NUMBER	5			この所属に登録可能なユーザ数の上限: 無制限
6		ログイン方法	LOGIN_TYPE	NUMBER	1			0: 通常ログイン 1: 認証サービス 2: 生体認証
7		認証サービス使用可否	AUTH_AVAILABLE_FLAG	NUMBER	1			【未使用】0: 認証サービスを使用しない 1: 認証サービスを使用する
8		認証サービス	AUTH_URL	CHAR	64			認証サービスを使用しない場合はnull
13		更新日時	UPDATED_AT	CHAR	17			【共通カラム】年月日時分秒ミ秒(yyyyMMddHHmmss)形式で保存
14		更新者	UPDATED_BY	VARCHAR2	32			【共通カラム】ユーザID
15		作成日時	CREATED_AT	CHAR	17			【共通カラム】年月日時分秒ミ秒(yyyyMMddHHmmss)形式で保存
16		作成者	CREATED_BY	VARCHAR2	32			【共通カラム】ユーザID
17		削除フラグ	DELETED	NUMBER	1			【共通カラム】1=削除 0=正常
18		レコードバージョン管理	LOCK_VERSION	NUMBER	9			【共通カラム】楽観ロック

収集

DataSpider
Servista

変換

インポート

ユーザ目線の論理情報とシステム
目線の物理情報をカタログ化

一連の流れを自動化、初期のカタログ立ち上げをクイックに

データプレビューとダウンロード

- カタログから発見したデータをプレビュー、ダウンロードすることができます。
今すぐにデータを活用したいニーズに応えることができます。

データプレビュー

リアルタイムにデータへ接続し、データの今の姿をプレビューします。

ダウンロード

実データをCSV形式でダウンロードできます。ユーザが欲しいタイミングでデータを入手できます。

抽出条件指定

データダウンロード時、抽出条件を指定する事ができます。SQLを書くことなく、目的のデータのみをダウンロードできます。

The screenshot displays the HULFT DataCatalog web application. The main interface shows a table preview for '有休未取得者マスタ' (Unpaid Leave Master) with columns: 社員No, 氏名, 部名, 課名, 有休取得日数, and 在籍. A 'ダウンロード' (Download) button is visible in the top right. A modal window titled 'ダウンロード' is open, showing a table with columns for '抽出条件を設定する' (Set extraction conditions). The modal includes a table with columns: 部名, 課名, 有休取得日数, 在籍, 氏名, and 抽出条件. The table contains the following data:

部名	課名	有休取得日数	在籍	氏名
HULFT	営業課	0	1	セゾン太郎
HULFT	営業課	0	2	セゾン次郎
HULFT	営業課	0	1	セゾン愛子
HULFT	営業課	0	1	セゾン美

Below the table, there is a section for '抽出条件' (Extraction conditions) with a table for '抽出条件を設定する' (Set extraction conditions). The table has columns: 抽出条件, 比較演算子, 値, and 論理演算子. The table contains the following data:

抽出条件	比較演算子	値	論理演算子
部名	=	HULFT	AND
課名	=	営業課	AND
有休取得日数	=	0	AND
在籍	=	1	AND
氏名	=	セゾン太郎	AND

A green arrow points from the 'ダウンロード' button in the main interface to the 'ダウンロード' modal window. Below the modal, the text 'SQLを記述する事なく、その場でデータ取得が可能' (Without writing SQL, data can be obtained on the spot) is displayed.

カスタムビュー

● 実用的なデータをオンデマンドで利用可能

自身に公開されているデータを自由にカスタマイズ（抽出や結合、並び替え）できます。
さらにそのI/Fを保存し、公開・再利用することができます。

社員テーブル

社員ID	名前	年齢	部署ID	課ID
202000001	鈴木一郎	44	1	1
202000002	佐藤次郎	32	1	1
202000003	田中三郎	24	1	2
202000004	伊東四郎	22	2	3
202000005	斎藤五郎	55	3	4

部署マスタ

部署ID	部署名
1	営業部
2	開発部
3	総務部

課マスタ

課ID	課名
1	第1課
2	第2課
3	システム課
4	経理課



それぞれの表から必要な要素だけを取り出して
仮想的な表を作る。

営業部の社員テーブル

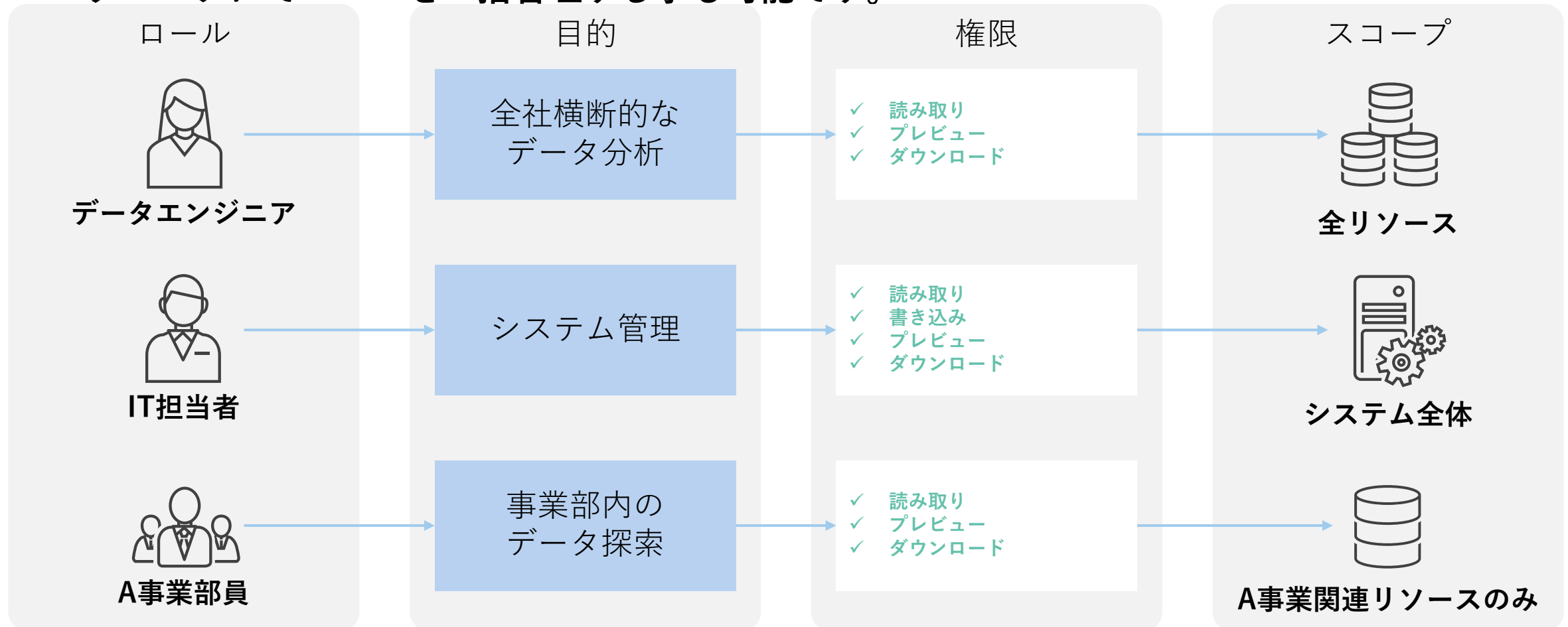
社員ID	名前	年齢	部署名	課名
202000001	鈴木一郎	44	営業部	第1課
202000002	佐藤次郎	32	営業部	第1課
202000003	田中三郎	24	営業部	第2課

データ利用者はカスタマイズ
したデータを自由に
レビュー／ダウンロード可能。



ユーザ権限管理

- ユーザの目的に応じて適切な権限を付与し、データ探索を管理支援する事ができます。
グループにてロールを一括管理する事も可能です。



アクティビティレポート

- HULFT DataCatalogではユーザのデータ探索状況を統計情報として記録します。この統計情報を用いることで、データに対するユーザニーズ特定にご活用いただけます。

HULFT DataCatalogで取得可能な統計情報



ログイン情報



検索情報



ダウンロード情報

定量的なデータ利活用状況を把握

よりデータ利活用が進む
次の一手の戦略立案に活用

Before

After



マネジメント

メンバにもっとデータを活用
した提案をしてもらいたいデータアクセスが増える毎にデータに
基づく論理的な提案になってきた

人財開発

データ活用人財を特定して
他メンバにも参考にしてもらいたいAさんがよく検索している
Aさんの考えをヒアリングしてみよう

IT担当

どんなデータにニーズがあるのか
知りたい、提供したい売上関連のデータが検索されている
過去のデータも提供できるように
整備してみよう

資料をご覧くださいありがとうございます。
お見積り・ご相談は[こちら](#)から

