

Informatica PowerCenter

メタデータ主導型データ統合プラットフォーム

Informatica PowerCenter は、データの価値を最大化します

『Informatica PowerCenter』は、大規模なデータ統合基盤として数多くの実績のある、インフォマティカ社のETL (Extract, Transform, Loading) ツールです。データ連携に必要なあらゆる機能を包括的に装備した、エンタープライズデータ統合プラットフォームです。

従来型のDWHにおける課題

ロジックのブラックボックス化

データ統合処理ロジックがDB内に埋め込まれ属人化・ブラックボックス化し、**運用負荷が増大**だけでなく、要件追加や将来の再構築など**変化対応の足かせ**に。

DBリソースの無駄遣い

加工前のローデータや、統合処理のための中間データ、似たようなデータの重複など、最終的には**不要なデータが貴重なDWHリソースを圧迫**

データガバナンスの欠如

どんなデータが存在するか、どういったロジックでデータが加工されているかが把握できず、**データの品質や信頼性、セキュリティも担保できない**。

PowerCenterで解決

ロジックの疎結合化と可視化

データ統合処理ロジックをDBから切り出すことで、属人化・ブラックボックス化を回避、**開発運用効率を向上**。ソース追加やDWH再構築にも**柔軟に対応可能**に。

DBリソースの最適化

必要なデータ統合・クレンジング処理をした上で、必要最小限のデータのみをロードすることで、データの冗長化を抑制、**DWHリソースを最大限活用**。

データガバナンスの強化

データの所在と流れを可視化、**来歴/影響分析を可能**にするとともに、**データの品質・信頼性、セキュリティも担保、ガバナンスを強化**。

Informatica PowerCenter データ統合を強力に支援する4つの性能

1. 接続性

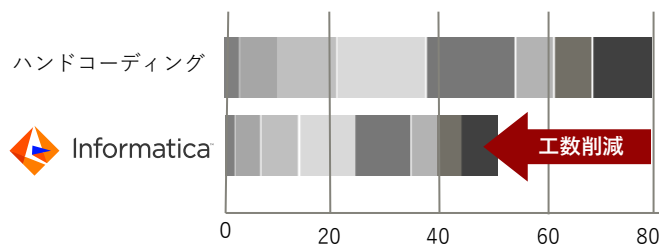
あらゆるデータへの容易なアクセスを実現

接続先は随時更新（追加）
掲載情報は2020年3月時点



2. 生産性

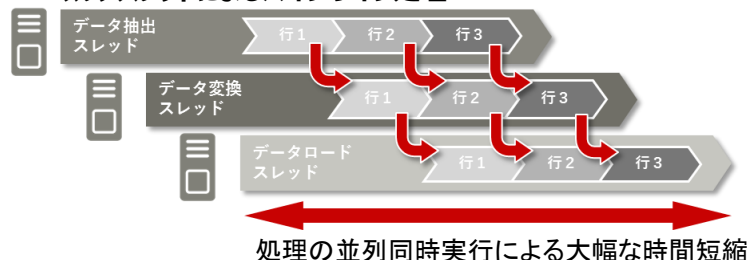
GUI開発、変換処理ロジックの再利用などによる開発・運用生産性の最大化



3. パフォーマンス

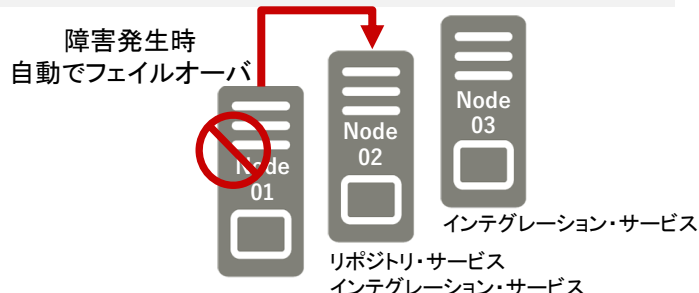
マルチコアCPUの性能を最大限に活かしたデータ連携処理の最速化

● マルチスレッドによるパイプライン処理



4. 可用性・拡張性

ミッションクリティカルな基盤としての堅牢性と柔軟な拡張性



PowerCenterの特長

企業のデータ連携を支えるデータ統合基盤ソリューションとして様々な特長があります。

➤ パフォーマンス

ETL処理で基本となるパイプライン処理を実装し、さらにプッシュダウン機能などによりお客様のデータ連携スピードを最速化します。さらにはスケールアップ、スケールアウトにも柔軟に対応でき、リニアな性能向上が期待できます。

➤ ユニバーサルアクセス

主要RDBMS製品および、各種ERP製品など多彩なデータ連携が可能です。またビジネス要求に合わせたタイミング（バッチ、リアルタイム、差分抽出）でのデータ連携を実現します。

➤ チームベース開発

開発に携わるユーザ管理、開発資産であるデータ連携ロジックのバージョン管理を実現します。チェックイン、チェックアウトの実装により排他的に操作し、リリース管理を実現しています。

➤ 開發生産性

開発環境はGUIとしてノンプログラミングで開発可能です。容易かつ視覚的に連携ロジックの把握、データ連携の見える化を実現します。さらに開発部品の再利用を可能にし、開發生産性を最大限に効率化します。

➤ メタデータ管理

データ連携開発の過程で収集した情報はすべてメタデータとして管理します。表定義だけではなく、データ項目、その属性を含めた情報が一元管理されます。

➤ 影響分析

メタデータ情報を元に、データ項目変更時などの影響分析、関連あるデータ連携の一覧表示、データ連携内で関連個所の見える化を実現します。

富士通の強み

- データ統合関連ビジネスに関して20年以上の実績を有し、経験豊富な技術者が多数在籍
- インフォマティカ再販パートナーとして15年以上のInformaticaソリューション取り扱い実績
- DWH/BI/BAといった情報利活用とともにお客様への最適なソリューションをご提案ご提供可能
- 「Informatica Partner Award」を2017年度、2018年度、2019年度の3年連続で受賞

富士通におけるInformaticaソリューション導入実績

業種を問わず様々なお客様への導入実績を有しております。

導入例	概略
【企業統合】 企業の統合に伴い、両会社のデータを統一して見られるDWHシステム用データの作成	お客様例：アパレル会社様 企業統合における全社を横断したデータ参照など、バラバラなコード体系／フォーマットを統一を実現
【共同利用システムからのデータ抽出】 複数会社で共同利用するシステムから、各社システムに合わせた形でデータを抽出	お客様例：カード会社様 共同利用システムのクラウド型システムから自システム用データ抽出を実現
【データハブ】 複数システム間の データ連携	お客様例：製薬会社様 「サイロ型のシステム」間連携の「データハブ」として実現
【大規模DWH作成】 データ量やテーブル数が多いDWHシステムの構築	お客様例：地方銀行様・生命保険会社様 大量データ、多数テーブル、短納期でのDWHの構築事例 手組での構築に比べ工数の低減およびスケジュールの短縮を実現

その他導入実績(業種)：物流・製造・金融(銀行)・金融(証券)・金融(クレジット)・商社・ヘルスケア・建設・不動産・造船・製薬・通信 など

※記載の会社名、商品名は、各社の商標または登録商標です。
※記載された情報は、予告なく変更することがあります。
※記載の内容は、2020年8月現在のものです。

お問い合わせ先

富士通コンタクトライン（総合窓口）
0120-933-200

受付時間 9:00～12:00 および 13:00～17:30
(土・日・祝日・当社指定の休業日を除く)

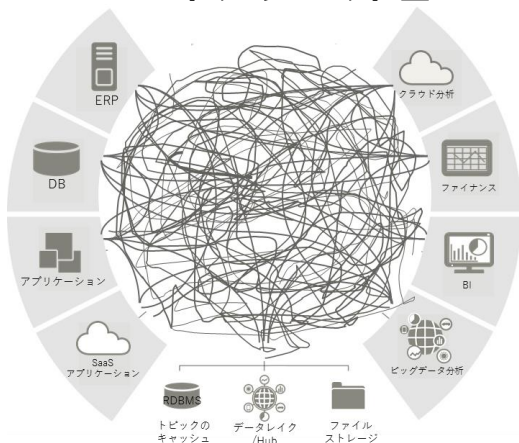
Informatica Data Integration Hub (DIH)

アプリケーション統合は、P2P型からハブ型へ

『Informatica Data Integration Hub』は、データ連携を送信(Publication)と受信(Subscription)に分けることで、システム間を疎結合化するとともに、データ連携の冗長性を排除、運用管理を一元化。データ連携開発・運用負荷の極小化とガバナンス強化を実現します。

従来型のデータ統合における課題

P2P (ピア・ツー・ピア) 型



とても複雑なデータ連携状態

- ✓ P2P型のファイル連携などによるスパゲティ状態
- ✓ システム間のデータの流が把握できない

拡張性の低さ

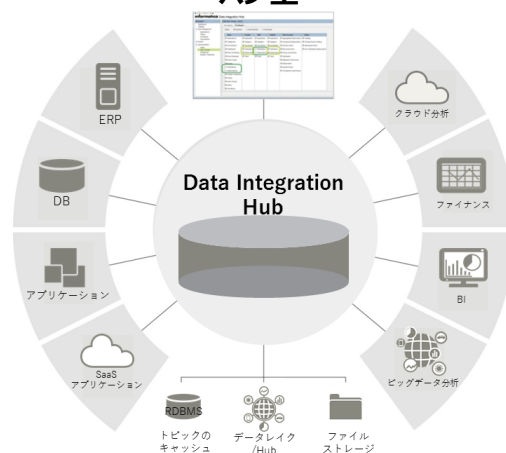
- ✓ 連携ロジックがブラックボックス化し、硬直化
- ✓ 新規の統合要件への対応に数か月かかる

ハイコスト

- ✓ 開発・運用コストが増大
- ✓ 障害時の復旧に多大な工数と時間を要する

InformaticaDIHで解決

ハブ型



簡明なデータ連携状態

- ✓ ハブ & スポーク型によるスパゲティからの脱却
- ✓ ハブを中心としたデータの流通を可視化

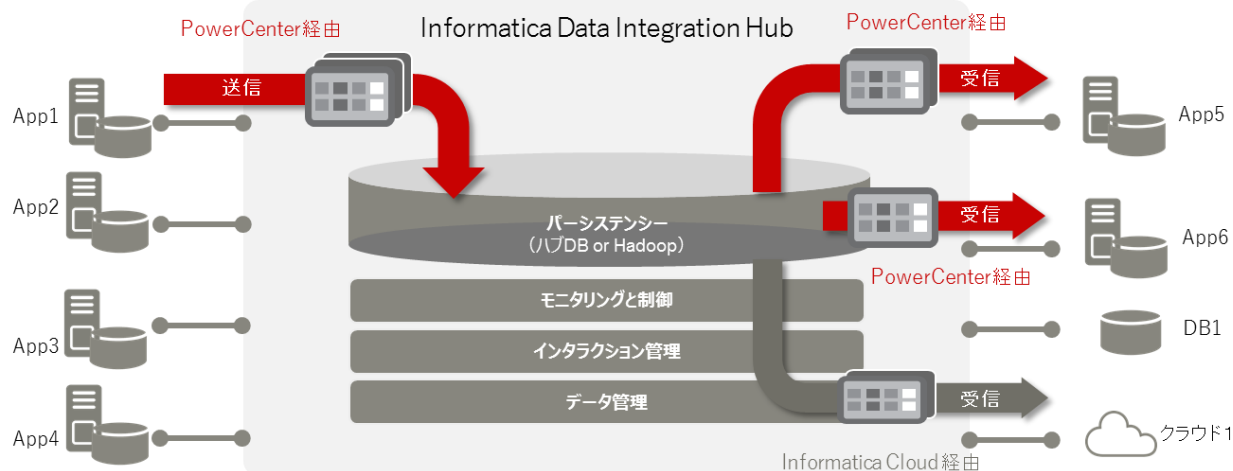
柔軟な拡張性

- ✓ アプリ間の依存性を下げ、柔軟に連携
- ✓ ハイブリッド環境を一元的に統合

ローコスト

- ✓ I/Fの冗長性排除による開発・運用コスト削減
- ✓ 障害時の復旧を迅速かつ容易に対応

システム間の疎結合化を実現するデータ統合プラットフォーム



DIHの優位性

①データの流通経路の見える化

- ・DIHを経由したデータフローの可視化が可能のため、関連システムの特定が容易になる。
- ・データの発生～活用先までのメタ情報を元にデータガバナンスの強化が図れる。

②開発・運用の効率化

- ・実際に使われている・いない、トピックの可視化がされ、連携すべきトピックの抜け漏れの確認が容易にできる。
- ・運用期間中に、既に不要になった連携の特定が容易にできる。

③変化対応力の向上

- ・新たに連携システムが増えた際、Hub上のトピックから必要なデータ項目をGUIで設定しSubscribeすることが可能。
- ・システムマイグレーション時に柔軟に並行稼働が可能。
- ・開発・運用時の、適切なアクセス制御が可能。
- ・機密情報を含むデータアクセスなど、権限を越えた不正アクセスを防ぐことが可能。
- ・単純連携の場合は、DIHのGUI画面から、自動マッピング生成の機能で、開発工数を大幅に削減可能。

④統合要件に応じた疎結合連携

- ・最新データを保持 or 過去分のデータを保持かを設定可能
- ・各システムの統合要件に応じたデータ同期が可能（夜間にバッチ連携、日中に差分同期、準リアルタイム同期など）
- ・障害時を考慮した運用が可能（リトライ・最新データ同期など）
- ・DIHはトピック単位でデータの削除期間が設定可能（期間指定、全てのSubscribeが完了したタイミングなど）

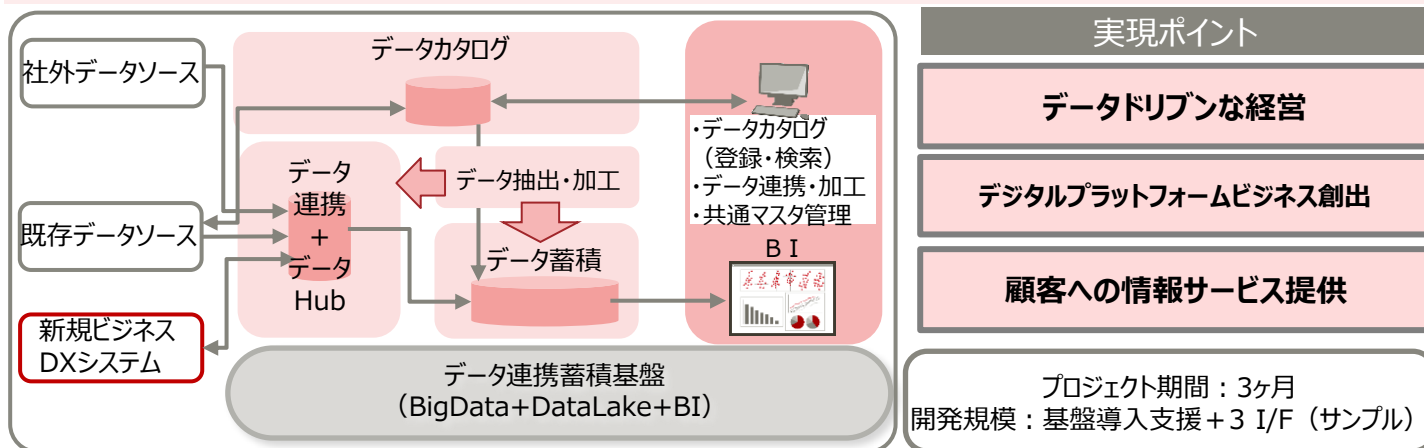
富士通の強み

- データ統合関連ビジネスに関して20年以上の実績を有し、経験豊富な技術者が多数在籍
- インフォマティカ再販パートナーとして15年以上のInformaticaソリューション取り扱い実績
- DWH/BI/BAといった情報利活用とともにお客様への最適なソリューションをご提案ご提供可能
- 「Informatica Partner Award」を2017年度、2018年度、2019年度の3年連続で受賞

富士通におけるDIH導入プロジェクト事例

データ連携蓄積基盤を構築して事業系システムのデータを集約し、
データドリブンなビジネスを展開する基盤として確立・活用することを目的とする

- ✓ データの論理的な集約／データカタログ化により社内システム情報を可視化
- ✓ 必要な時に必要なデータを連携してタイムリーに参照・利活用可能とするデータ連携プラットフォームを構築
- ✓ 全社的なデータ利活用に向けてデータモデル標準化・マネジメント規約を策定し、全社的なデータガバナンスを実現



※記載の会社名、商品名は、各社の商標または登録商標です。
※記載された情報は、予告なく変更することがあります。
※記載の内容は、2020年8月現在のものです。

お問い合わせ先

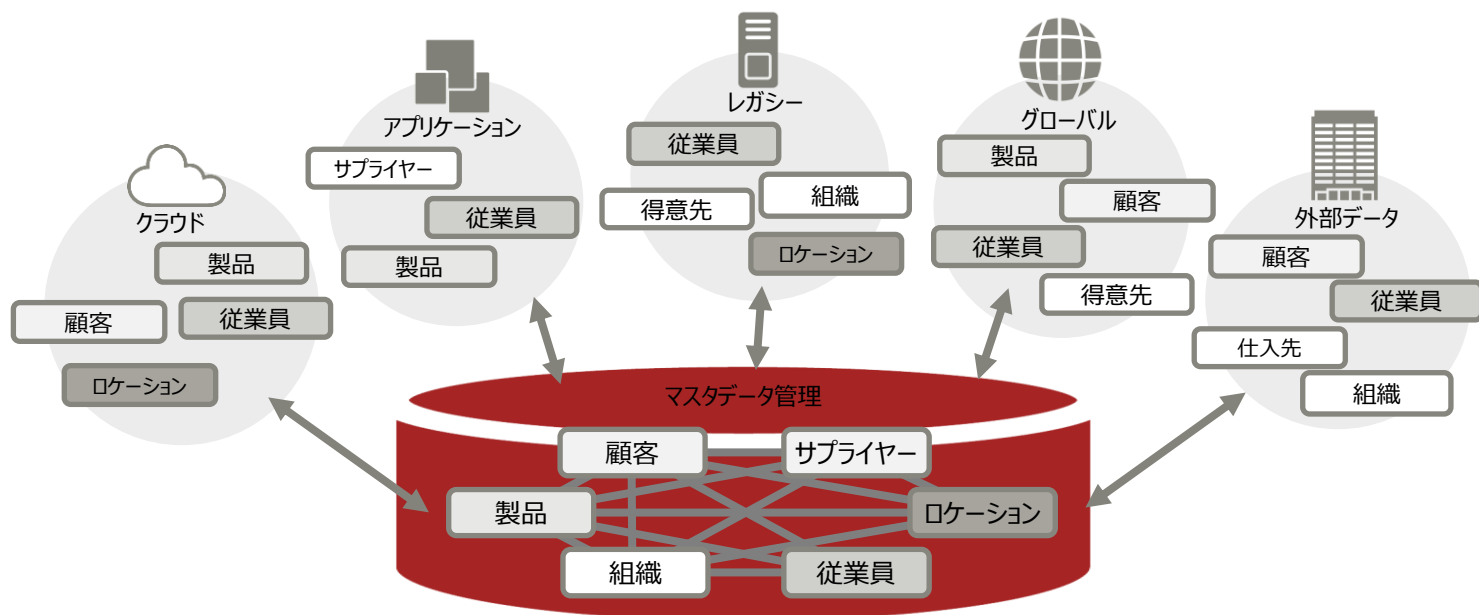
富士通コンタクトライン（総合窓口）
0120-933-200

受付時間 9:00～12:00 および 13:00～17:30
（土・日・祝日・当社指定の休業日を除く）

Informatica Master Data Management (MDM)

ビジネスに貢献するマスタデータ管理ソリューション

MDMのソリューションを導入することにより、ビジネスユーザーに対して、信頼できる品質の高い包括的なデータを提供し、「規制への対応」、「業務効率向上・コスト削減」によるマーケティング効率の改善や、「新規顧客獲得と維持」といった営業力の強化・CXの向上、「企業統合」「意思決定の向上」などビジネスに直結する価値を創出することが可能となります。

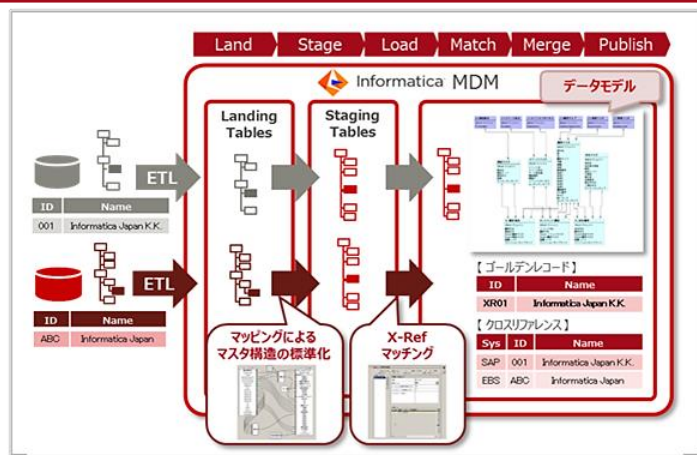


正確な業務・判断を実現するマスタデータ管理ソリューション

- ✓ データ集信、クレンジング、名寄せ、共通コード付与により、最も信頼できる「ゴールデンマスタ」を生成管理します
- ✓ 散在するマスタが統合／横串された一意のマスタとして利用可能となり、業務で扱うデータの精度が向上します
- ✓ お客様のビジネス状況がリアルタイムに可視化されるようになり、環境変化に強くタイムリーな判断が実現します

Informatica MDMのメリット

- マスタデータ管理に必要な全ての機能*1を共通データ・プラットフォームとして提供しています。
*1 データ連携、加工・クレンジング、名寄せ、マスタの登録・更新・削除画面、承認ワークフロー、更新履歴管理、データ品質監視、等
- お客様の業務要件や活用方法に合わせて、柔軟なデータモデルを構築することが可能です。
- お客様の目的やタイミングに応じて、様々なスタイルのマスタデータ統合アプローチを選択できます。



InformaticaMDMの特長

ビジネスに貢献するマスタデータ管理ソリューションとして様々な特長があります。

➤ **プラットフォーム型**

MDMの型にお客様のビジネスをはめ込むのではなく、お客様のデータをしっかりとMDMとして実装のできるプラットフォーム型のMDM環境構築を実現します。

➤ **マルチスタイル**

分析型、集中管理型、Hub 型など求められるMDM実装スタイルに合わせて、任意の実装が可能です。

➤ **マルチドメイン**

従来の1つのドメイン（PIM や CDI 等）を対象としたMDMではなく、1つのMDMプラットフォーム上で、ビジネスに必要なマスタデータのすべてを管理します。

➤ **マスタデータ管理コンソール**

MDM上に蓄積されたマスタデータを様々な観点から監視、管理するためのコンソールを提供します。MDM のデータ品質管理や、承認ワークフロー機能などを実現します

➤ **監査証跡**

MDM上で管理されるマスタデータでは、監査目的のための、変更履歴、証跡管理が必須となり、ビジュアルにデータ項目単位での監査証跡を実現します。

➤ **関連付け**

メタデータ情報を元に、データ項目変更時などの影響分析、関連あるデータ連携の一覧表示、データ連携内で関連個所の見える化を実現します。

富士通の強み

- データ統合関連ビジネスに関して20年以上の実績を有し、経験豊富な技術者が多数在籍
- インフォマティカ再販パートナーとして15年以上のInformaticaソリューション取り扱い実績
- DWH/BI/BAといった情報利活用とともにお客様への最適なソリューションをご提案ご提供可能
- 「Informatica Partner Award」を2017年度、2018年度、2019年度の3年連続で受賞

富士通におけるMDM導入プロジェクト事例

経営者や各国グループ企業のトップがタイムリーに意思決定するためのマネジメント指標が提示されていない
グローバルでの評価軸共通化と分析精度や柔軟性の向上が必要

課題と原因

- ・ グローバル横断でビジネスの状況が把握できない
- ・ 各拠点の入力データの粒度や集計単位がばらばらで、実態を把握するのに人手での集計が必要
- ・ 市場に対し、自社が強みを持つ分野での提供や商品展開ができていない
- ・ 入力データの粒度や集計単位のため、グローバル横断での分析に限界があり、市場やお客様の変化を捉えられない

課題解決に必要なこと

- グローバル同一基準で実態を把握するための評価軸共通化（マスタ整備・体制明確化）
- 横断分析するための共通コード整備と分析精度・柔軟性の向上

取り組み

グローバルマスタ統合管理プロジェクト

- ✓ グローバルで横串を通せる軸（顧客、商品）の整備
- ✓ マスタデータオーナーの明確化（顧客、商品）
- ✓ コード共通化

※記載の会社名、商品名は、各社の商標または登録商標です。
※記載された情報は、予告なく変更することがあります。
※記載の内容は、2020年8月現在のものです。

プロジェクト期間：15ヶ月
開発規模：79 I/F（連携先システム数 11）

お問い合わせ先

富士通コンタクトライン（総合窓口）
0120-933-200

受付時間 9:00～12:00 および 13:00～17:30
（土・日・祝日・当社指定の休業日を除く）

Informatica Data Quality (DQ)

全社レベルでデータ品質を管理する包括的なアプローチを実現

『Informatica Data Quality』を採用することで、データ品質の問題を発見・解消し、データから最大の価値を獲得します。

データの品質課題とは？

- ✓ データ品質は、完全性、適合性、関連性、一貫性、重複度などさまざまな観点で評価し、正確に課題を把握する必要があります。
- ✓ 外部データなどの信頼できるソースであっても、「株式会社」と「株」や電話番号等の表記フォーマットが異なる可能性があるためマッチング(名寄せ)の精度を高めるためにも、事前にプロファイリングし課題を把握、各課題に応じたクレンジングをすることが重要です。

【データ品質課題例】

顧客ID	会社ID	姓	名	業種	郵便番号	住所1	2	3	4	電話番号
A010010002	22001	園田	茂樹	20	324-42	栃木県大田原市大豆田3-15-17大田原センタービル				0375515395
A010010003	33001	馬場	清隆	20	370-2311	群馬県富岡市下高尾3-15				0307137216
A010020001	44001	井口	玲子	50	3670106	埼玉県児玉郡美里町南十条3-2-6				0369875175
A010020003	66001	浜崎	登	40	160-0005	東京都新宿区愛住町4-10-5				0323213105
	55001	清水	善次	60	375-0012	群馬県藤岡市下戸塚4-11-3				0387829328
...	...	...	...	...	...	...	...	...	...	...

顧客IDはプライマリーキーの役目を果たしているはずが、**データ移行時にNot Null制約に違反してロードに失敗**

業種は10、20のようなID値ではなく、「金融業」「製造業」などの実データとして**業種マスタ表**と連携しなければならなかった。さらに**60の値は業種マスタ表に存在しない**

フォーマットが異なるデータ
正しいフォーマット：999-9999

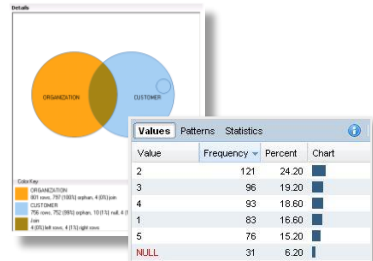
住所の分かち書きがなかったが、住所データの突合が技術的にクリアできずに住所データは妥協した

電話番号は10ケタのはずが、**データ移行時に桁あふれてロードに失敗**

データの品質を高める4つの主要機能

1. プロファイリング

現在のデータの品質を分析・評価



2. クレンジング

ルールを定義し、クレンジング部品でデータを標準化

< click to add a new row >

顧客顧客分類	顧客顧客タイプ
☐ Special	1
☐ Normal	5
☐ Platinum	2
☐ Gold	3
☐ Silver	4

☐ 10%以上
☐ トリプル
☐ パーサー
☐ マージ
☐ ラベラー
☐ 一致
☐ 加重平均
☐ 関連付け
☐ 大/小文字
☐ 変換
☐ 統合
☐ 比較
☐ 標準化
☐ 例外
☐ データ
☐ マッチング

4. スコアリング

定義したルールに基づき定常的にデータ品質を測定

Scorecard Before Scorecard

NAME	SCORE	% VALID CURRENT
CustomerName_Verify	87	
StateCd_Verify	90	
ZipCd_Accuracy	80	

3. マッチング

重複レコードの特定とゴールデンレコードの生成

顧客データ重複処理 - Duplicate Records

Cluster	Match	MATCH_SCORE	GROUP_ID	FILENAME	ADDRESS
1	1	1.0	1	1	埼玉県 さいたま市 中央区 さいたま市 中央区 さいたま市
2	1	1.0	2	2	埼玉県 さいたま市 中央区 さいたま市 中央区 さいたま市
3	1	1.0	3	3	埼玉県 さいたま市 中央区 さいたま市 中央区 さいたま市
4	1	1.0	4	4	埼玉県 さいたま市 中央区 さいたま市 中央区 さいたま市
Final	Records	0.00	10001	小宮 聡	埼玉県 さいたま市

データ品質

データ品質を支える高度なクレンジング処理機能

- 単純な数値変換、フォーマット変換では対応できないクレンジング処理を実現します。
- 住所クレンジング、重複データの検出、辞書を利用したデータ変換などに対応しています。

高度なクレンジング例
住所クレンジング

日本の住所クレンジング前

住所
東京都新宿区市ヶ谷本村町1-1
福岡県太宰府市国分二の七

- 住所の分割、修正、整合性検査
- 郵便番号の引当、逆引き対応
- 住所の分かち書き

クレンジング後

郵便番号	都道府県	市区郡	町村	丁目	番地	その他
1620845	東京都	新宿区	市谷本村町		1番地	1号
6340111	福岡県	太宰府市	国分	2丁目	2番地	7号

海外住所クレンジング例

100 Cardinal Street
Ledyard City, CA 94063-4755



100 Cardinal Way
Redwood City, CA 94063-4755

高度なクレンジング例
重複データの検出・排除

- 住所、氏名、生年月日などから同一と思われるデータを識別
- それらのデータを照合して、マージ

ソースデータ1

ID	姓	名	会社名	生年月日	郵便番号	都道府県	市区郡	町村
777	鈴木	一郎	インフォマティカ・ジャパン株式会社		1620845	東京都	新宿区	市谷本村町
11	山田	花子	株式会社 架空	1990年1月1日	8914207	鹿児島県	熊毛郡	屋久島町

ソースデータ2

ID	氏名	会社名	生年月日	郵便番号	都道府県	市区郡	町村
A500	鈴木 一郎	インフォマティカ株式会社		1020094	東京都	千代田区	紀尾井町
B333	佐藤 花子	株式会社 架空	1990年1月1日	8914207	鹿児島県	熊毛郡	屋久島町

マッチスコア

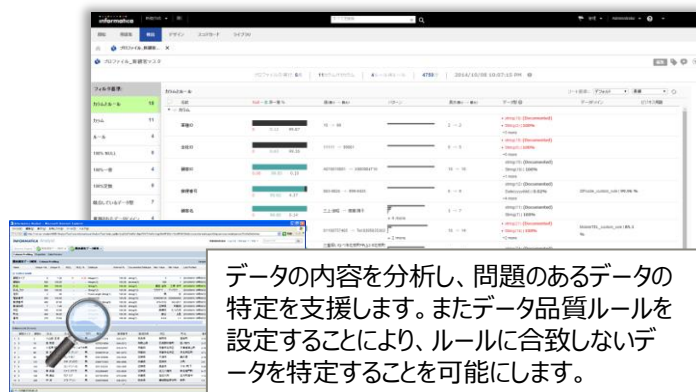
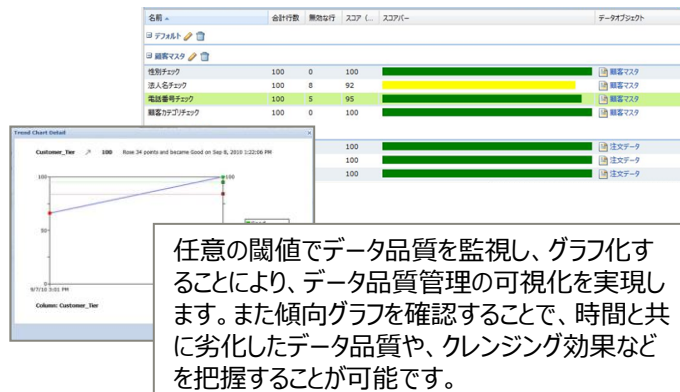
合計	氏名	会社名	生年月日	郵便番号	都道府県	市区郡	町村
70%	100%	80%	0%	0%	100%	0%	0%
90%	60%	100%	100%	100%	100%	100%	100%

生年月日に重み付け

データの品質を分析評価、監視するデータプロファイリングとスコアリング

データ品質管理を実現する上で重要なデータ内容の現状把握、継続的な分析機能を業務部門の方に使いやすいブラウザベースのツールとして提供します。

- プロファイリング : データの中身を分析し、迅速なデータ品質問題解決を支援
- スコアカード : 継続的に高いデータ品質を維持するためのデータ品質監視機能

高度なデータ分析
プロファイリングデータ品質監視
スコアカード

富士通の強み

- データ統合関連ビジネスに関して20年以上の実績を有し、経験豊富な技術者が多数在籍
- インフォマティカ再販パートナーとして15年以上のInformaticaソリューション取り扱い実績
- DWH/BI/BAといった情報利活用とともにお客様への最適なソリューションをご提案ご提供可能
- 「Informatica Partner Award」を2017年度、2018年度、2019年度の3年連続で受賞

※記載の会社名、商品名は、各社の商標または登録商標です。

※記載された情報は、予告なく変更することがあります。

※記載の内容は、2020年8月現在のものです。

お問い合わせ先

富士通コンタクトライン（総合窓口）
0120-933-200

受付時間 9:00～12:00 および 13:00～17:30
（土・日・祝日・当社指定の休業日を除く）

Informatica Enterprise Data Catalog

メリット

- AI搭載のカタログ機能により社内のあらゆるタイプのデータを自動的にカタログ化して分類
- インテリジェントなキュレーションでドメインやエンティティを特定
- 統制されたクラウドソースの注釈を使用してデータ資産をエンリッチ化
- Googleのようなパワフルなセマンティック検索によりデータ資産を発見
- リネージ、関係ビュー、データプロファイリング、品質統計などの総合的なビューでデータ資産を探索および把握
- データ環境の全体像を把握
- オープンAPIを自社環境に統合して、インテリジェントなメタデータをあらゆる場所で活用

インテリジェントなデータカタログでデータの力を最大限に引き出す

データは経済の生命線であり、データ主導型の企業は自社のデータ資産を売上や利益に変えています。データ主導のデジタルトランスフォーメーションの第一歩は、データを企業の資産として管理することです。

データは多様化しているだけでなく、さまざまな部門やアプリケーション、データウェアハウスに分散しています。そのため、どこにどのようなデータがあるのかを正確に把握することが難しくなっており、ビッグデータの時代では、この状況がさらに複雑になります。

Informatica Enterprise Data Catalog は、AI 搭載のデータカタログです。機械学習ベースの探索エンジンを通じて、クラウド、オンプレミス、ビッグデータなどソースに関係なくデータ資産を全社規模でスキャンしてカタログ化することが可能です。Enterprise Data Catalog の CLAIRE™ エンジンが、メタデータを活用して、データ管理タスクに関する推奨、提案、自動化をインテリジェントに実行します。これにより、IT 担当者の生産性が高まるとともに、業務担当者はデータの管理を全面的に協力できるようになります。

業務担当者と IT 担当者は、パワフルなセマンティック検索をはじめ、検索結果を絞り込む動的ファセット、データリネージ、プロファイリング統計、総合的な関係ビュー、データの類似性の提案、統合されたビジネスグロッサリを利用できます。また、企業データ資産の管理を簡素化し、全社規模でデータの価値を最大限に引き出せます。そして業務担当者は、データをすばやく見つけ出し、ビジネス用語や定義、参考データなどのライフサイクルを簡単に管理することができます。

Informatica Enterprise Data Catalog は、AI 搭載のデータカタログです。機械学習ベースの探索エンジンを通じて、クラウド、オンプレミス、ビッグデータなどソースに関係なくデータ資産を全社規模でスキャンおよびカタログ化することが可能です。

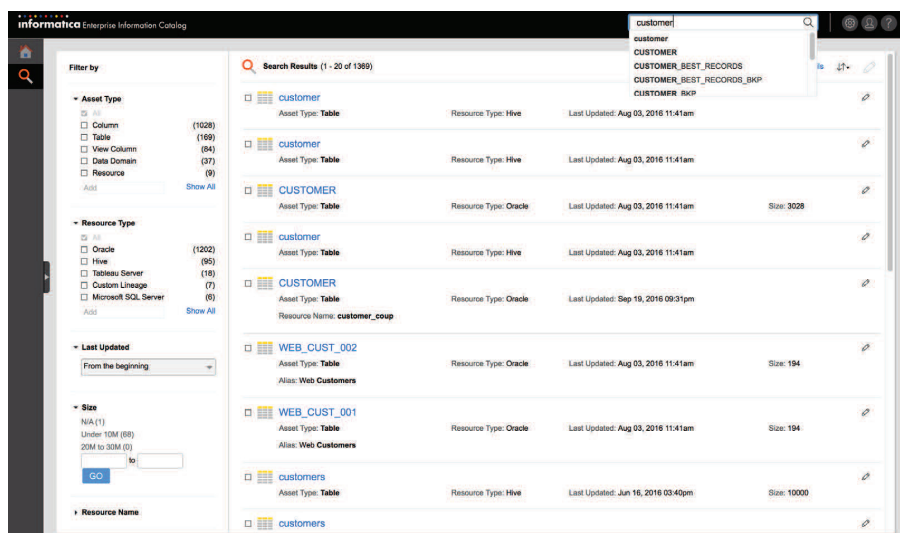


図 1：スマートセマンティック検索機能と動的ファセットにより、データセットを手早く発見

主な特長

メタデータ API を自社環境に統合

Enterprise Data Catalog を使用して REST ベースの API を自社環境に統合することにより、カタログのコンテンツをどこからでも利用できるようになります。また、業務担当者とインテリジェントなメタデータ（アプリケーション、BI レポート、ダッシュボード）を共有できるようになります。

Tableau 向けプラグインでデータのガバナンスと信頼性を確保

Informatica Enterprise Data Catalog for Tableau により、統制されたデータを使用した迅速なセルフサービスアナリティクスを実現します。Tableau ユーザーは、Enterprise Data Catalog のすべてのリソースを活用してデータを視覚化することができます。また、業務担当者は、ワークシート、ダッシュボード、データソースなどのビジネスコンテキストと技術コンテキストが並列表示された Tableau レポートを利用可能です。

インテリジェントな機能を実装したセマンティック検索

パワフルなセマンティック検索とインテリジェントな機能により、最も関連性の高いデータセットを検出、発見、分析することができます。また、トークンマッチングを活用した高度なキーワード検索では、カタログ内の最も関連性の高いデータ資産を見つけることができます。さらに、セマンティック検索は参照先のデータドメインにも適用されるため、データ資産が発見されずに見過ごされることはありません。検索結果に基づいてインテリジェント機能を活用することで、関心のあるデータセットに検索を絞り込むことができます。

データリネージと影響分析

情報をわかりやすくまとめたリネージビューを通して、データソースを相互に追跡できます。これらのビューではエンドポイントに焦点が置かれ、その間にある複雑な情報は表示されません。リネージビューはドリルダウンが可能で、すべてのリネージパスが展開できるため、さまざまな列やリネージダイアグラムの指標を確認することができま

す。また、上流／下流のデータ資産に関する詳細な影響分析を実行できます。

ドメインとエンティティをインテリジェントに認識して分類を自動化

すべての構造化／非構造化データに関して、ドメインやエンティティ（顧客、製品、注文など）を項目、列、テーブルレベルで自動的に分類および特定します。これは、企業がデータをカタログ化して管理し、価値を引き出す態勢を構築するための重要なステップです。データを分類することにより、検索、検索結果の絞り込み、ビジネスグロッサリの提案を強化できます。インフォマティカは、電子メールやクレジットカード番号、社会保障番号、国名や都市名、URL、会社名など、あらかじめ組み込まれた 60 以上のデータドメインを提供します。ユーザーが作成したカスタムドメインを追加することも可能です。データ資産は、データルール（ルールで定義された特定のロジックに一致するデータがある列を検出）、または列名ルール（ルールで定義された列名のロジックと一致する列を検出）を使用して分類できます。

統合データ品質統計

技術メタデータに加えて、データプロファイリング統計を確認し、データ資産の品質を把握してから、データを分析に使用できます。プロファイリング統計には、値の分布やパターン、データタイプやデータドメインの推測などがあります。

Informatica Axon のビジネスグロッサリの統合

Informatica Enterprise Data Catalog では、用語、ポリシー、分類などのビジネスグロッサリを簡単に Informatica Axon™ からインポートすることができます。ビジネス用語と適切な技術メタデータを関連付けることで、データに質の高いビジネスコンテキストを付加することができます。また、用語の関連付けを提案する機能により、業務部門と IT 部門のデータ管理者が連携しながら、人的ワークフローの効率的な自動化を含むビジネスメタデータを管理することが可能です。また、Informatica Business Glossary やサードパー

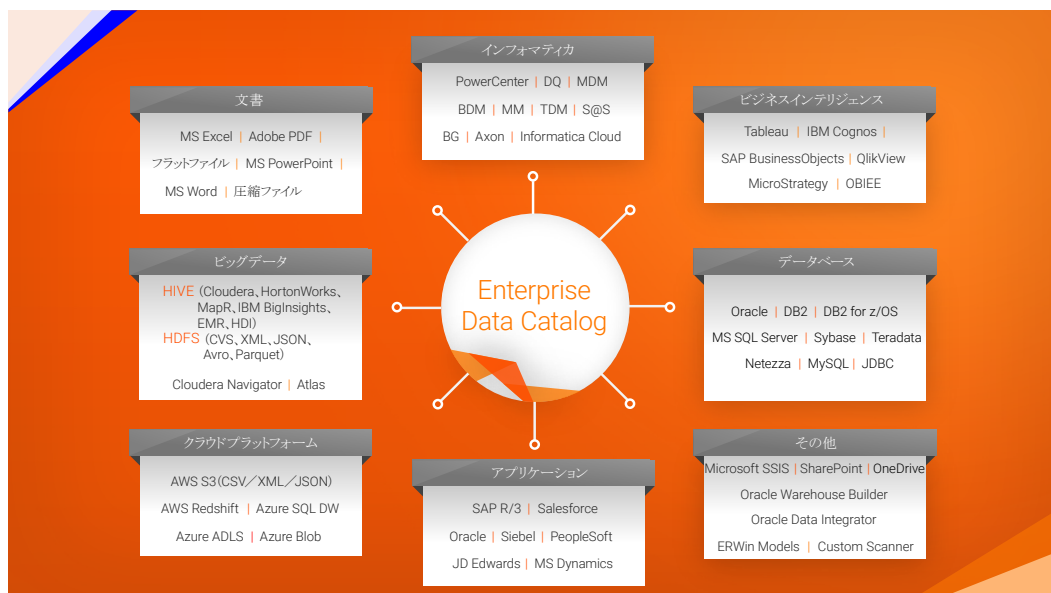


図 2 : Informatica Enterprise Data Catalog はユニバーサルなメタデータ接続性をサポート

ティツールからビジネスグロッサリ資産をインポートすることもできます。

インテリジェントな類似データの提案機能

高度な統計アルゴリズムと機械学習アルゴリズムが、類似するデータとデータサブセットを特定します。このパワフルな機能によって、最も関連性の高い、信頼できるデータを見つけ出すことができます。例えば、通信会社のアナリストが、顧客が他社への乗り換えに関する分析のため、当四半期のプリペイドカスタマーの活動を含むデータのクエリを実行するとします。この場合、Informatica Enterprise Data Catalog は、よりクリーンなデータ（データの置き換え）、前四半期の顧客の活動を含むデータ（結合可能なデータ）、データセットをエンリッチ化するための顧客情報テーブルを提案します。

ユニバーサルなメタデータ接続性

データベース、データウェアハウス、アプリケーション、クラウドデータストア、BI ツール、Hadoop、NoSQL など、企業全体で使用しているあらゆるタイプのデータソースからメタデータを抽出します。メタデータの抽出をサポートしているデータソースの例は図 2 のとおりです。

リソースレベルのセキュリティ

ユーザーやグループにリソースレベルで読み取り／書き込み権限を付与できます。ユーザーは、カスタム属性の閲覧／編集、ドメインキューレーションの実行、ビジネスグロッサリの用語の関連付けが可能です。

ビッグデータの大規模導入

Enterprise Data Catalog は、ビッグデータの大規模導入用に構築されており、Hadoop クラスタに実装できます。並列処理によるメタデータの取り込みと高速な分散インデクシングをサポートしているので、カタログコンテンツを迅速に更新し、卓越した検索パフォーマンスを実現します。また、24 時間 365 日のフォールトトレランス性（耐障害性）を提供します。

統合システム管理

カタログリソース、メタデータ抽出スケジュール、プロファイリング実行などのタスクを統合システム管理コンソールで管理および監視できます。ジョブコントロールダッシュボードでは、タスク監視ビューやリソースビューを表示するウィジェットを利用できます。また、システム管理者は電子メールアラートを設定することで、カタログの問題にプロアクティブに対応することが可能になります。

主なメリット

社内のあらゆるタイプのデータをインテリジェントにカタログ化

Informatica Enterprise Data Catalog は、企業内のさまざまなタイプのデータやその関係をインテリジェントに探索します。あらかじめ組み込まれたスキャナーが、データベース、データウェアハウス、アプリケーション、クラウドデータストア、BI ツール、Hadoop、NoSQL などからメタデータを収集します。すべてのメタデータは、迅速な更新やスマート検索、高速クエリを目的として設計された拡張可能なグラフデータベースでインデックス化され、カタログ化されます。生成されるデータ量が増えると共に企業組織全体に広がっていくのに伴い、類似データや重複データが増えてしまうことは避けられません。Informatica Enterprise Data Catalog は、高度な統計アルゴリズムと機械学習アルゴリズムを活用して類似データとデータのサブセットを探索することで、ユーザーが必要としている最も関連性の高い信頼できるデータを見つけ出します。

Google のようなパワフルなセマンティック検索でデータ資産を迅速に発見

何百もの企業システムの中から必要なデータを探すという作業は、時には実りのない無駄な労力に思えることがあるかも知れません。包括的なメタデータサービスと拡張可能なインフラストラクチャを基盤にしたパワフルなセマンティック検索機能がなければ、関連性のあるデータを見つけることさえできないでしょう。Informatica Enterprise

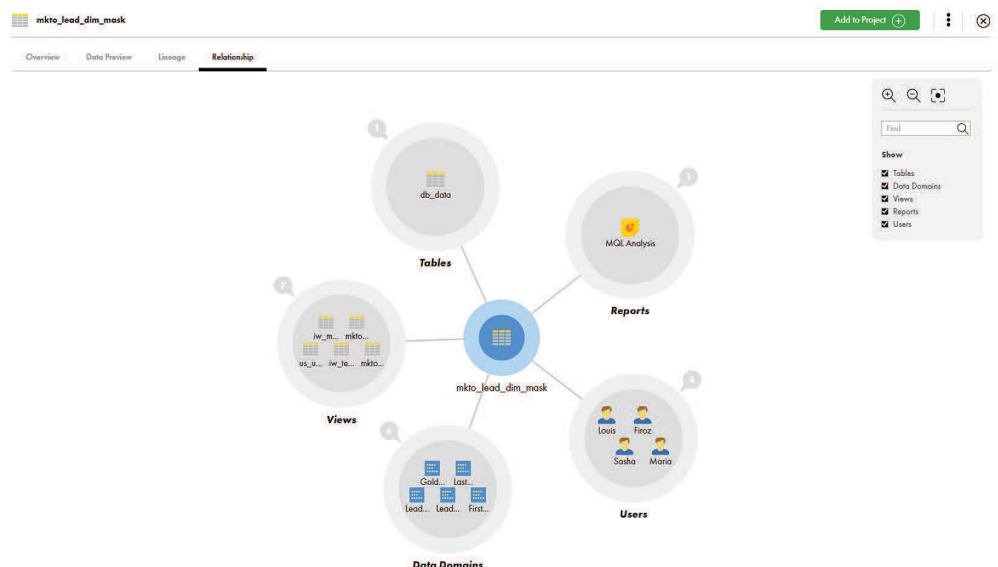


図 3：データの関係を示す総体的なビューでデータを把握

インフォマティカについて

インフォマティカは、エンタープライズクラウドデータ管理をリードする唯一の企業であり、データ主導型のデジタルトランスフォーメーションを推進しています。企業のイノベーションを加速し、俊敏性の向上や新たな成長機会の獲得によって、インテリジェントな破壊的イノベーションを支援しています。また、インフォマティカは企業のデータの力を25年以上にわたり引き出し、世界で9000社以上の企業がそのソリューションを利用しています。

インフォマティカに関する詳細は、インフォマティカ・ジャパン株式会社（代表：03-6403-7600）までお問い合わせいただくか、Webサイトwww.informatica.com/jp/をご覧ください。

Data Catalog では、セマンティック検索とインテリジェントな機能により、検索結果を詳細に絞り込むことができます。ビジネスメタデータ、技術メタデータ、運用メタデータが独自の方法で関連付けられているため、業務担当者はデータを探す際にビジネス用語を検索し、総体的な関係ビューを確認して、関連性のあるデータ資産を見つけることができます。

総体的な関係ビューとリネージでデータ資産を探索および把握

「測定できないものは管理できない」という格言がありますが、データ資産の管理についても同じことが言えます。データから価値を最大限に引き出すには、保有しているデータ、データの出所、データの変更履歴、その信頼性を把握する必要があります。Informatica Enterprise Data Catalog では、完全なエンドツーエンドのサマリーと詳細なリネージ、プロファイリング統計、総体的な関係ビューを通じてさらに多くの情報が提供され、ユーザーはデータを明確に把握できるようになります。

統制されたクラウドソースの注釈とビジネスコンテキストを使用してデータ資産をエンリッチ化

Informatica Enterprise Data Catalog は、企業データ資産を項目／列レベルまで自動分類し、データの再利用性と価値を最大化します。データの価値をさらに高めるために、クラウドソース対応のタグや注釈に加え、データの使用者や使用目的といったコンテキストも取得します。このような「集合知」を活用することでデータのエンリッチ化と取得を行い、全社規模でデータの有益性をさらに高めることが可能になります。また、Informatica Axon との統合により、ビジネス用語、定義、ポリシーなどのビジネスグロッサリを Axon から簡単にインポートできます。これらのビジネスメタデータは技術メタデータや運用メタデータと関連付けられるので、業務アナリストやデータスチュワード（データ管理／案内人）などのユーザーは、データを迅速に発見、把握し、協力してデータを活用することができます。

詳細はこちら

Informatica Enterprise Data Catalog の詳細については、<https://www.informatica.com/jp/products/big-data/enterprise-data-catalog.html> をご覧ください。



Informatica®

インフォマティカ・ジャパン株式会社

〒105-6226 東京都港区愛宕 2-5-1

愛宕グリーンヒルズ MORI タワー26 階

電話：03-6403-7600(代表) FAX：03-3433-1031

<http://www.informatica.com/jp/>

Informatica パートナー

富士通コンタクトライン(総合窓口) 0120-933-200

受付時間 9:00~12:00 および 13:00~17:30

(土・日・祝日・当社指定の休業日を除く)

Informatica

Enterprise Data Preparation

メリット

- 分析プロジェクトに最適なデータを発見
- データの信頼性を把握
- 必要なデータを迅速に準備して共有

分析データの発見、準備、管理をコラボレーティブに実現

21 世紀において、データが競争優位性を高める要素となっていることに疑いはありません。企業は、Apache Hadoop などの新しいデータ処理プラットフォームを活用し、これまで手にすることができなかった洞察を導き出そうとしています。Apache Hadoop とデータレイクという概念の登場により、企業のすべてのデータを蓄積するだけの余裕が生まれ、ユーザーはいつでもデータにアクセスしてあらゆるタイプの分析を行えるようになりました。

企業は、カスタマーエクスペリエンスの向上とビジネスの成長の可能性を探るために、顧客と市場のデータを収集しています。金融機関は、不正行為を検出する技術の強化、変化するグローバル規制への対応、自社サービスのセキュリティに対する顧客の信頼性向上のために、取引データやその他の予兆情報を収集、監視しています。医療機関は、個人に合わせた医療サービスを提供するために、電子カルテと請求データを保存しています。ビッグデータテクノロジーによって、データ活用のお機会はかつてないほど広がっているのです。

マイプロジェクト > CustomerActivityAnalysis_20190425

概要

所有者: Administrator

最終更新日時: Administrator - 2019年4月26日 12:32

入カソース: 3

ワークシート: 5

パブリケーション: 1

説明はありません。

コラボレータ

このプロジェクトにコラボレータは割り当てられていません。コラボレータを追加してください。

ワークシート

名前	ステータス	入カソース	パブリケーション
Agg_contents_view	作業進行中	t_contents	
Analysis_Result	パブリッシュ済み	Agg_contents_view, m_customer	analysis_result
m_customer	作業進行中	m_customer	
m_custtype	新規	m_custtype	
t_contents	新規	t_contents	

推奨事項 1追加、3代替

combination1_0201

リソース名: edp_hivescanner

作成者: Administrator

t_comments

リソース名: edp_hivescanner

作成者: Administrator

combination1_0125

リソース名: edp_hivescanner

作成者: Administrator

combination2_0201

リソース名: edp_hivescanner

作成者: Administrator

データ資産の提案を表示するプロジェクトワークスペースが他のアナリストとのコラボレーションを促進

負担か恩恵か？

IT 部門にとって、Hadoop システムに取り込まれるデータは手に余るほど圧倒的な量です。業務アナリストが Hadoop からの質の高いデータを待っている一方で、IT 部門にとっては手作業で時間をかけて未加工のデータを目的に合ったデータ資産にキュレーションするという負担がのしかかっています。データのキュレーションのための、拡張性に優れ、再利用可能でインテリジェントな仕組みがなければ、データレイクに期待されるすべての機会を実現できないリスクが生じます。データレイク自体は目的を達成する手段に過ぎず、そこから価値を引き出すことができないければ、単なるデータの泥沼となってしまいます。このような「データの泥沼化」の危機を解決するための鍵になるのが、データセットに関する情報を収集、インデックス化、プロファイリング、検索、推測を可能にする、ユニバーサルなメタデータにもとづくデータインテリジェンスです。データレイク内のデータの質と状態の両方をカタログ化することで得られる付加価値と、これらのデータに対するコラボレーティブなデータプレパレーション機能を通じて、はじめてビジネスインサイトを得ることができます。Informatica Enterprise Data Preparation は、多様なデータ利用者向けに、生のビッグデータを目的に合わせたデータセットへと系統立てて変換します。データサイエンティストおよびアナリストは、ファセット検索機能を備えたセマンティックサーチを利用して、目的のデータを迅速に見つ

けることができます。データのプロファイル、リネージ、その他の関係性を把握して、データを信頼できるかどうか、また対象とするアナリティクスプロジェクトに適しているかどうかを判断できます。

レコメンデーションとレシピ

データアナリストのチームは、プロジェクトのワークスペースで簡単にコラボレーションを実現し、結果を共有できます。プロジェクトワークスペースにデータセットを追加すると、機械学習アルゴリズムがバックグラウンドで機能し、チームの役に立つ可能性がある代替のデータセットを推奨します。プロジェクトワークスペース内にあるデータセットは、Excel ライクな使いやすい Informatica Enterprise Data Preparation のデータプレパレーションツールでいつでも開くことができます。

データセットを統合するための結合キーの推奨など、データインテリジェンスがデータプレパレーションをガイドします。データプレパレーションツール内のすべての加工手順は、レシピに記録されます。このレシピは、ビッグデータのインサイトを定常運用化する際に、本番環境に即実装可能なデータ処理ロジックを自動的に生成します。

このようにして、企業はビッグデータを持続可能な業務上の価値をもたらす信頼性の高い情報資産へと何度でも迅速に変換することが可能になります。

CustomerActivityAnalysis_20190425

カラム レシピ 材料

レシピステップ

1. テーブルからのサンプルデータのロード: m_customer

2. Edit all cells with value 'おとこ' to '男' in gender

3. Edit all cells with value 'Female' to '女' in gender

4. Edit all cells with value 'おんな' to '女' in gender

5. Lookup custtype_name for all rows from m_custtype on key...

6. Extract email domain from mail into new column mail_domain

#. 次のアクションがここに表示されます。

m_customer	t_contents	m_custtype	Agg_contents_view	Analysis_Result					
#	custno	name	namekana	gender	custtype	custtype_na...	customerstartdate	mail	mail_dom
1	T-0000...	小宮 聡	コミヤ サト...	男	1	Special	2016/01/23	omisawa@mrt...	mrt...
2	T-0000...	園田 茂樹	ソノダ シゲ...	男	4	Silver	2016/01/23	hana200@rmq...	rmq...
3	T-0000...	小山田 正道	オヤマダ マ...	男	5	Normal	2016/01/23	natsumi0033@uid...	uidq.kei
4	T-0000...	平松 建司	ヒラマツ ケ...	男	1	Special	2016/01/23	hinako_kakiuchi...	xopiuet...
5	T-0000...	堀口 元	ホリクチ ハ...	男	1	Special	2016/01/23	akotani@zioqit...	zioqit...
6	T-0000...	宮脇 昌利	ミヤウキ マ...	男	3	Gold	2016/01/23	hyl-usfxshana53...	awftuuz...
7	T-0001...	村木 秋夫	ムラキ アキ...	男	2	Platinum	2016/01/23	rika647@rpmu.w...	rpmu.w...
8	T-0001...	杉村 直	スギムラ タ...	男	2	Platinum	2016/01/23	yoshimi6753@xfj...	xfjppzzg...
9	T-0001...	北岡 文男	キタオカ フ...	男	3	Gold	2016/01/23	nozomi373@rjqvu...	rjqvukw...
10	T-0001...	石黒 崇	イシクロ タ...	男	4	Silver	2016/01/23	marumi_eguchi@x...	xrowhmi...
11	T-0001...	高坂 時雄	コウサカ ト...	男	2	Platinum	2016/01/23	Yoshio_Iwai@spf...	spfjj.q...
12	T-0001...	松本 太一	マツモト タ...	男	5	Normal	2016/01/23	wakamizuno@odgw...	odgunrx...
13	T-0001...	北沢 勇夫	キタザワ イ...	男	5	Normal	2016/01/23	daiki1972@cibyw...	cibyw.yl
14	T-0001...	藤崎 大造	フジザキ タ...	男	2	Platinum	2016/01/23	satoshi43916@vo...	vomg.bp
15	T-0001...	大森 大貴	オオモリ タ...	男	1	Special	2016/01/23	hana_sasamori@...	buyiokou
16	T-0001...	立川 運	タチカワ レ...	男	4	Silver	2016/01/23	Masae_Omata@npt...	nptcbzi...
17	T-0002...	宮内 秀	ミヤウチ ヒ...	男	2	Platinum	2016/01/23	yuukomiyachi@...	aurjtma
18	T-0002...	西 敦彦	ニシ アツヒ...	男	3	Gold	2016/01/23	aoda@vxzvglg.m...	vxzvglg
19	T-0002...	新川 淳三	シンカワ ジ...	男	1	Special	2016/01/23	ihonda@yny1mz.u...	yny1mz.i

ワークシートの概要

列の概要: mail

268 行

概要

タイプ: Email Address

ソース: テーブル: m_customer >...

一意: 99.63%

空白: 0%

長さの範囲: 15から41

値の頻度

ukishida@ijkgxbik...

Akane_Sasahara@...

Aki_Yasuda@fzwr.z...

Akiko_Kawase@aj...

Akiko_Sudou@bxv...

Ayaka_Yasui@qaer...

Chihiro_Kakizaki@...

提案

で分割

電子メールからドメイン名を抽出

大文字

小文字

分析用のデータセットを簡単に準備し、レシピによりデータ変換ステップを自動化

主な機能

スマート検索

データサイエンティストやアナリストは、スマートセマンティック検索機能と推測に基づく結果を使用して、データレイクや企業全体の他のシステム内のデータを短時間で見つけることができます。データ資産は、システム属性やカスタム定義の分類項目を切り口として、動的ファセットに基づいてフィルタリングできます。

360度のデータ探索

カスタム属性、データ品質のためのプロファイリング統計、ビジネスコンテンツ向けのデータドメイン、利用状況に関する情報など、データ資産の概要を把握できると共に、タグ付けによってデータセットに関する情報をクラウドソーシングし、メタデータを拡充してビジネス上の意義を追加できます。また、ユーザーの認証情報に基づいてサンプルデータをプレビューすることで、データを素早く把握でき、データリネージを確認することでデータがどこから来たのか、どこへ行くのかを把握できるので、データの信頼性が高まります。さらに、任意のデータ資産と、他のテーブルやビュー、ユーザー、BIレポート、データドメインなどの企業全体のオブジェクト間の関連性を把握できます。その他の資産も、リネージビューと関係ビューによって着実に特定できます。

ガイドつきデータプレパレーション

Excel ライクな直感的なインターフェイスで、データをフィルタリング、集計、統合、結合するためのあらかじめ組み込まれた変換機能を使用して、分析用のデータをインタラクティブに加工することができます。文字列、計算、日付、論理の各演算を使用して、列レベルでデータのクレンジングやデータ変換も行えます。ガイド式のインテリジェンス

が、データセットをブレンドする際の結合キーの提案など、データセットのプレパレーションを支援します。また、値の分布や数値と日付の分布など、シートレベルと列レベルの記述的な統計概要を確認でき、すべてのステップがレシピーに記録され、データ処理ロジックが自動生成されます。このデータフローは、分析のインサイトを定常運用化する際に、スケジュール設定をして繰り返し実行することもできます。

プロジェクトのコラボレーション

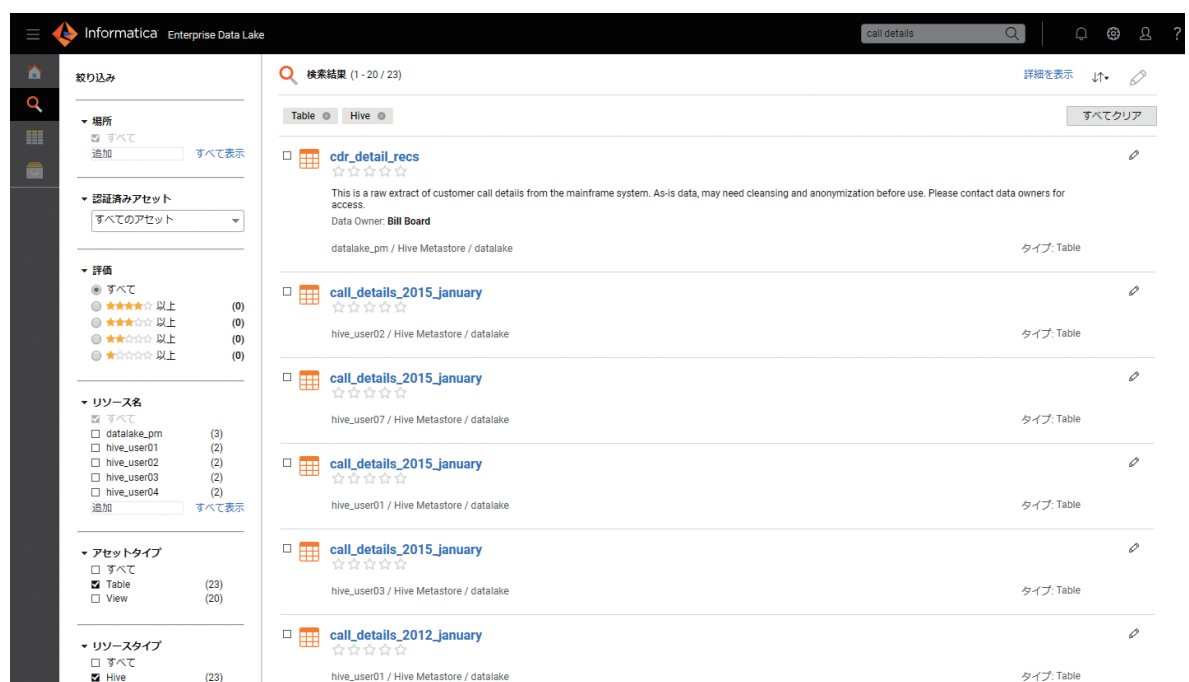
プロジェクトのワークスペースにデータ資産を追加することで作業を整理でき、共同所有者やエディター、閲覧者などさまざまな役割と権限のチームメンバーをプロジェクトに追加して、他のアナリストとのコラボレーションを実現できます。

データ資産に関する推奨

他のユーザーの行動履歴や共有された知識に機械学習アルゴリズムを適用した自動推奨機能によって生産性が高まり、信頼性の高い資産の再利用が促進されます。プロジェクトに追加されたデータ資産に基づいて、他の代替となるデータ資産や、追加で活用できるデータ資産が自動で提案されます。

ウィザード形式のデータアップロード

ウィザード形式のインターフェイスを使用して、個人がローカルに所有しているファイルをデータレイクにアップロードできます。アップロードでは、最適な形式でHiveテーブルが自動的に作成されます。アップロードしたデータをもとに、データ資産を作成または上書きしたり、データを追加したりできます。



スマートセマンティック検索機能と動的ファセットにより、データセットを手早く発見

インフォマティカについて

インフォマティカは、エンタープライズクラウドデータ管理をリードする唯一の企業であり、データ主導型のデジタルトランスフォーメーションを推進しています。企業のイノベーションを加速し、俊敏性の向上や新たな成長機会の獲得によって、インテリジェントな破壊的イノベーションを支援しています。また、インフォマティカは企業のデータの力を25年以上にわたり引き出し、世界で数千社以上の企業がそのソリューションを利用しています。

インフォマティカに関する詳細は、インフォマティカ・ジャパン株式会社（代表：03-6403-7600）までお問い合わせいただくか、Webサイト www.informatica.com/jp/ をご覧ください。

主なメリット

あらゆるデータを発見、アクセス

業務アナリストは、ビッグデータの特徴である「膨大な量、多様なタイプ、スピード」に合わせてデータを管理できる効率的な方法を切望しています。スマートセマンティック検索機能と動的ファセットを使用して結果をフィルタリングすることで、信頼できるデータ資産を簡単に発見することができます。また、自動化された機械学習ベースの探索プロセスでは、アナリストにとって関心が高いと考えられる、新たな関連データ資産が自動で推奨されます。これによって信頼性が大幅に高まると共に、類似したプロジェクトにおいて同じデータセットが重複して生成されることが少なくなります。

ガバナンスによる効率的なコラボレーション

業務アナリストは、プロジェクトワークスペースを活用してデータセットに関するコラボレーションを進めることで、ビッグデータアナリティクスプロジェクトの効率を高めることが可能になります。データセットがプロジェクトワークスペースに追加されると、アナリストは360度ビューでプロファイル統計、エンドツーエンドのデータセットリネージ、全関連データ資産、データドメイン、ユーザーなどを確認できます。これによって、データの品質を評価し、信頼できるデータを共有し、プロジェクトに役立つ他のデータセットを着実に発見できるようになります。また、役割ベースのセキュリティによって、プロジェクトに追加したアナリストだけにデータへのアクセスを認めることができます。

必要なデータを迅速に準備して共有

ビジネスサイクルが短くなり続ける中で、業務上の価値を高める上でデータアナリストが頼りにできる数少ない競争優位性の1つが、「スピード」です。優位な分析を行うのに必要なデータを短時間で準備し、共有することが重要です。インフォマティカのセルフサービスデータ準備機能は、馴染みのある使いやすいExcelライクなインターフェイスを備えており、業務アナリストはデータを迅速に結合、フィルタリング、統合し、必要なインサイトを手に入れることができます。クラウドソーシングしたデータ資産のタグ付けおよび共有機能により、業務アナリストはデータキュレーションプロセスを制御し、業務の生産性を向上させることができます。

データプレパレーションの工程を再利用可能なワークフローに変換

自動化ツールやセルフサービスツールを使用している、アナリストが新しいデータセットに対してまた同じ準備作業を繰り返さなければならないということも珍しくありません。これでは、せっかくの規模や再利用可能性から得られるメリットが無駄になってしまいます。Informatica Enterprise Preparation はデータプレパレーションの工程を記録し、繰り返し実行可能な自動化されたデータ統合処理としてスケジューリングすることができます。これによって、手作業のデータプレパレーションプロセスを、再利用可能で持続的かつ効率的なシステムに転換することができます。



インフォマティカ・ジャパン株式会社

〒105-6226 東京都港区愛宕 2-5-1

愛宕グリーンヒルズ MORI タワー26 階

電話：03-6403-7600(代表) FAX：03-3433-1031

<http://www.informatica.com/jp/>

Informatica パートナー

富士通コンタクトライン(総合窓口) 0120-933-200

受付時間 9:00~12:00 および 13:00~17:30

(土・日・祝日・当社指定の休業日を除く)

Informatica Big Data Management

メリット

- ビッグデータ統合に向けたアプローチを実現
- より多くのソースからデータを統合
- 開発者の生産性を向上
- Apache Hadoopで、より早く、柔軟で再利用可能なデータパイプラインを実行

機械学習を活用して、迅速かつ簡単にビッグデータを取り込む

ビジネススピードが速まり、競争に勝つための変革を進めなければならない今だからこそ、企業のデータアーキテクチャを最新化 / 最適化して、データを意思決定に役立つ戦略的資産に変えるチャンスです。現在は Apache Hadoop により、かつてないほど多くのデータを分析に利用できるようになっています。しかし、未だにハンドコーディングやコード生成などの旧式の手動アプローチにより、ほとんどのビッグデータを活かしきれていません。こうしたアプローチでは多くのビッグデータが分断され、不完全で一貫性のない状態になってしまいます。

より多くのデータソースからデータを迅速に統合

Informatica Big Data Management[®] は、高速で大量なデータの取り込みとデータ統合処理によって、ビジネスアナリストが必要なデータを素早く入手します。あらかじめ組み込まれた数百の高性能コネクタ、データ統合変換機能、およびパーサーにより、Apache Hadoop / NoSQL / MPP アプライアンスなどのビッグデータインフラストラクチャに、ほぼすべてのデータタイプを取り込み、処理します。また、動的マッピング、動的スキーマサポート、およびパラメーター化機能によって、テンプレートに基づいたデータ統合プロセスの自動化を実現します。

本製品は、MapReduce、Spark、Apache Tez、Informatica Blaze など複数の処理エンジンと互換性のあるスマートパフォーマンス最適化により、開発者の生産性や再利用性のある操作、データ統合のパフォーマンスを最大限に高め、ビジネスニーズに合った価値を短期間でもたらします。そして、ビッグデータ統合ソリューションにおける黄金律を提供することにより、より多くのビッグデータをビジネスバリューに迅速に転換することができます。

主な特長

ビジュアル開発インターフェイス

オープンソースコミュニティでは新たなイノベーションが絶え間なく登場し、Hadoop エコシステムは急速に変化しています。Informatica Big Data Management は、オープンソースの Hadoop フレームワークに基づいて構築されており、データパイプライン内にすべての変換ロジックを維持します。そのため開発者は、Hadoop の概念や言語に関する専門知識がなくても、データパイプラインを一度設計すれば、Hadoop の変更のたびに再構築することなく、簡単にデータパイプラインを実装することが可能です。その結果、本番システムへのインパクトやリスクを抑えつつ、Hadoop を素早く導入することができます。

あらゆるデータに接続

あらゆるトランザクションデータ（RDBMS、OLTP、OLAP、ERP、CRM、メインフレーム、クラウドなど）またはインタラクションデータ（ソーシャルメディアデータ、ログファイル、マシンセンサーデータ、Hadoop、NoSQL フォーマット、ドキュメント、電子メールなど）に素早くアクセスします。

HadoopやSparkでの高速データ統合

Hadoop または Spark でネイティブに実行される、あらかじめ組み込まれたデータ統合変換機能のライブラリにより、データをあらゆる規模で処理できます。ビジュアル開発インターフェイスと組み合わせ、すぐに使い始められる高度な変換機能によって、開発者はコードの作成ではなく、ビジネスロジックにより多くの時間を費やすことが可能です。

高速大量取り込みと抽出

Informatica Big Data Management は、大量取り込みとマッピングテンプレートをを用いて、少数の設計パターンに基づき数百のランタイムデータフローを生成します。こうしたデータフローを容易にパラメーター化することで、ビッグデータプロジェクトに共通する Web やマシンログファイルなどの動的スキーマを処理します。つまり、保守が容易でスキーマ変更にも対応できるデータフローを、素早く構築できるようになるのです。

Hadoopでのインテリジェントなデータ解析

Informatica Big Data Management により、Web ログ、JSON、XML、マシンデバイスデータなどの複雑な多構造化／階層化／非構造化／業界標準データに簡単にアクセスして、解析します。また、機械学習によってデータ構造を推測し、再利用するための特別なパーサーを作成します。さらに、市場データや業界標準（SWIFT、ACORD、HL7、HIPAA、EDI など）向けにあらかじめ組み込まれたパーサーも利用可能です。

スマートパフォーマンス最適マイザ

スマート最適マイザは、新しいテクノロジーの登場に合わせてデータパイプラインを再構築することなく、最高のパフォーマンス、拡張性、リソース使用率を実現します。たとえば、YARN 内の Informatica Blaze エンジン、Spark、Hive on Tez、または MapReduce を使用して、データ統合変換をすることができます。また、Informatica Blaze は YARN と統合されたデータ処理エンジンであり、データパイプライン処理、ジョブのパーティショニング、ジョブリカバリー、高パフォーマンスの拡張を可能にします。

SQLからマッピングへの変換

ANSI 準拠 SQL スクリプト、Informatica PowerCenter[®] Pre / Post-SQL、および SQL オーバーライドクエリを、Hadoop で実行される最適化されたビッグデータマッピングに変換することで、再利用性を最大限に高め、保守を簡素化し、エンドツーエンドのデータリネージを維持します。

柔軟なサーバーレス実装

オンプレミスかオフプレミスかを問わず (Amazon Web Services Elastic MapReduce や Microsoft Azure HDInsight)、分散するリソースを自動的に実装および管理します。

主なメリット

あらゆるデータの取り込みが可能に

あらかじめ組み込まれた広範なコネクタライブラリは、あらゆるデータの取り込みを可能にします。そして、データ統合ではなく新しいデータインサイトに焦点を絞り、革新的な製品やサービスの開発に活用することができます。

場所を問わずにデータを処理して提供

データを管理する担当者は、次世代のデータインフラストラクチャが蔓延する問題に直面しています。複数のオンプレミスおよびオフプレミスの Hadoop 配信を認定およびサポートしているのはインフォマティカだけであり、組織による場所を問わないデータの処理と提供を可能にします。

迅速な実行

複数のエンジンにわたって最適なランタイム処理とシンプルなモニタリングを実現し、開発と処理のスピード、柔軟性、再利用性を高めます。

インフォマティカについて

インフォマティカは、エンタープライズクラウドデータ管理をリードする唯一の企業であり、データ主導型のデジタルトランスフォーメーションを推進しています。企業のイノベーションを加速し、俊敏性の向上や新たな成長機会の獲得によって、インテリジェントな破壊的イノベーションを支援しています。また、インフォマティカは企業のデータの力を25年以上にわたり引き出し、世界で数千社以上の企業がそのソリューションを利用しています。

インフォマティカに関する詳細は、インフォマティカ・ジャパン株式会社（代表：03-6403-7600）までお問い合わせいただくか、Webサイトwww.informatica.com/jp/をご覧ください。



インフォマティカ・ジャパン株式会社

〒105-6226 東京都港区愛宕 2-5-1

愛宕グリーンヒルズ MORI タワー26 階

電話：03-6403-7600(代表) FAX：03-3433-1031

<http://www.informatica.com/jp/>

Informatica パートナー

富士通コンタクトライン(総合窓口) 0120-933-200

受付時間 9:00～12:00 および 13:00～17:30

(土・日・祝日・当社指定の休業日を除く)

Informatica Cloud – Cloud Data Integration

クラウドやオンプレミス環境のデータ連携・同期を一括処理する高性能なクラウドデータ統合

メリット

- 業務ユーザーがセルフサービスで使える操作性を提供し、ビジネス環境の素早い変化やビジネスプロセスの変化に対応
 - Amazon、Google、MS Azure、Salesforce、Marketo、Workdayなどの様々なクラウドサービスに対応した専用コネクタを提供、あらゆるデータへのアクセスが可能
 - 各種システムの環境をサポートした Informatica Cloud を利用することで、統合開発の生産性、パフォーマンス、拡張性を高めることが可能
 - クラウドとオンプレミスをつなぐハイブリッドアーキテクチャーを採用しており、オンプレミス、クラウド間をセキュアに接続
- SaaSアプリケーションが幅広く普及したことにより、IT部門だけでなく、業務部門のアプリ利用者や担当者もシステム間のデータ連携やリアルタイムのアプリケーション統合プロセスを作成する機会が増えています。また、データウェアハウスを始めとしたプラットフォームのクラウド化やビッグデータなど様々なデータ活用の普及から、かつてないほど膨大で多様なデータを統合するニーズが高まっています。
- 業務ユーザーにも使いやすいウィザード形式の Task Wizard により、データ同期や複製など比較的シンプルな連携要件を簡単に定義、実行
 - 開発ユーザー向けの Mapping Designer により、事前定義された変換部品を利用し、デザインキャンバス上でデータソースとターゲット設定や、データの結合、並べ替え、条件分岐、その他高度な変換をドラッグアンドドロップで簡単に設計
 - Salesforce や Workday などのクラウドアプリケーションに入力されたリアルタイムのデータを、Oracle や SAP といったオンプレミスに存在するデータソースと簡単に統合可能
 - データ複製サービスにより、Salesforce などのデータを簡単に他のクラウドサービスやオンプレミスのデータベースへ迅速に複製
 - Salesforce 向け Bulk API により、高スループットのデータ統合を実現
 - グローバルリポジトリを通じて、Mapping Designer で作成した統合テンプレートを全社で共有
 - REST API を通じて、Informatica Cloud の統合サービスを外部スケジュールに公開
 - SOAP や REST などの Web サービス間で、条件付きトリガーによるサービスのオーケストレーション（最適な運用管理）を実現
 - API 管理のフレームワークで、多様な統合パターンに応じた API を作成、活用を支援



Cloud Data Integration エディション別機能

		Basic	Advanced	Premium
データ統合	ハイパフォーマンスなバルクデータ統合	✓	✓	✓
リアルタイム統合	リアルタイムアプリケーション統合、サービスオーケストレーション、プロセスオートメーション	オプション	オプション	✓
ハイブリッド統合	Informatica Cloud と PowerCenter の相互連携		✓	✓
API	REST/SOAP API サービスの定義、生成、利用		✓	✓
B2B	EDI、HL7、HIPAA、SWIFT などの非構造化データフォーマットの利用			オプション
データ品質管理	クレンジング / 名寄せなどのデータ品質管理		オプション	オプション
運用管理	データ統合処理に関するログやステータス管理、ユーザー権限管理、運用モニタリング	✓	✓	✓
高度な管理	詳細なアクセス制御、管理の委任、ユーザー/グループ管理など高度な管理		✓	✓
Tier 1 コネクタ	接続コネクタ	1	1	2
Tier 2 コネクタ	接続コネクタ	1	1	3
Cloud Designer	開発者向け開発 Web 画面		✓	✓
Sandbox	開発・検証環境		✓	✓
Secure Agent	データ統合処理の実行エンジン (4 コア CPU 分)	✓	✓	✓
グローバルリポジトリ	開発者間のコラボレーション、開発資産の再利用	✓	✓	✓

インフォマティカについて

インフォマティカは、エンタープライズクラウドデータ管理をリードする唯一の企業であり、データ主導型のデジタルトランスフォーメーションを推進しています。企業がデータの力を引き出すことにより、俊敏性の向上、新たな成長機会の獲得、新しいソリューションの開発を実現するための洞察を通じて、インテリジェントに市場へ破壊的イノベーションをもたらし、ビジネスをリードできるよう支援します。

インフォマティカは、エンタープライズ向けクラウドデータ管理ソリューションにおける信頼性の高いリーディングカンパニーであり、世界 7,000 社以上の企業がそのソリューションを利用しています。

インフォマティカに関する詳細は、インフォマティカ・ジャパン株式会社（代表：03-6403-7600）までお問い合わせいただくか、Web サイト www.informatica.com/jp/ をご覧ください。

PowerCenter と Informatica Cloud のハイブリッド統合

PowerCenter と Informatica Cloud は相互運用が可能なため、Informatica PowerCenter 開発ツールで作成した統合処理ロジックを Informatica Cloud にインポートして使用できます。インポート可能な統合処理ロジックには PowerCenter マッピングやマップレット、テンプレート、プラグインなどが含まれます。

また、PowerExchange for Cloud Applications により、PowerCenter から Informatica Cloud のコネクタを透過的に利用し、各種 SaaS などのサービスに接続し、PowerCenter でハイブリッド統合処理を実行することも可能となります。



Informatica™

インフォマティカ・ジャパン株式会社

〒105-6226 東京都港区愛宕 2-5-1

愛宕グリーンヒルズ MORI タワー 26 階

電話：03-6403-7600(代表) FAX：03-3433-1031

<http://www.informatica.com/jp/>

Informatica パートナー

富士通コンタクトライン(総合窓口) 0120-933-200

受付時間 9:00~12:00 および 13:00~17:30

(土・日・祝日・当社指定の休業日を除く)