

# 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<sup>®</sup> 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オプティマイザにより、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化

スマート最適化は、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<sup>®</sup> 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 Informatica 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/](http://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)