

パンデミックに迅速かつ効果的に対処するためのデータ管理基盤

メリット

- ・住所、電子メール、電話番号データの検証とエンリッチ化
- ・データ環境全体を自動的にカタログ化
- ・データのインテリジェントプロファイリングにより、最高品質のデータを大規模に検索可能
- ・あらかじめ組み込まれた65を超えるデータ品質アクセラレータをクラウドまたはオンプレミスで活用
- ・ファイル、データベース、ストリーミングソースからペタバイト（PB）規模のデータを高速取り込み
- ・CCPA/GDPRへのコンプライアンスを確保
- ・機密性の高いPIIデータを検出して保護

データドリブンなDXで将来の危機に備える

組織の危機対応の成功は、関連する大量のデータを管理できるかどうかによって大きく左右されます。自然災害や経済破綻、また感染症のアウトブレイク（大流行）でも、対応するにはデータを資産として活用する能力が重要です。例えば、新型コロナウイルス感染症（COVID-19）の世界的なパンデミックのような危機に対応するには、政府機関は景気刺激対策資金の給付のために、医療機関は接触者の追跡のために、また高等教育機関は安全な学校再開のために、さまざまなデータを活用する必要があります。

適切なツールを導入して、多種多様なソースからのデータの流入を生産的に管理すれば、目下の危機に対してタイムリーで総合的な幅広い対応を調整し、回復への道を整え、将来の出来事に備えることができます。組織では、データの収集および保存方法を確立するとともに、データサイエンティストからポリシーの決定者まで誰もが高品質データに基づく分析や意思決定を行えるようにする必要があります。

使用に適したデータの確保

スプレッドシートやデジタル化されていない手作業でレポートを作成すると、ポリシーの決定に利用する予測モデル、人工知能（AI）、高度アナリティクスツール、ダッシュボードに提供するデータの鮮度や精度が損なわれるリスクが高まります。その結果、古いデータや不完全なデータに基づいて方針を決定することになり、対応の遅れ、不適切なリソース配分、社会的信頼の低下などにつながります。

使用に適したデータを確保する手段を整えれば、このようなリスクを避けられます。情報の探索、収集、クレンジング、ガバナンス、カタログ化のベストプラクティスに従い、組織全体にわたりプロセスを標準化することで、よりクリーンなデータセットを獲得することが可能になります。このため、品質、完全性、正確性に優れたデータに基づく信頼性の高いインサイトを意思決定に使用できます。

危機のさなかでは、こうした標準化は基本的には不可能な場合もあります。例えば、新型コロナウイルスの感染拡大では接触者の追跡を必要とする組織は多岐にわたります。多くの労力を要する手作業で追跡を行う政府機関やBluetooth対応アプリを導入する職場や学校、特注の追跡ソリューションを実装する医療機関などです。最新のデータ管理ソリューションでは、収集方法、データソース、データ形式の違いを調整し、異常値や重複データを自動的に識別して排除することができます。

場合によっては、収集したデータに含まれている情報が欠けていたり低品質だったりすることもあります。その場合、間違いは修正し、不完全な入力に補完する必要があります。最新のデータ管理システムには、データクレンジングツールも備わっています。このツールでは、統計分析を活用して、あらかじめ定義されたパラメータに基づきデータセットを検証します。パラメータに適合しなかったデータは、人間によるチェックに回されます。データのクレンジングに機械学習とAIを利用するシステムもあります。これらのシステムではプロセスが高速化され、不完全なデータセットがインピュテーションなどの統計処理によって自動的に補完されます。

データ品質管理、データガバナンス、メタデータ管理を自動化することで、流入する大量の構造化データと非構造化データ（さまざまなデバイスやセンサーからのIoTデータなど）をより適切に処理できます。さらに、危機対応、回復、将来への備えの包括的な計画をより効果的に作成して実行できます。

主な機能

個々の状況によっては、綿密なデータ管理危機対応戦略に広範なデジタルトランスフォーメーションが必要になる場合があります。特に、サイロ化したレガシーアプリケーション/システムを利用している場合は必要です。ただし、どの組織でも一定の手順に従うことでデータ管理プログラムを迅速に開始することができます。例えば、リアルタイムの情報へのアクセスは、十分な情報に基づく意思決定に欠かせません。自然災害や感染症のパンデミックのような危機では、状況が急変することがあります。避難所や適切な医療といった緊急のニーズに応じて必要な資源の配分を判断するためには、全体的な状況を可能な限り迅速に把握できなければなりません。

他に重要なデータ管理機能としては、次のようなものがあります。

自動的にデータを探索して準備

連絡先データ（一般市民や従業員の住所や電話番号など）を検証およびエンリッチ化します。また、電子メールアドレスの正確性に加え、送信時の安全性も素早くチェックします。自己関連と修正の機能もあります。ブラックリストとスパムトラップを自動検知する機能により、電子メールアドレスが正規のものであることを確認できます。グローバルジオコーディングを活用して、緯度と経度を正確に特定し、トレンド、ホットスポット、その他の問題を把握します。[Informatica® Data as a Service](#) (DaaS) では、電話番号のレコードが携帯電話の番号か固定電話の番号かを簡単に識別できます。これは、SMSアラートの受信者を登録する場合に重要です。APIベースの大容量SMS通知サービスでは、送受信メッセージとあらかじめ組み込まれたオプトイン/オプトアウトフローがサポートされ、SMSメッセージとSMSアラートをタイムリーに送信できます。

すべての構造化データ（スプレッドシートの入力情報など）と非構造化データ（文書など）を、コンテキストやグロッサリとともに、標準化およびカタログ化します。[Enterprise Data Catalog](#)のAIと機械学習の機能を使用して、クラウドおよびオンプレミスシステムでのデータの探索と把握を自動化することで、検索と取得を整流化します。インテリジェントプロファイリングにより、Google式の検索機能を使用して、最高品質のデータを大規模に検索できます。

大規模にデータを取り込んでクレンジング

オンプレミス環境、ハイブリッド環境、マルチクラウド環境（Microsoft Azure、Amazon Web Services、Google Cloud Platformなど）において、ファイル、データベース、ストリーミングソースのペタバイト（PB）規模のデータを高速で取り込んで統合できます。統合の新しいパターンや複雑なパターンへの容易に対応し、ミッションクリティカルなプロセスの、高性能で高い信頼性の準ユニバーサルな接続性を実現します。

Informatica Data Quality Accelerator for Crisis Responseを活用できます。このアクセラレータには、[Informatica Data Quality](#)と[Cloud Data Quality](#)の両方で機能する再利用可能な65以上のデータ品質ルールがあらかじめ組み込まれています。高速データクレンジングのサポートにより、オンプレミスとクラウドでのデータ処理が加速されます。データ品質ルールで要件に対応できない場合（値の欠落など）や、機密データをマージする場合は、高度な例外管理を通じてチーム全体にタスクを配布することで問題を解決できます。

Data Quality Accelerator for Crisis Response	
医療	診断結果の文字列から併存疾患を解析
	文字列からICDコードを解析
	医療機関施設名を標準化
	施設名から医療機関施設の種別を解析
地理空間	2つの座標から距離を算出
	国を特定
	国名を標準化
	ポイントがポリゴン内かどうかを確認
住所	米国の郡を検証
	米国の州を検証
	米国の郵便番号を検証
電子メール	文字列から電子メールアドレスを解析して検証
	電子メールアドレスの形式を検証
電話番号	米国の電話番号を解析
	米国の電話番号を標準化
	米国の電話番号の市外局番を検証
	米国の電話番号を解析、標準化、検証

氏名	対象者の氏名を特定
	愛称を正式名に変換
	文字列から複数の名前を解析
	姓と名の順序を解析
連絡先の 検証と重複排除	米国の運転免許証を検証
	米国の世帯連絡先データを検証
	米国の連絡先データの重複を削除
	人種を標準化
	性別を検証
日時	日付の完全性を確認
	文字列から日付を解析
	日付を検証および標準化
	日付間の日数を算出
数値	正数か負数かを検証
	数字を検証
	数字から文字列を解析
文字列解析	文字列の末尾または先頭の数字を解析
	正数を検証
	印字不能の文字と句読点を削除
	アクセント記号付き文字を通常文字に変換

図1：Informatica Data Quality Accelerator for Crisis Responseが基本的なデータ処理タスクの作業時間を大幅に短縮

データをマスター化し、関連付けて、唯一無比のソースを構築

トランザクション、インタラクション、その他一般のデータタイプを組み合わせて、既存のデータレイクやHadoop環境のあらゆるデータソースからの数十億のレコードを管理します。非構造化データの場合は、自然言語処理を使用して、一般市民、従業員、学生、患者の属性（健康上のリスク、予後、ライフイベントなど）を推論します。[Informatica Master Data Managementソリューション](#)では、すぐに使い始められる機能により、一般市民データ、患者データ、受給者データ、ソーシャルデータなどをリンクさせ、機械学習を使用して関係を構築することができます。関係、接点、家族／世帯、ソーシャルネットワーク、組織階層をグラフデータストアで大規模に視覚化します。

GDPR（EU一般データ保護規則）およびCCPA（米カリフォルニア州消費者プライバシー法）へのコンプライアンスの確保と、機密性の高いPII（個人を特定できる情報）の検出および保護も可能です。[Informatica Data Masking](#)のようなデータセキュリティソリューションでは、機密データを匿名化、マスキングすることで、アプリケーションユーザー、ミッションインテリジェンス、アプリケーションテスト、アウトソーシングによる不注意の漏えいや不正アクセスから機密データを保護します。

主なメリット

信頼できる意思決定基盤を提供

データの品質と正確性を確保することで、利用できる最良の情報に基づいた意思決定を確実に行えます。質が低いデータに基づく意思決定は、悪意がなくても見当違いの結果になり、社会的信頼の低下や不適切なリソース配分につながります。例えば、救命のための医薬品や保護具、検査機器の優先順位決定と配分を適切に行えないなどの問題が生じます。最も重要なのは、データの完全性が確保されていて、収集されるはずのデータと実際に収集されたデータに違いがないことです。重複するデータセットを自動排除する機能を使うことで、一定の限度がある帯域幅とデータストレージ容量を最大限に活用できます。

重要なタスクの実行を最適化

パンデミック時における接触者追跡のような重大なタスクを遂行する場合、感染者（および感染者と接触した可能性のある個人と場所）に関して収集した住所データを自動で検証およびエンリッチ化できます。これは、感染の疑いのある人の連絡や登録を確実に行うのに役立ちます。同様に、電子メールと電話番号の検証により、感染の疑いのある人に重要なメッセージや詳細情報を届ける際の確実性を高められます。電話番号の種類を識別することで、SMSアラートおよび通知プログラムに登録されている可能性がある個人を特定して、関連詳細情報を効率的に送信できます。

包括的なリアルタイムのリスク評価を作成することも可能です。例えば、影響を受けた地区のヒートマップを作成し、データポイント（感染の疑いのある人が検査を受けた、移動した、隔離されたなど）を追跡するとともに、感染のリスクがある他の一般市民に通知します。リスクにさらされている個人を複数のサービスにわたり管理するために、最低限のプロファイルを医師、保険会社、高等教育機関、その他の組織と共有することもできます。これにより、あらゆるニーズに対して質の高いサポートを提供できます。

不正を防止し、一般市民のプライバシーを保護

危機状況下では、多くの政府は危機対応の一環として緊急対策資金を給付し、不正、乱用、浪費を防ぐための報告と監視の要件を設けます。給付対象者は、一般市民、医療機関、教育機関、企業などです。目的は、危機対応リソースの確保や逸失利益の補填、危機下での支払能力の維持、危機後の回復の促進など多岐にわたります。給付対象者に関連付けられたすべてのレコードを、検索レコード内または機関のファイルのデータ内にエラーやバリエーションがあっても、リアルタイムで検索および取得できる機能があれば、住所、所在地、電話番号、過去の給付金申請、監査、調査などのデータを含む給付対象者の完全なプロファイルを確認することが可能です。この全体像の把握は、不完全なデータに基づく不適切な意思決定のリスクを低減するのに役立ちます。浪費や不正を低減することで、政府機関は適切な受益者に税金が使われていることについて国民からの信頼を得られます。

一般市民に関するデータの全体像を把握することで、米情報公開法（FOI）や裁判所命令などからの重大な要求への対応が容易になるだけでなく、他の部門に状況の変更を知らせやすくなります。また、一般市民に給付金の変更や新しい制度について通知しやすくなります。政府機関、医療機関、高等教育機関は機密データを保護するセキュリティ対策を実施することにより、社会的信頼を高め、公衆安全対策への協力を促進する環境づくりができます。

インフォマティカについて

デジタルトランスフォーメーションによって我々の期待値が変化しています。より良いサービスを、素早く、便利に、低コストで利用したいという期待が高まっているのです。企業も状況に応じて変化する必要があります。そしてそのヒントは「データ」にあります。

エンタープライズ向けクラウドデータ管理で世界をリードするインフォマティカは、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、あらゆる産業や分野の企業がインテリジェントに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、あらゆるデータを徹底的に重視し、企業の成功に必要とされる汎用性を提供します。

インフォマティカは、企業がこれからのインテリジェントな破壊的イノベーションを推進できるよう、当社が提供するあらゆるサービスを通じてデータの力を継続的に引き出すことを支援します。

次のステップ

ここで紹介したさまざまな機能を出発点として、危機への初期対応、回復、準備を整えるためのデータ管理戦略を策定して実行できます。将来にわたり高い回復力を確保するためには、企業・組織は自動化したインテリジェントなデータ管理を実装する包括的なデジタルトランスフォーメーションを実現する必要があります。詳細については、[こちら](#)をご覧ください。



Informatica

〒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：03-3433-1021

IN17_0620_03911

© Copyright Informatica LLC 2020. Informatica、Informaticaロゴは、米国およびその他の国におけるInformatica LLCの商標または登録商標です。インフォマティカの商標の最新版は、<https://www.informatica.com/jp/trademarks.html>をご覧ください。その他すべての企業名および製品名は、各社が所有する商号または商標です。本文書に記載されている情報は、予告なく変更されることがあり、現状のまま提供され、明示または黙示を問わず一切の保証を伴いません。