

スマートマスタリング

質の高いデータとガバナンスのためのアジャイルかつ効果的なアプローチ

クイックスタートガイド ・ 2021年5月

MarkLogic のスマートマスタリングクイックスタートサービスは、主要コンポーネントを使ってマスターデータ管理の問題を解決するものです。私たちのサービスは、情報管理の最も困難な要素に対してユニークなアプローチを提供し、完全に統合されたデータの「360 ビュー」（全体像）を提供します。

スマートマスタリング

スマートマスタリングは、MarkLogic データハブの一環として、異なるソースからのデータをマッチ&マージするという課題を取り扱います。従来、この課題はマスターデータ管理(MDM)ソリューションによって扱われています。しかし、長年使われてきている従来の MDM ソリューションは、変化し続けるビジネス要件や多様なデータへの対応に時間がかかり融通が利きません。

MarkLogic データハブフレームワークの機能の 1 つであるスマートマスタリングは、これまでとは異なるアーキテクチャや手法でデータの質を改善し、ガバナンスを強化し、ビジネスバリューを生み出します。私たちの技術では、事前定義されたスキーマや ETL がなくても分散するサイロからのデータを読み込むことができます。特許取得済みの検索インデックスを利用することにより、MarkLogic はサイロに分断されていたあらゆるデータソースに対して容易にクエリを実行できます。大きな初期投資なしで、データの質の改善、オーナーシップの割り当て、データアセットの収益化などを同一のプラットフォームで実現できます。

それではまず何をしたらよいのでしょうか。MarkLogic のコンサルティングは、データハブフレームワークの効果的な利用に関するベストプラクティスを提供し、お客様の環境においてスマートマスタリングが活用できるように準備します。さまざまな業界（金融 / 出版 / ヘルスケア / 情報機関）における複雑なデータガバナンスやデータの質の問題に長年取り組んできた経験を、ビジネスルールとして表現することで、ユーザー / 場所 / 製品といったビジネスエンティティを自動的に特定 / スコアリング / マッチ & マージします。エンドユーザーが今抱えている重要なビジネス課題にすぐに取り掛かることができ、事前の分析やデータモデリングは不要です。このアプローチでは事前の準備が最小限で済み、結果が数週間で得られます（月や年単位ではなくて）。ここで構築されるシステムは、すぐに拡張でき、設定は変更可能です。このためこのクイックスタートは完成後のリスクが小さく、将来に備えるものとなっています。

スマートマスタリングクイックスタート

「スマートマスタリングクイックスタート」コンサルティングサービスは、短期間でこの基礎を築き、このアプローチを実証します。また今後も時間を掛けて拡張していけるシステムを構築します。

クイックスタートのフェーズは、データハブパターン処理ステップに対応しています。これらのデータフローステップは、MarkLogic の一般的なデータの読み込みおよび業務への実装作業であり、データマスタリングの中核的処理であるマッチ&マージも含まれています。



私たちが行うこと

1. 主要なユースケースを特定し、そのために必要なAPIを定義する

ビジネスにおいて最初に要求されるのは、ポテンシャルのあるユースケースを複数特定し、これらにビジネス目標に応じた優先順位をつけることです。例えば、既存のプロジェクトにキューレーション済みデータを提供するプロジェクトや、完全に新規の MDM プロジェクトなどがあります。これらのユースケースごとにビジネスエンティティ（顧客 / 製品 / アセット / 場所など）を選びます。エンティティはこれらのプロジェクトの中核となります。これに基づいて API を定義し、明確なデータ要件を作成し、マッチ&マージ用のビジネスルールを決定します。MarkLogic コンサルタントは、各ステップにおいて技術的な指導を行い、このクイックスタートが確実に成功するようにお手伝いします。

2. これらのAPIにおいて必要とされるデータを一元化されたハブに読み込む

データソースを特定したならば、私たちのコンサルタントが MarkLogic に備わった業界標準のエンドポイント（REST など）の使い方と、特定されたソースシステムからデータを素早くかつ簡単に読み込むためのツールをご紹介します。御社のチームと一緒に、データディスカバリーのプロセスを素早く数日（数か月ではなく）で行います。またスマートマスタリングのプロジェクトでは、事前にデータモデルを完成させたり複雑な変換を行わなくてもよいことを説明します。

3. マッチ&マージのルールを定義・設定する

データの読み込みが終わったならば、マッチ&マージのステップの一部として、人 / 製品 / 薬 / 設備といったビジネスエンティティをマスタリングして 360 ビュー（全体像）の作成をお手伝いします。

これを実現するため、MarkLogic コンサルタントはレコード間の完全マッチ（およびほとんどマッチするもの）に関するさまざまな手法用にマッチング規則および重みをセットアップします。一番シンプルなのは ID によるマッチですが、ほとんどのシステムではこれに加えてファジー（曖昧）検索、ニックネーム、スペルミス、短縮形などを使ったマッチングも必要です。例えば、Person エンティティにおいて「William」と「Bill」はテキスト的には別のものですが、マッチさせる（同一のものとして扱う）ためには、マッチングアルゴリズムでニックネームのリストを使用できます。あるいは他のデータフィールド（誕生日や電話番号）も利用できます。

私たちのスマートマスタリング機能では、2つの対象がほぼ同じである（閾値を超えている）場合、自動的にマージできます。またマッチがそれほど確実ではない場合は、後から別途処理するようにも設定できます。それでは、どうしたら William と Bill を同一人物だと判断できるのでしょうか。姓、名、誕生日が一致したら十分でしょうか。このプロセスでは、私たちは皆さんと一緒に、御社における自動マージのビジネスルールを一緒に決めていきます（スコアリングアルゴリズムおよび閾値を含む）。またスコアが低いマッチにおける目検プロセスの導入についてもアドバイスさせていただきます。

ここでの目的は、データの 360 ビューを得ることです。その際にオリジナルソースおよび値を保持し、リネージや監査、コンテキストで活用できるようにします。

4. リアルタイムAPIおよび一括エクスポートによりデータを業務に実装する

MarkLogic のマスターデータシステムは、大量処理に対応して拡張し、またリアルタイムアクセスに対応できるという点でユニークです。また私たちの方から、MarkLogic サーバーの既存の REST API や SQL インターフェイスを使って、業務アプリケーションおよびレポート / 分析 / BI ツールで簡単に利用できるよう、データのマスタリングを行う方法をご紹介します。単一のデータベースおよび API サービスによるデータのキューレーションおよびリアルタイムのデータ活用は、業務の成長およびデータ需要に基づき、毎秒数十万件のリクエストへまで拡張できます。

成功への主な要因

「素早く軽く」行うというのが、私たちのアプローチです。このため大量のリソースや事前のデータモデリングは不要です。とはいえ、スマートマスタリングクイックスタートを成功させるうえで、重要なことがいくつかあります。

1. 最重要ユースケースを特定・定義する：私たちの「API ファースト」のアプローチでは、ビジネス価値が最も高いものを最初に開発します。ここにおいて、重要なデータ処理、モデリング、ビジネスロジックに注力しながら、後のフェーズで行われる他の作業の基盤も作っていきます。これによりアジリティと短期間でのデリバリーが可能となります。
2. データへのアクセス：完全なデータモデルは不要ですが、上流システムからのデータは必要です。
3. 統合された強力なチーム：御社チーム内のデータエキスパート / 開発者 / 業務担当者がデータを特定・定義し、結果を検証して、作業の優先順位付けを行う必要があります。クイックスタートの期間内でこのチームに対してオンザジョブトレーニングを行い、将来の作業を自分たちだけでできるようにしていきます。

私たちのやり方

MarkLogic コンサルティングとの活動に必要な上記の内容が準備されたら、スマートマスタリングプロジェクトでは最初の 30 日間に以下を行います。

1. プロセスおよびコミュニケーションの仕組みを定義・実装して、御社のビジネス分析 / 検証チームと効果的に連携できるようにします。
2. クイックスタート用に主要なユースケースを 1 つ（あるいは複数）選択して、ビジネスエンティティ / API 要件を特定します。
3. API 要件から遡って、このユースケースに必要なデータを特定します。
4. クイックスタートプロジェクトを技術的に指導します。御社チームに私たちのコンサルタントをビジネスアナリスト / 開発者 / データエキスパートとして配置します。これは短期間のデリバリーおよび効率的なナレッジトランスファーを目的とするものです。
5. MarkLogic をインストールして設定します。あるいは御社の運用チームと協力してインストールと設定を自動化します。
6. 重複解消 / マージ / 出自 / リネージ把握のためのデータマッピング / マッチ & マージルールを作成します。
7. 事前定義されたユースケース（業務システム / カスタマーポータル / 分析ダッシュボード / CRM システムなど）用の API を提供します。

30 日で、MDM システム稼働に関する上記のような基礎が実現されます。このクイックスタートは、短期間でのデリバリーを目的として対象を絞っていますが、今後、より多くのビジネスエンティティ / API / データソースに拡張できます。

MarkLogicのコンサルティングを選ぶ理由

MarkLogic のコンサルティングサービスチームは、連日データ統合に取り組んでいるエキスパート集団であり、一般的な課題とその解決策に関する知見があります。変更データの把握、マスターデータ管理、チーム編成、ハードウェア選定、クラウド移行、パフォーマンスチューニングなど、プロジェクトをスムーズに進め、取り組みの成功に必要な実績や方法論を持っています。

システムを短期間で立ち上げるだけでなく、御社チームに必要なスキルが備わるようお手伝いします。データやシステムのプロフェッショナルが集まった最高のチームであっても、MarkLogic のアジャイルな新しいアプローチ（事前モデリングなしでデータを読み込み / アジャイルなデータ統合処理 / 「データファースト」 から「API ベースの開発」へのシフトなど）に慣れるのに少し時間がかかるかもしれません。

MarkLogic コンサルティングは、これまでのプロジェクトから個人および部門全体としてスキルとベストプラクティスを習得しており、お客様の担当チームを喜んでお手伝いさせていただきます。

スマートマスタリングクイックスタートについてのお問い合わせは、MarkLogic-JP@marklogic.com までお願いいたします。

© 2021 MARKLOGIC CORPORATION. ALL RIGHTS RESERVED. このテクノロジーは、米国特許番号 7,127,469B2、米国特許番号 7,171,404B2、米国特許番号7,756,858 B2、米国特許番号7,962,474 B2で保護されています。MarkLogicは、米国およびその他の国におけるMarkLogic Corporationの商標または登録商標です。本書に記載されているその他の商標は、各企業の所有物です。

マークロジック株式会社 MARKLOGIC K.K.

150-0001 東京都渋谷区神宮前1-5-8 神宮前タワービルディング 13F
03 4540 0337 | jp.marklogic.com | MarkLogic-JP@marklogic.com



150-0001 東京都渋谷区神宮前1-5-8 神宮前タワービルディング 13F
03 4540 0337

jp.marklogic.com | MarkLogic-JP@marklogic.com