



# Smart Mastering

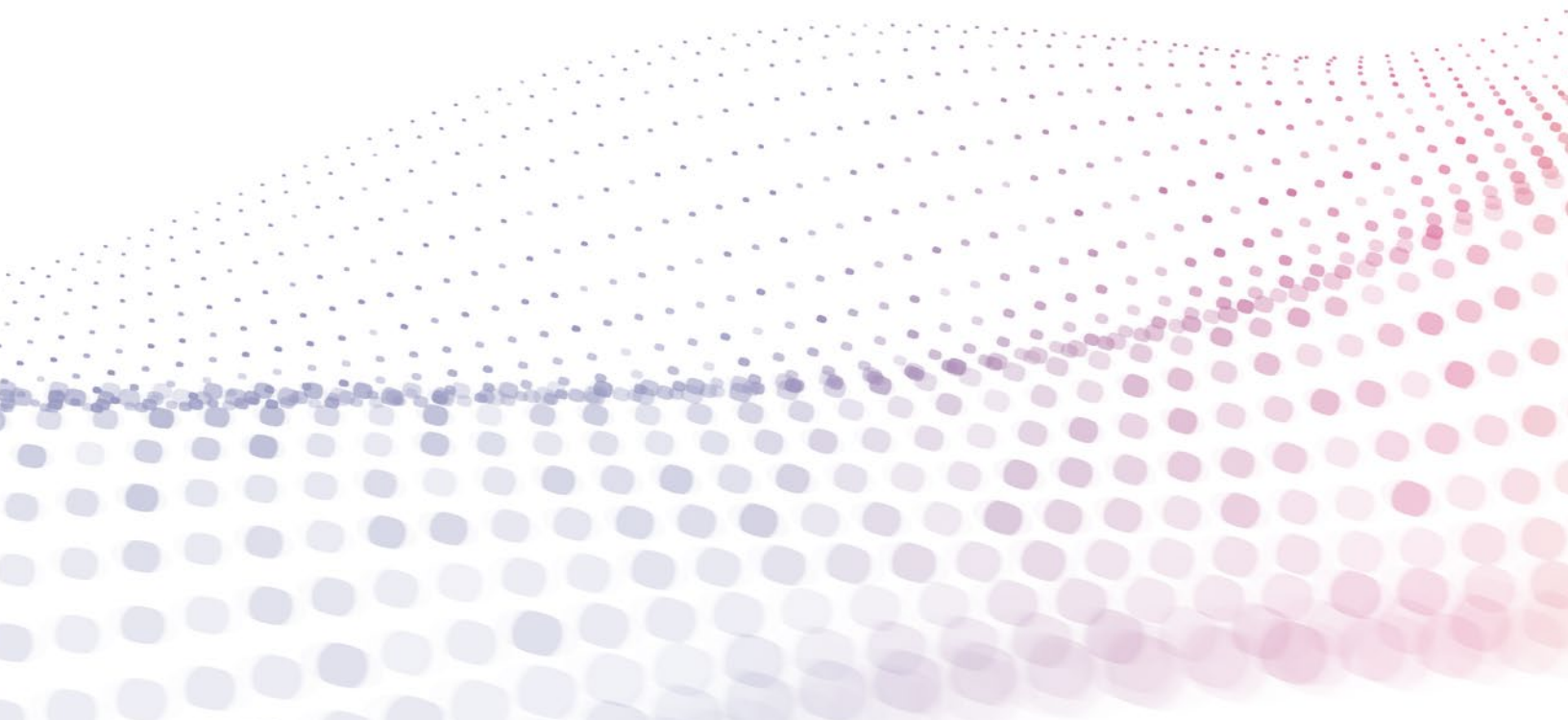
## Agile and Effective Approach for Data Quality and Governance

---

QUICK START GUIDE · JUNE 2020

---

The MarkLogic Smart Mastering Quick Start service is designed to get you up and running with the major components to tackle your master data management challenges. Our service provides a unique approach for one of the most challenging aspects in information management, creating a fully integrated 360-degree view of your data.



# Smart Mastering

Smart Mastering addresses the challenge of matching and merging data from different sources, as part of the MarkLogic Data Hub. Traditionally, this challenge is tackled by master data management (MDM) solutions. However, traditional MDM as a solution pattern has a long history for being slow and brittle in addressing the ever-changing requirements of the business and variety of data sources being used.

As a feature of the MarkLogic Data Hub Framework (DHF), Smart Mastering provides a different architecture and methodology to improve data quality, enforce governance, and create business value. Our technology can ingest disparate data silos without requiring predefined schemas or ETL. By leveraging our patented search indexes, MarkLogic makes it trivial to query across all data sources that were previously siloed. Without huge upfront investments, you can improve data quality, assign ownership and capitalize on your data assets, all from the same platform.

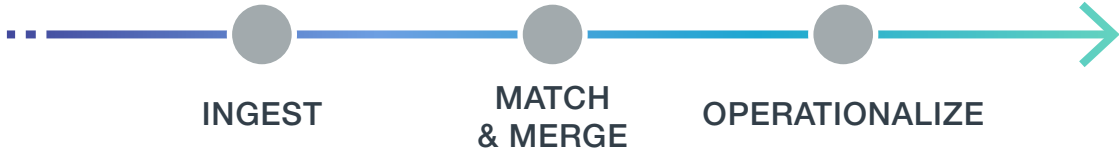
How to get started? MarkLogic Consulting provides best practices to effectively use the DHF and put Smart Mastering to work for you. We translate years of experience with complex data governance and quality issues in various domains (finance, publishing, healthcare, intelligence) into business rules to automatically identify, score, match and merge users, locations, products and other business entities. We start quickly and focus on initial, critical business objectives for end users, instead of getting mired in up-front analysis and data modeling. This approach minimizes upfront investment, producing results in weeks, not months or years. The resulting system is quick to extend and configuration-driven, enabling your team to move forward with less risk after the quick start is complete.

---

## Smart Mastering Quick Start

The Smart Mastering Quick Start consulting service is a short-term engagement to lay this foundation and prove out the approach, leaving behind a working system that can be extended over time.

The phases of the Quick Start engagement mirror the processing steps in the Data Hub pattern. These data flow steps are our typical data ingestion and operationalization, with matching and merging as the core processing phase to master data.



# What We Will Do

---

## 1. Identify Key Use Cases and Define APIs That Will Satisfy Them

First order of business is to identify potential use cases, and prioritize according to business objectives, either to support existing projects with curated data, or to kick-start your greenfield MDM effort. Based on the use cases, a business entity is chosen, such as a customer, product, asset, or location. This entity is central in our effort; it lays the groundwork for the API definitions, creates a clear set of data requirements and drives the business rules for matching and merging. MarkLogic consultants will provide technical leadership at every step of the process to ensure the success of the quick start.

## 2. Ingest Data Required for Those APIs Into a Centralized Hub

Once the data sources have been identified, our consultants will show you how to use our industry standard endpoints (including REST) and data loading tools for quick and easy integration of data from identified source systems. We'll work with your team on a fast data discovery process, executed in days rather than months, and help your team understand why you don't need to develop a full data model or complex transforms up-front for a Smart Mastering project.

## 3. Define and Configure Match and Merge Rules

Once the data has been loaded, as part of the Match and Merge step, we will help you create that 360-degree view of the person, the product, the drug, the piece of equipment, or other business entity being Mastered.

To achieve this, our consultants will set-up matching rules and weights to account for different ways records can match (or nearly match): ID-based matches are simplest, but most systems also require matching using fuzzy searches, nicknames, misspellings, and abbreviations. For instance, for Person entities, "William" and "Bill" are textually different yet should match, so the matching algorithm can be configured to use a nickname list, or consider other data fields such as date of birth and telephone number.

Our Smart Mastering feature can be configured to trigger automatic merge operations when matches are very close and above a configured merging threshold, or can mark less-certain matches for later processing. When should "William" and "Bill" become one person? Is full name together with date of birth enough? Within this process we work with you to determine your business rules for automatic merging, including scoring algorithms and thresholds. We also advise on and implement a human review process for lower-scoring matches.

The goal is to create a 360-degree view of your data while retaining original sources and values for lineage, audit and context.

## 4. Operationalize Your Data as Real-Time Apis and Bulk Exports

MarkLogic Master Data systems are unique in their ability to scale and handle real-time access at high loads. We'll show you how to leverage MarkLogic Server's existing REST APIs and SQL interfaces to allow master data to be easily consumed by operational applications as well as reporting, analytics and business intelligence tools. Curation and real-time operationalization of data from the same database, with API services that can scale to hundreds or thousands of requests per second, based on your business growth and data demand.

# Key Factors for Your Success

Our approach is fast and light, so you won't need a big resource commitment or upfront data modelling work; however, there are a number of things that we believe will be critical to any Smart Mastering Quick Start success:

1. Identify and define your most important use cases – our API first approach drives development focusing on the highest business value first. This focuses the effort on critical data processing, modelling, and business logic, while laying a foundation for other work to be done in a later phase – providing agility and fast delivery.
2. Access to data – we don't need a full data model, but we do need data from upstream systems.
3. A strong integrated team – data experts, developers and subject matter experts from your team are needed to identify and define data, validate results and prioritize work. This team will be trained on the job during the Quick Start to enable them to lead future efforts.

---

## How We Work

If you ensure that the above are in place when you start your engagement with MarkLogic Consulting, within 30 days of your Smart Mastering project, we will:

1. Define and implement processes and communication structures to ensure that we can work effectively with your business analysis and testing teams.
2. Review key use cases, select one or more for the Quick Start, and identify business entity / API requirements.
3. Trace API requirements back to identify the data required to fulfill the use case.
4. Provide technical leadership for the Quick Start effort, with your team embedded with our consultants as business analysts, developers and data experts. The goal is to deliver quickly, and also perform efficient knowledge transfer to your team.
5. Install and configure MarkLogic, or work with your Operations team to automate installation and configuration.
6. Configure data mapping, matching and merging rules for deduplication, merging and provenance and lineage tracking.
7. Deliver the API to support the predefined use case (e.g. line-of-business system, customer portal, analytics dashboard, CRM system).

The end result is a short, 30-day engagement that provides the baseline for a working MDM system. This Quick Start system is scoped to be fast to deliver, yet be extensible to more business entities, APIs and data sources going forward.

# Why MarkLogic Consulting

The MarkLogic Consulting Services team are experts who focus on data integration – all day, every day. We have insight into common challenges and their solutions. From change data capture to master data management, team structure, hardware selection, cloud migration and performance tuning – we have relevant experience and methodologies needed to smooth the path forward and make your effort successful.

In addition to standing up a system quickly, we aim to build required skills in your own team. Even the best team of data and systems professionals may need a little while to get their heads around new, agile, MarkLogic approaches – where you can ingest data without modeling it first, drive agility down to the very core of the data integration effort, switch from data-first to API based development.

MarkLogic Consulting have individually and collectively gathered skills and best practices from earlier projects, and look forward to supporting you on your initiative.

Contact us at [consulting@marklogic.com](mailto:consulting@marklogic.com) to learn more and request a Smart Mastering Quick Start engagement.

---

© 2020 MARKLOGIC CORPORATION. ALL RIGHTS RESERVED. This technology is protected by U.S. Patent No. 7,127,469B2, U.S. Patent No. 7,171,404B2, U.S. Patent No. 7,756,858 B2, and U.S. Patent No 7,962,474 B2. MarkLogic is a trademark or registered trademark of MarkLogic Corporation in the United States and/or other countries. All other trademarks mentioned are the property of their respective owners.

## MARKLOGIC CORPORATION

999 Skyway Road, Suite 200 San Carlos, CA 94070  
+1 650 655 2300 | +1 877 992 8885 | [www.marklogic.com](http://www.marklogic.com) | [sales@marklogic.com](mailto:sales@marklogic.com)



999 Skyway Road, Suite 200 San Carlos, CA 94070

+1 650 655 2300 | +1 877 992 8885

[www.marklogic.com](http://www.marklogic.com) | [sales@marklogic.com](mailto:sales@marklogic.com)