

SARS

BUSINESS REQUIREMENTS SPECIFICATION DOCUMENT

RFP 07/2026

**THE PROCUREMENT OF A MASTER DATA MANAGEMENT
AND DATA GOVERNANCE SOLUTION**

Table of Contents

1. Introduction	3
2. Objectives	3
3. Technical Requirements.....	4
3.1. Master Data Management.....	4
3.2. Data Cataloguing and Metadata Management	5
3.3. Visual Data Lineage and Impact Analysis	6
3.4. Unified Access.....	7
3.5. Scalability.....	7
3.6. User-Friendly Interface	7
3.7. Data Quality Management	7
3.8. Management of Data Privacy and Security	8
3.9. Policy and Compliance Management.....	9
3.10. Workflow and Collaboration	10
3.11. Reporting and Monitoring Capabilities	10
3.12. Version Control and Change Management	10
3.13. Data Governance Integration Capabilities	11
3.14. Real-Time Data Governance	12
4. Non-Technical requirements.....	12
4.1. Implementation Plan	12
4.2. Post-Implementation Support.....	12
4.3. Training and Knowledge Transfer	13
4.4. Licensing Packages	13

Adopting the Master Data Management and Data Governance Solution.

Ensuring Robust Data Management, Automated Metadata Management, Quality, Security, and Compliance

1. Introduction

The South African Revenue Service (SARS) is committed to implementing a robust master data management (MDM) and data governance (DG) agenda to optimise the management, quality, security, and compliance of its data assets. Recognising the critical role of data in driving analytics, artificial intelligence (AI), and operational excellence, SARS seeks to adopt a solution that integrates seamlessly across its diverse Information Technology landscape and meets stringent regulatory standards.

This Request for Proposal (RFP) invites qualified Suppliers to deliver an end-to-end master data and data governance solution that adheres to global best practices. The solution must enforce clear data ownership and stewardship, provide mechanisms for automated and integrated metadata management, ensure data lineage and traceability, and support comprehensive data quality controls. Additionally, it should offer robust security and privacy features to safeguard sensitive data, facilitate regulatory compliance, and enable proactive monitoring and auditing.

2. Objectives

The master data and data governance solution must enable SARS to:

- a. Establish a single source of truth for master data across key data domains.
- b. Implement data governance policies, processes, and controls in accordance with internal policies, regulatory requirements (POPIA, TAA, CRS, FATCA) etc., and industry standards.
- c. Automate and integrate metadata across SARS
- d. Enhance data quality, consistency, and accessibility across all data systems.
- e. Enable data stewardship, clear data ownership, and accountability frameworks.
- f. Ensure secure, auditable, and compliant management of master data throughout its lifecycle.
- g. Support business intelligence, analytics, and digital transformation initiatives by providing high-quality master data.
- h. Connect seamlessly with all existing systems, including but not limited to MAINFRAME (Adabas, DB2 etc.), SAP, Microsoft SQL Server, Microsoft Fabric etc, e-Filing, Enterprise Data Warehouse (EDW), SARS Customs Management System (CMS) etc.

3. Technical Requirements

3.1. Master Data Management

a. The solution must support the creation, maintenance, and management of master data entities without requiring the migration of data to a new environment.

I. This refers to the ability of a Master Data Management (MDM) solution to handle all aspects of master data directly within existing systems, rather than moving or copying that data to a separate platform or repository.

II. This implies:

- Creation: New master data entities are defined and added directly within the current operational systems or databases (e.g., MAINFRAME, SAP, Microsoft SQL Server, etc.), rather than having to transfer all data to a centralised MDM environment first.
- Maintenance: Updates, corrections, and enhancements to master data are performed in place, ensuring that changes are immediately reflected across connected systems without duplicating or moving the data.
- Management: Governance activities such as data stewardship, applying quality controls, and enforcing security policies are executed across all systems through integrated workflows and controls, ensuring consistent standards while maintaining data in its original location.

III. This capability streamlines operations, reduces disruption, and minimises the risks associated with large-scale data migration projects. It also supports real-time business intelligence and analytics initiatives, as business users can access up-to-date master data from all systems without delays caused by migration or synchronisation processes. Ultimately, it enables seamless integration and governance across diverse platforms, enhancing data quality, accessibility, and compliance while preserving the integrity and security of the original data environments.

b. The solution must have capabilities such as Data Virtualisation that integrates all enterprise data siloed across the diverse systems and delivers it to business users in real time, without physically storing the data.

c. The solution must provide workflow automation to facilitate master data approval, stewardship, and governance processes, ensuring all updates are subject to appropriate oversight.

d. Data validation and enrichment processes must be included to identify and resolve data quality issues at the point of entry.

e. Role-based access controls and Attribute-Based Access Controls must be implemented to restrict modifications and access to master data to authorised users only.

f. The solution must offer dashboards and reporting tools for real-time monitoring of data quality metrics, compliance status, and governance activities.

- g. The solution must provide data model management capabilities, enabling the definition, visualisation, versioning, and governance of master data models to ensure alignment with business requirements and standards.
- h. The solution must facilitate collaborative model development, allowing multiple stakeholders to propose, review, and approve modifications in a controlled, versioned environment that aligns with organisational governance policies.
- i. The solution should enable real-time collaborative editing, integrate easily with current systems, ensure data privacy compliance, and provide advanced analytics and reporting for monitoring model performance and usage.

3.2. Data Cataloguing and Metadata Management

The solution must be equipped with robust functionality to catalogue metadata through automated discovery methods. The solution must provide the following key data cataloguing and metadata management components:

- a. Integrated, centralised storage of metadata
 - I. Ensure that metadata assets remain unified in a central repository. The repository must be current, accessible, support discoverability and contextual understanding across multiple platforms and domains.
 - II. Offer a data glossary as a shared vocabulary to standardise terms across the organisation, reducing confusion and ensuring consistency in data use.
 - III. The solution must enable the integration of data catalogues, and business glossaries and to support e.g. compliance initiatives to enable users to find, understand, and trust organisational data.
- b. Automated metadata discovery
 - I. The solution must encompass a comprehensive suite of features designed to enhance automated metadata discovery. These features are essential for streamlining the management of metadata, reducing the reliance on manual processes, and ensuring the creation of a detailed and accurate data catalogue.
 - II. The solution must possess advanced capabilities for automatically identifying and categorising metadata across various data sources such as databases, files, data lakes, and applications while enabling user-driven manual curation to capture domain-specific context and expert insights.

c. Data classification and audit trail capabilities.

The solution should provide the following:

- I. Support configurable data classification schemes (e.g., Restricted, Confidential, Secret, Top Secret) and enable the enforcement of access controls, data handling rules, and compliance obligations based on classification level.
- II. Offer automated workflows for classification review, approval, and periodic re-evaluation should be incorporated to ensure ongoing accuracy and alignment with changing business and regulatory needs.
- III. Enable advanced search, classification, and tagging of metadata, thus improving data asset visibility and enabling efficient data stewardship.
- IV. Provide audit trails and reporting capabilities to demonstrate compliance and support governance monitoring activities.
- V. Provide role-based and attribute-based access and audit trails for metadata activities to ensure transparency, accountability, and compliance with internal and regulatory requirements.
- VI. Enable the maintenance of historical metadata snapshots and version control, allowing tracking of changes, restoration of previous versions, and impact analysis on downstream systems.

3.3. Visual Data Lineage and Impact Analysis

The solution must incorporate and facilitate both visual data lineage and robust impact analysis capabilities to ensure comprehensive oversight and management of data throughout its lifecycle.

The solution should offer the following:

- I. Provide intuitive graphical representations that illustrate the flow of data from its original source to the end user. This visual lineage is crucial for stakeholders to understand how data is transformed, processed, and utilised across various systems and applications.
- II. Offer a visualisation capability for lineage tracing and tracking, relationship mapping, and impact analysis, empowering stakeholders to make informed decisions regarding data usage, transformation, and retention.
- III. Offer customisable reporting mechanisms by summarising metadata movement trends and compliance in accordance with metadata standards. These dashboards and reports must enable transparent sharing of the metadata status to management and stakeholders.
- IV. The impact analysis feature should enable users to simulate and forecast the effects of proposed changes, whether it is alterations to data structures, updates to data sources, or modifications in processing logic on downstream processes and data consumers.

3.4. Unified Access

- a. The solution must allow users to access a single source of truth. This reduces redundancy and ensures consistency across different data systems.

3.5. Scalability

- a. The solution must be able to scale with SARS's data needs, accommodating growing volumes of data and metadata without compromising performance.
- b. It should support multi-domain master data management, allowing flexible modeling and management of diverse data domains.
- c. The solution must enable real-time or near-real-time synchronization of master data across distributed systems to minimize latency and support operational efficiency.

3.6. User-Friendly Interface

- a. The solution must have an intuitive interface that enables data stewards, analysts, and other stakeholders to navigate the system efficiently without encountering unnecessary difficulties.
- b. It should also provide self-service governance capabilities to allow business users, even those without extensive technical knowledge, to independently implement governance policies.
- c. They should be able to create, modify, and enforce data governance rules and standards without requiring assistance from IT or data specialists.
- d. The solution is required to be provided in English.

3.7. Data Quality Management

To ensure the reliability and usability of organisational data, the solution must incorporate comprehensive data quality management features:

- a. The solution must provide mechanisms for defining, tracking, and reporting on key quality metrics including completeness, validity, accuracy, timeliness, uniqueness, reliability, relevancy, consistency and integrity. using automated profiling tools and analytics. Data quality scoring dashboards covering the quality principles should give stakeholders a clear view of current quality levels and highlight areas that require attention and outlines recommendations.
- b. The solution must offer automated data cleansing functions, such as standardising formats, correcting errors, and removing duplicates, to enhance the overall reliability and integrity of information.

- c. Enable continuous monitoring of datasets with the ability to set thresholds and triggers for anomalies. Upon detection of quality issues, the system must automatically notify relevant users or administrators, ensuring rapid response and remediation.
- d. The solution must maintain data lineage, documenting the lifecycle of data from creation to consumption.
- e. Facilitate efficient investigation of data quality problems by providing detailed logs, lineage, and contextual information to identify root causes. Integrate workflows for tracking, assigning, and resolving quality issues to foster accountability and continuous improvement.
- f. Examine the data quality issues list and determine which are the highest priorities, based on how they are impacting revenue.
- g. Support data stewards and business users with collaborative tools for annotating datasets, documenting quality concerns, and proposing rule changes.
- h. Allow for both automated correction of certain types of quality issues and manual intervention where human judgement is required, ensuring flexibility and control over the remediation process.
- i. Provide customisable reports summarising data quality trends, remediation activities, and compliance with organisational standards. These reports enable transparent communication of data quality status to management and stakeholders.

3.8. Management of Data Privacy and Security

The solution must incorporate compliance and security features, which are critical for ensuring the integrity, confidentiality, and availability of data. These features should enable SARS to:

- a. Apply Data masking techniques to protect sensitive data from unauthorised access while still allowing for data analysis and processing.
- b. Implement end-to-end encryption. This approach ensures that data is encrypted before it leaves the source system and remains encrypted until it reaches its intended destination.
- c. Ensure that the masked data retains the same format and characteristics as the original data, allowing for accurate testing and analysis without compromising security.
- d. Adhere to data protection regulations such as POPIA by ensuring that sensitive data is adequately protected, thus minimising the risk of data breaches and associated penalties.
- e. Simplify the process of managing user access by enabling administrators to quickly adjust permissions as roles change or as new users are onboarded.

- f. Maintain a clear record of user roles and access permissions, making it easier to demonstrate compliance with data protection regulations during audits.
- g. Implement a data classification framework that categorises data based on sensitivity and compliance requirements, enabling organisations to apply appropriate security measures based on the classification level.
- h. Integrate authentication and authorisation mechanisms with existing enterprise identity management solutions, enabling federated authentication and single sign-on (SSO) to streamline and secure user access.
- i. Implement comprehensive logging and monitoring of all access, modification, and sharing activities related to all data assets, including metadata, ensuring traceability and rapid detection of suspicious activity.
- j. Keep a comprehensive audit trail that enable will enable the organisation to maintain detailed logs of all data interactions. This includes monitoring who accessed or altered data, the timing of these actions, and the methods used for access. Such logging is essential for compliance audits, as it allows organisations to demonstrate accountability and traceability in their data management practices.
- k. Apply detailed access policies utilising Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) models. This dual approach allows for a more granular and flexible access control mechanism.

3.9. Policy and Compliance Management

- a. The solution must facilitate the seamless integration of regulatory controls and policies.
- b. The solution must support automated workflows for policy enforcement. Automation is essential in streamlining processes, minimising human error, and enhancing efficiency.
- c. The solution must have the ability to produce comprehensive compliance reports. These reports should provide clear and actionable insights into the organisation's compliance status, highlighting areas of strength as well as those requiring attention.
- d. The reporting functionality should be flexible, allowing for customisation based on different regulatory requirements or internal needs. This capability is crucial for internal audits, external assessments, and for maintaining open communication with stakeholders regarding compliance efforts.

3.10. Workflow and Collaboration

- a. The solution must be capable of creating workflows for data owners, data stewards, and data governance team; enabling SARS to efficiently manage and oversee its data assets throughout their lifecycle. This includes the ability to define, implement, and monitor processes that ensure data quality, compliance, and security.

3.11. Reporting and Monitoring Capabilities

- a. The solution must provide advanced reporting and monitoring functionalities tailored for master data and data governance requirements. It should enable real-time dashboards and scheduled reports that track data quality metrics, such as completeness, accuracy, consistency, and timeliness of master data across the organisation.
- b. These dashboards should be configurable, allowing data owners, stewards, and governance teams to drill down into specific domains, datasets, or processes to identify anomalies, trends, or compliance gaps.
- c. The dashboards must provide visual representations of key metrics, such as data quality scores, metadata movement stats, compliance status, and workflow progress, allowing data owners and stewards to make informed decisions quickly.
- d. Comprehensive monitoring tools must allow for the automatic detection and alerting of policy violations, unauthorised data access, and data quality issues.
- e. The solution should support the generation of historical reports that show changes to data definitions, lineage, and stewardship activities, providing a clear audit trail for internal oversight and external regulatory review.

3.12. Version Control and Change Management

- a. The solution must incorporate robust Version Control and Change Management features to ensure effective Change Tracking throughout the data lifecycle. This entails the systematic maintenance of comprehensive logs that document all modifications made to data models, policies, or metadata.
- b. Such logs are essential for transparency and accountability, allowing stakeholders to trace the history of changes and understand the rationale behind each modification.
- c. It must provide the capability to revert changes when necessary, enabling users to restore previous versions of data models or policies in the event of errors or unintended consequences.

- d. This feature is crucial for mitigating risks associated with data governance, as it allows organisations to quickly address issues and maintain the integrity of their data assets.
- e. It should also facilitate streamlined Approval Processes through automated workflows for policy alterations.

3.13. Data Governance Integration Capabilities

- a. The solution must facilitate seamless integration with all existing information systems and workflows across various processes, ensuring that it can effectively communicate and operate alongside the diverse range of technologies and platforms currently in use.
- b. The solution must be configured to align with existing business processes, allowing users to incorporate data governance practices into their daily operations without disrupting established workflows.
- c. The solution must offer robust capabilities for monitoring and auditing data flows across systems, ensuring compliance with regulatory requirements and internal policies.
- d. The solution must enable integration for a hybrid (Cloud and On-Premises) environment, ensuring a seamless and cohesive approach to data management.
- e. The data generated by tool must be computed and stored in South African based data center, and such data centers must be compliant with SARS data governance and information protection policies, and other applicable SA sovereign legislations.
- f. The solution must also be flexible enough to adapt to the evolving landscape of data management, accommodating new technologies and methodologies as they emerge.
- g. The solution must seamlessly integrate or be able to connect with existing enterprise systems (e.g., MAINFRAME, SAP, Microsoft SQL Server, Microsoft Fabric, DB2, My SQL, SARS CMS, highlighting interoperability and data flow.
- h. The solution should be a single platform solution (without using multiple platforms) which cater for all aspects of Data Governance, Metadata and Data Quality.
- i. Suppliers are expected to deliver all required services directly and not rely on subcontractors for any aspect of the implementation process. Proposals must reflect the supplier's ability to fulfil all requirements without the use of subcontracting arrangements.

3.14. Real-Time Data Governance

- a. The solution must be designed to provide real-time notifications that alert users to any violations of established data policies or concerns related to data quality.
- b. The solution should incorporate automated remediation features. These features would allow the system to automatically resolve data issues or enforce governance policies without requiring human intervention.

4. Non-Technical requirements

4.1. Implementation Plan

- a. Proposed project methodology (e.g., Agile, Waterfall, hybrid).
- b. Comprehensive communication plan.
- c. Proposed timeline for deployment in a form a project plan which includes required milestones.
- d. Integration approach with existing systems.
- e. Testing and quality assurance plan.
- f. Resource allocation, roles and responsibilities.
 - I. The Supplier is required to provide the qualifications of the implementation team, number of years of experience in Master Data Management (MDM) and Data Governance, Skills, Previous project accomplishments in Master Data Management and Data Governance and ongoing professional development
 - II. The Supplier must provide at least two (2) client references from projects completed successfully within the last 5 years in South Africa in the sphere of Master Data Management and Data Governance.
- g. Risk management strategies.

4.2. Post-Implementation Support

- a. Description of support services.
- b. Solution Support and Maintenance:
 - The solution must provide standard support and maintenance packages.
 - The solution must be available 99.95% to 99.99% of the time. The acceptable downtime of planned maintenance must be 2 weeks notification prior maintenance and preferably after

working hours (off peak). Unplanned downtime acceptable time should be resolved within 2 hours.

- c. The solution must provide regular software updates and patches to address bugs and security vulnerabilities.
- d. Dedicated account manager and support team.
- e. Knowledge base and self-service portal for common issues.
- f. Post implementation support must include professional services.

4.3. Training and Knowledge Transfer

- a. The Supplier must be available for and to provide training on how to operate the solution:
 - Super User (Administrators) training – 20 users. Training tailored for administrator and technical training for platform configuration and maintenance
 - End-user training - 100 users. Training tailored to business users, data stewards, data analysts, data engineers and data scientists.
 - An internal hybrid "train the trainer" approach will be implemented, utilising both in-person and virtual sessions. The initial group of 100 trained users will subsequently provide training to the remaining team members.
 - Tailored Training Programmes: Design and deliver role-based training sessions, ensuring that different user groups receive relevant, practical instruction that matches their day-to-day responsibilities. Training should be available in virtual formats. (Virtual / Online documentation, video tutorials) to accommodate diverse learning preferences.
- b. The Supplier must provide training on new features or updates to the solution.
- c. The Supplier must conduct knowledge transfer and provide sustainability plan for SARS.
- d. The Supplier must implement comprehensive strategies that facilitate smooth transition, acceptance, and effective use of the new system across SARS.

4.4. Licensing, Subscription and once off costs (Packages)

The supplier must ensure that their pricing structure is inclusive of the following items and supported by a cost breakdown.

- a. SARS requires an enterprise license in accordance with the Supplier's licensing model in alignment with the SARS technology onboarding licenses.
 - I. Detailed pricing structure, including licensing fees, implementation costs, training, and ongoing support/maintenance.
- b. SARS requires a subscription pricing model.
- c. It is required that all SARS employees (approximately 14 000) are granted read-only access.
- d. Inclusion of compute once off costs subscription
- e. Storage costs – once off and subscription
- f. Any other technical (networking) costs