Nandini Gantayat

Individual data strategy report

Prof Arno De Caigny

MBD 2022-23

Contents

Summary of the conducted interviews

Company	**Position**	**Date of interview**
Sposea BV	Data analyst	26th June
Arcadis gen	Infrastructure engineer	27th June
Klarna	Analytics manager	28th June
Deloitte	Data scientist	4th July

Position I am interested in

Analytics manager

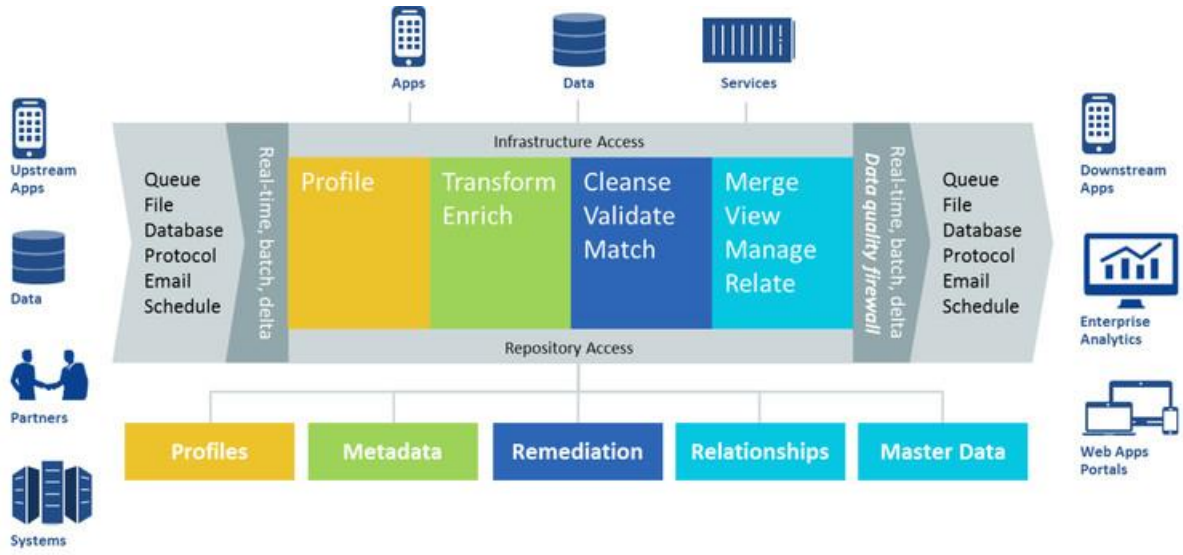
I was intrigued by the position of analytics manager specially because of its hyper dynamic nature. I belong to a business background while recently pivoting to a technology background thus working with cross functional teams really interests me. To begin on this path I need to follow a graduate program or specialized consultancy roles which allows me to do a rotation in different departments such as business intelligence, security and risk, project management and machine learning. Usually it has 3-4 different departments which will give me a real life experience of how these department work. This diverse experience will make me capable to work as an analytics manager.

The technical competencies require me to be fluent in Python and SQL, power BI for dashboarding, one scripting language like PowerShell, cloud technologies and containerization such as Kubernetes or Docker. On the management side project execution skills and agile sprints. It is necessary that I demonstrate leadership with the teams I manage and coherent communication with the stakeholders. It is also important to have a robust domain knowledge which I think I have acquired in the retail sector during my internships. Certifications such as Certified Analytics Professional (CAP), Certified Data Management Professional (CDMP) are some certification suggestions I received during the interview. Being aware of the legislative requirements is required too. I believe my key responsibilities would be to uncover hidden trends and patters from the data, optimizing the existing business processes, securing the data and setting key drivers for the organizations growth.

Reflecting upon the role I am interested in is portraying a very non linear growth path with lateral moves and deviations in its way. It's a complex role that involves data, technology, people and process. Managing it efficiently will come to me obviously with experience and continuous learning. Despite the complexity and steep learning curve I am inclined towards this role. To work at the intersection of all the crucial components of a company's growth and being a data driven decision maker really excites me to prioritize this position for my professional career.

Plan to improve operational efficiency

MASTER DATA MANAGEMENT



Master Data Management (MDM) architecture refers to the design and structure of the system and components that enable the management of master data within an organization. MDM architecture provides the framework for capturing, storing, organizing, integrating, and distributing consistent and accurate master data across different systems and applications.

The key components and concepts found in an MDM architecture are:

Centralized Repository: At the core of MDM architecture is a centralized repository or hub that serves as the authoritative source of master data. This repository stores and manages master data records, ensuring data consistency and integrity across the organization.

Data Integration: MDM architecture includes mechanisms for integrating master data from various sources and system such as data extraction, transformation, and loading (ETL), data synchronization, or real-time data integration through application programming interfaces (APIs) or web services.

Data Quality: MDM architecture incorporates data quality management capabilities to ensure the accuracy, completeness, and consistency of master data. This includes data cleansing, data validation, and data enrichment processes that improve data quality before it is stored in the MDM repository.

Data Governance: MDM architecture supports data governance practices to establish policies, rules, and processes for managing master data. Data governance ensures compliance with regulatory requirements, defines data ownership, and establishes data stewardship roles and responsibilities.

Data Modeling: MDM architecture includes data modeling techniques to define the structure and relationships of master data entities. This involves creating data models, entity-relationship diagrams, or ontologies that describe the attributes, hierarchies, and relationships of master data.

Data Integration Middleware: MDM architecture includes middleware components that facilitate data integration and interoperability between the MDM hub and other systems. This can involve

message queues, integration platforms, or enterprise service buses (ESBs) that enable seamless data exchange between different applications and systems.

Data Security and Privacy: MDM architecture incorporates measures to ensure data security and privacy. This includes access controls, authentication, encryption, and audit trails to protect sensitive master data from unauthorized access or breaches.

Data Syndication and Distribution: MDM architecture supports the distribution of master data to consuming systems and applications. This can involve data syndication mechanisms, such as data feeds, APIs, or web services, that provide real-time or batch access to master data for downstream systems.

User Interfaces: MDM architecture includes user interfaces that allow data stewards, administrators, and business users to interact with the MDM system. These interfaces enable data governance activities, data maintenance, data quality monitoring, and reporting on master data.

Scalability and Performance: MDM architecture is designed to handle large volumes of master data and support high-performance data processing. This can involve considerations such as data partitioning, indexing, caching, and load balancing to ensure optimal system performance.

Master Data Management (MDM) can be integrated with Continuous Integration and Continuous Deployment (CI/CD) practices, as well as rollback mechanisms.

CI/CD for MDM: CI/CD is a software development practice that emphasizes automation, collaboration, and frequent releases

Version control: MDM artifacts such as data models, transformations, or data quality rules can be version-controlled using tools like Git. This allows for traceability, collaboration, and the ability to roll back changes if needed.

Automated testing: Automated tests can be implemented to validate MDM processes and rules, ensuring that data is correctly managed and integrated. These tests can be integrated into the CI/CD pipeline, triggering tests on each change or before deploying new MDM configurations.

Automated deployment: MDM configurations, metadata, or reference data updates can be automated and deployed through CI/CD pipelines. This helps ensure consistency, reduces manual errors, and facilitates faster deployment cycles.

Rollback Mechanisms: In MDM, rollback mechanisms provide the ability to revert changes made to master data or MDM configurations in case of errors, data corruption, or unintended consequences. Rollback mechanisms can be implemented in different ways:

Versioning and auditing: Maintain a historical record of changes to master data and MDM configurations, including who made the changes and when. This allows for the identification of issues and the ability to revert to a previous known good state.

Point-in-time snapshots: Regularly capture snapshots of the MDM repository, allowing you to restore the system to a specific point in time if necessary. This ensures that you can revert to a previous state of master data in case of data corruption or errors.

Backup and restore: Implement regular backups of the MDM system, including the repository, configurations, and associated data. This enables the restoration of the entire MDM environment to a previous state in the event of catastrophic failures or data corruption.

Change management processes: Establish change management processes and approvals for MDM configurations and data changes. This ensures that changes are properly reviewed, tested, and approved before being deployed, reducing the likelihood of needing a rollback.

A combination of Consolidated Style MDM and Coexistent Style MDM is often beneficial for mid-size B2B analytics companies. Here's why:

- **Consolidated Style MDM**: Implementing Consolidated Style MDM can provide the company with a centralized repository for managing and governing critical master data. This approach allows for better control, data consistency, and data quality management. Consolidated MDM is particularly valuable if the company has significant data integration requirements, needs a single source of truth, or wants to ensure consistent reporting and analytics across the organization.
- **Coexistent Style MDM**: Coexistent Style MDM complements Consolidated MDM by enabling the coordination and synchronization of master data across multiple systems. This approach recognizes that some data may be managed in specific systems or departments, which can be the case in a mid-size organization with diverse data sources. Coexistent MDM facilitates data harmonization and integration while respecting existing data management practices and systems. It allows for a more agile and flexible approach to MDM implementation, accommodating existing data landscapes and reducing disruption.

By combining these two MDM approaches, a mid-size B2B analytics company can achieve the benefits of centralization, consistency, and data integration while respecting the autonomy and flexibility of various departments or systems. This approach enables the company to leverage its existing data infrastructure while progressively establishing centralized governance and consistency.

Implementing and maintaining a Master Data Management (MDM) initiative requires a range of teams and professionals with diverse skills and expertise.

Project Manager: Oversees the planning, execution, and monitoring of the MDM project, ensuring alignment with business goals, managing resources, and coordinating with stakeholders.

Data Architect: Designs the MDM system architecture, data models, and integration patterns to ensure effective management and integration of master data.

Data Stewards: Responsible for data governance, data quality management, and ensuring compliance with data policies and standards. They act as custodians of master data and ensure its accuracy, consistency, and integrity.

Data Analysts: Perform data profiling, analysis, and reporting to identify data quality issues, patterns, and trends. They assist in data cleansing, data validation, and data enrichment activities.

MDM Developers: Develop and configure the MDM system, including data integration workflows, data validation rules, business rules, and user interfaces.

MDM Administrators: Responsible for the day-to-day administration and maintenance of the MDM system, including user access management, system configuration, and performance monitoring.

In summary, Master Data Management (MDM) can help break an analytics gridlock by providing a unified and trusted source of master data. MDM ensures data consistency and integrity, facilitates data integration and aggregation, improves data quality, and establishes data governance and standardization. It enables easier data access and availability, promotes data collaboration and sharing, and supports scalability and adaptability. By addressing these data-related challenges, MDM empowers analytics teams to overcome gridlock, make better data-driven decisions, and drive business value through efficient and effective analytics processes.

References

https://www.linkedin.com/pulse/building-master-data-management-mdm-strategy-rana-banerji/

https://towardsdatascience.com/master-data-management-an-essential-part-of-data-strategy-db12411a05b2

https://www.credencys.com/blog/defining-challenges-of-unstructured-data-management-through-master-data-management/

https://www.talend.com/resources/what-is-master-data-management/#:~:text=Master%20data%20management%20(MDM)%20is,as%20a%20%E2%80%9Cgolden%20record.%E2%80%9D

https://www.mckinsey.com/capabilities/quantumblack/our-insights/how-to-unlock-the-full-value-of-data-manage-it-like-a-product

https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/breaking-through-data-architecture-gridlock-to-scale-ai

https://www.techtarget.com/searchdatamanagement/definition/master-data-management

https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/four-key-capabilities-to-strengthen-a-fraud-management-system