TOGAF[®] Series Guide

Information Architecture:
Customer Master Data Management (C-MDM)

Prepared by Céline Lescop and Jean-Baptiste Piccirillo

THE Open GROUP

TOGAF® Series Guide

Information Architecture: Customer Master Data Management (C-MDM)

Contents

Preface

The Open Group

The Open Group is a global consortium that enables the achievement of business objectives through technology standards. With more than 870 member organizations, we have a diverse membership that spans all sectors of the technology community – customers, systems and solutions suppliers, tool vendors, integrators and consultants, as well as academics and researchers.

The mission of The Open Group is to drive the creation of Boundaryless Information Flow™ achieved by:

- Working with customers to capture, understand, and address current and emerging requirements, establish policies, and share best practices

- Working with suppliers, consortia, and standards bodies to develop consensus and facilitate interoperability, to evolve and integrate specifications and open source technologies

- Offering a comprehensive set of services to enhance the operational efficiency of consortia

- Developing and operating the industry's premier certification service and encouraging procurement of certified products

Further information on The Open Group is available at www.opengroup.org.

The Open Group publishes a wide range of technical documentation, most of which is focused on development of Standards and Guides, but which also includes white papers, technical studies, certification and testing documentation, and business titles. Full details and a catalog are available at www.opengroup.org/library.

The TOGAF® Standard, a Standard of The Open Group

The TOGAF Standard is a proven enterprise methodology and framework used by the world's leading organizations to improve business efficiency.

This Document

This document is a TOGAF® Series Guide to Information Architecture: Customer Master Data Management (C-MDM). It addresses how to successfully implement C-MDM in an organization to increase the value generated to its customers' kernel customer data value. It has been developed by the Architecture Forum Information Architecture Work Group and approved by The Open Group.

This document is structured as follows:

- Chapter 1 (Introduction) introduces information architecture and data management capabilities in an organization, with a focus on the topic of this document: Customer Master Data Management

- Chapter 2 (Relationship to Existing Standards of The Open Group) shows how this document uses the ArchiMate® modeling language and the TOGAF® framework

- Chapter 3 (C-MDM Capability) presents the overall capability of C-MDM, including business and IT points of view

- Chapter 4 (Process and Methodology) extends the TOGAF ADM on the C-MDM topic; particularly, it proposes a focus on the Architecture Vision, key architecture principles, and migration planning to create a reference C-MDM journey

- Chapter 5 (Reference Models) proposes reference models: C-MDM business functions, C-MDM application functions, and C-MDM kernel data scope

- Chapter 6 (Integration Methodologies) is a focus on C-MDM integration styles, and related integration patterns to activate

The intended audience for this document is as follows:

- Chief Data Officers (CDO), staff of data management teams

- Chief Information Officers (CIO), Chief Technology Officers (CTO), Enterprise Architects, Business, Application, and Data Architects, and any IT professional involved in a C-MDM project

- Any business user involved in customer data management

It includes both executives and operational teams to support and align on a common vision across the organization on C-MDM capability.

More information is available, along with a number of tools, guides, and other resources, at www.opengroup.org/architecture.

About the TOGAF® Series Guides

The TOGAF® Series Guides contain guidance on how to use the TOGAF Standard and how to adapt it to fulfill specific needs.

The TOGAF® Series Guides are expected to be the most rapidly developing part of the TOGAF Standard and are positioned as the guidance part of the standard. While the TOGAF Fundamental Content is expected to be long-lived and stable, guidance on the use of the TOGAF Standard can be industry, architectural style, purpose, and problem-specific. For example, the stakeholders, concerns, views, and supporting models required to support the transformation of an extended enterprise may be significantly different than those used to support the transition of an in-house IT environment to the cloud; both will use the Architecture Development Method (ADM), start with an Architecture Vision, and develop a Target Architecture on the way to an Implementation and Migration Plan. The TOGAF Fundamental Content remains the essential scaffolding across industry, domain, and style.

Trademarks

ArchiMate, DirecNet, Making Standards Work, Open O logo, Open O and Check Certification logo, Platform 3.0, The Open Group, TOGAF, UNIX, UNIXWARE, and the Open Brand X logo are registered trademarks and Boundaryless Information Flow, Build with Integrity Buy with Confidence, Commercial Aviation Reference Architecture, Dependability Through Assuredness, Digital Practitioner Body of Knowledge, DPBoK, EMMM, FACE, the FACE logo, FHIM Profile Builder, the FHIM logo, FPB, Future Airborne Capability Environment, IT4IT, the IT4IT logo, O-AA, O-DEF, O-HERA, O-PAS, Open Agile Architecture, Open FAIR, Open Footprint, Open Process Automation, Open Subsurface Data Universe, Open Trusted Technology Provider, OSDU, Sensor Integration Simplified, SOSA, and the SOSA logo are trademarks of The Open Group.

AXA is a registered trademark of AXA.

All other brands, company, and product names are used for identification purposes only and may be trademarks that are the sole property of their respective owners.

About the Authors

(Please note affiliations were current at the time of approval.)

AXA® Data Architecture Community

This document is inspired by the data architecture deliverables donated to The Open Group by AXA. They are the results of the data architect community's work, with contributions from affiliates (mostly) in France, Germany, Belgium, UK, Japan, US, Hong Kong, Spain, Switzerland, and Italy from 2015 to 2020. The aim of this community is to share data architecture best practices across the AXA organization to improve and accelerate architecture delivery.

AXA is a global insurer (www.axa.com).

Céline Lescop – Author

Céline holds a Master's degree (1996) in Computer Science (www.ensiie.fr). She was a Data Project Leader in the pharmaceutical industry from 1999 until 2012 (Galderma & Astrazeneca), an Architect and TOGAF trainer from 2012 to 2015 at Arismore (The Open Group, France), and has been an Enterprise Architect at AXA leading the data architecture practice since 2016.

Jean-Baptiste Piccirillo – Author

Jean-Baptists holds a Master's degree (2010) in Computer Science (www.ensiie.fr). He was an Architect until 2015 at Arismore (The Open Group, France), a Data Scientist at AXA from 2015 to 2018, and he is currently a Data Architecture and Data Management Consultant at Rhapsodies Conseil (www.rhapsodiesconseil.fr) supporting AXA.

Acknowledgements

(Please note affiliations were current at the time of approval.)

The Open Group gratefully acknowledges members of The Open Group Architecture Forum past and present for their contribution in the development of this document:

- Julien Albert – AXA Data Transformation Program Lead

- Brendon Clark – AXA Enterprise Architect

- Sonia Gonzalez – The Open Group

- Chris Harding – Lacibus Ltd.

- Robert Weisman, University of Ottawa

The Open Group gratefully acknowledges the following reviewers who participated in the Company Review of this document:

- Chris Harding – Lacibus Ltd.

- Robert Weisman, University of Ottawa

Referenced Documents

The following documents are referenced in this TOGAF® Series Guide:

- An Information Architecture Vision: Moving from Data Rich to Information Smart, White Paper (W132), April 2013, published by The Open Group; refer to: www.opengroup.org/library/w132

- ArchiMate® 3.1 Specification, a standard of The Open Group (C197), published by The Open Group, November 2019; refer to: www.opengroup.org/library/c197

- Data Management Body of Knowledge (DMBoK), Version 2, DAMA International; refer to: https://www.dama.org/cpages/body-of-knowledge

- Information Architecture: Business Intelligence & Analytics and Metadata Management Reference Models, The Open Group Guide (G201), published by The Open Group, January 2020; refer to: www.opengroup.org/library/g201

- ISO 3166-1:2013: Codes for the Representation of Names of Countries and their Subdivisions – Part 1: Country Codes; refer to: https://www.iso.org/standard/63545.html

- ISO/TS 8000-1:2011: Data Quality – Part 1: Overview; refer to: https://www.iso.org/standard/50798.html

- IT4IT™ Reference Architecture, Version 2.1, a standard of The Open Group (C171), published by The Open Group, January 2017; refer to: www.opengroup.org/library/c171

- Lean ICT: Towards Digital Sobriety, The Shift Project, March 2019; refer to: https://theshiftproject.org/en/article/lean-ict-our-new-report/

- Michael Porter's Value Chain; refer to: https://en.wikipedia.org/wiki/Value_chain

- The TOGAF® Standard, 10th Edition, a standard of The Open Group (C220), published by The Open Group, April 2022; refer to: www.opengroup.org/library/c220

- TOGAF® Series Guide: Business Scenarios (G176), published by The Open Group, April 2022; refer to: www.opengroup.org/library/g176

TOGAF® Series Guide (2022)

1 Introduction

This document describes an approach for implementing Customer Master Data Management (C-MDM). The purpose of C-MDM in an organization is to increase the value of customer information shared across the organization. It includes people, process, organizations, and systems to manage this information as an asset.

1.1 Information Architecture Capability

The need for information architecture has existed since the 1980s when organizations started to invest in information systems to automate their operations. Initially focused on data modeling, data architects must now also focus on IT-based data management capabilities. Indeed, they play a key role in the architecture of the organization to reach a mature data management capability.

In society, usage of digital technologies is increasing daily. At the same time, the very dynamic and innovative IT market is rapidly proposing new solutions. The consequence is that the organization's information systems are continuously growing, storing much data in many different forms, and piling technologies. Another consequence is the increased information systems environmental impact that will have to be controlled in the coming years, the faster the better, as explained in The Shift Project report Lean ICT: Towards Digital Sobriety (see Referenced Documents). This growing dynamic must be stabilized to secure the support of the organization's business. Data management is a capability that can reduce data duplication and silos, and reduce the volume of data to be stored as well as improve the business's access to information.

With the rise of the Internet and platforms, data has become an asset that has to be managed in a financial and environmentally sustainable way. Data teams and initiatives lead data executives like Chief Data Officers (CDOs) and Chief Knowledge Officers (CKOs) that have been more or less recently created in many organizations to reinforce Chief Information Officers (CIOs) in the data management capability. Their purpose is to deliver high-quality data assets to support automation of operations, data science, and personalized relations with customers. Many organizations are currently racing to become data-driven at scale.

The information architecture practice must professionalize by developing approaches to organize related business operations and processes, IT systems, and data storage in a consistent holistic manner, but at the same time split into building blocks to allow decoupled but interoperable capabilities.

In this document, information architecture tools that were built to support architecture also work on the specific topic of C-MDM. It complements other documents that deal with IT-based data management capabilities, as described in Section 1.2. Best practices have been captured on the ground from specific projects and optimized to allow wide use in any organization. They aim to accelerate the delivery architecture study to make it Agile by not reinventing the wheel, bringing clarity to market offers, and providing a framework to train data architects and other data professionals in good practices. Last but not least, they aim to provide a common language for data professionals: CDOs, data managers, architects, data engineers, data scientists, etc.

1.2   Data Management Capabilities

For most organizations of a given size, major components of the value chain have been digitalized and are supported by human processes and information systems. Figure 1 describes the set of capabilities supported by the information system departments for business operations divided into six areas and organized like a value chain, splitting support activities and primary activities.
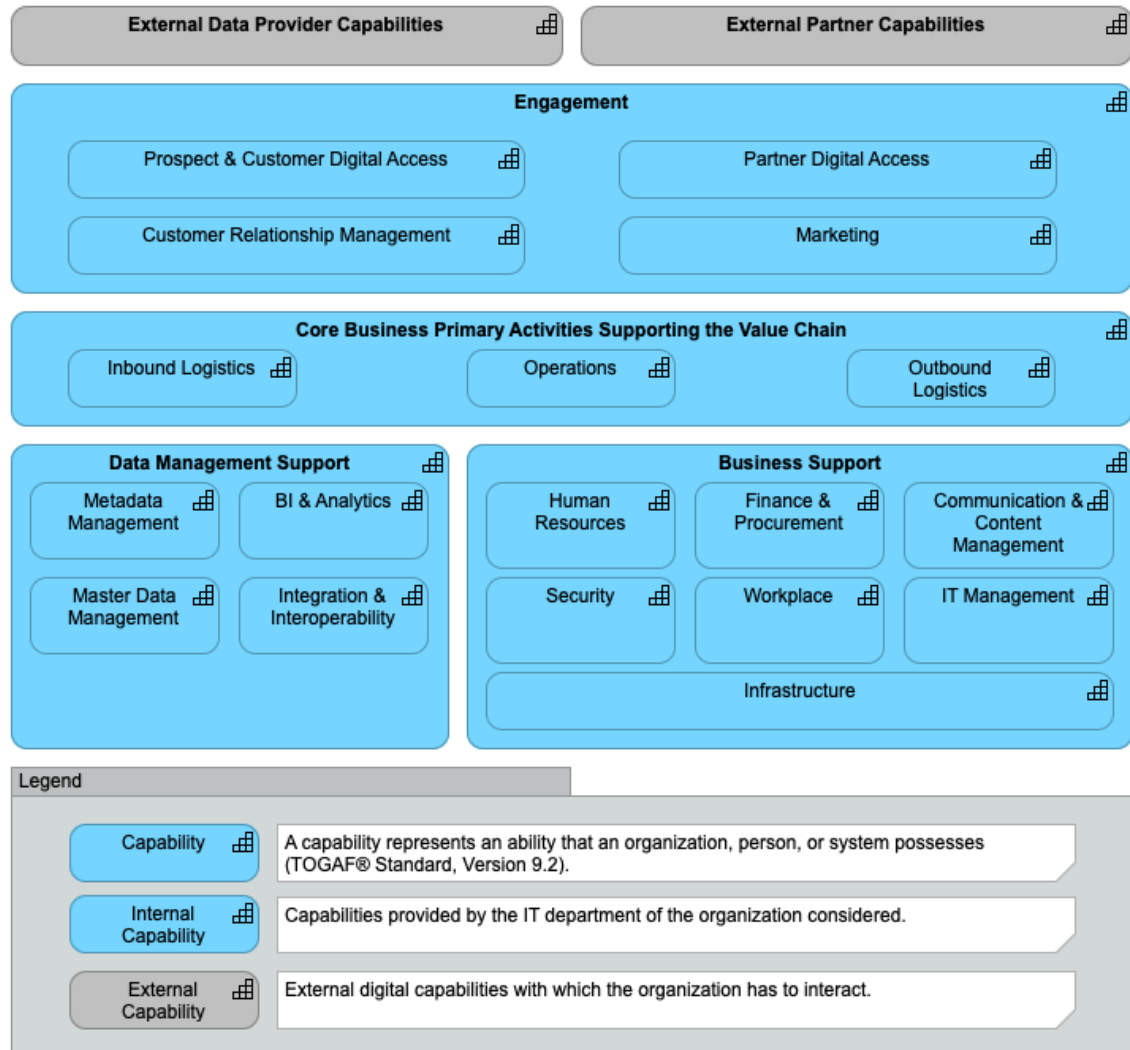

Figure 1: Generic Description of the Capabilities of the Organization[1]

---

[1] Figure 1 is inspired by Michael Porter's value chain (see Referenced Documents) with specific focus on data to become a "data-driven" organization.

Internal Capabilities

- Core Business Primary Activities Supporting the Value Chain

  This includes the capabilities that support the organization's value chain from Inbound Logistics (e.g., accept orders) to Outbound Logistics (e.g., deliver orders), passing through Operations (e.g., process orders). This is where the income is captured.

- Engagement

  This includes digital interactions with the value chain, such as the capabilities that are used to interact with the customer, business users, and partners. It covers the digital, service center, distribution, and partner channels Business to Business to Consumer (B2B2C). This category also includes Customer Relationship Management (CRM), marketing, servicing, and sales as key enablers for this layer.

- Data Management Support

  Becoming a data-driven company means a shift to a data-driven architecture approach. It leads to a specific focus on data management capabilities that support the business and that are the responsibility of the CDO. Metadata Management, Business Intelligence (BI) & Analytics, Master Data Management, and Integration & Interoperability are common data management capabilities shared across the businesses of the organization. They are the foundations that support the goal of Boundaryless Information Flow™, and are quoted in the DMBoK, Version 2 (see Referenced Documents).

- Business Support

  This includes other supporting capabilities that are required to support the value chain, such as Human Resources, Finance & Procurement, capabilities required to implement IT service capability management in alignment with the IT4IT™ Reference Architecture, Workplace, etc.

External Capabilities

- External Data Provider Capabilities

  These are data providers that can be open data or data acquired externally that is leveraged by the organization.

- External Partner Capabilities

  These are services with a digital front provided by other organizations in the market that must be combined with the digital services of the organization to deliver a more complete experience to customers.

As shown in Figure 2, the Data Management Support capabilities are mapped with the DMBoK, Version 2 Data Management Framework knowledge area of the DAMA wheel.
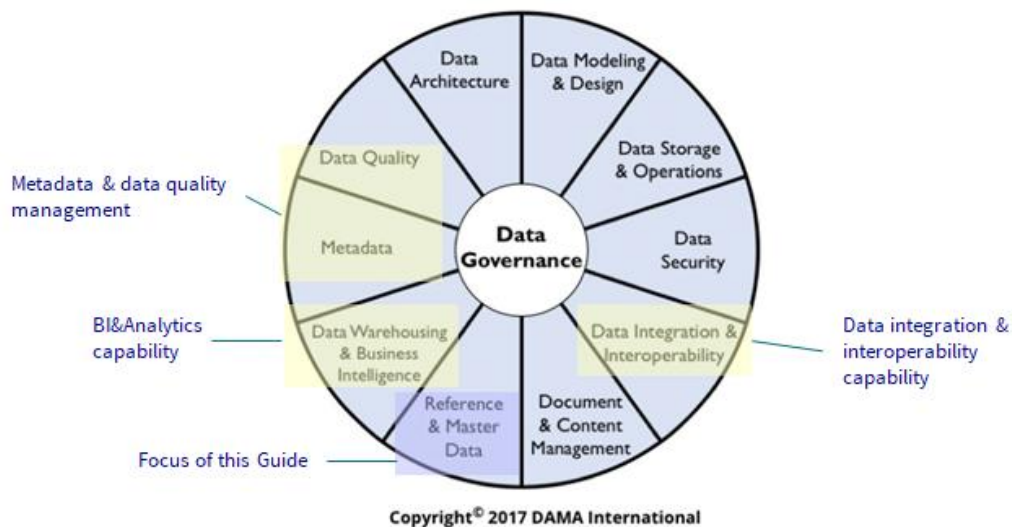
Copyright© 2017 DAMA International

Figure 2: Mapping of the Data Management Support Capabilities to the DMBoK DAMA Wheel

The Data Management Support capabilities are focused on this subset of the key DAMA activities. The other activities are covered as follows:

- Data Security is part of business support in the broader "Security" capability

- Document & Content Management is considered as a specific area in the "Workplace" capability

- Data Storage & Operations are part of the "Infrastructure" capability – they are technical transversal capabilities supporting the four core Data Management Support capabilities

- Data Modeling & Design is a subset of the "IT Management" capability

---

Business Intelligence (BI) & Analytics

A capability where all data available to the enterprise is accessible to any authorized user respecting security and privacy constraints. This enables evidence or fact-based decision-making. It covers the capabilities of data warehouses, data marts, BI, data lake, and data sciences. Indeed, these systems propose to make available all the data related to the organization in a single place, including the current and historical values. This capability aims to serve data to stakeholders needing reports for monitoring or regulatory purposes, queries, or more advanced processing like statistics or data science. Refer to The Open Group Guide: Information Architecture: Business Intelligence & Analytics and Metadata Management Reference Models (see Referenced Documents) for guidance on this capability.

Master Data Management (MDM)

This capability refers to processes, organizations, people, and systems put together to measure and continuously improve the value of master and reference data for an organization. It includes business value, data quality, availability, and security. It delivers accurate customer data to use-cases that need it. It ensures business value, data quality (uniformity, accuracy, semantic consistency, etc.), availability, security, stewardship, and customer data ownership. MDM delivers control over master data values and identifiers that enable consistent use in line with the business-defined data quality requirements.

Master Data

This is data that constitutes the most definitive, trusted values available to the enterprise and is to be both shared and used for the purposes of transactions, analytics, and/or records. For semi-structured data, it could be the approved enterprise policy on data lifecycle; for structured data, it could be the name and address of a client. Technically, master data is a consistent set of identifiers and extended attributes that is the reference source of information about shared entities used by the enterprise. Master data values are trusted and can be used with confidence by authorized users across the enterprise.

In the past, the following concepts preceded MDM and may well provide a basis for new MDM:

System of Record: where one information system provided the master data, whether the data was in a database or not (e.g., flat file).

Golden Record: where one data structure, within a specific data store, provided the master data.

Reference Data

These are standard data values defined external to the enterprise to facilitate interoperability, especially data integration. Reference data is often defined by standards organizations, such as country codes defined in ISO 3166-1 (see Referenced Documents) or NATO Stock Numbers (for supply chain). Specifically, this is any kind of data that is used solely to categorize other data found in a database, or solely for relating data in a database to information beyond the boundaries of the enterprise. This defines the set of values usually used to characterize, classify, or categorize other data.

Integration & Interoperability

This capability describes processes related to the movement and consolidation of data within and between data stores, applications, and organizations. It considers data semantics (business understanding), representation (e.g., encoding, content display, etc.), and data structure (e.g., relational, domain-specific, XML, API). It consolidates data into consistent forms, either physical or virtual.

Interoperability is the ability to "legally" share information and services. It focuses on achieving the ability to share information. (Source: The TOGAF Standard)

> Tools that support this capability include, for example: Enterprise Service Bus (ESB), Extract, Transform, Load (ETL), and Application Program Interface (API).
>
> For more information, refer to the Architecture Forum Data Integration Work Group.

> **Metadata Management**
>
> Metadata is data about data whether it is unstructured (book, bitmap images, etc.), semi-structured (e-doc), or structured (data tables, knowledge base, etc.). Metadata describes the data itself (e.g., creator, language, business description, security classification, namespaces), the concepts the data represents (e.g., business entities, semantics, business processes, technology aspects), and the connections (relationships) between the data and other concepts.
>
> Data quality can be considered as a specific type of metadata requiring specific monitoring through BI & Analytics and managing data through its lifecycle. It aims to set standards, build quality into the processes that create, transform, and store data, and measure data against standards.

1.3 C-MDM Capability

This document is about the Master Data and Reference Data management capability, focused on Customer Data.

It describes the C-MDM capability that many organizations must implement to maximize the value and efficiency of their activities. It is a data foundation for most organizations.

Indeed, without the correct data investments, data platforms will be ineffective. Being in a "customer-centric" mindset implies taking great care of customer data as an enterprise-wide discipline.

Robust data management organizations supported by a C-MDM capability should ensure that customer master data and its quality are under control by measuring and improving over time the intrinsic value of customer data (quality, accessibility, security) and the actual value of customer data by data usages (net promoter score, growth, efficiency, ROI, etc.).

Among organizations, data transformations are overwhelmed with new concepts and data platforms to master (e.g., data lake, cloud computing, data hubs, data science, Big Data, etc.). The "buzz" words obscure the essence of data transformation that is the data itself. This document suggests that fixing the basics of data management is mandatory before trying to set up advanced use-cases with data science.

> **Customer**
>
> Customer is a role that can be played by a party interacting with the organization. Each organization defines this role differently depending on their business and strategy. In this document, the term "client" is considered a synonym for "customer".

 TOGAF® Series Guide (2022)

In this document, the customers can include:

- A person: a unique human being who exists or existed in the past; they can have various relations to the organization:

  — Buyer and/or beneficiary of a service or product

  — Prospect (e.g., a person who has asked the organization for a quotation)

- An organization: a business concern or group of individuals that are systematically bound by a common purpose, with or without a legal status; organizations may be legal entities

  Various types of organization include and are not limited to:

  — Government (e.g., federal, state, regional, local, and various agencies)

  — Commercial (limited companies, publicly quoted multinationals, subsidiaries, etc.)

  — Organizational units (e.g., branches, departments, teams, etc.)

  — Informal groups (e.g., clubs, societies, charities, interest groups, etc.)

  — Other similar entities

2   Relationship to Existing Standards of The Open Group

2.1   The ArchiMate Specification

In order to properly describe the C-MDM architecture reference models, this document is based on the ArchiMate Specification (see Referenced Documents).

To cater for the specific modeling needs in this document, a subset of ArchiMate elements is chosen. This subset is given in Figure 3, followed by the definition of each concept.
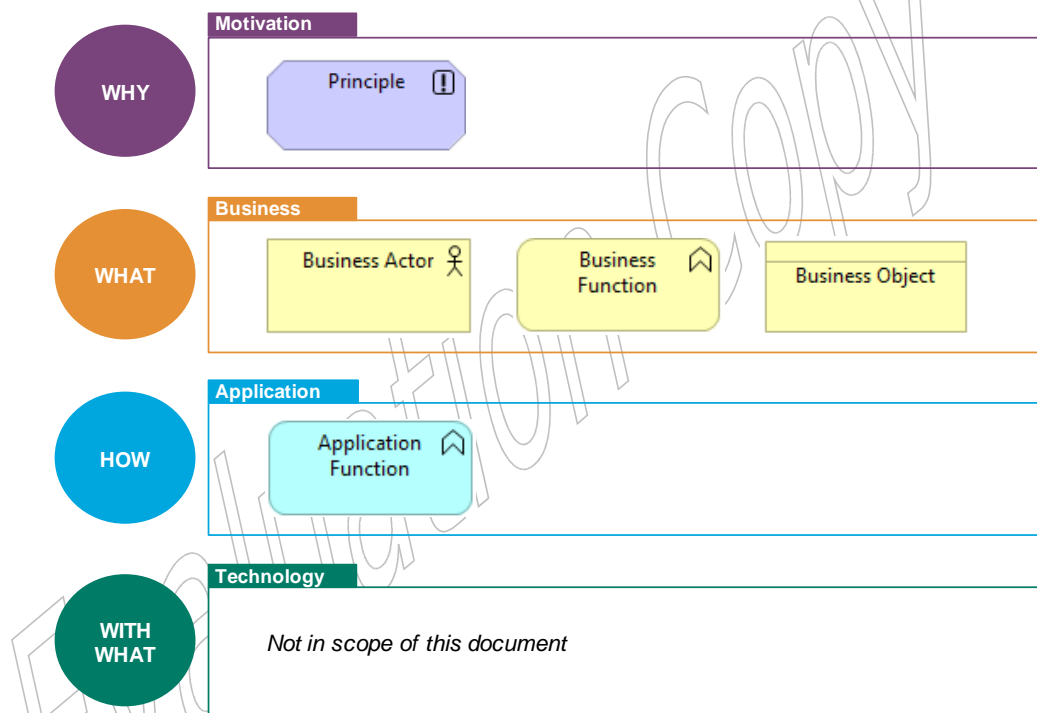

Figure 3: ArchiMate Modeling Notation Elements Used in this Document

The ArchiMate definitions of each concept are as follows:

- A principle represents a statement of intent defining a general property that applies to any system in a certain context in the architecture

    Principles are general rules or guidelines, intended to be enduring and seldom amended, that inform and support the way in which an organization sets about fulfilling its mission. They are a qualitative statement of intent that should be met by the architecture.

- A business actor represents a business entity that is capable of performing behavior

    A business actor is a person, an organization, or system that has a role that initiates or interacts with activities; for example, a sales representative who travels to visit customers.

TOGAF® Series Guide (2022)

Actors may be internal or external to an organization. A business actor is a business entity as opposed to a technical entity; i.e., it belongs to the Business Layer. Examples of business actors are humans, departments, and business units.

- A business function represents a collection of business behavior based on a chosen set of criteria (typically required business resources and/or competencies), closely aligned to an organization, but not necessarily explicitly governed by the organization

- A business object represents a concept used within a particular business domain.

  This matches with TOGAF "business information", representing a concept and its semantics used within the business.

- An application function represents automated behavior that can be performed by an application component

2.2   TOGAF Framework Extension

This document proposes:

- A focus on C-MDM to guide the design of the capability including reference models with business function and application function descriptions (see Chapter 5)

- An adaptation of the TOGAF Architecture Development Method (ADM) phases to support the execution of a C-MDM transformation program (see Chapter 4)

3   C-MDM Capability

For a given organization, C-MDM refers to people, processes, organizations, and systems put together to manage kernel customer data as an asset. It includes business value, data quality, availability, and security to deliver the required common customer data to relevant use-cases.

Kernel customer data is all the customer information shared across the organization's departments:

- Mandatory customer data for the organization to ensure at least minimal customer experience (name, surname, address, customer ID, etc.)

- Transverse and strategic customer data for the organization that should be shared across the value chain of the organization (customer type, age, professional activity category, etc.)

- Data often already existing outside the context of interaction with the organization

Kernel customer data is mainly about identification keys, personal identification data, contact data, preferences and consent data, and relationships data (relations between customers, hierarchy, etc.).

The C-MDM capability is composed of business and application functions. It requires well organized teams (Business Architecture) and adapted information systems (Information Systems Architecture). One of the challenges for the C-MDM capability is to find a subtle balance to allow adjustment between automation (in data quality, de-duplication) and manual processes when the algorithms do not produce certainty. On the one hand a lack of automation can generate excessive manual workflow, but on the other hand the algorithm cannot solve all the data quality issues, and missing human intelligence and workforce to deal with those cases prevents them from reaching the right data quality.

The C-MDM capability is under the responsibility of the customer data owner executive, who will ensure the contribution of concerned actors, as detailed in Figure 4.

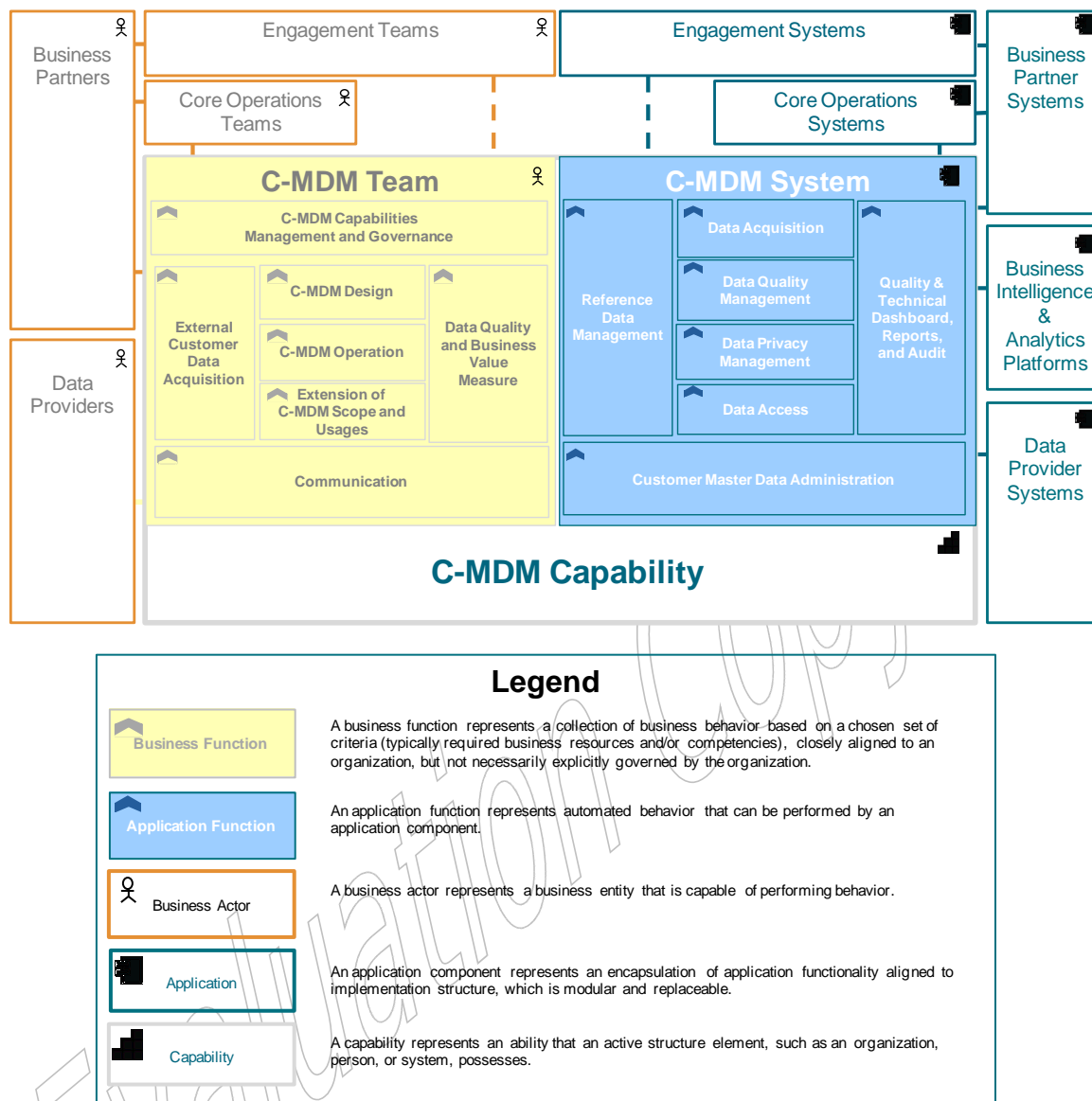**Figure 4: C-MDM Capability**

C-MDM business functions are under the responsibility of various actors: business actors like data owners or stewards, data managers, and architects collaborating to build the C-MDM vision and to operate it every day.

Table 1 provides the description of high-level business functions to leverage the capability. A complete detailed reference model is available in Chapter 5.

Table 1: C-MDM High-Level Business Function Descriptions

Business Function	Description
C-MDM Capabilities Management and Governance	Business functions supporting C-MDM strategy: governance organization, roadmap. This includes definition of customer data ownership and data stewardship and ensures that decisions are taken at the right place, with the relevant stakeholders and resources (financial and humans) to execute.
C-MDM Design	Business functions supporting design tasks: data business rules definition, customer master data model design, IT architecture design, delivery of documentation resulting from or needed by this design work. The design activity aims to identify mutualized processes across business lines implicated in C-MDM.
C-MDM Operation	Business functions supporting customer master data services delivery to consumers and producers of customer data. The purpose is to maintain and continuously improve the data quality and ensure the right level of Service-Level Agreement (SLA). It leverages relevant approaches to resolve data issues and respond to requests around data security, compliance, or data availability and accessibility.
Extension of C-MDM Scope and Usages	Business functions continuously supporting the extension of C-MDM use scope or data scope. For example, the process of integrating a new customer data producer system is the scope.
Communication	Business functions related to communication to share, advertise across the organization around C-MDM team services, ongoing roadmap/strategy, and actual results. These functions aim to show the value of customer data to the organization (use-cases, quality, etc.), and to make it obvious for everyone why this is critical for the organization to properly manage this data.
External Customer Data Acquisition	Business functions must consider enhancing customer data with external data sources, especially to accelerate data quality checks and continuous quality improvement.
Data Quality and Business Value Measure	Business functions needed to deliver and measure data quality, demonstrate continuous improvement, and show business value through Key Performance Indicator (KPI) definition and calculation to all involved stakeholders including business and data management actors.

The C-MDM system provides a single point of truth for customer data and allows the creation, update, deletion, and distribution of kernel customer data. It oversees the current version, past versions, and data lifecycle events such as mergers.

This system is "in the middle" of the information system, interconnected with all applications requiring customer data. It is involved in operational interactions with customers; it must therefore provide a very high quality of service and performance.

As this system is highly integrated with other applications, it must be decoupled from them in operations (e.g., stopping a Customer 360 (C360) service should not stop the C-MDM service), in maintenance (updating a C360 should not impact or require the test of the C-MDM service), and in design (data model) to be able to deliver the service described above.

The C-MDM system must also provide event data related to customers in a C360 view (e.g., contracts, history of interactions with the organization, claims, calls with organization services, etc.) to support customer data management.

Table 2 provides the description of high-level application functions to leverage the capability.

Table 2: C-MDM High-Level Application Function Descriptions

Application Function	Description
Data Acquisition	Application functions needed to capture customer data from a human user interface or other system inputs like data flows or an API.
Data Quality Management	Operational application functions ensure the quality of the data by implementing business and technical rules, and supporting business workflow and validation and merging processes.
Data Privacy Management	Application functions supporting data privacy constraints like retention rules.
Data Access	Application functions required by customer data consumers (human or other systems) to access the "single point of truth" master data according to visibility rules/access rights. It should provide: • Current information about a given customer • History of changes on this customer: when, who (user if manual/application if automated), what • Single point for create, update, delete, de-duplicate, etc. These operations can be decentralized in operational applications to fit different business contexts (e.g., an offline version of CRM). But in that case, the capability must be managed. The update/create event must reach the C-MDM as fast as possible to avoid data duplicates that would corrupt the master data version.
Reference Data Management	Application functions required to manage reference data (e.g., stable categories list, transcoding tables). Reference data lists are a key enabler of data quality for user interface controls or quality rules execution.
Quality & technical Dashboards and audit	Application functions needed to monitor technical and data issues, C-MDM data usage, data quality, and to ensure auditability.
Customer Master Data Administration	Administration functions like rights/roles management, data model management, logging management, etc.

4 Process and Methodology

In this chapter, an adaptation of the TOGAF ADM is proposed to support the execution of a C-MDM transformation program.

4.1 TOGAF ADM Specialization for the C-MDM Capability

This chapter describes a synthesis of key actions related to the TOGAF phases (Preliminary, A, B, C, D, E, and F) that should be done in the context of delivering a C-MDM capability.

Table 3: C-MDM Recommendations for the TOGAF ADM Phases

TOGAF ADM Phase	C-MDM Recommendation
Preliminary	Confirm that the appropriate Enterprise Architecture capabilities are in place to implement the model: repository, data sources, platforms, etc.
	Check that a multi-disciplinary team is in place with the requisite skills to architect an effective C-MDM business capability.
	Make sure that the Architecture Governance is well integrated, with teams having transversal responsibility on data in the organization (e.g., CDO teams, data governance teams, etc.).
	Ensure that the requisite privacy legislation and policies are available and understood, constraining the potential application of C-MDM. This is particularly relevant for multinational clients and/or companies that may have to comply to multiple standards; e.g., General Data Protection Regulation (GDPR), the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA), or US Federal Trade Commission (FTC) Privacy Program Criteria.

TOGAF ADM Phase	C-MDM Recommendation
Phase A: Architecture Vision	Develop an overview of existing pain points related to customer data from a customer perspective, as well as customer-facing employees or business users handling/using customer data. (The TOGAF Business Scenarios technique, as described in the TOGAF® Series Guide: Business Scenarios, can be used for this purpose; see Referenced Documents). Develop the value proposition for C-MDM considering existing or new business usage. Perform a high-level execution of Phases B, C, and D to get an overview of the main impacted systems, showing key macroscopic metrics. (How many applications can update kernel customer data today? How many applications can access kernel customer data in read-only mode? Overall, how many users can update data?) Create a conceptual C-MDM Target Architecture highlighting the value proposition for the key stakeholders (e.g., data quality improvement, expected income increase, expected customer experience increase, efficiency, risk and compliance, etc.). Master data is a direct function of the agreed level of data quality within the enterprise. There are standard metadata characteristics such as "accuracy" and "truthfulness" that must be agreed upon between the Business and Data Architects. The business has to specify the nature of data quality required in an SLA contract in Phases A and B to ensure that the master data, which can be acquired as data as a service, does not cost the enterprise more than it is worth. Examples of data quality are included in Annex A to The Open Group White Paper: An Information Architecture Vision, the DMBoK, Version 2, and in the latest version of ISO 8000: ISO/TS 8000-1:2011 (see Referenced Documents).
Phase B: Business Architecture	The business is accountable and responsible for data quality. Establish the governance, processes, and resources necessary to ensure that the data quality of the master data is achieved and sustained. Understand the usage of customer data in the daily activities of the business. Identify operational issues and related stories around customer data and understand the deep causes: organizational, process and tools, human or technical. Identify current operations or activities that are not effective due to lack of data quality. Identify future use-cases C-MDM should support (e.g., accurate C360 view, consent management, effective multi-channel marketing, etc.). Maintain a use-cases backlog and an issues backlog to fuel the C-MDM program. Formalize new requirements C-MDM should cover to develop these use-cases or correct current data issues. Formalize existing organization around managing kernel customer data. Propose target organizations and processes to increase C-MDM and contribute to cover the new identified requirements. Especially describe the match-and-merge target process from the Business Architecture point of view. Identify the level of change effort and mobilize adequate resources to roll out target organization and processes.

TOGAF ADM Phase	C-MDM Recommendation
Phase C: Information Systems Architectures	Have an overview of as-is applications creating, reading, updating, and/or deleting customer master data today. Qualify each of those applications with respect to customer data (volume of concerned customers, frequency of creation/updates of customer master data, data profiling to assess the existing quality of the data source, business value). Qualify each of those applications with respect to existing or planned projects or evolutions/roadmap (that could be an opportunity, for example, to move the mastership to future C-MDM). Identify current issues and improvement points around the current Information Systems Architecture. Define the logical target Information Systems Architecture to consistently manage kernel customer data. Review the match-and-merge customer process.
Phase D: Technology Architecture	Check that the technology choices leverage consistent transactions around kernel customer data (on which the system aims to be the point of truth). These capabilities should be loosely-coupled with the IT landscape to ensure consistency and availability of critical transactions and efficient maintenance.
Phase E: Opportunities and Solutions	Use the application reference models in this document to challenge solution providers around C-MDM. Reference models can be used as a basis to build selection criteria for the solutions.
Phase F: Migration Planning	Have a complete vision of planned projects/initiatives adjacent to C-MDM topics (e.g., projects link to applications managing customer data). Qualify dependencies and potential opportunities. Use the implied applications qualification (in Phase C) to prioritize integrations through time (e.g., regarding business value, customers volume and current quality, complexity), investigate if existing initiatives and planned projects could be used to cover the C-MDM program evolution. Define a migration roadmap and show how the mastery of kernel customer data evolves through this roadmap.
Phase G: Implementation Governance	Organize a compliance review with C-MDM projects to check that they respect C-MDM principles and the Target Architecture. (Examples of principles are given in Section 4.3.)

4.2 C-MDM Guidelines for an Architecture Vision (TOGAF ADM Phase A)

C-MDM projects are complex, impacting the scope of the entire entity with many actors to involve. So, it is imperative first to build a vision and to reach consensus among all concerned stakeholders at a high level. Table 4 shows the dimensions of this vision. This vision must be endorsed by an executive sponsor and/or the CEO.

Table 4: Architecture Vision Dimensions

C-MDM Team	**FROM** some data rules not documented **TO** one C-MDM team accountable for customer data • Build one polyvalent C-MDM team, with power to govern and decide • Define and operate ownership and stewardship – make businesses aware of their (new) responsibilities around the data and its quality • Define a value-added services catalog of the team with clear SLAs
Single point of reputed truth system	**FROM** multiple points of potential truth **TO** single point of truth for customer data • Have an ambitious and shared single point of truth target for kernel customer data • Be opportunist – have a good vision of the information system, integrate C-MDM in planned project solutions when relevant or possible • Take care of delivering business value to operational processes in the first steps of the C-MDM program – the delivery of data quality must solve business issues (such as mail, email not delivered, customer consent not properly managed, etc.) • C-MDM should save all instances of customer data and not just the current master version – some business usages may require them; as an example, date of birth is often used in pricing – in cases where the premium was calculated with a wrong date, the wrong date should be kept
Data quality and trust measures	**FROM** doubts about customer data quality and compliance **TO** trust and accuracy measured and improved continuously • Measure data quality and business value over time (KPI, ROI) and encourage the business to define quality dashboards and business KPIs • Use external data, documents, and evidence to improve quality (e.g., ID cards, external data sources from providers) • Ensure that C-MDM improves the capability to meet risks and regulation requirements
Business positive impacts ensured, from payer to partner	**FROM** difficulties to be reactive and proactive with customers **TO** real partnerships and mutual trust • Focus first on systems and people closest to the customer so they are well served (i.e., engagement) • Maintain continuous awareness of operational and strategic business evolutions – this can change the priorities of the project • Foresee the opportunity to integrate into the C-MDM approach an operational and usable C360 view and other value-added services

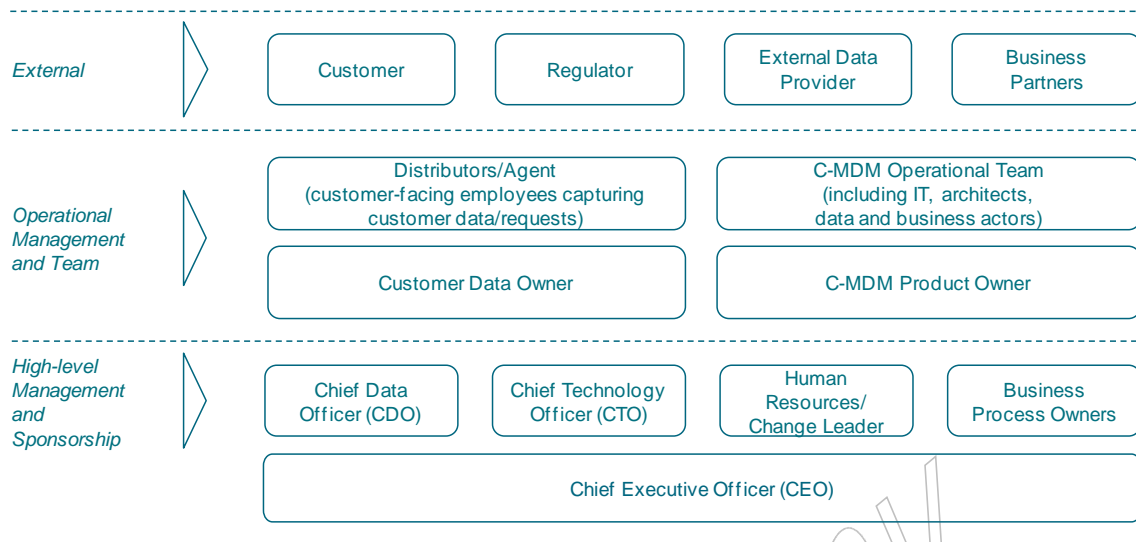The vision should imply key C-MDM stakeholders.

Figure 5: Architecture Vision Stakeholders

Finally, a transversal vision of the C-MDM approach is a must-have. The best executive sponsor is the CEO, able to break organizational/line of business silos for the benefit of the customer.
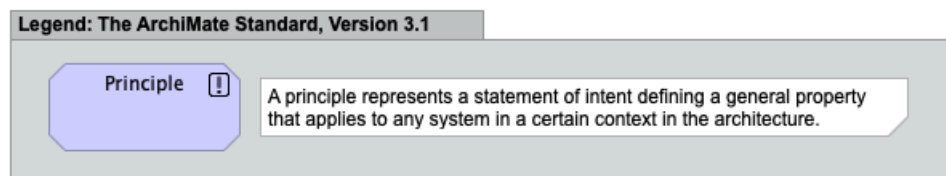
4.3 C-MDM Architecture Principles

Figure 6: ArchiMate Architecture Principle Definition

This section lists commonly observed architecture principles to set up a C-MDM capability in an organization. The principles are presented using the TOGAF ADM structure, starting with business principles, data principles, and application principles.

They can be used either as a first list to adapt to the specific organization, or to compare with an existing C-MDM principles list.

These principles are a first, reusable version. They should be fully formalized adding "Rationale" and "Implications" for each principle.

4.3.1 Business Principles

Business Principle ID	Principle Statement
P_BUSINESS_001	The C-MDM team is named and responsible for the customer data. This team has the power to decide about customer data (e.g., merge rules and process, stewardship, ownership) and to organize the dedicated governance with business and IT stakeholders.

TOGAF® Series Guide (2022)

Business Principle ID	Principle Statement
P_BUSINESS_002	The centricity of the C-MDM capability, both in the business scope and in the information system, requires high-level knowledge and expertise in multiple domains: lines of business, IT, data management. Those skills must be covered by the C-MDM team.
P_BUSINESS_003	A well-defined value-added services catalog of the C-MDM team should exist with related SLAs.
P_BUSINESS_004	The C-MDM team clearly identifies external data as a potential lever to improve de-duplication and data quality services, and external data market analysis is done on a regular basis.
P_BUSINESS_005	The C-MDM team should encourage use of documents and evidence in the data stewardship processes to improve data quality, when possible (e.g., ID cards, bills, etc.).
P_BUSINESS_006	Customer data quality indicators are defined, produced, and shared by the C-MDM team.
P_BUSINESS_007	C-MDM capability business value indicators are defined, produced, and shared by the C-MDM team.

4.3.2 Data Principles

Data Principle ID	Principle Statement
P_DATA_001	The targeted C-MDM data business scope should be defined (lines of business, customer roles, etc.).
P_DATA_002	A set of fields allowing unique identification of each customer must be identified and managed through the C-MDM capability, called kernel customer data. The quality of this set of fields must be the first focus of the C-MDM team.
P_DATA_003	Every unique customer identified must have a technical key. Links between technical and customer business keys should be managed in the C-MDM capability.
P_DATA_004	The kernel customer data should be line of business-agnostic and the data should be decoupled from other customer data to guarantee high availability, stability, performance, and scalability.
P_DATA_005	The C-MDM capability should provide the latest version of the customer data and track historical versions and operations, including contextual metadata such as source, author, date.
P_DATA_006	The customer master data model must include the following conceptual entities: customer data attributes, customer relationships, customer roles.
P_DATA_007	The customer data must be documented: data dictionary, glossary, model.

Data Principle ID	Principle Statement
P_DATA_008	C-MDM access must comply with the entity requirements and local regulation.
P_DATA_009	C-MDM is not the master of contract/policy data and will not impose any truth to contract data. Contract data is input to improve customer data quality, and customer data is input to improve contract data quality. The contract management process decides the data to be used to support its activities. The C-MDM process decides the data to be used to support its activities.

4.3.3   Application Principles

Application Principle ID	Principle Statement
P_APPICATION_001	The Application Architecture supports the C-MDM capability by being consistently integrated with points of acquisition and points of consumption. This integration ensures customer data consistency and integrity; access to all instances of customer data (e.g., current or old version).
P_APPICATION_002	Integration patterns are well-defined/documented and available to support data exchanges with the C-MDM system.
P_APPICATION_003	The Application Architecture should be compliant with the organization's security rules.
P_APPICATION_004	One or more user interfaces are implemented to support master customer data functions (stewardship). Depending on the defined stewardship, user interfaces are available on operational systems or on a dedicated C-MDM user interface.
P_APPICATION_005	The Application Architecture must provide: operational capabilities to support transactions on customer data (create, update, merge, delete, etc.); analytical capabilities to support potential complex de-duplication processes, quality indicators, and scores.
P_APPICATION_006	The system should have loosely-coupled capabilities around kernel C-MDM to ensure the data consistency, availability, accessibility, and stability of the master data scope.

4.4   Reference C-MDM Roadmap Steps (TOGAF ADM Phase F)

This section proposes reference C-MDM roadmap steps to support the architecture journey that can be used to structure TOGAF ADM Phase F: Migration Planning.
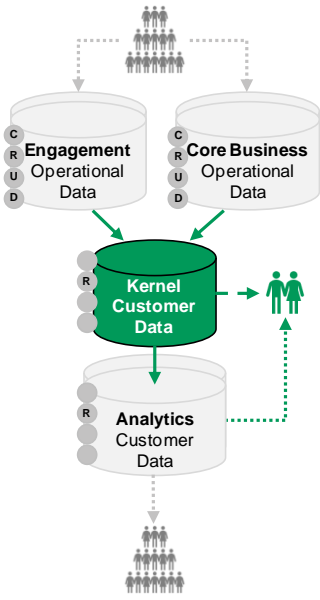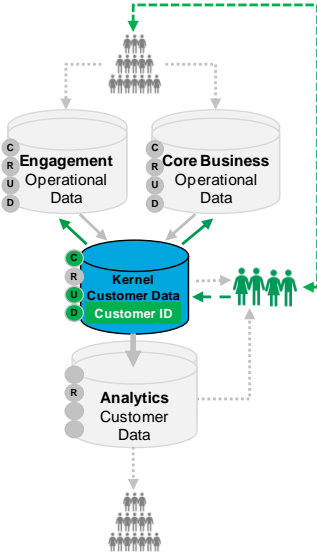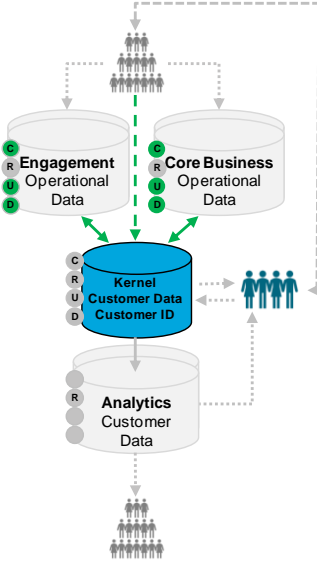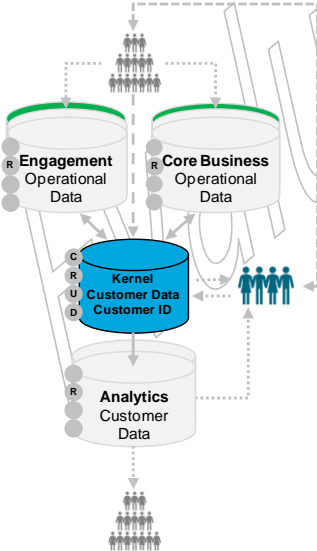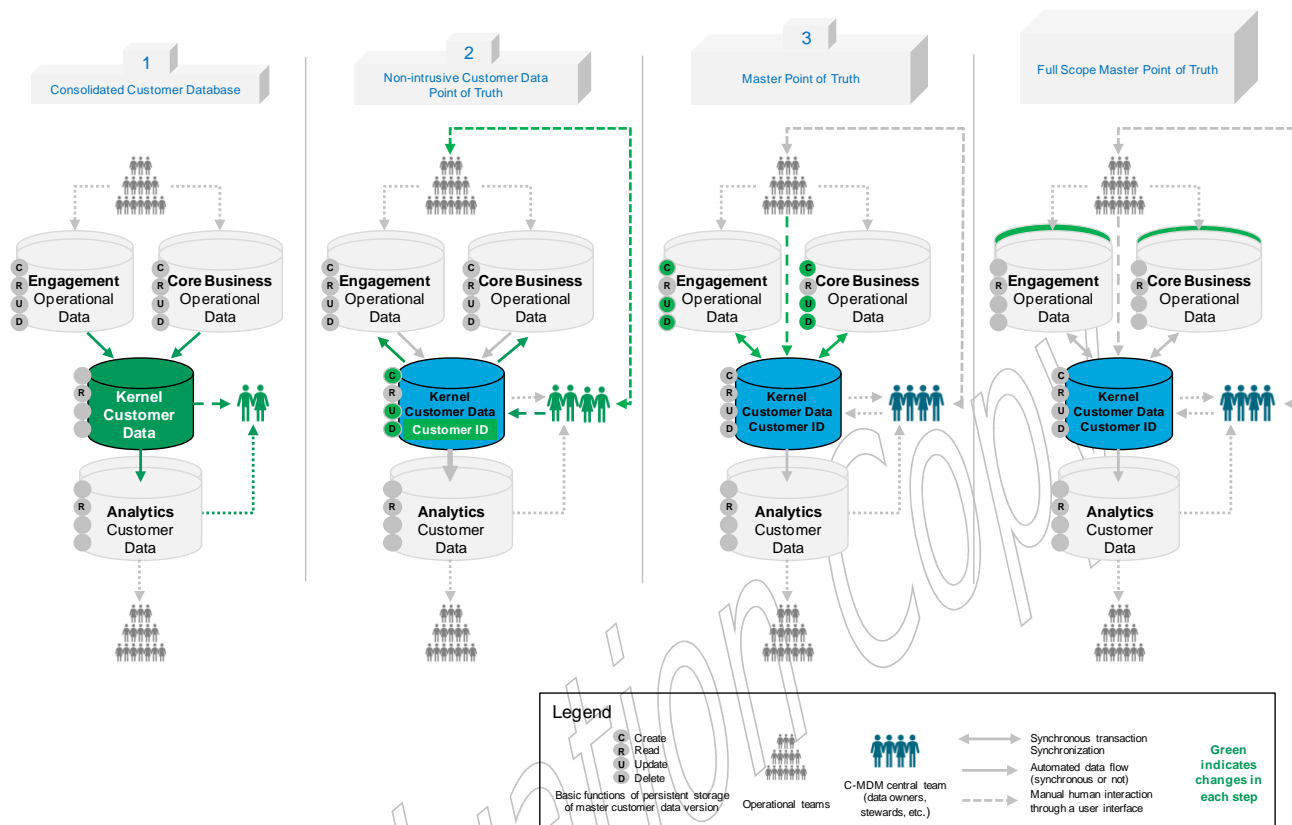

Figure 7: Reference C-MDM Roadmap Steps

The steps (plateaus) illustrated in Figure 7 are described in Table 5.

Table 5: C-MDM Roadmap Steps (Plateaus)

Step (Plateau)	Description
Step 1: **Consolidated Customer** **Database** 	Central customer data store, but without true golden record (no effective de-duplication, poor data quality) • Kernel customer data is updated in different operational systems • This kernel data is consolidated in an initial version of C-MDM as a first version of customer golden records Use-cases: • This first version enables use-cases such as segmentation, market share analysis, C360 view, some campaign management use-cases; these use-cases can tolerate errors on the matching process (100% accuracy of the de-duplication is not mandatory) Data quality: • Matching algorithms are executed to evaluate customer duplicate rates • Customers data quality dashboards are built to share insights on customer data quality: completeness, uniqueness, consistency, etc. Trigger for next step: • Need to improve data quality and support customer-facing business processes (Note: The "golden record" of a customer is the data values about the customer considered to be the truth for the organization, even if other versions of the data can exist in other systems.)

TOGAF® Series Guide (2022)

Step (Plateau)	Description
Step 2: **Non-intrusive Customer Data Point of Truth**	Analytical C-MDM focused on golden record, de-duplication, and data quality improvement • Kernel data is still updated in different operational systems • C-MDM provides a golden record for each identified unique person and attributes a stable organization customer ID; this golden record can be propagated in operational systems but is not synchronized • A dedicated central C-MDM team is responsible for data stewardship including confirming de-duplication, analyzing data quality issues, data correction, and completion Use-cases: • Address engagement use-cases: send automatic alerts to the CRM system on new address or bad address detection, display golden record to agents Data quality: • Some data quality issues are tackled at root cause in the operational source system thanks to manual feedback loops of the central C-MDM team • Data quality increases Trigger for next step: • Concurrent updates generate workload that creates an appetite for more integration; the value of shared customer data is obvious for lines of business and allows sponsorship for higher impact on the information system

Step (Plateau)	Description
Step 3: **Master Point of Truth** *(diagram: Engagement Operational Data and Core Business Operational Data flowing into Kernel Customer Data / Customer ID, with Analytics Customer Data below)*	Operational C-MDM progressively extending its mastership scope on the information system • Kernel customer data is updated in C-MDM (through different integration patterns: services, real-time synchronization, user interfaces, etc.) • Automated processes/workflows propagate data changes Use-cases: • Indicators are accurate, such as number of customers • The operational business processes and applications in the core business or engagement are reviewed to leverage the customer golden record; this allows consistent customer experience across channels • Customer data quality becomes a competitive advantage in business partnerships and allows advanced customer services Data quality: • Data quality is reaching expected target level – high integration comes with high data quality requirements: a wrong merge can lead to a crisis Trigger for next step: • Continuous communication on C-MDM value is done to secure sponsorship over several years – operational systems are integrated one by one, increasing the scope of customer data
C-MDM Target *(diagram: Engagement Operational Data and Core Business Operational Data with Kernel Customer Data / Customer ID, and Analytics Customer Data below)*	Operational C-MDM across the entire information system • Governance and stewardship on customer data are institutionalized • C-MDM covers the full scope of line of business customer populations: individuals, commercials, etc. • C-MDM is a central piece of the information system; its value is proven

TOGAF® Series Guide (2022)

5   Reference Models

Reference models aim to become a stable and common language to describe the functions that platforms should cover. They can be used:

- To describe the functional coverage of a given provider or to compare the functional coverage of two providers

- To assess the Baseline (as-is) Architecture of a given organization both from a business and technical point of view

- To design the Target (to-be) Architecture and Transition Architecture of a given organization

- To evaluate the maturity of a capability

- In Architecture Governance, to decide on a scenario for the Target Architecture
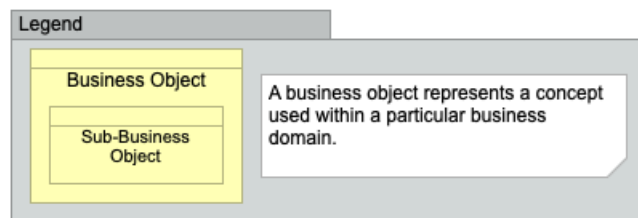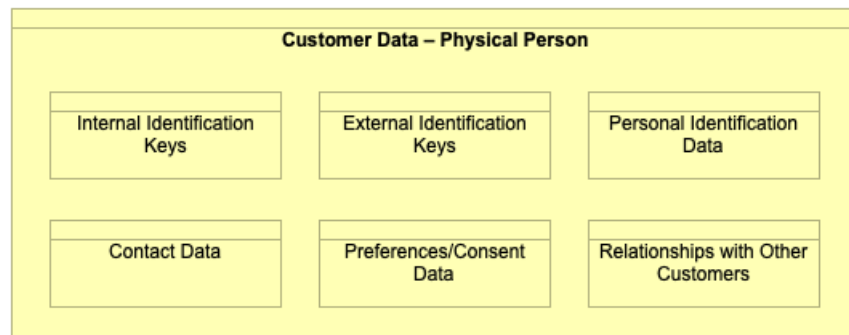
This chapter provides two types of reference model:

- Kernel customer data scope reference model

- C-MDM detailed business and application functions reference model

5.1   C-MDM Kernel Customer Data Scope

As seen previously, kernel customer data scope includes transversal data on the customers. This section provides a list of the most common fields considered as kernel data, covering physical person and organization/legal entity kernel data. It should not be used as a mandatory guideline, but as a first baseline or as a checklist to compare with the plan or existing capability in a specific organization.
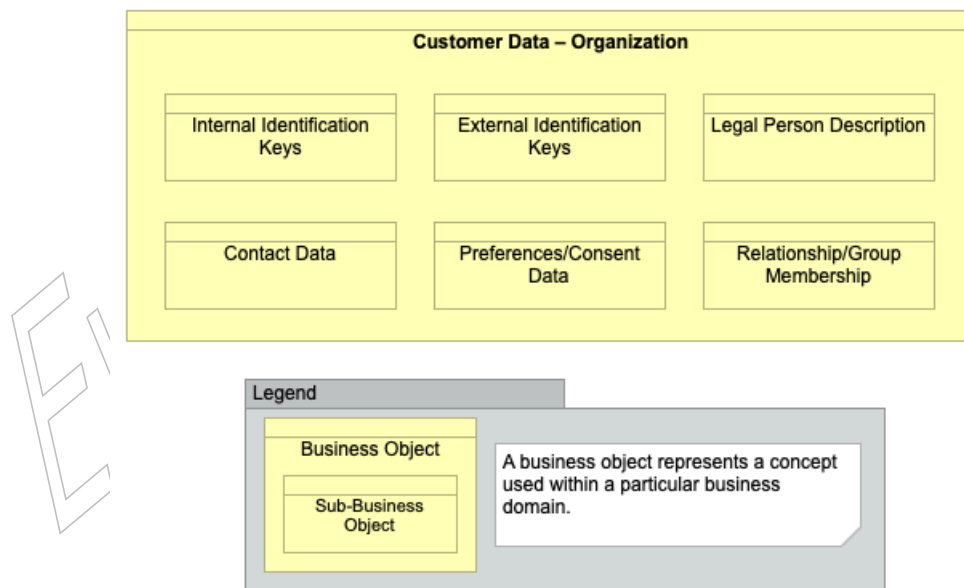
5.1.1 Physical Person Kernel Data

Type of Data	Information	Description (if needed)
Internal Identification Keys	Internal business keys (customer code).	Customer code(s)/key(s) making sense for the business and for the customer.
	Unique internal technical key (customer key).	Unique internal technical key of the customer managed by C-MDM.
	Other unique technical key managed by applications other than C-MDM.	Unique internal technical key of the customer managed by other applications (CRM, etc.).
External Identification Keys	External keys (e.g., ID card number, external data provider reference for this person, etc.). The organization must check that it is allowed to store and process this type of information.	External keys could be useful to ensure uniqueness of customers (e.g., ID card number). They also allow a process of collaboration with the external data provider to improve data quality.
Personal Identification Data	First name	—
	Last name	—
	Date of birth	—
	Place of birth	—
	Family/marital status	—
	Official professional category	—

TOGAF® Series Guide (2022)

Type of Data	Information	Description (if needed)
Contact Data	Address(es)	—
	Email(s)	—
	Phone number(s)	—
	(Other contact data: social network, etc.)	—
Preferences/ Consent Data	Opt-in/opt-out from customer authorizing or not authorizing the organization to use a scope of customer data for defined purposes.	In many organization this field is introduced in C-MDM and managed to comply with GDPR in 2018.[2]
Relationships with Other Customers	Relationships with other persons belonging to a group.	For example, father, mother, child, employee, etc.

5.1.2 Organization/Legal Entity Kernel Data

Most of the time, organizations are legal entities registered with competent authorities, depending on the country. They can be companies, administrations, Non-Governmental Organizations (NGOs), etc.

| Type of Data	Information	Description (if needed)
Internal Identification Keys	Internal business keys (customer code).	Customer code(s)/key(s) making sense for the business and for the customer.

---

[2] Refer to: https://gdpr.eu/gdpr-consent-requirements/.

Type of Data	Information	Description (if needed)
	Unique internal technical key (customer key).	Unique internal technical key of the customer.
External Identification Keys	External keys (e.g., organization official number, data provider references for this organization, national registry ID, etc.).	External keys could be useful to ensure uniqueness of customers. They also allow a process of collaboration with the external data provider to improve data quality. In the case of a legal entity, registration usually comes with a public ID.
Legal Person Description	First name	The official name of the organization.
	Alternative name(s)	Potential alternative names making sense for the business, usage name.
	Date of birth	Date of creation of the organization.
	Dimensioning data about the organization	For example, official categories to define the weight of the organization.
	Organization business sector	For example, official business sector name used in the country.
Contact Data	Referent physical (persons) and their contact data	Can be a link to another C-MDM physical person.
	Referent address(es)	Referent address(es) used to communicate with the organization.
	Referent email(s)	Referent email(s) used to communicate with the organization.
	Referent phone number	Referent phone numbers used to communicate with the organization.
Relationship/ Group Membership	Relationships with other persons (e.g., physical important persons related to this organization, CEO, etc.), belonging to a group (e.g., corporate center and affiliates).	

5.2   C-MDM Detailed Functions Reference Model

In this section, the reference model presented in Chapter 3 is completed with more detailed business and application functions. It can be used to deep-dive on specific questions during the architecture study, assess maturity of an existing capability, or provide a first framework for C-MDM projects to be customized.

5.2.1 C-MDM Detailed Business Functions Reference Model
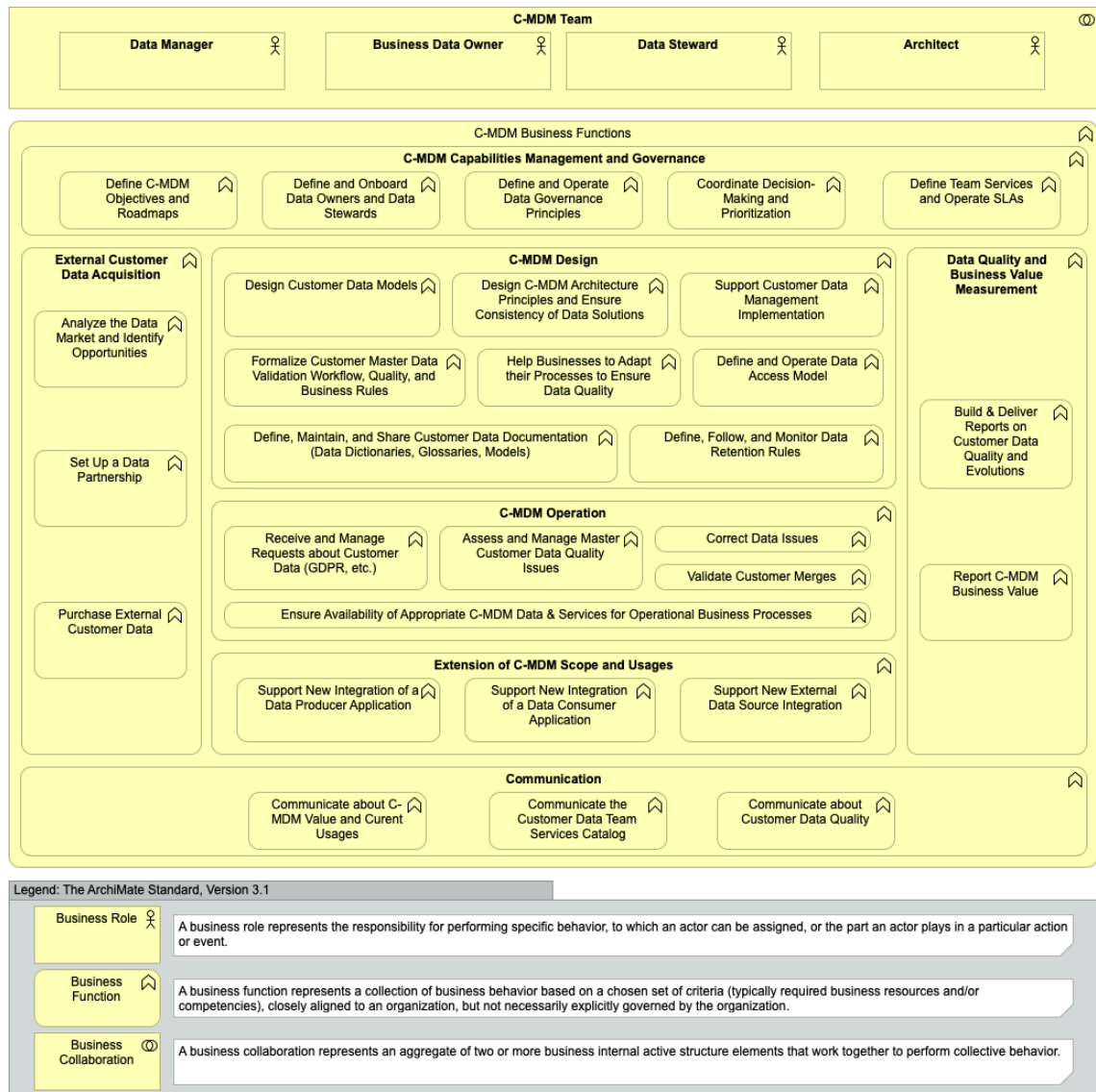
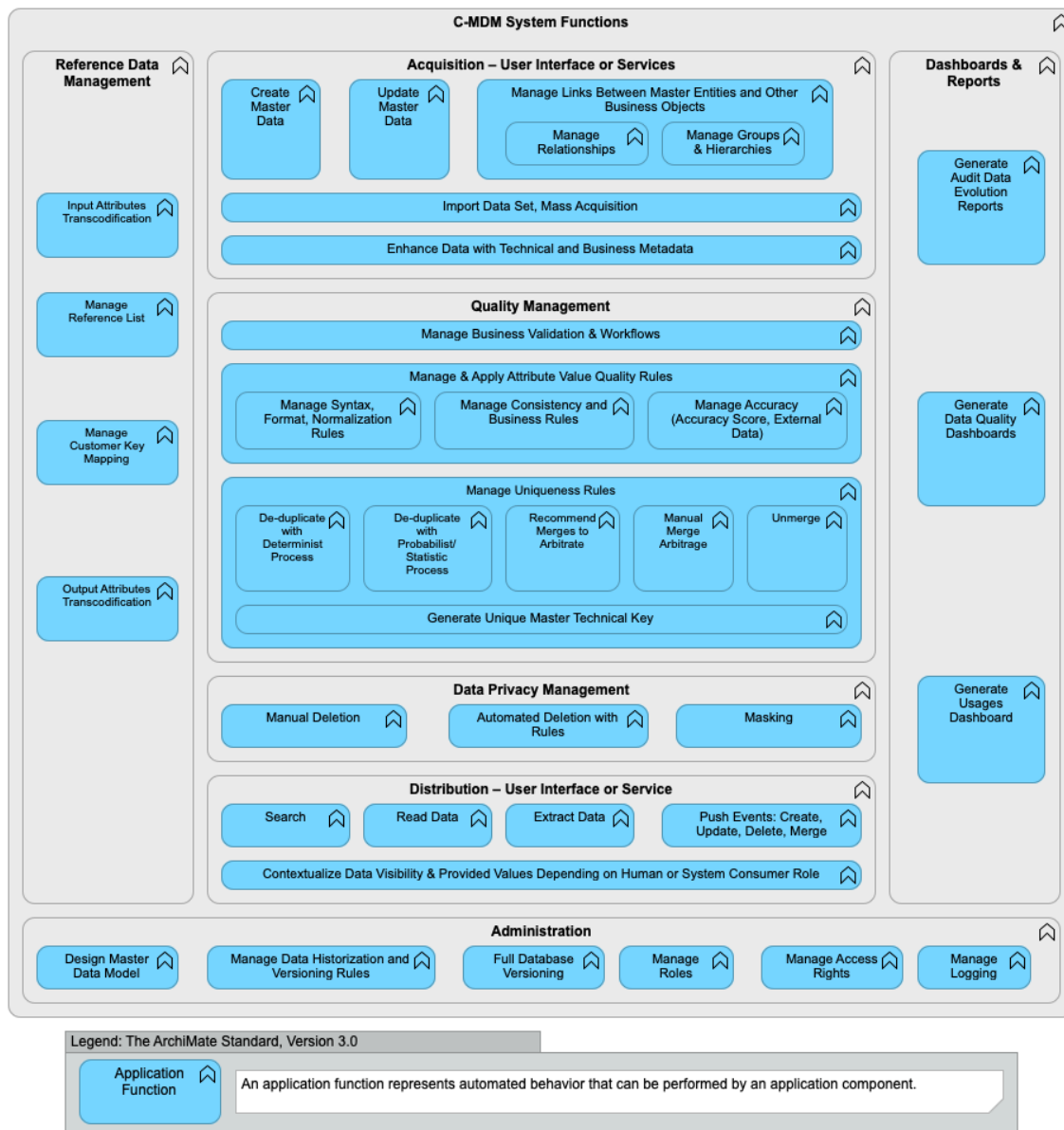Detailed descriptions of theses business functions are provided below.

Business Function	Description
C-MDM Capabilities Management and Governance	
Define C-MDM Objectives and Roadmaps	Define the transformation steps of the C-MDM capability with clearly defined objectives to achieve each of them and a plan to do so.
Define and Onboard Data Owners and Data Stewards	Identify and onboard business data owners in Business-as-Usual (BAU) mode to focus on kernel customer data.

Business Function	Description
Define and Operate Data Governance Principles	Define data governance principles about customer master data. Ensure that the organization and instances (e.g., boards) are set up to operate those principles.
Coordinate Decision-Making and Prioritization	Set up and animate data governance board(s) to ensure that the C-MDM roadmap is effectively implemented within the entity. This board ensures the consistency of decisions regarding customer data all over the entity and that all key concern/risk arbitration is reviewed and considered by stakeholders.
Define Team Services and Operate SLAs	Define the catalog of services of the C-MDM team. Define potential users of these services, and interfaces to deliver these services (e.g., ticketing tool, phone number, centralized email point, etc.). Define SLAs for each service and ensure that they are respected.
C-MDM Design	
Design Customer Data Models	Design and document the customer master data model (information model, database model, attributes definitions and types, etc.).
Design C-MDM Architecture Principles and Ensure Consistency of Data Solutions	Define customer master data requirements (e.g., data storage, data transfer, data publication, data transformation, data archiving, metadata management, data quality management, master and reference data management, etc.) in line with compliance standards. Define data architecture frameworks to meet these data requirements. Contribute to the IT roadmap and assessment of the proposed investments. Ensure that critical data is available in line with compliance and security standards.
Support Customer Data Management Implementation	Provide the resources and ensure that the C-MDM systems and team are set up to fit requirements over time. Secure separate implementation steps regarding the defined roadmap over time.
Formalize Customer Master Data Validation Workflow, Quality, and Business Rules	Formalize business rules and stewardship workflows to guarantee that quality of existing data will increase. The purpose is also to check that new data create and update processes are designed appropriately to ensure the quality of new data.
Help Businesses to Adapt their Processes to Ensure Data Quality	In the design of including new data scope in the C-MDM capability, or supporting new or existing use-cases, business processes could be impacted and are part of the overall solution. This activity is about helping businesses to adapt and even improve the processes that will be supported by C-MDM services. Change management efforts should not be underestimated.
Define and Operate Data Access Model	Define data access profiles according to security, privacy/compliance, and business value classification, and operate these access models through the C-MDM solution.

TOGAF® Series Guide (2022)

Business Function	Description
Define, Maintain, and Share Customer Data Documentation (Data Dictionaries, Glossaries, Models)	In a shared repository, maintain rules related to master data objects. Maintain data dictionaries and glossaries about customer master data. Capitalize on and share information and data models. Record the point of consumption, point of acquisition, and flows. Document data stewardship processes; i.e., who is responsible for what, where, when, how around customer data.
Define, Follow, and Monitor Data Retention Rules	Data retention rules are part of the data lifecycle. They make explicit the retention delays of all data, whether the organization should delete data after a defined period (e.g., data privacy) or should keep it for a while (e.g., tax reports).
C-MDM Operation	
Receive and Manage Requests about Customer Data (GDPR, etc.)	The C-MDM team receives a request from the business they serve, such as broker or agent reclamation, a customer request about their data, quality issues sent by operational services. They must then route it to the right person/steward.
Assess and Manage Master Customer Data Quality Issues	Detect data quality issues and manage them, respecting the C-MDM team's SLAs. This function is not part of the "Data Quality and Business Value Measurement" because it is not about quality measurement. This function delivers the process to detect concrete data quality issues and correct them, which is part of the core C-MDM operations.
Correct Data Issues	Restore the truth and update data according to it.
Validate Customer Merges	Manually arbitrate the merge of two customers: process the verification of the two customers so that they become one person and eventually merge.
Ensure Availability of Appropriate C-MDM Data & Services for Operational Business Processes	Ensure that operational business users (consumers or producers) have appropriate interfaces to optimally support their process. Be aware of business needs and pain points.
Extension of C-MDM Scope and Usages	
Support New Integration of a Data Producer Application	This activity is BAU. The way of supporting new integration of a data producer application is well managed and the process to extend sustainably of the C-MDM scope with new populations is mastered.
Support New Integration of a Data Consumer Application	This activity is BAU. The way of supporting new integration of a data consumer application is well managed. The new consumption of data is well contextualized with clearly-defined use-cases.
Support New External Data Source Integration	This activity is BAU. The way of supporting new integration of an external data source in the C-MDM ecosystem is well managed.

Business Function	Description
Communication	
Communicate about C-MDM Value and Current Usages	Communicate on KPIs demonstrating the value of C-MDM capabilities (cost/complexity reduction, generated qualitative/quantitative gains, satisfaction score from customer data users, etc.).
Communicate the Customer Data Team Services Catalog	Communicate the customer data team services catalog to the potential internal users of these services.
Communicate about Customer Data Quality	Communicate quality dashboard to operational or executive stakeholders.
Data Quality and Business Value Measurement	
Build & Deliver Reports on Customer Data Quality and Evolutions	Design a quality dashboard with executives, businesses, business data owners, and data stewards. Implement the quality dashboards according to their requirements.
Report C-MDM Business Value	Identify and implement relevant KPIs with businesses (including the IT team).
External Customer Data Acquisition	
Analyze the Data Market and Identify Opportunities	Analyze the market around providing party/contact data (potential partners or data providers). Manage and transform these opportunities to develop new ways of improving customer data quality.
Set Up a Data Partnership	Set up a partnership or extend an existing partnership to some data exchange, in order to improve data quality as a result of these external sources. Operationally manage data partnerships over time.
Purchase External Customer Data	Mutualize contact/party data purchasing management and integration of this external data in the customer data management processes.

5.2.2 C-MDM Detailed Application Functions Reference Model

Detailed descriptions of theses application functions are provided below.

Application Function	Description
Acquisition – User Interface or Services	
Create Master Data	Create a new customer master record including a unique internal organization key.
Update Master Data	Update an existing customer master record.

Application Function	Description
Manage Links between Master Entities and Other Business Objects	Create, update, and delete links between parties, and potentially between parties and other objects (e.g., roles in a contract, offer, etc.) that are managed as a replica in C-MDM. Delete operations must be conducted carefully to avoid breaking integrity or creating orphan records.
Import Data Set, Mass Acquisition	Data acquisition on more than one party (scheduled batch integrations or manual through the user interface, etc.).
Enhance Data with Technical and Business Metadata (especially data lineage)	When data is integrated, manage metadata, such as version of customer data instance, source of update/creation, date of last update.
Quality Management	
Manage Business Validation and Workflows	Provide functionalities to support validation processes on master data when this is relevant.
Manage & Apply Attribute Value Quality Rules	Add, update, and delete rules on attribute values to ensure quality and detect format, consistency, and business issues.
Manage Syntax, Format, Normalization Rules	Manage format validations, and set normalization rules (e.g., based on normalized code/categories list) or standardization rules (based on standard formats).
Manage Consistency and Business Rules	Manage consistency rules (e.g., consistency between two attribute values) and business rules.
Manage Accuracy	Manage accuracy level/confidence score (e.g., based on external source comparisons).
Manage Uniqueness Rules	Set of functions to detect duplicates, and propose best ways of de-duplicating rows.
De-duplicate with Determinist Process	De-duplicate with a designed automated workflow with rules.
De-duplicate with Probabilist/Statistic Process	De-duplicate based on pure statistical comparisons.
Recommend Merges to Arbitrate (matching)	Based on determinist or statistical process, system recommends potential merges to arbitrate (not 100% confidence).
Manual Merge Arbitrage	The decision of the merge is validated by an authorized person through a user interface.
Unmerge	Support a manual unmerge.
Generate Unique Master Technical Key	Manage unique technical key attributions.

Application Function	Description
Data Privacy Management	
Manual Deletion	Search and delete data manually (accessibility of this functionality should be very restricted) .
Automated Deletion with Rules	Automated deletion or proposals for deletions based on set rules.
Masking	Masking data depending on privacy requirements.
Distribution – User Interface or Service	
Search	Search person based on different criteria. Criteria could be based on attributes (e.g., business key, name, surname, contact data, etc.), or related object attribute (e.g., policy number, claim number, etc.).
Read Data	Read data values about a given person.
Extract Data	Extract data values about a scope of persons (based on filters, for example) by batch or manually.
Push Events: Create, Update, Define, Merge	Push events of data evolutions, configure distribution of the events to relevant consumers (e.g., to allow targeted system to update data locally, or to generate some alerts to be tackled over operational processes).
Contextualize Data Visibility & Provided Values	Configure the data visibility rules for each profile/role. Contextualize the given values depending on the consumer (e.g., mastering consumptions in a multiple language context).
Administration	
Design Master Data Model	Have tools to manage the physical design of the data model to fit C-MDM requirements.
Manage Data Historization and Versioning Rules	Configure the way the data is historized and manage versioning of master data, with potential purge of old data if it is relevant.
Full Database Versioning	Version overall database (structure and contents) and historize the version.
Manage Roles	Configure roles of users (human or systems).
Manage Access Rights	Manage and apply access rights to data and to operations (create, update, merge, etc.).
Manage Logging	Manage the way C-MDM systems log different events.
Dashboards & Reports	
Generate Audit Data Evolution Reports	Generate report to have a good understanding of data evolution, on a given scope of persons and time (e.g., indicators by data producer systems, by roles, by persons, etc.).

Application Function	Description
Generate Data Quality Dashboards	Generate dashboards about current quality, and evolution of data quality (completeness, accuracy, etc.).
Generate Usages Dashboard	Generate dashboard to understand how C-MDM is used (e.g., frequency of use, number of update events distributed, etc.).
Reference Data Management	
Input Attributes Transcoding	Manage transcoding rules from input attribute values to standardized MDM values.
Manage Reference List	Manage reference normalized lists (e.g., countries list, status list, etc.).
Manage Customer Key Mapping	Manage mapping between keys (e.g., operational application customer ID *versus* master customer ID).
Output Attributes Transcoding	Manage transcoding rules from standardized MDM attribute values to consumer application attribute values.

TOGAF® Series Guide (2022)

6 Integration Methodologies

6.1 Integration of C-MDM Systems

6.1.1 C-MDM Integration Styles

The three main C-MDM integration styles are: consolidation, cooperation, and centralized. They are described below.

These three styles are not mutually-exclusive and can be used in combination. If the organization has not started their IT journey with the centralized style, it will have to start with consolidation and evolve toward cooperation.

Consolidation Style (or Downstream MDM)

All systems have their own data. Updates are pushed to the MDM hub where they are consolidated. This is the simplest to implement because it has no impact on operational systems. It is the first step for organizations starting with the C-MDM capability. The BI & Analytics capability can be a simple solution here. It addresses the need for after-the-fact analyses and reporting that require customer data. Nevertheless, as there is no single point of truth for the customer and no impact on operational processes, data quality will be difficult to improve, and merge will not be possible.

Cooperation Style

All systems have their own data. Updates are pushed to the MDM hub where they are consolidated. The "new version of the truth" is then forwarded to all connected systems that can synchronize with the master through the customer ID attributed by the C-MDM application. The cooperation style is more complex from an IT point of view, but it allows a high level of control on customer data quality and governance in information systems that have been built in silos. In this style, the MDM capability is supporting user interfaces, workflows, and transactions to create, update, merge, or delete customer data that cannot be addressed by BI & Analytics platforms. A specific transactional application is therefore required.

Centralized Style

No system has its own data. The MDM is the only system that stores the master data. All connected systems retrieve the master data from the MDM: identifying the customer with the same data and when creating a new customer and checking capability if it already exists in the repository. In the same way as the cooperation style, this MDM capability must support user interfaces, workflows, and transactions and will be supported by a specific transactional application.
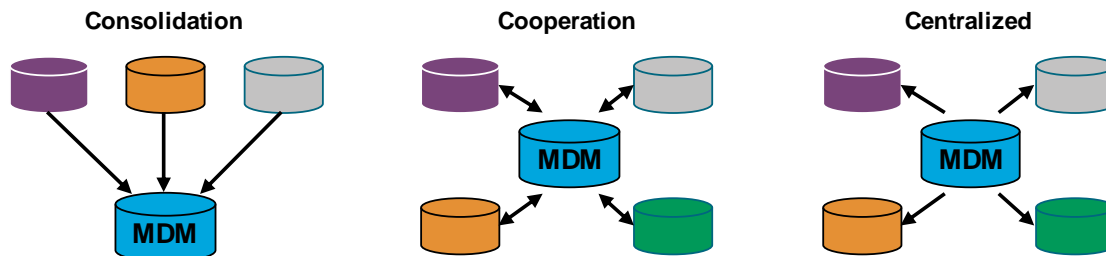
Figure 8: C-MDM Integration Styles

6.1.2   Related Integration Patterns

To support these different integration styles, activation of different integration patterns is needed for each style.

The examples below show exchange patterns between:

- Point of Acquisition (PoA) of customer data and the C-MDM system

- C-MDM system and the Point of Consumption (PoC) of customer data

The relevant pattern to use depends on the usage or business process it needs to support and might be different from one operational application to another.

Note: The scope of these patterns is about maintaining or consuming the latest updated information about the customer held by the C-MDM system. Nevertheless, applications could consume older instances/versions of customer data stored in the C-MDM system or keep their own version of the truth (e.g., age used to calculate the premium, etc.) if this is relevant for their specific usages.
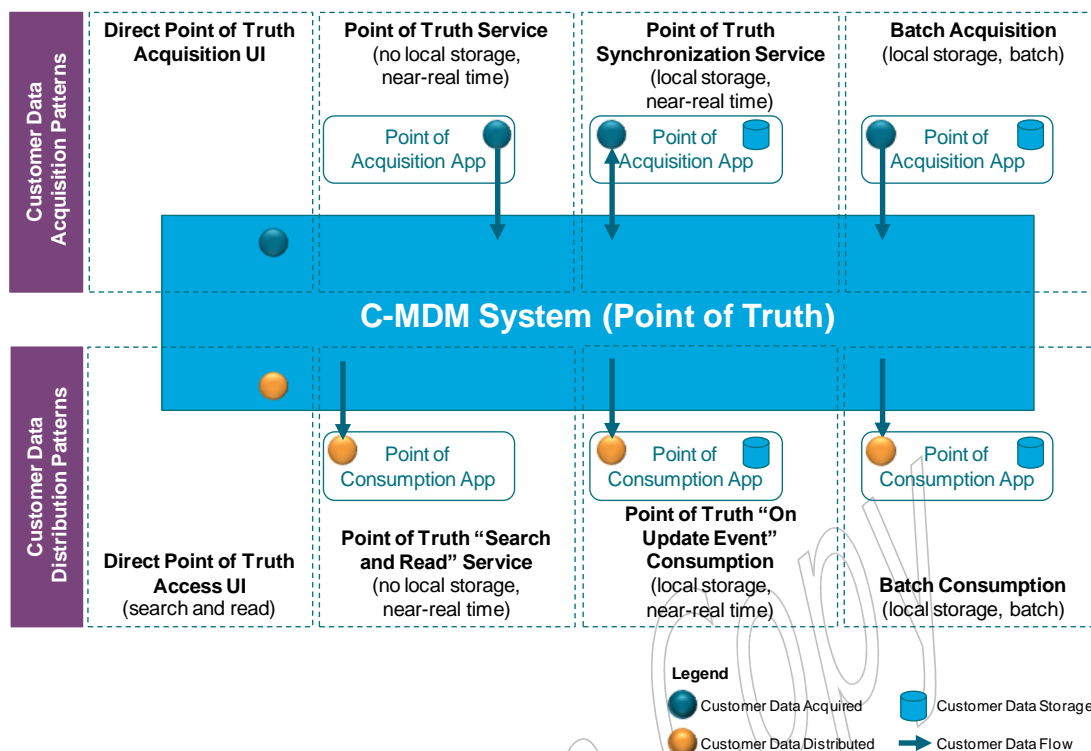
TOGAF® Series Guide (2022)

Figure 9: C-MDM Data Exchange Types Overview

These different kinds of patterns can be put in place step-by-step in the different stages of the C-MDM architecture roadmap.

Patterns to use for the consolidation style are the Batch Acquisition and Batch Consumption patterns (see Figure 10), which support customer data consolidation for analytical use-cases but not operational interactions with customers.
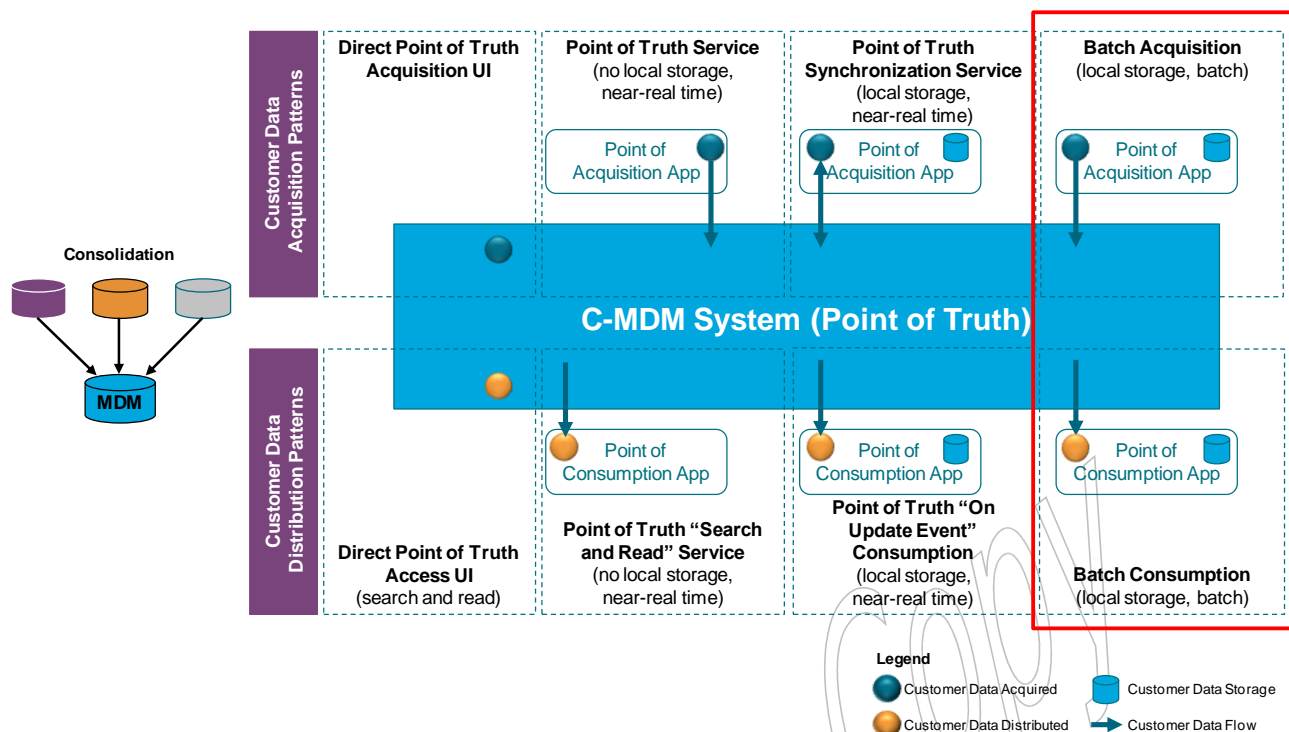
Figure 10: Consolidation Style Patterns

Patterns to use for the cooperation style are data "synchronization" management between primary source and replica sources (see Figure 11). The operational system should be aligned to the C-MDM truth, but potentially with some latency regarding exchange frequencies/architecture.
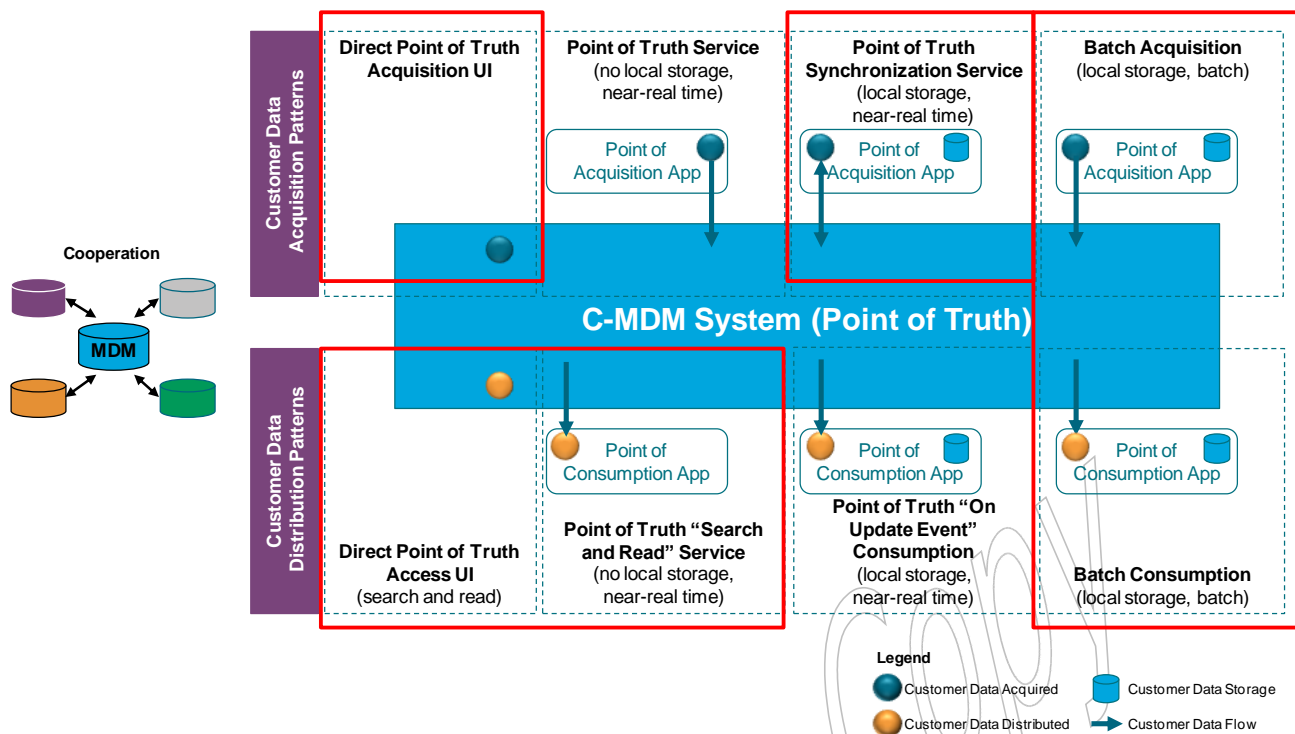
Figure 11: Cooperation Style Patterns

Patterns to use for the centralized style are:

- All updates and creation initiated through the C-MDM point of truth system and services

- The same functional coverage as the cooperation style, but less complex and more efficient from the architecture and information system operations point of view

New applications requiring customer data must rely on the C-MDM system from the start. Changing the mastership of customer data in an application that was not designed with this principle is far more complex and not always possible with reasonable effort.
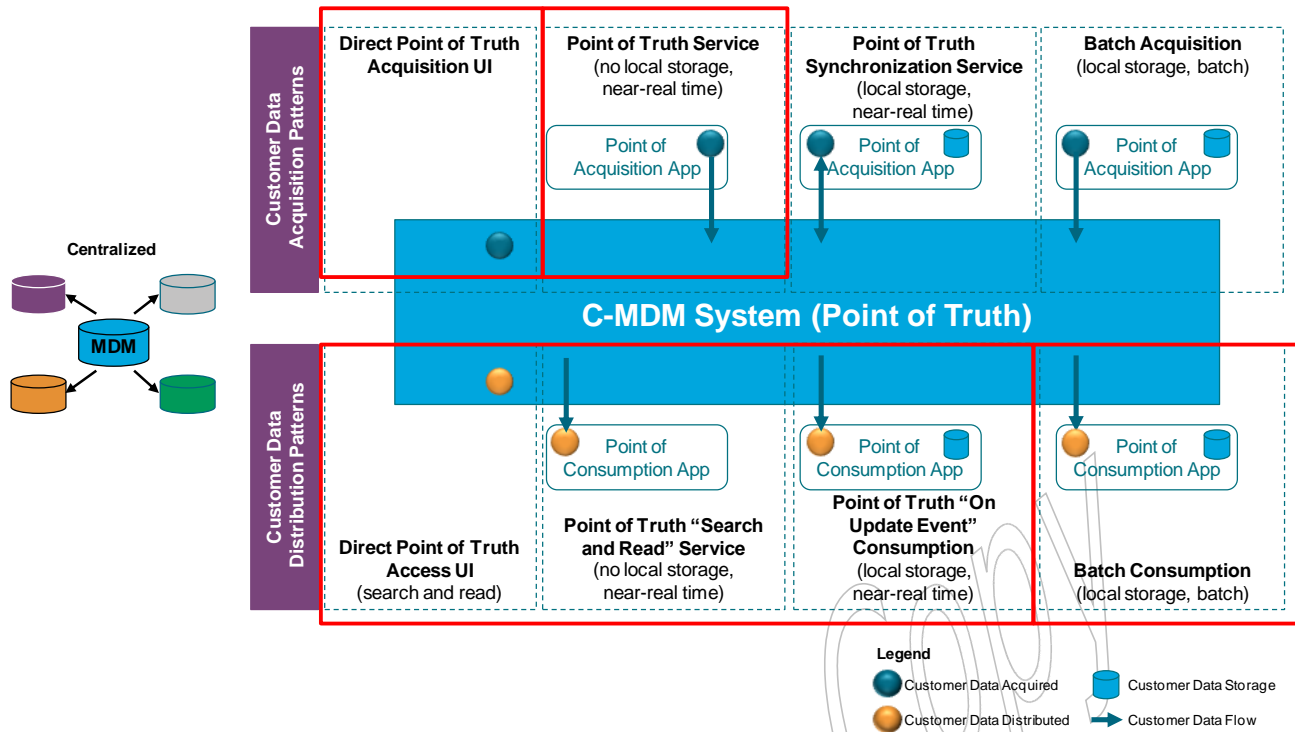
Figure 12: Centralized Style Patterns

TOGAF® Series Guide (2022)

A   Operational Challenges & Diagnostic

A.1   Operational Challenges

The intent of this section is to highlight the main challenges any organization will probably face to some extent when implementing a C-MDM capability. The six challenges listed below synthesize the observations made across several AXA entities. Despite the various business profile, organization, size, information system and legacy, geography, culture, data and IT teams in charge of C-MDM implementation projects, they all faced these to some extent.

- **Challenge 1**: Showing concrete value/benefits of C-MDM, first with a specific business scope and then extended

- **Challenge 2**: Having a dedicated team in charge of ensuring customer data quality, match-and-merge with appropriate governance and stewardship

- **Challenge 3**: Agreeing on customer data ownership between customer/company/distributors/brokers

- **Challenge 4**: Defining the list of kernel party data attributes to master in the MDM

- **Challenge 5**: Implementing customer data quality management through business processes review

- **Challenge 6**: Aligning multiple business, IT, and data roadmaps

A.2   Operational Diagnostic

The intent of this section is to share a list of ten questions that can be used to perform a high-level diagnostic of the level of implementation of a C-MDM capability in one organization.

1. Use-cases are managed and drive the C-MDM vision and roadmap.

2. C-MDM is identified as a key accelerator for consent and preferences data management in a centralized way. In addition, C-MDM accelerates customer requests management around data regulations (GDPR, local regulations, transparency).

3. Strategic vision and roadmap of C-MDM is defined.

4. A C-MDM team is in place and responsible for operational data management for customer data (composed of architects, data managers, data stewards, and data owners).

5. Business data owners are identified for customer data and contribute to C-MDM activities.

6. Data stewards for customer data are identified in the C-MDM team.

7. Data documentation around customer data is shared and maintained.

8.   Data issues, including quality issues, around customer data are captured and well managed.

9.   The level of quality of customer data is measured and used to manage the data quality improvement process.

10.   External data is a key axis of C-MDM improvement,

TOGAF® Series Guide (2022)

Acronyms and Abbreviations

API	Application Program Interface
B2B2C	Business to Business to Consumer
BAU	Business-as-Usual
BI	Business Intelligence
C360	Customer 360 (view)
CDO	Chief Data Officer
CIO	Chief Information Officer
CKO	Chief Knowledge Officer
C-MDM	Customer Master Data Management
CRM	Customer Relationship Management
CTO	Chief Technology Officer
ESB	Enterprise Service Bus
ETL	Extract, Transform, Load
FTC	Federal Trade Commission (US)
GDPR	General Data Protection Regulation
KPI	Key Performance Indicator
NGO	Non-Governmental Organization
PIPEDA	Personal Information Protection and Electronic Documents Act
PoA	Point of Acquisition
PoC	Point of Consumption
SLA	Service-Level Agreement

Index