



Master Data Management and Data Virtualization

CTO

08/03/2018



Dipartimento
dell'Amministrazione Generale,
del Personale e dei Servizi



Abstract

The document presents the main concepts related to Master Data Management (MDM) and Data Virtualization describing in particular, the MDM system and showing the reasons behind its adoption. The presentation highlights then the characteristics and phases of an MDM system summarizing the implications for Cloudify NoiPA.

The document finally shows the data virtualization solution.

Nel documento vengono presentati i concetti principali relativi al Master Data Management (MDM) e Data Virtualization.

In particolare, viene definito il sistema MDM e viene esplicitato il perché dotarsi di un tale sistema. Sono poi evidenziate le caratteristiche e le fasi di un sistema MDM e riassunte le implicazioni in Cloudify NoiPA.

In termini di data virtualization viene descritto cosa consente la soluzione di tale componente.

The Title and Abstract of the document have been translated with the aim to summarize the themes analyzed.

The following content will be presented instead in its original version (Italian)

Agenda

- 1 Definizione dei principali concetti
- 2 Inquadramento degli argomenti nel contesto Cloudify NopiPA

Definizioni

Master Data

I MD in una organizzazione sono una raccolta di **entità comuni**, insieme ai loro attributi e i loro valori, che sono considerati fondamentali per l'attività dell'organizzazione e che sono **necessari per l'uso in due o più sistemi o processi**.

Master Data Management (MDM)

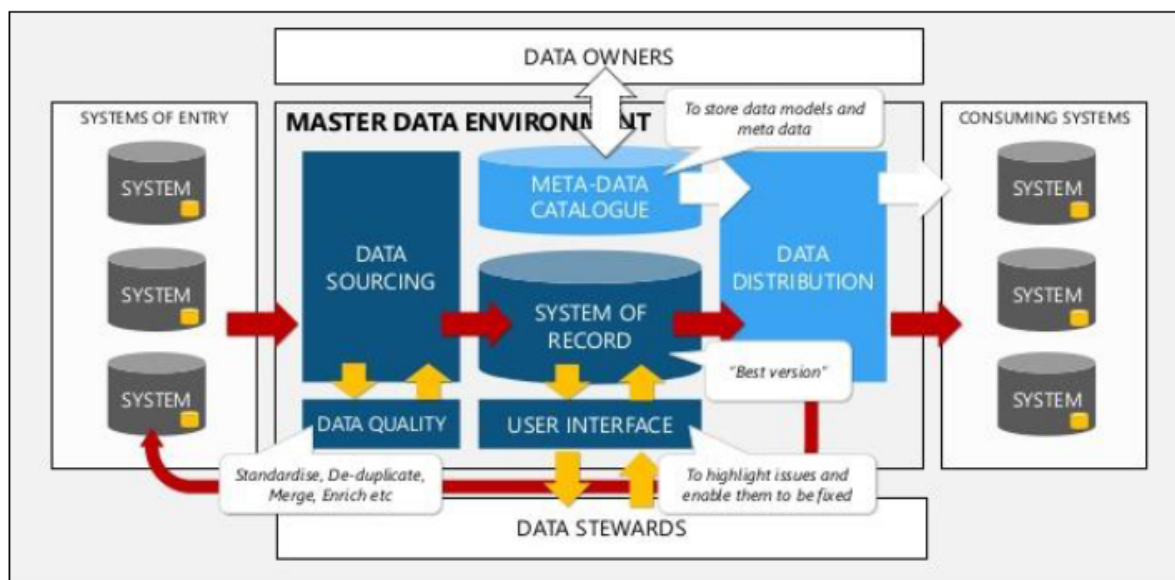
è il **processo** (strutturato e formalizzato) tramite il quale i dati master sono creati, mantenuti e gestiti come asset dell'organizzazione.

Un sistema di MDM è implementato per garantire che i dati master siano validati come **corretti, coerenti, completi e distribuiti** per l'utilizzo da parte dei processi aziendali, applicazioni o utenti interni o esterni.

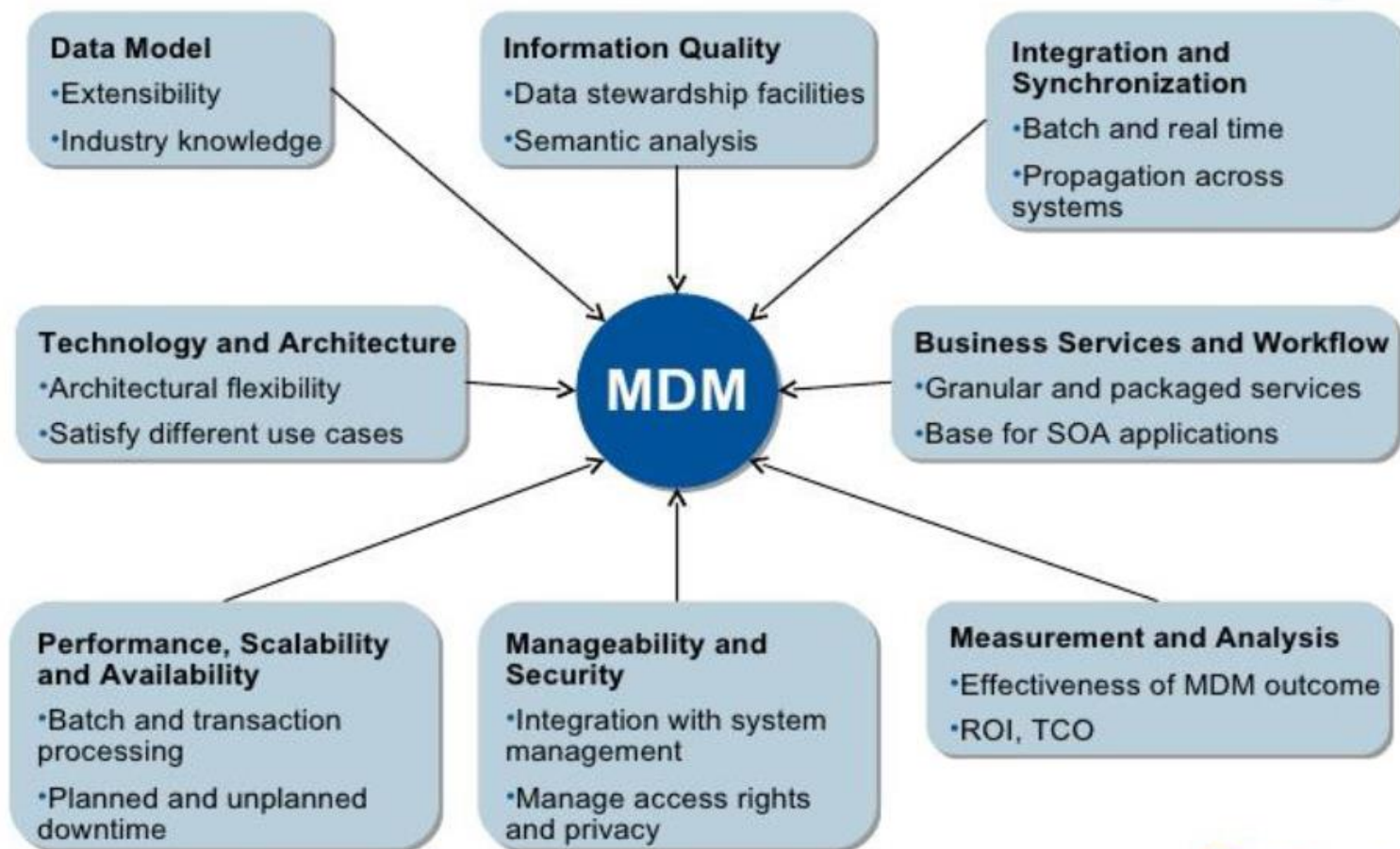
“Sistema” MDM

E' una **soluzione software** che:

- Supporta l'identificazione, il **collegamento e la sincronizzazione** di dati provenienti da / residenti in **fonti eterogenee**
- Crea e gestisce un sistema di record centrale e persistente per i dati master
- Consente la gestione di una «singola visione» di dati master **per ciascun dominio**
- Supporta i requisiti di gestione dei dati master e di governance attraverso il monitoraggio basato su **workflow** e le tecniche correttive
- Facilita l'**elaborazione distribuita** tra più architetture, piattaforme e applicazioni di sistema



Caratteristiche di un MDM



Gartner.

Perché un MDM ?

Dal punto di vista del "business" un MDM è utile quando i MD sono "dispersi" in più sistemi/applicazioni ed hanno tre caratteristiche:

- ❑ **Quantità:**

un'architettura MDM si giustifica per quantità elevate di dati da gestire

- ❑ **Validità nel tempo:**

i MD sono meno volatili dei dati transazionali

- ❑ **Valore:**

i MD hanno valore rilevante per l'organizzazione

Dal punto di vista tecnico un MDM è utile se la soluzione progettuale prevede:

1. Più istanze con più di un DB (con uno stesso schema o schemi diversi)
2. Una istanza con più DB
3. Più "sotto-sistemi" correlati tra loro (es.: gestionale, conoscitivo, documentale)

Fasi di un processo MDM

Step 1: Policy

Vengono individuati i responsabili della definizione dei domini e delle policy, che dovranno realizzare: policy, strategie, obiettivi, metriche, ed i processi di revisione

Step 2: Process

Vengono stabiliti i processi di gestione e le prassi operative per l'utilizzo, la memorizzazione/archiviazione e la protezione dei dati (per persone, applicazioni e servizi) .

Step 3: Controls

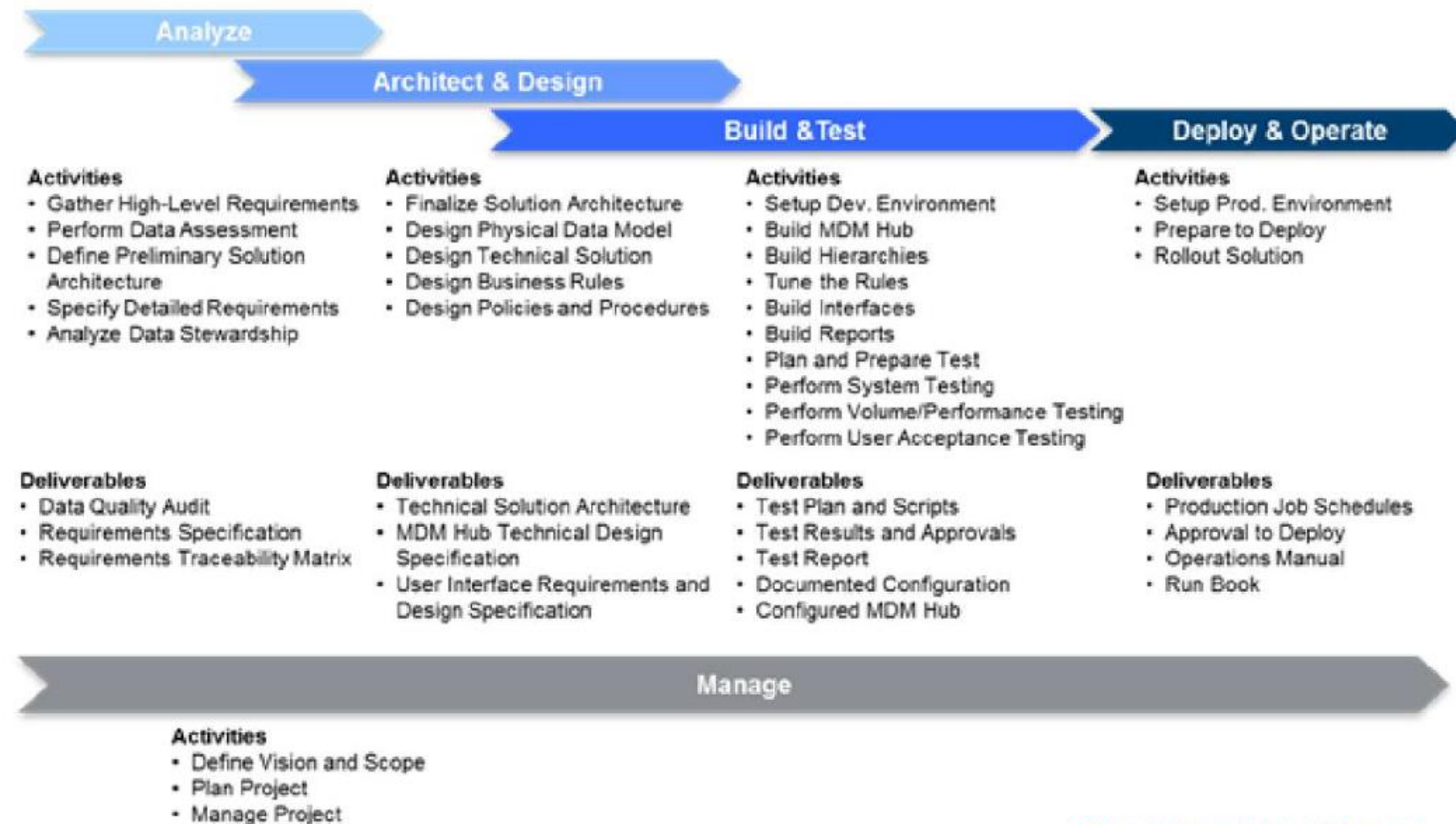
Il responsabile del processo effettua i controlli per l'applicazione delle policy e l'individuazione delle eccezioni

Step 4: Audit

Gli Auditors verificano le performance del Sistema, ed alimentano i processi di governance e revisione delle policy

Fonte: www.informatica.com

Fasi di un progetto MDM



Fonte: www.informatica.com

Cloudify NoiPA - MDM

Per Master Data Management (MDM) si intende un software che aiuta ad identificare **le informazioni fondamentali** all'interno di un'azienda e crea una **singola fonte di dati affidabili** per alimentare i processi di business. Inoltre, permette di **collegare tutti i dati critici** ad un file, chiamato file master, che fornisce un punto di riferimento comune. Se implementato correttamente, semplifica la condivisione dei dati tra personale e reparti. Il MDM **può facilitare l'elaborazione distribuita tra più architetture, piattaforme e applicazioni di sistema.**

Scopo dell'MDM è quello di gestire i dati in modo tale da renderli:

- Fruibili ed aperti ad integrazioni con applicazioni esistenti e future;
- Certificati e affidabili;
- Centralizzati e non ridondati, evitando al minimo i disallineamenti;
- Completi ed estendibili (es. dati di natura giuridica + natura economica);
- Monitorabili durante il ciclo di vita.

Scopo finale di un MDM è, quindi, quello di far diventare il sistema una **fonte autorevole ed attendibile di dati.**

Per raggiungere lo scopo previsto, il sistema comprenderà un insieme completo ed esaustivo di caratteristiche per la gestione dei dati, la maggior parte delle quali raggruppabili in quattro macro categorie: Data Services, Data Quality, Data Governance e Data Integration.

Strumenti di assicurazione qualità nel ciclo di sviluppo

Un progetto che include una soluzione MDM dovrebbe prevedere (esempi):

FASE DI PROGETTAZIONE HL:

- Verifica dei requisiti di base di utilità di una soluzione MDM (es.: multi-tenancy, segregazione dati/SW/infrastrutture)
- Verifica che nel piano del progetto sia prevista la definizione di processi, organizzazione e tecnologie per lo sviluppo e gestione del MDM

FASE DI PROGETTAZIONE DI DETTGLIO:

- Verifica dell'esistenza di una architettura HL e di specifiche line guida per orientare la progettazione di dettaglio e lo sviluppo delle applicazioni con MDM
- Verifica che siano stati previsti e realizzati deliverables specifici per la soluzione MDM, ad es: elenco e gerarchie MD, sottosistemi e fonti alimentanti, Schema ORS, regole di qualità dei dati
- Verifica (anche a campione, con audit e peer review) del rispetto dei requisiti HL

FASE DI SVILUPPO:

- Verifica (anche a campione, con audit e peer review) del rispetto delle line guida
- Verifica periodica delle regole MDM (es.: qualità dei dati)

Data Virtualization

- E' una soluzione che consente a un'applicazione di accedere ai dati senza conoscerne le caratteristiche tecniche, il formato o l'ubicazione fisica.
- Rende il SW indipendente dalla tecnologia sottostante (DBMS).
- Consente l'accesso a fonti eterogenee.

La componente di Data Virtualization permette di **disaccoppiare l'accesso ai dati** dai servizi applicativi, mettendo a disposizione dello strato software superiore delle API "ready-to-use".

Tale componente consentirà di:

- *accedere* a dati da più origini dati **eterogenee** tra loro;
- *combinare e trasformare* facilmente i dati in **viste e modelli di dati virtuali** riutilizzabili;
- *rendere* i dati **unificati e facilmente accessibili** tramite interfacce con standard "open".

Strumenti di assicurazione qualità nel ciclo di sviluppo

Un progetto che prevede la virtualizzazione dei dati dovrebbe prevedere:

FASE DI PROGETTAZIONE High Level:

- Verifica dei requisiti di base di utilità di una soluzione di virtualizzazione (es.: no vendor lock-in, molteplicità di fonti dati cui accedere per costruire informazioni complesse).

FASE DI PROGETTAZIONE DI DETTAGLIO:

- Verifica della presenza di adeguate line guida per l'accesso ai dati con le tecnologie adottate ad uso dei team di sviluppo.

FASE DI SVILUPPO:

- Verifica (anche a campione, con audit e peer review) del rispetto delle line guida.