
Gestió de dades mestres

PID_00251620

Josep Curto Díaz

Temps mínim de dedicació recomanat: 4 hores



Cap part d'aquesta publicació, inclòs el disseny general i la coberta, pot ser copiada, reproduïda, emmagatzemada o transmesa de cap manera ni per cap mitjà, sigui aquest elèctric, químic, mecànic, òptic, de gravació, de fotocòpia, o per altres mètodes, sense la prèvia autorització escrita dels titulars del copyright.

Índex

Introducció	5
Objectius	6
1. Gestió de dades mestres	7
1.1. La necessitat de la gestió de dades mestres	7
1.2. Què és <i>master data management</i> ?	8
1.3. Tipus de dades mestres	10
1.4. Quins beneficis aporta la gestió de dades mestres?	11
2. Programa de gestió de dades mestres	14
2.1. Fases d'un programa de gestió de dades mestres	14
2.2. MDM en el context de govern de la dada	17
2.3. Avaluació de la maduresa de la gestió de dades mestres	18
2.3.1. Nivell 1: inicial	18
2.3.2. Nivell 2: reactiu	19
2.3.3. Nivell 3: gestionat	20
2.3.4. Nivell 4: proactiu	20
2.3.5. Nivell 5: estratègic	21
3. Desenvolupant un programa de gestió de dades mestres	23
3.1. Enfocaments del programa MDM	23
3.2. Com decidir què cal mantenir i administrar	25
3.3. Crear i mantenir una llista mestra	27
3.3.1. Creació d'una llista mestra	27
3.3.2. Manteniment d'una llista mestra	28
3.4. Informació per domini	29
3.4.1. Client B2B	30
3.4.2. Client B2C	30
3.4.3. Producte	31
3.4.4. Dades de referència	31
3.4.5. Localització	32
3.4.6. Organització	32
3.4.7. Proveïdor	33
3.4.8. Empleat	33
4. Tècniques i tecnologies per a la gestió de dades mestres	34
4.1. Tècniques per a la gestió de dades mestres	34
4.2. Tecnologies per a la gestió de dades mestres	35
4.3. Exemple de cas d'ús	37

4.4. MDM i <i>big data</i>	38
Resum	40
Glossari	41
Bibliografia	42

Introducció

El model de dades de les organitzacions s'està fent més i més complex, en haver de donar suport al que coneixem com la transformació digital. Aquesta transformació es tradueix en què els actius digitals empresarials estan distribuïts en múltiples sistemes d'informació —dispositius i en el núvol— i desacoblats en diferents capes d'abstracció i serveis; al mateix temps, han de ser analitzats mitjançant tècniques analítiques i han de complir els marcs de regulació pertinents tant generals com vinculats al sector d'activitat.

Sense una estratègia adequada, freqüentment es generen diverses àrees en la informació que conformen una visió dispersa, fragmentada i incoherent de clients, productes, treballadors, proveïdors i processos.

Per donar resposta a la necessitat de gestionar aquesta complexitat, dins del govern de la dada existeix el que es coneix com la **gestió de dades mestres** (*master data management*), que permet netejar, racionalitzar i integrar les dades en un sistema de registre empresarial per donar suport a totes les funcions bàsiques de l'organització.

En una primera instància, podem entendre que la gestió de dades mestres posa el focus en el principi d'estandardització del govern de la dada.

En aquest mòdul estudiarem què és i en què consisteix la gestió de dades mestres; què és un programa de gestió de dades mestres; com es desenvolupa aquest programa; quins aspectes cal tenir en compte per portar a bon port aquesta iniciativa, i quines tècniques i tecnologies són necessàries per a la gestió de dades mestres.

Objectius

Aquest material didàctic s'adreça a:

- 1) Desenvolupadors i consultors que volen saber què significa gestió de dades mestres o *master data management*.
- 2) Desenvolupadors i consultors que volen ajudar al desenvolupament d'estratègies de negoci que incloguin gestió de dades mestres.
- 3) Gestors que estan interessats en la transformació digital de la seva organització i en la inclusió de gestió de dades mestres com un dels seus pilars fonamentals.

En els materials didàctics d'aquest mòdul trobarem les eines indispensables per assimilar els objectius següents:

- 1) Entendre el concepte de *master data management*, les situacions en què és necessari desplegar una solució d'aquest tipus i els avantatges que proporciona.
- 2) Saber en què consisteix un programa de gestió de dades mestres.
- 3) Enumerar i donar a conèixer millors pràctiques de gestió de dades mestres.
- 4) Conèixer tècniques i tecnologies per a la gestió de dades mestres.

Si bé l'obra és autocontinguda en la mesura del possible, els coneixements previs necessaris són:

- 1) Coneixements bàsics de *business intelligence* i *big data* (en català, dades massives).
- 2) Coneixements sobre estratègia i gestió de les tecnologies de la informació (TI).

S'introduiran els conceptes necessaris per al seguiment d'aquest material.

1. Gestió de dades mestres

1.1. La necessitat de la gestió de dades mestres

La gestió de dades mestres està rebent cada vegada més atenció en el context empresarial. Trobem diferents motius per aquest interès. D'una banda, la dada té un paper fonamental en la generació d'ingressos, en el control de costos i en la mitigació de riscos. Ser capaços de compartir, emmagatzemar, protegir i recuperar les dades rellevants quan és necessari pot derivar en importants avantatges competitiu. En definitiva, les organitzacions han de complir normatives (pròpies del sector i genèriques) i verificar requeriments estratègics de negoci com els següents:

- Amb l'objectiu de capturar, retenir i generar valor dels clients, totes les organitzacions necessiten una visió de 360° de les dades dels seus clients, des de les seves dades de perfil i contracte fins a les transaccions i interaccions amb cadascun dels departaments i canals de venda i suport.
- Amb l'objectiu d'establir una anàlisi completa de despeses, el departament financer ha d'accedir a tota la informació relacionada, amb independència del centre de cost propietari de la dada (per exemple, màrqueting).
- Des del juny del 2007, les organitzacions que realitzen processos industrials han de complir la Llei de la Unió Europea (UE) sobre les substàncies químiques i el seu ús i seguretat, coneguda com REACH (registre, avaluació, autorització i restricció de substàncies i preparats químics). Els objectius d'aquesta llei són: protegir millor la salut i el medi ambient dels possibles riscos derivats de les substàncies químiques; promoure l'ús de mètodes alternatius de comprovació de la seguretat, i millorar la manipulació i l'ús d'aquestes substàncies amb total seguretat en tots els sectors de la indústria.
- En el sector financer, les organitzacions que participen en el sector han de tenir en compte directives com MiFID II (*markets in financial instruments directive*), que harmonitza la regulació sobre els mercats de valors, els instruments financers que en ells es negocien i l'organització i la relació amb els clients de les entitats financeres que presten serveis d'inversió i protecció a l'inversor. En essència, aquesta directiva obliga a una transparència més gran dels mercats, en nom de la major projecció dels inversors, de manera que tot tipus de dada de l'entitat ha de ser consistent, actualitzada i completa.

Totes aquestes responsabilitats empenyen les organitzacions a dedicar temps per gestionar la dada. No obstant això, moltes organitzacions o bé no han desenvolupat aquestes estratègies de forma integrada i transversal, o bé només són iniciatives puntuals o manuals. D'aquesta manera, freqüentment es dediquen a arreglar i a millorar la dada en lloc de fer-la servir i prendre millors decisions. Quan s'arriba a aquesta situació, s'ha arribat al límit de les capacitats de manteniment manual i la dada es converteix en un actiu tòxic per a l'organització, tal com apunta Nassim Nicholas Taleb. El principal problema d'haver d'arreglar la dada de forma recurrent és la pèrdua de confiança dels usuaris de negoci, ja que no poden fer servir la dada per a la presa de decisions quan realment es necessita i té el seu màxim valor.

1.2. Què és *master data management*?

Hi ha moltes formes de classificar les dades. En el context d'aquests materials, introduïrem una de nova. Segons Reichert, Otto i Österle, les dades d'una organització poden dividir-se en dades mestres, dades transaccionals i dades d'inventari. Vegem-les:

- 1) **Dades mestres:** fan referència a les característiques dels objectes bàsics de negoci dins d'una organització. Aquest tipus inclou dades de materials, productes, proveïdors, clients, empleats i actius. Per exemple, respecte al client, el seu nom i cognoms formen part de les dades mestres.
- 2) **Dades transaccionals:** descriuen les transaccions comercials i representen l'entrada i la sortida dels processos empresarials. Es creen constantment en el curs de les activitats empresarials, fan referència a dades mestres i indiquen canvis en les dades d'inventari. Per exemple, imaginem que tenim una cafeteria: la compra d'un cafè i una galeta per esmorzar i les dades vinculades a aquesta compra (import, hora, etc.) són dades transaccionals.
- 3) **Dades d'inventari:** fan referència als nivells d'existències i comptes. Per exemple, si continuem amb el tema de la cafeteria, la quantitat en existència de cafè, en quina data es va comprar o la data de caducitat són dades d'inventari.

Cadascun d'aquests tipus de dades té característiques diferents que podem comparar i que ens permeten comprendre millor les diferències entre ells:

- 1) **Referència temporal:** és la dependència de la dada respecte a un determinat punt en el temps.
- 2) **Freqüència d'actualització:** fa referència a la freqüència amb què una dada s'actualitza o canvia durant un determinat període de temps.

Lectura complementària

N. N. Taleb (2016). *The Bed of Procrustes: Philosophical and Practical Aphorisms*. Penguin.

Lectura complementària

B. Otto; A. Reichert (2010). *Organizing master data management: findings from an expert survey. Proceedings of the 2010 ACM Symposium on Applied Computing* (págs. 106-110). ACM.

3) **Volatilitat de volum:** fa referència al creixement de la dada en volum respecte al temps.

4) **Independència existencial:** és el grau en què la dada fa referència a altres tipus de dades.

A continuació, resumim els tipus de dades i les seves característiques a la taula 1.

Taula 1. Tipus de dades i característiques

Tipus	Referència temporal	Freqüència d'actualització	Volatilitat de volum	Independència existencial
Dades mestres	Baixa	Baixa	Baixa	Alta
Dades transaccionals	Alta	Baixa	Alta	Baixa
Dades d'inventari	Alta	Alta	Baixa	Baixa

Exemple

Exemplifiquem aquestes característiques:

- 1) El nom i els cognoms d'un client, que sabem que són dades mestres, no solen canviar en el temps, pel que està clar que les tres primeres característiques són baixes i la quarta, alta.
- 2) En el cas de la compra del cafè i la galeta, aquesta dada està vinculada a un cert moment en el temps, no s'actualitzarà, depèn de les dades mestres (client i productes).
- 3) En el cas del nivell d'existència de cafè, depèn del temps, s'actualitza després de cada transacció en què es ven un cafè, depèn de les dades mestres i les transaccions, però tan sols canvia un registre.

Tenint en compte aquests tipus de dades, a més de les seves característiques, queda palès que necessitem una disciplina per poder gestionar les dades mestres.

S'entén com **gestió de dades mestres**, o *master data management* (MDM), el conjunt de processos, govern, polítiques, estàndards i eines que defineixen i gestionen constantment les dades crítiques d'una organització per proporcionar un únic punt de referència.

D'aquesta manera:

- La gestió de dades mestres és una peça fonamental per a les operacions i la presa de decisions analítiques.
- Les dades mestres es poden trobar disperses per tota l'organització, i fins i tot fora d'ella.
- La dada mestra estableix una definició estàndard per dades crítiques des de la perspectiva de negoci que es comparteix al llarg de l'organització, i representa l'única font de veritat.

En el context d'MDM, freqüentment es parla del registre d'or (*golden record*), o versió única de la veritat.

S'entén per *golden record* la versió única i ben definida de totes les entitats de dades en un ecosistema organitzacional.

La gestió de dades mestres busca generar aquest tipus de registres dins de l'organització, responent a les preguntes següents:

- Quines són les entitats importants de negoci que es fan servir com a referència repetidament?
- Quines dades descriuen la mateixa entitat de negoci?
- On són emmagatzemades les dades?
- Quines dades són les més exactes?
- Quines dades són més rellevants per a tasques específiques?
- Quines dades de diferents i múltiples fonts d'origen poden ser integrades per crear una vista més completa i comprensible de l'entitat de negoci?
- Quines regles de negoci es poden establir de forma automàtica per combinar, aparellar, identificar, etc.?
- Com identifiquem i restaurem dades que han estat combinades i aparellades de manera incorrecta?

1.3. Tipus de dades mestres

Encara que en general parlem de gestió de dades mestres, la realitat és que hi ha diferents tipus, més enllà de l'àmbit de negoci a què fan referència, que, com ja sabem, poden ser materials, productes, proveïdors, clients, empleats i actius. Podem diferenciar els següents:

- **Dades de referència:** és el conjunt de valors admissibles que han de ser utilitzats per altres camps de dades (mestres o transaccions). Les dades de referència normalment canvien lentament, reflectint canvis en els modes d'operació de l'empresa, en lloc de canviar en el curs normal dels negocis.
- **Dades mestres:** són una font única de dades empresarials comunes utilitzades en múltiples sistemes, aplicacions o processos.

- **Dades mestres corporatives:** fan referència a la font única de dades empresarials comunes utilitzades en tots els sistemes, aplicacions i processos de tota una empresa (tots els departaments, divisions, empreses i països).
- **Dades mestres de mercat:** són la font única de dades comercials comunes per a tot un mercat. Les dades mestres de mercat s'utilitzen entre empreses dins de la cadena de valor. Un exemple de dades mestres de mercat és l'UPC (*universal product code*), que es troba en els productes de consum.

Aquests tipus de dades mestres tenen associats metadades que són rellevants com, per exemple, el significat i el propòsit de la dada; en quines fonts d'origen apareix la dada; la versió de la dada en cada font d'origen; origen de la dada; versió disponible al sistema MDM; data d'actualització de la dada mestra; com s'efectua el manteniment de la dada, i qui és responsable de la gestió de la dada i de la metadada.

1.4. Quins beneficis aporta la gestió de dades mestres?

El *master data management* aporta múltiples beneficis a una organització, que van més enllà de tenir un visió única de la dada. La creació i gestió de dades mestres proporciona:

- Una comprensió consistent dels actius de dades més crítics i comuns de l'organització.
- Mecanismes per a l'ús consistent de les dades mestres al llarg de l'organització.
- Està dissenyat per a la gestió de canvi, cosa que ajuda a una organització a adaptar-se al mercat.

Ara entrarem en els detalls de tres beneficis que aporten valor en un àmbit **operatiu**.

El primer benefici fa referència a **optimitzar i evitar errors** en diferents processos de negoci, en treballar amb dades mestres consistents, fiables i compartides entre tots ells. Entre aquests processos, destaquem la millora de:

- Informes financers.
- Gestió i atenció de clients.
- Gestió de reclams.
- Gestió del compliment regulatori de processos i productes.

- Gestió de la privacitat i la seguretat.
- Gestió de la cadena de subministrament.
- Gestió de la seguretat.
- Vendes creuades i augment del tiquet mitjana.

El segon fa referència al fet que permet **guanyar eficiència i efectivitat**, atès que, per exemple:

- S'evita destinar recursos per gestionar problemes de dades de forma reactiva.
- S'evita danyar la credibilitat amb clients, proveïdors i socis, en poder confiar en les dades compartides amb ells.
- Es redueixen efectes negatius en els ingressos i beneficis, ja que les dades poden reflectir la realitat del negoci, i és possible prendre decisions més eficients.

En tercer lloc, la gestió de dades mestres és una peça angular perquè l'organització **treballi en contextos VUCA**. En particular, centrem-nos en alguns escenaris específics d'aplicació:

- **Gestió del risc:** la volatilitat del mercat, les bretxes en la seguretat cibernètica i fins i tot la inestabilitat de la política internacional estan contribuint a l'augment del risc empresarial. Per mitigar aquest risc, cal comprendre la dada des del seu origen i les seves modificacions. Per exemple, tenint traçabilitat de la dada es pot saber el punt on ha succeït la bretxa i, per tant, quin és el punt més feble de la cadena.
- **Compliment normatiu:** cada any, apareixen nous escenaris que requereixen l'actuació de les agències reguladores, desenvolupant o estenent mesures com HIPAA (*health insurance portability and accountability*, 1996), SOX (Sarbanes-Oxley, 2002) o GDPR (*general data protection regulation*, 2016). Ser capaç de produir informes precisos és més important que mai, sigui quin sigui l'àmbit que l'organisme regulador ocupi: medi ambient, impostos, ocupació, protecció del consumidor, importació/exportació o empleats.
- **Gestió d'inventaris:** els estàndards per a productes, que podem trobar a la taula 2, requereixen dades mestres netes per garantir que les empreses puguin mantenir-se al dia amb els seus inventaris i vendes. Per exemple, imaginem que som una empresa amb botigues físiques a Espanya, i que cadascuna de les botigues compra de forma independent al mateix proveïdor. El fet d'utilitzar la gestió de dades mestres donaria coneixement d'aquesta situació i, per tant, es podria haver negociat el preu amb el proveïdor.

VUCA

Quan parlem de VUCA, fem referència a un acrònim de les condicions de volatilitat (*volatility*), incertesa (*uncertainty*), complexitat (*complexity*) i ambigüïtat (*ambiguity*), que descriuen el mercat actual.

Taula 2. Estàndards en la gestió d'inventaris

Acrònim	Identificador	Es fa servir per identificar:	Exemples
GTIN	<i>Global trade item number</i>	productes i serveis	Llauna de sopa, barra de xocolata, àlbum de música
EAN	<i>European article number</i>	productes i serveis, en un àmbit europeu	Cereals, abrics, aerosol
JAN	<i>Japanese article number</i>	productes i serveis, en l'àmbit del Japó	Arròs, soja, sake
ASIN	<i>Amazon standard identification number</i>	productes i serveis, en el context d'Amazon	Llibre, videojoc, samarreta
ISBN	<i>International standard book number</i>	publicacions	Llibres, revistes
GLN	<i>Global location number</i>	localització	Empreses, magatzems, fàbriques, botigues
SSCC	<i>Serial shipping container code</i>	unitats de logística	Càrregues unitàries sobre palets, gàbies de rodets, paquets
GRAI	<i>Global returnable asset identifier</i>	actiu devolutiu	Palets, caixes
GIAI	<i>Global individual asset identifier</i>	actiu	Equips mèdics, de fabricació, de transport i de TI
GSRN	<i>Global service relation number</i>	relacions entre proveïdor de serveis i receptor	Metges en un hospital, membres de la biblioteca
GDTI	<i>Global document type identifier</i>	documents	Formularis d'enviament, llicències de conducció
GINC	<i>Global identification number for consignment</i>	consignacions	Unitats logístiques transportades juntes en un contenidor oceànic
GSIN	<i>Global shipment identification number</i>	enviaments	Unitats logístiques lliurades a un client de manera conjunta
GCN	<i>Global coupon number</i>	cupons	Cupons digitals
CPID	<i>Component/part identifier</i>	components i parts	Components d'un automòbil

2. Programa de gestió de dades mestres

Desenvolupar una iniciativa de gestió de dades mestres no només consisteix a implementar una solució tecnològica. Cal desenvolupar un programa perquè arribi a bon port, i s'ha de fonamentar en els principis rectors següents:

- Les dades compartides pertanyen a l'empresa i no a un departament.
- MDM és un esforç progressiu i continu.
- Els administradors de dades són l'autoritat responsable per controlar els valors de les dades de referència.
- Els valors dels *golden records* representen el millor esforç de l'organització per tenir la informació més **exacta**, **actual** i **rellevant** de les entitats de negoci clau.

En definitiva, s'entén com **programa de gestió de dades mestres** la metodologia estratègica i sistemàtica per a la identificació, la captura, l'emmagatzematge, la gestió i la propagació de dades mestres en una organització.

Com qualsevol altre projecte, el programa ha de tenir en compte els requisits, les prioritats, la disponibilitat de recursos, el temps i la mida del problema que s'està abordant.

Cal comentar que un programa de *master data management* té punts en comú amb el govern de la dada, amb la qualitat de la dada, i també amb la intel·ligència de negoci. Això vol dir que avenços en aquestes iniciatives repercuteixen positivament en el desenvolupament del programa.

2.1. Fases d'un programa de gestió de dades mestres

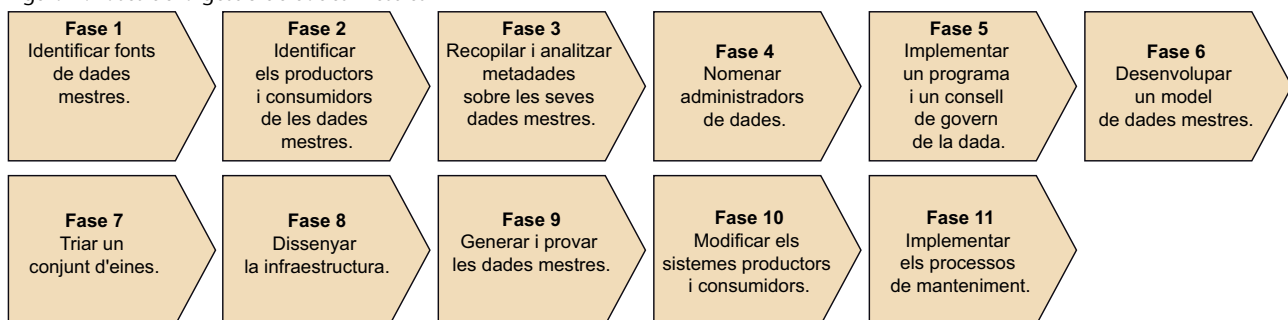
Encara que un programa MDM d'una empresa a una altra pot tenir diferents etapes, la majoria dels projectes MDM inclouen, almenys, aquestes fases:

- identificació de fonts de dades mestres,
- identificació dels productors i consumidors de les dades mestres,

- recopilació i anàlisi de metadades sobre les seves dades mestres,
- denominació d'administradors de dades,
- implementació d'un programa i d'un consell de govern de la dada,
- desenvolupament d'un model de dades mestres,
- elecció d'un conjunt d'eines,
- disseny de la infraestructura,
- generació i prova de les dades mestres,
- modificació dels sistemes productors i consumidors,
- implementació dels processos de manteniment.

Aquestes fases mínimes són seqüencials, tal com s'il·lustra a la figura 1.

Figura 1. Fases de la gestió de dades mestres



Font: Josep Curto

A continuació, explicarem amb detall aquestes etapes:

- **Identificar fonts de dades mestres:** en aquest pas, es crea un catàleg de totes les fonts de dades (bases de dades, fitxers, etc.) que contenen les dades de materials, productes, proveïdors, clients, empleats i actius. Aquest pas freqüentment resulta molt revelador, atès que es descobreix que les dades solen estar disperses en múltiples fonts que el departament de les tecnologies d'informació (TI) no sabia que existien.
- **Identificar els productors i consumidors de les dades mestres:** després de tenir el catàleg de dades, el pas següent és saber quins actors en l'organització generen i consumeixen dades. Per exemple, quines aplicacions/-persones produeixen les dades mestres identificades en el primer pas i quines aplicacions/persones les utilitzen. En alguns casos, aquest punt pot ser més senzill si tots els canvis es detecten i es manegen en l'àmbit de base de dades.
- **Recopilar i analitzar metadades sobre les seves dades mestres:** aquest pas consisteix a generar coneixement sobre les fonts de dades. Per a totes les fonts identificades en el primer pas, quines són les entitats i els atributs de les dades, i què signifiquen? Respondre a aquesta pregunta significa aconseguir el nom de l'atribut, el tipus de dades, els valors permesos, les restriccions, les opcions per defecte, les dependències i qui posseeix la

definició i el manteniment de les dades. Freqüentment, determinar qui és el propietari de la dada resulta summament difícil i posa de manifest una necessitat.

- **Nomenar administradors de dades:** en l'etapa anterior, haurà quedat patent quines dades tenen propietaris i administradors. Un administrador de dades, també conegut com *data steward*, és una persona amb el coneixement de les dades font actuals i la capacitat de determinar com transformar la font en el format de dades mestres. En general, els administradors han de ser nomenats pels propietaris de cada font mestra de dades, pels arquitectes responsables dels sistemes MDM i per representants dels usuaris empresarials de les dades mestres.
- **Implementar un programa i un consell de govern de la dada:** la implementació d'MDM necessita del govern de la dada. És summament aconsellable que hi hagi un procés de presa de decisions respecte a la dada ben definit, ja que caldrà prendre decisions sobre com es mantenen les dades mestres, què contenen, quant de temps es mantenen i com s'autoritzen i auditen els canvis.
- **Desenvolupar un model de dades mestres:** normalment, aquesta és una de les fases més importants i difícils. Aquí cal determinar com són les dades mestres. És a dir, quins atributs s'inclouen, quina mida i tipus de dades són, quins valors es permeten, etc. També aquí s'inclou l'assignació entre el model de dades mestres i les fonts de dades actuals. La dificultat, en aquest punt, rau en el fet que si s'inclouen tots els atributs d'origen en l'entitat mestra, sovint s'acaba amb un model de dades mestres massa complex i molest com per ser útil.
- **Triar un conjunt d'eines:** la implementació de la gestió de dades mestres requereix eines de suport per a l'emmagatzematge de la dada, el tractament de la dada, la seva neteja i la seva gestió. Hi ha eines especialitzades en MDM (que integren diferents components necessaris), o que s'apliquen tan sols en un dels conjunts de dades, com pot ser pel client, conegudes com *customer data integration* (CDI) o pel producte, conegudes com *product information management* (PIM); o bé és possible aprofitar eines existents en l'organització d'àmbits com la integració de dades o la qualitat de la dada. Com veurem en l'apartat 4, el conjunt d'eines ha de suportar múltiples funcions, entre les quals destaquem també mantenir versions i jerarquies de dades mestres.
- **Dissenyar la infraestructura:** les dades mestres s'han d'emmagatzemar en una infraestructura que permeti exposar-les a les aplicacions i proporcionar al mateix temps processos per administrar-les i mantenir-les. Hi ha diferents opcions per al disseny, com veurem en l'apartat 3. El focus d'aquesta infraestructura és la fiabilitat i l'escalabilitat.

- **Generar i provar les dades mestres:** en aquest pas, s'utilitzen les eines que han estat seleccionades i es combinen les dades d'origen per confirmar les llistes de dades mestres. Durant el procés, freqüentment es requereix l'ajust de regles de negoci en descobrir, per exemple, excepcions. Tot i que les eines han avançat considerablement, la inspecció manual pot ser necessària per assegurar que els resultats són correctes i complir amb els requisits establerts per al projecte. Aquest procés es coneix com **cura de dades**.
- **Modificar els sistemes productors i consumidors:** la implementació d'un programa MDM en una organització implica freqüentment canviar els sistemes que produeixen, mantenen o consumeixen dades mestres per treballar amb la nova font de dades mestres. Si les dades mestres s'utilitzen en un sistema separat dels sistemes d'origen (un magatzem de dades, per exemple), és possible que els sistemes font no hagin de canviar. No obstant això, si els sistemes font faran servir les dades mestres, probablement hi haurà canvis requerits. Els sistemes font hauran d'accedir a les noves dades mestres o les dades mestres hauran de sincronitzar-se amb els sistemes font, de manera que els sistemes font tinguin una còpia de les dades mestres netes que s'utilitzaran. En aquest punt, les tècniques d'integració de dades tindran un paper fonamental.
- **Implementar els processos de manteniment:** MDM és un procés continu en el qual s'ha d'assegurar i mantenir la qualitat de les dades mestres. Aquesta fase consisteix a establir mecanismes per reconèixer els problemes (per exemple, assenyalar coincidències qüestionables), simplificar les correccions (per exemple, l'eina ha de proposar solucions i ajudar a la ràpida identificació dels errors) i revisar l'historial de canvis (per exemple, per aïllar la font d'errors i desfer canvis incorrectes o revisar els elements que es van afegir com a nous, perquè els criteris de coincidència eren a prop però per sota del llindar).

2.2. MDM en el context de govern de la dada

En el context de govern de la dada, la gestió de dades mestres és una funció més. Com ja sabem, cada funció té diferents activitats (planificació, control, desenvolupament i operatives), cadascuna de les quals és efectuada pel rol corresponent.

Per a la gestió de dades mestres, aquestes activitats són:

- Gestió de les dades mestres i de referència (activitat de planificació).
- Comprensió de les necessitats d'integració de dades mestres i de referència (activitat de planificació).
- Definició de l'arquitectura d'MDM (activitat de desenvolupament).
- Implementació de la solució d'MDM (activitat de desenvolupament).
- Validació de dades de control i de referència (activitat de control).

- Integració de dades mestres (activitat operativa).
- Replicació de dades mestres i de referència (activitat operativa).
- Manteniment de les jerarquies dimensionals (activitat operativa).

El programa de gestió de dades mestres cobreix aquestes funcions, que formen part del marc més general.

2.3. Avaluació de la maduresa de la gestió de dades mestres

Quan hem parlat de les diferents fases que componen un programa de gestió de dades mestres, hem comentat que no totes les organitzacions passaran per les mateixes etapes. Això és així perquè freqüentment ja han començat a desenvolupar algunes de les capacitats necessàries per dur a terme el programa.

Aplicar un model de maduresa en aquest àmbit permet saber quines capacitats ja han estat desenvolupades i en quina mesura, i suggerir quines no són presents o s'han de sofisticar per continuar el desenvolupament del programa.

Presentarem un model de maduresa per a la gestió de dades mestres fonamentat en el tradicional model de maduresa de capacitats (CMM), desenvolupat per l'Institut d'Enginyeria de Programari de la Universitat Carnegie Mellon i creat per David Loshin.

El model de maduresa té cinc nivells: inicial, reactiu, gestionat, proactiu i estratègic; i cada nivell s'avalua respecte a sis capacitats: arquitectura, govern, gestió, identificació, integració i gestió de processos de negoci. Una organització pot trobar-se respecte a cadascuna de les capacitats en un nivell diferent, i no necessàriament totes en el mateix nivell.

2.3.1. Nivell 1: inicial

En aquest nivell, les capacitats es caracteritzen per:

- **Arquitectura:** l'arquitectura d'aplicacions es defineix per a cada aplicació empresarial. Hi ha una consolidació empresarial limitada de models representatius i col·leccions de diccionaris de dades en diverses formes, però sense models de dades mestres.
- **Govern:** es porta a terme una neteja limitada de dades per aplicació / línia de negoci, amb finalitats específiques (per exemple, normalització d'adreces), però hi ha una absència de models definits de propietat de dades. Emergeix el reconeixement de la necessitat de supervisió.

- **Gestió:** es porta a terme la gestió d'identitats per aplicació quan és necessari (per exemple, amb clients). Es fa la configuració de l'aplicació, però no es coordina per mitjà de la gestió centralitzada.
- **Identificació:** ús limitat de la gestió d'identitat per línia de negoci. S'intenta consolidar les dades del client segons ho requereixin les aplicacions (per exemple, actualitzacions de programari o transició d'aplicacions comptables).
- **Integració:** hi ha còpies replicades de dades de referència, però amb una reutilització limitada de dades, i no hi ha reutilització de serveis d'aplicacions.
- **Gestió de processos de negoci:** participació de negoci limitada o no existent, excepte en el nivell més alt de definició de requisits.

2.3.2. Nivell 2: reactiu

En aquest nivell, les capacitats es caracteritzen per:

- **Arquitectura:** s'intenta reunir els diccionaris de dades en un únic repositori. Comença l'exploració inicial en serveis d'aplicacions de nivell baix i la revisió d'opcions per compartir informació (per exemple, integració d'informació empresarial o integració d'aplicacions empresarials).
- **Govern:** es fan servir aplicacions externes per administrar metadades, i s'introdueix la gestió de la qualitat de les dades per a l'anàlisi, la normalització i la consolidació.
- **Gestió:** s'assignen recursos per administrar l'ús de conjunts d'eines introduïts per a la gestió de dades mestres. La capacitat per al desplegament d'eines i tecnologia de l'empresa fa que les capacitats estiguin disponibles en una base més àmplia. Hi ha una administració centralitzada de metadades i índexs mestres.
- **Identificació:** s'habilita la cerca d'identitats i coincidències utilitzades per reduir la duplicació, a més de la cerca d'identitat i concordança utilitzades per a la vinculació de registres rudimentaris amb fins domèstics.
- **Integració:** exploració inicial de la consolidació de dades per a aplicacions analítiques recentment desenvolupades (per exemple, CRM). El magatzem de dades és utilitzat com un repositori central per a dades mestres. No hi ha integració de nou en les aplicacions contribuents.
- **Gestió de processos de negoci:** es descriuen models de processos empresarials conceptuals i emergeix l'ús inicial de regles empresarials integrades

CRM

Quan parlem de CRM, *customer relationship management*, fem referència a les estratègies, a les pràctiques i a les tecnologies que habiliten la gestió de les interaccions amb un client al llarg del seu cicle de vida.

en aplicacions. També apareix la integració analítica d'aplicacions de dades consolidades.

2.3.3. Nivell 3: gestionat

En aquest nivell, les capacitats es caracteritzen per:

- **Arquitectura:** hi ha un model de dades de mestres definit per la persistència, i una arquitectura fonamental per al repositori mestre compartit. També hi ha un marc operacional identificat per a activitats de cicle de vida de les dades mestres de baix nivell, i serveis definits per a la integració amb el repositori mestre.
- **Govern:** eines de qualitat de dades al seu lloc. Polítiques i procediments per a la gestió de la qualitat de les dades. Seguiment de problemes de qualitat de dades. Processos d'estàndards de dades en el lloc. Línia d'administració de dades empresarials.
- **Gestió:** gestió d'identitats centralitzada en l'índex mestre i utilitzada en nombroses aplicacions. Jerarquies identificades (dins d'una classe de dades) i utilitzades per les aplicacions analítiques. Configuració avançada i administració de l'ús de les dades mestres. Un pla de migració està disponible per a aplicacions seleccionades.
- **Identificació:** serveis de recerca d'identitat disponibles per a les aplicacions. Enllaç de registre integrat a la capa de servei MDM. Normes de fusió i consolidació estandarditzades i gestionades sota control centralitzat. Fusió i consolidació de processos establerts i repetibles.
- **Integració:** serveis de components disponibles per a la integració d'aplicacions. Els serveis sincronitzen les aplicacions amb el repositori.
- **Gestió de processos de negoci:** integració de regles de negoci amb operacions de dades mestres. Connectivitat fonamental entre aplicacions empresarials i objectes de dades bàsiques. Els analistes de processos de negoci participen en els requisits d'enginyeria de dades mestres.

2.3.4. Nivell 4: proactiu

En aquest nivell, les capacitats es caracteritzen per les raons següents:

- **Arquitectura:** s'estableixen models mestres. Existeix la capacitat per passar del marc de dades mestres a la transacció. Així mateix, hi ha una arquitectura orientada a serveis (SOA), establerta per l'arquitectura d'aplicacions, que es combina amb la gestió centralitzada de metadades empresarials.

- **Govern:** programa de gestió de dades empresarials en vigor amb normes de dades empresarials i gestió de metadades en el lloc. El monitoratge proactiva per al control de qualitat de dades alimenta el programa de governabilitat.
- **Gestió:** gestió d'identitats completament integrada i establerta a tota l'empresa, amb identificació única de totes les instàncies d'objectes mestres. La gestió de la jerarquia de cicle complet dona suport tant a les activitats analítiques com a les operacionals. L'administració de la jerarquia permet el restabliment d'errors de consolidació de falsos positius.
- **Identificació:** els serveis per al cicle de vida de dades integren recerca, coincidència i resolució d'identitats. Totes les operacions del cicle de vida de les dades són estructurades sobre els serveis de fusió i consolidació. Aquesta consolidació es produeix majoritàriament de manera automàtica i, de forma puntual, acurada per experts de negoci.
- **Integració:** sincronització completament integrada en els serveis de cicle de vida. La capa de components admet la integració d'aplicacions en un àmbit d'objecte mestre. SOA impulsa la integració d'aplicacions empresarials.
- **Gestió de processos de negoci:** la lògica empresarial es reutilitza. Les regles de negoci estan integrades en un motor de regles i disponibles en el nivell de processos empresarials. Analistes de negoci integrals per al desenvolupament d'aplicacions. Relacions personalitzades amb els clients. Processos de negoci automatitzats.

2.3.5. Nivell 5: estratègic

En aquest nivell, les capacitats es caracteritzen per:

- **Arquitectura:** hi ha una completa integració de tota l'arquitectura empresarial amb el sistema MDM que propaga els canvis de forma automàtica entre les diferents fonts.
- **Govern:** el govern de dades mestres està integrat amb les altres iniciatives de govern de la dada com una funció més, i assegura un intercanvi d'informació d'alta qualitat.
- **Gestió:** hi ha una gestió transparent de la identitat de tots els objectes de dades sincronitzades, amb representacions internes i externes.
- **Identificació:** hi ha mecanismes automàtics per a la resolució de conflictes entre dades mestres. S'ha vinculat el rendiment empresarial amb les dades mestres.

- **Integració:** el desenvolupament d'aplicacions de negoci es vincula amb les dades mestres, i es tenen en compte les necessitats d'integració amb el sistema MDM.
- **Gestió de processos de negoci:** MDM completament integrat en els processos de negoci, tant en un àmbit operacional com analític. Destaca l'ús per al perfilat de clients, proveïdors i productes.

3. Desenvolupant un programa de gestió de dades mestres

Tot i que és cert que l'aplicació d'un programa de gestió de dades mestres serà més efectiva si s'aplica a totes les dades mestres, l'abast, el cost, l'esforç i els riscos associats a un projecte en un àmbit corporatiu són barreres infranquejables i difícils de justificar en una fase inicial. Per això, és molt més senzill començar amb només algunes fonts de dades mestres i anar expandint progressivament el projecte. Aquest enfocament permet demostrar l'èxit de la iniciativa i, al mateix temps, consolidar el coneixement i les lliçons apreses.

D'aquesta manera, desenvolupar un programa d'MDM comporta:

- Identificar un enfocament per al programa.
- Identificar quines dades mestres s'inclouen dins de la iniciativa.
- Crear i mantenir una llista mestra.

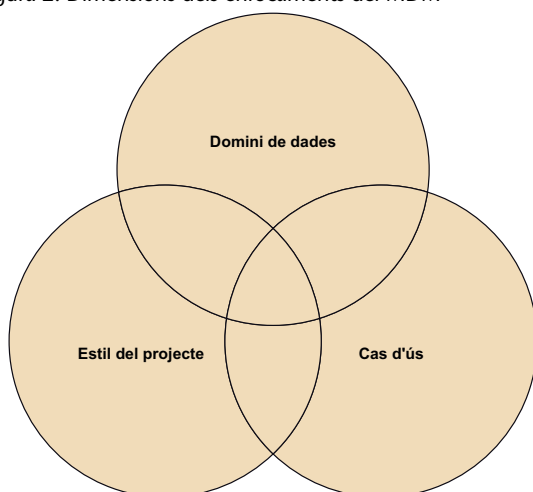
Tractem amb detall aquests punts en els subapartats següents.

3.1. Enfocaments del programa MDM

Tenint en compte la necessitat d'iniciar el programa de forma acotada, cal identificar l'enfocament que tindrà el programa de gestió de dades mestres.

Troblem diferents dimensions per l'enfocament, tal com il·lustra la figura 2.

Figura 2. Dimensions dels enfocaments del MDM



Font: Josep Curto

Lectura complementària

A. Haug; J. Stentoft; F. Zachariassen; J. Schlichter (2013). «Master data quality barriers: an empirical investigation». *Industrial Management & Data Systems* (núm. 113, vol. 2, pàg. 234-249).

Definir el programa comporta definir un enfocament combinat d'aquestes tres dimensions:

- **Domini de dades (mestres):** aquesta dimensió s'enfoca en el domini de dades que cal considerar en el projecte, si ha de ser acotat o complet.
- **Acotat:** fa referència a un conjunt limitat, però tancat de dades. Aquestes dades poden ser de clients (B2B o B2C), de producte, de dades de referència, de localitzacions, de proveïdors, etc.
- **Complet:** fa referència a totes les dades mestres de l'organització.
- **Estil del projecte:** aquesta dimensió posa el focus en l'estil del projecte MDM. Tenim, principalment, sis estils:
 - **Estil de registre:** en aquest tipus de projecte, els diferents sistemes d'origen publiquen les seves dades i el sistema MDM emmagatzema només els identificadors (ID) de sistema d'origen, les claus foranes (ID de registre en sistemes d'origen) i els valors de dades clau necessaris per a la coincidència. El sistema executa els algorismes de neteja i aparellament, i assigna identificadors globals únics als registres coincidents, però no envia cap dada de nou als sistemes d'origen. En aquest cas, el sistema s'apalanca a la federació de dades per crear una vista de les dades mestres.
 - **Estil de consolidació:** els registres mestres es consoliden físicament en un sistema central. La creació de les dades segueix estant distribuïda per mitjà dels sistemes operacionals, i les dades mestres poden actualitzar-se a partir d'esdeveniments, però no es garanteix que estiguin actualitzats. Les dades mestres, en aquest cas, generalment no s'usen per a transaccions, sinó que donen suport a la generació d'informes.
 - **Estil de coexistència:** el sistema MDM inclou dades mestres escrites i emmagatzemades en nombrosos sistemes operacionals, a més d'un registre mestre físicament emmagatzemat en el sistema central i dades mestres harmonitzades en tota la cartera d'aplicacions de l'organització.
 - **Estil de transacció:** el sistema MDM emmagatzema, millora i manté tots els atributs de dades (mestres) rellevants. Es converteix en la font autoritzada de la veritat, i publica aquesta informació valuosa de nou en els respectius sistemes d'origen. El sistema MDM publica i escriu els diferents elements de dades en els sistemes font, després que els algorismes de vinculació, de neteja, d'aparellament i d'enriquiment hagin fet la seva feina. Els sistemes operacionals estan subscriptes a actualitzacions.
 - **Estil de transacció adaptable:** és similar a l'estil de transacció, però a més proporciona la capacitat de respondre a diferents sol·licituds d'informació i de procés en tota l'empresa. El sistema consolida informació interna i ex-

B2B i B2C

Quan parlem d'un negoci i els seus clients, podem trobar-nos empreses que només comercialitzen amb empreses (B2B, acrònim de *business to business*) o que ho fan amb consumidors (B2C, acrònim de *business to consumer*).

terna, i proporciona serveis a sistemes operacionals, analítics o tots dos. Proporciona informació en temps real i en llinatge (o traçabilitat) de canvis.

- **Estil de confederació:** trobem diferents sistemes MDM que es mantenen en un àmbit departamental/territori/agència, i que estan connectats a un sistema central. Cada sistema, en un àmbit de domini, pot dissenyar-se amb un dels estils anteriors, però el central sol ser de registre. Aquest enfocament acostuma a aplicar-se en el sector públic, on no és possible emmagatzemar dades de diferents agències de forma conjunta.
- **Cas d'ús:** hi ha dos casos d'ús:
 - **Operacional:** que proporciona serveis d'MDM per a sistemes operacionals.
 - **Analític:** que proporciona serveis d'MDM per a sistemes analítics, com el magatzem de dades.

L'enfocament de la iniciativa MDM és la suma, per tant, dels tres tipus de dimensions anteriors, tenint en compte sempre en quin sector s'opera i si es proporciona el servei a una àrea determinada o a tota l'organització.

3.2. Com decidir què cal mantenir i administrar

Un dels punts rellevants és quines entitats de dades mestres han de formar part del programa, ja que no totes les dades que s'ajusten a la definició de dades mestres s'han de gestionar necessàriament com a tals.

Hi ha diversos criteris que cal tenir en compte per justificar si una dada ha de ser tractada com una dada mestre:

- **Comportament:** les dades mestres es poden identificar per la forma en què interactuen amb altres dades. Les dades mestres normalment estan relacionades amb les dades transaccionals, seguint una relació substantiu/verb. Les dades transaccionals capten verbs (per exemple, una venda) i els mestres, els substantius (per exemple, client i producte).
- **Cicle de vida:** les dades mestres també es poden identificar per la forma en què es creen, es llegeixen, s'actualitzen, s'eliminen i es busquen. Aquest cicle de vida es coneix com CRUD, acrònim de *create*, *read*, *update* i *delete*. Aquest procés depèn de les regles empresarials de l'empresa, el segment de la indústria i els sistemes de dades. La taula 3 il·lustra cicles CRUD per a quatre dominis de dades mestres comuns.

Taula 3. Dades mestres segons CRUD

CRUD	Client	Producte	Actiu	Empleat
Crear	Visita del client (via web o a la botiga); compte creat	Producte comprat o creat	Unitat adquirida; procés d'aprovaçió	Contractació, orientació, etc.
Llegir	Vistes contextualitzades basades en els credencials de l'usuari	Inventari, catàleg de producte	Informes periòdics, verificació, depreciació	Accés a l'oficina, revisions, etc.
Actualitzar	Adreça, descompte, telèfon, preferències, compte de crèdit	Canvis en els materials o embalatge	Transferència, accidents, manteniment	Matrimoni, promoció, transferència
Esborrar	Mort, fallida, liquidació	Cancel·lat, reemplaçat, no disponible	Obsolet, venut, destruït, robat	Acomiadament, mort
Cercar	CRM, centre d'atenció telefònica, gestió de contactes	ERP, sistema de comandes	Gestió d'actius	Sistema de RR. HH.

- **Cardinalitat:** a mesura que la cardinalitat d'una dada decreix, la probabilitat que sigui considerada mestra descendeix. Per exemple, si una empresa té una quantitat limitada i reduïda de clients, els beneficis d'implementar una solució de gestió de dades mestres seran limitats.
- **Vida de la dada:** les dades mestres tendeixen a ser menys volàtils que les dades transaccionals. La durada de la dada depèn del sector. Un tipus de dada en una indústria particular pot considerar-se una dada mestra (en termes de baixa volatilitat), mentre que, en d'altres, no. Per exemple, un contracte per pintar una casa és una transacció, ja que és un acord únic i de curta durada, mentre que un contracte per representar un artista pot ser considerat una dada mestra —ja que es tracta d'un acord a llarg termini— i pot incloure dades no volàtils.
- **Complexitat:** les dades mestres solen ser complexes, pel que requereixen més atenció i cura en la seva gestió. Aquells actius que són simplement guardats i emmagatzemats (el que es pot entendre com a baixa complexitat) no necessiten ser considerats dades mestres.
- **Valor:** com més valor tingui una dada, més gran és la necessitat de protegir-la i mantenir la seva qualitat, i també la probabilitat que sigui considerada com una dada mestra.
- **Volatilitat:** encara que les dades mestres són menys volàtils, no vol dir que no siguin mestres. Una dada que no presenta aquesta característica no necessita una solució per gestionar-la.
- **Reusabilitat:** les dades mestres es caracteritzen per la seva alta reusabilitat en l'organització. Per exemple, considerem les dades de client que es comparteixen en múltiples aplicacions i departaments. Sense una solució d'MDM, existiran múltiples versions del mateix concepte.

Cardinalitat

Quan parlem de cardinalitat, fem referència al nombre d'elements en un conjunt.

Resumint, el principal repte d'una organització és identificar quines de les seves dades han de considerar com a dada mestra. D'aquesta manera, crear una llista de dades mestres no només es fonamenta en aplicar la definició del concepte, sinó en categoritzar-les en funció del seu comportament i atributs.

3.3. Crear i mantenir una llista mestra

Crear i mantenir les dades mestres és la funció més important del programa de gestió de dades mestres. Cal comprendre de manera detallada aquestes dues tasques.

3.3.1. Creació d'una llista mestra

Per crear dades mestres, hi ha dos passos bàsics:

- Netejar i estandarditzar les dades.
- Combinar les dades de totes les fonts per consolidar-ne duplicats.

Abans, però, de netejar i normalitzar les dades, és absolutament necessari comprendre el model de dades per a les dades mestres. És a dir, cal identificar per a cada dada mestra i de referència el seu contingut i la seva assignació amb el sistema d'origen. Aquesta informació s'utilitza per definir les transformacions necessàries per netejar les dades d'origen.

Per a la neteja de dades i la seva transformació en el model de dades mestres, necessitem tecnologies i processos de qualitat de dades i d'integració de dades. En particular, són similars als processos d'extracció, transformació i càrrega de dades (ETL) utilitzats per carregar les dades a la factoria d'informació i els seus elements (com el *data warehouse*, conegut en català com magatzem de dades).

De fet, en algunes organitzacions, es fan servir freqüentment les eines ETL disponibles, en lloc d'implementar o aprendre una nova eina. Sigui quin sigui l'escenari, algunes de les funcions típiques que cal fer seran:

- **Normalitzar formats de dades:** consisteix a establir i mantenir un format comú per a les dades mestres. Per exemple, que tots els números de telèfon segueixin el mateix format (codi del país i número).
- **Reemplaçar els valors que falten:** consisteix a completar les dades mestres amb valors predeterminats. Per exemple, completar els codis postals d'una direcció a partir d'una base de dades especialitzada.
- **Estandarditzar els valors:** consisteix a unificar i estandarditzar valors amb l'objectiu de poder comparar-los. Per exemple, transformar els preus en una moneda comuna.
- **Mapa d'atributs:** consisteix a analitzar una dada per identificar els diferents atributs que conté. Per exemple, en el nom de contacte, el nom i cognom.

Lectura complementària

J. Curto (2017). *Introducción al Business Intelligence*. Barcelona: Editorial UOC.

Hi ha situacions en què no serà possible automatitzar completament el procés de neteja de dades. En aquest moment, serà necessari un processament manual, però aquest procés s'ha de fonamentar en la cura per part dels experts de negoci.

Les dades de qualitat formaran part de la taula (o les taules) de dades mestres. Els errors formaran part d'una altra taula i, a mesura que cada font es neteja, la sortida ha de ser examinada per assegurar que el procés estigui funcionant correctament.

3.3.2. Manteniment d'una llista mestra

Hi ha diferents aproximacions per gestionar i utilitzar dades mestres. Revisarem tres dels escenaris més comuns:

- **Còpia única:** en aquest enfocament, només hi ha una còpia mestra de les dades mestres. Tots els canvis en el model de dades mestres es duen a terme sobre aquesta còpia i, a partir d'aquesta, es propaguen a les aplicacions.
- **Múltiples còpies, manteniment únic:** en aquest enfocament hi ha una còpia mestra, però en els sistemes font s'emmagatzema una còpia. Aquestes aplicacions no poden canviar o afegir dades mestres, però poden actualitzar les parts de les dades que no formen part de les dades mestres.
- **Combinació contínua:** en aquest enfocament, les aplicacions poden canviar la seva còpia de les dades mestres. Aquests canvis es remetent a la còpia mestra, on es fusionen amb el model mestre.

Cadascun d'aquests enfocaments té beneficis i inconvenients que es resumeixen en la taula 4.

Taula 4. Escenaris de manteniment

Enfocament	Beneficis	Inconvenients
Còpia única	Es garanteix la consistència de la dada.	El procés de còpia a les fonts d'origen pot derivar en la modificació dels esquemes que, en certs casos, és impossible.
Múltiples còpies, manteniment únic	Es garanteix la consistència de la dada, s'evita la modificació d'esquemes.	Serà necessari modificar les aplicacions per evitar la modificació de les dades mestres, que pot significar la pèrdua de funcionalitats.
Combinació contínua	Es minimitzen els canvis en aplicacions, al mateix temps que s'evita la modificació d'esquemes i se'n garanteix la consistència.	Poden augmentar els conflictes d'actualització, consistència i reconciliació de dades, i s'han de definir més regles.

Sense importar com es dugui a terme el manteniment, és important entendre d'on venen els canvis, quin procés s'ha seguit, quin és l'estat anterior i a quina jerarquia de dades pertany. Això vol dir que el manteniment requereix capacitats de **gestió de versions**, **auditoria** i **gestió de jerarquies**.

El fet que el sistema MDM gestioni els punts anteriors permet:

- Identificar errors i tornar a versions anteriors vàlides de la dada mestra.
- Tenir traçabilitat del procés de canvi de les dades.
- Fer servir el sistema MDM per propagar un canvi de jerarquia en una dada mestra.

3.4. Informació per domini

Quan parlem dels dominis de dades mestres, cal considerar que tenim dades en diferents nivells i que conformen el coneixement sobre un cert domini. Aquesta estratificació va des de la informació bàsica del domini, que podem entendre com les dades mestres bàsiques, a les observacions, les transaccions i les interaccions que fan referència al comportament. Tenim:

- Informació bàsica del domini.
- Coneixement estès del domini.
- Informació financera, legal i reguladora.
- Observacions, transaccions i interaccions.

Aquestes dades normalment s'obtenen de les fonts d'origen següents:

- **Dades financeres:** aquest tipus de dades es pot obtenir de l'ERP (*enterprise resource planning*), que és un sistema informàtic destinat a l'administració de recursos en una organització.
- **Dades de clients:** aquest tipus de dades es pot obtenir del CRM (*customer relationship management*), que és un sistema informàtic per a l'administració i gestió de la relació amb clients. No obstant això, amb l'explosió de les xarxes socials i la informació de tercers, és possible estendre aquesta informació més enllà dels sistemes interns de l'organització.
- **Dades d'empleats:** aquest tipus de dades es pot obtenir de l'HRM (*human resources management*), que és un sistema informàtic per a l'administració i gestió de la relació amb empleats. També és possible trobar informació en sistemes per a la gestió de projectes. El mateix passa amb la informació del client que és possible d'estendre amb informació de tercers.
- **Dades de productes:** aquest tipus de dades es pot obtenir del PLM (*product lifecycle management*), que és un sistema per a la gestió del cicle de vida d'un producte, des del seu disseny fins a la fi de la seva vida útil i el seu reciclatge.
- **Dades de localització:** que poden incloure dades geopolítiques (com països, estats, províncies, etc.) i dades de negoci (com adreces d'oficines, de magatzems, etc.).

3.4.1. Client B2B

En el cas de client B2B, la figura 3 n'il·lustra els nivells de coneixement.

Figura 3. Informació de client B2B

Observacions, transaccions i interaccions	Oportunitats Vendes actuals, previsió, ràtio de conversió, etc.	Social Esdeveniments, notícies, blogs, alertes, canals, PR	Servei Nivell de servei, nivell de risc, etc.	Vendes Factures, terminis, línia de crèdit, etc.	Ordres Quantitat, productes, comandes, etc.	Centre d'atenció telefònica Trucades, satisfacció, mètriques	Contracte Tipus, data, risc, etc.
Informació financera, legal, regulatòria	Legal Propietaris, estructura del capital, etc.	Financera Informes financers, ràtings, etc.	Risc Ràtings, avaluació creditícia, etc.	Mercats Segments, geografia, etc.	Regulació Certificats, regulació, etc.		
Coneixement estès	Canals Web, correu electrònic, telèfons, LinkedIn, etc.	Contactes Decisors, administrador, etc.	Jerarquies Empresa mare, mercat, ràting, etc.	Localitzacions Tipus, localització, certificacions, etc.	Adreces D'enviament, de facturació, etc.		
Informació bàsica	Client B2B Nom legal, any de creació, IBEX 35, etc.						

Font: Josep Curto

3.4.2. Client B2C

En el cas del client B2C, la figura 4 n'il·lustra els nivells de coneixement.

Figura 4. Informació de client B2C

Observacions, transaccions i interaccions	Oportunitats Vendes actuals, previsió, ràtio de conversió, etc.	Social Esdeveniments, notícies, blogs, alertes, canals, PR	Servei Nivell de servei, nivell de risc, etc.	Vendes Factures, terminis, línia de crèdit, etc.	Ordres Quantitat, productes, comandes, etc.	Centre d'atenció telefònica Trucades, satisfacció, mètriques	Contracte Tipus, data, risc, etc.
Informació financera, legal, regulatòria	Conformitat Registre criminal, frau, seguretat, etc.	Salut Registres, malalties conegudes, etc.	Risc Ràtings, avaluació creditícia, etc.	Ocupació Dates, posicions, empleadors, etc.	Legal Identificador, tipus, any, entitat certificadora, etc.		
Coneixement estès	Canals Canal de contacte favorit	Xarxes socials Facebook, Twitter, LinkedIn, Google, etc.	Adreces D'enviament, de facturació, etc.	Relacions Amics, família, fills, etc.	Punts de contacte Correu electrònic, telèfon, etc.		
Informació bàsica	Client B2C Nom, cognom, data de naixement, etc.						

Font: Josep Curto

3.4.3. Producte

En el cas de producte, la figura 5 n'il·lustra els nivells de coneixement.

Figura 5. Informació de producte

Observacions, transaccions i interaccions	Vendes Vendes actuals, previsió, ràtio de conversió, etc.	Social Esdeveniments, notícies, blogs, alertes, canals, PR	Inventari Inventari, magatzems, disponibilitat, etc.	Logística Capacitat, traçabilitat, etc.	Recomanació Detecció de productes similars, etc.	Centre d'atenció telefònica Trucades, satisfacció, mètriques	Campanyes Activitat de màrqueting, etc.
Informació financera, legal, regulatòria	Mercat i preu Preus, models, disponibilitat, etc.	Cicle de vida Cicle de fabricació, empaquetat	Marca i PI Marca, patents, propietat intel·lectual, etc.	Contracte Contractes, terminis, <i>resellers</i>	Compliment regulatori Ambient, salut, etc.		
Coneixement estès	Taxonomia Famílies, categories, classificació, etc.	Variants Colors, mides, dimensions, variacions, etc.	Màrqueting Imatges, <i>media</i> , descripcions, documents, etc.	Atributs de negoci Tipus, localització, certificacions, etc.	Materials Llista de materials, jerarquia, etc.		
Informació bàsica	Producte Identificació, nom, descripció, components, etc.						

Font: Josep Curto

3.4.4. Dades de referència

En el cas de dades de referència, la figura 6 n'il·lustra els nivells de coneixement.

Figura 6. Informació de dades de referència

Observacions, transaccions i interaccions	Anàlisi d'impacte Ús en aplicacions		
Informació financera, legal, regulatòria	Estàndards de la indústria NAICS, SWIFT, ISO, etc.	Regulació Certificats, regulació, etc.	Mapa extern Mapeig de dades internes davant d'externs
Coneixement estès	Taxonomies Jerarquies, classificacions, diccionaris, etc.	Glossari de negoci Definicions, rols, etc.	Mapa intern Mapeig de dades, traduccions, etc.
Informació bàsica	Informació bàsica Codis interns geogràfics, estats, etc.		

Font: Josep Curto

3.4.5. Localització

En el cas de localització, la figura 7 n'il·lustra els nivells de coneixement.

Figura 7. Informació de localització

Observacions, transaccions i interaccions	Vendes Factures, termes, línia de crèdit, etc.	Social Esdeveniments, notícies, blogs, alertes, canals, PR	Campanyes Activitat de màrqueting, mètriques, etc.	Planificació Capacitat, previsió, pressupost, etc.	Web Trànsit, <i>page rank</i> , SEO, etc.	Producció Rendiment, indicadors, etc.
	Informació financera, legal, regulatòria	Marca i PI Marca, patents, propietat intel·lectual, etc.	Contractes Acords, lloguer, socis, etc.	Compliment tributari Local, nacional, exportacions, etc.	Regulació Certificats, regulació, etc.	Costos Inversió, manteniment, amortitzacions, etc.
Coneixement estès	Geografia Localització, adreça, etc.	Horaris Horaris, calendari, obertura, tancament, etc.	Màrqueting Imatges, <i>media</i> , descripcions, documents, etc.	Contactes Mànager, correu electrònic, telèfons, etc.		
	Informació bàsica	Localització Identificador, nom, descripció, etc.				

Font: Josep Curto

3.4.6. Organització

En el cas d'organització, la figura 8 n'il·lustra els nivells de coneixement.

Figura 8. Informació de l'organització

Observacions, transaccions i interaccions	Talent i habilitats Formació, certificacions, etc.	Activitats socials Esdeveniments, notícies, articles, etc.	Rendiment Informes, avaluacions, etc.	Nòmina Quantitat, data de pagament, etc.	Viatges i allotjament Quantitat, data, freqüència, etc.	Contracte Tipus, data, risc, etc.
	Informació financera, legal, regulatòria	Assegurança i beneficis Assegurances, beneficis, etc.	Sou i bons Rang salarial, impostos, bo, etc.	Salut Revisió anual, malalties conegudes, etc.	Conformitat Registre criminal, frau, seguretat, etc.	Legal Identificador, tipus, any, entitat certificadora, etc.
Coneixement estès	Cost de centre Cost de centre, jerarquia, etc.	Unitat de negoci Unitat, jerarquia, etc.	Historial Dates, posicions, empleadors, etc.	Feines Descripció, responsabilitats, etc.	Jerarquia RH Mànager, responsable, equip, etc.	
	Informació bàsica	Empleat Nom, cognoms, correu electrònic, DNI, passaport, etc.				

Font: Josep Curto

Un possible enfocament consisteix a centrar-se en les persones com a eix central de l'organització, de manera que aquests nivells coincideixen amb els d'empleat.

3.4.7. Proveïdor

En el cas de proveïdors, la figura 9 n'il·lustra els nivells de coneixement.

Figura 9. Informació de proveïdors

Observacions, transaccions i interaccions	Financer	Social	Centre d'atenció telefònica	Vendes	Ordres
	Preu de les accions, actes financers	Esdeveniments, notícies, blogs, alertes, canals, PR	Trucades, satisfacció, mètriques	Factures, terminis, línia de crèdit, etc.	Quantitat, productes, comandes, etc.
Informació financera, legal, regulatòria	Legal	Financera	Risc	Acords	Regulació
	Propietaris, estructura del capital, etc.	Informes financers, ràtings, etc.	Ràtings, avaluació creditícia, etc.	Propietat intel·lectual, SLA, NDA, etc.	Certificats, regulació, etc.
Coneixement estès	Jerarquies	Emplaçament	Contactes	Banc	Adreces
	Mercat, ràtings, grup empresarial	Tipus, localització, certificació, etc.	Decisors, administrador, etc.	Compte, autoritzacions, etc.	Web, correus electrònics, adreça, etc.
Informació bàsica	Proveïdor				
	Nom legal, any, SIC, NAIC, etc.				

Font: Josep Curto

3.4.8. Empleat

En el cas d'empleats, la figura 10 n'il·lustra els nivells de coneixement.

Figura 10. Informació d'empleats

Observacions, transaccions i interaccions	Talent i habilitats	Activitats socials	Rendiment	Nòmina	Viatges i allotjament	Contracte
	Formació, certificacions, etc.	Esdeveniments, notícies, articles, etc.	Informes, avaluacions, etc.	Quantitat, data de pagament, etc.	Quantitat, data, freqüència, etc.	Tipus, data, risc, etc.
Informació financera, legal, regulatòria	Assegurança i beneficis	Sou i bons	Salut	Conformitat	Legal	
	Assegurances, beneficis, etc.	Rang salarial, impostos, bo, etc.	Revisió anual, malalties conegudes, etc.	Registre criminal, frau, seguretat, etc.	Identificador, tipus, any, entitat certificadora, etc.	
Coneixement estès	Cost de centre	Unitat de negoci	Historial	Feines	Jerarquies RH	
	Cost de centre, jerarquia, etc.	Unitat, jerarquia, etc.	Dates, posicions, empleadors, etc.	Descripció, responsabilitats, etc.	Màner, responsable, equip, etc.	
Informació bàsica	Empleat					
	Nom, cognoms, correu electrònic, DNI, passaport, etc.					

Font: Josep Curto

4. Tècniques i tecnologies per a la gestió de dades mestres

Hem revisat fins ara els aspectes vinculats als processos, govern, polítiques i estàndards en el context de la gestió de dades mestres.

En aquest apartat, ens centrarem en les eines, les tècniques i les tecnologies que suporten la gestió de dades mestres.

Les eines per a la gestió de dades mestres han evolucionat de forma considerable en els últims anys. El concepte emergeix vinculat al CRM, l'ERP i el *data warehouse*, que necessiten informació fidedigna i compartida per a la creació d'informes i l'anàlisi de les entitats més rellevants de negoci. No es considera una disciplina independent de la intel·ligència de negoci, i no hi ha eines dedicades. A poc a poc, a mesura que més aplicacions de negoci i fins i tot sistemes operacionals necessiten l'ús de dades mestres, comencen a emergir eines integrades per MDM, vinculades a la informació de clients o la del producte.

Actualment, existeixen ja plataformes dedicades a la gestió de dades mestres. Per aquest motiu, en parlar de tècniques i tecnologies dins d'aquestes plataformes, moltes d'elles ja deuen ser conegudes. En certa mesura, podem considerar la gestió de dades mestres com un cas específic i combinat de totes aquestes, que ha aconseguit la suficient entitat pròpia com per escindir-se de la intel·ligència de negoci.

4.1. Tècniques per a la gestió de dades mestres

Tenint en compte l'evolució de la gestió de dades mestres, no és estrany descobrir que les tècniques que suporten MDM siguin una combinació de tècniques que freqüentment poden trobar-se de forma independent en una organització però que, en aquest context, han de treballar de forma orquestrada i integrada.

Aquestes tècniques inclouen:

- **Tècniques d'integració de dades:** permeten la captura, el processament i la distribució de dades des de fonts d'origen cap al repositori central de dades mestres, i viceversa. Aquestes tècniques inclouen la propagació, la consolidació, la federació de dades i el *change data capture*.

Change data capture

Quan parlem de *change data capture*, fem referència a una tècnica que propaga els canvis en una base de dades en el moment en què ocorren cap a les fonts destinació.

- **Tècniques d'emmagatzematge de dades:** permeten l'emmagatzematge de dades mestres i de referència en un repositori central (i aquelles dades que siguin necessàries en funció de l'enfocament). Aquestes tècniques inclouen la creació de models de dades i l'establiment de regles de negoci.
- **Tècniques de qualitat de dades:** permeten mantenir i millorar la qualitat de les dades mestres i de referència. Aquestes tècniques inclouen el perfilat de dades, l'anàlisi semàntica i l'estadística de dades, l'aparellament per regles de negoci, la semàntica i estadística de dades, i la normalització i estandardització de dades. Part d'aquestes tècniques es fonamenten en algorismes d'aprenentatge automàtic.
- **Tècniques de jerarquies:** possibiliten ordenar i categoritzar les dades mestres de manera flexible. Aquestes tècniques tenen punts en comú amb les tècniques OLAP (en el sentit de definició).
- **Tècniques d'exploració de dades, aplicacions i sistemes:** permeten identificar automàticament les dades d'una organització (presents en múltiples sistemes i aplicacions) i, a més, proposar quines dades han de formar part del repositori MDM. Part d'aquestes tècniques també es fonamenten en algorismes d'aprenentatge automàtic.

OLAP

Quan parlem de OLAP, *online analytical processing*, fem referència al mètode per organitzar i consultar dades sobre una estructura multidimensional.

4.2. Tecnologies per a la gestió de dades mestres

En aquest subapartat, farem referència a la plataforma de gestió de dades mestres com a eina integrada. Aquest tipus de plataformes han de tenir els components següents:

- **Modelització de dades:** component que s'encarrega de la creació i el manteniment del model de dades mestres. Inclou el disseny fonamentat en regles de negoci, el glossari de negoci, el suport per a regles complexes de negoci, etc.
- **Semàntica i qualitat de dades:** component que s'encarrega de la qualitat de dades, incloent-hi la normalització, l'estandardització, l'ús de capacitats semàntiques i estadístiques per l'aparellament i la qualitat de dades, l'enriquiment de dades, etc.
- **Fluxos i processos de negoci:** component per a la gestió del flux de treballs vinculats amb dades mestres, incloent-hi assignacions basades en rols i usuaris, API per a la integració amb sistemes externs, mètriques per a la comprensió dels fluxos i activitats fonamentades en regles de negoci.
- **Govern i definició de polítiques de dades:** component que habilita la creació i el manteniment de polítiques de dades, incloent-hi el perfilat i el descobriment de dades, la creació col·laborativa de regles de dades, el

glossari de definicions, la versió de polítiques i el quadre de comandament de control.

- **Gestió de jerarquies de dades:** component per a la gestió de jerarquies de dades mestres que inclou l'autodescobriment de jerarquies, les jerarquies basades en relacions predefinides, la cerca per jerarquies, el suport per jerarquies múltiples per rol i dada mestre, a més de jerarquies balancejades, desequilibrades i recursives.
- **Administració i aplicació de polítiques de dades:** component que és l'únic punt per a l'aplicació de polítiques de dades mestres, incloent-hi l'autoria i la combinació de dades, l'aplicació de polítiques en un àmbit de registre o atribut, la gestió de cues d'aplicació de polítiques, la creació i la gestió d'esdeveniments i notificacions i mètriques i el quadre de comandament de control.
- **Integració de dades:** component que s'encarrega de la integració de dades mestres, incloent-hi connectors a tot tipus de fonts de dades (bases de dades relacionals, NoSQL, fitxers, aplicacions, API, etc.), ETL, CDC, etc.
- **Arquitectura de dades:** component que s'encarrega de l'emmagatzematge de dades mestres que ha de suportar el desplegament en el núvol o al centre de dades propi de l'organització, a més de l'escalabilitat.
- **Seguretat i confiança de dades:** component que s'encarrega de la seguretat de la dada, incloent-hi la seguretat basada en rols, en un àmbit de registre i camp; la integració amb sistemes corporatius de la gestió de protocols d'accés com LDAP (*lightweight directory access protocol*); l'auditoria de dades; la traçabilitat de la dada, i l'anàlisi de l'impacte.

En el mercat, podem trobar plataformes genèriques que inclouen la gran majoria dels components (com Semarchy*), plataformes que a partir de la integració de dades han evolucionat cap a MDM (com Informatica**), solucions comercials de codi obert (com la de Talend***), i també solucions específiques per a dades de producte (com la de inRiver****) o de client (com la de Pitney Bowes*****).

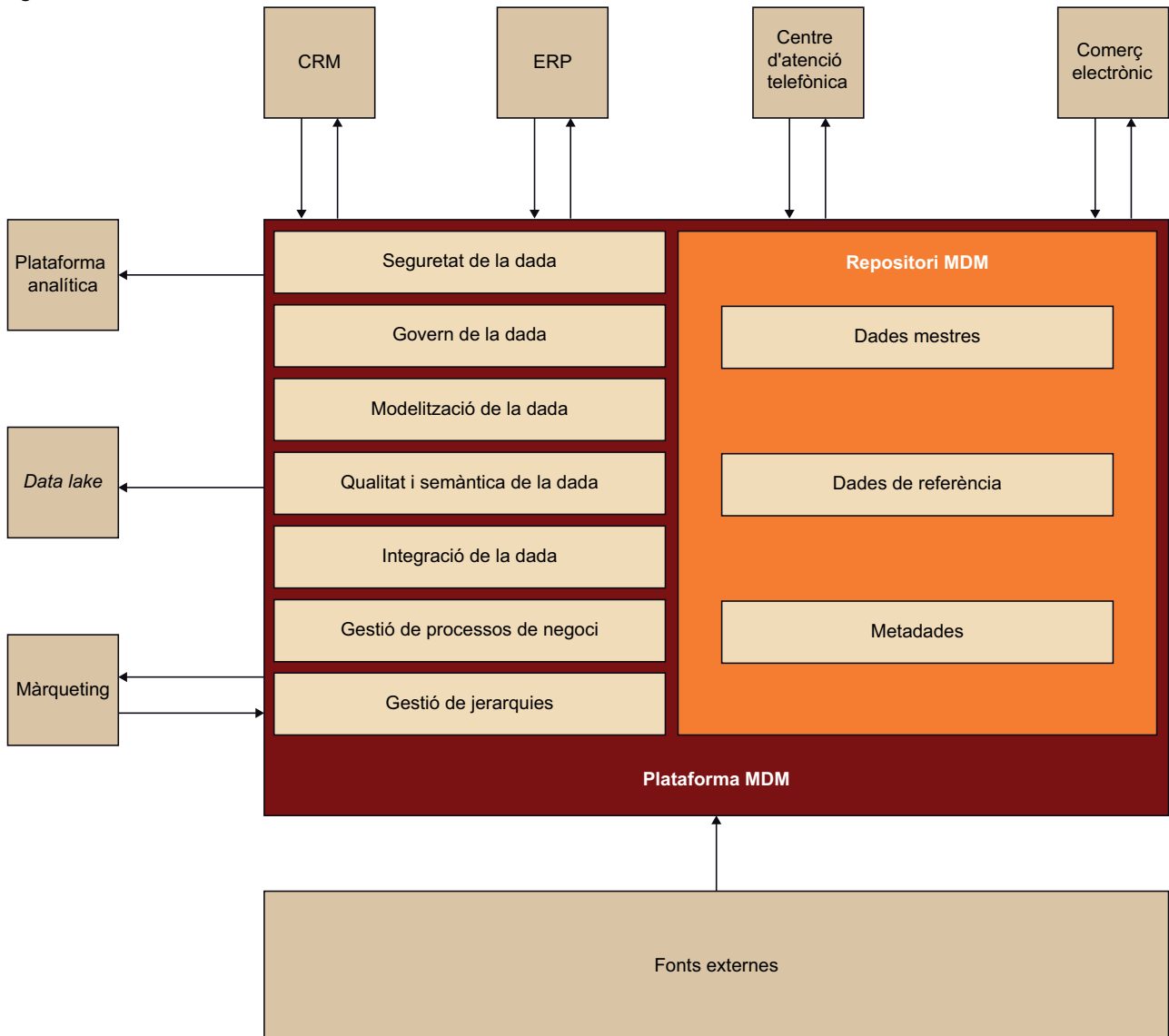
La figura 11 il·lustra la plataforma MDM que es complementa amb el repositori per a les dades mestres, de referència i metadades. La plataforma es connecta amb aquells sistemes d'informació (interns i externs) que tenen dades rellevants per al sistema MDM. Aquesta comunicació, com ja sabem, és bidireccional per poder propagar les dades des del sistema d'emmagatzematge de dades mestres (freqüentment anomenat *hub*).

LDAP

Quan parlem de LDAP, fem referència a un protocol en un àmbit d'aplicació que permet l'accés a un servei de directori ordenat i distribuït per buscar diferent informació en un entorn de xarxa.

* www.semarchy.com
 ** www.informatica.com
 *** www.talend.com
 **** www.inriver.com
 ***** www.pitneybowes.com

Figura 11. Plataforma MDM



Font: Josep Curto

4.3. Exemple de cas d'ús

Imaginem una empresa de comerç electrònic que s'ha adonat de la necessitat de millorar la informació dels seus productes per:

- Millorar el seguiment de productes, amb l'objectiu de reduir incidències vinculades amb la regulació associada al seu sector.
- Evitar la frustració de clients i empleats, fonamentada en l'existència de dades de productes erronis.

Com pot començar la seva iniciativa d'MDM?

Imaginem que l'organització disposa d'una eina d'MDM. L'organització ha de seguir aquests passos:

- **Descobrir** quines aplicacions contenen dades rellevants de producte i susceptibles de formar part de les dades mestres. Així mateix, cal avaluar el seu nivell de qualitat. En aquest punt, es fan servir els components d'integració i de qualitat de dades de la plataforma.
- **Crear** el model de dades mestres per mitjà del consens amb les diferents unitats de negoci implicades. En aquest punt, es fa servir el component de modelització de dades mestres.
- **Definir** la regles de validació (de negoci), de qualitat de dades i de convergència entre fonts. En aquest punt es fan servir diferents components, com el d'administració de les polítiques.
- **Identificar** i definir les interfícies d'accés d'usuaris a la plataforma MDM. En aquest punt, intervé el component d'integració de dades i la seguretat de la dada i el govern.
- **Definir** les polítiques de govern de dades mestres i els seus fluxos de treball. Aquí és rellevant el component d'administració i definició de polítiques.
- **Integrar** MDM amb les aplicacions de negoci necessàries. Aquí hi juga un paper fonamental el component d'integració de dades.
- **Posar** en producció i gestionar el govern de la dada mestra, fonamentat en el component de govern.

En essència, la plataforma MDM esdevé un mecanisme per articular qualsevol iniciativa. És clar que com més integració hi hagi entre els components de la plataforma, aquests passos seran més fluidos.

4.4. MDM i *big data*

Com ja sabem, *big data* és una de les iniciatives actuals més rellevants per generar valor a partir de les dades. En general, múltiples estudis, com els de Capgemini*, suggereixen que tan sols el 35% dels projectes de *big data* es poden descriure com d'èxit o de molt èxit i, d'altra banda, Gartner** indica que tan sols el 15% d'empreses són capaces de posar les seves iniciatives de *big data* en producció. En definitiva, aquests estudis confirmen que aquests projectes són complexos i que poden acabar en fracàs.

Dins de les iniciatives que s'emmarquen al *big data*, ens trobem amb aquelles vinculades amb el *data lake* (llac de dades). El propòsit d'aquesta iniciativa és presentar una visió no refinada de les dades, que permeti explorar lliurement tècniques d'anàlisi de dades i investigar-ne casos d'ús. Inicialment, aquestes iniciatives s'han fonamentat en bases de dades NoSQL del tipus clau-valor.

Big data

Big data fa referència al conjunt d'estratègies, tecnologies i sistemes per a l'emmagatzematge, el processament, l'anàlisi i la visualització de conjunts de dades complexes.

* M. Colas i altres (2014). *Cracking the Data Conundrum: How Successful Companies Make Big Data Operational*. Capgemini.

** N. Heudecker, J. Hare. (2016). *Survey Analysis: Big Data Investments Begin Tapering in 2016*. Gartner.

Data lake

Data lake fa referència a la col·lecció d'instàncies d'emmagatzematge de tots els actius de dades referents a una organització emmagatzemats en una còpia gairebé exacta, o fins i tot exacta, del format font.

Desafortunadament, les iniciatives de *data lake* no estan donant els fruits esperats*. El focus inicial ha estat a la captura i l'emmagatzematge de dades, sense tenir en compte capacitats de gestió de dades i la seva posterior explotació.

A mesura que ha resultat difícil la implantació del *data lake*, els fabricants han començat a crear plataformes integrades que van més enllà de l'emmagatzematge massiu de dades. Estem parlant de la integració de les capacitats següents:

- **Ingerir dades:** capacitat per a la ingesta de dades d'autoservei amb neteja de dades, validació i perfils automàtics.
- **Preparar dades:** capacitat per manipular i preparar dades mitjançant SQL, i transformació interactiva per mitjà d'una interfície d'usuari simple.
- **Descobrir dades:** capacitat per a explorar dades i metadades, i comprendre el seu llinatge i el seu perfilat.
- **Monitoritzar dades:** capacitat per comprendre el rendiment i funcionament del *data lake* (és a dir, qui el fa servir, com, etc.), a més del seguiment de SLA i resolució de problemes de rendiment.
- **Disseny de fluxos de dades:** capacitat per al disseny de flux de dades per a l'autoservei.

Com es pot apreciar, aquests components tenen punts en comú amb els components d'una gestió de dades mestres que hem revisat anteriorment com, per exemple, la integració, la qualitat, la validació i el govern de la dada.

Una altra aproximació que està emergint és l'ús d'una plataforma MDM que estén el *data lake* i el complementa, és a dir, la plataforma de *big data* es converteix en una altra font de dades que pot rebre la propagació de dades mestres.

En essència, MDM crea un context per *big data* proporcionant informació fiable sobre com les dades no estructurades i entrants encaixen en l'entorn empresarial. D'altra banda, el *big data* crea un nou context de dades per expandir les dades mestres, cosa que ajuda a l'organització a incrementar el valor de la dada.

* N. Heudecker; A. White. (2014). *The Data Lake Fallacy: All Water and Little Substance*. Gartner.

SQL

Quan parlem de SQL, *structured query language*, fem referència al llenguatge de consulta de bases de dades relacionals.

SLA

Quan parlem de SLA, *service level agreement*, fem referència al contracte entre un proveïdor de serveis (intern o extern) i l'usuari final que defineix el nivell de servei esperat per part del proveïdor.

Resum

En aquest mòdul didàctic hem presentat el concepte de gestió de dades mestres, que té com a objectiu crear definicions estàndard de les dades crítiques i estables d'una organització. Les dades mestres es transformen en un únic punt de referència per a múltiples sistemes, incloent-hi la intel·ligència de negoci.

Primer hem introduït el concepte, els diferents tipus de dades mestres existents i els beneficis que aporta aquesta estratègia per a una organització.

A continuació, hem revisat en què consisteix un programa de gestió de dades mestres, com encaixa dins del govern de la dada, les diferents metodologies existents per al seu desenvolupament i les millors pràctiques. Així mateix, hem revisat què significa MDM per a diferents tipus de dades, com de client o de producte.

Finalment, hem revisat les tècniques i la tecnologia que componen la gestió de dades mestres i els components que formen part d'una plataforma.

Glossari

big data *m* Conjunt d'estratègies, tecnologies i sistemes per a l'emmagatzematge, el processament, l'anàlisi i la visualització de conjunts de dades complexes.

business intelligence *f* Conjunt de metodologies, aplicacions, pràctiques i capacitats enfocades a la creació i administració d'informació i que permet prendre millors decisions als usuaris d'una organització.

CRM *m* Sistema informàtic per a l'administració i gestió de la relació amb clients.

data quality *f* Tècniques per a la identificació, el control, l'increment i el manteniment de la qualitat de dades en una organització.

data warehouse *m* Repositori de dades que proporciona una visió global, comú i integrada de les dades de l'organització, de manera independent de com seran utilitzades posteriorment pels consumidors o els usuaris, amb les propietats següents: estable, coherent, fiable i amb informació històrica.

ERP *m* Sistemes informàtics destinats a l'administració de recursos en una organització.

ETL *m* Processos que permeten l'extracció, la transformació i la càrrega de dades des de fonts d'origen fins a la destinació per al seu correcte consum.

Gestió de dades mestres *f* Conjunt de processos, govern, polítiques, estàndards i eines que defineixen i gestionen constantment les dades crítiques d'una organització per proporcionar un únic punt de referència.

golden record *f* Versió única i ben definida de totes les entitats de dades en un ecosistema organitzacional.

HRM *m* Sistema informàtic per a l'administració i la gestió de la relació amb empleats.

PLM *m* Sistema per a la gestió del cicle de vida d'un producte.

Bibliografia

Berson, A.; Dubov, L. (2010). *Master Data Management and Data Governance, 2/E*. Nova York: McGraw Hill Professional.

Brackett, M.; Earley, P. S. (2009). *The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide)*. Nova York: DAMA.

Curto, J. (2017). *Introducción al Business Intelligence (nueva edición ampliada y revisada)*. Barcelona: Editorial UOC.

Loshin, D. (2010). *Master data management*. Nova York: Morgan Kaufmann.

Mosley, M. (2009). *DAMA-DMBOK functional framework*. Nova York: DAMA.