



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

FACULTAD DE INGENIERÍA

Gestión de datos maestros y de referencia  
basada en el marco de una arquitectura  
empresarial

T E S I S

QUE PARA OBTENER EL TÍTULO DE:  
INGENIERÍA EN COMPUTACIÓN

P R E S E N T A :

FLOR REYNALDA AGUILAR HERNÁNDEZ



DIRECTOR DE TESIS:  
Dra. María del Pilar Ángeles  
2014

*“Imagina la vida como un juego en el cual tú te encuentras haciendo malabares con 5 pelotas en el aire, imagínate que esas pelotas son: el trabajo, la familia, la salud, los amigos y el espíritu y las mantienes todas en el aire... La pelota del trabajo es de hule, y si la dejas caer, regresará a ti, pero las otras 4 pelotas, familia, salud, amigos y espíritu son de cristal; si dejas caer alguna de ellas... jamás volverá a ser lo mismo.”*

*Brian Dyson*

## **Agradecimientos**

A mis compañeros de vida:

Papa, gracias por creer en mí, por transmitirme el amor al deporte. También por ser valiente y divertido.  
Mama, gracias por ser el mejor apoyo, el mejor consejo, y la mejor amiga. Gracias por enseñarme lo importante de la vida y guiarme en ella, gracias a ti soy la persona de ahora.  
Carla y Clemen, gracias por ser mis confidentes y por bridarme tanta felicidad en estos 23 años. ¡Ustedes son fantásticas!  
Tío Polo, gracias por ser un ejemplo de éxito, por nunca soltarme y apoyarme incondicionalmente en el logro de mis metas. Gracias por enseñarme a ir siempre por lo mejor.  
Gracias abuelita Chela, tí@s y prim@s por brindarme esa unión familiar que nos caracteriza.  
Vladimir, gracias por el gran apoyo y amor que me has dado estos 3 años. Te admiro no solo por tus cualidades intelectuales y profesionales, también por la bella persona que eres. ¡Te amo!

A mis mentores de la facultad:

Gracias a la profesora Pilar Ángeles por apoyarme en este proyecto tan importante para mí. Gracias por impulsarme a ser cada día más disciplinada, por los conocimientos adquiridos en la materia de Bases de Datos Distribuidas y darme la oportunidad de obtener mi primera certificación de TI.  
A mí querida profesora Laura Sandoval, más que una tutora una amiga, gracias por escucharme y darme consejos los 9 semestres de mi carrera.  
Gracias a los grandiosos profesores que tenido en la Facultad de Ingeniería, porque ellos me han hecho amar cada vez más la ingeniería: Luis López, Balmori, Carlos Crail, Erik Castañeda, Vicente Flores, Lucilla, Margarita Fournier, Sergio Noble, Osorio Compartan, Rubén Anaya, Jorge Alberto Campos, Gloria Correa, Patricia del Valle.

A mis mentores de la vida laboral:

Al mejor jefe, Eduardo Jallath, gracias por darme la oportunidad de colaborar contigo y aprender del brillante ejemplo que eres en todos los aspectos. Gracias por enseñarme a hacer las cosas bien, por explotar mi capacidad profesional y gracias por presionarme para finalizar mi tesis.  
Oswaldo Baez, gracias por tenerme paciencia e instruirme con los conocimientos y herramientas necesarias para mi crecimiento profesional.  
Pilar Trujillo, gracias por compartirme tus consejos y conocimientos de estructura y redacción.  
Muchas gracias Oswaldo, Pilar y Eduardo, por sus enseñanzas y comentarios al desarrollo de este trabajo.

También doy gracias por tener a personas que me hacen feliz día a día, personas que quiero tanto e inspiran mi vida, hablo de mis amig@s. You know who you are ;)

**GRACIAS UNAM, POR DARME UNA EDUCACIÓN EXCEPCIONAL, POR HACERME CONOCER A PERSONAS GRANDIOSAS, POR HABERME DADO LA OPORTUNIDAD DE REALIZAR DIVERSAS ACTIVIDADES QUE RUTO, POR CREAR EN MÍ LA PASIÓN DE AMAR A MI CARRERA, POR DARME LAS BASES PARA TENER UN CRECIMIENTO PERSONAL Y PROFESIONAL, GRACIAS POR DARME TANTAS EXPERIENCIAS MARAVILLOSAS.**

**TÚ, MI AMADA UNAM, ME HAS DADO TANTO QUE SIEMPRE TENDRE UNA DEUDA ETERNA CONTIGO.**

También agradezco, por hacerme este camino más divertido, a las personas que inventaron y han hecho posible:



Gestión de datos maestros y de referencia basada en el marco de una arquitectura empresarial by Flor R. Aguilar Hernández is licensed under a Creative Commons Reconocimiento-NoComercial 4.0 Internacional License .



## **Contenido**

1. INTRODUCCIÓN .....	11
1.1 Objetivo.....	13
1.2 Objetivos particulares .....	13
2. ANTECEDENTES.....	13
2.1 Definición de entidades maestras.....	14
2.2 Definición de entidades de referencia.....	16
2.3 Diferencia entre entidades maestras y de referencia .....	17
2.4 Situación actual en las organizaciones.....	18
2.5 Beneficios de la gestión de entidades maestras y de referencia .....	19
3. MARCO TEÓRICO .....	21
3.1 Marco de la arquitectura empresarial (TOGAF) .....	21
3.2 Organizaciones y metodologías de datos .....	23
3.2.1 Data Management Association .....	23
3.2.2 Master Data Management Institute.....	30
3.2.3 Enterprise Data World .....	31
3.3 Esquemas comerciales para soluciones de datos .....	33
3.3.1 Oracle.....	34
3.3.2 IBM .....	35
3.3.3 Microsoft.....	38
3.3.4 Informatica.....	40
3.4 Modelo de madurez de la gestión de datos maestros .....	41
4. ARQUITECTURA DE DATOS MAESTROS Y DE REFERENCIA .....	43
4.1 Fase 0: Preliminar .....	46
4.1.1 Evaluación de madurez.....	48
4.2 Fase A: Visión de la arquitectura .....	55
4.2.1 Estrategia .....	55
4.3 Fase B: Arquitectura de negocio .....	60
4.3.1 Principios de datos .....	61
4.3.2 Procesos .....	65
4.3.3 Funciones .....	72
4.3.4 Servicios .....	76

4.3.5 Roles .....	79
4.4 Fase C: Arquitectura de datos .....	80
4.4.1 Artefactos del diseño conceptual .....	81
4.4.2 Artefactos del diseño lógico .....	85
4.5 Fase E: Oportunidades y soluciones .....	88
4.5.1 Restricciones para la implementación .....	88
4.5.2 Estrategia de migración e implementación.....	88
4.6 Fase F: Planeación de migración.....	90
4.6.1 Plan de migración e implementación.....	91
4.7 Meta-modelo de la arquitectura .....	93
4.8 Bloques de construcción de la arquitectura.....	94
<b>5. CASO DE ESTUDIO: USUARIOS Y PROVEEDORES DE UN ÁREA FINANCIERA .....</b>	<b>96</b>
5.1 Fase 0: Preliminar .....	96
5.1.1 Evaluación de madurez para usuarios y proveedores .....	97
5.2 Fase A: Visión de la arquitectura .....	109
5.2.1 Nivel de madurez objetivo .....	110
5.2.2 Estrategia .....	111
5.3 Fase B: Arquitectura de negocio .....	112
5.4 Fase C: Arquitectura de datos .....	112
5.4.1 Conceptual .....	112
5.4.1 Lógica .....	117
5.6 Fase E: Oportunidades y soluciones .....	121
5.7 Fase F: Planeación de migración.....	123
5.7.1 Iniciativa 1: Detección y análisis de datos maestros y de referencia .....	123
5.7.2 Iniciativa 2: Descripción y modelado de datos maestros y de referencia .....	136
<b>6. CONCLUSIONES .....</b>	<b>150</b>
<b>7. ANEXOS.....</b>	<b>151</b>
7.1 Relación entre datos maestros y datos transaccionales .....	151
7.2 Diferencia entre un DWH y un repositorio maestro .....	153
7.3 Proceso general de vinculación de registros .....	153
7.4 Definición de servicios SOA.....	155
<b>8. BIBLIOGRAFÍA Y REFERENCIAS .....</b>	<b>157</b>

## **Índice de Figuras**

- Figura 2.1 Ejemplo de datos maestros y de referencia
- Figura 2.2 Situación actual de DMyR
- Figura 3.1 Método de desarrollo de la arquitectura. Traducida de TOGAF, 2011.
- Figura 3.2 Diagrama de la administración de DMyR. Traducida de DAMA, 2009.
- Figura 3.3 Arquitectura de DM. Adaptada de DAMA, 2009.
- Figura 3.4 Arquitectura de DR. Adaptada de DAMA, 2009.
- Figura 3.5 Arquitectura de un ODS central de datos. Adaptada de DAMA, 2009.
- Figura 3.6 Componentes de la administración de datos maestros. Traducido de MDM Institute, 2007.
- Figura 3.7 Administración central de datos maestros. Traducido de EDW, 2013.
- Figura 3.8 Administración de datos. Traducida de EDW, 2013.
- Figura 3.9 Servicios para la gestión de datos maestros. Adaptado de Oracle, 2011.
- Figura 3.10 Solución para la gestión de datos maestros. Tomado de IBM, 2007.
- Figura 3.11 Arquitectura de servicios para datos maestros. Adaptada de IBM, 2011.
- Figura 3.12 Arquitectura de la administración de datos maestros. Traducida de Microsoft, 2006
- Figura 3.13 MDM Madurity Model. Traducida de Loshin, 2009.
- Figura 4.1 Alcance del desarrollo de la arquitectura de DMyR. Adaptado de TOGAF, 2011
- Figura 4.2 Alcance del desarrollo de la Arquitectura de DMyR
- Figura 4.3 Cuestionario de Arquitectura. Adaptado de Loshin, 2009
- Figura 4.4 Cuestionario de Gobierno. Adaptado de Loshin, 2009.
- Figura 4.5 Cuestionario de Administración. Adaptado de Loshin, 2009
- Figura 4.6 Cuestionario de Identificación. Adaptado de Loshin, 2009
- Figura 4.7 Cuestionario de Integración. Adaptado de Loshin, 2009.
- Figura 4.8 Cuestionario de Admón. de procesos de negocio. Adaptado de Loshin, 2009
- Figura 4.9 Niveles de abstracción para la obtención de Iniciativas
- Figura 4.10 Iniciativas que soportan las capacidades de la Arquitectura
- Figura 4.11 Iniciativas que soportan las capacidades del Gobierno
- Figura 4.12 Iniciativas que soportan las capacidades de la administración
- Figura 4.13 Iniciativas que soportan las capacidades de la Identificación
- Figura 4.14 Iniciativas que soportan las capacidades de la Integración
- Figura 4.15 Iniciativas que soportan las capacidades de Admón. de procesos de negocio
- Figura 4.16 Estrategia para desarrollar las capacidades
- Figura 4.17 Proceso de creación y mantenimiento de las entidades
- Figura 4.18 Proceso de calidad de entidades
- Figura 4.19 Proceso de actualización de registros
- Figura 4.20 Proceso de intercambio de registros
- Figura 4.21 Funciones de la gestión de DMyR
- Figura 4.22 Funciones para el proceso: Creación y mantenimiento de entidades

- Figura 4.23 Funciones para el proceso: Calidad de entidades  
Figura 4.24 Funciones para el proceso: Actualización de los registros  
Figura 4.25 Funciones para el proceso: Intercambio de registros  
Figura 4.26 Servicios para el proceso: Creación y mantenimiento de entidades  
Figura 4.27 Servicios para el proceso: Calidad de entidades  
Figura 4.28 Servicios para el proceso: actualización de registros  
Figura 4.29 Servicios para el proceso: Intercambio de registros  
Figura 4.30 Niveles de abstracción de la arquitectura de datos  
Figura 4.31 Ejemplo de un diagrama de clases de una Universidad. Tomado de Wikipedia.  
Figura 4.32 Diagrama del ciclo de vida del DMyR. Adaptada de Microsoft, 2006.  
Figura 4.33 Modelo de datos  
Figura 4.34 Diagrama de diseminación de datos  
Figura 4.35 Estructura de desglose del trabajo para desarrollar la arquitectura de DMyR  
Figura 4.36 Meta-modelo de la arquitectura. Traducida y adaptada de TOGAF, 2011.  
Figura 4.37 Bloques de construcción de la arquitectura  
Figura 5.1 Alcance del ADM para el caso de estudio  
Figura 5.2 Evaluación del componente de Arquitectura  
Figura 5.3 Evaluación del componente de gobierno  
Figura 5.4 Evaluación del componente de administración  
Figura 5.5 Evaluación del componente de Identificación  
Figura 5.6 Evaluación del componente de integración  
Figura 5.7 Evaluación de administración de procesos de negocio  
Figura 5.8 Estrategia acotada del caso de estudio  
Figura 5.9 Diagrama de clases para Usuarios  
Figura 5.10 Diagrama de clases para Proveedores  
Figura 5.11 Modelo de datos de usuarios y proveedores  
Figura 5.12 Jerarquía de usuarios finales  
Figura 5.13 Jerarquía para usuarios operativos  
Figura 5.14 Niveles de representación de usuarios  
Figura 5.15 Relación proveedor-contacto  
Figura 5.16 Niveles de representación de proveedores  
Figura 5.17 Flujo de control para las validaciones de usuarios y proveedores  
Figura 5.18 Flujo de datos para la depuración de usuarios y proveedores  
Figura 5.19 Modelo de información  
Figura 5.20 Diagramas de clases para usuarios  
Figura 5.21 Diagrama de clases para Proveedores  
Figura 5.22 Modelo de datos de usuarios  
Figura 5.23 Modelo de datos de Proveedores  
Figura 5.24 Modelo maestro de usuarios y proveedores  
Figura 7.1 Ejemplo de datos maestros y transaccionales

Figura 7.2 Vinculación de datos

Figura 7.3 Diseño de servicios para DMyR

## Índice de Tablas

Tabla 2.1 Definiciones de datos maestros

Tabla 2.2 Ejemplos de datos maestros

Tabla 2.3 Catálogo de tipo de operación

Tabla 2.4 Catálogo de monedas

Tabla 3.1 Actividades, entregables y roles. Traducida de DAMA, 2009.

Tabla 4.1 Iniciativa 1 “Detección y análisis de DMyR”

Tabla 4.2 Iniciativa 2 “Descripción y modelado de datos maestros”

Tabla 4.3 Iniciativa 3 “Distribución de los DMyR a través de servicios SOA”

Tabla 4.4 Iniciativa 4 “Colaboración de los recursos humanos”

Tabla 4.5 Servicios para el proceso: Creación y mantenimiento de entidades

Tabla 4.6 Servicios para el proceso: Calidad de entidades

Tabla 4.7 Servicios para el proceso: Actualización de registros

Tabla 4.8 Servicios para el proceso: Intercambio de registros

Tabla 4.9 Catálogo de referencias cruzadas

Tabla 4.10 Catálogo de reglas de negocio

Tabla 4.11 Catálogo de reglas de correspondencia

Tabla 4.12 Matriz Entidad/Función

Tabla 4.13 Catálogo de metadatos

Tabla 4.14 Catálogo de relaciones entre objetos

Tabla 4.15 Matriz Objeto/ Aplicación

Tabla 4.16 Matriz de seguridad del dato

Tabla 4.17 Proyectos de la Iniciativa “Detección y análisis de DMyR”

Tabla 4.18 Proyectos de la Iniciativa “Descripción y modelado de datos”

Tabla 4.19 Proyectos de la Iniciativa 3 “Distribución de los DMyR”

Tabla 4.20 Proyectos de la Iniciativa 4 “Colaboración de los recursos humanos”

Tabla 4.21 Plan de implementación para la Iniciativa 1 “Detección y análisis de DMyR”

Tabla 4.22 Plan de implementación para la Iniciativa 2 “Descripción y modelado de DM”

Tabla 4.23 Plan de implementación para la Iniciativa 3 “Distribución de DMyR”

Tabla 4.24 Plan de implementación para la Iniciativa 4 “Colaboración de los recursos humanos”

Tabla 5.1 Resumen de la evaluación

Tabla 5.2 Componentes y capacidades del nivel de madurez cinco “Desempeño estratégico”

Tabla 5.3 Descripción de la Iniciativa 1 “Detección y análisis de datos maestros y de referencia”

Tabla 5.4 Descripción de la Iniciativa 2 “Descripción y modelado de datos maestros”

Tabla 5.5 Catálogos de reglas de negocio para Usuarios

Tabla 5.6 Catálogos de reglas de negocio para Proveedores

- Tabla 5.7 Catálogos de reglas de correspondencia  
Tabla 5.8 Matriz Entidad/Función  
Tabla 5.9 Catálogo de metadatos  
Tabla 5.10 Matriz de las relaciones entre objetos de datos  
Tabla 5.11 Nivel inicial y objetivo de la arquitectura de datos  
Tabla 5.12 Proyectos de la Iniciativa 1  
Tabla 5.13 Proyectos de la Iniciativa 2  
Tabla 5.14 Plan de trabajo de la iniciativa 1  
Tabla 5.15 Artefacto para la recolección de usuarios y proveedores  
Tabla 5.16 Regla de correspondencia de usuarios y proveedores  
Tabla 5.17 Resultados de la depuración de datos  
Tabla 5.18 Selección de registros  
Tabla 5.19 Plan de trabajo para la iniciativa 2  
Tabla 5.20 Matriz Entidad/Función  
Tabla 5.21 Reglas de negocio para usuarios  
Tabla 5.22 Reglas de negocio para proveedores  
Tabla 5.23 Artefacto de Metadatos para Usuarios y Proveedores  
Tabla 5.24 Matriz de relaciones entre objetos de datos

## 1. INTRODUCCIÓN

Actualmente la infraestructura de datos de las organizaciones necesita alinearse a las mejores prácticas de Tecnologías de la Información (TI) para tener un mayor aprovechamiento de los recursos humanos y tecnológicos. Es común que diversas áreas de cada organización utilicen los mismos datos pero con diferentes propósitos. A estos datos compartidos se les denominan datos maestros y de referencia.

Los datos maestros son las representaciones de los elementos primordiales para el funcionamiento de las organizaciones como clientes, productos y proveedores. Los datos de referencia son listas de claves que contienen sus significados de negocio, como por ejemplo los catálogos de tipos de productos, tipos de operaciones, entre otros.

La importancia de gestionar estas dos categorías de datos radica en tenerlos unificados para compartirlos y sincronizarlos con todos los procesos del negocio. Para tener una visión unificada de estos datos es necesario centralizarlos en un repositorio único. La centralización de los datos maestros y de referencia disminuye el riesgo de operar información sin calidad debido a que, estos datos son utilizados por los procesos para validar los datos de entrada y generar productos de información basados en información confiable.

Dentro de las organizaciones, cada área tiene diferentes expectativas de la calidad de los datos maestros y de referencia, y como no se tiene un control de ellos, surgen múltiples versiones creando “islas de datos” incompatibles en formato, estructura y contenido.

Aunado a esto, existen otros vicios como la escasa cultura para compartir datos entre diferentes áreas o incluso entre compañeros de un mismo equipo de trabajo, surgiendo así la apropiación de la información.

Aun cuando las estructuras actuales de datos arrojen los productos de información requeridos, su contenido puede ser erróneo. Trabajar con datos duplicados y obsoletos impacta negativamente la calidad de los datos de salida que proveen las organizaciones y por lo tanto, la toma de decisiones basadas en esta información.

Dado lo anterior esta tesis propone el diseño de una arquitectura que facilite las prácticas de la gestión de datos maestros y de referencia, con el fin de crear una visión unificada de estos datos para compartirlos, sincronizarlos y mantenerlos actualizados en beneficio de los procesos del negocio.

Para el diseño de esta arquitectura se revisaron diferentes fuentes literarias centradas en el estudio del tema como el Data Management Association (DAMA, 2009), Master Data Management Institute (MDM Institute, 2013), Enterprise Data World (EDW, 2013), así como empresas comerciales de TI.

Así mismo para estructurar y dirigir la información recopilada de las diferentes referencias literarias, se eligió el marco de referencia The Open Group Architecture Framework (TOGAF, 2011) por las siguientes razones:

- Proporciona fases y directrices para la definición de los elementos primordiales del dominio de negocio, datos, aplicaciones e infraestructura tecnológica.
- Facilita la integración de otros marcos metodológicos.
- Define las relaciones entre los principales activos de las organizaciones como procesos, personas e información. La representación de estas relaciones proporciona visibilidad del uso de estos recursos.
- Proporciona fases para la planeación, implementación y gobierno de la arquitectura definida en las fases previamente mencionadas.

Esta tesis se ha estructurado de la siguiente manera: en el *Capítulo 2* como marco teórico básico se describen las definiciones necesarias que sustentan al presente trabajo, así como la situación actual del tratamiento de los datos maestros y de referencia, y los beneficios de su gestión; en el *Capítulo 3* se presenta el marco conceptual de investigación, el cual contiene la teoría recopilada de metodologías y artículos sobre dicho tema; posteriormente en el *Capítulo 4* se presenta la arquitectura-propuesta a través de las siguientes secciones:

- Las principales necesidades de las empresas en materia de estos datos (Fase: Preliminar).
- Alcance, misión y visión (Fase: Visión de la arquitectura).
- Principios, procesos y roles primordiales para su gestión (Fase: Arquitectura de negocio).
- Artefactos a nivel conceptual y lógico para la documentación, estandarización e implementación de los datos (Fase: Arquitectura de datos).
- Una estrategia integral que contempla cuatro iniciativas para la recolección, análisis, depuración, modelado de datos maestros y de referencia, así como la puesta en marcha de las anteriores arquitecturas a través de los recursos humanos y tecnológicos (Fase: Oportunidades y soluciones).
- Actividades de estas iniciativas en un plan de trabajo (Fase: Planeación de migración).

Posteriormente, para aplicar la arquitectura propuesta en el *Capítulo 5* se utilizaron como caso de estudio los datos de usuarios y proveedores de una Institución Financiera. En el *Capítulo 6* se exponen las conclusiones obtenidas del desarrollo y aplicación del modelo; y finalmente se presentan los anexos en el *Capítulo 7*.

Advertencia. El desarrollo de esta tesis se enfocó en resolver una problemática para una Institución Financiera, sin embargo es posible adaptar la arquitectura a cualquier otro organismo debido a que se cuidó que la definición de los elementos arquitectónicos del modelo fuera estandarizada y apegada a las mejores prácticas. El presente trabajo está

dirigido a profesionales de negocio y de TI, así como a estudiantes e investigadores que deseen adentrarse en el tema de la gestión de datos maestros y de referencia.

## **1.1 Objetivo**

Diseñar los elementos arquitectónicos de un modelo de gestión de datos maestros y de referencia, utilizando el marco de Arquitectura Empresarial The Open Group Architecture Framework (TOGAF). Esto con el fin de propiciar y facilitar las prácticas de gestión de este tipo de datos para obtener una versión unificada de ellos, sincronizarlos, compartirlos y mantenerlos actualizados en beneficio de los procesos de negocio.

## **1.2 Objetivos particulares**

- Recopilar y analizar las metodologías y soluciones relacionadas con el tema de datos maestros y de referencia.
- Diseñar los elementos de la arquitectura de negocio para la gestión de este tipo de datos.
- Diseñar los elementos de la arquitectura de datos para la gestión de este tipo de datos.
- Definir la estrategia para implantar los elementos mencionados en el punto anterior.
- Aplicar la arquitectura propuesta a un caso de estudio para mejorar los diseños y comprobar su uso.

## **2. ANTECEDENTES**

En este capítulo se describen los precedentes que justifican el desarrollo de este trabajo. Se inicia por describir la definición para los datos maestros y de referencia con base en diversas

fuentes de referencia, así como sus diferencias. Posteriormente se describe la situación actual en las organizaciones respecto al tratamiento de estos datos, y la importancia de su gestión.

## 2.1 Definición de entidades maestras

Se comienza por comprender el concepto central del desarrollo de este documento: dato maestro. Para ello en la Tabla 2.1 se enlistan sus definiciones provenientes de organizaciones, metodologías y empresas de tecnologías de la información (TI).

Fuente	Definición
Organización internacional	Son datos sobre entidades de negocio que proporcionan un contexto para las transacciones de negocio (DAMA, 2009).
Esquemas comerciales	Información crítica del negocio que da soporte a las operaciones transaccionales y analíticas de la empresa (Oracle, 2011).
	Datos de alto valor, información core que fluye a través de los procesos clave de una organización. (IBM, 2011).
	Datos que son críticos, compartidos y utilizados por varias aplicaciones, y son a menudo los principales activos de la empresa (Microsoft, 2006).
	Registro único que sirve de referencia para toda la empresa, el nombre de un cliente, el código de un producto (Informatica, 2006).

Tabla 2.1 Definiciones de datos maestros

Retomando las partes esenciales de las definiciones, se propone la siguiente adaptación para definir a los datos maestros:

Son conjuntos de datos clave que representan los elementos o actores más importantes de una organización. Estos datos sirven para:

- Proporcionar un contexto confiable en los procesos que verifican la calidad de las operaciones transaccionales.
- También se usan como parte de los contenidos de los informes analíticos para la toma de decisiones.

Esta definición quiere decir que los datos maestros (DM) deben contener registros únicos y vigentes para proporcionar un contexto confiable a los procesos de la organización y por lo tanto realizar la toma de decisiones con información correcta.

Existe una clasificación para el dominio de datos bajo los criterios de: formato, propósito, alcance de integración, precisión, integridad, consistencia, oportunidad, relevancia y confianza (IBM, 2010).

- Para el primer criterio, *formato*, los DM cumplen con las características de datos estructurados, que son los que se utilizan, por ejemplo, en un modelo de datos relacional.
- Respecto al segundo criterio, *propósito*, los DM tienen relación con los datos operativos; esta relación se da porque los datos transaccionales dependen de los datos maestros.
- El *alcance de integración* de los DM es a nivel organización, ya que por ser datos primordiales del negocio se utilizan en todas las áreas u oficinas.
- La *precisión* debe ser muy alta, si los datos maestros presentan errores la calidad de la información se ve afectada.
- La *integridad y consistencia* de estos datos deben responder con valores verdaderos a las necesidades que se tengan, es decir, la información debe ser la misma en cualquier lugar y momento solicitado.
- La *oportunidad* de los DM significa que al realizar un cambio en ellos, el retraso de la actualización no debe afectar su calidad para que sigan siendo oportunos.
- La *relevancia* se basa en satisfacer las necesidades del usuario de acuerdo a un determinado contexto.
- Finalmente el último criterio, la *confianza*, se genera cuando se sabe que los DM cumplen con altos niveles en los criterios antes mencionados.

Con base en el DAMA los datos maestros se agrupan en: personas, cosas, lugares y conceptos. Las personas pueden ser clientes, empleados y vendedores; puede haber productos, almacenes u otros activos; los lugares hacen referencia a oficinas y divisiones geográficas; y los conceptos se refieren a contratos, garantías o licencias.

Todos estos ejemplos son datos maestros y algunas de estas agrupaciones se pueden dividir aún más. Por ejemplo, una empresa puede tener clientes nuevos, usuales, y ejecutivos o un producto puede ser segmentado aún más por el sector y la industria. A estas clasificaciones se les denomina jerarquías.

En la Tabla 2.2 se muestran ejemplos de datos maestros de diferentes ámbitos o sectores.

Ámbitos	Ejemplos
Comercial	<ul style="list-style-type: none"><li>▪ Productos</li><li>▪ Clientes</li><li>▪ Proveedores</li><li>▪ Socios</li><li>▪ Empleados</li></ul>

	<ul style="list-style-type: none"> <li>▪ Sucursales</li> </ul>
Legal	<ul style="list-style-type: none"> <li>▪ Sospechosos</li> <li>▪ Testigos</li> <li>▪ Victimas</li> <li>▪ Juez</li> </ul>
Organizaciones sin fines de lucro	<ul style="list-style-type: none"> <li>▪ Miembros</li> <li>▪ Donantes</li> </ul>
Asistencia sanitaria	<ul style="list-style-type: none"> <li>▪ Pacientes</li> <li>▪ Proveedores</li> <li>▪ Hospitales</li> </ul>
Educación	<ul style="list-style-type: none"> <li>▪ Estudiantes</li> <li>▪ Escuelas</li> <li>▪ Maestros</li> </ul>
Financiero	<ul style="list-style-type: none"> <li>▪ Centros de costos</li> <li>▪ Centros de beneficio</li> <li>▪ Cuentas de libro mayor</li> <li>▪ Presupuestos</li> <li>▪ Proyecciones</li> </ul>

Tabla 2.2 Ejemplos de datos maestros

Los valores de los DM también son llamados datos empresariales o datos de autoridad porque son los datos más exactos y disponibles del negocio utilizados para establecer el contexto de las operaciones.

## 2.2 Definición de entidades de referencia

La definición de las entidades de referencia adoptada en este trabajo es la siguiente:

Son las entidades que representan listas de valores válidos para un dominio del negocio. Su definición puede realizarse al interior o de forma externa cuando se emplean fuentes de proveedores ajenos a las organizaciones (DAMA, 2009).

Con base en el DAMA estos valores forman un conjunto acotado, a este conjunto se le llama dominio de valores. Las reglas de negocio son las que determinan el dominio de valores válidos para conformar a las entidades de referencia. La importancia de los datos de referencia (DR) radica en que sin un significado de los códigos del negocio la relación con otras entidades sería difícil de comprender.

Ejemplos de estos datos pueden ser:

- Códigos únicos
- Vocabularios o términos del negocio

Como se mencionó previamente los valores de los datos de referencia se pueden definir dentro o fuera de la organización. En la Tabla 2.3 se muestra como ejemplo de definición interna el catálogo de tipo de operación.

código	nombre de la operación
C	Compra de divisas
V	Venta de divisas

Tabla 2.3 Catálogo de tipo de operación

Los estándares de la industria son ejemplos de valores de datos que se pueden definir externamente. En la Tabla 2.4 se muestra un ejemplo de una entidad de referencia externa tomada de la lista de códigos de monedas (ISO 4217,2008).

código	Nombre de la moneda
MXN	Peso mexicano
JPY	Yen japonés
USD	Dólar estadounidense
UYU	Peso uruguayo

Tabla 2.4 Catálogo de monedas

### 2.3 Diferencia entre de referencia

### entidades maestras y

El dominio de los datos no es homogéneo, consiste en conjuntos de categorías diferentes, y cada una de ellas tiene sus propias características. Es necesario tomar en cuenta las necesidades de cada categoría de datos para:

- Establecer el correcto entendimiento y comunicación dentro de las organizaciones.
- Desarrollar una solución clara.
- Realizar la adecuada integración y almacenamiento de los datos.
- Explotar y usar correctamente los datos.

Existe una categoría que tiene dependencias con los datos maestros, esta categoría corresponde a los datos de referencia. Comúnmente se comete el error de considerar a estas dos clasificaciones iguales porque al ser primordiales en una arquitectura de datos y no reconocer sus diferencias coloca en desventaja a cualquier organización.

Con base en el DAMA, a continuación se enlistan sus principales diferencias:

1. Los volúmenes de datos de referencia son menores que los volúmenes de datos maestros.
2. El cambio de los datos de referencia es más lento que el de los datos maestros.
3. Los datos de referencia a menudo consisten en una entidad que contiene un código, una descripción y algunos campos más. Sin embargo, la entidad de datos maestros está compuesta por más de una entidad.

4. Los datos maestros son primordiales en el negocio, por lo tanto estos deberán tener más restricciones para compartirse.
5. El conjunto de valores que pueden tomar los datos maestros no se limitan en rangos predefinidos en su dominio, los datos de referencia sí.
6. La gestión de datos de referencia consiste en dar mantenimiento a una lista de valores válidos y sus significados. En la gestión de datos maestros se requiere identificar, analizar y aplicar métodos para desarrollar registros únicos.

En la Figura 2.1 se presenta un ejemplo de entidades maestras y de referencia, donde la entidad de Usuario está compuesta de datos maestros y las entidades Tipo de Usuario y Estado Civil están compuestas de datos de referencia. Los campos “Tipo de usuario” y “Estado civil” de la entidad maestra (Usuario), contienen claves que hacen referencia a las otras entidades.

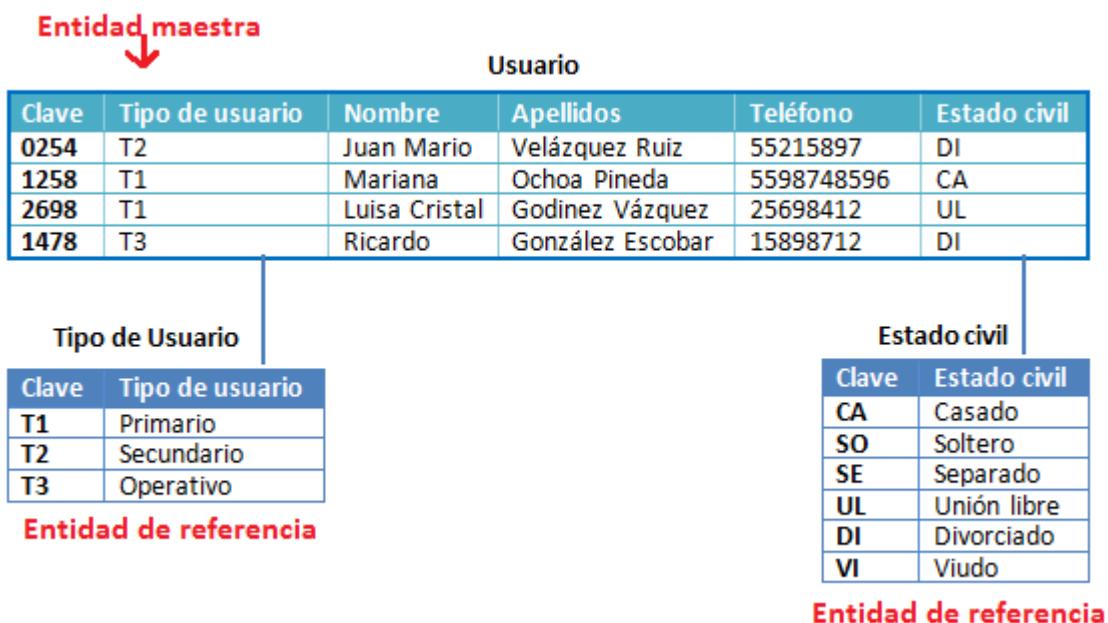


Figura 2.1 Ejemplo de datos maestros y de referencia

## 2.4 Situación actual en las organizaciones

Los requerimientos específicos del uso de datos de diferentes áreas de una organización, han conducido a crear aplicaciones de propósito específico. La situación actual respecto al almacenamiento de los datos maestros y de referencia se representa en uno o más de los siguientes casos:

- Tener valores similares de datos y compatibles en formatos.
- Tener valores similares de datos pero incompatibles en formatos.
- Tener valores diferentes de datos e incompatibles en formatos.

Los datos maestros están en constante cambio, por ejemplo: los clientes de una empresa cambian de dirección, de trabajo, hacen nuevas compras, se divorcian, cambian de cuenta, entre otras actividades. Cuando se tiene información que el cliente ha cambiado, esta es obsoleta y no sirve.

La mayoría de las organizaciones no hacen uso de una infraestructura que administre sus datos, prefieren invertir en ganancias más tangibles a corto plazo. Si la información no se actualiza en todos los sistemas que comparten estos datos, la empresa puede perder a sus clientes al contar con información verdadera.

La Figura 2.1 pretende mostrar la situación actual del tratamiento de estos datos, donde cada área de la empresa X maneja sus propios datos maestros, por ejemplo; la oficina de “Moneda Extranjera y Derivados” tiene una versión diferente del mismo proveedor que también es de la oficina “Mercado de Dinero, Captación y Comisiones”.

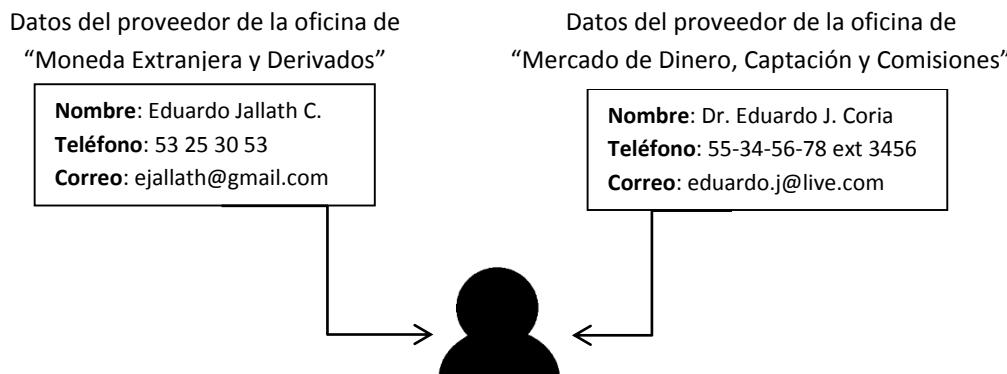


Figura 2.2 Situación actual de DMyR

¿Las fuentes de información corresponden a la misma persona?, si es así ¿Cuál es el dato correcto?, estos cuestionamientos se podrán resolver con el modelo propuesto en el *Capítulo 4*.

## 2.5 Beneficios de la gestión de entidades maestras y de referencia

Es necesario llevar un tratamiento adecuado de los datos maestros y de referencia ya que permiten verificar que la información que entra y sale a los procesos de negocio sea de calidad suficiente para hacer un análisis y tomar decisiones.

La importancia de definir una gestión para las entidades maestras y de referencia radica en que estas entidades contienen los principales valores de datos que fluyen en los procesos clave del negocio. Para tener una dirección ágil y efectiva de los datos maestros y de referencia deben compartirse, distribuirse y sincronizarse con todos los procesos del negocio.

Además, se necesitan procesos para establecer y mantener la consistencia de los datos maestros y de referencia. Los beneficios que se obtendrán al gestionar adecuadamente estos datos son:

- Crear una visión unificada de las múltiples versiones que se tienen de los datos maestros y de referencia, es decir, obtener una perspectiva integral.
- Controlar, aprovechar y mejorar el valor de los datos maestros y de referencia, por medio de:
  - Supervisar el uso de los catálogos por medio de una interfaz.
  - Establecer un método para distribuir las versiones consolidadas de los catálogos.
  - Lograr que todos los integrantes de la dirección que hagan uso de esos catálogos se comuniquen y compartan necesidades o cambios.
- Disminuir el riesgo de proveer información de baja calidad, a través de la supervisión de indicadores de calidad.
- Reducir costos operativos a través la agilidad de los procesos de negocio.
- Cumplir con estándares de la industria y regulaciones gubernamentales.

### **3. MARCO TEÓRICO**

Para la definición del modelo de gestión de datos maestros y de referencia se revisaron diferentes fuentes literarias como: arquitecturas de referencia, artículos realizados por empresas comerciales y publicaciones de libros. Estas fuentes se describen en este capítulo con el fin de dar a conocer los avances que se han realizado en relación a este tema para sustentar en el siguiente capítulo la propuesta de la arquitectura-solución.

Cabe aclarar que el tamaño del contenido de los diferentes marcos dependió de la factibilidad para obtener información referente al tema y su relevancia respecto al objetivo de este trabajo.

#### **3.1 Marco de la arquitectura empresarial (TOGAF)**

Por sus siglas en inglés significa The Open Group Architecture Framework, es un marco de referencia que proporciona un enfoque para el diseño, planificación, implementación y gobierno de una arquitectura empresarial.

Este marco contiene un modelo iterativo denominado Método de Desarrollo de la Arquitectura (ADM). Este método está compuesto de diferentes etapas a realizar de forma cíclica. De tal forma que en cada ciclo de ejecución se incremente la madurez de la organización y el valor que aporta al negocio.

El punto de partida del ciclo ADM es la Fase Preliminar, seguido de la Fase A (Visión de Arquitectura) hasta la Fase H (Gestión de Cambios de Arquitectura). Para tener una idea más clara de que es lo que implica cada una de estas fases, a continuación se describe cada una de ellas brevemente.

- Fase Preliminar. Dado que TOGAF es un marco con el objetivo de adaptarse a cualquier organización, en esta etapa deber ser adaptado a los requisitos específicos de este trabajo.

- Fase A (Visión de Arquitectura). En esta fase se establece el alcance, se deben identificar las partes interesadas y se desarrolla la visión general de los cambios que se llevarán a cabo en la organización.
- Fase B, C Y D (Arquitectura de Negocio, Sistemas de Información y de Tecnología). En estas tres fases, se desarrolla la arquitectura base y objetivo para cada dominio de arquitectura: negocio, datos, aplicaciones y tecnología para posteriormente realizar el plan de trabajo de la arquitectura.
- Fase E (Oportunidades y Soluciones). Se define la planificación inicial para la puesta en marcha de la arquitectura objetivo, identificándose posibles arquitecturas de transición, es decir, arquitecturas intermedias hacia la arquitectura objetivo.
- Fase F (Planificación de Migración). En esta fase, los proyectos de migración son priorizados y el plan de implementación y migración debe ser finalizado.
- Fase G (Gobernanza de la Implementación). En esta fase, se confirma y supervisa el cumplimiento de la post-implementación para validar cualquier proyecto respecto a la arquitectura definida.
- Fase H (Gestión de Cambios de Arquitectura). En esta fase deben estar establecidos los procedimientos necesarios para poder gestionar el cambio.
- Gestión de Requerimientos. Identifica, da seguimiento y documenta los requerimientos.

En la Figura 3.1 se muestra el ADM y sus fases iterativas.

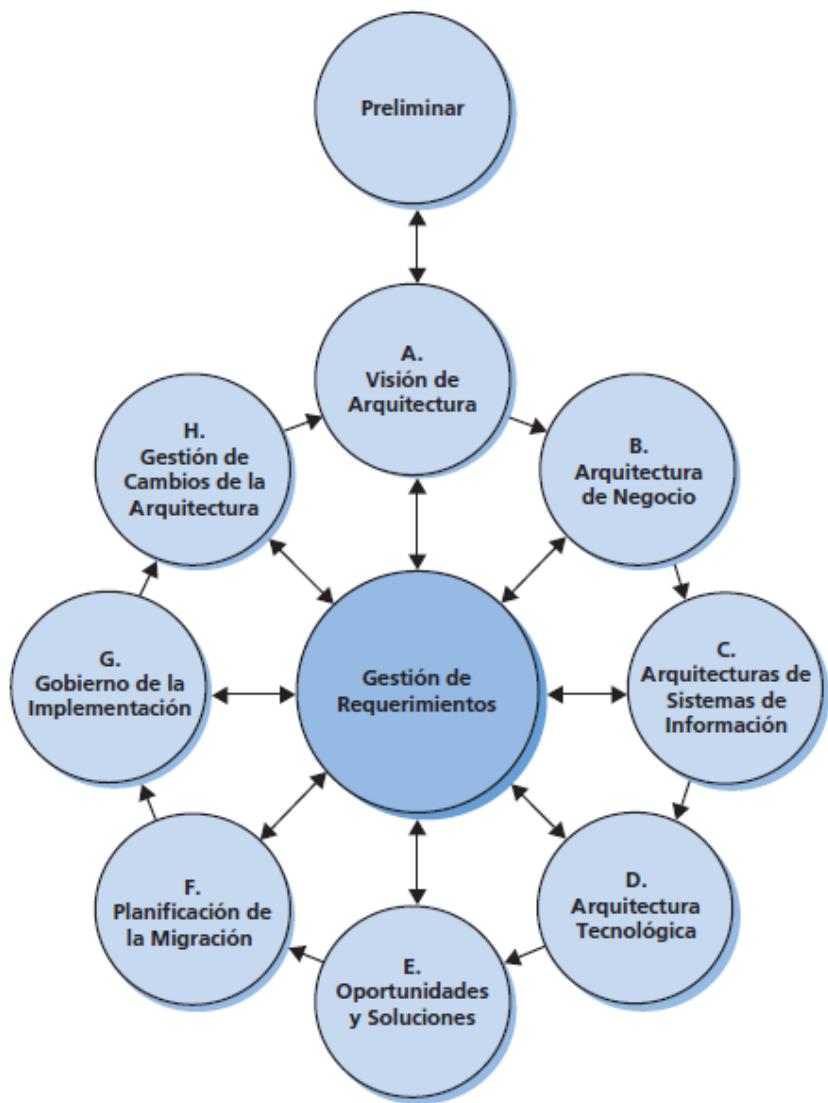


Figura 3.1 Método de desarrollo de la arquitectura. Traducida de TOGAF, 2011.

## 3.2 Organizaciones y metodologías de datos

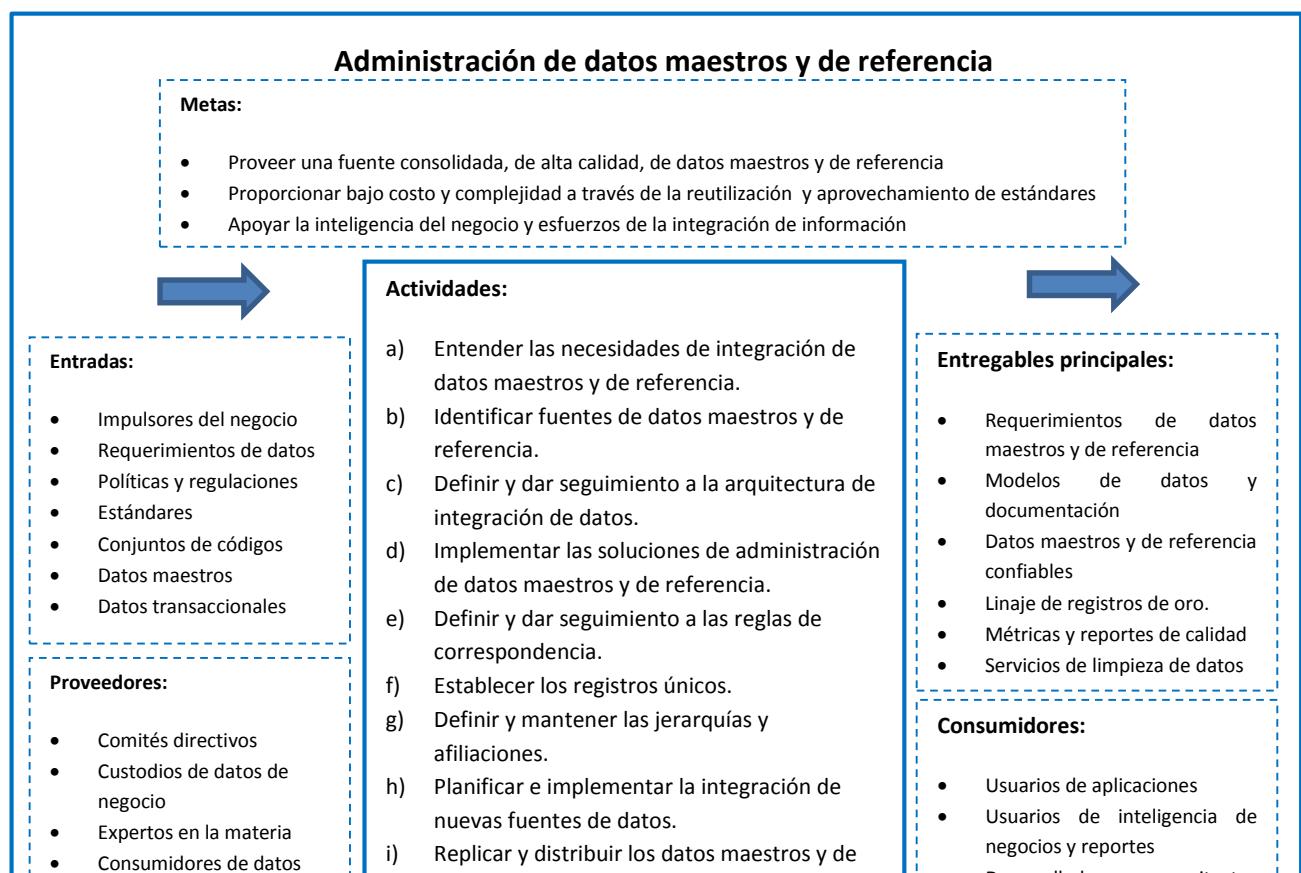
### 3.2.1 Data Management Association

Por sus siglas en inglés significa Data Management Association, es una organización internacional dedicada a las mejores prácticas de gestión de datos. Esta organización plantea sus iniciativas y conocimientos en la guía DAMA-DMBOK.

La gestión de datos maestros es el control de los valores más precisos, oportunos y relevantes de las entidades esenciales de la organización. Estos valores permiten el uso coherente, compartido y contextual a través de los sistemas (DAMA, 2009).

La gestión de datos de referencia es el control de los valores definidos como vocabularios, códigos, términos normalizados y otros identificadores únicos para clasificar y categorizar datos (DAMA, 2009).

El DAMA considera que los programas de gestión de datos maestros y de referencia se basan en la planificación, implementación y control de actividades. En la Figura 3.2 se muestra el diagrama que contiene metas, entradas, entregables, proveedores, consumidores, participantes, métricas, herramientas y actividades que se requieren para lograr la gestión de datos maestros y de referencia.



**Figura 3.2 Diagrama de la administración de DMyR. Traducida de DAMA, 2009.**

Cada una de las diez actividades centrales para la administración de datos maestros y de referencia recibe una de las siguientes clasificaciones: planeación (P), desarrollo (D), control (C) y operacional (O).

En la Tabla 3.1 se muestra a detalle los entregables generados por cada actividad, así como los roles que intervienen agrupándose en: responsables, de aprobación y de contribución.

Actividades	Entregables	Roles responsables	Roles de aprobación	Roles de contribución
1. Entender las necesidades de la integración de datos de referencia (P)	Requerimientos de datos maestros y de referencia	Analista de negocios	Consejo de gobierno de datos e Interesados	Administradores de datos de negocios y expertos en la materia
2. Identificar las fuentes y colaboradores de datos de referencia (P)	Descripción y evaluación de: fuentes y contribuidores.	Arquitectos de datos y administradores de datos	Consejo de gobierno de datos	Analista de datos , expertos en la materia
3. Definir y mantener la integración de la arquitectura de datos (P)	Integración de arquitectura de datos maestros y de referencia, y ruta a seguir (roadmap).	Arquitectos de datos	Consejo de gobierno de datos	Arquitectos de aplicaciones, administradores de datos
	Especificaciones del diseño de servicios de integración de datos.	Arquitectos de datos y arquitectos de aplicaciones	Administración TI	Otros profesionales de TI, interesados

4. Implementar la solución de administración de datos maestros y datos de referencia (D)	Bases de datos y aplicaciones de datos maestros y de referencia, bases de datos y aplicaciones de la administración de datos maestros.	Arquitectos de aplicaciones , arquitectos de datos	Consejo de gobierno de datos	Otros profesionales de TI
	Servicios de calidad de datos.	Arquitectos de aplicaciones, arquitectos de datos	Consejo de gobierno de datos	Analistas de datos, otros profesionales de TI
	Replicación de datos y accesos de servicios para aplicaciones.	Arquitectos de datos, arquitectos de aplicaciones,	Consejo de gobierno de datos	Analistas de datos, otros profesionales de TI.
	Replicación de datos de servicios para el Datawarehouse .	desarrolladores de integración		
4. Definir y mantener las reglas de coincidencia	Reglas de coincidencias de registros (Especificaciones funcionales).	Analistas de negocio, arquitectos de datos, administrados de datos de negocio	Consejo de gobierno de datos	Arquitectos de aplicaciones, expertos en la materia
5. Establecer los registros de oro (C)	Datos maestros y de referencia exactos.	Administradores de datos	Interesados	Analistas de datos, arquitectos de datos, expertos en la materia
	Datos de referencia cruzados.	Administradores de datos	Interesados	Analistas de datos, expertos en la materia
	Reportes de linaje de datos.	Arquitectos de datos	Administradores de datos	Analistas de datos
	Reportes de calidad de datos.	Analistas de datos	Administradores de datos, interesados	Arquitectos de datos
7. Definir y mantener las jerarquías y afiliaciones (C)	Definir jerarquías y afiliaciones.	Administradores de datos	Interesados	Analistas de datos, proveedores de datos
8. Planear e implementar la integración de nuevas fuentes (D)	Calidad de datos fuentes y evaluaciones de integración.	Analistas de datos, arquitectos de datos, arquitectos de aplicaciones	Administradores de datos, administrador de TI	Proveedores de datos, expertos en la materia
	Nuevas fuentes de datos integradas.	Arquitectos de datos, arquitectos de aplicaciones	Administradores de datos, interesados	Analistas de datos, otros profesionales de TI
9. Replicar y distribuir los datos maestros y de referencia (O)	Datos replicados.	Arquitectos de datos, arquitectos de aplicaciones	Administradores de datos, interesados	Analistas de datos, otros profesionales de TI
10. Administrar cambios para datos maestros y de referencia (C)	Procedimientos de requerimientos de cambios.	Arquitectos de datos	Consejo de gobierno de datos, administradores de datos	Otros profesionales de TI, interesados
	Requerimientos de cambios y respuestas.	Administrados de datos	Consejo de gobierno de datos	Interesados, analistas de datos, arquitectos de datos, arquitectos de aplicaciones
	Métricas de requerimientos de cambios.	Arquitectos de datos	Administradores de datos y Consejo de gobierno de datos	Analistas de datos, otros profesionales de datos.

Tabla 3.1 Actividades, entregables y roles. Traducida de DAMA, 2009

Respecto a los enfoques arquitectónicos, el DAMA propone para la integración de datos maestros y de referencia tener una base central que contenga estos datos. En la Figura 3.3 se muestra que algunas aplicaciones pueden leer datos o realizar actualizaciones en la base central de datos maestros.

Un sistema de registros proporciona los datos definitivos sobre una instancia (DAMA, 2009). Dado esto, un sistema de registros es una aplicación que usa y actualiza información de la base de datos maestros y de referencia, es decir, recopila y suministra datos. Este sistema deberá proveer de servicios para la verificación y autorización de un nuevo dato o una actualización de datos.

Por otro lado un mercado de datos, o en inglés conocido como Data Mart, está compuesto de subconjunto de datos con el propósito de ayudar a un área específica dentro del negocio para la toma de decisiones.

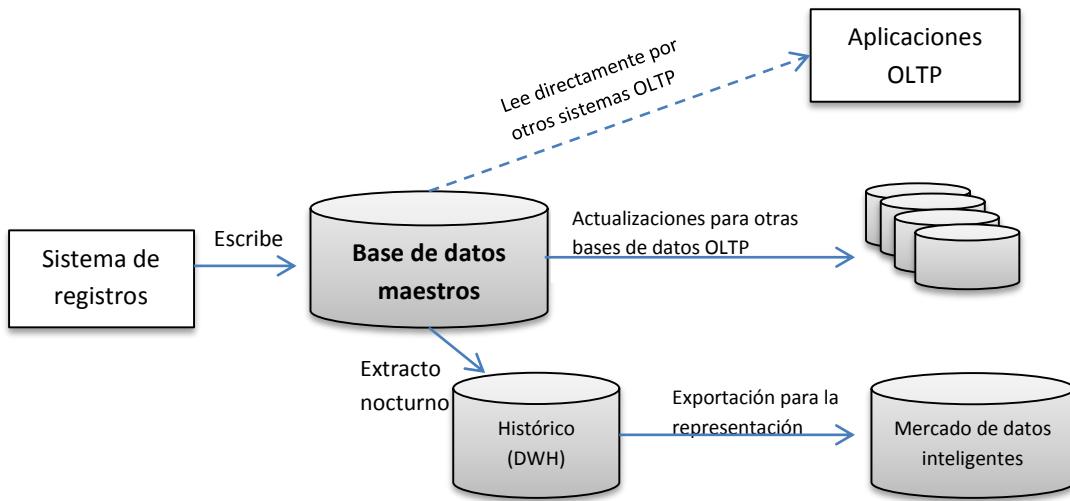
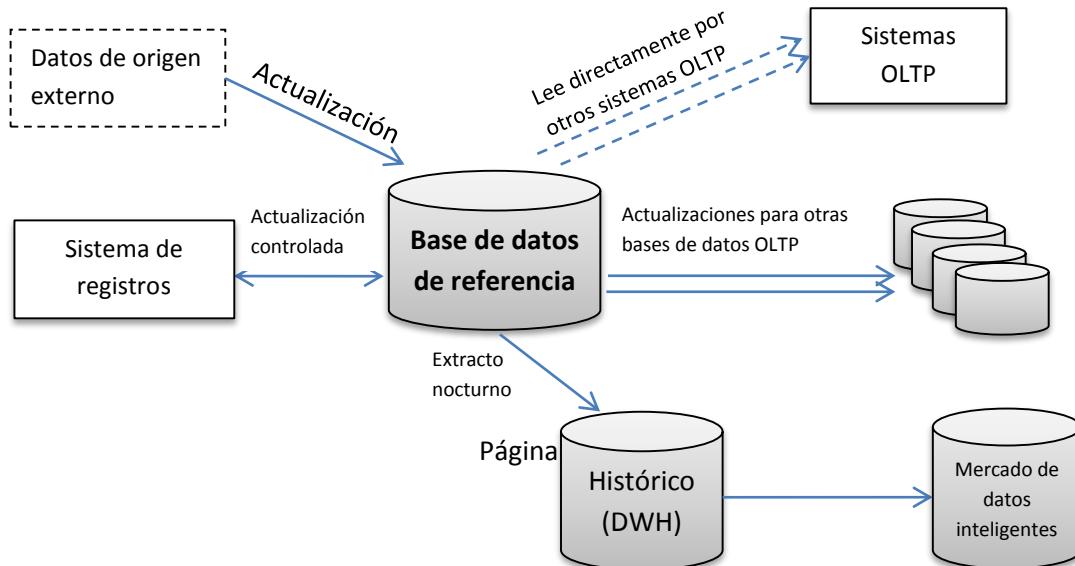


Figura 3.3 Arquitectura de DM. Adaptada de DAMA, 2009.

Los datos maestros y de referencia se usan a través de aplicaciones. En la Figura 3.4 se muestra que para proveer estos datos se pueden 1) replicar desde una base de datos central a otras bases ó 2) otras aplicaciones pueden leer directamente de la base de datos maestros y de referencia.

La actualización de datos de origen externo se deberá controlar mediante la definición de un proceso que establezca si la nueva versión del dato es adaptable al negocio.



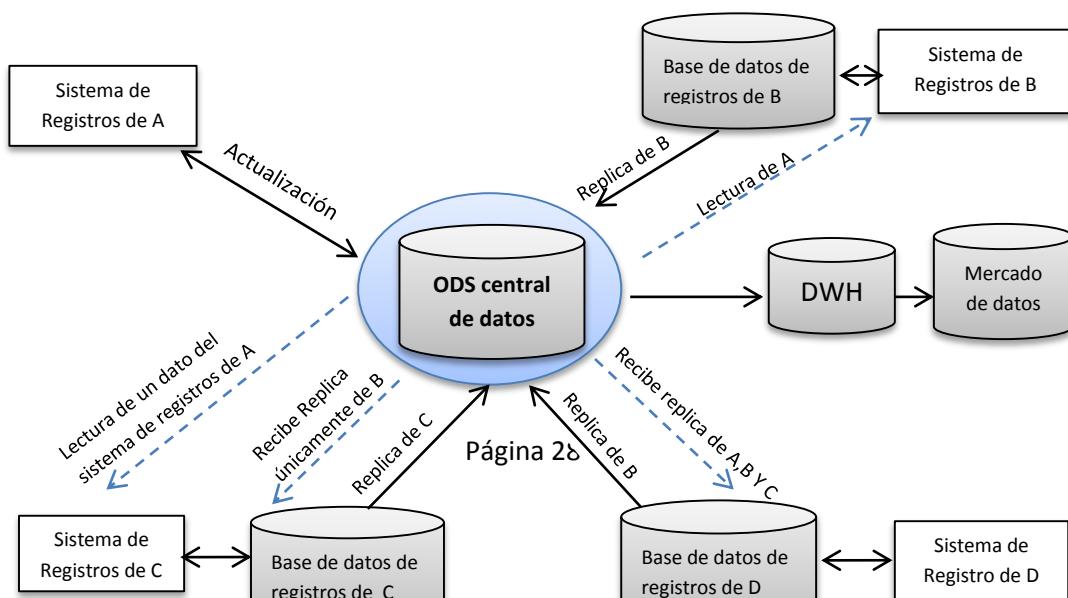
Exportación para la representación

Figura 3.4 Arquitectura de DR. Adaptada de DAMA, 2009.

Un Almacén de Datos Operativos, o por sus siglas en inglés Operational Data Store ODS, es una base de datos que integra información de múltiples fuentes. Debido a que un almacén de datos operativos recibe información que proviene de múltiples fuentes, su integración implica la limpieza y comparación de reglas de negocio para estos datos.

En la Figura 3.5 hay cuatro sistemas de registros (A, B, C, D), suponiendo que cada letra corresponde a una temática diferente en una organización, se presentan los siguientes casos:

- El sistema A proporciona actualizaciones al Almacén de Datos Operativos central sin tener base de datos y no necesita datos de los sistemas B, C y D.
- Los sistemas B, C Y D cuentan con sus propias bases de datos.
- El sistema B realiza una lectura de los datos que usa el sistema A y replica sus datos maestros al Almacén de Datos Operativos.
- EL sistema C realiza una lectura de los datos del sistema A y en su base de datos recibe datos del sistema B. Finalmente replica sus datos a través del Almacén de Datos Operativos.
- La base de datos del sistema D recibe datos replicados de A, B y C para posteriormente replicar sus datos a través del Almacén de Datos Operativos.



**Figura 3.5 Arquitectura de un ODS central de datos. Adaptada de DAMA, 2009.**

Con base en el DAMA una arquitectura de integración de datos debe proporcionar servicios que incluyan:

- Cambio en el proceso de requerimientos, incluyendo la revisión y aprobación.
- Calidad de los datos de referencia (adquiridos externamente) y maestros.
- Aplicación de normas de calidad de datos y reglas de correspondencia.
- Patrones consistentes de procesamiento.
- Meta-datos coherentes de las asignaciones, las transformaciones, programas y flujos de trabajo.
- Auditoría consistente, solución de errores y monitoreo de datos.
- Enfoques consistentes para la replicación de datos, incluyendo suscripción y publicación.

Finalmente el DAMA proporciona los siguientes principios para los datos maestros y de referencia:

1. Los datos maestros y de referencia son compartidos y pertenecen a la organización, no a una aplicación o servicio en particular.
2. La administración de datos de datos maestros y de referencia es un programa para mejorar la calidad de datos, sus objetivos no pueden ser alcanzados por un solo proyecto.
3. Los administradores de datos del negocio son las autoridades responsables de controlar los valores de los datos maestros y de referencia. Este rol trabaja con profesionales de datos para mejorar su calidad.
4. Los valores de oro representan los esfuerzos de los datos más exactos, actuales y relevantes para su uso contextual. Los nuevos datos pueden resultar falsos, por lo tanto, se aplican las reglas de correspondencia con precaución y seguridad de que cualquier cambio que se realiza es reversible.
5. Se puede replicar valores de datos maestros y de referencia sólo de la base de datos de registro.
6. Es importante solicitar, comunicar, y aprobar cambios de los valores de datos maestros y de referencia antes de la implementación.

Debido a que las necesidades de una arquitectura empresarial es posible considerar las actividades, roles, servicios y principios que provee esta organización para la gestión de datos maestros y de referencia. Para la definición de estos elementos también es posible considerar los diferentes enfoques arquitectónicos que se discutieron previamente.

### 3.2.2 Master Data Management Institute

Fue fundado en 2004 y proporciona consultorías de investigación exclusivamente en la gestión de datos maestros a empresas de TI y organismos gubernamentales. El Instituto tiene su sede en San Francisco y tiene clientes principalmente en América del Norte, Europa y Asia. Patrocina conferencias en Londres, Nueva York, San Francisco, Singapur, Sydney, Tokio y Toronto, siendo sus especialidades gestión de datos maestros y de referencia, gobierno de datos e integración de datos de clientes.

La administración de datos maestros es la autorización para que los datos maestros sean utilizados en muchas aplicaciones con el objetivo de proporcionar una visión única de la verdad, sin importar dónde se encuentran (MDM Institute, 2013).

Las sub definiciones para la administración de datos maestros son:

- Operacional: Se refiere a la definición, creación y sincronización de los datos maestros necesarios para los sistemas transaccionales y entregados a través de la arquitectura orientada a servicios. Un ejemplo son los centros de datos de clientes actualizados en tiempo real.
- Analítica: Se refiere a la definición, creación y análisis de datos maestros. Ejemplos de esto son las aplicaciones de gestión de riesgo de contraparte e información financiera, gasto global o plan de cuentas de consolidación.
- Colaborativa: Se refiere a la definición, creación y sincronización de datos de referencia a través de flujos de trabajo, servicios de salida. Ejemplo gestión de la información de los productos.

Las siguientes fases para un proyecto para la gestión de datos maestros son las siguientes (MDM Institute, 2009):

- Fase cero: Definir el alcance de la siguiente fase, capturar requerimientos, proyectar el retorno sobre la inversión, vender y evaluar el proyecto.
- Fase uno: Desarrollar una entidad maestra por cada división o departamento del negocio, ejemplo cliente o producto.
- Fase dos: Elegir una entidad maestra para la empresa.
- Fase tres: Compartir los datos maestros.

El diseño de la solución para la gestión de datos maestros debe tener una arquitectura orientada a servicios (SOA) y las aplicaciones enfocadas a componentes. La figura 3.6 ilustra

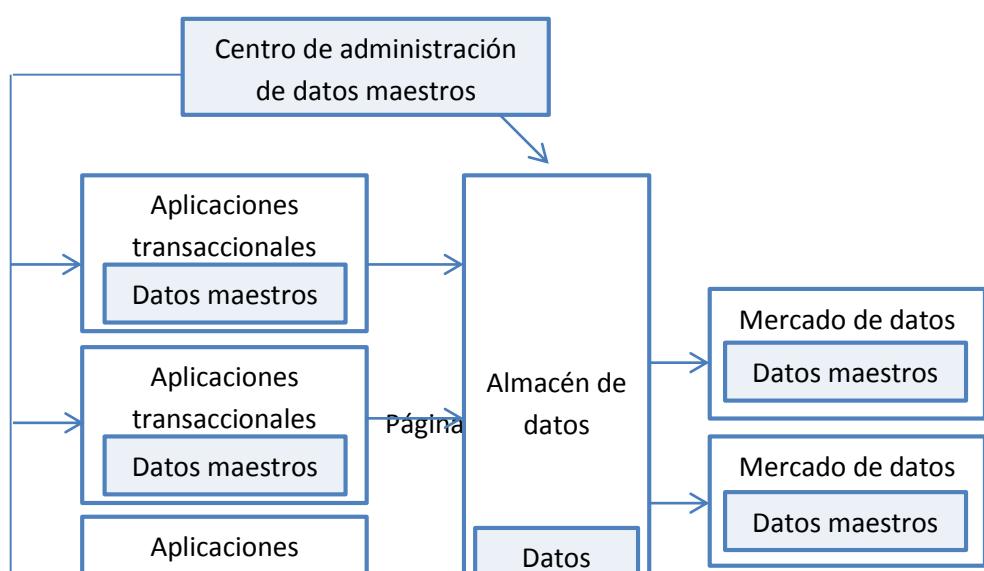
los componentes que comprenden la administración de datos maestros necesario para ofrecer una solución robusta.

Es posible considerar para el desarrollo de la arquitectura de datos maestros y de referencia las fases que nos proporciona este marco de referencia, así como los componentes que se muestran en la figura previa. El desarrollo de la arquitectura del presente trabajo se orientara hacia el diseño de servicios SOA.

### 3.2.3 Enterprise Data World

Es un evento educativo sobre el manejo de datos. Durante cinco días, este evento presenta un calendario diverso de programación incluyendo conferencias magistrales, talleres, tutoriales, estudios de casos y debates. En el evento que se llevó a cabo el mes de abril del 2013 se mostró información relacionada con la arquitectura de datos maestros, la cual se expone a continuación.

En la Figura 3.7 se muestra la arquitectura de la administración central de datos maestros y almacenes de datos (DWH).

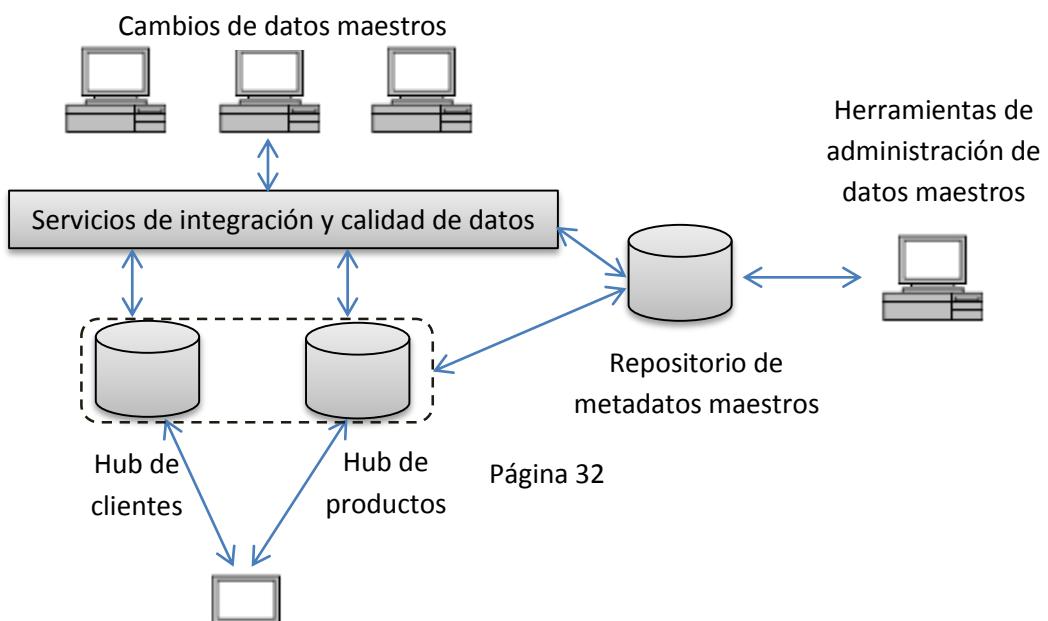


La arquitectura contempla un centro de administración de datos maestros, estos se comparten con el ámbito transaccional e histórico. No se comparten directamente datos maestros, del centro de administración, a los mercados de datos. Esto se debe a que el almacén de datos es quien genera los mercados de datos. Ahí se pueden encontrar datos maestros analizados.

En este mismo evento la empresa Intelligent Business Strategies compartió la siguiente definición para la administración de datos maestros:

Conjunto de políticas, servicios, procesos y tecnologías usadas para crear, mantener y administrar los datos asociados con las entidades primordiales de negocio como un sistema de registros (EDW, 2013).

La administración de datos maestros provee servicios comunes que se pueden compartir a través de múltiples aplicaciones. En la Figura 3.8 existe un repositorio de metadatos maestros, un hub para cada elemento maestro (cliente, producto, etc.), servicios de integración y calidad de datos, y aplicaciones.



El repositorio de metadatos maestros se relaciona con: los hubs maestros, servicios de integración y calidad de datos, y con la herramienta de administración de datos maestros. Los servicios de integración y calidad de datos se relacionan con los hubs de datos maestros y con aplicaciones que recogen los cambios a estos datos. Finalmente para generar informes de datos los hubs de datos maestros se relacionan directamente con una aplicación, no existe una relación intermedia con los servicios de integración. La integración de datos maestros no se debe realizar en el ambiente histórico.

Por otro lado, la empresa Intelligent Business Strategies define que las políticas son las que limitan la capacidad de gobernanza de los datos maestros. Estas políticas se deben enfocar en:

- Restricciones para la integridad
- Validación de datos
- Limpieza de datos
- Integración de datos
- Formato del suministro de datos
- Control de acceso
- Privacidad de datos
- Expiración de datos
- Respaldo y almacenamiento de datos

Para el desarrollo del capítulo 4 es posible retomar del EDW que la arquitectura de datos maestros y de referencia debe considerar los repositorios centrales de estos datos, sus metadatos, herramientas para la gestión y servicios de integración. Como parte de los requerimientos para el desarrollo de funciones y procesos de una arquitectura empresarial, se consideraran las políticas que se proveen en esta sección.

### **3.3 Esquemas comerciales para soluciones de datos**

La información conceptual que proveen cada uno de los siguientes esquemas se obtuvo de sus artículos publicados.

### 3.3.1 Oracle

Es una compañía de software que se enfoca al desarrollo de productos que van desde bases de datos hasta sistemas de gestión. Cuenta con herramientas propias de desarrollo para realizar potentes aplicaciones.

La gestión de datos maestros es la combinación de las aplicaciones y tecnologías que consolidan, limpian e incrementan la colección de datos maestros, y sincronizan esto con todas las aplicaciones, procesos de negocio y herramientas de análisis (Oracle, 2011).

La administración de datos maestros consiste en elaborar perfiles, consolidar, sincronizar, limpiar y enriquecer los datos maestros estructurados y no estructurados. Las aplicaciones de la administración de datos maestros deben soportar las siguientes características:

- Capacidad para administrar metadatos.
- Capacidad para administrar fuentes de datos maestros a través de referencias cruzadas.
- Capacidad para gestionar la migración de datos.
- Tener funciones de limpieza de datos.
- Tener una interfaz que ayude a prevenir errores de datos al ingresarlos.
- Tener una interfaz para los usuarios y administradores de datos.
- Contar con una plataforma única y escalable de alta disponibilidad.
- Contar con un mecanismo interno para activar algún cambio en los datos a los sistemas que los utilizan.
- Contar con un sistema de seguridad que controle y monitoree el acceso a los datos, así como el historial de cambios.
- Contar con un modelo flexible para mantener las necesidades de datos maestros (el modelo de datos debe soportar trabajos de carga OLTP).

Todo lo mencionado previamente se resume en la arquitectura que se muestra en la Figura 3.9 a continuación.

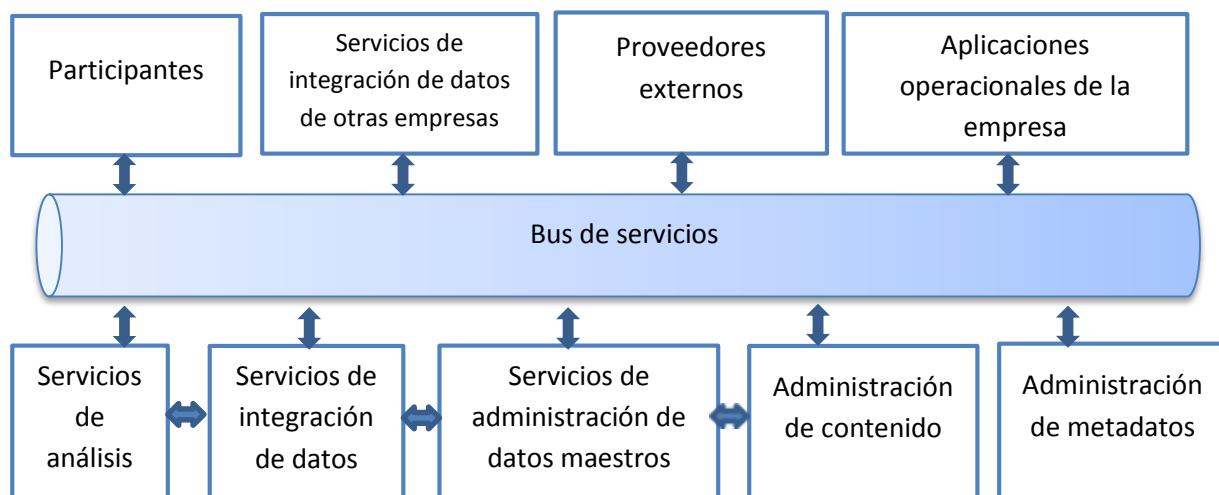


Figura 3.9 Servicios para la gestión de datos maestros. Adaptado de Oracle, 2011.

El software que proporciona Oracle para la gestión de datos maestros recibe el nombre de Master Data Management Suite (Oracle, 2014). La suite de productos Oracle Enterprise Master Data Management sirven para consolidar y mantener completos, exactos y acreditados los datos maestros de toda la empresa, y distribuir esta información a todas las aplicaciones operativas y analíticas como un servicio compartido.

Las aplicaciones MDM gestionan datos de clientes, proveedores, productos y finanzas con servicios de gobierno de datos, y son compatibles con los mejores componentes de inteligencia empresarial e integración.

Este esquema comercial sugiere, como parte de la gestión de datos maestros, elaborar perfiles para acceder a ellos, así como consolidarlos, sincronizarlos, limpiarlos y enriquecerlos. También nos proporciona los servicios y capacidades con las que deben cumplir las aplicaciones MDM.

Para el desarrollo del capítulo 4 es posible considerar las actividades de consolidación, sincronización y limpieza de datos debido a que uno de los objetivos de la arquitectura empresarial es definir los procesos que regirán, en este caso, a los datos maestros y de referencia . También se tomará en cuenta la existencia de fuentes externas e internas para la definición de estos datos.

### **3.3.2 IBM**

Es una empresa que se dedica al desarrollo de tecnología, fabrica y comercializa hardware y software para computadoras. También ofrece servicios de infraestructura tecnológica.

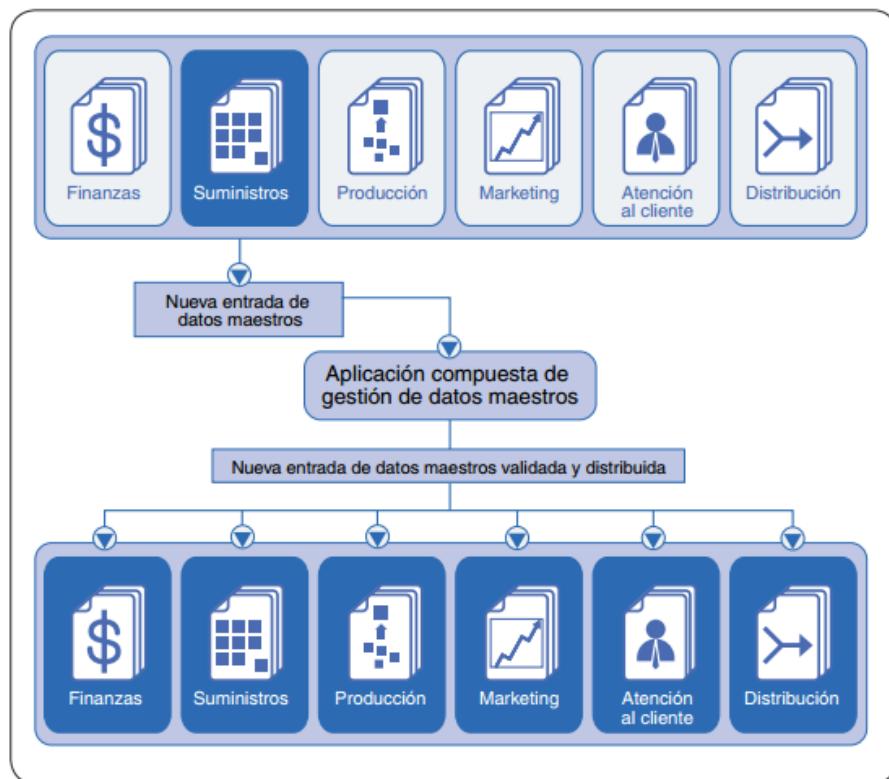
La gestión de datos maestros es el conjunto de disciplinas, tecnológicas y soluciones utilizadas para crear y mantener datos coherentes, completos, contextuales y precisos de todas las partes: usuarios, aplicaciones, almacenes de datos, procesos y socios (IBM, 2007).

Esto quiere decir que el objetivo de la administración de datos maestros es conformar una vista única de la información para proporcionarla a las personas y los procesos necesarios.

La gestión de datos maestros proporciona un método mediante el cual una organización puede gestionar de manera eficaz datos ya presentes en distintos sistemas. Las características que debe tener la gestión de datos maestros son:

- Proporcionar un proceso de autorización para crear, definir y sincronizar los datos.
- Proporcionar la recuperación, actualización, distribución y distintos usos de los datos.
- Usar los datos de manera operativamente para la integración de diferentes entornos en tiempo real.
- Distribuir los datos a través de procesos y servicios.
- Soportar el uso analítico de los datos
- Con el fin de acelerar la entrada, recuperación y análisis de datos maestros, tener la información unificada de clientes, productos y proveedores.
- Soportar varios usuarios para añadir, actualizar y consultar los procesos que mantienen los datos maestros.

En la Figura 3.10 se muestra una solución para la gestión de datos maestros, la aplicación central denominada “Aplicación compuesta de gestión de datos maestros” recopila toda la información introducida por un área de negocio (Suministros) para replicarla a los demás rubros del negocio (Finanzas, Suministros, Producción, Marketing, Atención al cliente, Distribución). Esto quiere decir que los métodos y procesos de los datos maestros deben funcionar centralizadamente, independientemente de la línea de negocio, para responder a distintos usos.



**Figura 3.10 Solución para la gestión de datos maestros. Tomada de IBM, 2007.**

Un sistema de administración de datos maestros requiere la capacidad de distribuir los datos en un entorno heterogéneo cumpliendo con requisitos de negocio y técnicos. En la Figura 3.11 se muestra la arquitectura lógica para la gestión de datos maestros que propone IBM.

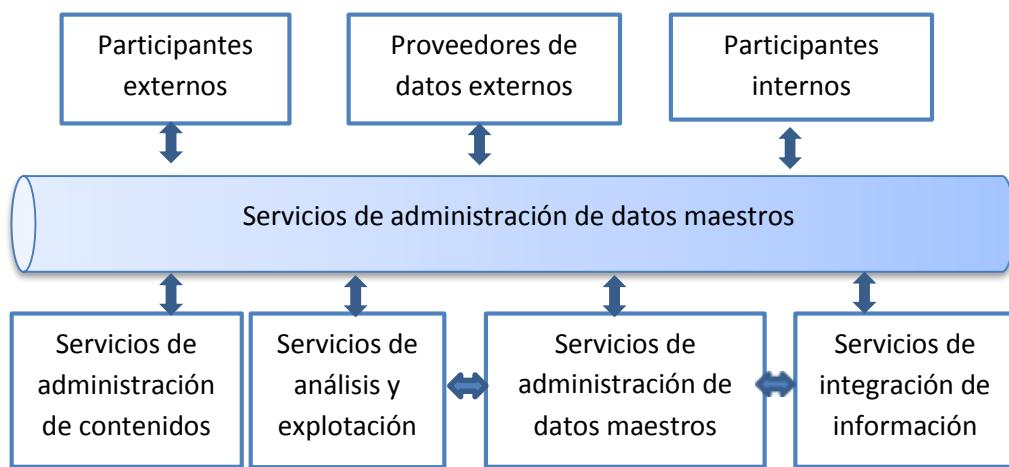


Figura 3.11 Arquitectura de servicios para datos maestros. Adaptada de IBM, 2011.

Los participantes externos, proveedores de datos externos y los participantes internos pueden acceder y actualizar datos a través de canales de distribución.

Respecto a los principios para la gestión de datos maestros, IBM los define como una ley integral y fundamental que proporciona una guía para el desarrollo de una solución. Los siguientes puntos son principios fundamentales para guiar el desarrollo de la gestión de datos maestros:

- Disponer de los datos maestros como un activo estratégico para su uso.
- Proporcionar una fuente autorizada de datos maestros
- Proporcionar flexibilidad para adaptarse a cambios
- Preservar la integridad y seguridad de los datos.
- Basarse en estándares para apoyar el uso de múltiples tecnologías
- Basarse en un marco arquitectónico y servicios reutilizables para aprovechar las tecnologías existentes.

La solución que proporciona IBM para la gestión de datos maestros es InfoSphere Master Data Management V11.3 (IBM, 2014). Este software comprende con un repositorio maestro físico que presenta una única versión de las entidades fundamentales de datos de una empresa como clientes, productos, proveedores, etc.

Es posible considerar para el desarrollo del presente trabajo la integración, sincronización, recuperación, actualización y distribución de datos maestros a través de la definición de

procesos y servicios. Al igual que Oracle hace mención de la influencia de los participantes externos e internos para la definición del contenido de estos datos.

También se tomaran en cuenta tanto los servicios y principios que proporciona IBM para la gestión de datos maestros y de referencia.

### 3.3.3 Microsoft

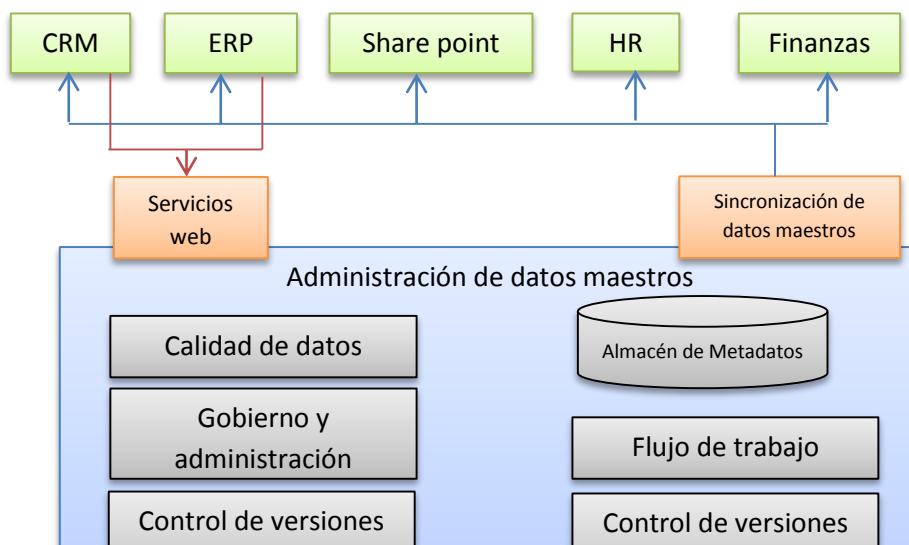
Empresa de origen estadounidense que desarrolla, fabrica, licencia y respalda productos y servicios relacionados con la informática.

La administración de datos maestros es la tecnología, herramientas y procesos requeridos para crear y mantener consistente y exacto listas de datos maestros. La administración de datos maestros incluye tanto la creación como el mantenimiento de datos maestros (Microsoft, 2006).

El plan de trabajo para la administración de datos maestros contiene las siguientes fases:

1. Identificar las fuentes de datos maestros
2. Identificar los productores y consumidores de datos maestros
3. Reunir y analizar metadatos sobre datos maestros
4. Asignar administradores de datos.
5. Implementar un programa y comité de gobierno de datos
6. Desarrollar un modelo de datos maestros
7. Elegir un conjunto de herramientas
8. Diseñar la infraestructura
9. Generar y probar los datos maestros
10. Modificar los sistemas productores y consumidores

En la Figura 3.12 se muestra que el hub para la administración de datos maestros contiene funciones y herramientas requeridas para mantener consistentes y exactas las entidades de datos maestros. Como se observa, la función de sincronización es responsable de vigilar y coordinar el dato maestro con sistemas transaccionales.



**Figura 3.12 Arquitectura de la administración de datos maestros. Traducida de Microsoft, 2006**

Microsoft proporciona tres estilos de arquitectura: registro, repositorio e híbrido.

- En el enfoque del repositorio, los datos maestros se almacenan en una base de datos única. Las características de este estilo son:
  - El modelo de datos incluye todos los atributos requeridos por las aplicaciones.
  - Las aplicaciones usan datos del hub central y no de su propia base de datos.
  - El hub central es previamente mantenido por la base de datos de la aplicación.
  - No hay problema con el mantenimiento de múltiples versiones de un registro, también se asegura la sincronización de múltiples aplicaciones.
- El enfoque de registro es opuesto al enfoque de repositorio. Sus características son:
  - Los registros se almacenan en las bases de cada aplicación, por lo tanto cada aplicación mantiene sus propios datos.
  - El hub central contiene listas del mapeo de las claves.
  - La desventaja de este modelo es que se tiene que realizar consultas distribuidas.
- El enfoque híbrido incluye características de los anteriores enfoques. Sus características son:
  - Los registros de datos maestros se almacenan en las bases de cada aplicación y las claves se mantienen en el hub central.
  - Los atributos más importantes para cada entidad maestra se replican en el hub central, los atributos menos comunes se dejan en la base de datos de la aplicación.
  - Los conflictos que se tienen es la actualización, replicación de latencia, definición de que atributos y en qué formato se almacenaran en el hub.

En los tres enfoques se debe incluir un modelo de datos común para la base de datos central.

Por otro lado, el software que proporciona Microsoft para la gestión de datos maestros es una solución de SQL Server denominada Master Data Services (Microsoft, 2014). Donde se realizan esfuerzos para detectar y definir listas no transaccionales de datos con el objetivo de compilar listas maestras que se puedan mantener. Un proyecto de datos maestros incluye la evaluación y reestructuración de los procesos corporativos internos junto con la implementación de tecnología.

Entre otras características de Master Data Services se incluyen jerarquías, seguridad específica, transacciones, control de versiones de datos y reglas de negocios.

Como parte del desarrollo de la arquitectura de datos maestros y de referencia es posible considerar las fases que proporciona este esquema comercial, así como orientar esta arquitectura hacia un estilo de repositorio. También es posible considerar la sincronización como una de las principales funciones de la gestión de estos datos.

### **3.3.4 Informatica**

Empresa dedicada al software de integración de datos, fue fundada en 1993 y actualmente cuenta con 2690 empleados. Se maneja a la administración de datos maestros como el enfoque que usa gente, procesos y tecnología en la empresa para proveer y administrar una vista de unificada, consistente y exacta de datos maestros.

La visión de Informatica respecto a la administración de datos maestros no es una aplicación, solución tecnológica o conjunto de herramientas que se desarrolla para solventar un problema tecnológico en la empresa.

Respecto a los principios de datos maestros, Informática define que son esenciales para la toma de decisiones, se basan en la industria, estrategia y requerimientos de negocio. A continuación se muestran algunos principios para los programas de administración de datos maestros:

- Administrar diligentemente los datos maestros como otros activos estratégicos de la empresa.
- El negocio posee a los datos maestros y es responsable de su integridad.
- Establecer estándares y gobierno para los procesos de datos maestros.
- Configurar el marco para habilitar un entorno de tecnología.
- Establecer una versión única.
- Comunicar que la calidad de los datos depende de la excelencia de los procesos.
- Garantizar la coherencia y la alineación en todas las actividades relacionadas.
- Establecer el proceso de gestión del cambio.

Una solución de datos maestros bien diseñado asegura el éxito de la implementación de SOA, que conduce a la mejora de la productividad del desarrollo de aplicaciones basados en estándares.

La selección de la tecnología debe realizarse basado en procesos para ayudar a manejar el riesgo, horario, costo y esfuerzo. Para la gestión d datos maestros puede implementarse una variedad de estilos arquitectónicos y la selección de la arquitectura depende de los requerimientos del negocio, estado de los datos y el nivel de preparación.

El software que proporciona esta empresa para la gestión de datos maestros se denomina Informatica MDM (Informatica, 2014). Este software cuenta con tecnología multidominio, las implantaciones son rápidas y se pueden ampliar a toda la empresa. Modela todos los dominios de datos maestros para satisfacer las necesidades específicas y únicas de su negocio sin imponer definiciones de negocio fijas o predefinidas. También mantiene el historial y el linaje de todos los datos maestros relevantes con fines de auditoría y cumplimiento de normativas.

Para el desarrollo de la arquitectura empresarial de datos maestros y de referencia es posible considerar los principios que nos provee este esquema comercial. Se orientara esta arquitectura principalmente hacia el diseño de procesos y la implementación de servicios SOA.

### **3.4 Modelo de madurez de la gestión de datos maestros**

David Loshin, presidente de Knowledge Integrity Inc. es reconocido como uno de los expertos en las áreas de calidad de datos, administración de datos maestros e inteligencia de negocios.

La gestión de datos maestros es la recopilación o colección de las mejores prácticas que incorporan aplicaciones de negocio, métodos y herramientas de gestión de datos para aplicar políticas, procedimientos, servicios e infraestructura que soporte la captura, integración y uso compartido de datos maestros consistentes, completos, precisos y oportunos. (Loshin, 2009)

Loshin es el autor del modelo por sus siglas en inglés Master Data Management Maturity Model, el cual evalúa la madurez de la gestión de datos maestros. Este modelo consta de seis componentes: arquitectura, gobierno, administración, identificación, integración y administración de procesos de negocio. En la Figura 3.13 se presentan las siguientes capacidades y servicios que debe abarcar una solución madura para cada componente:

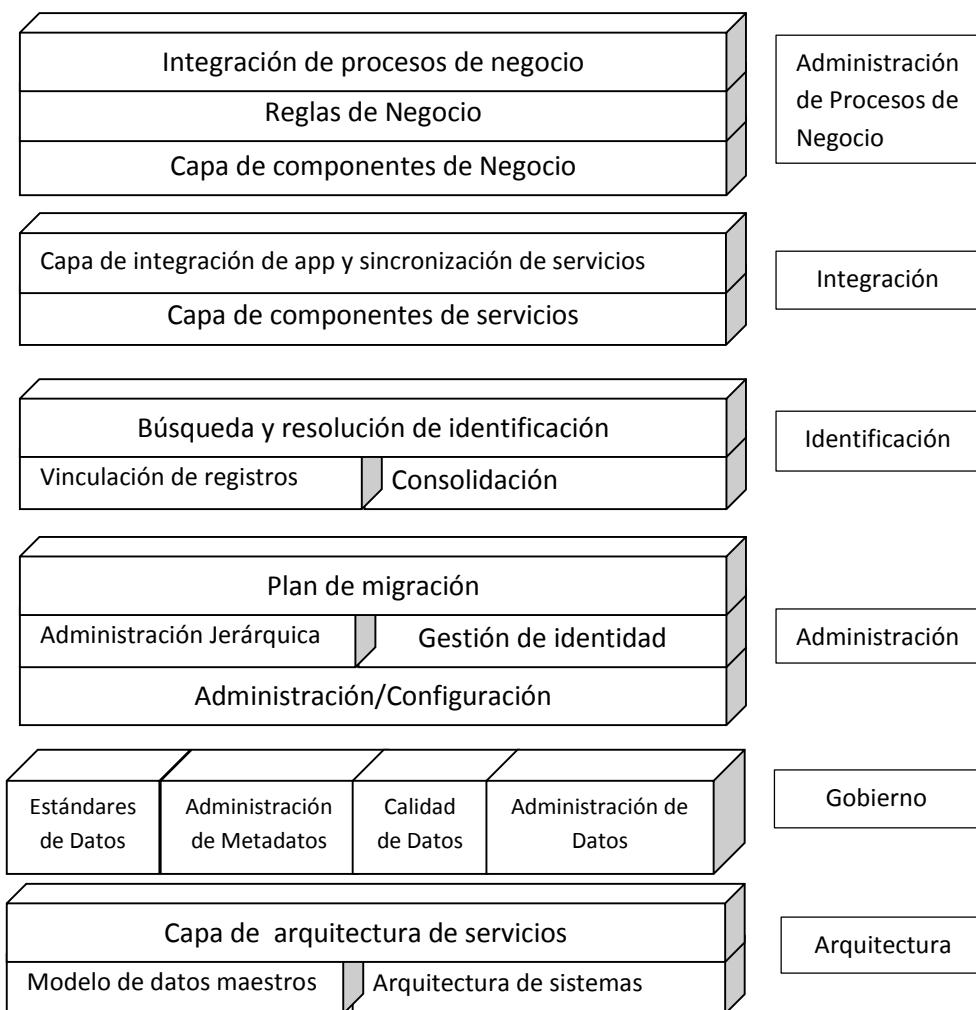


Figura 3.13 MDM Madurez Model. Traducida de Loshin, 2009.

Las capacidades de estos componentes se representan en función de su nivel de madurez. Para ello, los niveles de madurez son: inicial, reactivo, administrado, proactivo y desempeño estratégico.

Las características del cumplimiento de cada nivel dependen de cada componente. El objetivo de este modelo es medir las buenas prácticas que se llevan a cabo en una organización respecto a los datos maestros.

#### **4. ARQUITECTURA DE DATOS MAESTROS Y DE REFERENCIA**

Como se ha mencionado en capítulos anteriores, los datos maestros son los elementos clave de una organización y sirven para proporcionar un contexto confiable a los procesos que verifican la calidad de las operaciones transaccionales, también se usan como parte de los informes analíticos para la toma de decisiones. Ejemplos de datos maestros son usuarios, proveedores y productos de una empresa.

Por su parte los datos de referencia representan listas de claves que contienen significados de negocio. Ejemplos de datos de referencia son catálogos de monedas, catálogos de países.

Debido a que en las organizaciones no existe un tratamiento y control específico de los datos maestros y de referencia, es posible que estos datos no estén actualizados. Por ejemplo un cliente puede cambiar de trabajo, teléfono, dirección, etc. Entonces ¿Cómo mantener actualizados estos conjuntos de datos para no tomar decisiones con información obsoleta?

Hoy en día existen soluciones comerciales por parte de empresas dedicadas a las tecnologías de la información para la gestión de datos maestros, sin embargo son poco accesibles en cuanto a presupuesto. Por ello, el presente capítulo desarrolla una propuesta para la gestión de los datos maestros y de referencia enfocada a una Arquitectura Empresarial (AE).

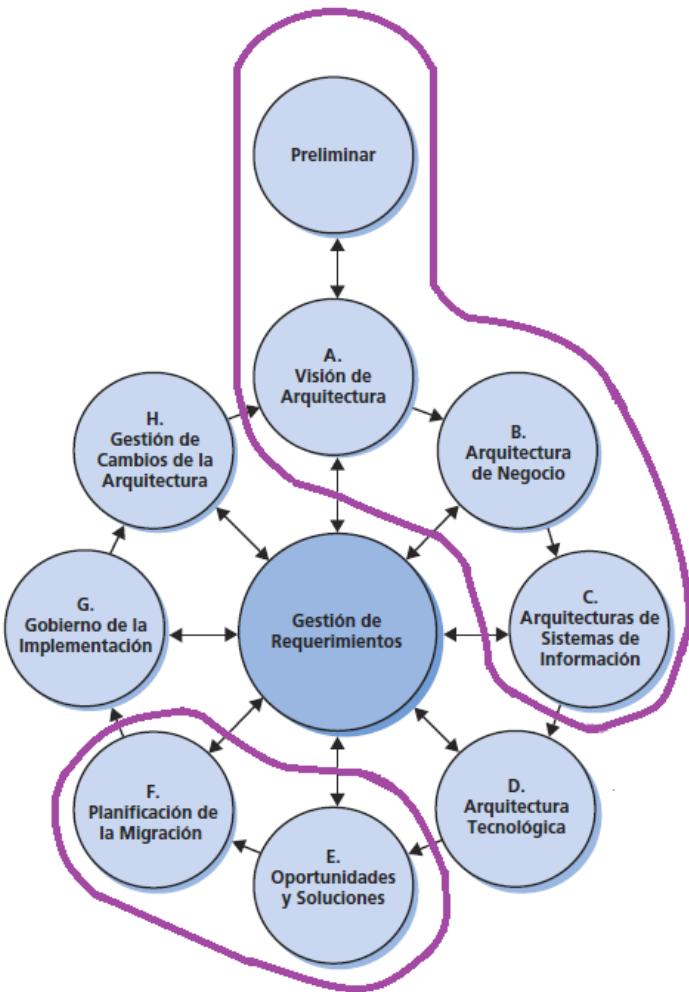
¿Por qué diseñar una solución basada en una AE? Una arquitectura empresarial permite representar y entender las necesidades de las organizaciones enfocada a una gestión de procesos. La principal razón por la cual se decidió basar su diseño en una AE es porque alinea estratégicamente las necesidades del negocio con la tecnología por medio de la definición de relaciones entre procesos, personas, servicios, aplicaciones y activos de información.

Dado lo anterior, se ha seleccionado el marco de referencia TOGAF (The Open Group Architecture Framework) que contiene una metodología para desarrollar los elementos de una arquitectura empresarial. A través del Método de Desarrollo de la Arquitectura o por sus siglas en inglés ADM, TOGAF establece diferentes iteraciones y fases para el desarrollo de los siguientes elementos:

- La Iteración de *Contexto* comprende la fase preliminar y visión de la arquitectura, donde los elementos a desarrollar son la evaluación de madurez, la selección de marcos metodológicos a utilizar, identificación de interesados, el alcance, la misión y la visión de la capacidad arquitectónica a diseñar.
- Para la Iteración de *Definición* se comprenden las siguientes fases:
  - Arquitectura de Negocio: Los elementos a desarrollar de este dominio son procesos, funciones, servicios, roles, principios y políticas de negocio.
  - Arquitecturas de Sistemas de Información: Artefactos para datos a nivel conceptual, lógico y físico, así como el desarrollo de servicios y componentes SOA.
  - Arquitectura tecnológica: este dominio contempla los requerimientos del hardware necesario (como routers y servidores) para soportar la anterior fase.
- Para la Iteración de Planeación se comprenden las fases de Oportunidades y Soluciones, y Planeación de Migración. En estas fases se desarrollan las iniciativas para cumplir implementar la arquitectura definida en los dominios previamente mencionados, así como se desarrolla un plan de migración.
- Para la Iteración de Gobierno se comprenden las fases de Gobierno de la Implementación y Administración de Cambios. Los elementos a desarrollar son principalmente mecanismos para auditar y vigilar que se ha implementado correctamente la arquitectura diseñada.

El alcance de este trabajo contempla las siguientes iteraciones basadas en TOGAF, las cuales se pueden visualizar en la Figura 4.1.

- Iteración de contexto: Fase Preliminar y A (Visión).
- Iteración de definición: Fase B (Arquitectura de negocio) y Fase C (Arquitectura de datos).
- Iteración de planeación: Fase E (Oportunidades y soluciones) y Fase F (Planeación de migración).



**Figura 4.1 Alcance del desarrollo de la arquitectura de DMyR. Traducida de TOGAF,2011.**

El alcance de este trabajo no contempla el desarrollo de las Fases C (arquitectura de aplicaciones), D (Infraestructura Tecnológica), G (Gobierno de la implementación), H (Arquitectura de gestión de cambios) y la fase de Administración de Requerimientos debido a las siguientes razones:

- El principal impulsor para desarrollar este trabajo es el estudio y mejora de la Arquitectura de Datos.
- Para llegar al dominio de la arquitectura de datos de la Fase C del marco TOGAF “Arquitecturas de Sistemas de Información”, es necesario desarrollar las fases previas a esta, es decir, Fase Preliminar, A Y B.
- Puesto que esta tesis debe contemplar una implementación y con el fin de verificar la funcionalidad del diseño de las fases A, B y C (únicamente la parte de datos), se han desarrollado la Fase E y F.

Para el dominio de datos de la Fase C, se desarrollará la arquitectura a nivel conceptual y lógico. No se especificará la arquitectura física de datos debido a que no se contempla el desarrollo de la arquitectura de aplicaciones.

En la Figura 4.8 se muestra un esquema de las arquitecturas que incluyen la gestión de datos maestros y de referencia, como se mencionó anteriormente, para la arquitectura de datos este trabajo solamente contempla el nivel conceptual y lógico, sin embargo, la arquitectura de aplicaciones y de tecnología no se desarrollarán en este trabajo.

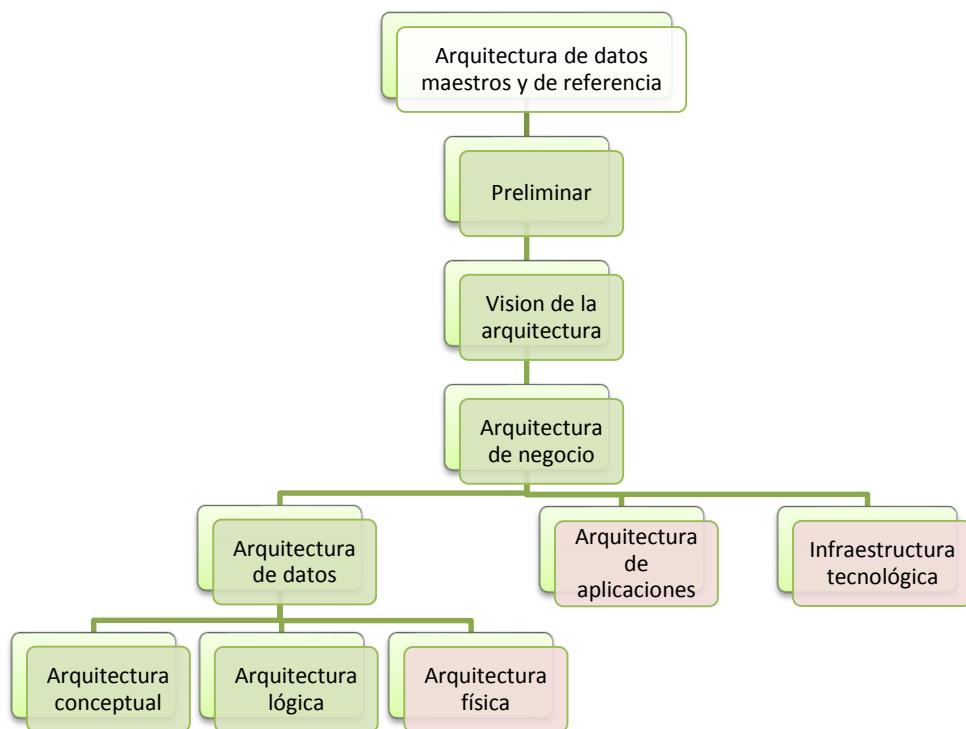


Figura 4.2 Alcance del desarrollo de la Arquitectura de DMyR

El esquema muestra con color verde las arquitecturas de definición que se desarrollan en este trabajo, con color rojo se muestran los dominios que se no se contemplan en el alcance.

#### 4.1 Fase 0: Preliminar

Con base en la metodología TOGAF en esta fase se pretende definir las metas, marcos de referencia a utilizar en conjunto y las directrices para llevar a cabo una evaluación de madurez.

Las principales metas que conducen este trabajo se resumen en las siguientes:

- Contar con datos maestros y de referencia verídicos en todos los sistemas. Esto proporcionará un contexto confiable a los datos origen de cada formulario para su validación y enriquecimiento.
- Realizar informes de, o que contemplen, datos maestros y de referencia para el análisis y la toma de decisiones.

Para la aceleración y reutilización de soluciones eficaces, el desarrollo de las arquitecturas conceptuales y lógicas de datos se basa en arquitecturas de referencia que fueron descritas en el marco teórico (*Capítulo 3*) de este trabajo.

Cabe aclarar que la guía que regirá el desarrollo de esta arquitectura será TOGAF, y se utilizarán los siguientes marcos para el diseño del contenido de cada fase que propone el ADM o Método de Desarrollo de la Arquitectura:

- Del DAMA se tomarán las actividades, roles, servicios y principios para utilizarse en la *Fase de Arquitectura de Negocio*. Para la *Fase de Arquitectura de Datos* es posible considerar los diferentes enfoques arquitectónicos que provee este marco metodológico.
- Del MDM Institute se consideraran las actividades que se definen en para el desarrollo de un proyecto de gestión de datos maestros, así como la orientación de esta arquitectura hacia el diseño de servicios SOA. Algunas de las fases proporcionadas por el MDM Institute vienen implícitas en la *Iteración de Contexto del ADM*.
- Del EDW se tomara en cuenta las políticas para el desarrollo de procesos en la *Fase de Arquitectura de Negocio*, así como el uso de repositorios centrales y metadatos en la *Fase de Arquitectura de Datos*.
- Del esquema de Oracle se consideraran las actividades de consolidación, sincronización y limpieza de datos para el desarrollo de servicios en la *Arquitectura de Negocio*.
- Del esquema de IBM se retomara a la integración, sincronización, recuperación, actualización y distribución de datos maestros como parte de la definición de procesos y servicios en la *Arquitectura de Negocio*. También se tomaran en cuenta tanto los principios que proporciona.
- Del esquema de Microsoft se consideran las etapas que proporciona para la gestión de datos maestros, estas etapas se utilizaran en la definición e procesos en la *Fase de Arquitectura de Negocio*. También se retomara el *Estilo de Repositorio* para la orientación de esta arquitectura hacia un repositorio único y centralizado.
- Del esquema de Informática se retomaran los principios que provee hacia el diseño de procesos en la *Fase de Arquitectura de Negocio*.
- El MDM Madurity Model se tomará como marco de referencia para realizar la evaluación de madurez en la *Fase Preliminar*.

#### **4.1.1 Evaluación de madurez**

La evaluación de madurez de datos maestros y de referencia se realizará con base en el MDM Madurity Model. Se seleccionó este modelo principalmente por que se tuvo fácil acceso y su enfoque está orientado únicamente al dominio de los datos maestros.

El modelo define seis componentes, y cada uno de ellos contiene ciertas capacidades. Así mismo establece cinco niveles de madurez: inicial, reactivo, gestionado, proactivo y rendimiento estratégico.

Con base en las características que proporciona el modelo de madurez mencionado y con la intención de hacer más flexible la forma de evaluar la madurez para los datos maestros y de referencia en cualquier organización, se ha desarrollado en el presente trabajo un *Cuestionario de evaluación* para cada componente.

Cabe aclarar que las preguntas de los cuestionarios se desarrollaron con la información del MDM Madurity Model para presentarlas en un Test con el fin de obtener los niveles de madurez de una manera más sencilla. A continuación se presentan los seis cuestionarios de evaluación que se proponen para evaluar en cualquier organización las capacidades de cada componente.

##### ***Cuestionario 1: Arquitectura***

## Cuestionario de evaluación: Arquitectura

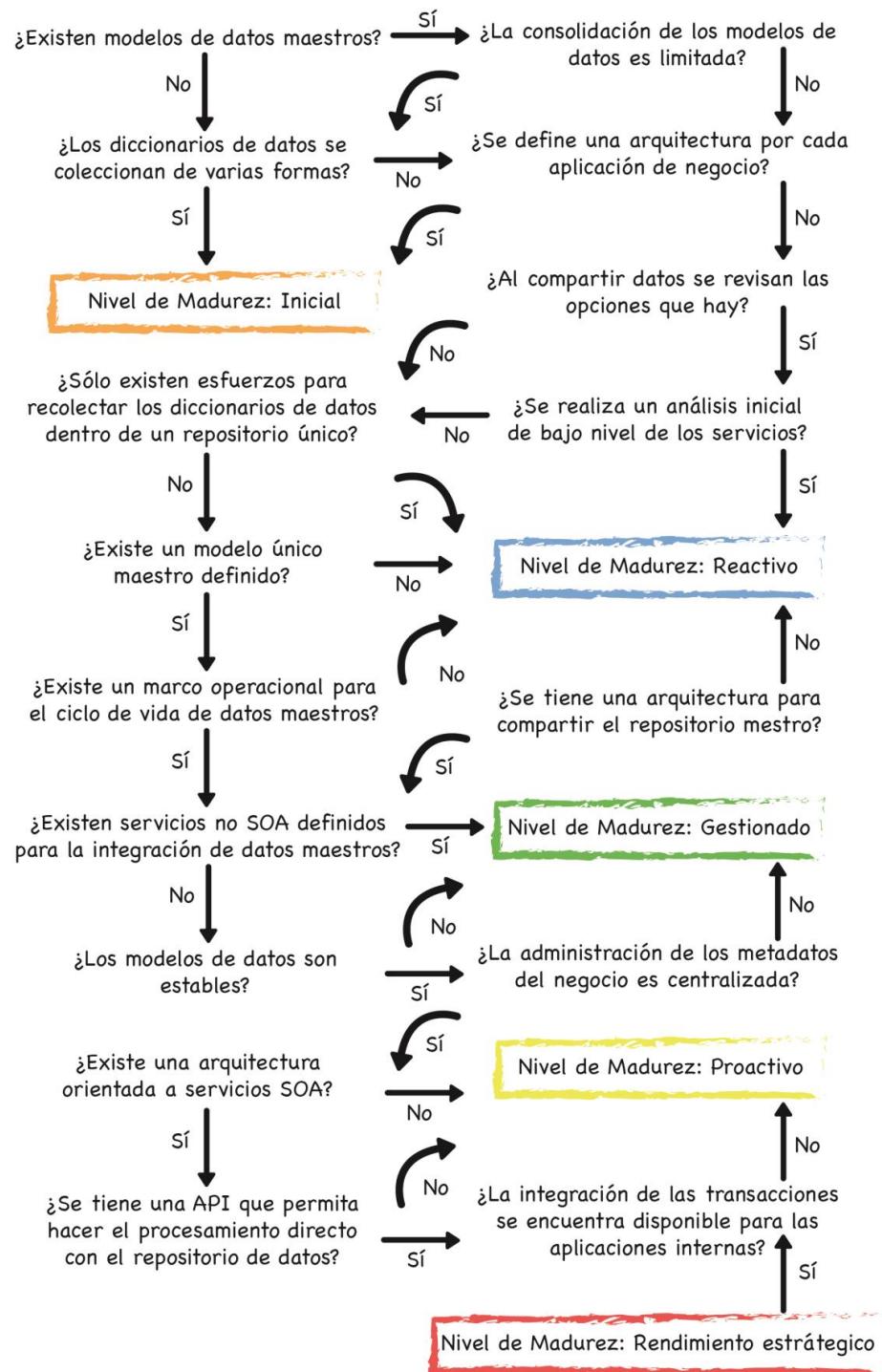


Figura 4.3 Cuestionario de Arquitectura. Adaptado de Loshin, 2009.

## Cuestionario 2: Gobierno

### Cuestionario de evaluación: Gobierno

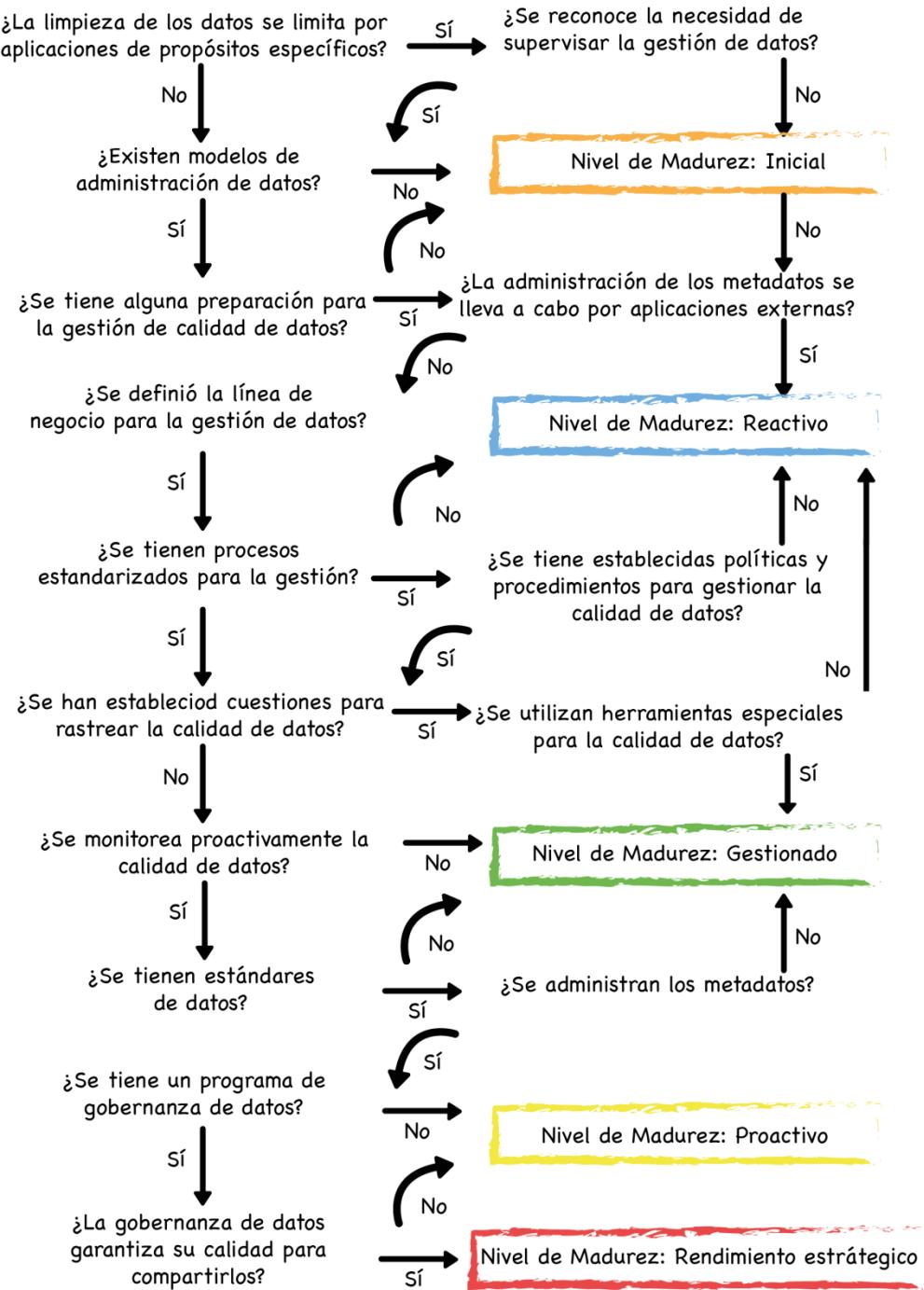


Figura 4.4 Cuestionario de Gobierno. Adaptado de Loshin, 2009.

### Cuestionario 3: Administración

#### Cuestionario de evaluación: Administración

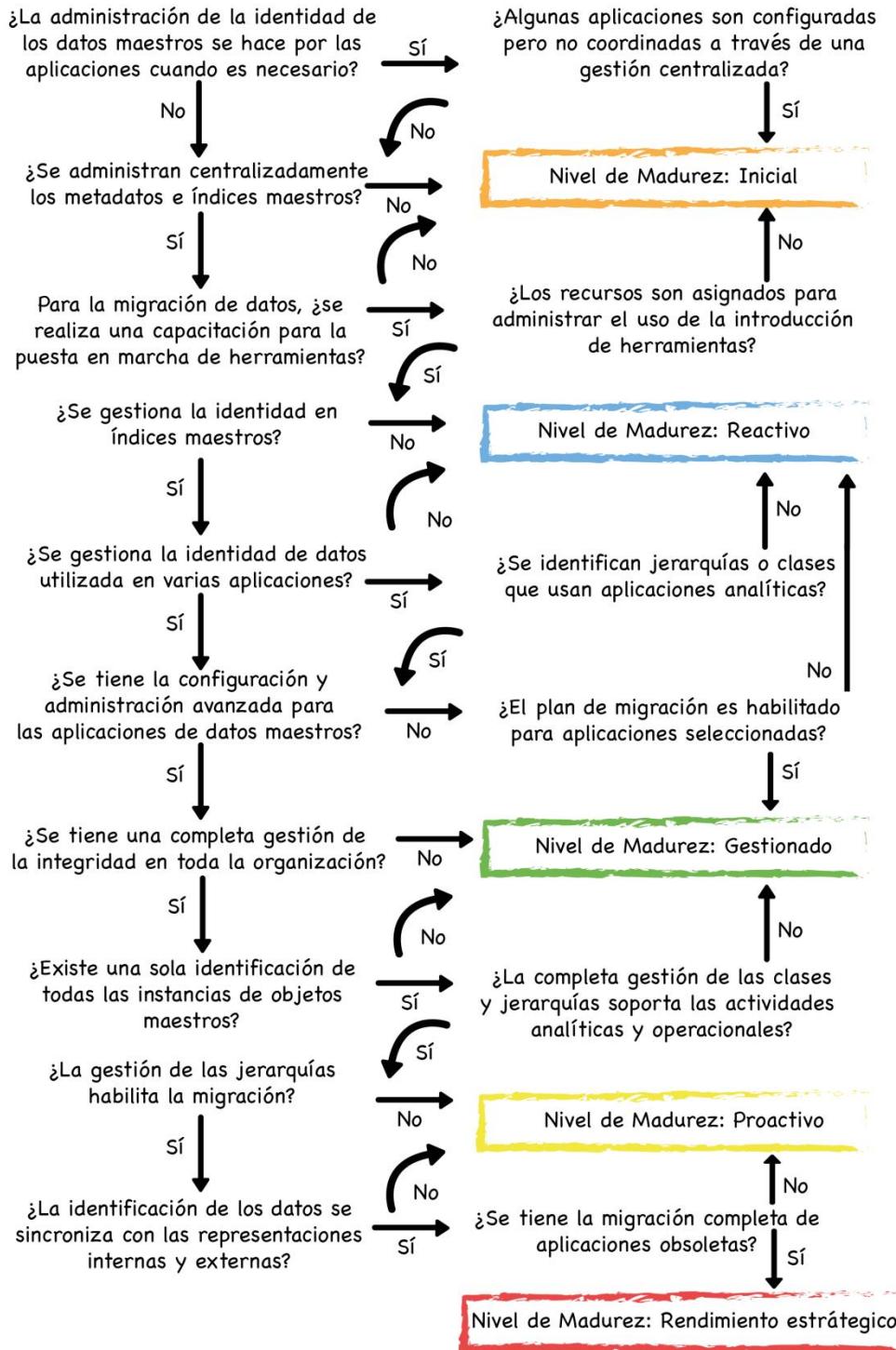
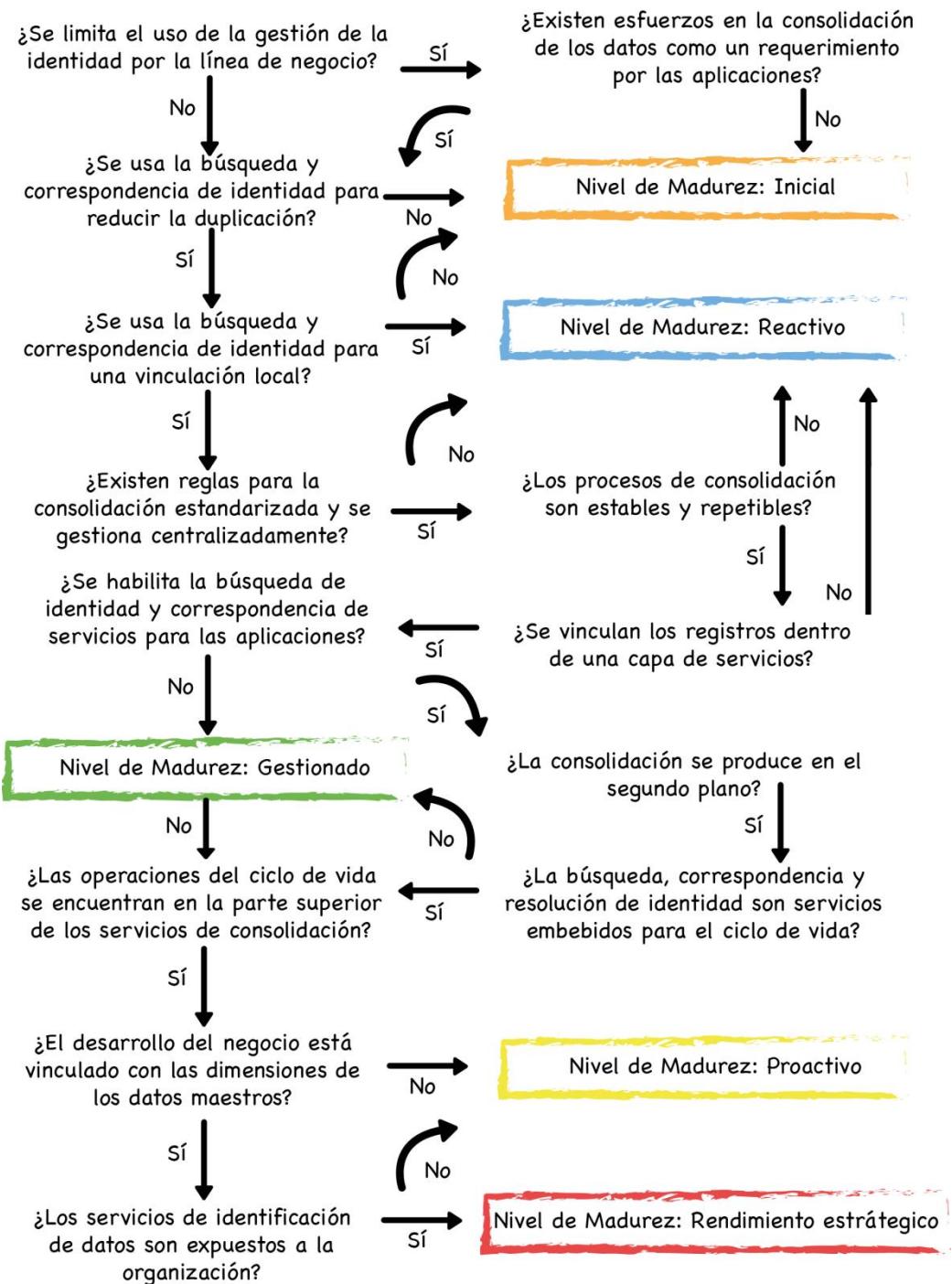


Figura 4.5 Cuestionario de Administración. Adaptado de Loshin, 2009.

## *Cuestionario 4: Identificación*

## Cuestionario de evaluación: Identificación



**Figura 4.6 Cuestionario de Identificación.** Adaptado de Loshin, 2009.

Cuestionario 5: Integración

## Cuestionario de evaluación: Integración

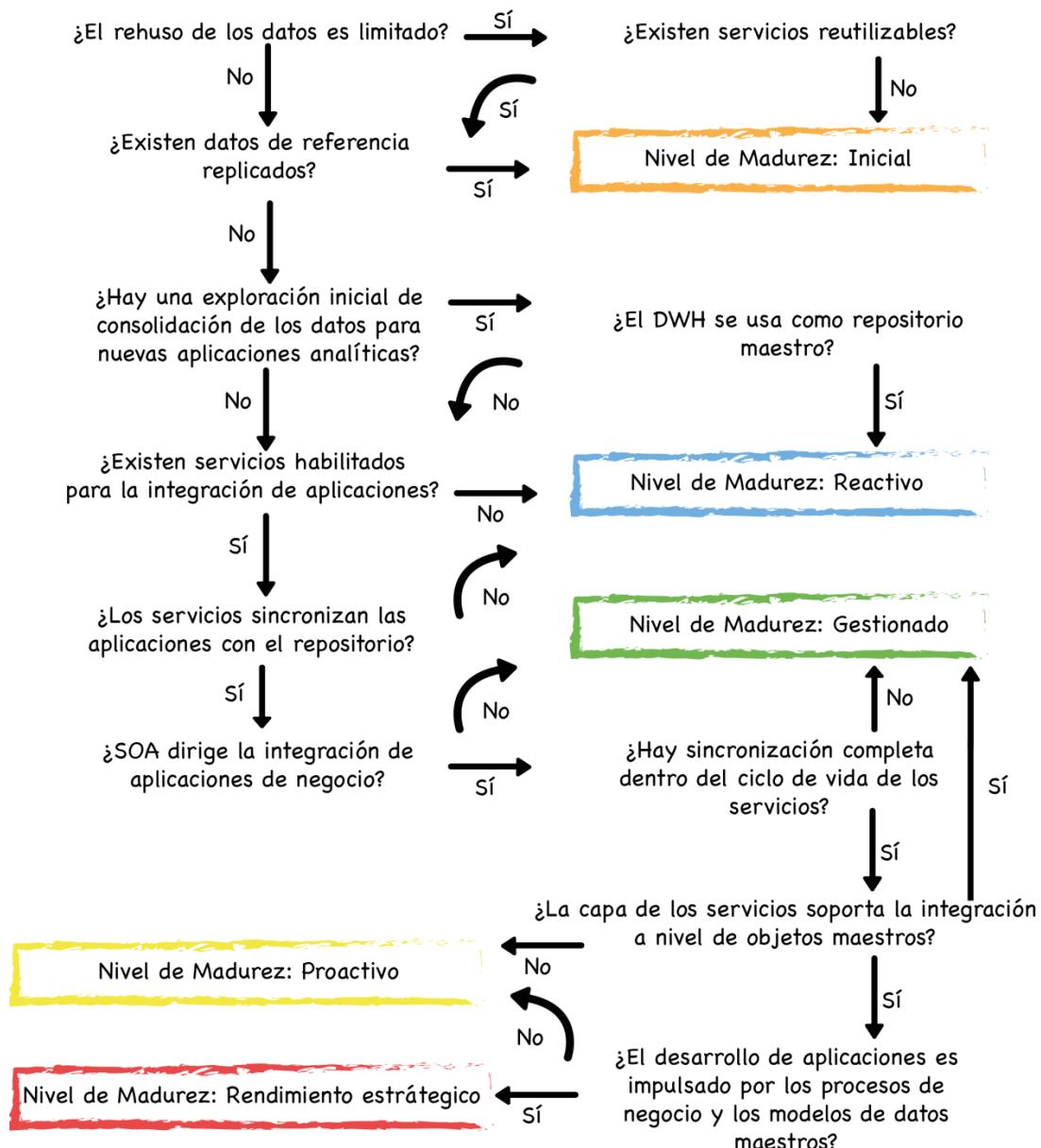


Figura 4.7 Cuestionario de Integración. Adaptado de Loshin, 2009.

## Cuestionario 6: Administración de procesos de negocio

### Cuestionario de evaluación: Administración de procesos de negocio

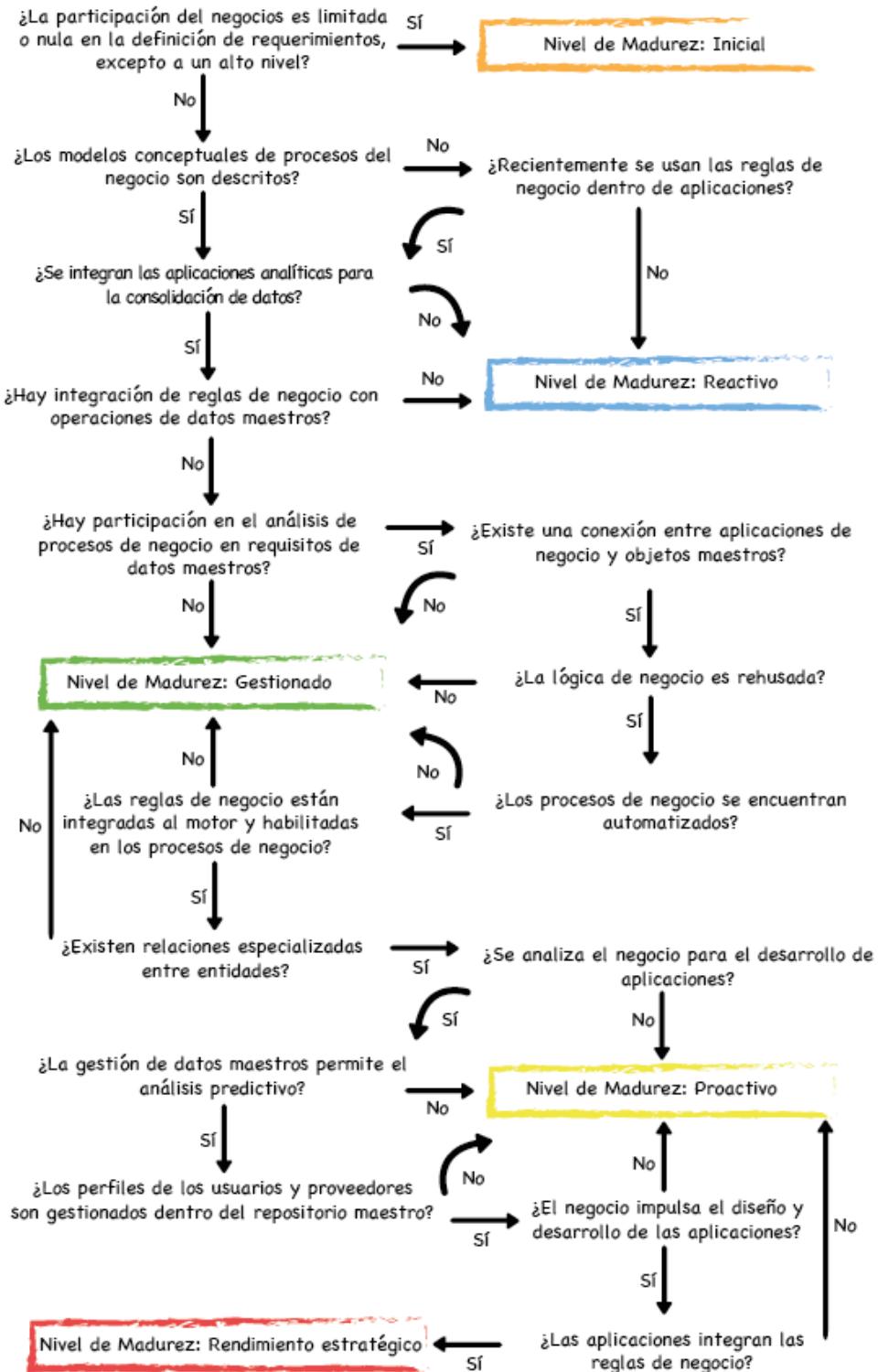


Figura 4.8 Cuestionario de admón. de procesos de negocio. Adaptado de Loshin, 2009

## 4.2 Fase A: Visión de la arquitectura

En esta fase se desarrolla el alcance, la misión, la visión, las expectativas y la estrategia.

Los interesados del desarrollo de este trabajo son aquellas personas que requieran implementar la gestión de datos maestros y de referencia, siendo que sus principales preocupaciones radican en el uso de datos duplicados que afectan la calidad de la información que se proporciona o analiza para la toma de decisiones.

La misión de esta arquitectura es: Contar con información veraz, completa y homologada de datos maestros y de referencia para administrar sus cambios.

La visión de esta arquitectura es: Gestionar los datos maestros y de referencia enfocada al diseño de una arquitectura empresarial para contar con datos íntegros, estandarizados y medibles.

Las expectativas de este trabajo son:

- Compartir la importancia de los datos maestros y de referencia.
- Proporcionar una definición clara para los datos maestros y de referencia.
- Proporcionar un esquema que muestre cómo son utilizadas las entidades maestras y de referencia por elementos del negocio.
- Desarrollar una arquitectura funcional del dominio de negocio y datos.
- Comunicar dentro de la dirección la solución para la gestión de las entidades maestras y de referencia.

### 4.2.1 Estrategia

Las directrices de la estrategia se basan en el modelo de madurez MDM Madurity Model. Por lo que los componentes a considerar son seis: Arquitectura, Gobierno, Administración, Identificación, Integración y Administración de procesos de negocio. Este modelo contiene capacidades que debe cumplir la arquitectura objetivo. Con base en los requerimientos de estas capacidades se detectaron iniciativas.

La Figura 4.9 muestra que las iniciativas de la estrategia se obtuvieron de las capacidades de cada componente.

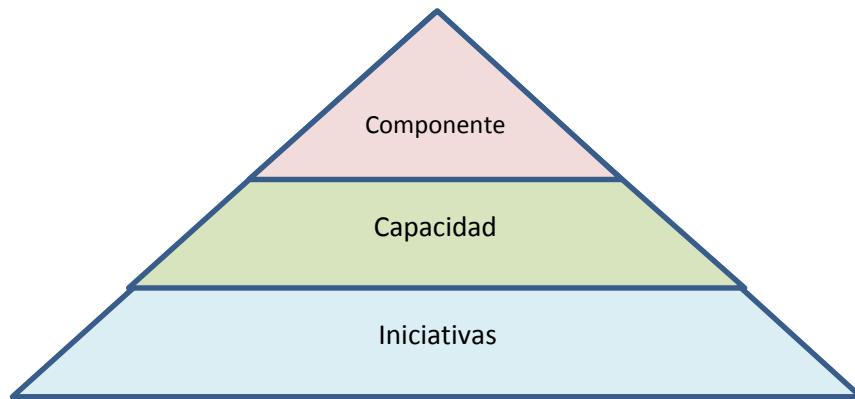


Figura 4.9 Niveles de abstracción para la obtención de Iniciativas

Para el componente de *Arquitectura* las iniciativas que soportan sus capacidades se ilustran en la Figura 4.10.

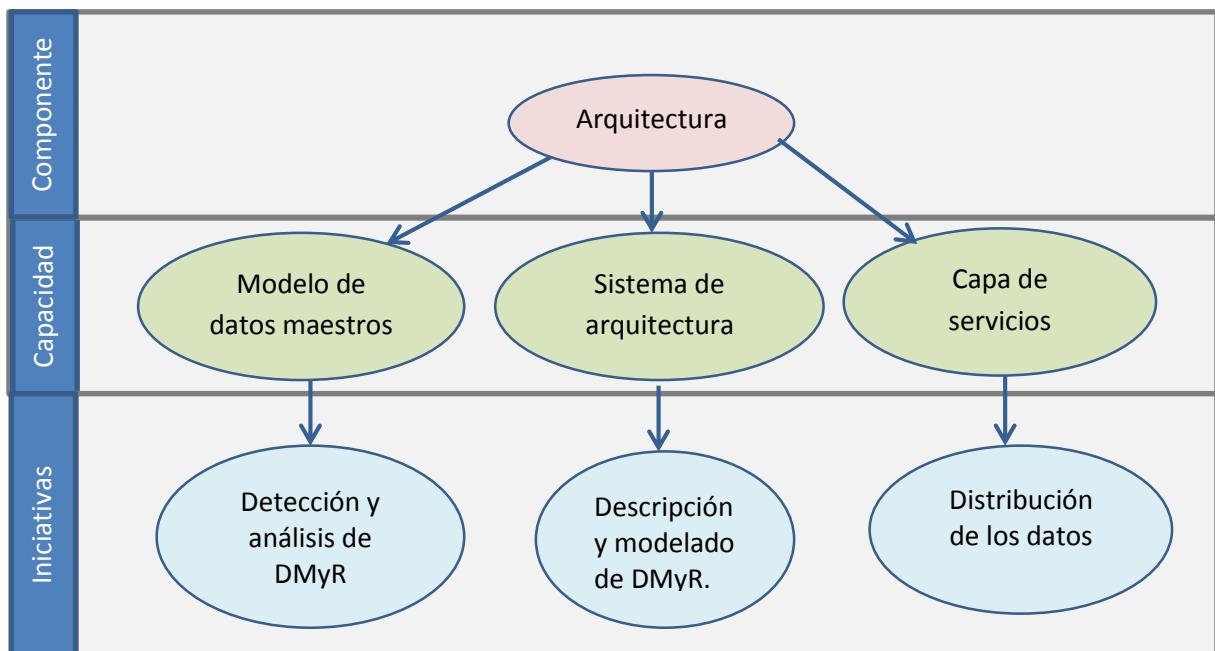


Figura 4.10 Iniciativas que soportan las capacidades de la Arquitectura

Para el componente de *Gobierno* las iniciativas que soportan sus capacidades se ilustran en la Figura 4.11.

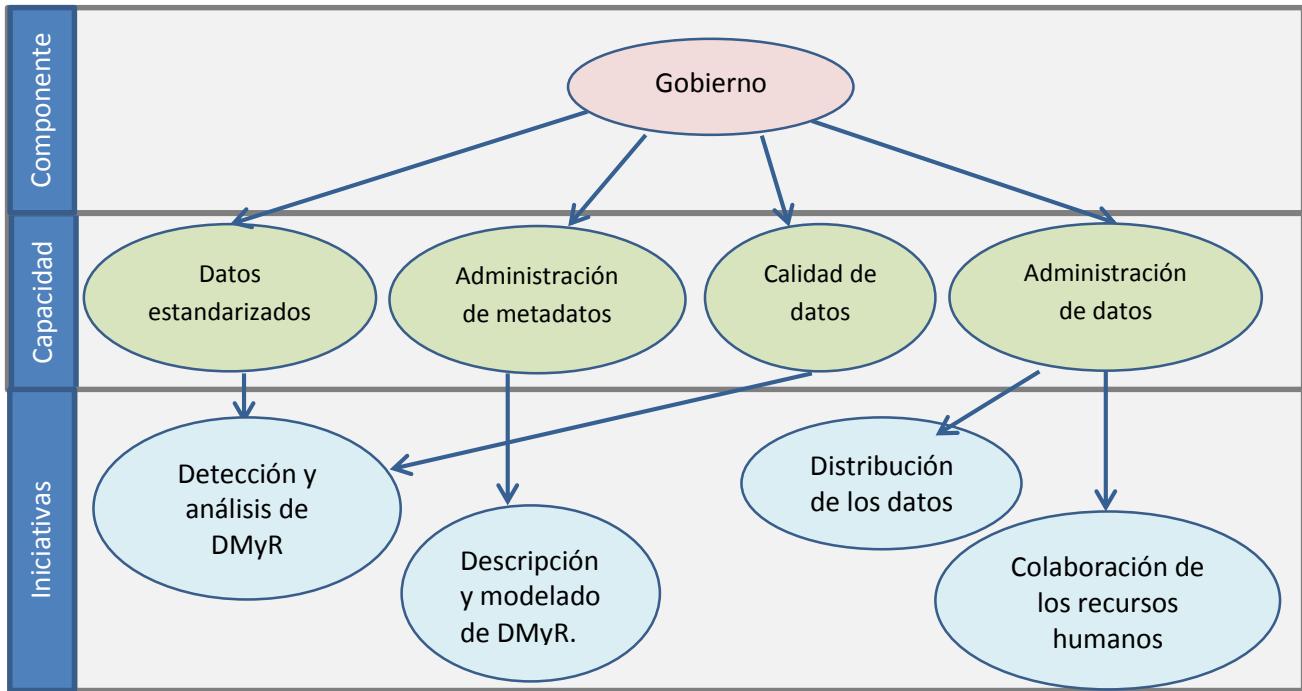


Figura 4.11 Iniciativas que soportan las capacidades del Gobierno

Para el componente de *Administración* las iniciativas que soportan sus capacidades se ilustran en la Figura 4.12.

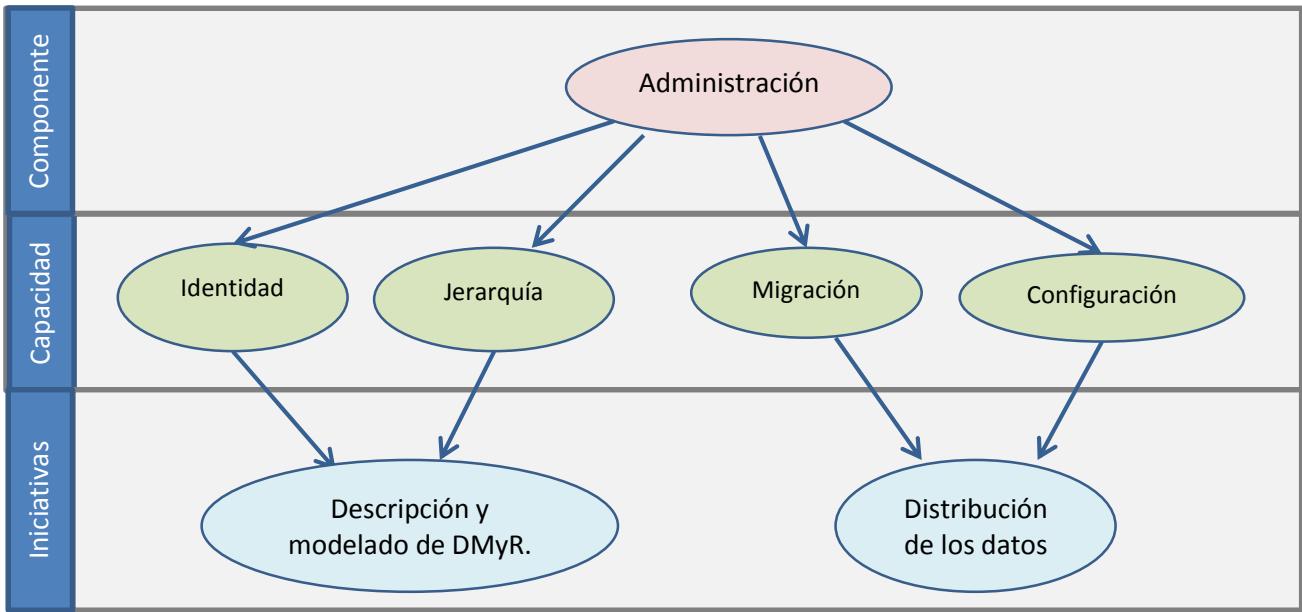


Figura 4.12 Iniciativas que soporta las capacidades de la administración

Para el componente de *Identificación* las iniciativas que soportan sus capacidades se ilustran en la Figura 4.13.

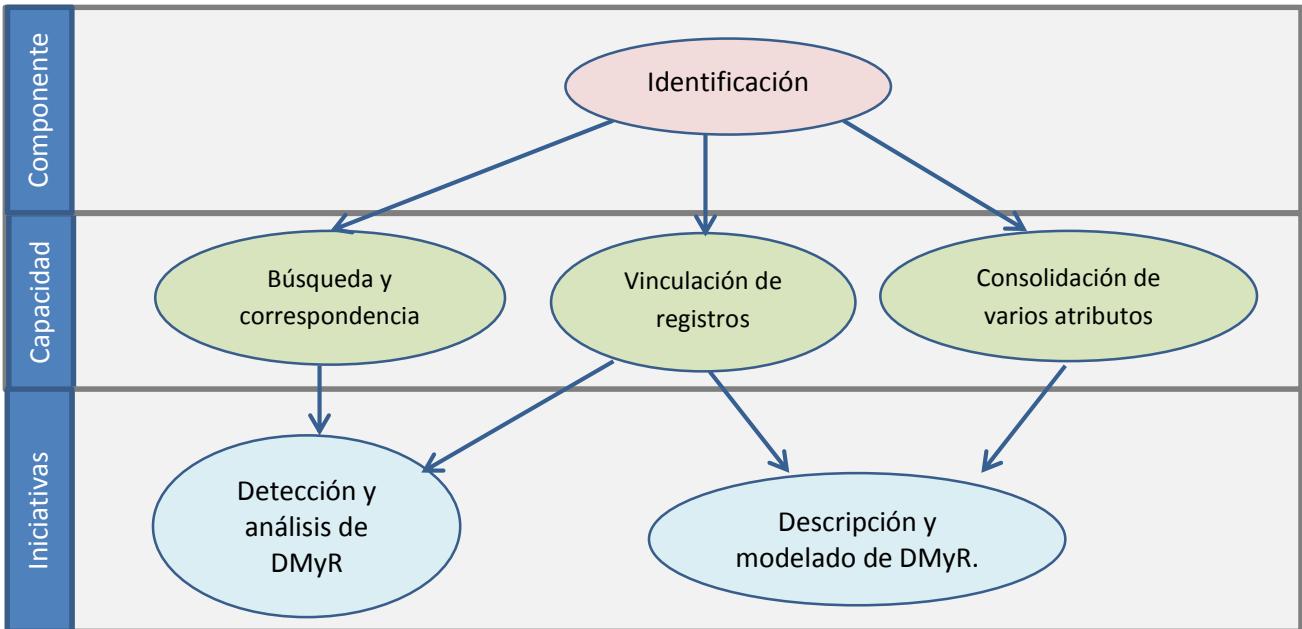


Figura 4.13 Iniciativas que soportan las capacidades de la Identificación

Para el componente de *Integración* las iniciativas que soportan sus capacidades se ilustran en la Figura 4.14.

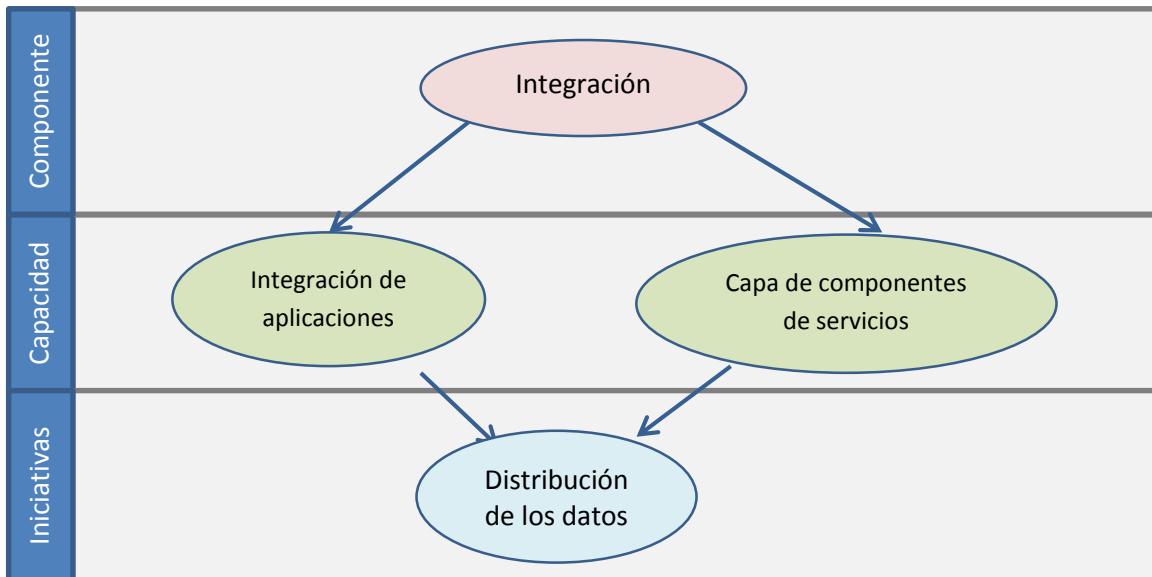


Figura 4.14 Iniciativas que soportan las capacidades de la Integración

Para el componente de *Administración de procesos de negocio* las iniciativas que soportan sus capacidades se ilustran en la Figura 4.15.

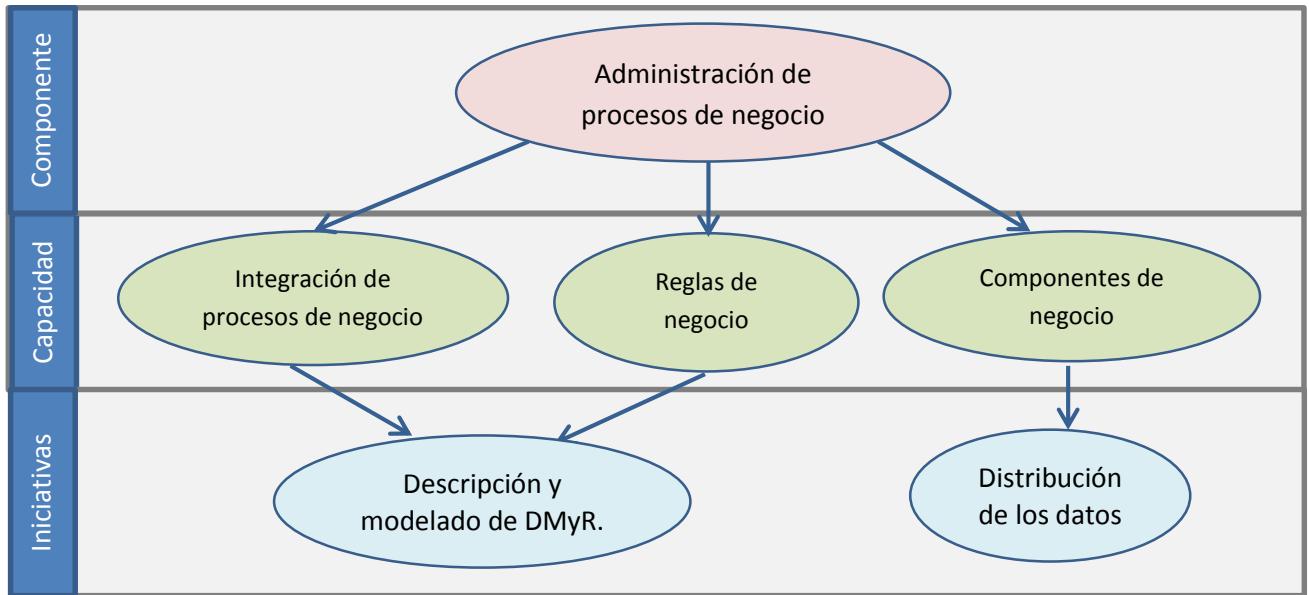


Figura 4.15 Iniciativas que soportan las capacidades de Admón. de procesos de negocio

Por lo tanto la estrategia se compone de 4 principales iniciativas, esta se ilustra en la Figura 4.16.

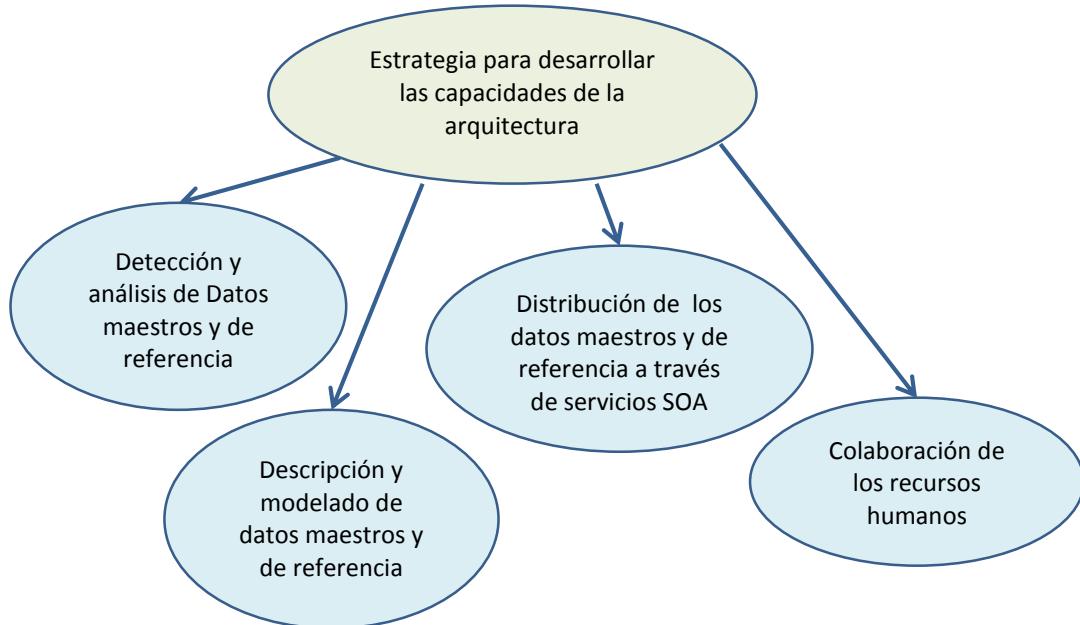


Figura 4.16 Estrategia para desarrollar las capacidades

En las siguientes tablas se presenta la descripción y objetivo para cada una de las iniciativas.

<b>Objetivo</b>
Recopilar y consolidar los datos maestros y de referencia por medio de artefactos estandarizados.
<b>Descripción</b>
Esta iniciativa comprende los proyectos necesarios para recopilar y depurar diferencias sintácticas de los datos maestros. Esto se realizará a través de métodos de búsqueda, correspondencia y limpieza de datos. Esta iniciativa será la base para el desarrollo de la siguiente.

**Tabla 4.1 Iniciativa 1 “Detección y análisis de datos maestros y de referencia”**

<b>Objetivo</b>
Establecer el uso de los datos maestros y de referencia dentro de los procesos de negocio, así como definir una estructura maestra para su gestión central.
<b>Descripción</b>
Esta iniciativa comprende los proyectos necesarios para desarrollar un modelo conceptual de información por cada proceso de negocio que utilice a los datos maestros y de referencia. Esta iniciativa comprende los proyectos necesarios para generar una estructura estandarizada para su gestión central partiendo de los datos que fueron depurados en la iniciativa anterior. La iniciativa también contempla la definición de los subconjuntos de datos que se tienen que replicar a diferentes procesos.

**Tabla 4.2 Iniciativa 2 “Descripción y modelado de datos maestros y de referencia”**

<b>Objetivo</b>
Compartir datos a través de la vinculación de servicios SOA con el repositorio central.
<b>Descripción</b>
Esta iniciativa comprende los proyectos necesarios para establecer la relación entre las entidades de datos y los servicios requeridos para su distribución a través de la vinculación con el repositorio maestro.

**Tabla 4.3 Iniciativa 3 “Distribución de los datos maestros y de referencia a través de servicios SOA”**

<b>Objetivo</b>
Comunicar y capacitar al recurso humano de la organización para poner en marcha la arquitectura definida.
<b>Descripción</b>
Esta iniciativa comprende los proyectos necesarios para llevar a cabo las responsabilidades definidas en los roles definidos en las fases de la arquitectura. Las definiciones, principios, procesos y servicios de negocio se deben comunicar a través de toda la organización. Para la ejecución adecuada de dichos procesos el personal deberá estar capacitado.

**Tabla 4.4 Iniciativa 4 “Colaboración de los recursos humanos”**

### **4.3 Fase B: Arquitectura de negocio**

Esta arquitectura describe las directrices que necesita la dirección para operar y lograr los objetivos de negocio. Sus principales elementos son procesos, funciones, servicios, roles,

principios y políticas. A continuación se describe cada elemento para gestionar las entidades maestras y de referencia.

#### 4.3.1 Principios de datos

TOGAF define una plantilla estándar para la descripción de los principios de cada dominio (negocio, datos, aplicaciones y tecnología). Esta plantilla promueve que el principio sea entendible y aceptado. Los atributos de la plantilla son:

- Nombre
- Declaración
- Razón fundamental
- Implicaciones

El nombre debe representar la esencia de la regla, así como debe ser fácil de recordar. La especificación de las plataformas tecnológicas no debe mencionarse en el nombre o declaración. La declaración debe comunicar sin ambigüedades la regla fundamental.

La razón fundamental debe usar la terminología del negocio enfocándose en los beneficios de la adopción de este principio. Este atributo también puede describir la relación con otro principio. Finalmente las implicaciones deben enfocarse en los requerimientos tanto de negocio como de TI para llevar a cabo el principio.

Con base en la plantilla que propone TOGAF para los principios de una arquitectura de datos, se describirán los principios de entidades maestras y de referencia.

IBM nos proporciona el siguiente principio “*Disponer de los datos maestros como un activo estratégico para su uso*”. Por su parte Informática proporciona lo siguiente “*Administrar diligentemente los datos maestros como otros activos estratégicos de la empresa*”. La adaptación de este principio es:

#### Principio 1

##### Las EMyR son un activo estratégico

- Declaración: Las entidades maestras y de referencia son usadas como un activo estratégico para la dirección.
- Razón fundamental: Las entidades maestras y de referencia son un activo de información que proporcionan valor a las organizaciones porque representan valiosas fuentes de conocimiento con las cuales se pueden tomar decisiones. Estos datos se deben gestionar cuidadosamente para saber obtenerlos donde y cuando se necesiten.
- Implicaciones:
  - Las organizaciones debe tener una estrategia impulsada por el negocio para guiar la gestión de entidades maestras y de referencia.

- Los datos maestros y de referencia deben ser actuales, oportunos, confiables y explicables.
- Los administradores de datos maestros deben gestionarlos al nivel de la organización.

El DAMA define que “*Los valores de oro de las EMyR representan los esfuerzos realizados para obtener los datos más exactos, actuales y relevantes*”, por otro lado uno de los principios de IBM es el siguiente: “*Proporcionar una fuente autorizada de datos maestros*”. Sobre este mismo tema Informatica proporciona como uno de sus principios “*Establecer una versión única*”. La adaptación es:

## **Principio 2**

### **Las EMyR contienen valores de oro**

- Declaración: Los valores de los registros de las entidades maestras y de referencias son únicos.
- Razón fundamental: Los datos maestros son el recurso de mayor valor, son datos únicos y sirven de referencia para toda la dirección. Parte de la gestión de las entidades maestras y de referencia identifica los valores de oro; los más actuales y verídicos. Las entidades maestras contienen valores de oro porque sus definiciones se estandarizan y sus registros se unifican a través de toda las organizaciones.
- Implicaciones:
  - Se requiere tener una versión unificada de los datos maestros y de referencia. Estos datos son utilizados por sistemas transaccionales y de inteligencia de negocios.
  - Los registros consolidados de las entidades maestras y de referencia deberán poblar y respetar la organización centralizada del repositorio maestro.
  - Los valores de oro se replicarán a través de servicios del repositorio central a otros repositorios.
  - Para mantener los valores de oro se aplicarán reglas de coincidencia a los nuevos registros, esto con el fin de asegurar su consistencia.

El DAMA define el siguiente principio “*Los datos maestros y de referencia son compartidos y pertenecen a la organización*”. La adaptación es la siguiente:

## **Principio 3**

### **Las EMyR pertenecen a las organizaciones y son compartidas**

- Declaración: Los usuarios tiene acceso a las entidades maestras y de referencia para llevar a cabo sus funciones.
- Razón fundamental: Las entidades maestras y de referencia pertenecen a las organizaciones, no a una persona, aplicación o servicio en particular. Al ser compartidas, las definiciones de las entidades maestras y de referencia son comprensibles y están a la disposición de los usuarios.
- Implicaciones:
  - Entender la relación entre valor, uso y acceso a los datos.
  - El intercambio de datos requiere un cambio cultural significativo.
  - Los usuarios requieren acceso al subconjunto de entidades maestras y de referencia que necesiten para desempeñar sus funciones.
  - El uso e intercambio de entidades maestras y de referencia se realizará por medio de procesos, funciones y servicios de negocio.

Otro de los principios de datos que el DAMA plantea es el siguiente “*La administración de los datos maestros y de referencia es un programa para mejorar la calidad de los datos*”. Similar a este Informatica presenta “*Comunicar que la calidad de los datos depende de la excelencia de los procesos*”. La adaptación es la siguiente:

#### **Principio 4**

##### **La gestión de las EMyR mejora la calidad de la información**

- Declaración: La administración de datos maestros es un programa que contiene la mejora de la calidad de la información.
- Razón fundamental: La gestión de entidades maestras y de referencia es un programa de mejora, sus objetivos no pueden ser alcanzados por un solo proyecto. Los usuarios de las entidades maestras y de referencia, a través de reglas de negocio, especifican las expectativas de la calidad de estas entidades.
- Implicaciones:
  - Los procesos deben garantizar la calidad y la integridad de las entidades maestras y de referencia.
  - Se requiere aprovechar la gobernanza de las entidades maestras y de referencia para el control y rendimiento de su calidad.
  - Se informa a los administradores del negocio y datos los niveles de calidad que tienen las entidades maestras y de referencia.

Uno de los principios del DAMA es “*Proporcionar flexibilidad para adaptarse a cambios*”. La adaptación es:

#### **Principio 5**

### **La arquitectura de EMyR es flexible a cambios**

- Declaración: La arquitectura para la gestión de datos maestros y de referencia debe proporcionar flexibilidad para adaptarse a nuevos cambios.
- Razón fundamental: La mejora de los elementos de la arquitectura a nivel, negocio, datos y aplicaciones se dará en la medida que sea una solución de toda las organizaciones: negocio y TI.
- Implicaciones:
  - La tecnología no debe mandar la definición de la arquitectura. Esta arquitectura debe cubrir el uso de la tecnología.
  - La arquitectura de entidades maestras y de referencia deben proveer una solución basada en bloques de construcción.
  - Una solución que no se tiene acceso no muestra el valor del negocio.

El DAMA proporciona el siguiente principio “*Es importante solicitar, comunicar y aprobar los cambios en los datos maestros y de referencia*”. Informatica nos comparte el siguiente principio “*Establecer el proceso de gestión del cambio*”. La adaptación es la siguiente:

### **Principio 6**

#### **Los cambios en las EMyR se solicitan, comunican y aprueban**

- Declaración: La gestión de los datos maestros y de referencia está basado en un proceso que solicita, comunica y aprueba los cambios que surjan.
- Razón fundamental: Se deben solicitar, comunicar y aprobar los cambios que surjan en las entidades maestras y de referencia antes que se implementen. Si los cambios no se solicitan y comunican antes de aprobarse es posible que se tenga que hacer un nuevo cambio.
- Implicaciones:
  - La naturaleza de los procesos de la arquitectura de entidades maestras y de referencia requieren definir el flujo de las actividades y los participantes a interactuar.
  - La arquitectura de entidades maestras y de referencia en las organizaciones debe establecer un vocabulario común del negocio para comunicar sus necesidades a los responsables del negocio.
  - Los expertos de datos y del negocio deben participar en la aprobación de cualquier cambio.

El DAMA define “Los administradores de las EMyR son responsables de controlar sus valores”, IBM proporciona “Preservar la integridad y seguridad de los datos”, Informatica “El negocio posee a los datos maestros y es responsable de su integridad”.

### **Principio 7**

### **Los administradores de las EMyR son responsables de sus valores e integridad**

- Declaración: La responsabilidad de la gestión de entidades maestras y de referencia se comparte con los administradores del negocio y custodios expertos.
- Razón fundamental: Las entidades maestras y de referencia son protegidas del uso y divulgación no autorizada. La toma de decisiones compartida es el sello distintivo de su gestión. Por lo tanto, los administradores del negocio y expertos de datos son las autoridades responsables de controlar los valores de las entidades maestras y de referencia.
- Implicaciones:
  - La creación o actualización de los registros de las entidades maestras y de referencia deben ser verificadas por más de un rol.
  - Las decisiones tomadas se deben comunicar a toda la dirección.

#### **4.3.2 Procesos**

Para la especificación de los procesos que gestionaran los datos maestros y de referencia, del DAMA se consideraron las actividades de administrar cambios, replicar y distribuir datos, así como los entregables como requerimientos, modelos, servicios de limpieza de datos, métricas y reportes de calidad.

Del esquema de Oracle se tomaron cinco características que proporciona en su enfoque orientado a procesos y que debe cumplir la arquitectura de este trabajo: elaborar perfiles de datos, consolidar los datos en un repositorio central, gobernar los datos de acuerdo a las reglas de negocio, compartir los datos sincronizándolos con procesos de negocio y aplicaciones dependientes.

Del modelo comercial de IBM se tomaron las siguientes características:

- Proporcionar un proceso de autorización para crear, definir y sincronizar los datos
- Proporcionar la recuperación, actualización, distribución y distintos usos de los datos
- Distribuir los datos a través de procesos y servicios
- Soportar el uso analítico de los datos
- Soportar varios usuarios

Los procesos de la administración de datos maestros son bidireccionales, es decir deben garantizar la sincronización de los datos entre el repositorio único y los sistemas origen y destino asociados. En base a estas características se definen cuatro procesos que debe tener en esta arquitectura.

Los procesos que interactúan con los participantes internos a las organizaciones son:

- A. Proceso de creación y mantenimiento de las entidades
- B. Proceso de calidad de entidades

Los procesos que tienen interacción tanto con los participantes externos como internos a las organizaciones son:

- C. Proceso de actualización de registros
- D. Proceso de intercambio de registros

A continuación se explica el objetivo y la secuencia de actividades que contiene cada uno de ellos.

#### *Creación y mantenimiento de las entidades*

Este proceso proporciona las directrices para generar o modificar características de las entidades maestras y de referencia. Este proceso no se enfoca en la gestión de cambios del contenido de las entidades, estas funciones son parte del proceso de actualización de registros.

A nivel de entidad se puede realizar lo siguiente:

- Definir la estructura y demás características de nuevas entidades maestras y de referencia.
- Modificar entidades maestras y de referencia ya definidas.
- Eliminar entidades maestras y de referencia

Las entradas de este proceso son:

- I. Requerimientos
- II. Entidades de datos
- III. Reglas de correspondencia y métodos de búsqueda

Las salidas de este proceso son:

- I. Nuevas entidades o entidades modificadas
- II. Respuesta al requerimiento

Las etapas de este proceso son:

- ✓ Recibir requerimiento
- ✓ Confirmar recepción de requerimiento
- ✓ Revisar y validar la procedencia del requerimiento
- ✓ Atender requerimiento de acuerdo a su naturaleza:
  - Para actualizar entidades
- ✓ Verificar la existencia de la entidad. Si la entidad no existe notificar rechazo de la actualización de la entidad
- ✓ Validar reglas, analizar resultados y evaluar la entidad existente
- ✓ Tomar decisión; aprobar o rechazar autorización. Si se rechaza el requerimiento notificar. Si la actualización de la entidad se aprueba:

- ✓ Notificar decisión a interesados
- ✓ Actualizar entidad
- ✓ Almacenar actualización de entidad
- ✓ Respaldar histórico de actualización
- ✓ Replicar actualización
- ✓ Notificar cambio
  - Para crear entidades
    - Verificar la no existencia de la entidad. Si la entidad existe entonces de notificar el rechazo del requerimiento. Si la existencia no existe:
      - Analizar y evaluar impactos de creación de entidad
      - Tomar decisión. Si se rechaza la creación se deberá notificar. Si se acepta
        - Notificar la decisión a interesados
        - Crear una nueva entidad
        - Almacenar la entidad creada
        - Respaldar el histórico de la creación. Esta acción se realiza para una futura recuperación.
        - Notificar la creación de la entidad

En la Figura 4.17 se muestra una vista del flujo de este proceso realizado en Archimate 2.0.

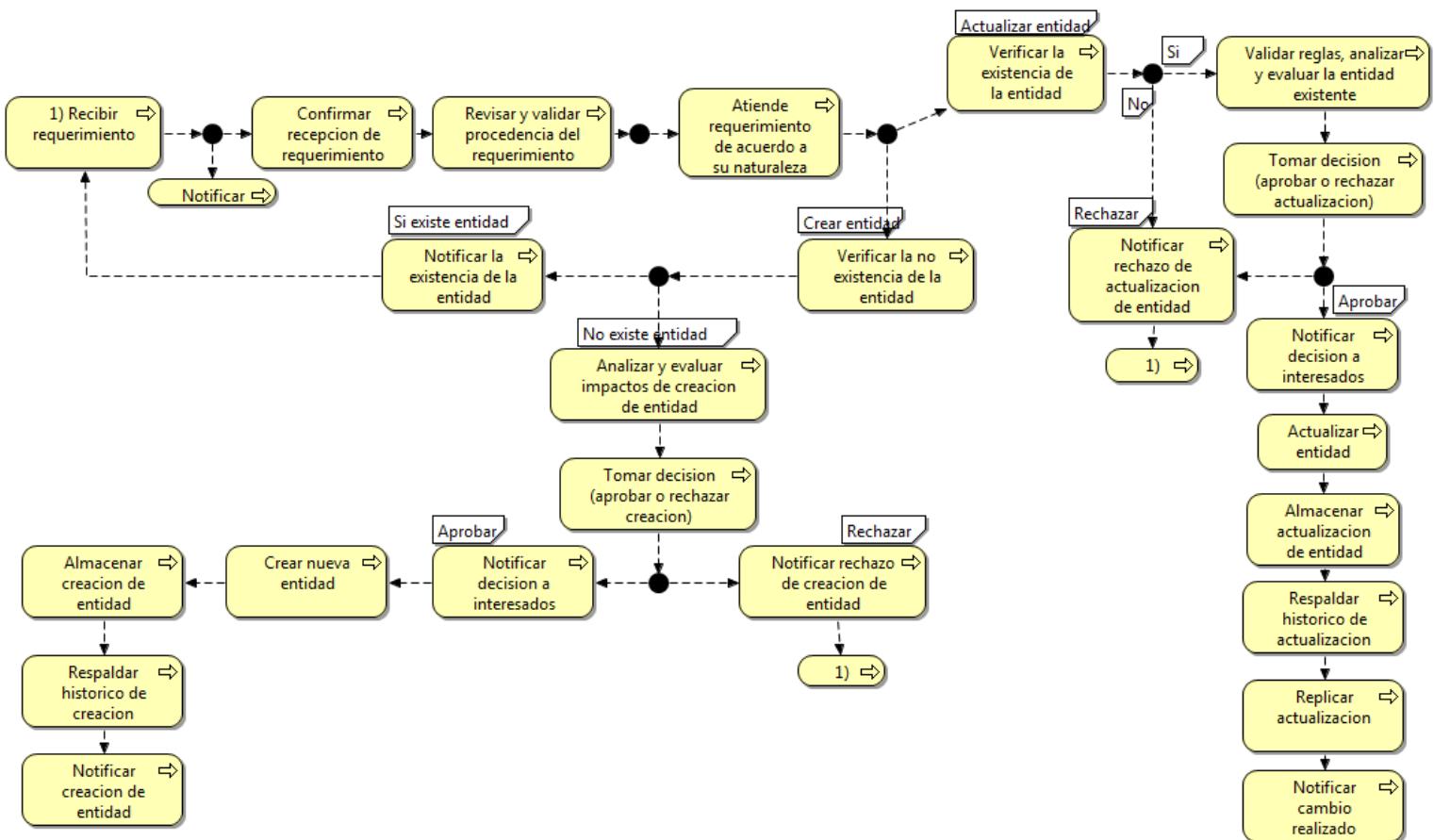


Figura 4.17 Proceso de creación y mantenimiento de las entidades

### *Calidad de entidades*

El objetivo de este proceso es medir y controlar la calidad de las entidades maestras y de referencia. Este proceso forma parte de una evaluación periódica.

A través de un servicio se mostrarán los parámetros de calidad a considerar.

Las principales entradas son los propios datos de las entidades, indicadores y métodos definidos para la evaluación de la calidad.

Las entradas de este proceso son:

- I. Datos
- II. Indicadores
- III. Métodos para evaluar la calidad

Las salidas de este proceso son:

- I. Diagnósticos

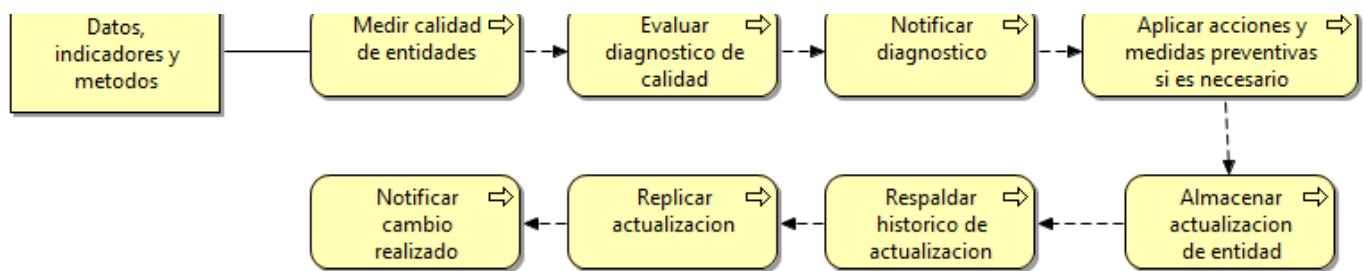
## II. Datos depurados

Las etapas de este proceso son:

- ✓ Medir la calidad de las entidades y obtener un diagnóstico
- ✓ Generar y analizar el diagnóstico
- ✓ Notificar diagnóstico a responsables del negocio y medidas preventivas
- ✓ Aplicar acciones y medidas preventivas las entidades que lo ameriten
- ✓ Almacenar actualizaciones de entidades
- ✓ Respaldar histórico para una futura recuperación
- ✓ Replicar actualización
- ✓ Notificar cambio realizado

En la figura 4.18 se muestra una vista del flujo de este proceso realizada en Archimate 2.0

Figura 4.18 Proceso de calidad de entidades



### Actualización de registros

El objetivo de este proceso es gestionar los cambios que surjan en cuanto a contenido de las entidades maestras y de referencia. En este proceso se puede:

- Dar de alta un registro maestro y de referencia
- Dar de baja un registro maestro y de referencia
- Modificar un registro maestro y de referencia

Las entradas de este proceso son:

- I. Requerimientos
- II. Entidades de datos
- III. Reglas de correspondencia y métodos de búsqueda

Las salidas de este proceso son:

- I. Nuevas registros o registros modificados
- II. Respuesta al requerimiento

Las etapas de este proceso son:

- ✓ Recibir requerimiento

- ✓ Confirmar recepción de requerimiento
- ✓ Revisar y validar la procedencia del requerimiento. Si no se valida la procedencia del requerimiento notificar rechazo.
- ✓ Atender requerimiento de acuerdo a su naturaleza:
  - Para actualizar registros
- ✓ Validar reglas de correspondencia entre registros
- ✓ Analizar resultados y evaluar impactos de modificación.
- ✓ Tomar decisión; aprobar o rechazar autorización. Si se rechaza el requerimiento notificar. Si la actualización del registro se aprueba:
  - ✓ Notificar decisión a interesados
  - ✓ Actualizar registro
  - ✓ Almacenar actualización del cambio del registro.
  - ✓ Respaldar histórico de actualización
  - ✓ Replicar actualización
  - ✓ Notificar cambio
    - Para dar de alta registros
      - Verificar la no existencia del registro en una entidad determinada. Si el registro existe entonces notificar el rechazo del requerimiento. Si no existe el registro:
        - Analizar y evaluar impactos de alta de registros
        - Tomar decisión. Si se rechaza la creación se deberá notificar. Si se acepta:
          - Notificar la decisión a interesados
          - Dar de alta un nuevo registro
          - Almacenar un nuevo registro
          - Respaldar el histórico de la creación. Esta acción se realiza para una futura recuperación.
          - Notificar la creación de la entidad

A continuación se muestra una vista del flujo de este proceso realizada en Archimate 2.0

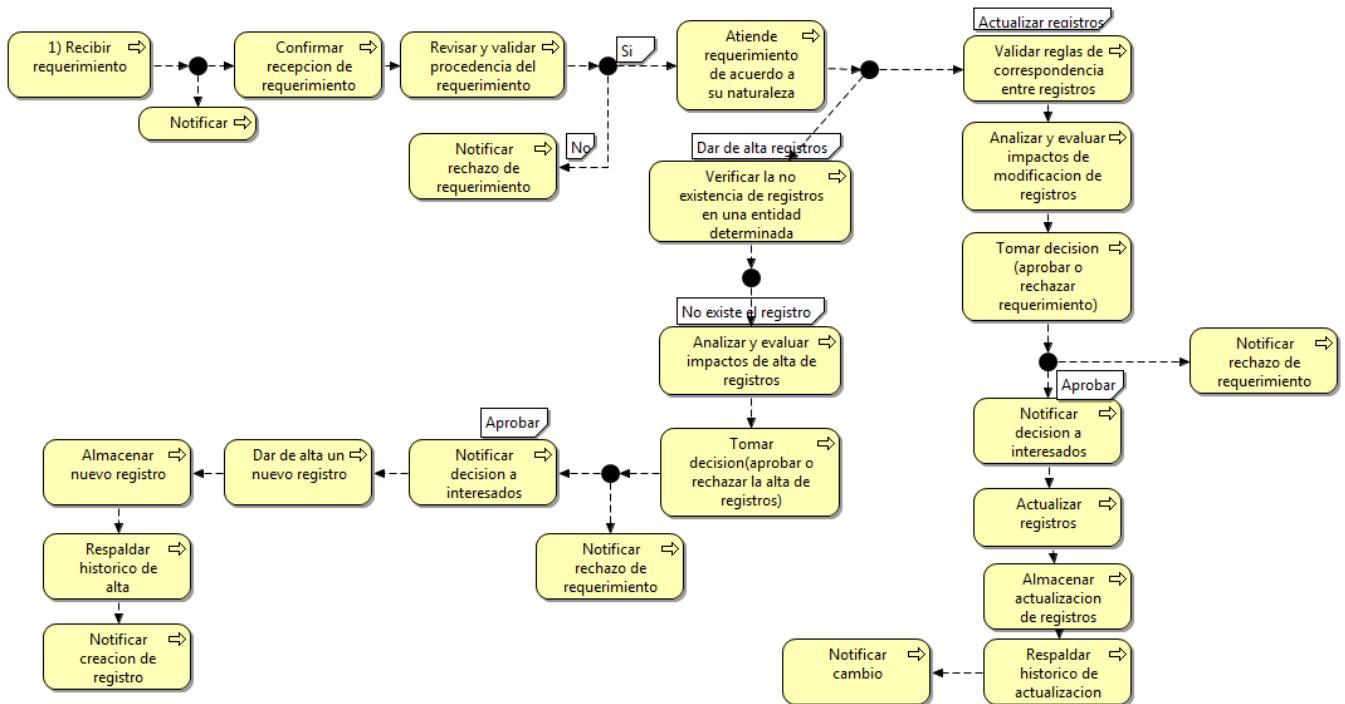


Figura 4.19 Proceso de actualización de registros

### *Intercambio de registros*

El objetivo de este proceso es compartir los datos maestros y de referencia para satisfacer las necesidades de los usuarios de ellas. Entiéndase por usuario o consumidor a una persona, proceso operativo o sistema.

El principal uso de las entidades maestras y de referencia es por el ambiente transaccional, descriptivo e histórico. También pueden ser usadas por participantes internos y externos a las organizaciones.

Las entradas de este proceso son:

#### I. Requerimientos

Las salidas de este proceso son:

- I. Productos de información (reportes, vistas )
- II. Respuesta al requerimiento

Las etapas de este proceso son:

- ✓ Recibir requerimiento
- ✓ Notificar recepción de requerimiento

- ✓ Revisar y validar la procedencia
- ✓ Identificar y validar usuario y registros de diseminación.
- ✓ Tomar decisión. Si la decisión de rechaza se deberá notificar. Si se aprueba:
  - Notificar decisión a interesados
  - Definir medio y diseminar registros
  - Respaldar histórico de acción
  - Notificar acción

A continuación se muestra una vista del flujo de este proceso realizada en Archimate 2.0

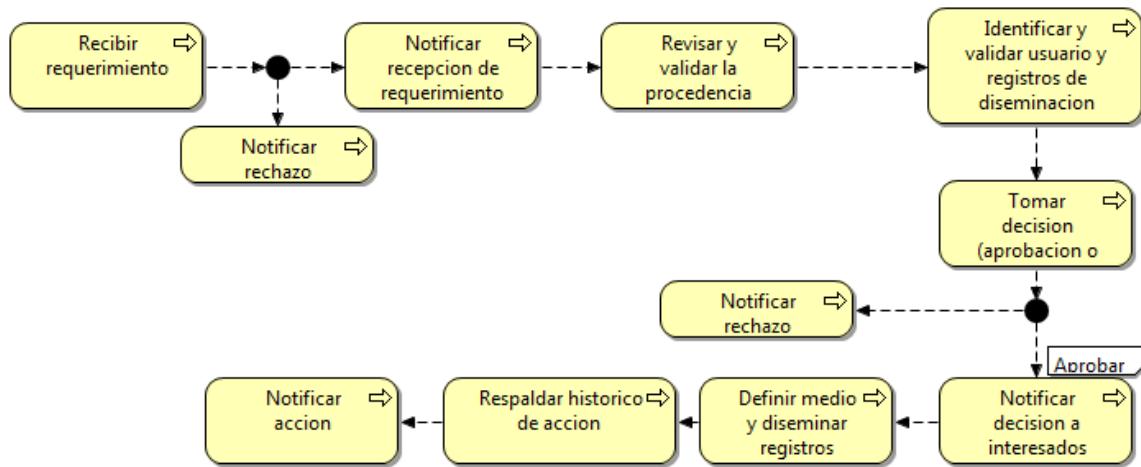


Figura 4.20 Proceso de intercambio de registros

#### 4.3.3 Funciones

Las funciones de negocio entregan capacidades, por lo tanto la función principal para esta arquitectura es la de administración de datos maestros y de referencia. Esta función contempla los cuatro procesos para llevar a cabo esta administración.

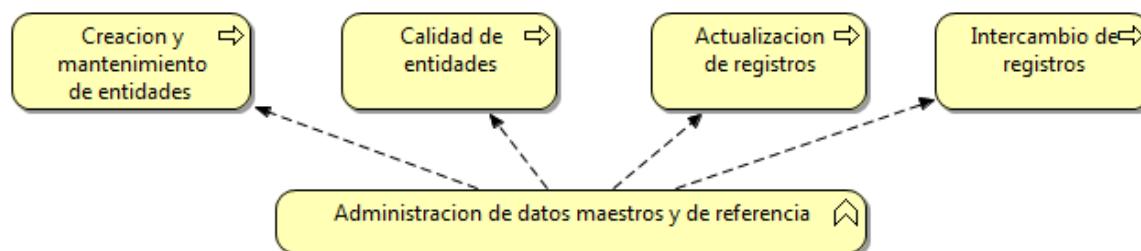
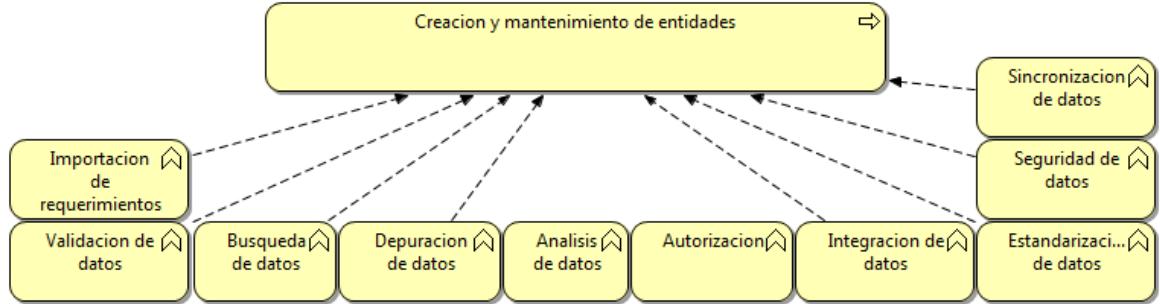


Figura 4.21 Funciones de la gestión de DMyR

Las funciones de negocio para el proceso de creación y mantenimiento de entidades son las siguientes:

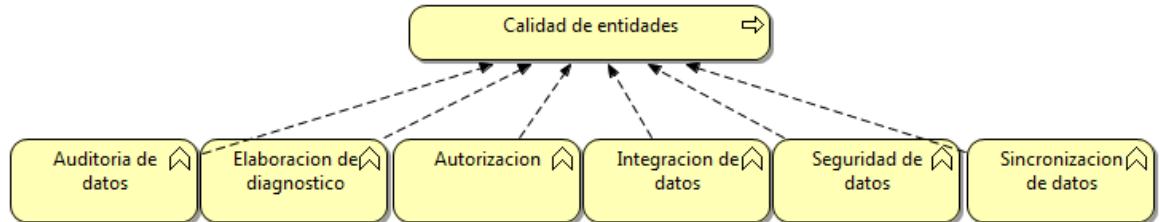
- Importación de requerimientos  
Las solicitudes de ingresarán al sistema de gestión de datos por medio de un formulario web.
- Identificación de datos  
Tiene la funcionalidad para reconocer información en el requerimiento ingresado.
- Búsqueda de datos  
Busca información sobre meta datos de entidades. La búsqueda de información se basa en algoritmos de indexación, codificación y comparación.
- Depuración de datos  
Cumple con el objetivo de proveer la solución más idónea para cada caso de dato. Esta función permite identificar datos incompletos, incorrectos, inexactos y no pertinentes. Un sinónimo de esta función es limpieza de datos. Este proceso cubre la función de análisis de datos.
- Análisis de datos  
Tiene la finalidad de realizar un análisis humano sobre el dato a evaluar. Este análisis se basa en el resultado de las anteriores funciones para tomar una decisión.
- Autorización  
Cumple con la función de tomar una decisión sobre la aprobación o rechazo de un requerimiento.
- Integración de datos  
Cumple con la finalidad de obtener datos exactos, íntegros, consistentes, uniformes, densos y unificados.
- Estandarización de datos  
Tiene la finalidad para almacenar la información sobre las entidades maestras o de referencia en formatos estandarizados a través de toda la dirección.
- Seguridad de datos  
Cumple con el objetivo de almacenar e intercambiar la información bajo medidas de seguridad de la información.
- Sincronización de datos  
Cumple con la función de replicar datos a todos los procesos de negocio que utilicen alguna entidad maestra o de referencia que ha sido modificada.



**Figura 4.22 Funciones para el proceso: Creación y mantenimiento de entidades**

Las funciones para el proceso de calidad de entidades son:

- Auditoría del dato  
Esta función tiene el objetivo de medir, evaluar y realizar acciones sobre la calidad del dato.
- Elaboración de diagnóstico  
Cumple con la funcionalidad de obtener reportes sobre la calidad de las entidades.
- Autorización
- Integración de datos
- Seguridad de datos
- Sincronización de datos

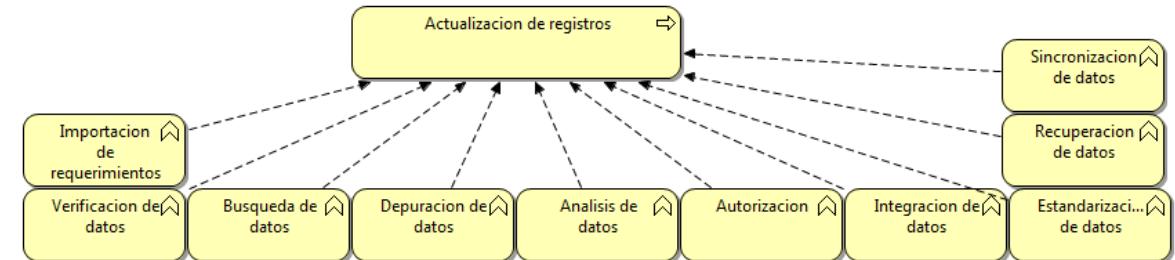


**Figura 4.23 Funciones para el proceso: Calidad de entidades**

Las funciones de negocio para el proceso de actualización de registros son:

- Validación de claves
- Importación de requerimientos
- Identificación
- Verificación de datos
- Búsqueda de datos
- Depuración de datos
- Análisis de datos
- Autorización

- Integración de datos
- Estandarización de datos
- Recuperación de datos
- Sincronización de datos



**Figura 4.24 Funciones para el proceso: Actualización de los registros**

Las funciones para el proceso de intercambio de registros son:

- Validación de claves
- Importación de requerimientos
- Identificación de datos
- Autorización
- Exportación de datos
- Tiene el objetivo de transmitir datos de manera rápida y segura. Se garantía que los datos cumplan con las medidas de seguridad de información.
- Disponibilidad de datos
- Asegura que un conjunto de registros se puedan poner a disposición de las necesidades de los usuarios. El valor de esta función es cumplir de manera oportuna, precisa y eficaz los productos de datos maestros y de referencia.
- Recuperación de datos
- Sincronización de datos

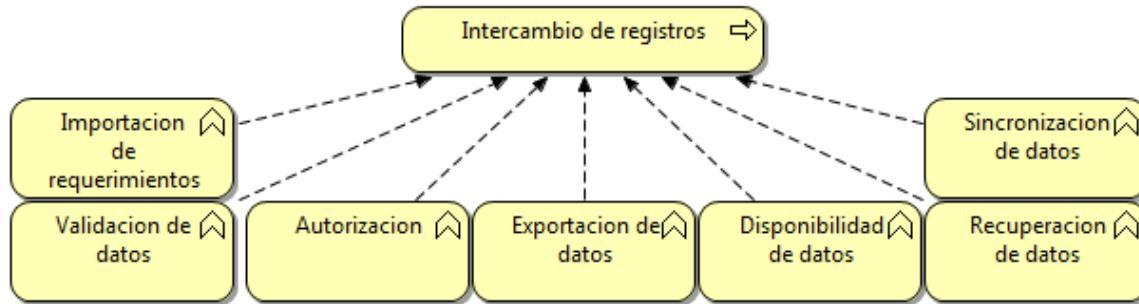


Figura 4.25 Funciones para el proceso: Intercambio de registros

#### 4.3.4 Servicios

Los servicios soportan las capacidades que generan las funciones a través de una interfaz. Los servicios identificados para el proceso de creación y mantenimiento de entidades son:

Función de negocio	Servicio de negocio
Importación de requerimientos	Servicio de despliegue de solicitud de requerimientos
	Servicio para crear una entidad
	Servicio para modificar una entidad
Validación de datos	Servicio de despliegue de información identificada
Búsqueda de datos	Servicio de resultados de búsqueda
Depuración de datos	Servicio de limpieza de datos
Análisis de datos	Servicio de despliegue de selección de datos
Autorización	Servicio de notificación de aprobación o rechazo de requerimiento
Integración de datos	Servicio de información integrada
Estandarización de datos	Servicio de información almacenada en formato definido
Seguridad de datos	Servicio de respaldo de información
Sincronización de datos	Servicio de información sincronizada
	Servicio de proceso finalizado

Tabla 4.5 Servicios para el proceso: Creación y mantenimiento de entidades

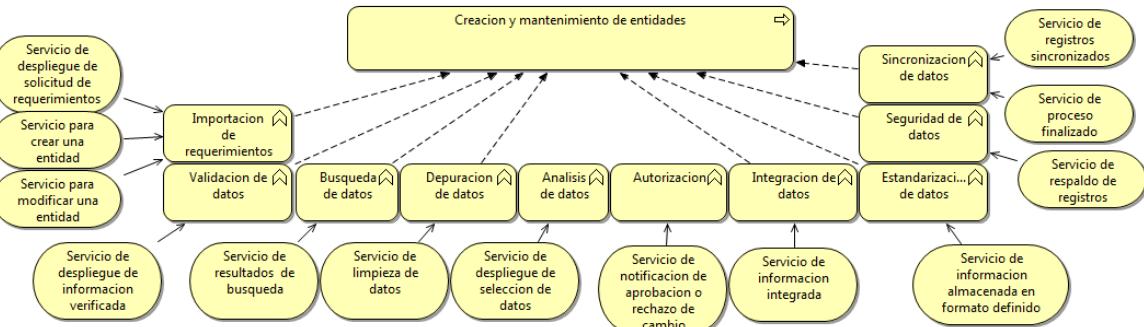


Figura 4.26 Servicios para el proceso: Creación y mantenimiento de entidades

Los servicios identificados para el proceso de calidad de entidades son:

Función de negocio	Servicio de negocio
<b>Auditoría de datos</b>	Servicio de medición de calidad a través de indicadores y fórmulas
<b>Elaboración de diagnóstico</b>	Servicio de despliegue de información de diagnóstico
<b>Autorización</b>	Servicio de aprobación o rechazo de cambio
<b>Integración de datos</b>	Servicio de información integrada
<b>Seguridad de datos</b>	Servicio de respaldo de información
<b>Sincronización de datos</b>	Servicio de información sincronizada Servicio de proceso finalizado

Tabla 4.6 Servicios para el proceso: Calidad de entidades

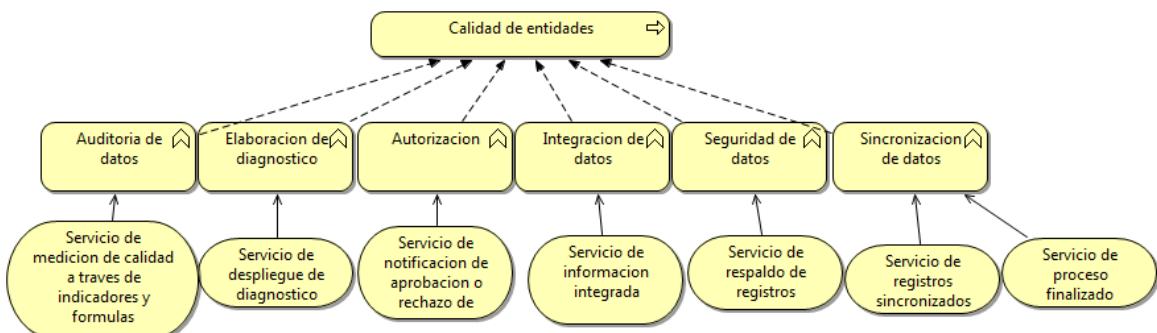


Figura 4.27 Servicios para el proceso: Calidad de entidades

Los servicios identificados para el proceso de actualización de registros son:

Función de negocio	Servicio de negocio
<b>Importación de requerimientos</b>	Servicio de despliegue de solicitud de requerimientos
	Servicio para dar de alta un registro
	Servicio para modificar un registro
	Servicio para eliminar un registro
<b>Verificación de datos</b>	Servicio de despliegue de información verificada
<b>Búsqueda de datos</b>	Servicio de resultados de búsqueda
<b>Depuración de datos</b>	Servicio de limpieza de datos
<b>Análisis de datos</b>	Servicio de despliegue de selección de datos
<b>Autorización</b>	Servicio de notificación de aprobación o rechazo de requerimiento
<b>Integración de datos</b>	Servicio de información integrada
<b>Estandarización de datos</b>	Servicio de información almacenada en formato definido
<b>Seguridad de datos</b>	Servicio de respaldo de información
<b>Sincronización de datos</b>	Servicio de información sincronizada
	Servicio de proceso finalizado

Tabla 4.7 Servicios para el proceso: Actualización de registros

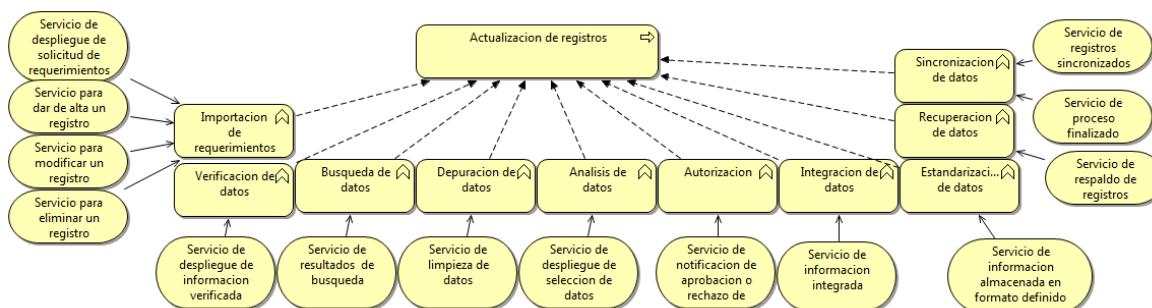


Figura 4.28 Servicios para el proceso: actualización de registros

Los servicios identificados para el proceso de intercambio de registros son:

Función de negocio	Servicio de negocio
<b>Importación de requerimientos</b>	Servicio de despliegue de solicitud de requerimientos
	Servicio para usar un conjunto de registros
<b>Validación de datos</b>	Servicio de despliegue de información verificada
<b>Autorización</b>	Servicio de aprobación o rechazo de requerimiento
<b>Exportación de datos</b>	Servicio para publicar registros
<b>Disponibilidad de datos</b>	Servicio de consultar un conjunto de registros
<b>Recuperación de datos</b>	Servicio de respaldo de registros
<b>Sincronización de datos</b>	Servicio de registros sincronizados

Tabla 4.8 Servicios para el proceso: Intercambio de registros

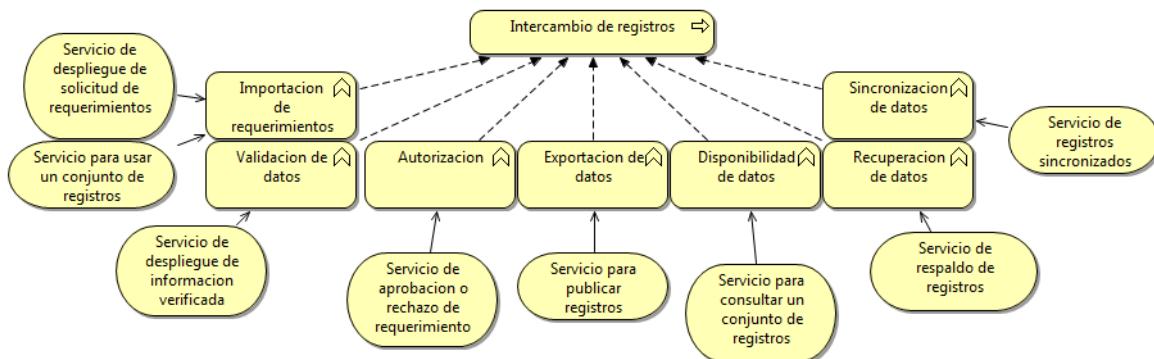


Figura 4.29 Servicios para el proceso: Intercambio de registros

#### 4.3.5 Roles

Esta solución contempla roles para llevar a cabo las actividades de los procesos descritos. El DAMA define tres agrupaciones de roles: responsables, de aprobación y de contribución. En esta arquitectura los roles se clasifican en: de contribución y responsables. No se incluye la agrupación de aprobación ya que la acción de autorizar se considera en los roles responsables.

1. **Roles de contribución.** Son los participantes que suministran información para la actualización del contenido de las EMyR. Los roles de esta agrupación son:

- Usuario de datos
- Proveedor de datos
- Interesado.

2. **Roles responsables.** Son aquellas personas expertas que garantizan el control y uso eficaz de las EMyR. Aprueban, validan y realizan los cambios necesarios para las entidades maestras y de referencia contengan valores de oro y se compartan adecuadamente. Los roles de esta agrupación son:
  - Administrador del negocio
  - Custodio de datos
  - Administrador de datos.
  
3. **Roles de aprobación.** Son aquellas personas que revisan y validan los cambios que surjan de la arquitectura, buscando el mejor interés para las organizaciones, o afecten al contenido de las entidades.
  - Arquitecto de datos
  - Arquitecto del negocio
  - Responsable del negocio
  - Consumidor

#### 4.4 Fase C: Arquitectura de datos

Esta arquitectura soporta la arquitectura de negocio y describe las características que deben tener las entidades de datos maestras y de referencia. También describe el ciclo de vida de estos datos a lo largo de los procesos de negocio.

En una arquitectura de datos existen cuatro niveles de abstracción, estos niveles son los que se muestran en la Figura 4.30.

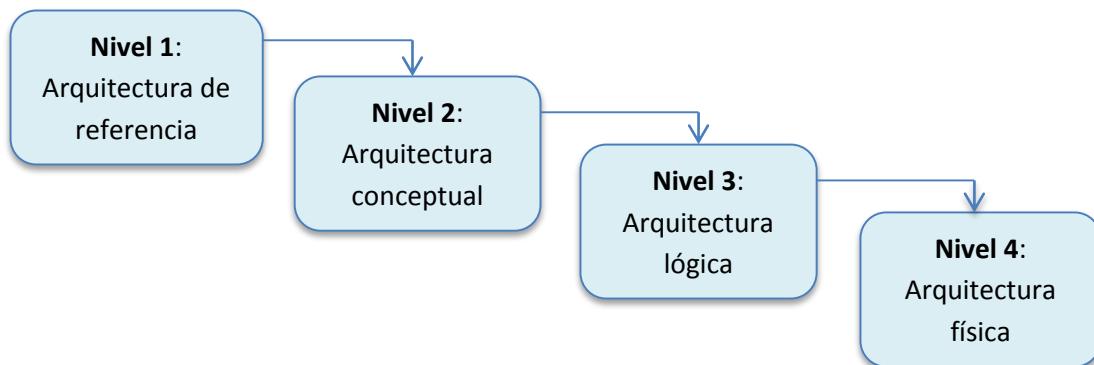


Figura 4.30 Niveles de abstracción de la arquitectura de datos

Una arquitectura de referencia proporciona una plantilla de la generalización de un conjunto de mejores prácticas de la industria de tecnologías de la información. Estas soluciones son estructuradas con base en la recolección de un conjunto de patrones que se han observado en

una serie de implementaciones exitosas. Ejemplos de arquitecturas de referencia son las proporcionadas por TOGAF, DAMA, MDM Institute y EDM Council.

Una arquitectura conceptual es la selección de los componentes y modelos, es decir; en base a una arquitectura de referencia se puede particularizar conforme a las necesidades de cada negocio.

Una arquitectura lógica es el diseño detallado de la arquitectura conceptual. La arquitectura lógica contiene los componentes lógicos (subsistemas, o macro-funciones) que participan en la solución, y la relación entre ellos. Una arquitectura física contempla implementaciones actuales conforme a la arquitectura lógica. Como se mencionó en la *Fase de Visión* de la arquitectura en este trabajo no se desarrollan los artefactos de la arquitectura física.

Como se

#### **4.4.1 Artefactos del diseño conceptual**

La definición de la estructura de una entidad maestra o de referencia depende del propósito de cada una de ellas. El DAMA sugiere que se asignen identificadores globales una vez que se consoliden los registros de diferentes fuentes. Con base en esto y siguiendo el *Principio 2* definido en la *Arquitectura de Negocio* “Las EMyR contienen valores de oro” toda entidad deberá tener un identificador global.

A continuación se describen los catálogos, matrices y diagramas para el nivel conceptual de la arquitectura de datos maestros y de referencia de las organizaciones.

#### **Catálogo de entidades de referencias cruzadas**

Algunos datos de referencia establecen referencias cruzadas de múltiples valores de códigos que representan el mismo concepto. Esto quiere decir que diferentes aplicaciones pueden usar varios conjuntos de códigos para representar el mismo atributo conceptual. Este artefacto es de gran utilidad en las arquitecturas transitorias.

Los datos de referencia cruzada permiten la traducción de un código a otro. La estructura propuesta para el catálogo de entidades de referencias cruzadas es la siguiente:

Atributos	Descripción
-----------	-------------

<b>Identificador</b>	Clave única asignada a cada registro.
<i>[Identificador del proceso A]</i>	Clave de referencia a un proceso determinado.
<i>[Identificador del proceso B]</i>	Clave de referencia a un proceso determinado.
...	...

Tabla 4.9 Catálogo de referencias cruzadas

#### Catálogo de reglas de negocio

Las reglas de negocio son importantes porque diseñan el modelo maestro. Toda regla de negocio se debe definir y almacenar en el siguiente formato estandarizado.

Atributos	Descripción
<b>Clave de la regla de negocio</b>	Clave única asignada a cada regla.
<b>Descripción de la regla de negocio</b>	Descripción de la regla de negocio.

Tabla 4.10 Catálogo de reglas de negocio

#### Catálogo de reglas de correspondencia

En el proceso definido en la Arquitectura de Negocio “Actualización de registros” es necesario para crear, actualizar o borrar un registro maestro o de referencia, validar reglas de correspondencia. Con base en el DAMA las reglas de correspondencia pueden coincidir en los siguientes escenarios:

- Para nuevos datos, se deberá asegurar que no existan en el repositorio central.
- Para editar o borrar datos, se deberá comprobar su edición conforme a reglas de negocio y realizar un análisis del impacto que puede tener el cambio.

La estructura propuesta para recopilar las reglas de correspondencia es la siguiente:

Atributos	Descripción
-----------	-------------

<b>Clave de la regla</b>	Identificador único asignado a cada regla.
<b>Nombre</b>	Nombre corto de la regla de correspondencia.
<b>Descripción</b>	Breve descripción de la regla de correspondencia.
<b>Entidades que les aplique la regla</b>	Enlista las entidades maestras o de referencia que interactúen con la regla.
<b>% Mínimo de correspondencia entre parámetros</b>	Porcentaje para que la ejecución de la regla sea verdadera.
<b>Atributos/Parámetros</b>	Parámetros de las entidades de datos para verificar su validación.
<b>Comentarios</b>	Incluir notas o comentarios importantes para la implementación de las reglas.

Tabla 4.11 Catálogo de reglas de correspondencia

Debido a que nunca se puede evitar que las reglas resulten incorrectas se deberá guardar el histórico para poder deshacer el cambio cuando se requiera, para requiere un atributo extra denominado Estado.

#### *Matriz Entidad/Función de negocio*

Con base al artefacto propuesto por TOGAF el propósito de este artefacto es representar las relaciones entre las entidades maestras o de referencia y las funciones establecidas en la arquitectura de negocio.

	<i>[Nombre de la entidad maestra o de referencia 1]</i>	<i>[Nombre de la entidad maestra o de referencia 2]</i>
<i>[Función de Negocio 1]</i>		
<i>[Función de Negocio 2]</i>		

Tabla 4.12 Matriz Entidad/Función

### *Diagrama de clases*

El diagrama de clases es aquel que contiene las relaciones entre las entidades maestras y de referencia. El propósito de este artefacto es abstraer las relaciones entre las entidades de datos dentro de la empresa a nivel conceptual. A continuación se muestra un ejemplo de un diagrama de clases utilizando la nomenclatura UML:

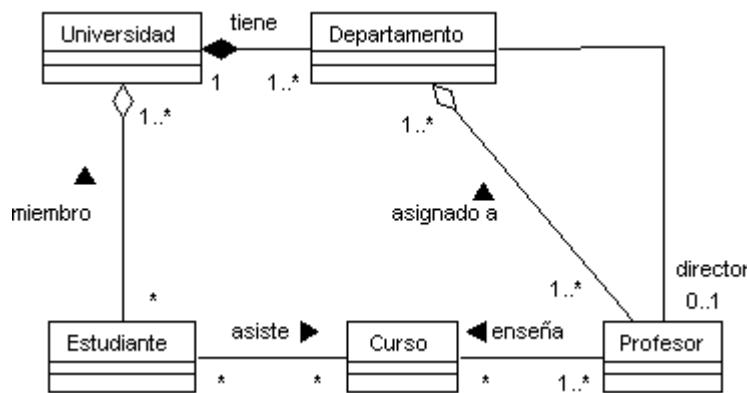


Figura 4.31 Ejemplo de un diagrama de clases de una Universidad. Tomado de Wikipedia.

### *Diagrama del ciclo de vida del dato*

Con base en la metodología de TOGAF, este diagrama es parte esencial de la administración de datos. El ciclo de vida representa las etapas por las que el dato pasa a lo largo de su vida, desde su concepción hasta su eliminación.

Por ejemplo, el cómo se crea un usuario, depende en gran medida de los procesos definidos en el negocio. El ciclo de los datos maestros puede abarcar la forma en que se crea, lee, actualiza y elimina (Microsoft, 2006).

Retomando las partes esenciales de los diagramas del DAMA, TOGAF y Microsoft se ha realizado la siguiente propuesta de las etapas que deben tener el ciclo de vida únicamente para las entidades maestras y de referencia:



Figura 4.32 Diagrama del ciclo de vida del DMyR. Adaptada de Microsoft, 2006.

- **Crear:** para que un dato maestro o de referencia se pueda crear necesita haberse validado previamente, de lo contrario se podrían crear duplicados.
- **Almacenar:** una vez que el dato es creado, actualizado o deshabilitado es necesario resguardarlo en un sitio. Este sitio puede ser un repositorio centralizado. El repositorio no debe permitir inconsistencias semánticas y sintácticas.
- **Distribuir:** cuando el dato se encuentra almacenado en un repositorio centralizado, este se comparte a través de servicios de negocio para su uso.
- **Usar:** Todas las etapas del ciclo de vida del dato tiene asociado riesgos, pero solo la etapa de uso añade valor al negocio, es decir; el dato tiene valor cuando es usado actualmente o en el futuro.
- **Actualizar o deshabilitar:** el dato puede sufrir cambios a través del tiempo, si esta actualización no se almacena y distribuye se estarán usando datos obsoletos.

#### **4.4.2 Artefactos del diseño lógico**

La arquitectura de datos lógica incluye las estructuras de datos utilizadas por la tecnología. A continuación se describen los catálogos, matrices y diagramas para el nivel conceptual de la arquitectura de datos maestros y de referencia de las organizaciones.

##### *Catálogo de metadatos*

Este artefacto contiene la descripción y caracterización de los datos de cada objeto maestros o de referencia. Este catálogo permitirá recuperar, interpretar, usar y administrar dichos activos. La estructura propuesta es la siguiente:

Atributos	Descripción
<b>Clave del objeto maestro o de referencia</b>	Identificador único asignado a cada objeto maestro o de referencia.
<b>Nombre del objeto de datos</b>	Nombre del objeto de datos.
<b>Descripción del objeto de datos</b>	Breve descripción del objeto de datos
<b>Categoría</b>	Clasificación del objeto de datos, puede ser maestra o de referencia.
<b>Nombre del atributo</b>	Nombre de los atributos que estructuran al objeto maestro o de referencia.
<b>Descripción del atributo</b>	Descripción de cada uno de los atributos que conforman el objeto maestro o de referencia.

Tabla 4.13 Catálogo de metadatos

### *Catálogo de relaciones entre objetos*

Es recomendable documentar todas las relaciones entre entidades maestras y de referencia, independientemente si en la arquitectura física se implementan mediante procedimientos almacenados o en el modelo de datos se establece la vinculación. El siguiente artefacto sirve para analizar el impacto que tendrá el cambio de un objeto de datos en los otros que dependen de este.

	<i>[Nombre del objeto maestra o de referencia 1]</i>	<i>[Nombre del objeto maestra o de referencia 2]</i>
<i>[Nombre del objeto maestra o de referencia 1]</i>		
<i>[Nombre del objeto maestra o de referencia 2]</i>		

**Tabla 4.14 Catálogo de relaciones entre objetos**

### *Matriz Objeto/Aplicación*

Con base en TOGAF el siguiente artefacto representa las relaciones de componentes lógicos de aplicaciones contra objetos de datos, sean maestras o de referencia.

	<i>[Nombre del objeto maestra o de referencia 1]</i>	<i>[Nombre del objeto maestra o de referencia 2]</i>
<i>[Componente lógico de App 1]</i>		
<i>[Componente lógico de App 2]</i>		

**Tabla 4.15 Matriz Objeto/ Aplicación**

### *Matriz de seguridad del dato*

Con base en TOGAF el propósito de esta matriz es representar que actores pueden acceder a los datos maestros y de referencia. El artefacto original de TOGAF tiene los siguientes atributos: actor, clase de rol, función, servicio de negocio, locación y tipo de acceso. La personalización de dicho artefacto para este trabajo es la siguiente:

<b>Rol</b>	<b>Función</b>	<b>Servicio</b>	<b>Tipo de acceso</b>

**Tabla 4.16 Matriz de seguridad del dato**

### *Modelo de datos maestros y de referencia*

Un modelo de datos a nivel lógico identifica y refleja los atributos necesarios para cada una de las entidades de datos que surgen de las entidades del negocio representadas en el Diagrama de clases.

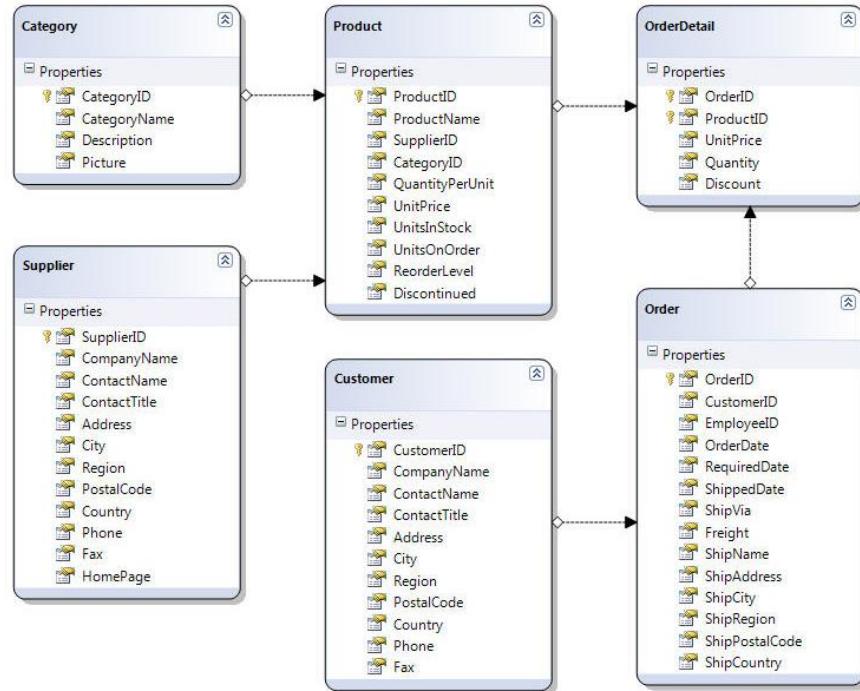


Figura 4.33 Ejemplo de un modelo de datos lógico

#### Diagrama de diseminación de datos

Con base en el artefacto de TOGAF, el propósito de este diagrama es mostrar las relaciones entre las entidades maestras y de referencia, servicios de negocio y las aplicaciones. El diagrama debe mostrar como las entidades lógicas son utilizadas por los componentes de aplicación.

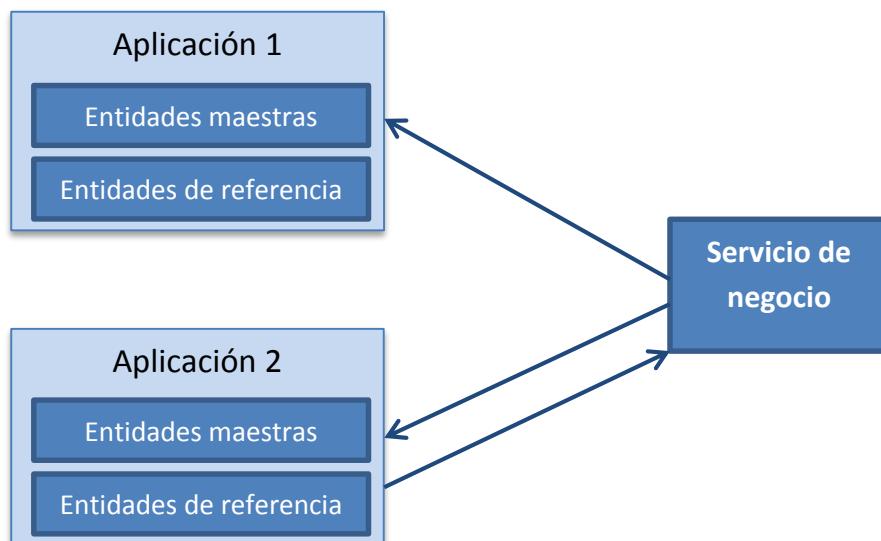


Figura 4.34 Diagrama de diseminación de datos

Adicionalmente el diagrama puede mostrar la replicación. Con base en el DAMA la replicación de datos se puede llevar a cabo por las siguientes formas:

1. Los datos se pueden leer directamente desde una base
2. Los datos pueden ser replicados a otras bases de datos

La segunda opción tiene la ventaja de que en la base central se puede garantizar fácilmente la integridad referencial puesto que no se hacen consultas distribuidas.

#### **4.5 Fase E: Oportunidades y soluciones**

Con base en la metodología de TOGAF la implementación de la arquitectura definida para la gestión de datos maestros requiere de una planeación estratégica. Esta planeación deberá contener un plan de trabajo para el desarrollo de la infraestructura que soporta la gestión de datos maestros y de referencia. El plan de trabajo deberá cumplir las expectativas. Esto se basa en el análisis sobre las brechas que existen con la arquitectura objetivo.

##### **4.5.1 Restricciones para la implementación**

- Los datos maestros y de referencia se deben almacenar en un repositorio único
- Los datos maestros y de referencia no se deberán duplicar
- El intercambio de datos se deberá realizar mediante la implementación de las aplicaciones definidas en la arquitectura de aplicaciones.

##### **4.5.2 Estrategia de migración e implementación**

La estrategia de migración conduce la implementación para obtener como resultado a la arquitectura objetivo. En este trabajo la estrategia que se presenta a continuación se basa en una estrategia de convergencia de acuerdo con TOGAF. Se eligió la estrategia de convergencia porque las directrices se pueden definir por fases para introducir nuevas capacidades.

La estrategia se desarrolló enlistando las actividades que se deben realizar para llegar a obtener la arquitectura negocio, datos y aplicaciones (definidas anteriormente) y las capacidades del modelo de madurez. La propuesta de la estrategia aborda cada una de estos componentes considerados en el MDM Madurity Model y sus capacidades mediante programas y proyectos.

A continuación se muestran los programas y sus respectivos proyectos de la estrategia para poner en marcha las anteriores fases del ciclo de TOGAF y cumplir con las capacidades del modelo de madurez.

La estrategia para poner en marcha la definición de las arquitecturas de negocio, datos y aplicaciones contiene programas que materializan las capacidades de este modelo de madurez.

En la siguiente figura se muestra que a través de los proyectos de cada programa se materializan las capacidades de los componentes del modelo de madurez.

Partiendo de la estrategia desarrollada en la *Visión de la arquitectura*, a continuación se muestra que proyecto de que iniciativa cubre las capacidades cubre de cada componente del modelo de madurez MDM.

Iniciativa 1: Detección y análisis de datos maestros y de referencia	
Proyectos	Descripción:
1.1 Abstracción y conceptualización de los datos	Este proyecto contempla las actividades necesarias para definir el metadato del dato maestro y establecer sus jerarquías. Así mismo definir sus reglas de negocio.
1.2 Recolección de datos	Este proyecto contempla las actividades para ubicar, extraer a los datos, definir un artefacto para su recolección y validarlos de acuerdo al artefacto.
1.3 Consolidación del contenido de datos maestros y de referencia	Este proyecto contempla las actividades para realizar un análisis de los datos recopilados a través de métodos de limpieza de datos.

Tabla 4.17 Proyectos de la Iniciativa “Detección y análisis de Datos maestros y de referencia”

Iniciativa 2. Descripción y modelado de datos maestros y de referencia	
Proyectos	Descripción:
2.1 Definición de las representaciones maestras	Este proyecto contempla las actividades para definir conceptualmente el uso de los datos maestros en el negocio, así como definir las entidades y el contenido de ellas que se tienen que replicar a diversos procesos.
2.2 Integración de las representaciones maestras en un modelo de datos	Este proyecto contempla las actividades para diseñar un modelo central con las entidades establecidas del proyecto anterior.

Tabla 4.18 Proyectos de la Iniciativa “Descripción y modelado de datos maestros y de referencia”

Iniciativa 3. Distribución de los datos maestros y de referencia	
Proyectos	Descripción:
3.1 Implementación de los servicios	Este proyecto contempla las actividades para identificar las aplicaciones que proveerán los datos a través de la vinculación del repositorio y servicios SOA.

Tabla 4.19 Proyectos de la Iniciativa 3 “Distribución de los Datos maestros y de referencia”

Iniciativa 4. Colaboración de los recursos humanos	
Proyectos	Descripción:
4.1 Comunicación y promoción de la arquitectura definida	Este proyecto contempla las actividades para difundir a la organización lo establecido para la gestión de datos maestros.
4.2 Capacitación del personal necesario	Este proyecto contempla las actividades para capacitar al personal necesario para la puesta en marcha y operación de la arquitectura definida.

Tabla 4.20 Proyectos de la Iniciativa 4 “Colaboración de los recursos humanos”

#### 4.6 Fase F: Planeación de migración

Una estructura de desglose de trabajo o por sus siglas en inglés WBS Work Breakdown Structure es la descomposición jerárquica de entregables para cumplir con los objetivos de un proyecto. Su representación jerárquica permite una identificación más fácil de los paquetes de trabajo que se tienen que generar.

En la Figura 4.35 se muestra la Estructura de Descomposición de Trabajo para el desarrollo de la arquitectura de DMyR. Esta estructura primero se descompone en cuatro iniciativas, las que a su vez en proyectos, y posteriormente para cada uno de estos proyectos se plantearán actividades para conformar el *Plan de migración e implementación*.

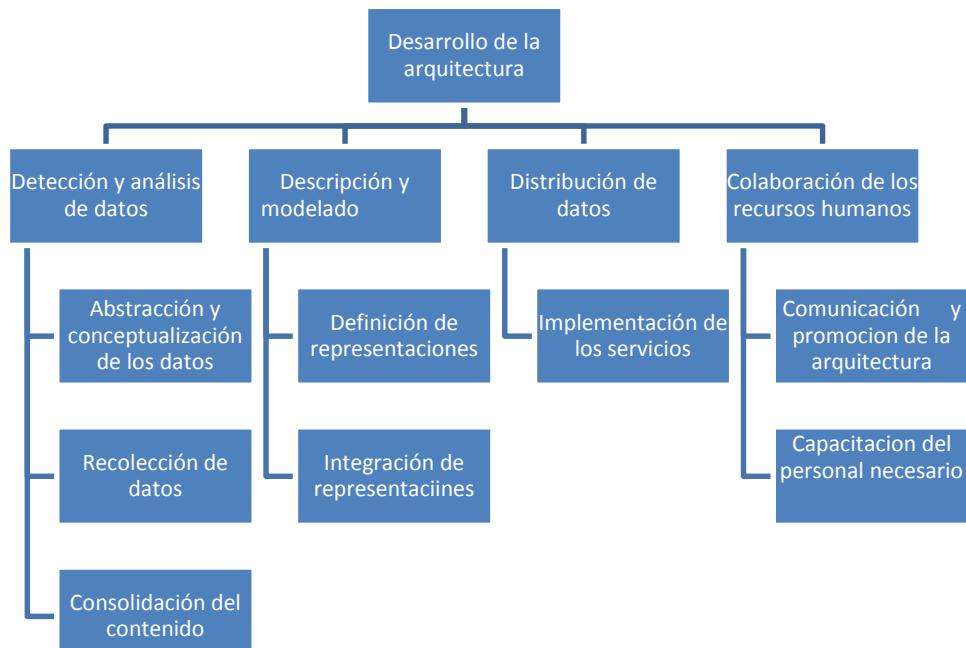


Figura 4.35 Estructura de desglose del trabajo para desarrollar la arquitectura de DMyR

#### **4.6.1 Plan de migración e implementación**

Este plan contiene las actividades y el orden de secuencia en que estas se deberán ejecutar. Debido a que cada organización puede tener sus propios requerimientos, a continuación se presentan las etapas generales de desarrollo para incorporarlas y detallarlas en el plan de migración e implementación para cada caso específico.

<b>Iniciativa</b>
1. Detección y análisis de datos maestros y de referencia
<b>Proyectos</b>
1.1 Abstracción y conceptualización de los datos
<b>Actividades que materializan las capacidades:</b>
- Definir definición y clasificaciones del dato maestro o de referencia
1.2 Recolección de datos
- Localizar los datos maestros y de referencia en procesos de negocio, bases de datos - Adoptar algoritmos de búsqueda de datos - Definir artefacto estandarizado para su recolección - Recopilar o extraer datos para llenar el artefacto - Validar los datos para conservar su estandarización - Almacenar los datos y prepararlos para su análisis
1.2 Consolidación del contenido de datos maestros y de referencia
- Elegir herramienta para su análisis - Elegir métodos de depuración de datos - Definir parámetros y reglas de correspondencia - Aplicar métodos de limpieza de datos - Resolver manualmente diferencias lingüísticas

Tabla 4.21 Plan de implementación para la Iniciativa 1 “Detección y análisis de DMyR”

<b>Iniciativa</b>
2. Descripción y modelado de datos maestros
<b>Proyectos</b>
2.1 Definición de las representaciones maestras
<p><b>Actividades que materializan las capacidades:</b></p> <ul style="list-style-type: none"> <li>- Definir su ubicación dentro de los procesos de negocio</li> <li>- Desarrollar una modelo conceptual de información</li> <li>- Diseñar una arquitectura de datos conceptual. Dentro de los artefactos de esta arquitectura se incluirán diagramas del ciclo de vida del dato, así como diagramas de los procesos que gestionen a los datos.</li> <li>- Definir el subconjunto de atributos acotado que se requieren mostrar para cada proceso o requerimiento de negocio</li> <li>- Definir el subconjunto de datos para cada proceso de negocio</li> <li>- Definir los atributos para representar la entidad maestra contemplando todos los subconjuntos definidos anteriormente.</li> <li>- Definir sus metadatos</li> </ul>
2.2 Integración de las representaciones maestras en un modelo de datos
<ul style="list-style-type: none"> <li>- Definir reglas de negocio para modelar</li> <li>- Establecer las relaciones con otras entidades maestras</li> <li>- Desarrollar catálogos de referencia</li> <li>- Mapear sus claves a un catálogo de referencia</li> <li>- Consolidar la representación de un modelo maestro</li> <li>- Definir los modelos de datos mediante un lenguaje que incluya taxonomías, ontologías u otras representaciones conceptuales y lógicas de alto nivel.</li> <li>- Implementar el modelo maestro en un repositorio</li> </ul>

Tabla 4.22 Plan de implementación para la Iniciativa 2 “Descripción y modelado de DM”

<b>Iniciativa</b>
3. Distribución de los datos maestros y de referencia
<b>Proyectos</b>
3.1 Implementación de los servicios
<p><b>Actividades que materializan las capacidades:</b></p> <ul style="list-style-type: none"> <li>- Llenar el artefacto de tipo matriz entidad versus aplicación y esta información representarla en diagramas.</li> <li>- Llenar todos los artefactos de matrices</li> <li>- Desarrollar los servicios e interfaces definidos en la fase C (Arquitectura de aplicaciones )</li> <li>- Vincular los servicios al modelo físico</li> <li>- Hacer las pruebas para insertar, actualizar y replicar los datos maestros a través de los servicios</li> </ul>

Tabla 4.23 Plan de implementación para la Iniciativa 3 “Distribución de DMyR”

<b>Iniciativa</b>
4. Colaboración de los recursos humanos
<b>Proyectos</b>
4.1 Comunicación y promoción de la arquitectura definida
<b>Actividades que materializan las capacidades:</b>
<ul style="list-style-type: none"> <li>- Preparar los contenidos audio visuales para difundir la arquitectura</li> <li>- Definir mecanismos y medios para comunicar la arquitectura</li> <li>- Difundir los contenidos desarrollados</li> <li>- Establecer un seminario o taller informativo</li> <li>- Evaluar el nivel de comprensión de los contenidos comunicados</li> </ul>
4.2 Capacitación del personal necesario
<b>Actividades que materializan las capacidades:</b>
<ul style="list-style-type: none"> <li>- Evaluación de conocimientos</li> <li>- Establecer los cursos necesarios</li> <li>- Definir su prioridad</li> <li>- Enviar convocatoria de cursos</li> </ul>

Tabla 4.24 Plan de implementación para la Iniciativa 4 “Colaboración de los recursos humanos”

## 4.7 Meta-modelo de la arquitectura

Un meta-modelo expresa, a un nivel abstracto, la interacción de varios elementos de un mismo tema. Con base en el alcance de la arquitectura desarrollada en el presente trabajo a continuación se muestra la acotación del meta-modelo que proporciona TOGAF.

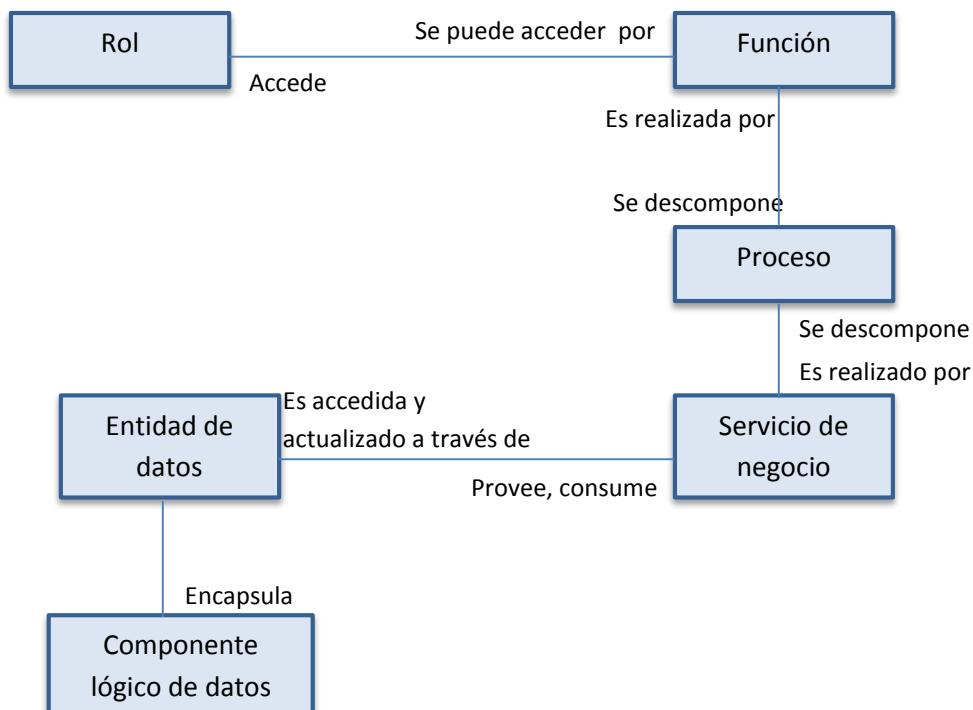


Figura 4.36 Meta-modelo de la arquitectura. Traducida y adaptada de TOGAF, 2011.

## 4.8 Bloques de construcción de la arquitectura

TOGAF define a un bloque de construcción como un paquete que satisface ciertas necesidades del negocio a través de una funcionalidad específica. Un bloque de construcción puede contener otros bloques de construcción, es decir, se pueden definir bloques a varios niveles de detalle.

En la Figura 4.37 se muestran los bloques de construcción que se desarrollaron como parte del alcance de este trabajo.

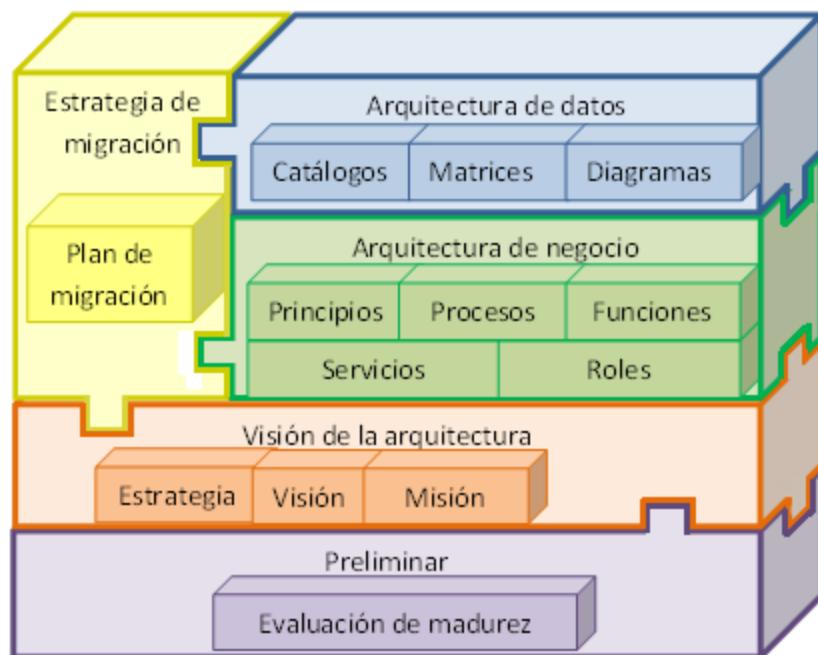


Figura 4.37 Bloques de construcción de la arquitectura

Se puede visualizar que el bloque de construcción de la *Arquitectura de datos* está conformado por los sub bloques *Catálogos*, *Matrices* y *Diagramas*.

El bloque de la arquitectura de datos es soportado por el bloque de *Arquitectura de negocio*, que contiene los sub bloques *Principios*, *Procesos*, *Funciones*, *Servicios* y *Roles*. Este bloque es soportado el bloque *Visión de la arquitectura*, que contiene la *Estrategia general*, la *Visión* y *Misión*.

A su vez los bloques *Visión de la arquitectura*, *Arquitectura de datos* y *Arquitectura de negocio* soportan al bloque de *Estrategia de migración*, el cual contiene el sub bloque *Plan de migración*. Esto es porque la Fase E y F de TOGAF contempla todas las arquitecturas de definición de las fases B, C y D.

Finalmente la definición de los cuatro bloques que se han mencionado (*Estrategia de migración*, *Arquitectura de datos*, *Arquitectura de negocio* y *Visión de la arquitectura*) es

soportada por el desarrollo del primer bloque *Preliminar*, que contiene la evaluación de madurez.

## **5. CASO DE ESTUDIO: USUARIOS Y PROVEEDORES DE UN ÁREA FINANCIERA**

Actualmente un área de un banco provee información financiera, que recibe y procesa, a unidades administrativas del mismo banco, otras autoridades financieras y al público en general. El procesamiento de esta información contempla su validación, y en algunos casos su transformación.

Para que el Área Financiera (AF) genere y provea productos de calidad es indispensable verificar y enriquecer la información recibida con datos maestros y de referencia. Los principales macro-procesos de negocio del Área Financiera para la recepción y entrega de información son: acopio, validación, transformación y diseminación. Estos macro-procesos son los que utilizan catálogos, por ejemplo, de moneda, país, divisa, plazo, proveedores, contrapartes y filiales, entre otros, para verificar la información de entrada.

A través de los macro-procesos mencionados previamente, es posible operar más de 30 Formularios<sup>1</sup> en el área financiera. Estos formularios corresponden a cada temática del sistema financiero y de acuerdo a su naturaleza, se tienen diferentes expectativas del contenido de los datos maestros y de referencia. Por ejemplo cada uno de los *Formularios* ocupa un subconjunto igual o diferente de los valores del *Catálogo de moneda*.

Este caso de estudio se enfoca en aplicar el diseño de la arquitectura del Capítulo 4 a los datos de proveedores y usuarios de un área financiera.

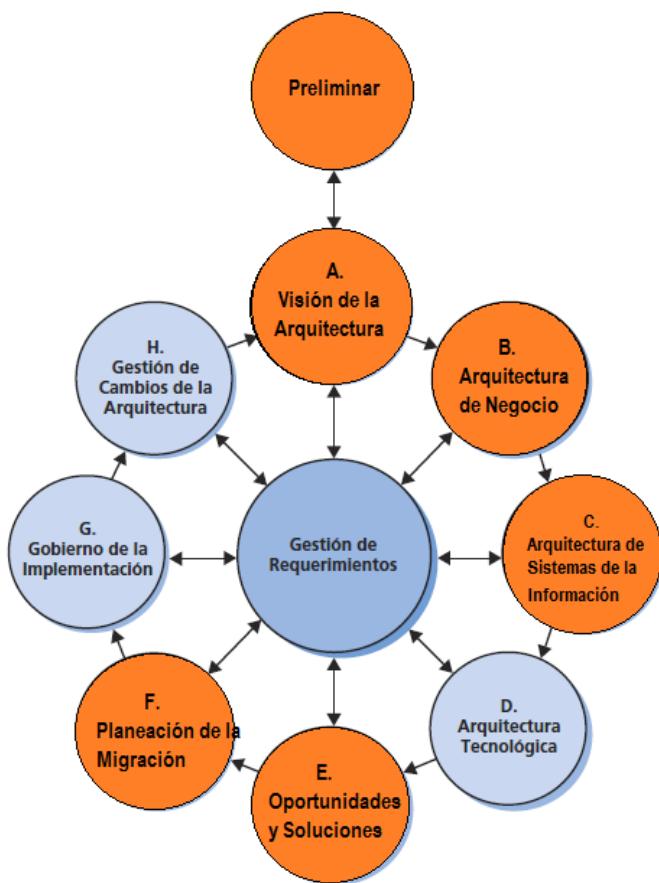
### **5.1 Fase 0: Preliminar**

El alcance de este caso de estudio involucra la implementación de dos iniciativas: 1.*Detección y análisis de DMyR* y 2.*Descripción y modelado de DMyR* en la Fase F (Planeación de la Migración).

De la estrategia definida en la Visión de la Arquitectura no se implementaron las iniciativas 3 (*Distribución de los DMyR a través de servicios SOA*) y 4 (*Colaboración de los recursos humanos*) debido a que no se diseñó la Arquitectura de Aplicaciones. En la Figura 5.1 muestra las fases abarcadas en este caso de estudio.

---

<sup>1</sup> Un formulario en el área financiera corresponde a un proceso operativo del negocio.



**Figura 5.1 Alcance del ADM para el caso de estudio**

### 5.1.1 Evaluación de madurez para usuarios y proveedores

Para medir el nivel de madurez respecto al tratamiento de los datos de usuarios y proveedores en el Área Financiera se utilizaron los cuestionarios de evaluación de madurez adaptados en el *Capítulo 4*, como se ha comentado estos cuestionarios se obtuvieron con base al modelo creado por David Loshin MDM Madurity Model.

#### *Componente 1: Arquitectura*

Las tres capacidades a evaluar como parte del componente Arquitectura son los modelos de datos maestros, el sistema de arquitectura y la capa de servicios.

Para el componente de Arquitectura el Área Financiera se tiene un modelo de datos físico únicamente para proveedores, pero no se tiene para usuarios. Sin embargo, este modelo físico para proveedores no contempla todos los casos de negocio. Por lo que respecto a este primer componente no se tiene un modelo consolidado y centralizado de estos datos maestros que soporte los diferentes formatos y estructuras de datos.

Con base en lo descrito anteriormente el nivel de madurez sobre el tratamiento de usuarios y proveedores para el componente arquitectura es:

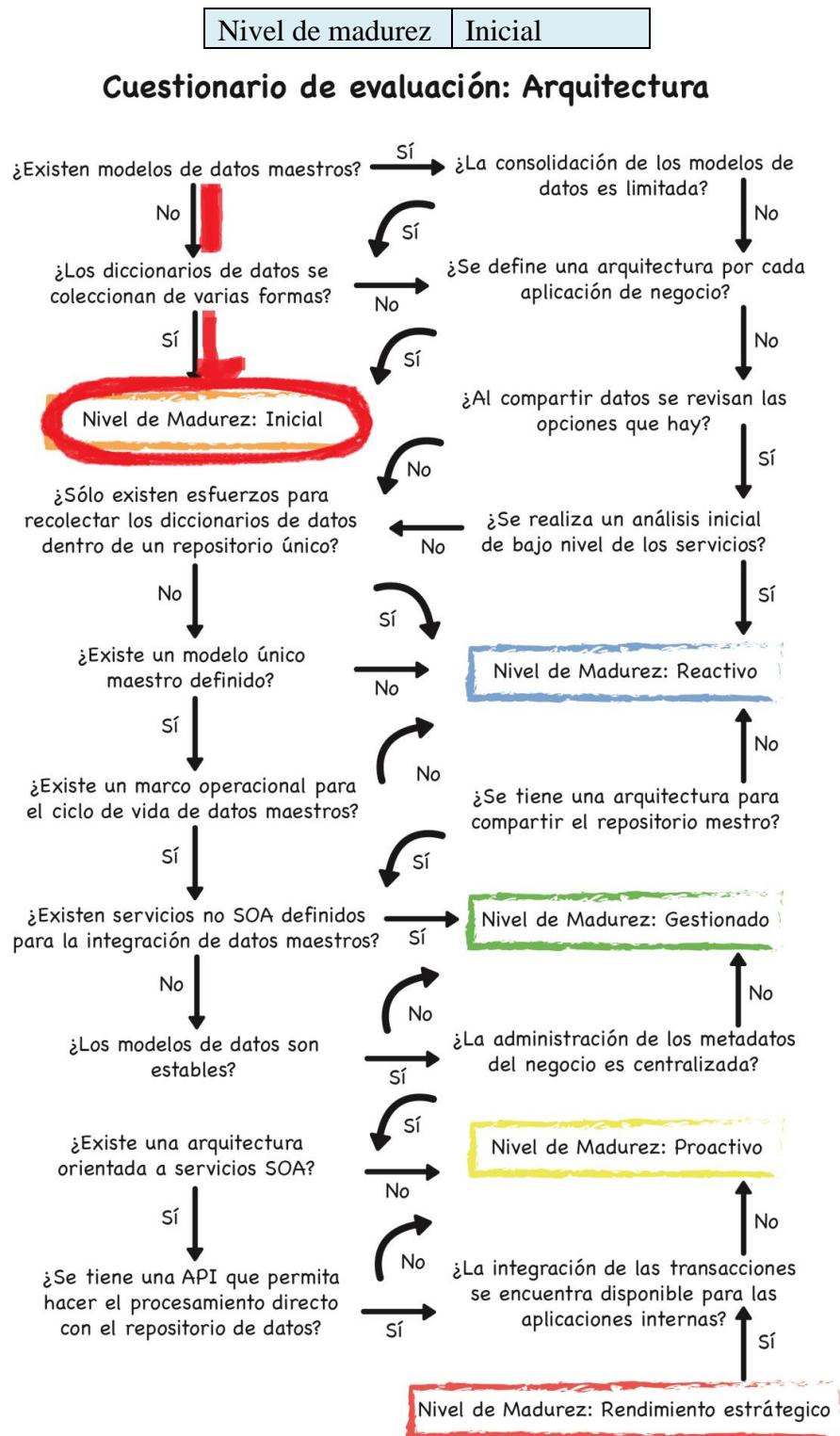


Figura 5.2 Evaluación del componente de Arquitectura

Las características para determinar este nivel de madurez son las siguientes:

- No existen modelos conceptuales de usuarios y proveedores definidos por el área financiera. Físicamente existen modelos implementados sin una estandarización de su estructura ni contenido; muchos datos se encuentran duplicados en varias bases de datos.
- No se ha consolidado ni establecido conceptualmente un modelo de estos datos maestros por parte de las tres gerencias del área financiera.
- Se tienen colecciones de diccionarios de estos datos maestros en formatos particulares de cada gerencia.

No se puede considerar a la arquitectura de datos maestros en un nivel 2 (Reactiva) debido a que no se coleccionan diccionarios de datos en un repositorio central, ni tampoco se comparten los valores de estos datos maestros que tiene cada gerencia a través de toda el área financiera.

Tampoco se tiene un ciclo de vida definido para estos datos maestros, evidentemente tampoco servicios que apoyen a este ciclo. La información que se recolecta de usuarios y proveedores en el área financiera se hace de forma aislada por cada gerencia; por diferentes métodos y momentos. La capa de servicios no se tiene definida.

#### *Componente 2: Gobierno*

Las tres capacidades a evaluar como parte de este componente son definiciones estandarizadas, administración de metadatos, calidad de datos y administración de datos.

En el Área Financiera respecto a la primera capacidad no se tiene un método definido para resolver las diferencias lingüísticas de los datos. Para la segunda capacidad no se tienen descripciones de usuarios y proveedores ni mucho menos de sus atributos. Para la capacidad de la calidad no se tienen establecidos indicadores que la midan. Finalmente, para la administración tampoco existen lineamientos definidos para la gobernanza de usuarios y proveedores, evidentemente tampoco existen métodos aun definidos para garantizar su cumplimiento.

Con base en esto el nivel de madurez que se ha determinado para el segundo componente es:

Nivel de madurez	Inicial
------------------	---------

## Cuestionario de evaluación: Gobierno

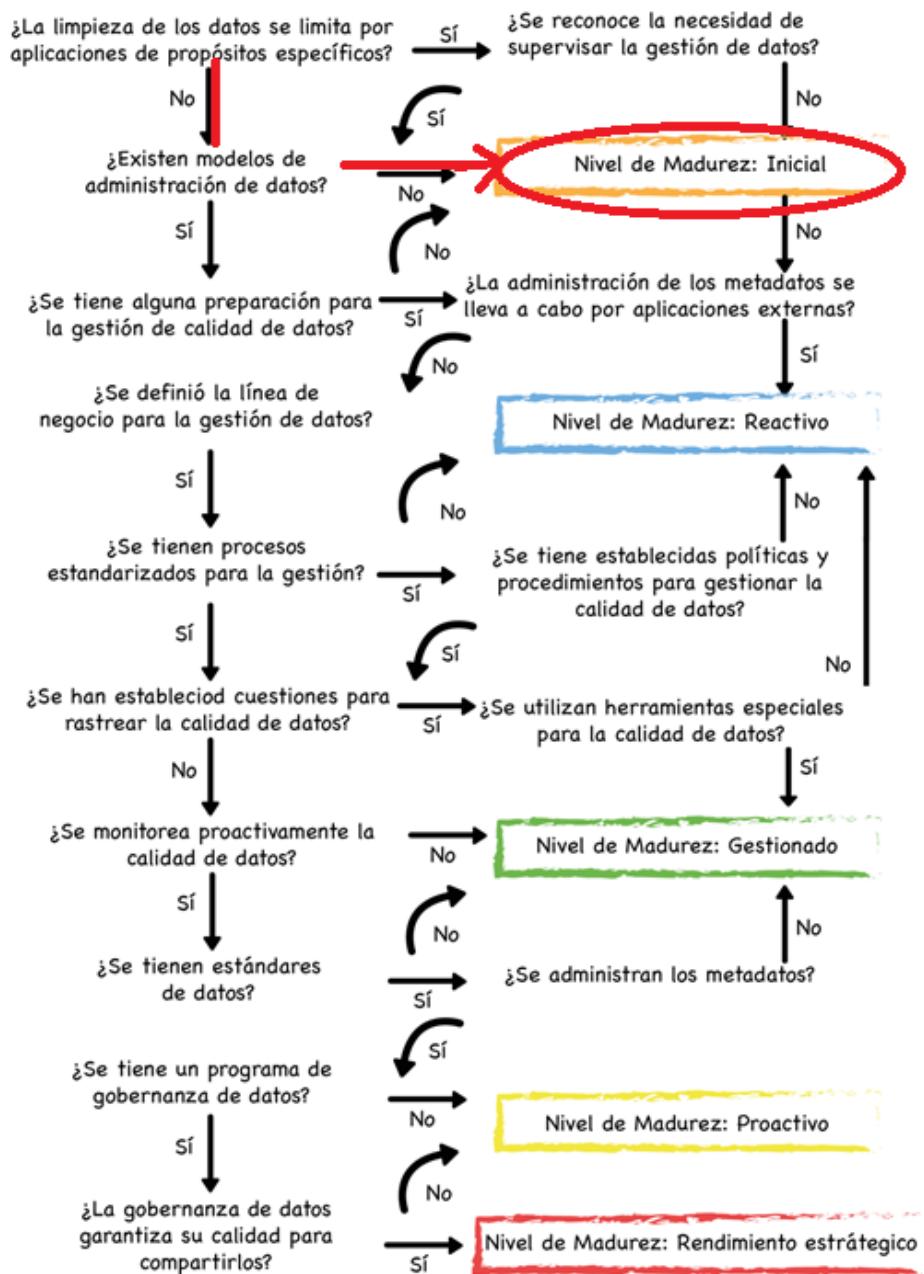


Figura 5.3 Evaluación del componente de gobierno

Las características para determinar el nivel de madurez de este componente son las siguientes:

- No se tiene una cultura, ni mucho menos proceso de realizar la limpieza de datos en un determinado periodo.
- No se han estandarizados los atributos de las entidades maestras como usuarios y proveedores.

- No se tiene una estructura para la custodia de estos datos maestros, sin embargo si se reconoce la necesidad definir un rol.

En este componente no se considera el nivel de madurez dos (Reactivo) por que no se usan aplicaciones externas para administrar los metadatos, no se administra calidad de datos para el análisis, estandarización y consolidación.

### *Componente 3: Administración*

Las capacidades a evaluar como parte de este componente son: identidad, jerarquía, migración y configuración.

Para el área financiera estudiada no se tienen aún desarrolladas ya que como se ha mencionado previamente no se tiene un repositorio único para el almacenamiento de proveedores y usuarios por lo que no se tiene una administración de la identificación única de atributos. El nivel de madurez para este componente es el siguiente:

Nivel de madurez	Inicial
------------------	---------

## Cuestionario de evaluación: Administración

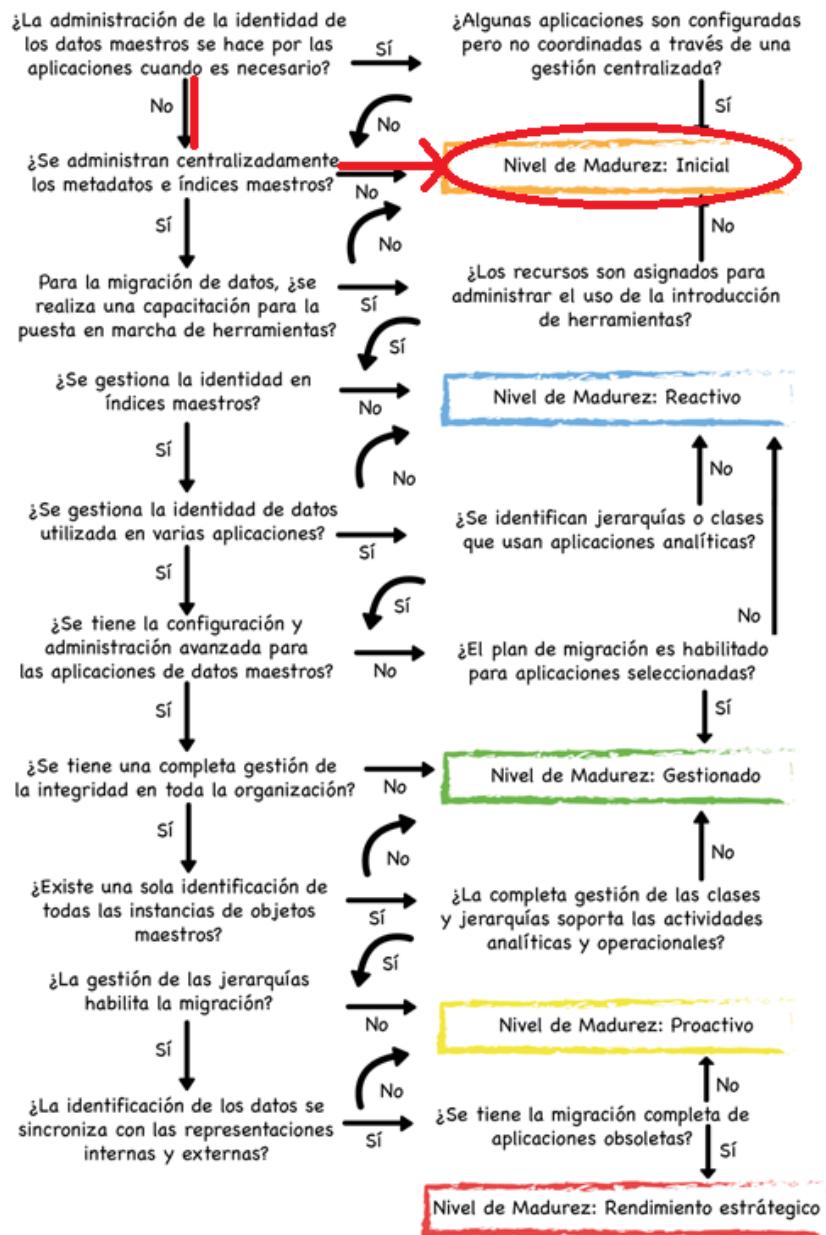


Figura 5.4 Evaluación del componente de administración

Las características para determinar este nivel de madurez de este componente son las siguientes:

- La administración de proveedores y usuarios se hace por cada aplicación cuando es necesario y no por entidades.
- La administración no se coordina través de una administración centralizada.

El nivel de madurez no se puede considerar dos (Reactivo), puesto que no hay roles definidos encargados especialmente para la administración de datos maestros, tampoco se administran centralizadamente los metadatos e índices maestros.

#### *Componente 4: Identificación*

Las capacidades a evaluar como parte de este componente son: búsqueda y correspondencia, vinculación de registros, consolidación de varios atributos.

Para la primera capacidad, en el área financiera de estudio la habilidad para determinar que dos o más representaciones sean resueltas en una única se realiza mediante un archivo Excel al ingresar un nuevo dato, sin embargo esto no garantiza la duplicación o la mala inserción de un proveedor o usuario.

Para la segunda y tercera capacidad, actualmente no se aplican algoritmos de limpieza que establezcan límites para tener conocimiento si se hace referencia a la misma instancia de una entidad maestra.

Por lo tanto, el nivel de madurez para el componente de identificación es:

Nivel de madurez	Inicial
------------------	---------

## Cuestionario de evaluación: Identificación

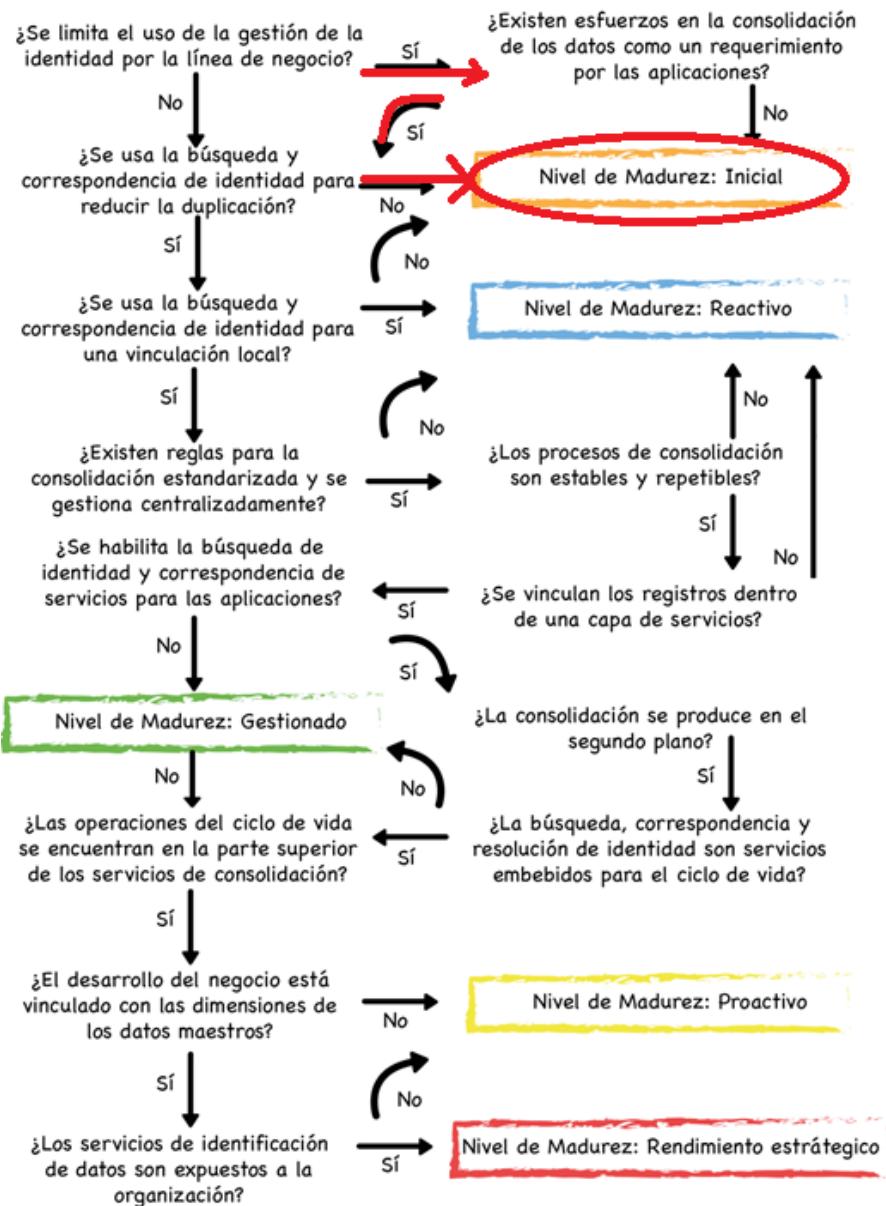


Figura 5.5 Evaluación del componente de Identificación

Las características que fundamentan este nivel de madurez son las siguientes:

- La gestión de los datos de proveedores y usuarios se ha limitado al uso por cada línea de negocio; cada proceso operativo de la dirección.
- Se han hecho intentos para consolidar los datos de los usuarios, sin embargo la falta de planeación de gestión ha hecho que esta información se encuentre obsoleta.

No se considera el nivel dos (Reactivo) porque no se realiza la búsqueda y correspondencia entre los registros para evitar la duplicación de ellos. Tampoco se vinculan los registros cuando se inserta o actualiza un dato maestro.

***Componente 5: Integración***

Las capacidades a evaluar como parte de este componente son: integración de aplicaciones y capa de componentes de servicio.

Respecto a estas capacidades los datos deben ingresar fácilmente al repositorio y su acceso a él debe ser accesible. Para el área financiera de estudio estas capacidades aún no se desarrollan. Por lo que el nivel de madurez es:

Nivel de madurez	Inicial
------------------	---------

## Cuestionario de evaluación: Integración

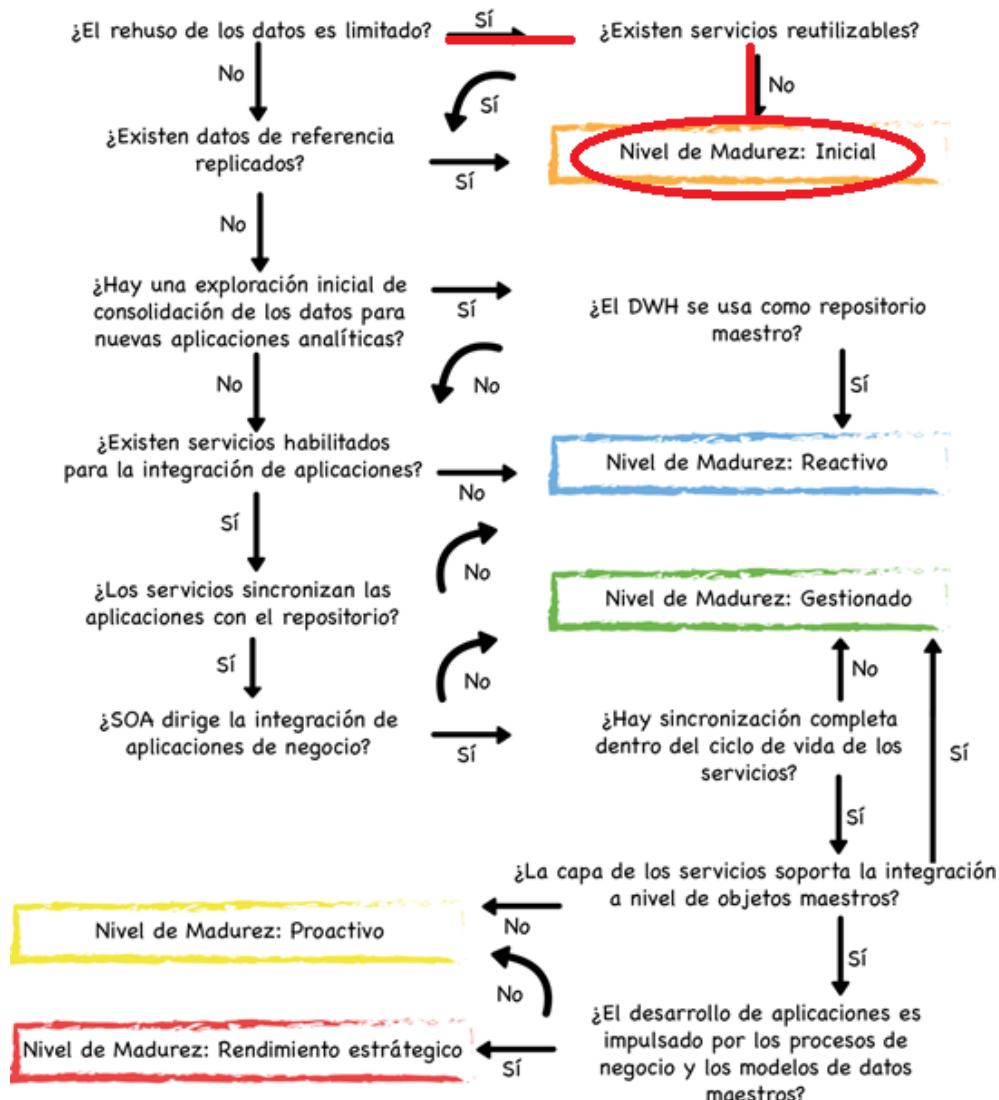


Figura 5.6 Evaluación del componente de integración

Las características para determinar el nivel de madurez de este componente son las siguientes:

- Existen islas de información para los catálogos de proveedores y usuarios en cada proceso operativo del área financiera. No se encuentran unificados ni relacionados entre sí.
- Debido a que estos catálogos no se encuentran unificados ni relacionados no se replican sus actualizaciones mediante catálogos de referencia cruzada.
- No se reutilizan servicios de datos puesto que no se tienen definidos servicios para la transición entre aplicaciones y repositorio.

No se considera el nivel dos (Reactivo) porque no hay una consolidación de los datos, como el desarrollo de las relaciones de los usuarios de la dirección. Físicamente existe una base de datos maestros, sin embargo ese repositorio se ha poblado con catálogos de configuración. En el Data warehouse no existen datos almacenados de proveedores y usuarios.

#### ***Componente 6: Administración de procesos de negocio***

Las capacidades a evaluar como parte de este componente son: integración de procesos de negocio, reglas de negocio y componentes de negocio.

En el área financiera analizada se ha comenzado a desarrollar un modelo de negocio que recopila los procesos que intervienen, así como sus entradas, salidas y eventos. Respecto a las reglas de negocio se ha comenzado con la estandarización de las reglas de negocio utilizadas para sus datos transaccionales no maestros. Por lo tanto, el nivel de madurez para este componente es:

Nivel de madurez	Reactivo
------------------	----------

## Cuestionario de evaluación: Administración de procesos de negocio

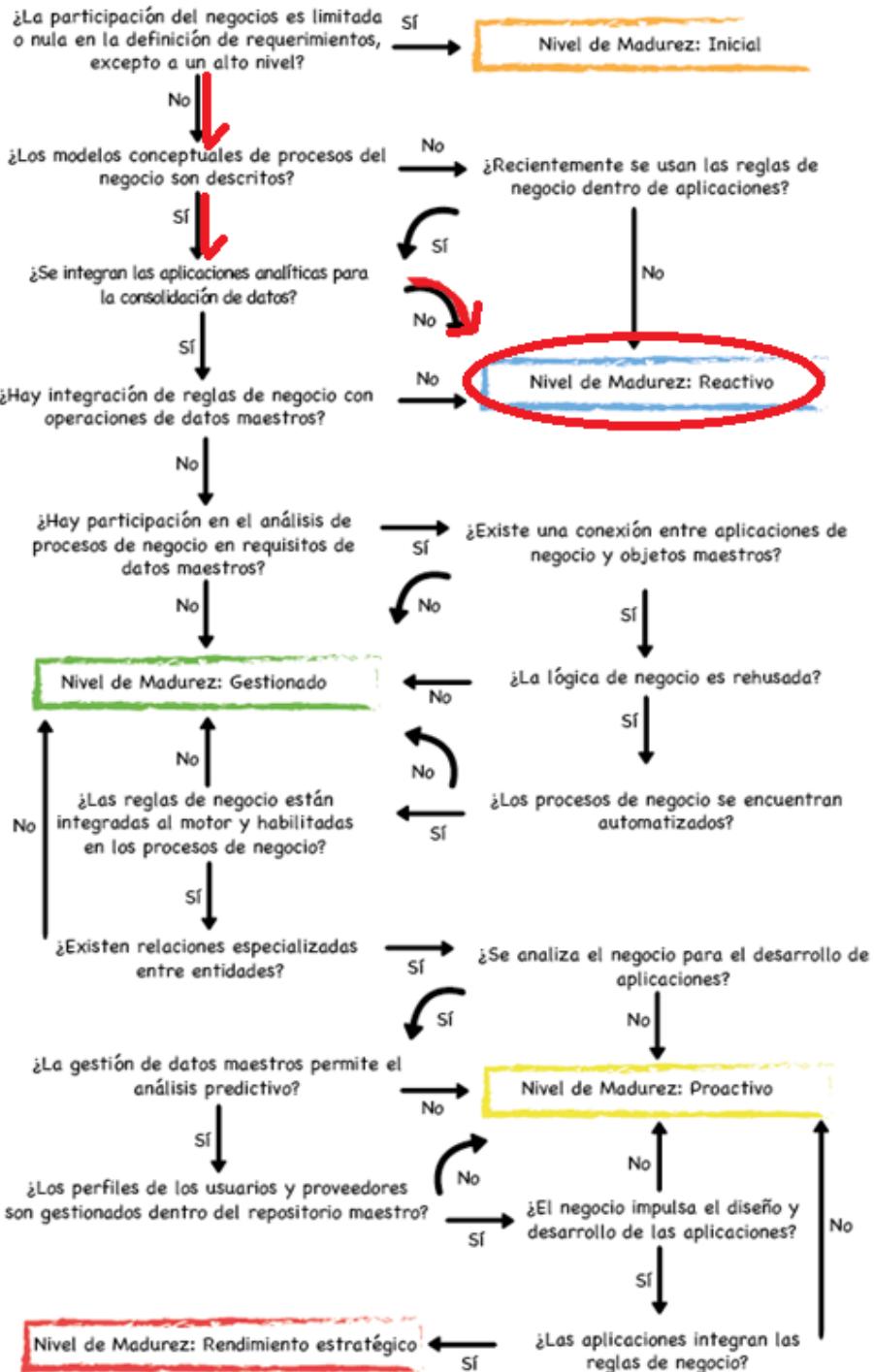


Figura 5.7 Evaluación de admón. de procesos de negocio

Las características para determinar el nivel de madurez de este componente son las siguientes:

- La participación del negocio no es limitado para la definición de los requerimientos.

- Para la mayoría de *Formularios* contenidos en el área financiera se ha comenzado con la descripción de sus modelos de negocio conceptuales.
- Uso de las reglas de negocio dentro de aplicaciones.

Se considera el nivel de madurez dos (Reactivo) porque a pesar de que no se tienen aplicaciones que consoliden los datos maestros, la participación del negocio no es limitado para los requerimientos como se describe en el nivel uno (Inicial).

#### **Resumen de la evaluación**

Mediante esta evaluación se determinó y confirmó el alcance de los objetivos a lograr y la estrategia particular de este caso. Como se ha mencionado previamente la implementación de esta estrategia se realizará a través de un plan de trabajo.

A continuacion se muestra en la siguiente tabla el resumen de la evaluacion de madurez realizada por cada componente para usuarios y proveedores de un área financiera.

<b>Componentes del modelo de madurez</b>	Nivel 1: Inicial	Nivel 2: Reactivo	Nivel 3: Administrado	Nivel 4: Proactivo	Nivel 5: Desempeño estratégico
Arquitectura	X				
Gobierno	X				
Administración	X				
Identificacion	X				
Integracion	X				
Administracion de procesos de negocio		X			

**Tabla 5.1 Resumen de la evaluación**

Como se puede observar en la tabla el nivel de madurez respecto al tratamiento de usuarios y proveedores es el nivel 1: Inicial. Este nivel se caracteriza por la ausencia de capacidades y explotacion de datos maestros, así como la existencia de las multiples copias de un catálogo maestro. Respecto a la gente, no todos lo colaboradores están preparados para consolidar los conjuntos de datos con fines analíticos.

#### **5.2 Fase A: Visión de la arquitectura**

Una vez que se definió el nivel de la situación inicial, a continuación se determinará la misión, visión y estrategia para este caso de estudio.

La misión es: contar con información veraz, completa y homologada de usuarios y proveedores para administrar sus cambios.

La visión de esta arquitectura es: gestionar los datos de usuarios y proveedores en un área financiera de un banco central mediante definiciones estandarizadas, indicadores y métodos aplicables para la medición de la calidad.

### 5.2.1 Nivel de madurez objetivo

Con base en el nivel de madurez evaluado, para cumplir con las necesidades del Área Financiera estudiada se pretende llegar al quinto nivel: Desempeño estratégico. Este nivel tiene las siguientes características por cada componente:

Componente	Capacidades
<b>Architectura</b>	<ul style="list-style-type: none"><li>• La integración de las transacciones se encuentra disponible para aplicaciones internas.</li><li>• La interfaz del API permite hacer el procesamiento directo con el repositorio de los datos maestros.</li></ul>
<b>Gobierno</b>	<ul style="list-style-type: none"><li>• La gobernanza garantiza la calidad de datos para compartirlos.</li></ul>
<b>Administracion</b>	<ul style="list-style-type: none"><li>• La identificación de los datos se sincroniza con las representaciones internas y externas.</li><li>• Migración de aplicaciones obsoletas.</li></ul>
<b>Identificacion</b>	<ul style="list-style-type: none"><li>• Los servicios de identificación de datos son expuestos a la organización.</li><li>• El desarrollo del negocio esta vinculado con las dimensiones de los datos maestros.</li></ul>
<b>Integracion</b>	<ul style="list-style-type: none"><li>• El desarrollo de aplicaciones es impulsado por los procesos de negocio y los modelos de datos maestros.</li></ul>
<b>Administracion de proceso de negocio</b>	<ul style="list-style-type: none"><li>• El negocio impulsa el diseño y desarrollo de las aplicaciones.</li><li>• Las aplicaciones integran las reglas de negocio</li><li>• Las vistas para los proveedores se gestionan dentro del repositorio principal.</li><li>• La gestión de datos maestros permite el análisis predictivo.</li></ul>

Tabla 5.2 Componentes y capacidades del nivel de madurez cinco “Desempeño estratégico”

Este nivel contempla una arquitectura orientada a servicios, tambien abarca el desarrollo ágil de aplicaciones que apoyan la parte operativa y analítica de la empresa. Los colaboradores del negocio ayudarán de definir las expectativas de los procesos para su posterior implementacion.

Por otro lado, la misión del área financiera analizada no se basa en ganancias económicas. Hasta ahora su visión no contempla el análisis de las expectativas del usuario; análisis predictivo. Esta es una única característica del nivel 5 que no se pretende realizar a mediano plazo.

### 5.2.2 Estrategia

Con base en la estrategia definida en el Capítulo 4 en la Fase Visión de la arquitectura y para este caso de estudio, se han elegido para la estrategia dos de las iniciativas propuestas. Se adoptan la misma descripción y objetivo para cada iniciativa definidas previamente.

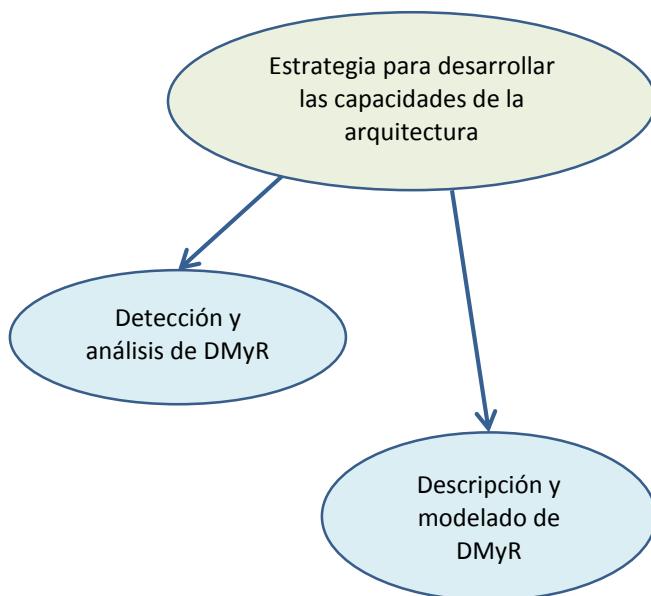


Figura 5.8 Estrategia acotada del caso de estudio

Iniciativa	Objetivo
1. Detección y análisis de datos maestros y de referencia	Recopilar y consolidar los datos maestros y de referencia por medio de artefactos estandarizados en el <i>Área Financiera</i> .
Descripción	<p>Esta iniciativa comprende los proyectos necesarios para recopilar y depurar diferencias sintácticas de los datos maestros. Esto se realizará a través de métodos de búsqueda, correspondencia y limpieza de datos. Esta iniciativa será la base para el desarrollo de la siguiente iniciativa.</p>

Tabla 5.3 Descripción de la Iniciativa 1 “Detección y análisis de datos maestros y de referencia”

Iniciativa	Objetivo
2. Descripción y modelado de datos maestros	Establecer el uso de los datos maestros y de referencia dentro de los procesos de negocio, así como definir una estructura maestra para su gestión central.
Descripción	
Esta iniciativa comprende los proyectos necesarios para desarrollar un modelo conceptual de información por cada proceso de negocio que utilice a los datos maestros y de referencia. Esta iniciativa comprende los proyectos necesarios para generar una estructura estandarizada para su gestión central partiendo de los datos que fueron depurados en la iniciativa anterior. La iniciativa también contempla la definición de los subconjuntos de datos que se tienen que replicar a diferentes procesos del Área Financiera.	

Tabla 5.4 Descripción de la Iniciativa 2 “Descripción y modelado de datos maestros”

### 5.3 Fase B: Arquitectura de negocio

La definición de los principios, procesos, funciones, servicios y roles definidos en el *Capítulo 4* se adoptan para este caso de estudio.

### 5.4 Fase C: Arquitectura de datos

Como se ha especificado en la Visión de la arquitectura el alcance de este caso de estudio no contempla la arquitectura de aplicaciones por lo tanto para este caso de estudio se han omitido los artefactos, diagramas o matrices que involucren la especificación de los componentes lógicos de aplicaciones.

Cabe aclarar que en la Fase F se especifica junto con la estrategia el llenado de cada uno de los artefactos presentados a continuación.

#### 5.4.1 Conceptual

De los artefactos que se diseñaron en la *Arquitectura de datos* del *Capítulo 4*, únicamente se utilizaron los siguientes.

##### *Catálogo de reglas de negocio*

El catálogo de reglas de negocio para las entidades de usuarios son las siguientes:

Clave de la regla	Descripción
DM.USUARIO.01	Un usuario es representado por una entidad financiera, unidad administrativa o proceso operativo.
DM.USUARIO.02	Un sector financiero debe tener al menos una entidad financiera
DM.USUARIO.03	Una entidad financiera pertenece a un sector financiero
DM.USUARIO.04	Una entidad financiera debe tener al menos una unidad administrativa.
DM.USUARIO.05	Una unidad administrativa pertenece a una entidad financiera
DM.USUARIO.06	Una unidad administrativa tiene al menos una persona encargada de ser el contacto como usuario
DM.USUARIO.07	Una persona pertenece al menos a una entidad financiera, por lo tanto pertenece al menos a una unidad administrativa
DM.USUARIO.08	Al menos una persona es el contacto como usuario de un proceso operativo
DM.USUARIO.09	Un contacto debe ser usuario de al menos un proceso operativo
DM.USUARIO.10	Un contacto usuario puede tener al menos un correo electrónico
DM.USUARIO.11	Un contacto usuario puede tener al menos un número telefónico
DM.USUARIO.12	Un contacto usuario puede tener al menos un número de fax
DM.USUARIO.13	Un proceso operativo debe tener por lo menos un usuario
DM.USUARIO.14	Un usuario debe ser alguno de los 2 tipos establecidos en la Área Financiera (operativo, final), pero no puede ser los dos tipos al mismo tiempo.
DM.USUARIO.15	Un proceso operativo puede tener como usuarios operativos a uno o más procesos operativos
DM.USUARIO.16	Un proceso operativo puede ser usuario de más de un proceso operativo.
DM.USUARIO.17	Un usuario operativo debe pertenecer a la AF
DM.USUARIO.18	Un usuario final externo a la Área Financiera se representa por medio de la clave CUIC de la entidad financiera
DM.USUARIO.19	Un usuario final interno pertenece a la Institución pero no al Área Financiera.
DM.USUARIO.20	Para los contactos que sean usuarios operativos o internos, se tomaran los datos que se tienen de los empleados de la institución financiera.

Tabla 5.5 Catálogos de reglas de negocio para Usuarios

El catálogo de reglas de negocio para las entidades de proveedores son las siguientes:

<b>Clave de la regla</b>	<b>Descripción</b>
DM.PROVEEDOR.01	Un proveedor es representado por una entidad financiera, unidad administrativa y/o proceso operativo
DM.PROVEEDOR.02	Una entidad financiera pertenece a un sector financiero
DM.PROVEEDOR.03	Un sector financiero debe tener al menos una entidad financiera.
DM.PROVEEDOR.04	Una entidad financiera debe tener al menos una unidad administrativa.
DM.PROVEEDOR.05	Una unidad administrativa pertenece a una entidad financiera
DM.PROVEEDOR.06	Una unidad administrativa tiene al menos una persona encargada que es el contacto como proveedor
DM.PROVEEDOR.07	Una persona pertenece al menos a una entidad financiera, por lo tanto pertenece al menos a una unidad administrativa
DM.PROVEEDOR.08	Al menos una persona es el contacto que representa a un proveedor.
DM.PROVEEDOR.09	Un contacto debe ser proveedor de al menos un proceso operativo
DM.PROVEEDOR.10	Un proceso operativo puede tener más de un proveedor
DM.PROVEEDOR.11	Un proveedor debe ser de alguno de los tres tipos de proveedores establecidos en la Área Financiera (primario, secundario, operativo) pero no puede ser los tres tipos al mismo tiempo
DM.PROVEEDOR.12	Un proveedor primario/secundario externo se representa por medio de la clave CUIC de la entidad financiera
DM.PROVEEDOR.13	Un contacto que es proveedor primario/secundario externo a la Área Financiera debe tener un correo electrónico
DM.PROVEEDOR.14	Un contacto que es proveedor primario/secundario externo a la Área Financiera debe tener al menos un número telefónico
DM.PROVEEDOR.15	Un contacto que es proveedor primario/secundario externo a la Área Financiera puede tener al menos un número de fax
DM.PROVEEDOR.16	Un proveedor puede ser secundario interno y pertenecer a la Institución siempre y cuando sea externo al Área Financiera.
DM.PROVEEDOR.17	Un proveedor secundario interno a la Área Financiera se representa por la clave Institución Financiera de su unidad administrativa
DM.PROVEEDOR.18	El contacto como proveedor secundario interno será identificado por la clave de empleado de la institución financiera.
DM.PROVEEDOR.19	Un proveedor operativo pertenece a la Área Financiera
DM.PROVEEDOR.20	Un proceso operativo puede tener como proveedores operativos a otros procesos operativos
DM.PROVEEDOR.21	Un proceso operativo puede ser proveedor de más de un proceso operativo

**Tabla 5.6 Catálogos de reglas de negocio para Proveedores**

*Catálogo de las reglas de correspondencia*

Clave	Nombre	Descripción	Entidades que les aplica la regla	% Mínimo de correspondencia entre parámetros	Atributos/Parámetros	Comentarios
<b>1</b>	Personas duplicadas	Analizar un conjunto de atributos que identifican de forma única cada persona que puede ser usuario y/o proveedor.	Artefacto de usuarios y proveedores	80%	Nombres	Ignorar mayúsculas y minúsculas, ignorar acentos, ignorar caracteres sin espacio
					Apellidos	Ignorar mayúsculas y minúsculas, ignorar acentos, ignorar caracteres sin espacio

**Tabla 5.7 Catálogos de reglas de correspondencia**

*Matriz Entidad/ Función de negocio*

	<i>Usuario</i>	<i>Proveedor</i>
<i>Función de Acopio</i>		X
<i>Función de Validación</i>		X
<i>Función de Transformación</i>		X
<i>Función de Diseminación</i>	X	

**Tabla 5.8 Matriz Entidad/Función**

### Diagramas de clases

Para los usuarios su diagrama de clases es el siguiente:

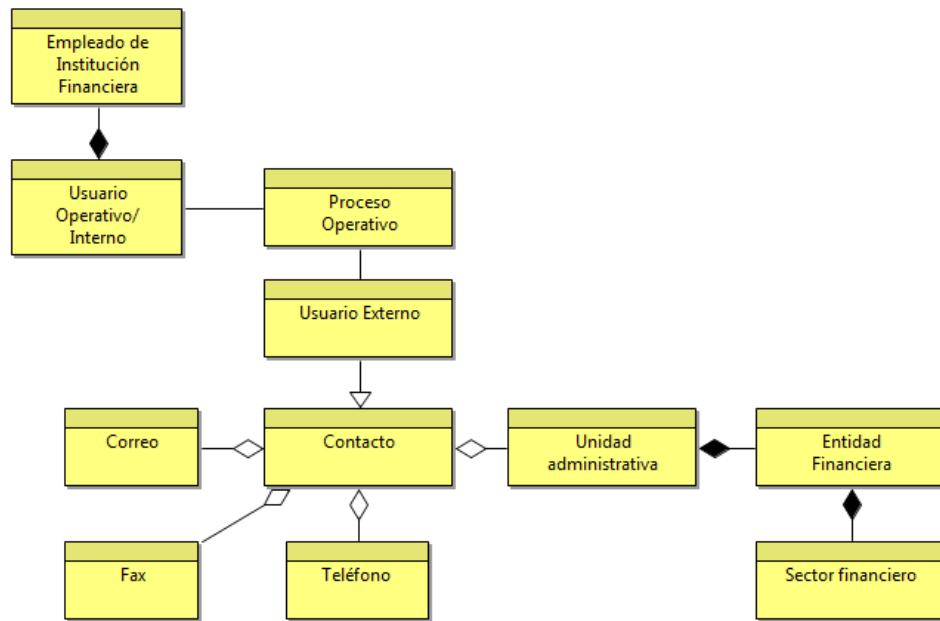


Figura 5.9 Diagrama de clases para Usuarios

Para los proveedores su diagrama de clases es el siguiente:

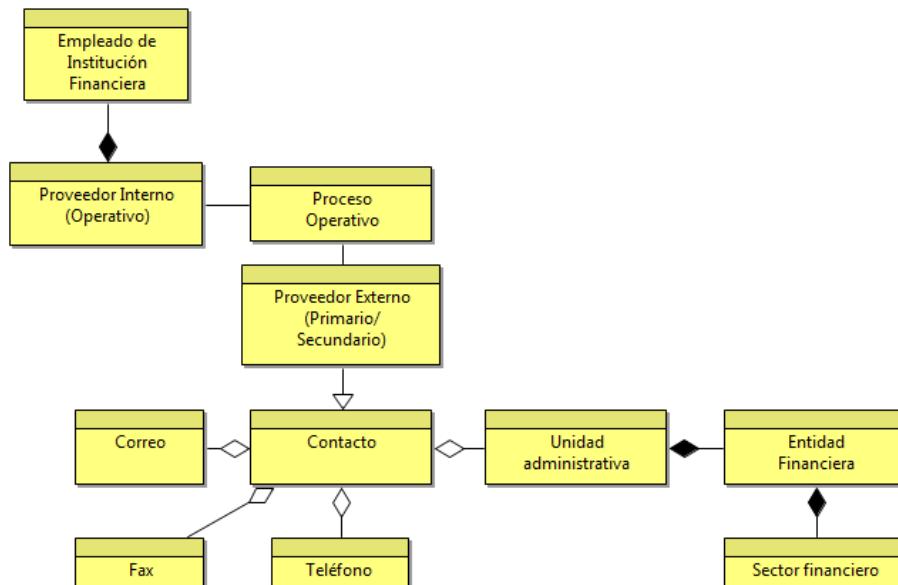


Figura 5.10 Diagrama de clases para Proveedores

## 5.4.1 Lógica

### Catálogo de metadatos

Los metadatos descritos para cada entidad son los siguientes:

Clave del objeto	Nombre del objeto	Descripción	Atributos	Descripción del atributo
AD.DMyR.01	Contacto Externo	Entidad que contiene la información que se forma de la relación que tiene una persona con una institución financiera.	Identificador del contacto externo	Clave numérica única que reconoce cada registro de un contacto.
			Puesto del contacto	Puesto del contacto que ejerce en su unidad administrativa externa.
			Identificador de la Unidad Administrativa.	Clave numérica única que referencia cada unidad administrativa.
			Identificador de la persona externa	Clave numérica única que referencia cada persona.
			Dirección	Ubicación donde ejerce sus responsabilidades.
			Identificador de la entidad financiera	Clave numérica y única que referencia cada entidad financiera.
AD.DMyR.02	Correo Electrónico	Es la tabla que contiene la información de los correos que puede tener el contacto. Esta tabla tiene como atributos el identificador del correo electrónico, el correo electrónico y la clave del contacto al que pertenecen los datos.	Identificador del correo electrónico	Clave numérica única que reconoce cada correo electrónico asociado al contacto.
			Correo electrónico	Correo del contacto.
			Identificador del contacto externo	Clave numérica única que referencia cada registro de un contacto.
AD.DMyR.03	Entidad Financiera	Es el catálogo que contiene las instituciones financieras que componen el sistema financiero	Identificador de la Entidad Financiera	Clave numérica única que reconoce cada registro de una Entidad Financiera.
			Clave de la Entidad Financiera	Clave única de tipo alfanumérica que se asigna dentro del área financiera a cada entidad.
			Nombre de la Entidad Financiera	Nombre largo de la Entidad Financiera.
			Descripción de la entidad financiera	Breve descripción sobre la principal razón de ser de la entidad.
			Dirección Fiscal de la Entidad Financiera	Dirección Fiscal de la Entidad Financiera.
			Clave del tipo de sector financiero	Clave única de tipo alfanumérica que se asigna a cada Sector Financiero. Esta es la clave que referencia a la tabla “Tipo de Sector Financiero”.
AD.DMyR.04	Empleado del área financiera	Entidad que contiene los datos de los colaboradores de la organización, los colaboradores enlistados en la tabla deberán de ser usuarios o proveedores del área financiera	Clave del empleado	Identificador único de cada empleado de la organización. Esta clave se obtendrá del catálogo único que administra el área de recursos humanos.
			Clave de la unidad administrativa	Identificador único de la unidad administrativa de la organización. Este identificador se obtendrá del catálogo que administra el área de recursos humanos.
AD.DMyR.05	Fax	Es la tabla que contiene la información de los números del fax de cada contacto. Esta tabla tiene como atributos el identificador del fax, el número de fax, la extensión del fax y la	Identificador del fax	Clave única que se le asigna a cada número de fax.
			Numero	Número de fax de 8 dígitos.
			Extensión	Extensión numérica del fax.
			Identificador del contacto externo	Clave numérica única que referencia a cada contacto.

		clave del contacto al que pertenecen los datos.		
AD.DMyR.06	Empleado responsable de operar el proceso operativo	Entidad que guarda la relación del responsable de cada formulario del área financiera. La existencia de esta tabla se basa en la regla de negocio que dicta que un colaborador es responsable de un formulario y ese formulario puede tener usuarios que son responsables de otro formulario.	Identificador del operador del área financiera	Clave única que representa cada instancia de las entidades.
			Clave de empleado	Identificador único contenido de cada colaborador de la organización.
			Identificador del proceso operativo	Clave única del formulario del área financiera.
AD.DMyR.07	Persona Externa	Es la entidad que contiene los datos de los actores que participan únicamente por la función de ser proveedor o usuario de los procesos operativos del área financiera.	Identificador de la persona externa	Identificador numérico único que reconoce a cada persona.
			Nombres	Nombre(s) de la persona.
			Apellidos	Apellido paterno y materno de la persona.
AD.DMyR.08	Proceso Operativo	Catálogo que contiene los procesos operativos del área financiera.	Identificador del Proceso Operativo del área financiera	Identificador numérico único que reconoce cada formulario del área financiera.
			Clave del Proceso Operativo	Clave que se asigna dentro de la a cada proceso operativo. Ejemplo: "ACM" (Anexo al Catálogo de Mínimo).
			Tipo del Proceso Operativo	Clasificación del tipo de Proceso Operativo. Este valor debe de ser 1,2 o 3.
			Nombre del Proceso Operativo	Nombre largo del Proceso Operativo
AD.DMyR.09	Proveedor operativo	Entidad que contiene las personas que pertenecen al área financiera y son proveedores operativos.	Identificador del proveedor operativo	Identificador numérico único que reconoce a nivel físico cada registro de un proceso operativo .
			Identificador del proceso operativo	Clave que se asigna dentro de la a cada proceso operativo. Ejemplo: "ACM" (Anexo al Catálogo de Mínimo).
			Identificador del operador del área financiera	Clave única que representa cada instancia de las entidades
AD.DMyR.10	Proveedor interno	Entidad que contiene las personas que pertenecen a la institución financiera y son proveedores internos.	Identificador del proveedor interno	Identificador numérico único que reconoce a nivel físico cada registro de un proceso operativo .
			Clave del empleado del área financiera	Identificador único de cada empleado de la organización. Esta clave se obtendrá del catálogo único que administra el área de recursos humanos.
			Identificador del proceso operativo	Clave que se asigna dentro de la a cada proceso operativo. Ejemplo: "ACM" (Anexo al Catálogo de Mínimo).
AD.DMyR.11	Proveedor externo	Entidad que contiene los contactos externos que son proveedores de información al área financiera.	Identificador del proveedor externo	Identificador numérico único que reconoce a nivel físico cada registro de un proceso operativo .
			Identificador del contacto externo	Clave numérica única que reconoce cada registro de un contacto.
			Identificador del proceso operativo	Identificador numérico único que reconoce cada formulario del área financiera.
AD.DMyR.13	Teléfono		Identificador del teléfono	Clave única que identifica a la tabla.

		Es la tabla que contiene la información de los teléfonos que puede tener el contacto. Esta tabla tiene como atributos el identificador del teléfono, el número de teléfono, la extensión del teléfono y la clave del contacto al que pertenecen los datos	Numero Extensión	Número telefónico. Número de la extensión telefónica.
AD.DMyR.14	Tipo de Sector Financiero	Es el catálogo que contiene la clasificación por sector del conjunto de instituciones que componen el sistema financiero	Identificador del tipo de sector financiero	Identificador numérico único que reconoce a nivel físico cada registro de un sector financiero.
			Clave del tipo de sector financiero	Clave única de tipo alfanumérica que se asigna dentro del área financiera.
			Nombre del sector financiero	Nombre largo del Sector Financiero.
AD.DMyR.15	Usuario Operativo	Es la tabla que contiene la información de la relación que se forma cuando un proceso operativo tiene como usuarios a otros procesos operativos dentro del área financiera.	Identificador del usuario operativo	Clave única
			Identificador del proceso operativo	Identificador numérico único que reconoce cada formulario del área financiera
			Identificador del operador	Clave única que representa cada instancia de las entidades
AD.DMyR.16	Usuario interno	Entidad que representa los usuarios del área financiera; estas contactos son internos a la organización pero externos al área financiera.	Identificador del usuario interno	Clave única que reconoce a cada usuario
			Clave de empleado	Identificador único de cada empleado de la organización.
			Identificador del proceso operativo	Identificador numérico único que referencia cada formulario del área financiera.
AD.DMyR.17	Usuario externo	Entidad que representa la relación de los contactos externos que son usuarios de los formularios del área financiera.	Identificador del usuario externo	Clave única que reconoce cada usuario externo al área financiera.
			Identificador del contacto externo	Clave numérica única que referencia cada registro de un contacto.
			Identificador del proceso operativo	Identificador numérico único que referencia cada formulario del área financiera.
AD.DMyR.18	Unidad Administrativa de Área Financiera	Es el catálogo que contiene las unidades administrativas de las entidades financieras que participan como proveedores o usuarios de los procesos del área financiera	Identificador de la Unidad Administrativa	Identificador numérico único que reconoce cada unidad administrativa que pertenece a la organización donde se encuentra el área financiera.
			Nombre de la Unidad Administrativa	Nombre corto de la Unidad Administrativa de la organización a la que pertenece el área financiera.
AD.DMyR.19	Unidad administrativa externa	Es el catálogo que contiene las unidades administrativas de las entidades financieras que participan como proveedores o usuarios de los procesos del Área Financiera	Identificador de la unidad administrativa externa	Identificador numérico único para representar a cada unidad administrativa de las entidades financieras externas.
			Nombre de la unidad administrativa	Nombre de la Unidad Administrativa de las entidades financieras externas.

Tabla 5.9 Catálogo de metadatos

*Matriz de las relaciones entre objetos de datos*

	<i>Empleado del área financiera</i>	<i>Operador del AF</i>	<i>Proceso operativo</i>	<i>Proveedor operativo</i>	<i>Proveedor interno</i>	<i>Proveedor externo</i>
<i>Empleado del área financiera</i>		X			X	
<i>Operador del Área Financiera</i>	X		X	X		
<i>Proceso operativo</i>		X				X
<i>Proveedor operativo</i>		X	X			
<i>Proveedor interno</i>	X		X			
<i>Proveedor externo</i>			X			

**Tabla 5.10 Matriz de las relaciones entre objetos de datos**

## Modelo de datos maestros y de referencia

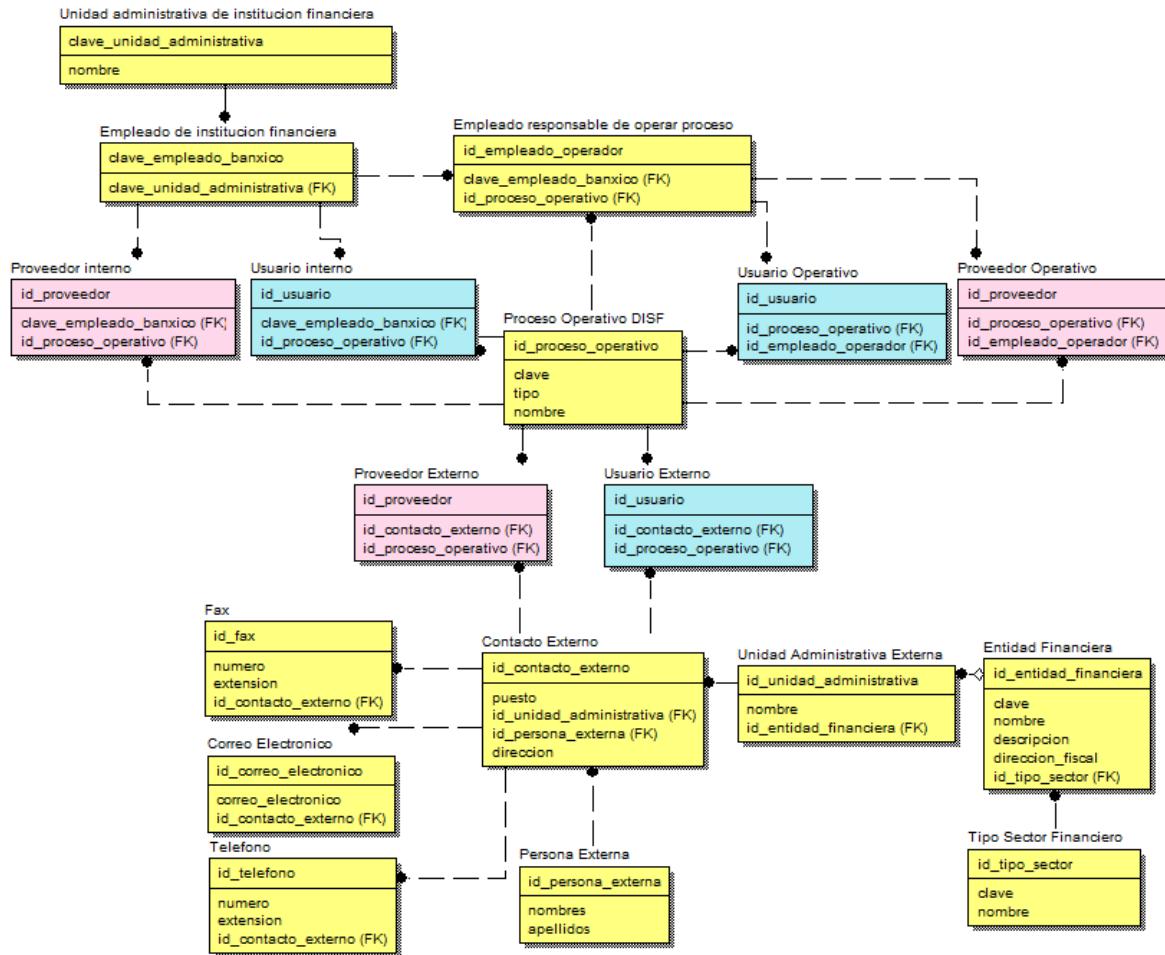


Figura 5.11 Modelo de datos de usuarios y proveedores

## 5.6 Fase E: Oportunidades y soluciones

Las directrices de la estrategia permitirán alcanzar el nivel de madurez objetivo partiendo del nivel actual como se muestra en la siguiente tabla.

Componentes	Nivel 1: Inicial	Nivel 2: Reactivo	Nivel 3: Administrado	Nivel 4: Proactivo	Nivel 5: Desempeño estratégico
<b>Arquitectura</b>	X				X
<b>Gobierno</b>	X				X
<b>Administración</b>	X				X
<b>Identificación</b>	X				X
<b>Integración</b>	X				X
<b>Administración de procesos de negocio</b>		X			X

Tabla 5.11 Nivel inicial y objetivo de la arquitectura de datos

Partiendo de la estrategia desarrollada en el *Capítulo 4, Fase Oportunidades y soluciones*, a continuación se muestran las iniciativas que se cubrirán en este caso de estudio.

<b>Iniciativa 1</b>	
Detección y análisis de datos maestros y de referencia	
Proyectos	Descripción:
1.1 Abstracción y conceptualización de DMyR	Este proyecto contempla las actividades para comprender y establecer la definición del dato maestro o de referencia estudiado.
1.2 Recolección de DMyR	Este proyecto contempla las actividades para definir un artefacto para su recolección y validarlos de acuerdo a su estructura.
1.3 Consolidación del contenido de DMyR	Este proyecto contempla las actividades para realizar un análisis de los datos recopilados del <i>Área Financiera</i> a través de métodos de limpieza de datos.

Tabla 5.12 Proyectos de la Iniciativa 1

<b>Iniciativa 2</b>	
Descripción y modelado de datos maestros y de referencia	
Proyectos	Descripción:
2.1 Definición de las representaciones maestras	Este proyecto contempla las actividades para definir conceptualmente el uso de los datos maestros en el <i>Área Financiera</i> , así como definir las entidades y el contenido de ellas que se tienen que replicar a diversos procesos.
2.2 Integración de las representaciones maestras en un modelo de datos	Este proyecto contempla las actividades para diseñar un modelo central con las entidades establecidas del proyecto anterior.

Tabla 5.13 Proyectos de la Iniciativa 2

## **5.7 Fase F: Planeación de migración**

El plan de trabajo se basa en el que se describió en el Capítulo 4 (*Fase Planeación de migración*) a excepción de algunas actividades que se omitieron. A continuación se describe la realización de cada una de las actividades incluidas en este plan de trabajo de las dos iniciativas que se implementaron en el *Área Financiera*.

### **5.7.1 Iniciativa 1: Detección y análisis de datos maestros y de referencia**

A continuación se expone el plan de trabajo que se siguió para esta iniciativa.

Nombre de tarea	Duración
<b>1. Detección y análisis de datos maestros y de referencia</b>	<b>104 días</b>
<b>1.1 Abstracción y conceptualización de DMyR</b>	<b>20 días</b>
-Definir usuario y clasificaciones	20 días
-Definir proveedor y clasificaciones	20 días
<b>1.2 Recolección de datos</b>	<b>65 días</b>
-Definir el artefacto de recolección de datos de usuarios y proveedores	5 días
<b>1.3 Consolidación del contenido de entidades de usuarios y proveedores</b>	<b>60 días</b>
-Proceso de recolección de datos	60 días
Recopilar los datos para llenar el artefacto	60 días
Almacenar los datos	60 días
Validar los datos para conservar su estandarización recopilada	60 días
-Elegir herramienta para analizar los datos	1 día
-Elegir métodos de depuración de datos	5 días
-Definir parámetros y reglas de correspondencia	3 días
-Aplicar métodos de limpieza de datos	3 días
-Resolver manualmente diferencias lingüísticas	7 días
<b>1.4 Análisis de datos maestros y de referencia</b>	<b>19 días</b>

**Tabla 5.14 Plan de trabajo de la iniciativa 1 “Detección y análisis de datos maestros y de referencia”**

### **Proyecto 1.1 Abstracción y conceptualización de DMyR**

Respecto a la conceptualización de Usuarios y Proveedores a continuación se comienza con su definición y posibles categorías de negocio para posteriormente representarlo en las siguientes arquitecturas.

#### **▪ Definir usuario y clasificaciones**

Se define a un usuario de un área financiera como aquella persona con privilegios, concedidos por el personal autorizado del proceso de que se trate, para consultar o utilizar los productos respectivos. Como ejemplo, un usuario puede ser un funcionario o analista del Banco Central, o un empleado de algún organismo o autoridad financiera nacional (CNBV, SHCP, CONDUSEF, etc.) o extranjera, o el público en general (vía internet).

Como resultado de la revisión conceptual y las clasificaciones establecidas en el artefacto de este elemento, un usuario externo a un área financiera puede pertenecer a una unidad

administrativa, y esta a su vez, pertenece a una entidad financiera, que a su vez pertenece a un sector financiero. Esto manifiesta la importancia de organizar jerárquicamente la información de los usuarios a fin de facilitar su explotación en distintos niveles.

La clasificación para usuarios se estableció en función al tipo de productos que tienen derecho a consumir. Por lo cual, se puede establecer que existen dos tipos de usuarios: finales y operativos. Los usuarios finales pueden o no ser externos al banco central, pero si al área financiera. La jerarquía para este tipo de usuarios es la siguiente:



Figura 5.12 Jerarquía de usuarios finales

Por el otro lado, los usuarios operativos son aquellos que pertenecen al área financiera y para los cuales se propone la siguiente jerarquía:



Figura 5.13 Jerarquía para usuarios operativos

Una de las principales problemáticas que surge en los directorios de usuarios y proveedores, es que estos datos pueden cambiar fácilmente debido a la frecuencia de rotación del personal

en ellos. Por lo cual, no se ve conveniente relacionar los usuarios con los procesos operativos al nivel de contacto. Para ello se propone representar a los usuarios con los siguientes niveles:

- Los usuarios finales de entidades externas al banco central se pueden representar por la entidad financiera.
- Los usuarios finales, pertenecientes al banco central, pero externos al área financiera, se pueden representar por las unidades administrativas a las que pertenecen.
- Finalmente, los usuarios operativos, se pueden identificar por su proceso operativo.

Una esquematización de los niveles propuestos para los usuarios es la siguiente:

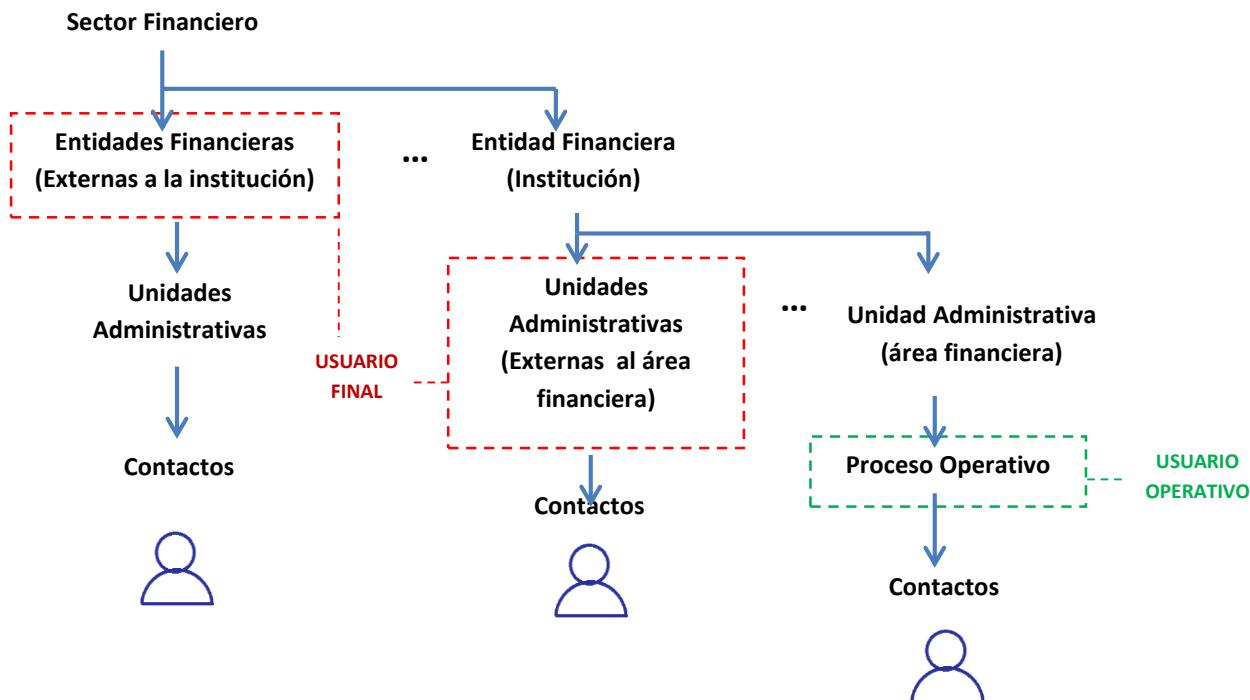


Figura 5.14 Niveles de representación de usuarios

#### ▪ Definir proveedor y clasificaciones

Se define a un proveedor de un banco central como aquella entidad, y su respectivo conjunto de contactos, que suministra de forma recurrente, insumos a cualquiera de los tres principales tipos de procesos operativos del área financiera. Como ejemplo, el contacto de un proveedor puede ser un funcionario o analista de alguna entidad financiera.

Se le conoce como proveedor a toda aquella entidad, y su respectivo conjunto de contactos, que suministra de forma recurrente, insumos a cualquiera de los tres principales tipos de procesos operativos del AF.

Toda entidad proveedor puede clasificarse en sectores según su naturaleza, y cada uno de estos sectores contiene entidades financieras, las cuales a su vez contienen unidades administrativas.

Toda entidad tiene asociados contactos, que son las personas que tienen responsabilidades de proveer información representando a un sector financiero, entidad financiera, unidad administrativa y/o proceso operativo.



Figura 5.15 Relación proveedor-contacto

Los proveedores se clasifican en función del tipo de insumos que suministran y de su pertenencia o no al área financiera en: primarios, secundarios y operativos.

- Los proveedores primarios son todas las entidades financieras externas al banco central, a quienes el área financiera tiene facultad de requerir información conforme a lo señalado en la legislación aplicable vigente. Por lo tanto, estos proveedores se pueden identificar por medio de la entidad financiera.
- Los proveedores secundarios son todas las entidades o autoridades (externas o internas) que, por acuerdo o contrato, comparten información con el área financiera. El proveedor secundario no pertenece a el área financiera pero si puede pertenecer al banco central. Estos proveedores se representan por la entidad financiera si son externos a la el área financiera y si son internos al banco central por la unidad administrativa.

De igual manera estos dos tipos de proveedores, se representan jerárquicamente como la *Figura 5.12 Jerarquía de usuarios*.

- Los proveedores operativos se refieren a otros procesos operativos de la Área Financiera que conforme a sus responsabilidades suministran información necesaria. Estos proveedores se identificaran por los procesos operativos.

De igual manera ese tipo de proveedor se representa jerárquicamente como la *Figura 5.13 Jerarquía para usuarios operativos* de usuarios.

En el siguiente diagrama se muestra el dato jerárquico que representa a los proveedores:

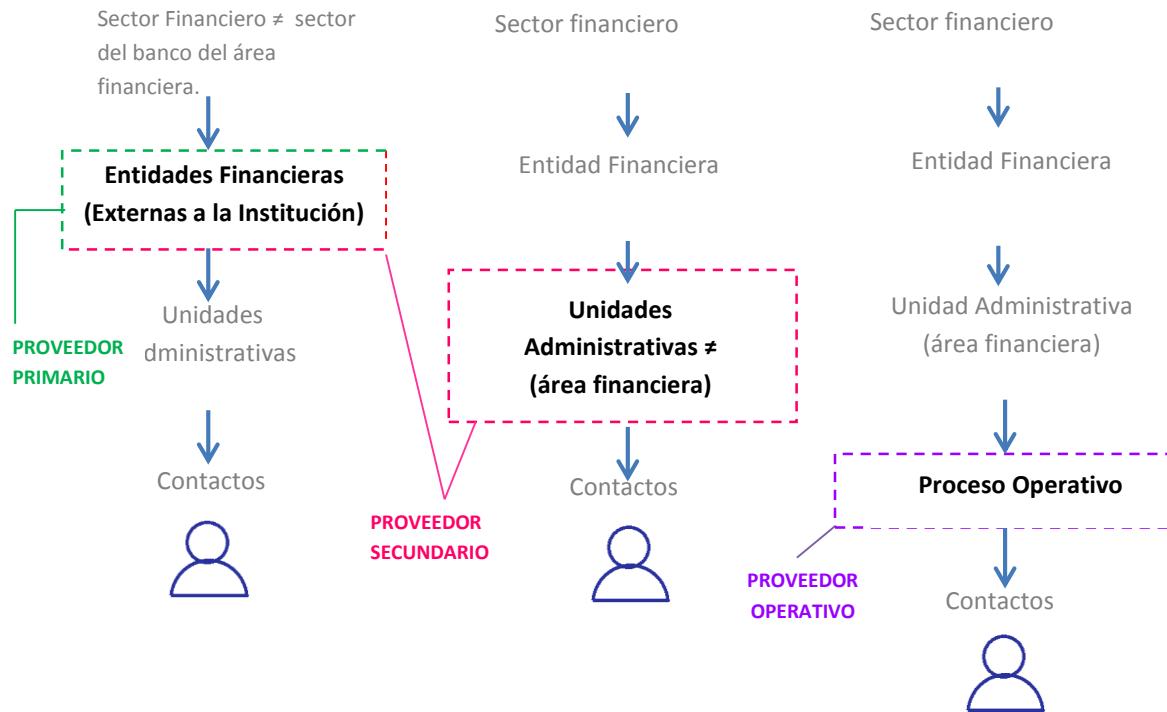


Figura 5.16 Niveles de representación de proveedores

La arquitectura de negocio de este caso de estudio se apoya de los principios, procesos, funciones, servicios y roles definido en el *Capítulo 5, Sección 4.3*.

#### Proyecto 1.2 Recolección de datos

El Área Financiera únicamente cuenta con un catálogo que contiene la clave y nombre de las instituciones financieras. Puesto que no se tienen estructuras de datos detalladas de sus usuarios y proveedores se omitieron las siguientes actividades: *Localizar los datos maestros y de referencia en procesos de negocio y bases de datos, Adoptar algoritmos de búsqueda de datos*. A continuación se describe la realización de las actividades que si se contemplaron en este proyecto.

- **Definir el artefacto de recolección de datos de usuarios y proveedores**

En Octubre de 2012 se comenzó a realizar el levantamiento de la información de proveedores y usuarios por cada *Formularios* del área financiera. Esto se realizó a través de la especificación y adaptación del siguiente artefacto para usuarios y proveedores.

<b>Directorio de usuarios/ proveedores</b>	<b>Metadatos de los atributos</b>
Clave del proceso operativo	Clave del proceso operativo del Área Financiera que genera los productos (conforme al catálogo de Procesos Operativos). Este campo no puede ser vacío y sólo puede tener un valor por registro.
Clave de entidad financiera	Clave de la Entidad a la que pertenece el contacto del usuario de que se trate. Esta clave se obtendrá del catálogo CUIC. Este campo no puede ser vacío.
Identificador de usuario	Clave numérica secuencial que identifica de forma única al contacto del usuario. Este campo no puede ser vacío y la clave no puede estar duplicada para dos o más registros.
Tipo de usuario	Conforme a lo definido en este documento, se deberá indicar el número que corresponda: 1. Usuario final. 2. Usuario operativo. Este campo no puede ser vacío. Sólo aplica un número.
Procesos operativos asociados	Clave del proceso operativo del Área Financiera que consulta o utiliza los productos (Anotar la clave correspondiente conforme al catálogo respectivo). Sólo aplica para usuarios operativos. En caso contrario, anotar NA (No aplica).  En caso que un mismo contacto operativo consulte o utilice dos o más productos que pertenezcan a diferentes procesos operativos, debe registrarse conforme a lo siguiente: NombreProcesoOperativo1 NombreProcesoOperativo2.
Unidad administrativa del contacto	Nombre Oficial completo (en mayúsculas y con acentos) de la Unidad Administrativa a la que pertenece el contacto del usuario. Este campo no puede ser vacío para contactos de usuarios externos de la Institución Financiera. En caso que el contacto sea empleado del Área Financiera debe anotarse NA (No aplica).
Clave del contacto	En caso que el contacto del usuario final u operativo sea empleado de la Institución Financiera, se anotará la clave de empleado correspondiente. En caso contrario, se anotará NA (No aplica).
Nombres	Nombre completo del contacto que utiliza los productos de la área financiera. En caso que el contacto sea empleado de la Institución Financiera debe anotarse NA (No aplica).
Apellido paterno	Apellido paterno en mayúsculas y con acentos del contacto que utiliza los productos del área financiera.
Apellido materno	Apellido materno en mayúsculas y con acentos del contacto que utiliza los productos del área financiera.
Puesto del contacto	Indicar el puesto (en mayúsculas y con acentos) que ocupa el contacto del usuario, dentro de la unidad administrativa a la que pertenece. En caso que el contacto sea empleado de la Institución Financiera debe anotarse NA (No aplica).

**Tabla 5.15 Artefacto para la recolección de usuarios y proveedores**

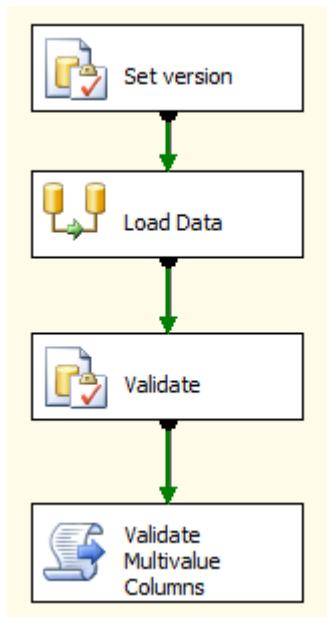
Correos electrónicos	Correo electrónico del contacto del usuario. Anotar sólo caracteres alfanuméricos sin espacios. Cuando existan varias cuentas de correo electrónico para un contacto, el formato a seguir para su documentación será: eMail1 eMail2. Por ejemplo, <a href="mailto:usuario@banamex.com">usuario@banamex.com</a>   <a href="mailto:usr@banamex.com">usr@banamex.com</a> . En caso que no exista cuenta de correo se anotará ND (No disponible). En caso que el contacto sea empleado de la Institución Financiera debe anotarse NA (No aplica).
Teléfonos del contacto	<p>Teléfono y extensión del contacto del usuario. Anotar sólo números sin espacios. Cuando el número telefónico a reportar contenga un número de extensión, el formato a seguir para su documentación será: NúmeroTelefónico.NúmeroExtensión. Por ejemplo, 9838374738.3373. Cuando existan varios números telefónicos a reportar, el formato a seguir para su documentación será: NúmeroTelefónico1.Extensión1  NúmeroTelefónico2.Extensión2. Por ejemplo, 9838374738.3373 12121213.556.</p> <p>En caso que un contacto tenga dos o más extensiones, con el mismo número de conmutador, el formato a seguir será: NúmeroTelefónico1.Extensión1 NúmeroTelefónico1.Extensión2. En caso que el contacto sea empleado de la Institución Financiera debe anotarse NA (No aplica).</p> <p>Nota: El carácter pipe ( ) se localiza en la tecla del lado izquierdo del número 1.</p>
Fax del contacto	<p>Fax del contacto del usuario. Anotar sólo números sin espacios. Cuando el número de fax a reportar contenga una extensión, el formato a seguir para su documentación será: NúmeroFax.Extensión. Por ejemplo, 9838374738.3373. Cuando existan varios números de fax a reportar, el formato a seguir para su documentación será: NúmeroFax1.Extensión1 NúmeroFax2.Extensión2. Por ejemplo, 9838374738.3373  12121213.556. En caso que no exista número de fax se anotará NA (No aplica).</p> <p>En caso que un contacto tenga dos o más extensiones, con el mismo número de conmutador, el formato a seguir será: NúmeroTelefónico1.Extensión1 NúmeroTelefónico1.Extensión2.</p> <p>En caso que el contacto sea empleado de la Institución Financiera debe anotarse NA (No aplica).En caso que no exista número de fax se anotará ND (No disponible).</p>
Dirección	Dirección del contacto del usuario. En caso que el contacto sea empleado de la Institución Financiera debe anotarse NA (No aplica). En caso que se desconozca la dirección se anotará ND (No disponible).

## ■ Proceso de recolección de datos

Este proceso se aplicó iterativamente debido a que a través de varias revisiones se recopiló la información con los responsables del negocio.

- ✓ Recopilar los datos para llenar el artefacto. Esta información se recabó en un archivo Excel para cada uno de los 32 *Formularios* del AF, teniendo cada uno de ellos su propio responsable de información.
- ✓ Almacenar los datos. Los datos que se iban recopilando se almacenaron en el manejador de bases de datos SQL Server. En este manejador se crearon dos tablas, una para la información de usuarios y otra tabla para la información de proveedores. Estas tablas no se normalizaron ya que el fin del almacenamiento de estos datos es validarlos y posteriormente depurarlos.
- ✓ Validar los datos para conservar su estandarización recopilada. Para la validación de los datos recabados se eligió la herramienta SQL Server Integration Services. En esta herramienta se desarrolló un ETL para validar principalmente los siguientes puntos:
  - Todos los atributos del artefacto no deberán tener valores (no nulos ni vacíos).
  - Los atributos de referencia del artefacto deberán existir en los catálogos de respectivos.
  - No deberá existir duplicidad en los identificadores únicos.

A continuación se muestra el flujo de control que se definió en el ETL para la implementación de estas validaciones.



**Figura 5.17 Flujo de control para las validaciones de usuarios y proveedores**

- Set version. Es el script SQL que establece un control de versión para el artefacto.
- Load Data. Flujo de datos que extrae la información del artefacto desde el archivo de Excel y lo carga en la BD staging (Staging). Omite filas nulas o vacías del Excel.
- Validate. Es el script SQL que se encarga de validar individualmente cada registro del artefacto de usuarios. Si existe algún error en la validación, se lleva a una bitácora.  
Para llevar un control de los errores cometidos anteriormente se desarrolló un script que además de ejecutar las respectivas validaciones cada vez que detecte un error este se muestre para su corrección, creando una nueva versión a la siguiente ejecución. De esta forma se garantiza la mejora de la recolección de datos y el histórico de archivos que se llevó acabo en el proceso.
- Validate Multivalue Columns. Es el script .NET que se encarga de validar columnas con datos multivalores, separados por un carácter especial “|” (pipe). Se prefirió usar un script .NET en vez de SQL para un mejor performance y evitar el manejo de cursores en SQL.

#### **Proyecto 1.3 Consolidación del contenido de entidades de usuarios y proveedores**

Este proyecto contempla las actividades para realizar un análisis de los datos recopilados del Área Financiera a través de métodos de limpieza de datos.

- **Elegir herramienta para analizar los datos**

Las herramientas disponibles en el Área Financiera para consolidar los datos de usuarios y proveedores son SSIS y BIDS.

- Integration Services o sus siglas en SSIS es una plataforma para la integración de datos. Sirve para resolver problemas mediante la copia o descarga de archivos, la actualización de almacenamientos de datos, la limpieza y minería de datos, la administración de datos de SQL Server () y otros proveedores. Las tareas del flujo de datos permite hacer las siguientes transformaciones entre otras: Conditional Split, Multicast, Union-All, Merge, and Merge Join, Sort, Fuzzy Grouping.
- Business Intelligent Development Studio o sus siglas BIDS es el entorno de desarrollo para soluciones y proyectos de Integration Services, Analysis Services y Reporting Services, disponiendo de plantillas de proyecto específicas (Microsoft, 2011). Principalmente es una interfaz para crear paquetes ETL haciendo uso de SSIS.

#### ▪ **Elegir métodos de depuración de datos**

Los métodos de depuración de datos principalmente se establecen dentro de la configuración de la herramienta. En el caso de SSIS es posible utilizar la transformación Fuzzy Grouping, el cual permite agrupar datos por una similitud configurable. Esta similitud se logra a través de técnicas de lógica difusa para la limpieza de datos.

La lógica difusa es una extensión de la lógica booleana, utiliza conceptos de pertenencia de conjuntos. La lógica difusa es una forma de multi-valuada lógica; se trata de un razonamiento que es aproximado y no fijo y exacto. En comparación con los conjuntos binarios tradicionales, las variables de lógica difusa pueden tener un valor de verdad que oscila en grados entre 0 y 1.

#### ▪ **Definir parámetros y reglas de correspondencia**

Los parámetros a utilizar para establecer la coincidencia entre registros similares son los atributos de nombre y apellido. En el siguiente artefacto “*Catálogo de reglas de correspondencia*”, el cual fue definido en la *Arquitectura de datos* se muestra la regla de negocio utilizada para esta iniciativa.

Clave	Nombre	Descripción	Entidades que les aplica la regla	% Mínimo de correspondencia entre parámetros	Atributos/Parámetros	Comentarios
1	Personas duplicadas	Analizar un conjunto de atributos que identifican de forma única cada persona que puede ser usuario y/o proveedor.	Artefacto de usuarios y proveedores	80%	Nombres	Ignorar mayúsculas y minúsculas, ignorar acentos, ignorar caracteres sin espacio
					Apellidos	Ignorar mayúsculas y minúsculas, ignorar acentos, ignorar caracteres sin espacio

Tabla 5.16 Regla de correspondencia de usuarios y proveedores

El % mínimo de coincidencia entre los atributos “Nombre” y “Apellido” será de 80%. Esto quiere decir que se pueden presentar coincidencias en los siguientes casos:

- A) Nombres 60% y apellidos 100%
- B) Nombre 80% y apellidos 80%
- C) Nombres 100% y apellidos 60%

Se estableció el 80% porque este valor garantiza coincidencias seguras. Este valor pueden excluir algunos registros que en realidad pertenezcan a un conjunto X de datos, sin embargo el trabajo de ubicar a estos registros será mínimo en comparación con obtener agrupaciones no reales de datos producidos por un valor más bajo del % mínimo de coincidencias, ejemplo 60%, 70%.

- **Aplicar métodos de limpieza de datos**

Una vez que se han establecido los valores de coincidencia se implementó el siguiente flujo de datos en la herramienta BIDS.

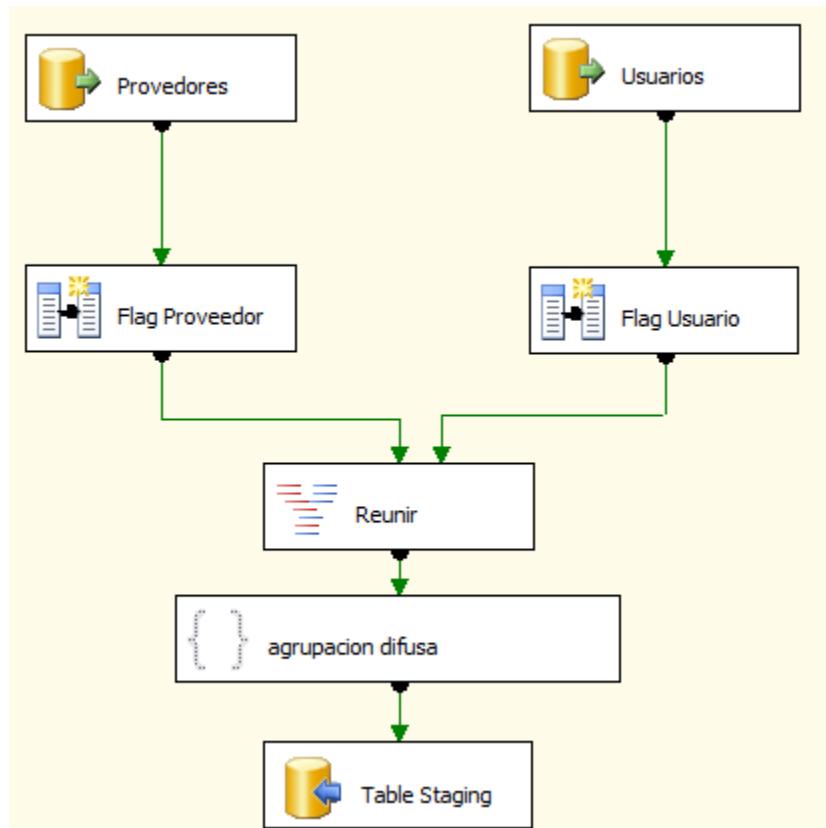


Figura 5.18 Flujo de datos para la depuración de usuarios y proveedores

Cada caja de este ETL representa lo siguiente:

- ✓ Proveedores- Usuarios. Se toman los datos del artefacto correspondiente en la base de datos donde se almacenaron.
- ✓ Flag Proveedor- Flag Usuario. Se añadieron indicadores (flags) para identificar el origen del artefacto de cada registro.
- ✓ Reunir. Tanto los datos recopilados como usuarios y proveedores se unen en un conjunto de datos ya que como se mencionó anteriormente un usuario del Área Financiera también puede ser proveedor de la misma.
- ✓ Agrupación difusa. Se ejecuta la función Fuzzy Grouping de SSIS conforme a los valores establecidos.
- ✓ Table Staging. Una vez que se obtuvieron las agrupaciones de los datos de usuarios y proveedores se exportan a una tabla.

- **Resolver manualmente diferencias lingüísticas**

El resultado de la aplicación de lógica difusa da como resultado agrupaciones de posibles coincidencias de datos. Esta actividad consiste en la revisión humana de dichas agrupaciones para su aprobación o posible reubicación de las instancias de usuarios y proveedores.

En el siguiente caso se muestra la agrupación del mismo registro en dos agrupaciones. Cabe aclarar que se recortaron atributos de la siguiente vista para fines de una representación más clara.

Agrupación de Registro Coincidente	Nombres	Apellido Paterno	%Coincidencia	Nombre de la Entidad Financiera	Tipo	Teléfono	Correo	Fax
121	JUAN C.	AGUILA R	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	54512565
121	JUAN C.	AGUILA R	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	54512565
121	JUAN C.	AGUILA R	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	54512565
122	JUAN CARLOS	AGUILA R	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	55585666 54512565
122	JUAN CARLOS	AGUILA R	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	55585666
122	JUAN CARLOS	AGUILA R	100.00%	CB ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	55585666
122	JUAN CARLOS	AGUILA R	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	juanito@zign.com.mx	55585666

**Tabla 5.17 Resultados de la depuración de datos**

Al identificar manualmente que los anteriores registros pertenecen a la misma persona “Juan Carlos Aguilar”, se observa que esta persona colabora dos entidades financieras: BANCA ZIG y Casa de Bolsa ZIG. Con base en la abstracción y conceptualización de usuarios y proveedores que se realizó en la Iniciativa 1 se han seleccionado los siguientes registros:

Agrupación de Registro	Nombre s	Apellido Paterno	%Coincidencia	Nombre de la	Tipo	Teléfono	Correo	Fax

Coincidente				Entidad Financiera					
121	JUAN C.	AGUILAR	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	54512565	X
121	JUAN C.	AGUILAR	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	54512565	X
121	JUAN C.	AGUILAR	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	54512565	X
121	JUAN CARLOS	AGUILAR	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	55585666 54512565	<input type="checkbox"/>
121	JUAN CARLOS	AGUILAR	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	55585666	X
121	JUAN CARLOS	AGUILAR	100.00%	CB ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	55585666	<input type="checkbox"/>
121	JUAN CARLOS	AGUILAR	100.00%	BANCA ZIG MEXICO	Proveedor	58954865	Juanito@zig.com.mx	55585666	X

Tabla 5.18 Selección de registros

### 5.7.2 Iniciativa 2: Descripción y modelado de datos maestros y de referencia

Partiendo de que el Área Financiera tiene estandarizados sus procesos de negocio y contemplando que estas iniciativas pretenden representar, almacenar y usar estandarizadamente los datos de usuarios y proveedores no se contempla en este caso de estudio la actividad “Definir el subconjunto acotado de atributos y registros para cada requerimiento de negocio”. Esta actividad aplica sin excepción para los datos de referencia y en algunos casos puede aplicar para los datos maestros, pero este no es el caso.

El plan de trabajo que se siguió para esta iniciativa es el siguiente:

Nombre de tarea

Duración

<b>2. Descripción y modelado de datos maestros</b>	<b>61 días</b>
<b>2.1 Definición de la representación de usuarios y proveedores</b>	<b>30 días</b>
-Definir su ubicación dentro de los procesos de negocio	15 días
-Desarrollar un modelo de información	15 días
<b>2.2 Integración de las representaciones maestras en un modelo de datos</b>	<b>21 días</b>
-Definir reglas de negocio para modelar	5 días
- Desarrollar un diagrama de clases donde se reflejen conceptualmente las reglas descritas anteriormente	3 días
-Desarrollar el modelo de datos mediante un lenguaje que incluya taxonomías, ontologías u otras representaciones conceptuales y lógicas de alto nivel	5 días
-Definir sus metadatos	2 días
-Establecer las relaciones con otras entidades maestras	3 días
-Implementar el modelo maestro en un repositorio	3 días

Tabla 5.19 Plan de trabajo para la iniciativa 2 “Descripción y modelado de datos maestros y de referencia”

#### *Proyecto 2.1 Definición de las representación de usuarios y proveedores*

Una vez analizados los datos y obtenidos los valores reales de usuarios y proveedores se desarrollaron las siguientes actividades:

- **Definir su ubicación dentro de los procesos de negocio**

¿En qué procesos del Área Financiera se utilizan estos datos? Para representar la respuesta de esta pregunta se ha utilizado el artefacto de la Arquitectura de datos “*Matriz Entidad/Función de negocio*”, en donde se han enlistado las principales funciones de negocio del Área Financiera que son acopio, validación, transformación y diseminación.

Como se puede ver en la siguiente matriz la función de acopio, validación y transformación hace uso de los datos del proveedor para recibir, verificar y generar productos de información. La función de negocio –diseminación- deberá considerar los datos del usuario para la publicación de los productos de información generados.

Estas relaciones se acordaron con los interesados y responsables de negocio del AF.

	<i>Usuario</i>	<i>Proveedor</i>
<i>Función de Acopio</i>		X
<i>Función de Validación</i>		X
<i>Función de Transformación</i>		X
<i>Función de Diseminación</i>	X	

Tabla 5.20 Matriz Entidad/Función

- **Desarrollar un modelo de información**

Una vez que se ha establecido la ubicación de estas representaciones maestras en los principales procesos del Área Financiera, para cada uno de los *Formularios* se desarrolló un modelo de información como el que se muestra a continuación.

- El siguiente ejemplo representa el *Formulario* denominado Activos y Pasivos en Moneda Extranjera y Posición en Divisas, se muestra la entidades proveedoras de información a este formulario y las entidades usuarias.

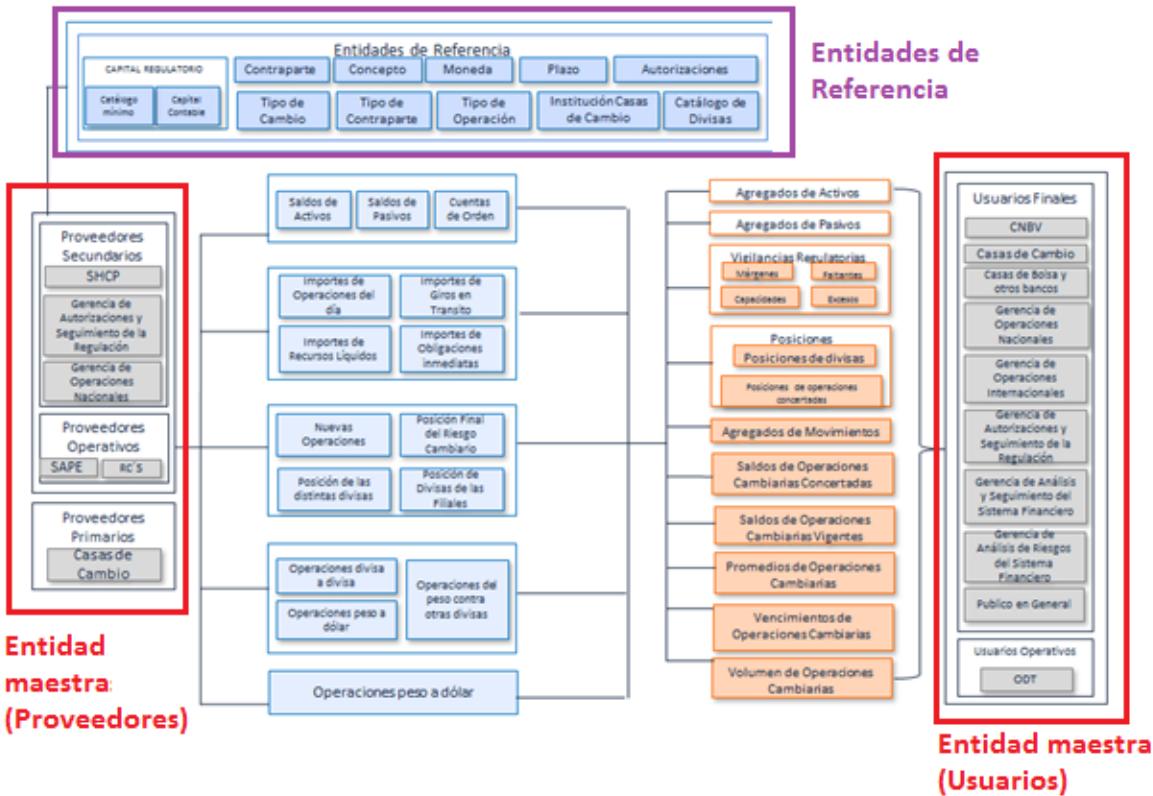


Figura 5.19 Modelo de información

### Proyecto 2.2 Integración de las representaciones maestras en un modelo de datos

Este proyecto contempla las actividades para diseñar un modelo central con las entidades establecidas del proyecto anterior. A continuación se describe el desarrollo de cada una de las siguientes actividades.

- **Definir reglas de negocio para modelar**

El desarrollo de la iniciativa 1 y el proyecto 2.1 de la Iniciativa 2 hacen posible la definición puntual de las reglas de negocio para el modelo de usuarios y proveedores. Se utilizó el artefacto definido en la arquitectura de datos denominado “Catálogo de reglas de negocio”.

El catálogo de reglas de negocio para las entidades de usuarios son las siguientes:

Clave de la regla	Descripción
-------------------	-------------

DM.USUARIO.01	Un usuario es representado por una entidad financiera, unidad administrativa o proceso operativo.
DM.USUARIO.02	Un sector financiero debe tener al menos una entidad financiera
DM.USUARIO.03	Una entidad financiera pertenece a un sector financiero
DM.USUARIO.04	Una entidad financiera debe tener al menos una unidad administrativa.
DM.USUARIO.05	Una unidad administrativa pertenece a una entidad financiera
DM.USUARIO.06	Una unidad administrativa tiene al menos una persona encargada de ser el contacto como usuario
DM.USUARIO.07	Una persona pertenece al menos a una entidad financiera, por lo tanto pertenece al menos a una unidad administrativa
DM.USUARIO.08	Al menos una persona es el contacto como usuario de un proceso operativo
DM.USUARIO.09	Un contacto debe ser usuario de al menos un proceso operativo
DM.USUARIO.10	Un contacto usuario puede tener al menos un correo electrónico
DM.USUARIO.11	Un contacto usuario puede tener al menos un número telefónico
DM.USUARIO.12	Un contacto usuario puede tener al menos un número de fax
DM.USUARIO.13	Un proceso operativo debe tener por lo menos un usuario
DM.USUARIO.14	Un usuario debe ser alguno de los 2 tipos establecidos en la Área Financiera (operativo, final), pero no puede ser los dos tipos al mismo tiempo.
DM.USUARIO.15	Un proceso operativo puede tener como usuarios operativos a uno o más procesos operativos
DM.USUARIO.16	Un proceso operativo puede ser usuario de más de un proceso operativo.
DM.USUARIO.17	Un usuario operativo debe pertenecer a la AF
DM.USUARIO.18	Un usuario final externo a la Área Financiera se representa por medio de la clave CUIC de la entidad financiera
DM.USUARIO.19	Un usuario final interno pertenece a la Institución pero no al Área Financiera.
DM.USUARIO.20	Para los contactos que sean usuarios operativos o internos, se tomarán los datos que se tienen de los empleados de la institución financiera.

Tabla 5.21 Reglas de negocio para usuarios

El catálogo de reglas de negocio para las entidades de proveedores son las siguientes:

Clave de la regla	Descripción
-------------------	-------------

DM.PROVEEDOR.01	Un proveedor es representado por una entidad financiera, unidad administrativa y/o proceso operativo
DM.PROVEEDOR.02	Una entidad financiera pertenece a un sector financiero
DM.PROVEEDOR.03	Un sector financiero debe tener al menos una entidad financiera.
DM.PROVEEDOR.04	Una entidad financiera debe tener al menos una unidad administrativa.
DM.PROVEEDOR.05	Una unidad administrativa pertenece a una entidad financiera
DM.PROVEEDOR.06	Una unidad administrativa tiene al menos una persona encargada que es el contacto como proveedor
DM.PROVEEDOR.07	Una persona pertenece al menos a una entidad financiera, por lo tanto pertenece al menos a una unidad administrativa
DM.PROVEEDOR.08	Al menos una persona es el contacto que representa a un proveedor.
DM.PROVEEDOR.09	Un contacto debe ser proveedor de al menos un proceso operativo
DM.PROVEEDOR.10	Un proceso operativo puede tener más de un proveedor
DM.PROVEEDOR.11	Un proveedor debe ser de alguno de los tres tipos de proveedores establecidos en la Área Financiera (primario, secundario, operativo) pero no puede ser los tres tipos al mismo tiempo
DM.PROVEEDOR.12	Un proveedor primario/secundario externo se representa por medio de la clave CUIC de la entidad financiera
DM.PROVEEDOR.13	Un contacto que es proveedor primario/secundario externo a la Área Financiera debe tener un correo electrónico
DM.PROVEEDOR.14	Un contacto que es proveedor primario/secundario externo a la Área Financiera debe tener al menos un número telefónico
DM.PROVEEDOR.15	Un contacto que es proveedor primario/secundario externo a la Área Financiera puede tener al menos un número de fax
DM.PROVEEDOR.16	Un proveedor puede ser secundario interno y pertenecer a la Institución siempre y cuando sea externo al Área Financiera.
DM.PROVEEDOR.17	Un proveedor secundario interno a la Área Financiera se representa por la clave Institución Financiera de su unidad administrativa
DM.PROVEEDOR.18	El contacto como proveedor secundario interno será identificado por la clave de empleado de la institución financiera.
DM.PROVEEDOR.19	Un proveedor operativo pertenece a la Área Financiera
DM.PROVEEDOR.20	Un proceso operativo puede tener como proveedores operativos a otros procesos operativos
DM.PROVEEDOR.21	Un proceso operativo puede ser proveedor de más de un proceso operativo

Tabla 5.22 Reglas de negocio para proveedores

- Desarrollar un diagrama de clases donde se reflejen conceptualmente las reglas descritas anteriormente

Para el desarrollo de esta actividad se usó el artefacto definido en la arquitectura de datos denominado “Diagrama de clases”. Con base en las reglas de negocio descritas anteriormente se obtuvieron los siguientes diagramas de clases.

Para los usuarios su diagrama de clases es el siguiente:

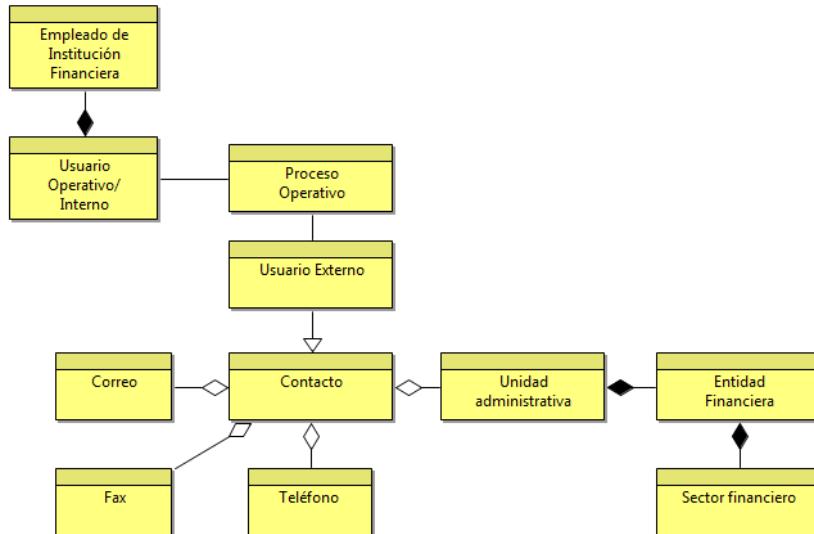


Figura 5.20 Diagramas de clases para usuarios

Para los proveedores su diagrama de clases es el siguiente:

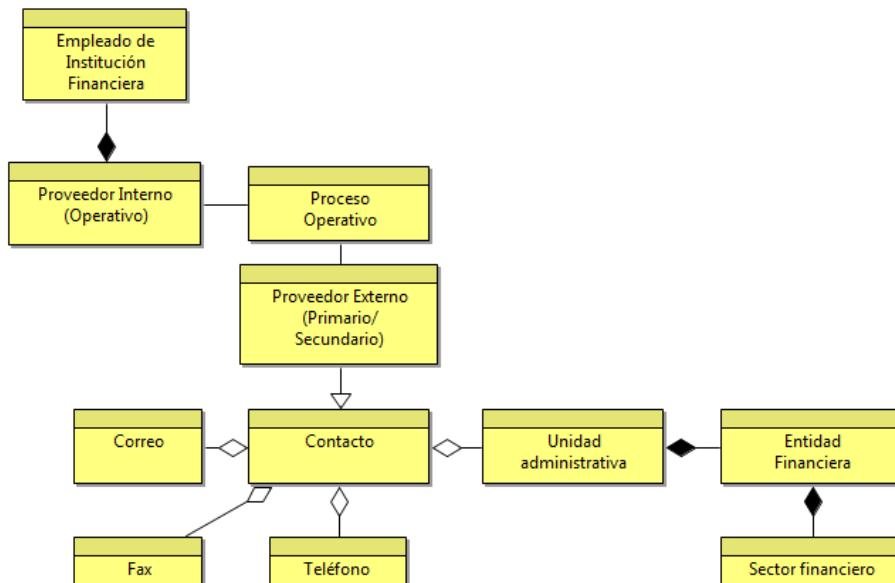


Figura 5.21 Diagrama de clases para Proveedores

- Desarrollar el modelo de datos mediante un lenguaje que incluya taxonomías, ontologías u otras representaciones conceptuales y lógicas de alto nivel

Contemplando todas las actividades previas se tienen el siguiente modelo lógico considerando únicamente usuarios.

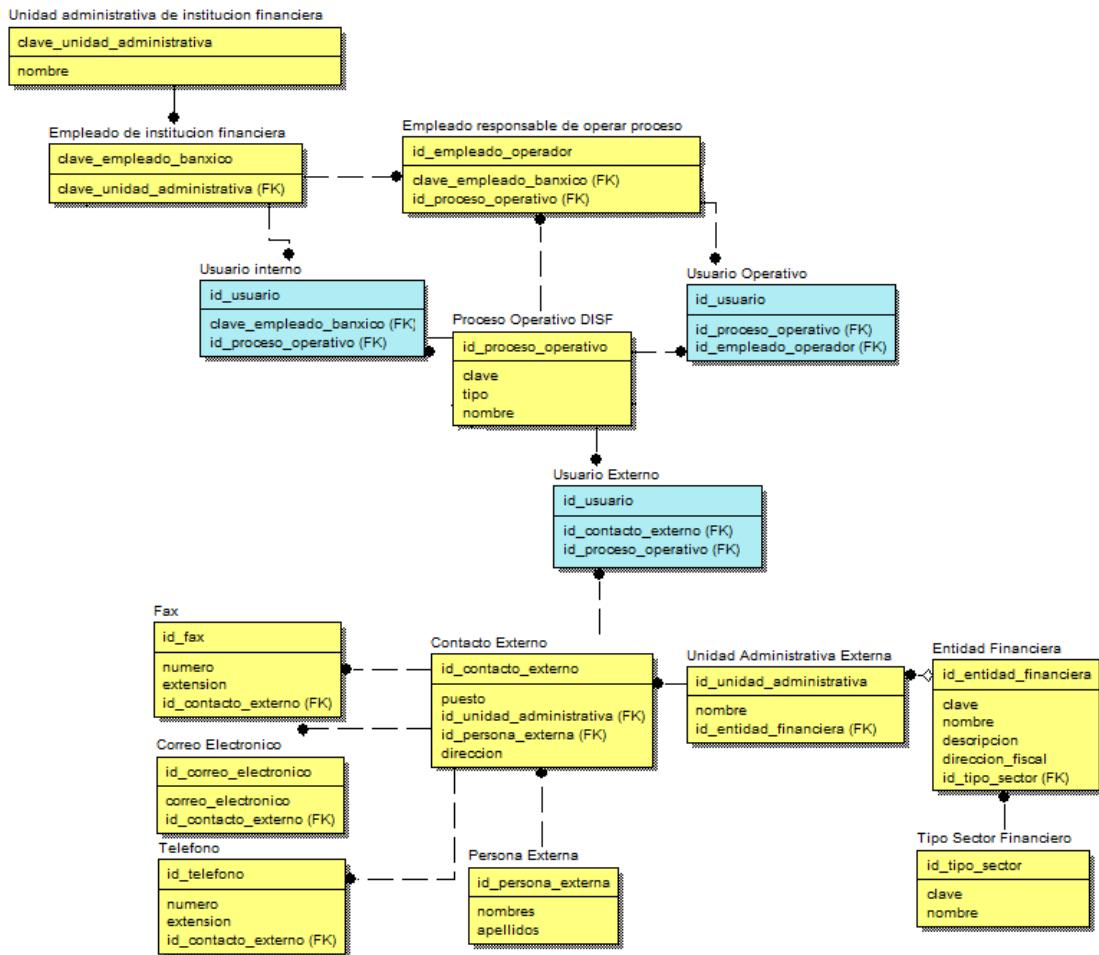


Figura 5.22 Modelo de datos de usuarios

Contemplando todas las actividades previas se tienen el siguiente modelo lógico considerando únicamente a los proveedores.

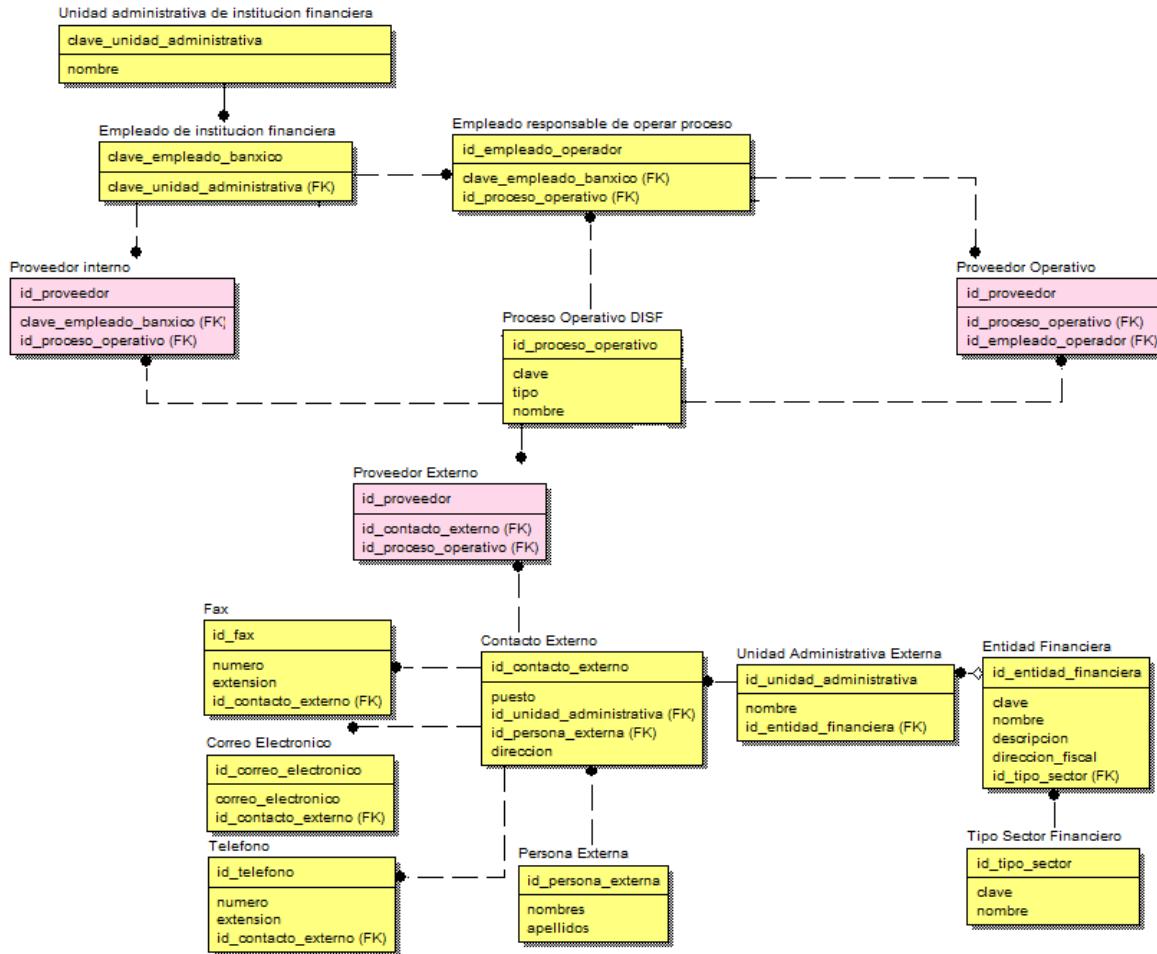


Figura 5.23 Modelo de datos de Proveedores

#### ▪ Definir sus metadatos

Los metadatos descritos para cada entidad son los siguientes:

Clave del objeto	Nombre del objeto	Descripción	Atributos	Descripción del atributo
AD.DMyR.01	Contacto Externo	Entidad que contiene la información que se forma de la relación que tiene una persona con una institución financiera.	Identificador del contacto externo	Clave numérica única que reconoce cada registro de un contacto.
			Puesto del contacto	Puesto del contacto que ejerce en su unidad administrativa externa.
			Identificador de la Unidad Administrativa.	Clave numérica única que referencia cada unidad administrativa.
			Identificador de la persona externa	Clave numérica única que referencia cada persona.
			Dirección	Ubicación donde ejerce sus responsabilidades.
			Identificador de la entidad	Clave numérica y única que referencia cada entidad financiera.
AD.DMyR.02	Correo Electrónico	Es la tabla que contiene la información de los correos que puede tener el contacto. Esta tabla tiene como atributos el identificador del correo electrónico, el correo electrónico y la clave del contacto al que pertenecen los datos.	Identificador del correo electrónico	Clave numérica única que reconoce cada correo electrónico asociado al contacto.
			Correo electrónico	Correo del contacto.
			Identificador del contacto externo	Clave numérica única que referencia cada registro de un contacto.
AD.DMyR.03	Entidad Financiera	Es el catálogo que contiene las instituciones financieras que componen el sistema financiero	Identificador de la Entidad Financiera	Clave numérica única que reconoce cada registro de una Entidad Financiera.
			Clave de la Entidad Financiera	Clave única de tipo alfanumérica que se asigna dentro del área financiera a cada entidad.
			Nombre de la Entidad Financiera	Nombre largo de la Entidad Financiera.
			Descripción de la entidad financiera	Breve descripción sobre la principal razón de ser de la entidad.
			Dirección Fiscal de la Entidad Financiera	Dirección Fiscal de la Entidad Financiera.
			Clave del tipo de sector financiero	Clave única de tipo alfanumérica que se asigna a cada Sector Financiero. Esta es la clave que referencia a la tabla “Tipo de Sector Financiero”.
AD.DMyR.04	Empleado del área financiera	Entidad que contiene los datos de los colaboradores de la organización, los colaboradores enlistados en la tabla deberán de ser usuarios o proveedores del área financiera	Clave del empleado	Identificador único de cada empleado de la organización. Esta clave se obtendrá del catálogo único que administra el área de recursos humanos.
			Clave de la unidad administrativa	Identificador único de la unidad administrativa de la organización. Este identificador se obtendrá del catálogo que administra el área de recursos humanos.
AD.DMyR.05	Fax	Es la tabla que contiene la información de los números del fax de cada contacto. Esta tabla tiene como atributos el identificador del fax, el número de fax, la extensión del fax y la clave del contacto al que pertenecen los datos.	Identificador del fax	Clave única que se le asigna a cada número de fax.
			Número	Número de fax de 8 dígitos.
			Extensión	Extensión numérica del fax.
			Identificador del contacto externo	Clave numérica única que referencia a cada contacto.
AD.DMyR.06	Empleado responsable de operar el	Entidad que guarda la relación del responsable de cada formulario del área financiera. La	Identificador del operador del área financiera	Clave única que representa cada instancia de las entidades.
			Clave de empleado	Identificador único contenido de cada colaborador de la organización.

	proceso operativo	existencia de esta tabla se basa en la regla de negocio que dicta que un colaborador es responsable de un formulario y ese formulario puede tener usuarios que son responsables de otro formulario.	Identificador del proceso operativo	Clave única del formulario del área financiera.
AD.DMyR.07	Persona Externa	Es la entidad que contiene los datos de los actores que participan únicamente por la función de ser proveedor o usuario de los procesos operativos de la	Identificador de la persona externa	Identificador numérico único que reconoce a cada persona.
			Nombres	Nombre(s) de la persona.
			Apellidos	Apellido paterno y materno de la persona.
AD.DMyR.08	Proceso Operativo	Catálogo que contiene los procesos operativos de la .	Identificador del Proceso Operativo del área financiera	Identificador numérico único que reconoce cada formulario del área financiera.
			Clave del Proceso Operativo	Clave que se asigna dentro de la a cada proceso operativo. Ejemplo: "ACM" (Anexo al Catálogo de Mínimo).
			Tipo del Proceso Operativo	Clasificación del tipo de Proceso Operativo. Este valor debe de ser 1,2 o 3.
			Nombre del Proceso Operativo	Nombre largo del Proceso Operativo
AD.DMyR.09	Proveedor operativo	Entidad que contiene las personas que pertenecen al área financiera y son proveedores operativos.	Identificador del proveedor operativo	Identificador numérico único que reconoce a nivel físico cada registro de un proceso operativo .
			Identificador del proceso operativo	Clave que se asigna dentro de la a cada proceso operativo. Ejemplo: "ACM" (Anexo al Catálogo de Mínimo).
			Identificador del operador del área financiera	Clave única que representa cada instancia de las entidades
AD.DMyR.10	Proveedor interno	Entidad que contiene las personas que pertenecen a la institución financiera y son proveedores internos.	Identificador del proveedor interno	Identificador numérico único que reconoce a nivel físico cada registro de un proceso operativo .
			Clave del empleado del área financiera	Identificador único de cada empleado de la organización. Esta clave se obtendrá del catálogo único que administra el área de recursos humanos.
			Identificador del proceso operativo	Clave que se asigna dentro de la a cada proceso operativo. Ejemplo: "ACM" (Anexo al Catálogo de Mínimo).
AD.DMyR.11	Proveedor externo	Entidad que contiene los contactos externos que son proveedores de información al área financiera.	Identificador del proveedor externo	Identificador numérico único que reconoce a nivel físico cada registro de un proceso operativo .
			Identificador del contacto externo	Clave numérica única que reconoce cada registro de un contacto.
			Identificador del proceso operativo	Identificador numérico único que reconoce cada formulario del área financiera.
AD.DMyR.13	Teléfono	Es la tabla que contiene la información de los teléfonos que puede tener el contacto. Esta tabla tiene como atributos el identificador del teléfono, el número de teléfono, la extensión del teléfono y la clave	Identificador del teléfono	Clave única que identifica a la tabla.
			Numero	Número telefónico.
			Extensión	Número de la extensión telefónica.
			Identificador del contacto externo	Clave numérica única que reconoce cada registro de un contacto.

		del contacto al que pertenecen los datos		
AD.DMyR.14	Tipo de Sector Financiero	Es el catálogo que contiene la clasificación por sector del conjunto de instituciones que componen el sistema financiero	Identificador del tipo de sector financiero	Identificador numérico único que reconoce a nivel físico cada registro de un sector financiero.
			Clave del tipo de sector financiero	Clave única de tipo alfanumérica que se asigna dentro del área financiera.
			Nombre del sector financiero	Nombre largo del Sector Financiero.
AD.DMyR.15	Usuario Operativo	Es la tabla que contiene la información de la relación que se forma cuando un proceso operativo tiene como usuarios a otros procesos operativos dentro del área financiera.	Identificador del usuario operativo	Clave única
			Identificador del proceso operativo	Identificador numérico único que reconoce cada formulario del área financiera
			Identificador del operador	Clave única que representa cada instancia de las entidades
AD.DMyR.16	Usuario interno	Entidad que representa los usuarios del área financiera; estas contactos son internos a la organización pero externos al área financiera.	Identificador del usuario interno	Clave única que reconoce a cada usuario
			Clave de empleado	Identificador único de cada empleado de la organización.
			Identificador del proceso operativo	Identificador numérico único que referencia cada formulario del área financiera.
AD.DMyR.17	Usuario externo	Entidad que representa la relación de los contactos externos que son usuarios de los formularios del área financiera.	Identificador del usuario externo	Clave única que reconoce cada usuario externo al área financiera.
			Identificador del contacto externo	Clave numérica única que referencia cada registro de un contacto.
			Identificador del proceso operativo	Identificador numérico único que referencia cada formulario del área financiera.
AD.DMyR.18	Unidad Administrativa de Área Financiera	Es el catálogo que contiene las unidades administrativas de las entidades financieras que participan como proveedores o usuarios de los procesos del área financiera	Identificador de la Unidad Administrativa	Identificador numérico único que reconoce cada unidad administrativa que pertenece a la organización donde se encuentra el área financiera.
			Nombre de la Unidad Administrativa	Nombre corto de la Unidad Administrativa de la organización a la que pertenece el área financiera.
AD.DMyR.19	Unidad administrativa externa	Es el catálogo que contiene las unidades administrativas de las entidades financieras que participan como proveedores o usuarios de los procesos del Área Financiera	Identificador de la unidad administrativa externa	Identificador numérico único para representar a cada unidad administrativa de las entidades financieras externas.
			Nombre de la unidad administrativa	Nombre de la Unidad Administrativa de las entidades financieras externas.

Tabla 5.23 Artefacto de Metadatos para Usuarios y Proveedores

- **Establecer las relaciones con otras entidades maestras**

Para documentar las relaciones lógicas que existen entre cada entidad que se utilizó para la representación maestra de usuarios y proveedores se usó el artefacto definido en la arquitectura de datos denominado “Matriz de las relaciones entre objetos de datos”. Un ejemplo acotado del artefacto es el siguiente:

	<i>Empleado del área financiera</i>	<i>Operador del AF</i>	<i>Proceso operativo</i>	<i>Proveedor operativo</i>	<i>Proveedor interno</i>	<i>Proveedor externo</i>
<i>Empleado del área financiera</i>		X			X	
<i>Operador del AF</i>	X		X	X		
<i>Proceso operativo</i>		X				X
<i>Proveedor operativo</i>		X	X			
<i>Proveedor interno</i>	X		X			
<i>Proveedor externo</i>			X			

**Tabla 5.24 Matriz de relaciones entre objetos de datos**

- **Implementar el modelo maestro en un repositorio**

La última actividad de esta iniciativa es la implementación física de este modelo de datos maestros para usuarios y proveedores. Para ello, se utilizó el manejador de bases de datos Sybase y SQL Server.

Se eligió SQL Server debido a que provee ventajas sobre la estructura de datos como la elaboración de índices en cadenas o el tipo de dato “strings”, esto ayudar a la unicidad de los datos lo cual es indispensable en los datos maestros y de referencia.

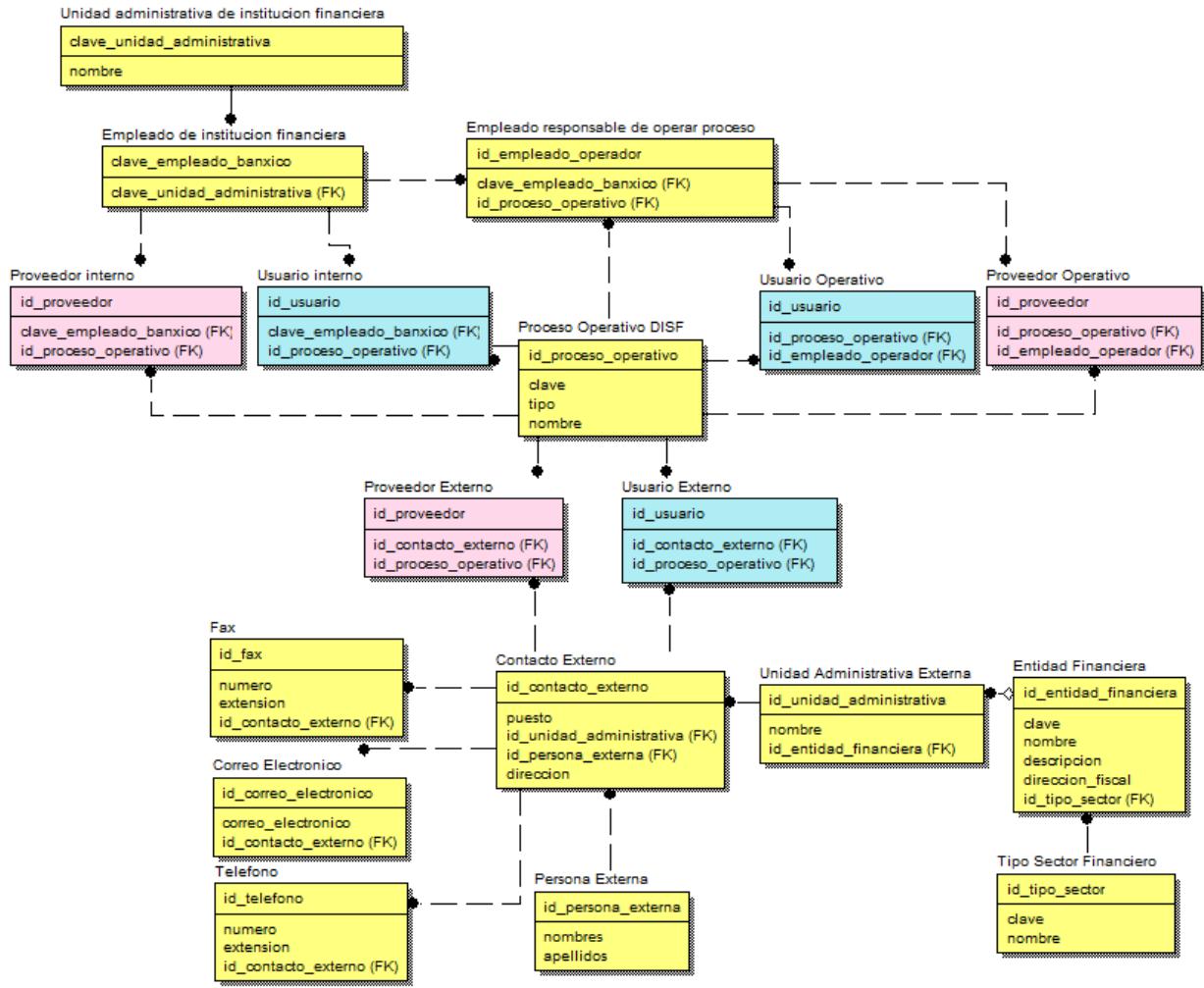


Figura 5.24 Modelo maestro de usuarios y proveedores

## 6. CONCLUSIONES

Derivado del desarrollo y aplicación del modelo-solución para la gestión de datos maestros y de referencia es posible afirmar que el uso de un marco como TOGAF facilita la definición del contenido del modelo-solución, debido a que provee fases para el diseño, planificación e implementación de una arquitectura empresarial.

En la aplicación del caso de estudio abordado se confirmó que para la fase *Preliminar* es importante realizar una evaluación de madurez porque de esta manera se tiene un punto de partida para llegar a la arquitectura objetivo, y puesto que cada organización tiene un nivel de madurez inicial es necesario identificar ese valor para definir la estrategia a seguir en la fase E *Oportunidades y soluciones*.

Para la fase *Visión de la arquitectura* se confirmó que este modelo cubrió las siguientes necesidades de información: contar con una versión única y estandarizada de los datos, y cumplir con estándares de la industria de TI. Es posible afirmar esto debido a los artefactos documentados en la Iniciativa 1 (*Detección y análisis de datos*) para el área financiera, en específico el proyecto 1.3 (*Consolidación del contenido de usuarios y proveedores*).

Se confirmó que los artefactos establecidos en la fase C *Arquitectura de datos* ayudan a la conceptualización y estandarización de usuarios y proveedores. Esto se reflejó gracias a que hubo una mayor comprensión, tanto para la gente del negocio como de TI, después de la ejecución de la Iniciativa 1 y 2.

Para la fase E *Oportunidades y soluciones* se encontró que los conflictos humanos y metodológicos se ven reducidos cuando se sigue una estrategia, de lo contrario se tiene la posibilidad de perder el principal objetivo y desgastarse en otras tareas que son necesarias, pero no primordiales en ese momento. También al implementar las iniciativas se descubrieron otras áreas de oportunidad para la área financiera.

La aplicación del modelo confirmó que la centralización de los datos maestros y de referencia aportó una visión holística al Área Financiera del caso de estudio porque proporciona: beneficios que se pueden tener sobre todos los elementos de la arquitectura-solución, visualización de los impactos y escalabilidad de su arquitectura.

Al aplicar la Iniciativa 1 llamada *Detección y análisis de usuarios y proveedores*, se uno de los grandes retos fue y sigue siendo la resistencia al cambio por parte de los colaboradores. Esto se debe principalmente a su negligencia acompañada del miedo a un cambio estructural.

Así mismo se observó que el personal, tanto de negocio como de TI, de la institución financiera no tiene conocimiento de la definición e importancia que presentan los datos

maestros y de referencia. Fue y es un desafío plantearle a la gente que es necesario cambiar los hábitos en su forma de trabajar. Confirmándose con esto que un modelo por perfecto que sea, no aportará ventajas sin que sea comprendido por la gente.

Se observó que las Iniciativas 1 y 2 de este modelo funcionan bajo las siguientes condiciones:

- Organizaciones que cuenten con el apoyo de sus patrocinadores e interesados para convencer al personal del negocio con su colaboración para el levantamiento de información.
- Organizaciones que cuenten con herramientas de TI como manejadores de bases de datos, ETL y sitios de colaboración.

El caso de estudio es una aproximación de un gobierno de datos, por lo que en estudios posteriores deberán hacerse esfuerzos por implementar a un caso de estudio las dos iniciativas restantes de la estrategia definida en la fase E, así como finalizar la definición de las fases C, E, F y G con base en TOGAF. Sin embargo es posible anticipar las siguientes hipótesis:

- Si el personal de cualquier organización no comienza por comprender las definiciones de datos maestros y de referencia, no se ejecutará con éxito cualquier esfuerzo relacionado con la gestión de estos datos.
- Los principios, procesos y roles definidos en la *Arquitectura de negocio* facilitarán la eliminación de vicios como la apropiación de la información debido a que se pondrá su definición a disposición de todos los colaboradores.
- La puesta en marcha de las cuatro iniciativas definidas en la estrategia del modelo permitirán aprovechar mejor los recursos humanos y tecnológicos.
- La definición de indicadores y procesos que supervisen el cumplimiento de la gestión de datos maestros y de referencia incrementará la calidad de información transaccional y analítica.
- Compartir, sincronizar y actualizar los datos maestros y de referencia con los procesos de negocio, hará más perceptibles las ventajas de la arquitectura-solución.
- Una vez que se finalice la definición de todas las fases de la arquitectura-solución para la gestión de datos maestros y de referencia se podrá asegurar el cumplimiento de la calidad de información que se maneje.

## 7. ANEXOS

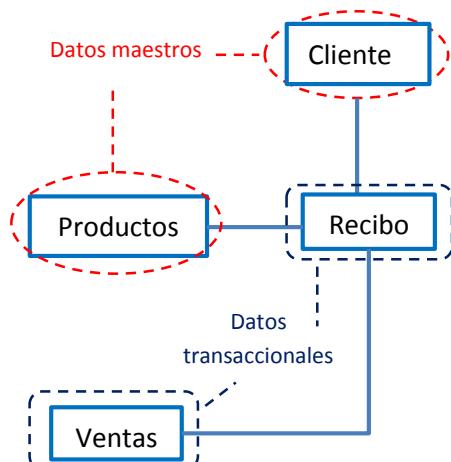
### 7.1 Relación entre datos maestros y datos transaccionales

Los datos transaccionales, también conocidos como datos operacionales son aquellos que derivan de las transacciones diarias del negocio. Como ejemplo de esta categoría es la información que se captura mediante:

- Pedidos
- Facturas
- Recibos
- Órdenes de servicio

Estos datos describen lo que ocurrió en el negocio y se relacionan con aplicaciones de centros de llamadas, gestión de relaciones con el cliente (CRM), entre otros. Retomando la definición que se propuso para datos maestros, la relación que existe entre estos dos tipos radica en que los datos transaccionales dependen de los datos maestros.

Los datos maestros son la base para muchos procesos operacionales, su contenido y precisión son cruciales ya que involucran de alguna forma a los datos transaccionales (Kontenix, 2010). Para exemplificar esta relación, situémonos en un negocio que vende celulares y cuenta con un sistema de ventas. El cliente al comprar su teléfono se le da un ticket o factura, para hacerse este recibo se tienen que introducir, y por lo tanto también verificar, datos del cliente y las características del teléfono (producto de la empresa). El sistema de ventas debe almacenar y sumar todos los recibos para determinar las ventas totales que se han hecho en el negocio en un determinado periodo, por lo que también existe una relación con la entidad ventas.



En este ejemplo tenemos las entidades de clientes, productos (en este caso son teléfonos), recibos y ventas. Como se ha descrito anteriormente producto y cliente son datos maestros y, recibos y ventas, datos transaccionales.

Figura 7.1 Ejemplo de datos maestros y transaccionales

Los datos transaccionales carecen de veracidad cuando se basan en datos maestros obsoletos, esto provoca reportes incorrectos, a su vez estos reportes impactan en el análisis y toma de decisiones basadas en esta información. Por lo tanto la calidad de los datos maestros proporciona un contexto seguro para los datos transaccionales.

## **7.2 Diferencia entre un DWH y un repositorio maestro**

Un data warehouse es un repositorio que permite el almacenamiento de datos históricos de la empresa para el análisis, toma de decisiones y posterior publicación de esta información. Un data warehouse es una copia estructurada de las transacciones de datos para la consulta y el análisis (Kimball, 2009).

En un data warehouse solo se puede leer información una vez que se almacena, por el contrario un repositorio de datos maestros es aquel que almacena de manera centralizada estos datos, y las actualizaciones que surjan en estos datos se deben modificar constantemente.

Una vez definidos estos elementos las principales diferencias son:

- Un repositorio de datos maestros almacena los datos críticos del negocio (datos maestros) un data warehouse almacena los datos transaccionales históricos.
- La estructura y organización de la información en un repositorio de datos maestros se almacena en un modelo relacional normalizado, por el contrario la información de un data warehouse se almacena en un modelo estrella.
- Un data warehouse emplea un proceso mono direccional (una vez que los datos se ingresan solo sirven de consulta), mientras que un repositorio de datos maestros necesita un proceso bidireccional; actualización constante (Informática, 2008).

Recordando el ejemplo del negocio que vende celulares, los datos maestros (productos, clientes, etc.) sirven para diferentes contextos. Por ejemplo un dato maestro que se utiliza en un sistema de pedidos para crear una “orden” tiene una finalidad operativa para un sistema en particular, proporcionar un contexto confiable a los datos transaccionales. Al finalizar la “orden” esta información se almacena en un data warehouse y los datos maestros son parte del análisis.

## **7.3 Proceso general de vinculación de registros**

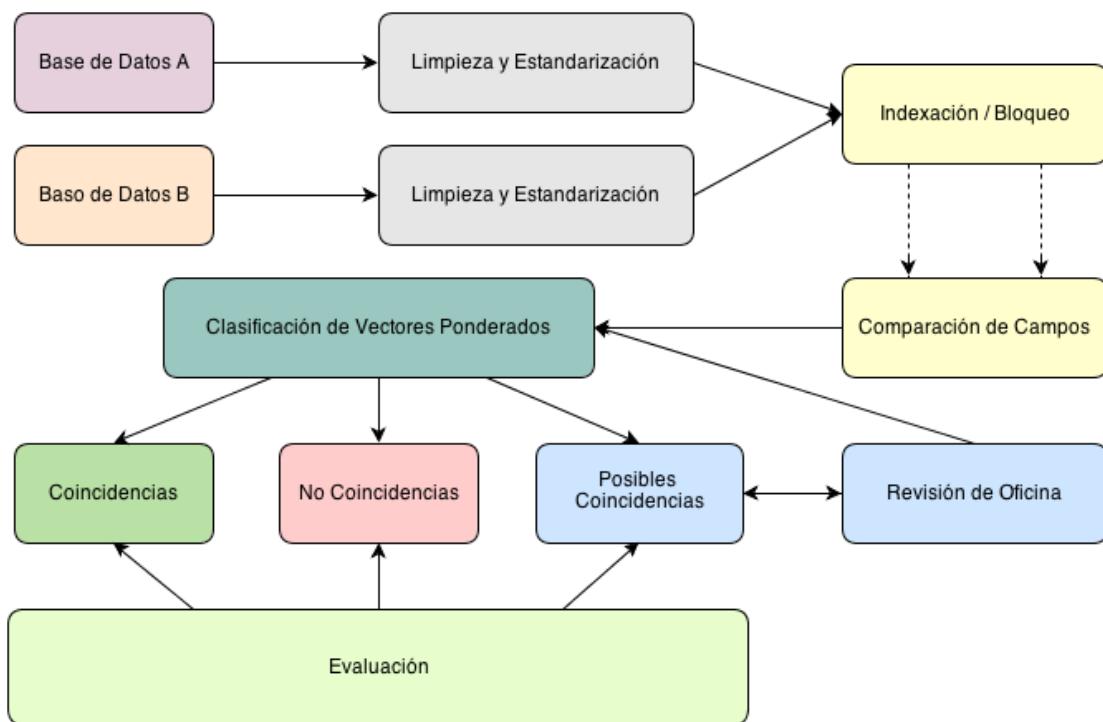
En la mayoría de los sectores se ha experimentado un crecimiento en los datos, gran parte de estos datos se refieren a información de clientes. La vinculación de Datos Maestros entre diferentes sistemas es una tarea importante porque permite obtener un panorama general de todas las áreas de la empresa para poder realizar un análisis del negocio.

Al mismo tiempo un dato vinculado puede ayudar a mejorar políticas, descubrir fraudes y reducir costos en la organización.

En el área de sistemas de vinculación de registros e integración de datos se han tenido avances, muchos de estos han sido implementados en software de licencia comercial. Estos sistemas son como una “caja negra” (Christen, 2007). No se permite a los usuarios ver los detalles de la tecnología implementada y por lo tanto es difícil aprender nuevas técnicas para permitir comparar la vinculación de registros.

Para un estudio más exhaustivo sobre la integración de los datos, en este subcapítulo se pretende entender cómo trabaja internamente un sistema las comparaciones entre diferentes sistemas de datos, como: nombres y direcciones.

Para comprender las ventajas y limitaciones respecto a las técnicas de vinculación de registros, a continuación se muestra la siguiente figura de un proceso general de vinculación de registros entre dos bases de datos:



**Figura 7.2 Vinculación de datos**

Esta figura representa el proceso de vinculación para la integración de dos bases de datos. Primero se realiza una limpieza y estandarización en los datos, después se ejecuta una indexación en cada registro. A la salida de este bloqueo o indexación se les denomina “pares de registros candidatos”, por lo que se codifican.

Posteriormente, de acuerdo al código asignado se realiza una comparación de cada campo de la base de datos en base a vectores, y de esta comparación se obtienen “vectores ponderados” con consideraciones numéricas. Estos vectores se pueden clasificar en los siguientes tipos:

- Coincidencias
- No coincidencias
- Posibles coincidencias

Para la agrupación de posibles coincidencias se le asigna a una persona la tarea de descartar elementos que puedan estar duplicados (Coincidencias) y/o para que se verifique los datos que no coinciden (No coincidencias). Finalmente la persona asignada realiza una evaluación final de los 3 tipos de clasificaciones.

El fusionar grupos de registros que han sido clasificados como “matches” o correspondientes (es decir, que se supone que se refieren a la misma entidad) en un registro limpio y consistente para una sola base de datos se llama Deduplication

#### **7.4 Definición de servicios SOA**

Con base en el marco TOGAF, se recomienda realizar el diseño de aplicaciones orientado a SOA. SOA es un modelo de arquitectura y se basa en el paradigma orientado a objetos donde su unidad fundamental es el servicio.

TOGAF define a un servicio como la representación lógica de una actividad repetible de negocio que tiene un resultado específico.

Existen tres capas de modelado de servicios SOA, las cuales se describen brevemente a continuación:

- Tarea: servicios que expresan capacidades de negocio que tienen funcionalidad específica, no tiene potencial de ser reutilizable.
- Entidad: servicios con contextos de negocio que tienen funcionalidad genérica; pueden ser reutilizables. Ejemplo taxonomías, ontologías, información, modelos de datos.
- Utilidad: servicios que soportan y procesan capacidades de negocio y son genéricas. Ejemplo: Acceso a datos, mensajería, log, correo electrónico, seguridad.

Con base en esta clasificación y el modelamiento SOA a continuación se diseñan los servicios genéricos para los procesos definidos en la *Arquitectura de Negocio* de este capítulo.

Para el diseño de servicios se siguieron los siguientes pasos (Erl, 2008):

1. Descomponer el proceso de negocio
2. Filtrar lógica no apta para ser encapsulada en un servicio
3. Identificar servicios agnósticos candidatos
4. Identificar lógica específica de proceso
5. Aplicar el paradigma de orientación a servicios

## 6. Identificar composiciones de servicios candidatos

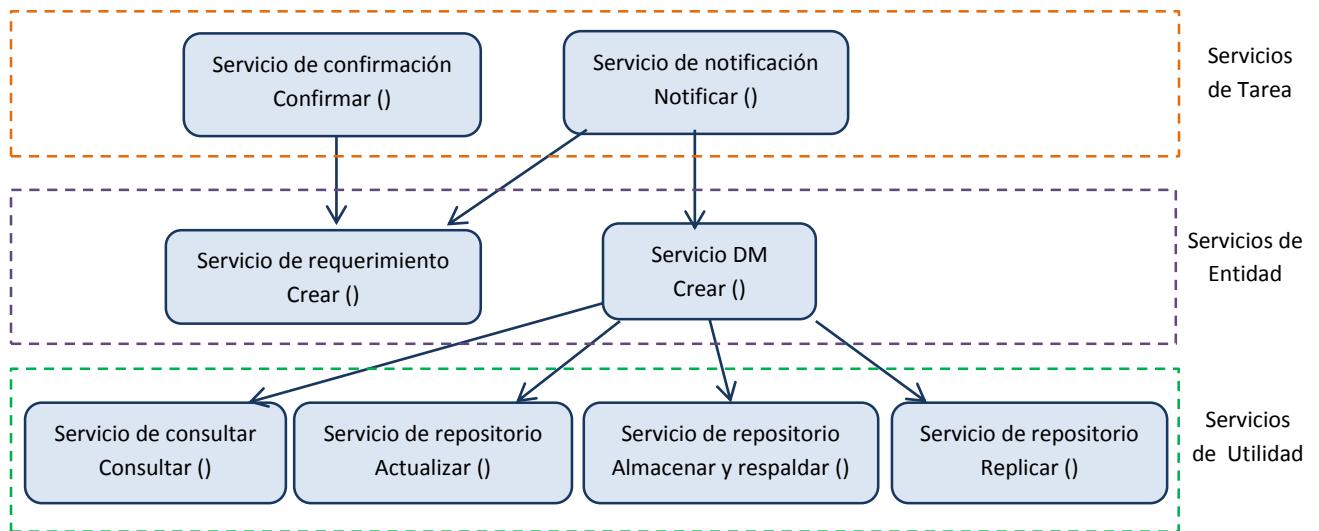


Figura 7.3 Diseño de servicios para DMyR

El servicio de confirmación apoya todas las tareas sobre:

- ✓ Confirmar recepción de requerimiento
- ✓ Revisar y validar la procedencia del requerimiento

El servicio de notificación apoya todas las tareas sobre:

- ✓ Si la entidad no existe notificar rechazo de la actualización de la entidad
- ✓ Tomar decisión; aprobar o rechazar autorización. Si se rechaza el requerimiento notificar
- ✓ Notificar decisión a interesados
- ✓ Notificar cambio

El servicio de consultar apoya tareas sobre:

- ✓ Validar reglas, analizar resultados y evaluar la entidad existente

Verificar la existencia de la entidad

El servicio de actualizar repositorio apoya tareas sobre:

- ✓ Actualizar entidad

El servicio de almacenar y respaldar repositorio apoya tareas sobre:

- ✓ Almacenar actualización de entidad
- ✓ Respaldar histórico de actualización

El servicio de replicar repositorio apoya tareas sobre:

- ✓ Replicar actualización

Estos servicios SOA se han descrito conforme a las actividades del proceso de negocio.

## 8. BIBLIOGRAFÍA Y REFERENCIAS

- Christen, 2007, Febrl Freely Available Record Linkage System with a Graphical User Interface, Universidad Nacional de Australia.  
[<http://datamining.anu.edu.au/linkage.html>](http://datamining.anu.edu.au/linkage.html)
- DAMA, 2009, The DAMA Guide to The Data Management Body of Knowledge, U.S.A., Technics Publications, LLC, pág. 176,174, 171, 173.
- EDW, 2013, Data Architecture Workshop, U.S.A,
- Erl, 2008, SOA Principles of Service Design, Prentice Hall, U.S.A, pág. 524

- IBM, 2014, InfoSphere Master Data Management, <http://www-03.ibm.com/software/products/es/infosphere-master-data-management>
- IBM, 2011, How master data management serves the business, U.S.A, pág. 2, <<http://www-01.ibm.com/software/data/master-data-management/overview.html>>
- IBM, 2010, The Art of Enterprise Information Architecture, U.S.A Boston, IBM Press, pág. 57-59
- IBM, 2007, Gestión de datos maestros: superar la visión única para encontrar la visión adecuada, EE.UU. , pág. 2,8, <[http://www-05.ibm.com/services/es/cio/pdf/CIO\\_Series\\_0402.pdf](http://www-05.ibm.com/services/es/cio/pdf/CIO_Series_0402.pdf)>
- Informatica, 2014, Master Data Management, <http://www.informatica.com/es/products/master-data-management/mdm/>
- Informatica, 2006, Gestión de datos maestros (MDM), España, <[http://www.informatica.com/INFA\\_Resources/br\\_mdm\\_es.pdf](http://www.informatica.com/INFA_Resources/br_mdm_es.pdf)>
- Informatica, 2008, <<http://www.informatica.com/us/products/master-data-management/mdm/>>
- Informatica, 2008 ,Gestión de datos maestros (MDM) <[http://www.informatica.com/INFA\\_Resources/br\\_mdm\\_es.pdf](http://www.informatica.com/INFA_Resources/br_mdm_es.pdf)>
- ISO 4217,2008, Currency codes, <<http://www.currency-iso.org/en/home/tables/table-a1.html>>
- Kimball, 2009, Data Warehouse Implementations: Critical Implementation Factors Study. VDM Verlag
- Kontenix, 2010, Datos Maestros VS Transaccionales. <[http://www.kontenix.com/site/index.php?option=com\\_content&view=article&id=193:datos-maestros-vs-transaccionales&catid=64:blog&Itemid=178](http://www.kontenix.com/site/index.php?option=com_content&view=article&id=193:datos-maestros-vs-transaccionales&catid=64:blog&Itemid=178)>
- Loshin, 2009, Master Data Management, U.S.A., Morgan Kaufmann, pág. 43-65
- MDM Institute, 2013, ¿Qué es MDM?, <<http://0046c64.netsolhost.com/whatIsMDM.html>>
- MDM Institute, 2009, MDM & DATA GOVERNANCE READINESS ASSESSMENT: A Summary Practical Guide to Systems Integrators & Consultancies, U.S.A, <<http://www.tcdii.com/PDF/A%20Summary%20Practical%20Guide%20to%20System%20Integrators%20&%20Consultancies%20for%20MDM%20&%20DG.pdf>>
- MDM Institute, 2007, Multi-Entity Master Data Management: The 4th Generation of MDM, pág. 7 <[http://0046c64.netsolhost.com/PDF/Multi-Entity\\_MDM\\_white\\_paper.pdf](http://0046c64.netsolhost.com/PDF/Multi-Entity_MDM_white_paper.pdf)>
- Microsoft, 2014, Master Data Services, <http://msdn.microsoft.com/es-mx/library/ee633763.aspx>
- Microsoft, 2011, Microsoft Business Intelligence: vea el cubo medio lleno, SolidQ Press, España, pág. 23

- Microsoft, 2006, The What, Why, and How of Master Data Management, <<http://msdn.microsoft.com/en-us/library/bb190163.aspx>>
- Oracle, 2014, Oracle Master Data Management Suite , <http://www.oracle.com/es/products/applications/master-data-management/index.html?ssSourceSiteId=null>
- Oracle, 2011, Master Data Management, U.S.A. , pág. 5 <<http://www.oracle.com/us/products/applications/master-data-management/018876.pdf>>
- Rahm, 2013, Data Cleaning: Problems and Current Approaches. University of Leipzig, Germany , <<http://dbs.uni-leipzig.de>>
- SAS,2013 , Compatibilidad con la gestión de datos maestros , <<https://www.sas.com/offices/latinamerica/mexico/software/data-management/entdiserver/master-data-mgmt.html>>
- TOGAF, 2011, TOGAF Version 9.1 Enterprise Edition, U.S.A, ISBN 9789087536794, <<http://pubs.opengroup.org/architecture/togaf9-doc/arch/chap06.html>>