



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA

**Uso de Visualizaciones de datos para
fortalecer la gestión de las
organizaciones a través de una
arquitectura empresarial**

TESIS

Que para obtener el título de
Ingeniero en Computación

P R E S E N T A

Alberto Álvarez González

DIRECTOR DE TESIS

Ing. Jorge Alberto Rodríguez Campos



Ciudad Universitaria, Cd. Mx., 2018

Agradecimientos

Al término de esta etapa, quiero expresar un profundo agradecimiento a quienes con su ayuda y apoyo me alentaron a lograr alcanzar esta meta

A mis padres:

Alberto Álvarez Valdés y Yolanda González Del Ángel; Sabiendo que no existirá una forma de agradecer una vida de sacrificio y esfuerzo; Gracias por todo su cariño, guía y apoyo que me han brindado a lo largo de mi vida; Gracias por los valores que me han inculcado, por haberme dado una educación siendo esta la herencia más importante que puedo recibir. Gracias ya que es por ellos soy la persona que soy ahora.

A mi hermano

Alejandro Álvarez González, por ser parte de mi vida siendo un apoyo y motivación en los momentos más difíciles.

A la UNAM:

Por haberme brindado de una distinguida formación dándome las herramientas para defenderme en el ámbito profesional y laboral, dándome la oportunidad de conocer a grandes profesores, de quienes he adquirido gran conocimiento enseñándonos no solo como ser mejores profesionistas, sino mejores personas. Gracias por todo UNAM y Facultad de Ingeniería.

Gracias a todos los que participaron en la construcción de este trabajo, en especial al Ing. Jorge Alberto Rodríguez Campos agradeciendo su dirección a lo largo de este, asimismo agradecer a los miembros del jurado por tomarse el tiempo de leer el presente trabajo y por los comentarios hechos al respecto

Gracias a todos.

INDICE

1. INTRODUCCIÓN	10
1.1 OBJETIVO.....	12
1.2 OBJETIVOS PARTICULARES	12
2. ANTECEDENTES	13
2.1. SITUACIÓN ACTUAL	13
3. MARCO TEÓRICO	16
3.1 TECNOLOGÍAS WEB	16
3.2 VISUALIZACIONES DE DATOS	17
3.2.1 BUENAS PRÁCTICAS PARA LA CREACIÓN DE VISUALIZACIONES	19
3.2.1.1. ELECCIÓN DE LAS MÉTRICAS IMPORTANTES	20
3.2.1.2. VISUAL.....	22
3.2.1.3. INTERACTIVO.....	23
3.2.1.4. MANTENERLO ACTUALIZADO	25
3.2.1.5. ACCESIBILIDAD	25
3.2.2. ERRORES MÁS COMUNES AL DESARROLLAR VISUALIZACIONES.....	25
3.2.2.1. COMENZAR CON DEMASIADA COMPLEJIDAD.....	26
3.2.2.2. USAR MÉTRICAS QUE NADIE ENTIENDE	26
3.2.2.3. ATIBORRAR LA VISUALIZACIÓN CON GRÁFICOS DE BAJA CALIDAD.	27
3.2.2.4. ESPERAR POR TECNOLOGÍA COMPLEJA Y GRANDES PROYECTOS DE IMPLEMENTACIÓN DE INTELIGENCIA COMERCIAL.....	27
3.2.2.5. SUBESTIMAR LA NECESIDAD DE MANTENIMIENTO.	28
3.2.2.6. QUE LAS MÉTRICAS NO COINCIDAN CON LAS METAS.	28
3.2.2.7. USAR GRÁFICO INEFICIENTE Y MAL DISEÑADO.	30
3.3 ARQUITECTURA DE SERVICIOS.....	30
3.3.1 SERVICIO.	31
3.3.2 WEB SERVICES.	32
3.3.3 WEB SERVICE BASADOS EN REST Y SOAP.....	32
3.3.4 BENEFICIOS DE USAR SOA.....	34
3.3.4.1 MEJORAR LA TOMA DE DECISIONES.....	35
3.3.4.2 MEJORAR LA PRODUCTIVIDAD DE LOS EMPLEADOS.	35
3.3.4.3 POTENCIAR LAS RELACIONES CON CLIENTES Y PROVEEDORES.	35
3.3.4.4 APLICACIONES MÁS PRODUCTIVAS Y FLEXIBLES.	35
3.3.4.5 DESARROLLO DE APLICACIONES MÁS RÁPIDO Y ECONÓMICO.	36
3.3.4.6 APLICACIONES MÁS SEGURAS Y MANEJABLES.	36
3.4. BUSINESS PROCESS MANAGEMENT.	36
3.5 MARCO DE LA ARQUITECTURA EMPRESARIAL (TOGAF).....	38

3.6 GESTIÓN ORGANIZACIONAL	42
4. CASO DE ESTUDIO.....	45
4.1 PRIMERA ITERACIÓN (LIMPIEZA DE LOS DATOS)	47
4.1.1 ITERACIÓN DE CONTEXTO	47
4.1.1.1 FASE 0: PRELIMINAR	47
4.1.1.2 FASE A: VISIÓN DE LA ARQUITECTURA	69
4.1.2 ITERACIÓN DE DEFINICIÓN	69
4.1.2.1 FASE B: ARQUITECTURA DE NEGOCIO.....	69
4.1.2.2 FASE C: ARQUITECTURA DE SISTEMAS DE INFORMACIÓN	74
4.1.2.3 FASE D: ARQUITECTURA DE TECNOLOGÍA.....	83
4.1.3 ITERACIÓN DE PLANEACIÓN	89
4.1.3.1 FASE E: OPORTUNIDADES Y SOLUCIONES.....	89
4.1.3.2 FASE F: PLANEACIÓN DE MIGRACIÓN	90
4.1.4 CONCLUSIÓN DE PRIMERA ITERACIÓN.....	90
4.2 SEGUNDA ITERACIÓN	92
4.2.1 ITERACIÓN DE CONTEXTO	92
4.2.1.1 FASE 0: PRELIMINAR.....	92
4.2.1.2 FASE A: VISIÓN DE LA ARQUITECTURA	93
4.2.2 ITERACIÓN DE DEFINICIÓN	93
4.2.2.1 FASE B: ARQUITECTURA DE NEGOCIO.....	93
4.2.2.2 FASE C: ARQUITECTURA DE SISTEMAS DE INFORMACIÓN	97
4.2.2.3 FASE D: ARQUITECTURA DE TECNOLOGÍA.....	120
4.2.3 CONCLUSIONES DE SEGUNDA ITERACIÓN	128
4.2.3 ITERACIÓN DE PLANEACIÓN	121
4.2.3.1 FASE E: OPORTUNIDADES Y SOLUCIONES.....	121
4.2.3.2 FASE F: PLANEACIÓN DE MIGRACIÓN	122
4.2.4 ANALISIS DE RESULTADOS.....	123
5. CONCLUSIONES.....	129
6. BIBLIOGRAFÍA.....	131

Índice de Figuras

Figura 3.1 Características del diseño visual

Figura 3.2 Modelo de base de Datos Escolar

Figura 3.3 Patrones a seguir al crear una visualización

Figura 3.4 Ejemplo Visualización Interactiva

Figura 3.5 Ciclo de Vida para implementación de BI

Figura 3.6 Visualización con sobrecarga de datos

Figura 3.7 Ilustración de Arquitectura de Servicios

Figura 3.8 Interacción a través de Web Service SOAP

Figura 3.9 Componentes arquitectura empresarial

Figura 3.10 Componentes arquitectura empresarial

Figura 3.11 Método de desarrollo de Arquitectura, TOGAF 2011

Figura 4.1 Iteraciones del modelo ADM

Figura 4.2 Estructura Jerárquica de la Institución

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 Administración de Procesos de negocio. Adaptado de Loshin, 2009

Figura 4.9 Evaluación de arquitectura

Figura 4.10 Evaluación de Gobierno

Figura 4.11 Evaluación de Administración

Figura 4.12 Evaluación de Identificación

Figura 4.13 Evaluación de Integración

Figura 4.14 Evaluación de Administración de Proceso de Negocio

Figura 4.15 Proceso para controlar calidad de datos

Figura 4.16 Funciones para el proceso: Calidad de entidades

Figura 4.17 Servicios para el proceso de control de calidad de datos

Figura 4.18 .Niveles de abstracción de arquitectura de datos

Figura 4.19 Escenario de los proveedores

Figura 4.20 Diagrama conceptual de la base de datos

Figura 4.21 Modelo Lógico de la base de datos

Figura 4.22 Cuadrante Gartner para herramientas de calidad de datos 2012

Figura 4.23 Carga de Archivos

Figura 4.24 Consulta de coincidencias

Figura 4.25 Carga de Archivos

Figura 4.26 Estructura de desglose de trabajo

Figura 4.27 Proceso para la creación de nuevas visualizaciones

Figura 4.28 Funciones para el proceso: Creación de visualizaciones de datos

Figura 4.29 Servicios para el proceso de creación de visualizaciones de datos

Figura 4.30 Modelo en Cascada para desarrollo de Software

Figura 4.31 Grafica de Barras

Figura 4.32 Grafica de líneas

Figura 4.33 Grafica circular

Figura 4.34 Grafica de dispersión o burbujas

Figura 4.35 Infografía

Figura 4.36 Grafico de Árbol

Figura 4.37 Grafico Jerárquico Circular

Figura 4.38 Grafico de árbol Horizontal

Figura 4.39 Grafico Jerárquico Circular mostrando relación

Figura 4.40 Modelo de Arquitectura de la Aplicación

Figura 4.41 Diseño Front de la visualización

Figura 4.42 Diseño Front del formulario.

Figura 4.43 Modelo de datos necesario para la visualización de datos

Figura 4.44 Visualización de árbol estándar.

Figura 4.45 Visualización personalizada a tipo de nodos

Figura 4.46 Visualización adicional creada con D3

Figura 4.47 Comportamiento de la Visualización adicional creada con D3

Figura 4.48 Modelo de Técnico para el desarrollo de la aplicación

Figura 4.49 Estructura de desglose de trabajo para la creación de la visualización

Figura 4.50 Visualización tipo Tree obtenida

Figura 4.51 Visualización mostrando los nodos de segundo nivel

Figura 4.52 Visualización mostrando los nodos de tercer nivel

Figura 4.53 Visualización mostrando los contactos a atender la visualización

Figura 4.54 Formulario para envío de reporte

Figura 4.55 Mensaje de envío correcto de reporte

Índice de Tablas

Tabla 3.1 Tabla comparativa entre SOAP y REST

Tabla 4.1 Resumen de Evaluación de Madurez

Tabla 4.2 Servicios de Negocio

Tabla 4.3 Entidad de datos

Tabla 4.4 Relaciones de Tablas

Tabla 4.5 Datos de la base de datos productiva

Tabla 4.6 Datos nuevos a ingresar

Tabla 4.7 Datos finales

Tabla 4.8 Proyectos de implementación de limpieza de datos

Tabla 4.9 Servicios de Negocio para visualización de datos

Tabla 4.10 Servicios SOA

Tabla 4.11 Proyectos para la implementación de la visualizaciones

1. INTRODUCCIÓN

Es a partir de 1960 que la informática se introduce en las organizaciones con el objetivo de automatizar tareas administrativas repetitivas (contabilidad, facturación y nómina, principalmente). La tecnología se basa en grandes ordenadores o mainframes. El hardware y el software son extraordinariamente caros. Sólo las grandes organizaciones con enormes volúmenes diarios de trabajo administrativo pueden permitirse dichos costes. La justificación de la inversión radica en cálculos sobre los posibles ahorros salariales que se podían realizar en las áreas administrativas. En realidad, no se producen ahorros, sino que, con recursos humanos no crecientes, se absorbe una cantidad mucho mayor de trabajo.

Para los años 70 se consolida la informática en las grandes transnacionales, sin embargo, el cambio de paradigma trae también sus consecuencias y algunos supervisores comienzan a experimentar los cambios llamando la atención en relación con los sistemas de información y es que les llega demasiada información. La explosión de información avanza, cruzándose y entrecruzándose por las mesas de los jefes, con una enorme cantidad de datos. La mayoría de toda esta avalancha se asimila sólo parcialmente.

En los años 80 empieza a detectarse que el cambio introducido por las TIC (Tecnologías de la Información y Comunicación) en las organizaciones va mucho más lejos de lo comprendido hasta el momento. Las TIC no son sólo un instrumento de reducción de costes y de mejora de la información para la gestión. Las TIC tampoco son sólo un servicio de apoyo a la actividad principal de la organización. Las TIC están cambiando la naturaleza de los productos, de los procesos de producción y servicios, así como la naturaleza de la competencia y de los sectores económicos mismos.

El progresivo desarrollo tecnológico ha ido penetrando y transformando la manera de cómo los seres humanos se comunican, cómo piensan, cómo se distribuye el tiempo, la manera de cómo se hacen las cosas y hasta cómo consiguen sus objetivos las personas en la actualidad.

Hoy en día las empresas deben de contar con sistemas de información que les permita obtener información confiable y que les ayude a la toma de decisiones. Los progresos en las denominadas TI, que abarcan los equipos de cómputo y las aplicaciones informáticas y de telecomunicaciones, están teniendo un gran efecto. De hecho, se dice que estamos en un nuevo tipo de sociedad llamada Sociedad de la información o Sociedad de Conocimiento, que viene a reemplazar a los dos modelos socioeconómicos precedentes.

Con el empleo de la tecnología web, puede gestionarse la información de la organización relacionada con los recursos humanos, las finanzas, los procedimientos a implantar en la organización, la información financiera interna y hasta noticias de interés para todos los trabajadores.

Brindan, además, la posibilidad de la captura, almacenamiento, transformación, recuperación y difusión de la información, así como la digitalización y la conmutación de ésta, para colocarla al alcance de los usuarios. Permite la organización de la información que posee la institución y la pone a disposición de todo el personal, para utilizarla en el momento oportuno para la toma de decisiones. De manera que esto ha conducido al desarrollo de sistemas de información para el manejo, tratamiento y uso de la información en las organizaciones que posibilitan la agilidad y facilidad de acceso necesaria para ejecución de los procesos organizacionales y la toma de decisiones.

Dado lo anterior, esta tesis propone el uso de diversos tipos de visualizaciones e interfaces, montados a través de medios de tecnología web, con el fin que los datos puedan ser compartidos, analizados y se puedan mantener actualizados por los miembros activos de la organización, en beneficio de los procesos del negocio.

Esta tesis se ha estructurado de la siguiente manera: en el capítulo 2 como antecedentes, se realiza una descripción de la problemática actual a la que se enfrentan las empresas en este tema; en el capítulo 3 como marco teórico básico se describen las definiciones necesarias que sustentan al presente trabajo; en el Capítulo 4 se presenta como caso de estudio de la implementación de visualizaciones en la cual se presenta la arquitectura que se seguirá para la resolución del problema a través de las siguientes secciones:

- ✚ Las principales necesidades de las empresas (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).
- ✚ 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, en el Capítulo 5 se exponen las conclusiones obtenidas del desarrollo y aplicación del modelo.

1.1 OBJETIVO

Mostrar la importancia de las visualizaciones de datos, para la representación de la información, y cómo es que éstas contribuyen al fortalecimiento de la gestión organizacional en una empresa, desarrollando acciones aplicables a los procesos administrativos en contextos organizacionales, buscando la sostenibilidad y la competitividad en los entornos complejos de los mercados nacionales y globales.

1.2 OBJETIVOS PARTICULARES

- ✚ Recopilar y analizar los diferentes tipos de visualizaciones de datos, mostrando cuál es la más adecuada ante los objetivos que busque alcanzar la organización con esta, así como mostrar las buenas prácticas que se deben de seguir al realizar éstas.
- ✚ Diseñar una estrategia para la implementación de visualizaciones en alguna empresa o institución, que permita el logro de los objetivos organizacionales de corto y largo plazo.
- ✚ Aplicar una propuesta de visualización a un caso de estudio, buscando generar un desarrollo en la gestión organizacional de la institución.

2. ANTECEDENTES

En este capítulo se realiza una descripción de la problemática actual en las empresas, siendo el principal punto la forma en que se realiza el análisis de los datos para su mejor entendimiento. Exponiendo cómo es que, a través de las visualizaciones de datos, podemos atacar este problema, asimismo mostrando la importancia y ventajas que puede traer el tener una buena estrategia de implantación de estas, contribuyendo a mejorar y fortalecer la gestión organizacional.

2.1. SITUACIÓN ACTUAL

Cada vez más las organizaciones dependen de su infraestructura de IT (por sus siglas en inglés Information Technology, o bien Tecnología de la información y Comunicación TIC) para alcanzar sus objetivos. Pero en un entorno competitivo como el actual, aprovechar las oportunidades de negocio exige moverse con rapidez. Sin embargo, con frecuencia las TIC no permiten estas respuestas rápidas ni disponen de la flexibilidad necesaria para competir de forma efectiva.

Por lo mismo, un alto porcentaje de las ineficiencias organizativas tienen un mismo origen: el predominio de procesos manuales con un nivel de error elevado, sistemas ineficaces para compartir la información en el seno de la organización; la incapacidad de hacer un correcto seguimiento de los procesos de negocio desde principio a fin; conjuntamente el cumplimiento con las normativas legales aplicables obliga a manejar grandes cantidades de información en formatos complicados de utilizar y entender, lo cual convierte en un proceso pesado el manejo de la información, haciendo que la productividad de los empleados sea negativa, por consiguiente poniendo en riesgo la capacidad de crecimiento y competencia de la propia empresa.

En la raíz de todas estas deficiencias está la información. No es un problema de escasez de información, pues hoy en día nos encontramos en un mundo donde cada acción que realizamos nos genera una serie de datos. Sin embargo, por si mismos, no nos dan información completa. Para poder crear contenido a partir de ellos se necesita interpretarlos, agruparlos y comprenderlos, presentando la información de forma sencilla y útil a los usuarios y directivos de una manera coherente y sistemática.

En última instancia, esto se debe a que las aplicaciones de línea de negocio y otras aplicaciones antiguas normalmente dan soporte a funcionalidades avanzadas en áreas como la gestión financiera, marketing, control de clientes, etc. Pero sin poder compartir información entre ellas y, por consiguiente, no pueden aportar una visión general de los procesos de negocio cuando éstos abarcan varias áreas funcionales.

Estamos en una época en que todo entra por los ojos, en un mundo más visual, donde el principal sentido es la vista, y obtenemos el 80% de la información a través de este sentido, y el 20% de los sentidos restantes. ^[1]

Hemos visto cómo desde hace miles de años, desde la aparición del primer hombre, los seres humanos han dibujado para poder comunicarse, incluso, desde niños antes de aprender a leer o a escribir lo primero que hacen es dibujar para transmitir una idea.

Se ha comprobado que la información es más fácil de captar y recordar a través de una imagen. Una sobrecarga de información textual solo nos llevaría a caer en una saturación lo que representa uno de los principales obstáculos para lograr el objetivo de compartición y entendimiento de la información.

En la actualidad, ya no es complicado conseguir cualquier tipo de información, dentro de una empresa, existen gran cantidad de medios que la capturan y renuevan minuto a minuto, pero ¿Cuántos datos de los que nos son presentados diariamente podemos retener en un solo día? ¿Logramos entender realmente lo que significa esa información?, estas son algunas de las preguntas que nos muestra la problemática que existe en empresas respecto a la información.

Día a día y en distintas circunstancias, buscamos usar la información para comunicar una idea de forma clara y objetiva. La información no es nada más que un conjunto de datos organizados que logramos comprender y asimilar a través diversos medios, ya sea textual, a través del uso de una visualización o infografía. Ambas posibilidades nos sirven a nosotros mismos para explorar, analizar y contextualizar la información para poder transmitirla en nuestras propias palabras.

El diseño de la información que utilizamos y consumimos es muy importante para nuestra comprensión. (Friedman, 2008) menciona “El objetivo principal de la visualización de los datos es transmitir información clara y eficaz a través de medios gráficos. Esto no significa que la visualización de datos tenga que ser aburrida para ser funcional o extremadamente sofisticada para ser bella.”.

Es imprescindible preocuparse en el diseño del formato estético de la información pues como se ha dicho previamente, el ser humano recibe el mayor porcentaje de información a través del sentido de la vista, lo cual quiere decir que cualquier mensaje

[1] BRUNO MUNARI, Diseño y Comunicación Visual: Contribución a una Metodología Didáctica, Ed. Gustavo Gili.

que se reciba a través de los ojos, tendrá mayor captación en el cerebro que el resto de los mensajes que pudieran captarse a través de oídos, olfato o tacto, convirtiendo a estos últimos en meros acompañantes de la información visual.

Retomando, el hecho de que la vista capta el 80% de la información se desprenden las teorías semióticas en la que se menciona que la estructura más importante se da a través de símbolos o signos visuales. De tal forma que para poder hacer que un mensaje se capte con mayor facilidad y rapidez, deberán utilizarse símbolos gráficos, tales como colores o formas dotadas previamente con significados convencionales.

Esta es una de las razones por la cual la publicidad gráfica es la más socorrida al momento de crear estrategias de publicidad, pues el mensaje llegará directa y rápidamente al público meta elegido.

En esta época donde a cada segundo nos llega información es importante recibir y leer la más relevante. No debemos confundir cantidad con calidad pues lo más probable es que terminemos saturados de información que solamente nos va a confundir y desinformar. Como hace mención, Ignasi Alcalde(2015), en su libro Visualización de la información "La imagen nunca es menospreciada por ninguna persona. Cualquiera, hasta la menos indicada puede mostrar un mínimo de interés hacia una figura, pues se llega a interpretar la información visual más que la escrita".

Una de las revelaciones que hace Stephen Few es que los humanos solo logran recordar el 20% de lo que leen; es decir, logran más recordar lo que observan que lo que se consume de forma textual.

Como medio de comunicación, si logramos compartir contenido visualmente atractivo y de información relevante de la institución, esta no solo logrará ser transmitida a los empleados de la empresa o institución, sino que también significará que estos lo vean como un aporte relevante de información que puede ser entendido a primera vista por todos los usuarios.

El potencial de la información visual no ha hecho nada más que despegar en los últimos años. Esto se debe a que una visualización de datos explica historias, simplifica, mide, compara, explora y descubre nuevas cosas con el único objetivo de transformar los datos en información y en mero conocimiento. [2]

[2] IGNASI ALCALDE, Visualización de la información de los datos al conocimiento, Ed. Tapa Blanda

3. MARCO TEÓRICO

En este capítulo se describen los precedentes que justifican el desarrollo de este trabajo, para el cual se revisaron diferentes fuentes literarias como: arquitecturas de referencia, artículos realizados por empresas comerciales, publicaciones de libros, publicaciones web, etc.

Estas fuentes se describen en este capítulo con el fin de dar a conocer su importancia y relación con el tema de esta investigación, asimismo para sustentar en el siguiente capítulo la propuesta de la arquitectura solución.

3.1 TECNOLOGÍAS WEB

Se comienza por comprender el concepto central del desarrollo de este documento: tecnologías web. Pues en diversas ocasiones nos hemos preguntado, por que considerar a la Web como una tecnología, y el por qué la fuerza que han tomado en los últimos años, es algo de lo que trataremos de explicar en este apartado.

Las tecnologías Web se definen como las herramientas que nos sirven para acceder a los recursos de conocimiento disponibles en Internet o intranets utilizando un navegador, facilitando la realización de diversas tareas.

Actualmente han tomado gran fuerza, pues es una tecnología que nos facilita el desarrollo de sistemas de Gestión del Conocimiento, además de su flexibilidad en términos de escalabilidad.

Internet, Intranet o extranet permiten a los usuarios el acceso a una gran cantidad de información, tales como lectura publicaciones periódicas, búsqueda de información, realización de compras electrónicas entre otras muchas funciones. Gracias a la forma en que está organizada la *World Wide Web* (WWW), los usuarios pueden saltar de un recurso a otro con facilidad.

3.2 VISUALIZACIONES DE DATOS

En la sociedad de la información y el conocimiento en la que estamos inmersos la idea de que “la información es poder” está muy extendida en nuestros días, pero es errónea ya que sería más adecuado el aforismo del filósofo americano David Lewis “El conocimiento es poder, la información no”.

En nuestro día a día, estamos inundados de gran cantidad de información, es decir la información nos rodea, nos desborda y apenas somos capaces de vislumbrar su poder. Aunque pensemos que esta forma de consumir información, impulsiva, masiva, sesgada y en tiempo real, nos genera seguridad y conocimiento más bien nos lleva a nuevas formas de ignorancia. La moneda del siglo XXI ya no es la información sino el conocimiento.

Así mismo, si añadimos el concepto de Big Data el cual maneja grandes conjuntos de datos que generamos, podremos vislumbrar que la dificultad más habitual no es la captura y almacenamiento de los datos, sino más bien el análisis y su posterior representación visual, estática o interactiva. Algunos datos simplemente no tienen ningún sentido si no somos capaces de visualizarlos y contextualizarlos. A menos que los transformemos en información comprensible, no nos van a permitir acceder al conocimiento.

Las visualizaciones se llevan a cabo dentro de la disciplina que es el diseño visual, es una disciplina profesional que estudia los sistemas de información, con el objeto de convertir los datos en formas visuales, teniendo en cuenta los procesos perceptivos. Consiste en la creación de imágenes funcionales con fines netamente comunicacionales, para esto se hace uso de las nuevas tecnologías para un desarrollo más estructurado.

A nivel específico, el Diseño Visual propone:

- ✚ Brindar las herramientas conceptuales necesarias para identificar y comprender el proceso de constitución del lenguaje visual.
- ✚ Aplicar los conocimientos teóricos, técnicos y científicos; así como el manejo de los medios, instrumentos y materiales en la creación de imágenes para la información y la comunicación.
- ✚ Proyectar sobre la comunidad en general los conocimientos adquiridos, buscando que el diseño incida en el mejoramiento de la calidad de vida que el ser humano establece con el entorno.

- ✚ Aportar al sistema de producción regional y nacional, desde la construcción de nuevos sistemas de información, productos visuales competitivos que se validen a nivel nacional e internacional.

Como se muestra en la Figura 3.1, para que se pueda tener un producto de diseño visual exitoso, debe orientarse a tres vertientes:

- ✚ El público a quien va dirigido
- ✚ El contenido que se desea presentar
- ✚ La presentación de la información.

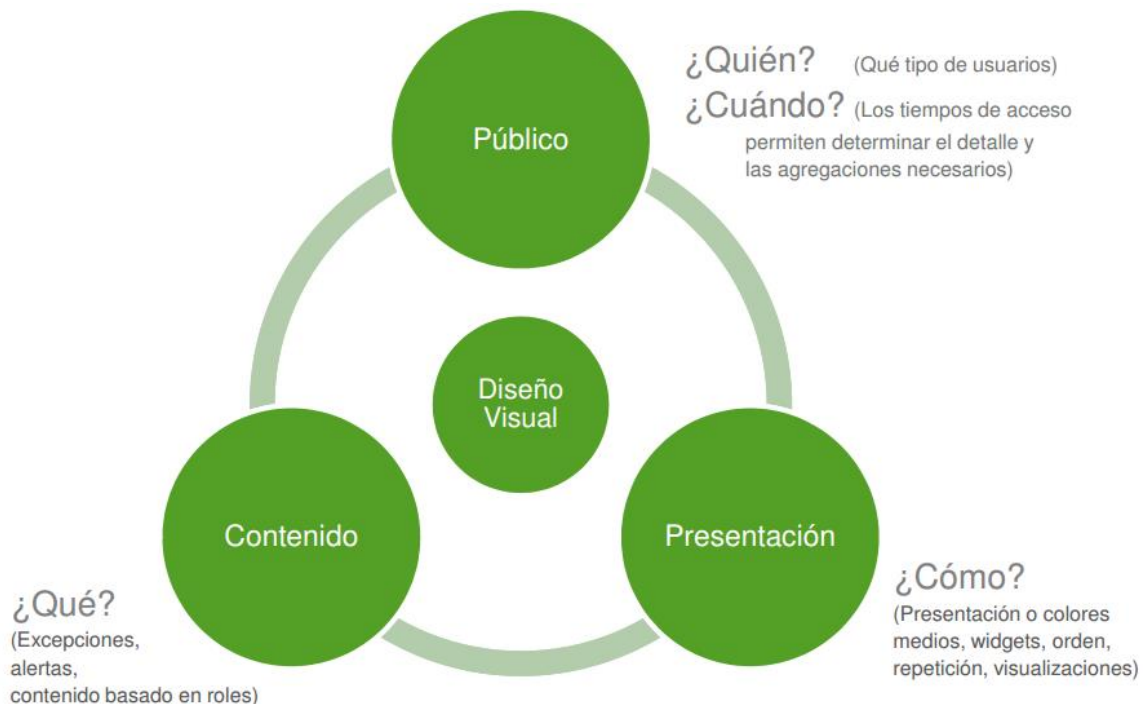


Figura 3.1 Características del diseño visual

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

Como un proceso de búsqueda, interpretación, contrastación y comparación de datos que permite un conocimiento en profundidad y detalle de los mismos de tal forma que se transformen en información comprensible para el usuario.

Es necesario conocer las diferencias entre infografía y visualización ya que en muchas ocasiones se considera que son los mismo; sin embargo, en el mundo del diseño de la información son cosas totalmente diferentes.

Tal vez el término infografía nos suene más familiar que la palabra visualización pues la primera la encontramos en revistas, diarios, periódicos murales, etc. Es un recurso muy utilizado en el periodismo y educación pues se utiliza para expresar diversos temas. La información se representa mediante símbolos e imágenes a los que se le suma un conjunto de datos para completar una historia. La infografía permite explicar acontecimientos, exponer procesos detalladamente y describir situaciones de una manera más visual y poco compleja que leer un texto.

Debemos dejar en claro que no a cualquier diseño se le puede llamar infografía pues para poder ser denominada como tal debe de aportar gran cantidad de información a la persona que la está visualizando. A comparación de una visualización, esta es estática y no permite al consumidor interactuar con la imagen como si la visualización de datos.

Estas dos herramientas nos ayudan a filtrar los datos irrelevantes centrándose únicamente en lo más pertinente para nuestro uso y permitiéndonos así recordar lo importante.

3.2.1 BUENAS PRÁCTICAS PARA LA CREACIÓN DE VISUALIZACIONES

Cuando las condiciones de visualización de datos son las adecuadas, los usuarios pueden interactuar con los datos de forma mucho más efectiva, logrando:

- ✚ Explorar las relaciones entre los datos.
- ✚ Identificar los vínculos entre la información presentada y fuentes comerciales y demográficas estándar.
- ✚ Integrar fuentes no relacionales, tales como los datos geo-espaciales.
- ✚ Dotar a su trabajo de mayores dosis de creatividad, que les permiten explorar nuevos horizontes.

Las visualizaciones de datos ayudan a la empresa y sus colaboradores a tener un mejor control de los datos, uno de sus activos más importantes y que, a menudo, no son tomados en cuenta, pues ayudan a crear una alineación organizacional, ya que todos los empleados ven lo mismo.

En el libro de Information Dashboard Design (Diseño del cuadro de mandos de información) y Show Me the Numbers (Muéstrame los números) de Stephen Few, nos muestran 5 puntos que se deben seguir, para tener una visualización buena, y que pueda llegar a cumplir con su objetivo de transmitir la información de forma sencilla y clara a las personas a las que se les está presentando. Estos cinco puntos se describen a continuación.

3.2.1.1. ELECCIÓN DE LAS MÉTRICAS IMPORTANTES

Seleccionar métricas importantes para incluir en la visualización es fundamental. Ante todo, deben ser métricas importantes y relevantes para lo que se desea comunicar. Sin embargo, esto no significa que se deba incluir cada métrica; al contrario, debe ser muy selectivo al determinar qué métricas merecen un espacio en la visualización.

Para encontrar las métricas correctas que debe incluir, tiene que considerar lo siguiente:

- ✚ ¿De qué forma contribuye cada métrica a los objetivos?
- ✚ ¿Tiene datos, internos o externos, que contribuyan a clarificar los objetivos?
- ✚ ¿Puede diseñar alguna métrica significativa para medir esas contribuciones?
- ✚ ¿Esa métrica es realmente necesaria para contribuir al logro de los objetivos?
- ✚ ¿Puede crear un método de medición sistemático y continuo?

Por ejemplo, en una institución educativa, en la cual se desea realizar una serie de visualizaciones de datos, con el objetivo de mostrar los factores que influyen en el aprovechamiento escolar, se presenta el siguiente modelo de base de datos, tal como se muestra en la Figura 3.2

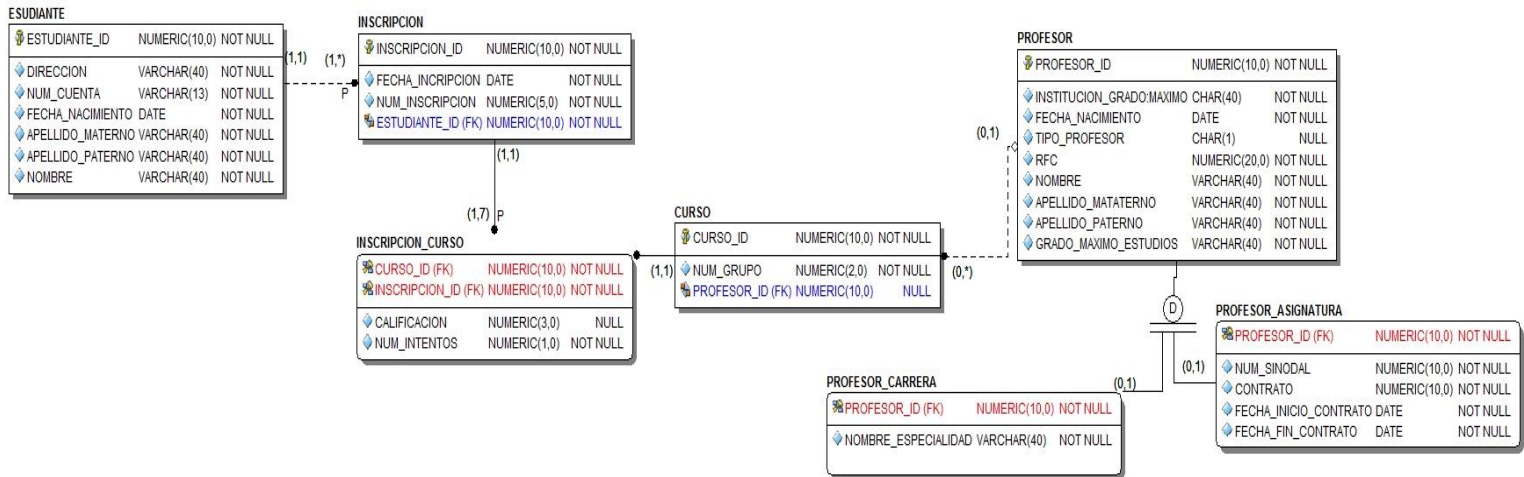


Figura 3.2 Modelo de base de Datos Escolar

Primero se debe analizar de qué forma contribuyen las métricas a los objetivos que se desean alcanzar, pues son algo que siempre se debe de tener en mente, en este caso el objetivo es mostrar los factores que influyen en el aprovechamiento escolar de los alumnos, por lo tanto, se debe de ser cuidadosos al seleccionar las métricas, que conformaran la visualización, pues si se seleccionan métricas que no aportan ningún valor se tendrá una visualización con sobre carga de datos. Por lo cual, como primer punto vemos que los datos personales del profesor son irrelevantes y no nos ayudan a nuestra visualización.

Como segundo punto se debe de checar si se tiene datos externos que ayuden a clarificar los objetivos, esto es importante, pues al momento de seleccionar las métricas para la visualización debemos de considerar aquellas ya sean de datos internos o externos a la institución, que ayuden a simplificar y entender de mejor manera los objetivos de esta, pues en ocasiones estos solo son claros para algunos, cuando lo que se busca con la visualización es que pueda ser entendida por todos, sin tener gran contexto de la institución, para este ejemplo aprovecharemos el dato de Institución donde se obtuvo el profesor obtuvo su máximo grado de estudios, el cual es un dato totalmente externo a la institución, pero puede contribuir a conocer cómo puede afectar este en el aprovechamiento de los alumnos que tuvo a su cargo.

Como tercer punto se debe checar si se puede diseñar alguna métrica para medir las contribuciones, ya que en ocasiones los datos por si solos que son almacenados en la base de datos, no dan ninguna información valiosa, pero al combinar 2 o más campos, podemos generar un dato que si nos proporcione información realmente útil y por lo

tanto, generamos una métrica de mayor importancia para la visualización de datos, en el caso del ejemplo, podemos ver que a partir de la dirección del alumno, podemos sacar el campo derivado de la distancia hacia la escuela, lo cual puede ser un factor que influya en su aprovechamiento.

Como cuarto punto, se debe verificar si la métrica seleccionada en verdad aporta a la contribución de los objetivos pues como hemos mencionado, el meter demasiadas métricas al generar una visualización, lleva a generar una, con sobrecarga de información, por lo que al ir integrando métricas se debe analizar si en verdad es una métrica útil, o solo distraerá nuestra atención, sin generar un valor importante al mismo, por ejemplo, si empezamos a considerar el número de grupo, fecha de inscripción o número de cuenta del alumno, son datos irrelevantes que no generan un beneficio alguno a la visualización, en cambio sobrecarga de información la visualización haciéndola más difícil de interpretar y analizar.

Como quinto punto se debe ver si se puede crear un método de medición sistemático y continuo, debido a que en toda visualización se deben de mostrar datos que sean actualizados, pues con datos desactualizados, no se tendrá en realidad un contexto de la situación actual, un ejemplo es buscar mantener los datos actualizados por semestre, así como llevar un control de las calificaciones por mes, en lugar de bimestre o semestre.

La prueba para saber si se han seleccionado las métricas importantes, consiste en saber si se puede explicar con claridad cómo cada métrica seleccionada en la visualización se relaciona con los objetivos de la organización.

3.2.1.2. VISUAL

Las visualizaciones o cuadros de mandos deben ser rápidos y fáciles de leer. Las tablas basadas en números no son rápidas ni fáciles de leer. Este es un caso en el que una imagen vale más que mil palabras.

El cerebro humano procesa números e imágenes como “fragmentos” discretos de información. Por lo tanto, las tablas y los informes repletos de números son difíciles y lentas de procesar. Sin embargo, es posible comprender un gráfico casi inmediatamente.

Cuando las personas tienen la opción de ver los datos como una imagen o como una tabla, se concentran en lo que dice la visualización. Esta es la razón subyacente por

la cual las visualizaciones o cuadros de mandos deben enfatizar la visualización de datos en gráficos y no el exceso de tablas.

Cuando se diseñe una visualización de datos, de deben incluir colores, formas, líneas, grosores, grados de sombreado, y otros tratamientos que mejoren la percepción visual, como se muestra en la figura 3.3. Entre las cosas que se debe evitar se incluyen los widgets demasiado atractivos, tratamientos de gráficos en tres dimensiones y paletas de colores que distraigan la atención.



Figura 3.3 Patrones a seguir al crear una visualización

3.2.1.3. INTERACTIVO

Como hemos visto las visualizaciones de datos debe tener métricas importantes y ser atractivamente visual; todo el mundo mirará la misma página. Pero una vez que estén en la misma página, los visores tendrán preguntas acerca de lo que ven.

Por lo que se debe buscar crear una visualización, de forma que cada visor pueda interactuar con ella y obtener las respuestas que buscan. Las visualizaciones interactivas permiten que la audiencia realice tareas estadísticas básicas, como:

✚ Filtrar vistas.

✚ Ajustar parámetros.

✚ Profundizar en la examinación de datos subyacentes.

Todo con selecciones intuitivas, pues con esta interactividad los visores pueden pasar de una visión general a una más particular, teniendo comprensión de cómo pueden usar los datos para tomar mejores decisiones en cuestión de uno a dos clics.

Un ejemplo donde se aplican estos 3 puntos lo podemos ver en la figura 3.4, donde en la parte superior tenemos la opción de ajustar los parámetros de la visualización, de igual forma al posar el apuntador sobre uno de los estados o distritos, según se haya seleccionado en los parámetros, se profundiza en la información de dicho estado o distrito, aplicando diferentes filtros, de forma más sencilla sobre la visualización.

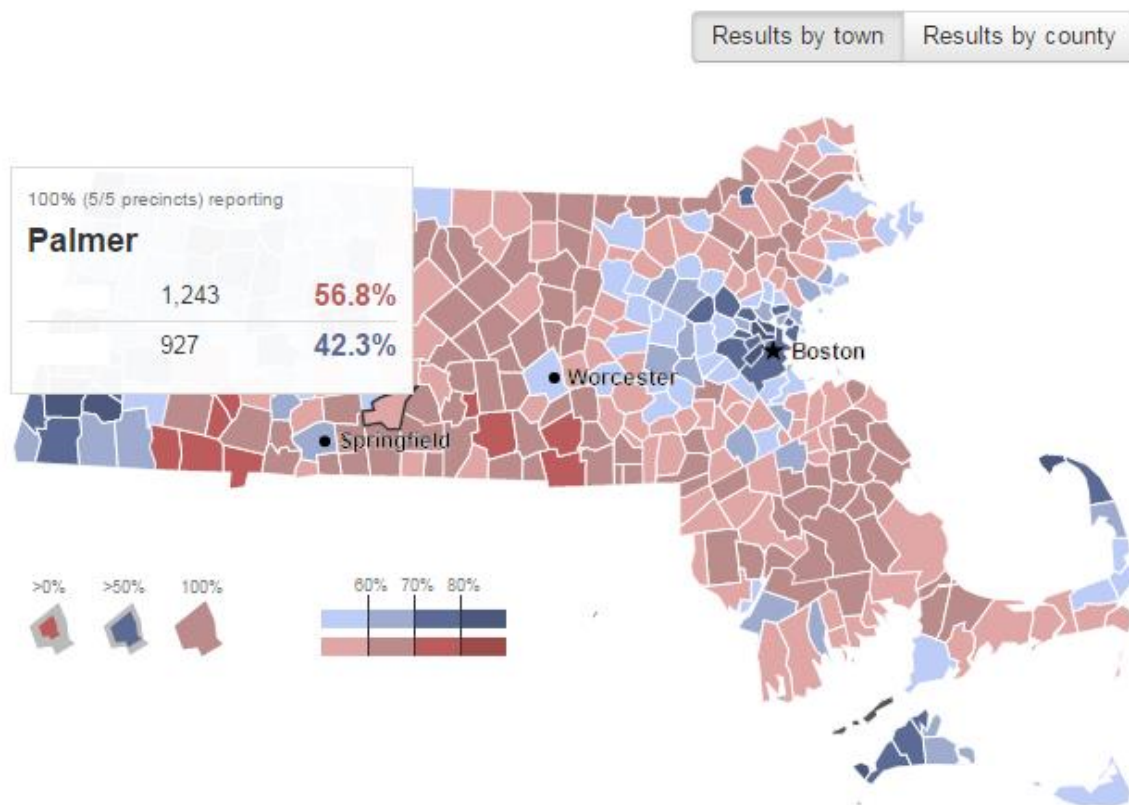


Figura 3.4 Ejemplo Visualización Interactiva

3.2.1.4. MANTENERLO ACTUALIZADO

Se debe asegurar de que los datos subyacentes en la visualización de datos estén actualizados y de que las métricas seleccionadas reflejen los desafíos actuales del negocio. Debido a que los datos viejos pueden prestarse para que haya una falsa impresión de seguridad en la toma de las decisiones. Pensará que está tomando una decisión basada en datos, pero estos ya no son representativos o importantes para la situación actual. De igual forma considerar que con el tiempo, las métricas clave también pueden cambiar, por lo que se deberían de actualizar las métricas seleccionadas.

3.2.1.5. ACCESIBILIDAD

Esta mejor práctica es fundamental. No importa si se han cumplido los puntos anteriormente mencionados, nadie utilizará la visualización si no tienen acceso a ella. La distribución mediante uso de tecnologías web es ideal, en especial si esto nos permite extraer datos actuales, pues nos permite tener una actualización constante de los datos, así como adherirse a una seguridad propia de la empresa donde se esté distribuyendo.

La distribución basada en archivos presentará problemas de sincronización de datos (es decir, las personas leen versiones antiguas) y es posible que la seguridad sea más engorrosa.

El resultado de la aplicación de estos puntos como buenas prácticas a la hora de la creación de visualizaciones de datos, es el enriquecimiento de la organización, que disfruta de un nivel de conocimiento como nunca antes teniendo una excelente herramienta para tomar decisiones más inteligentes y mejorando la productividad de la empresa.

3.2.2. ERRORES MÁS COMUNES AL DESARROLLAR VISUALIZACIONES

Por el contrario, las visualizaciones mal creadas pueden tener efectos muy negativos para el negocio llegando a:

- ✚ Confundir a los usuarios
- ✚ Dificultar la comprensión de los datos expuestos.

✚ Complicar el procesamiento de la avalancha de datos diarios

✚ Hacer perder la confianza usuaria en su inteligencia de negocios

Es, por ello, imprescindible garantizar las mejores condiciones para la visualización de datos de forma que asegure que logre cumplir con sus objetivos y logre aumentar el rendimiento organizacional, en vez de suponer una traba para su desarrollo.

Caer en alguno de estos errores alejará a la institución de las metas propuestas, dificultando el fortalecimiento de la gestión organizacional en una empresa. A continuación, se mencionan algunos ejemplos.

3.2.2.1. COMENZAR CON DEMASIADA COMPLEJIDAD

Es fácil volverse demasiado ambicioso y querer crear visualizaciones altamente detalladas que cubran cada desafío para el negocio y que ofrezcan a los usuarios una gran cantidad de opciones para profundizar.

La solución a este obstáculo es no pasar grandes cantidades de tiempo buscando crear la visualización “perfecta”. Por el contrario, trabajar en ciclos cortos e ir fortaleciendo la visualización, en cada iteración, tal como se es propuesto en este trabajo.

3.2.2.2. USAR MÉTRICAS QUE NADIE ENTIENDE

Es posible que la forma en que muestra y etiqueta las métricas tenga sentido para el creador de la visualización, pero ¿Las pueden entender los demás “los usuarios finales”?

Las métricas pueden ser tan familiares que el encargado de la creación de la visualización, olvide el paso de validarlas para los demás. Por lo que se debe asegurar de que las métricas escogidas no solo soporten los objetivos de la visualización, sino que también tengan sentido para la audiencia que esta tendrá.

3.2.2.3. ATIBORRAR LA VISUALIZACIÓN CON GRÁFICOS DE BAJA CALIDAD.

Se debe evitar realizar una visualización demasiado suntuosa o con un diseño exagerado, con gráficos y widgets muy complejos. A pesar de la información que puedan presentar, le impedirán lograr el objetivo de la visualización, que es el informar a la audiencia de manera rápida y fácil.

Se debe recordar y tener siempre presente que el atractivo visual de una visualización debe ser simple, para que esta pueda ser entendida de una sola vista.

3.2.2.4. ESPERAR POR TECNOLOGÍA COMPLEJA Y GRANDES PROYECTOS DE IMPLEMENTACIÓN DE INTELIGENCIA COMERCIAL.

Con frecuencia, las implementaciones tradicionales de inteligencia comercial (BI Business Intelligence) tardan mucho más de lo que se anticipa en un principio. Por lo que esperar a que se materialice un proyecto de BI tradicional puede significar meses o años de retraso, en comparación de la implementación de visualizaciones de datos, pues este es un proceso más complejo, cuyo ciclo de vida comprende varias etapas, como podemos observar en la Figura 3.5 y cada una requiere un análisis previo para su implementación, por lo que llega a ser más tardado que implementar una estrategia para visualizaciones de datos.

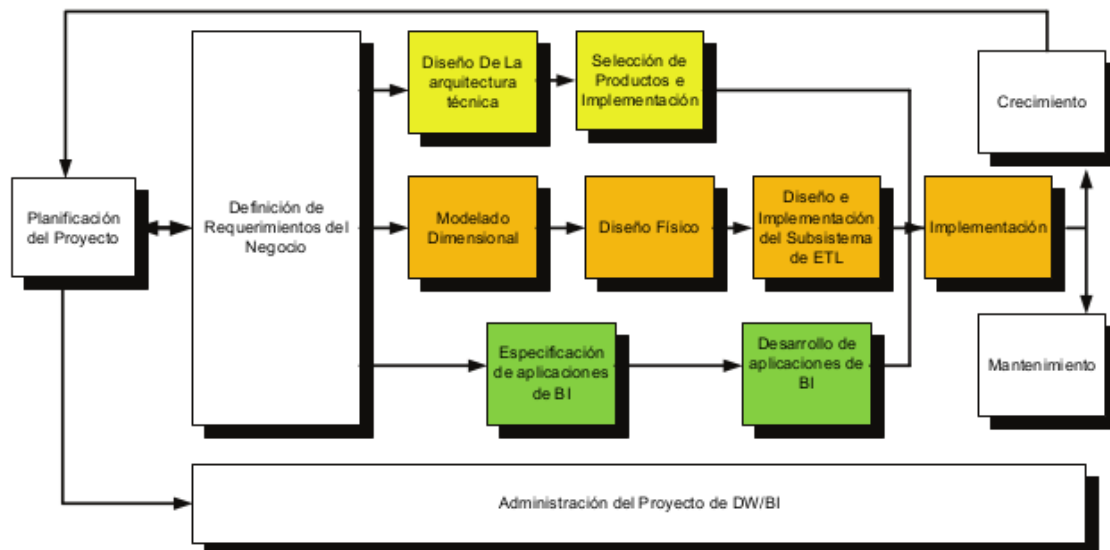


Figura 3.5 Ciclo de Vida para implementación de BI

Y es que para poder crear una visualización que ayude en la toma de decisiones y ayude a fortalecer la gestión empresarial en una compañía, se deben seleccionar las métricas verdaderamente importantes, las cuales son analizadas previamente con un proceso de BI; siendo este análisis un proceso que busca que las métricas estén estrechamente alineadas con la estrategia de la compañía y con sus capacidades esenciales, buscando abarcar un conjunto equilibrado de indicadores que señalan áreas que requieren atención y anticipen oportunidades (leading) así como los indicadores que miden el pasado (lagging).

Es por ello, que la solución es hacer uso de aplicaciones que ayuden a construir y distribuir cuadros de mandos con rapidez, pues con esta nos trae grandes beneficios como:

- ✚ Ayuda en la toma de decisiones.
- ✚ Da respuestas más rápidas.
- ✚ Se puede obtener información importante sobre el comportamiento de la institución o factores externos.
- ✚ Mayor facilidad de acceso y distribución, etc.

3.2.2.5. SUBESTIMAR LA NECESIDAD DE MANTENIMIENTO.

Desarrollar una visualización no es un ejercicio único. Si no se registra con la audiencia y se validan las métricas que aún son relevantes y los datos no se actualizan, no se utilizará la visualización, quedando inservible para conseguir los objetivos con los que fue hecho.

Por lo que debe de asegurarse de que se esté validando la utilidad de la visualización sobre una base regular.

3.2.2.6. QUE LAS MÉTRICAS NO COINCIDAN CON LAS METAS.

Los cuadros de mandos son más útiles cuando están relacionados con objetivos más generales. No se debe caer en la trampa de crear cuadros de mandos que no reflejen cómo las métricas clave de su departamento influyen en el éxito de la organización en general.

3.2.2.7. USAR GRÁFICO INEFICIENTE Y MAL DISEÑADO.

Los gráficos deben ser diseñados cuidadosamente. Por ejemplo, los gráficos en tres dimensiones no aumentan la comprensión del visor. Los colores chillones pueden interferir con la interpretación. A menudo, usar un gráfico circular es ineficiente, en especial si se representan más de seis valores.

3.3 ARQUITECTURA DE SERVICIOS.

La capacidad para responder rápidamente ante los cambios y optimizar los procesos de negocio es un factor clave para la competitividad y el crecimiento de las organizaciones. La agilidad de éstas puede verse cuestionada si se apoya en entornos de las TIC que no pueden responder de forma flexible a los cambios que afectan a la actividad de negocio. Liberar el potencial que poseen las aplicaciones y recursos de las TIC y hacerlo disponible de forma general a toda la organización facilita la optimización de procesos y mejora la agilidad empresarial

La arquitectura orientada a servicios (SOA, siglas del inglés *Service Oriented Architecture*) no se trata de software o de un lenguaje de programación, SOA es un marco de trabajo conceptual que permite a las organizaciones unir los objetivos de negocio con la infraestructura de TIC integrando los datos y la lógica de negocio de sus sistemas separados, permitiendo a empleados, clientes y socios comerciales responder de forma más rápida y adaptarse adecuadamente a las presiones del mercado.

La Arquitectura SOA supone una estrategia general de organización de los elementos de las TIC, de forma que una colección abigarrada de sistemas distribuidos y aplicaciones complejas se pueda transformar en una red de recursos integrados, simplificada y sumamente flexible.

SOA establece un marco de diseño para la integración de aplicaciones independientes de manera que desde la red pueda accederse a sus funcionalidades, las cuales se ofrecen como servicios.

La forma más habitual de implementarla es mediante Servicios Web (*Web Services*), una tecnología basada en estándares e independiente de la plataforma, con la que SOA puede descomponer aplicaciones monolíticas en un conjunto de servicios e implementar esta funcionalidad en forma modular.

Un proyecto SOA bien ejecutado permite alinear los recursos de TIC de forma más directa con los objetivos de negocio, ganando así un mayor grado de integración

con clientes y proveedores, proporcionando una inteligencia de negocio más precisa y más accesible con la cual se podrán adoptar mejores decisiones, ayudando a las empresas a optimizar sus procesos internos y sus flujos de información para mejorar la productividad individual. El resultado neto es un aumento muy notable de la agilidad de la organización.

3.3.1 SERVICIO.

Un servicio es una funcionalidad concreta que puede ser descubierta en la red y que describe tanto lo que puede hacer como el modo de interactuar con ella.

Desde la perspectiva de la empresa, un servicio realiza una tarea concreta, puede corresponder a un proceso de negocio tan sencillo como introducir o extraer un dato. Pero también los servicios pueden acoplarse dentro de una aplicación completa que proporcione servicios de alto nivel, con un grado de complejidad muy superior.

La estrategia de orientación a servicios permite la creación de servicios y aplicaciones compuestas que pueden existir con independencia de las tecnologías subyacentes. En lugar de exigir que todos los datos y lógica de negocio residan en un mismo ordenador, el modelo de servicios facilita el acceso y consumo de los recursos de IT a través de la red, buscando como se muestra en la Figura 3.7, un menor acoplamiento entre el FrontEnd y el BackEnd, ya que se tiene la capa intermedia de los servicios, siendo esta la que se entera de los cambio que se llegarán a realizar en el BackEnd, modificando solo este si fuera necesario, sin tocar la parte Frontal de la aplicación, es por esto su importancia y como un servicio puede ser consumido por diferentes aplicaciones.

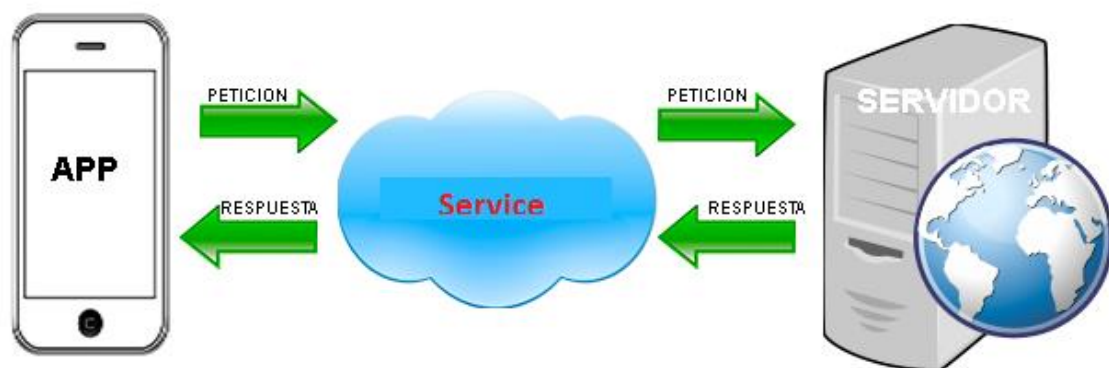


Figura 3.7 Ilustración de Arquitectura de Servicios

Los servicios están diseñados para ser independientes, autónomos y para interconectarse adecuadamente. Pueden combinarse con suma facilidad en aplicaciones complejas que respondan a las necesidades de cada momento en el seno de una organización.

Las organizaciones que adoptan la orientación a servicios pueden crear y reutilizar servicios y aplicaciones y adaptarlos ante los cambios evolutivos que se producen dentro y fuera de ellas, y con ello adquirir la agilidad necesaria para ganar ventaja competitiva.

3.3.2 WEB SERVICES.

Los Web Services o Servicios Web son aplicaciones que utilizan estándares para el transporte, codificación y protocolo de intercambio de información.

Los servicios Web permiten la intercomunicación entre sistemas de cualquier plataforma y se utilizan en una gran variedad de escenarios de integración, tanto dentro de las organizaciones como con socios (*partners*) de negocios.

La adopción de una solución de diseño basada en SOA no exige implantar servicios Web. No obstante, los servicios Web son la forma más habitual de implementar SOA.

Los servicios Web se basan en un conjunto de estándares de comunicación, como son XML para la representación de datos, SOAP (Simple Object Access Protocol) para el intercambio de datos y el lenguaje WSDL (Web Services Description Language) para describir las funcionalidades de un servicio Web.

A través de la descripción del servicio se obtiene la información que un consumidor necesita para considerar si usa o no el servicio. Con esta información se puede informar al consumidor: si el servicio existe (alcance), qué funciones realiza (funcionalidad), qué restricciones se aplican a su uso y cómo se debe interactuar con el servicio (interfaz) tanto en cuanto a formato como a secuencias.

La posibilidad de extender sistemas anexándole nuevos servicios y haciendo reuso de los ya existentes, con el objetivo de lograr interoperabilidad entre tecnologías y aplicaciones heterogéneas, nos permite prolongar la vida de los sistemas

3.3.3 WEB SERVICE BASADOS EN REST Y SOAP

REST (Representational State Transfer) es un estilo de arquitectura de software para sistemas hipermedias distribuidos tales como la Web, es un término utilizado en su

mayoría para describir a cualquier interfaz que transmite datos específicos de un dominio sobre HTTP sin una capa adicional, como lo hace SOAP.

La arquitectura REST debe cumplir con seis principios esenciales:

- ✚ Cliente – Servidor
- ✚ Interface Uniforme
- ✚ Capaces de almacenar en caché
- ✚ No manejan estado
- ✚ Sistemas en capas
- ✚ Código baja demanda

SOAP es un protocolo para el intercambio de mensajes sobre redes de computadoras, generalmente usando HTTP. Está basado en XML, esto facilita la lectura, pero también los mensajes resultan más largos y, por lo tanto, considerablemente más lentos de transferir, este tipo de servicio, hace uso de las UDDI, que son las siglas del catálogo de negocios de Internet denominado *Universal Description, Discovery and Integration*; tal como se muestra en el ejemplo de la Figura 3.8.

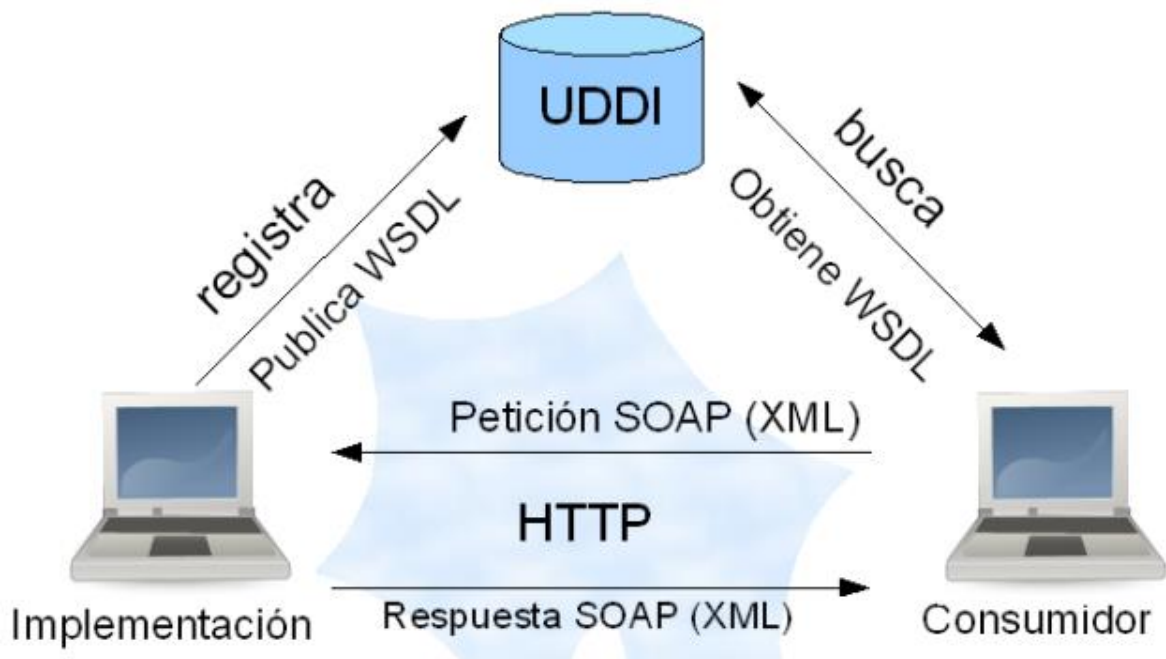


Figura 3.8 Interacción a través de Web Service SOAP

En la tabla mostrada en la Tabla 3.1, podemos ver una comparación entre los servicios REST y los SOAP

REST	SOAP
Las operaciones se definen en los mensajes.	Las operaciones son definidas como puertos WSDL.
Una dirección única para cada instancia del proceso.	Dirección única para todas las operaciones.
Cada objeto soporta las operaciones estándares definidas.	Múltiples instancias del proceso comparten la misma operación.
Componentes débilmente acoplados.	Componentes fuertemente acoplados.
Pocas operaciones con muchos recursos	Muchas operaciones con pocos recursos.
Mecanismo consistente de nombrado de recursos (URI)	Falta de un mecanismo de nombrado.
Se centra en la escalabilidad y rendimiento a gran escala para sistemas distribuidos hipermedia.	Se centra en el diseño de aplicaciones distribuidas.
Protocolo: HTTP GET, HTTP POST, HTTP PUT, HTTP DEL	Protocolo: SMTP, HTTP POST, MQ
Seguridad: HTTPS.	Seguridad: WS-Security.
Identificar recursos a ser expuestos como servicios	Listar las operaciones del servicio en el documento WSDL
Técnicas para añadir sesiones: Cookies	Técnicas para añadir sesiones: Cabecera de sesión (no estándar)

Tabla 3.1 Tabla comparativa entre SOAP y REST

3.3.4 BENEFICIOS DE USAR SOA

Los beneficios de SOA para una organización se plasman a dos niveles distintos: al del usuario corporativo y a nivel de la organización de IT. Desde el punto de vista de la empresa, SOA permite el desarrollo de una nueva generación de aplicaciones dinámicas que resuelven una gran cantidad de problemas de alto nivel, fundamentales para el crecimiento y la competitividad. El uso de SOA, nos ofrece beneficios tales como:

3.3.4.1 MEJORAR LA TOMA DE DECISIONES

Al integrar el acceso a los servicios e información de negocio dentro de un conjunto de aplicaciones dinámicas compuestas, los directivos disponen de más información y de mejor calidad (más exacta y actualizada). Las personas, procesos y sistemas que abarcan múltiples departamentos pueden introducirse de forma más directa en una panorámica unificada, lo que permite conocer mejor los balances de costes y beneficios que se producen en las operaciones de negocio que se realizan a diario. Y al disponer de mejor información en un tiempo menor, las organizaciones pueden reaccionar de manera más ágil y rápida cuando surgen problemas o cambios.

3.3.4.2 MEJORAR LA PRODUCTIVIDAD DE LOS EMPLEADOS.

Un acceso óptimo a los sistemas y la información y la posibilidad de mejorar los procesos permiten a las empresas aumentar la productividad individual de los empleados.

3.3.4.3 POTENCIAR LAS RELACIONES CON CLIENTES Y PROVEEDORES.

Los beneficios que ofrece SOA trascienden los límites de la propia organización. Los procesos de fusión y compra de empresas se hacen más rentables al ser más sencilla la integración de sistemas y aplicaciones diferentes. Con SOA se puede conseguir mejorar la capacidad de respuesta a los clientes. Si los clientes y proveedores externos pueden disponer de acceso a aplicaciones y servicios de negocio dinámicos, no solamente se permite una colaboración avanzada, sino que se aumenta la satisfacción de clientes y proveedores.

3.3.4.4 APLICACIONES MÁS PRODUCTIVAS Y FLEXIBLES.

La estrategia de orientación a servicios permite a las TIC conseguir una mayor productividad de los recursos existentes y obtener mayor valor de ellos de cara a la organización sin necesidad de aplicar soluciones de integración desarrolladas ex profeso para este fin. La orientación a servicios permite además el desarrollo de una nueva generación de aplicaciones compuestas que ofrecen capacidades avanzadas y

multifuncionales para la organización con independencia de las plataformas y lenguajes de programación que soportan los procesos de base.

3.3.4.5 DESARROLLO DE APLICACIONES MÁS RÁPIDO Y ECONÓMICO.

El diseño de servicios basado en estándares facilita la creación de un repositorio de servicios reutilizables que se pueden combinar en servicios de mayor nivel y aplicaciones compuestas en respuesta a nuevas necesidades de la empresa. Con ello se reduce el coste del desarrollo de soluciones y de los ciclos de prueba, se eliminan redundancias y se consigue su puesta en valor en menos tiempo.

3.3.4.6 APLICACIONES MÁS SEGURAS Y MANEJABLES.

Conforme van evolucionando las necesidades de negocio, SOA facilita la posibilidad de añadir nuevos servicios y funcionalidades para gestionar los procesos de negocio críticos. Se accede a los servicios y no a las aplicaciones, y gracias a ello la arquitectura orientada a servicios optimiza las inversiones realizadas en TIC potenciando la capacidad de introducir nuevas capacidades y mejoras.

3.4. BUSINESS PROCESS MANAGEMENT.

El concepto de BPM (*Business Process Management* o gestión de procesos de negocio) está también muy ligado a SOA. BPM es una disciplina de gestión que combina una visión centrada en procesos y de integración de funcionalidades que pretende mejorar la efectividad de las organizaciones.

El campo que abarca la gestión de los procesos de negocios incluye dos comunidades: la de la administración del negocio (analistas de negocio) y las de la TIC.

La gestión de procesos de negocio se basa en la idea de que cada producto es el resultado de un conjunto de actividades realizadas para obtenerlo. La tecnología de la información en general y los sistemas de información en particular, juegan un rol importante en la gestión de procesos de negocio dado que muchas de las actividades son soportadas por sistemas de información. El éxito en el alcance de los objetivos y su logro eficaz y eficiente depende del trabajo coordinado de los recursos que intervienen (sean estos humanos o tecnológicos).

Según Mathias Weske un proceso de negocio es un conjunto de actividades que se realizan en coordinación en un ambiente organizacional y técnico. Estas actividades

alcanzan el objetivo del negocio. Cada proceso de negocio representa una única organización, pudiendo interactuar con otras. ¹

BPM ha adquirido una atención considerable recientemente tanto por las comunidades de administración de negocios como las de ciencia de la computación. Los miembros de estas comunidades están identificados por diferentes soportes e intereses educacionales; entre ellos encontramos los analistas de negocios, que están interesados en mejorar las operaciones de las compañías

La gestión de procesos de negocio incluye conceptos, métodos y técnicas para soportar el diseño, administración, configuración, representación y análisis de los procesos de negocio. Su objetivo último es representar el proceso de negocio con sus actividades y las restricciones de ejecución entre ellas.

Podemos concluir como definiciones de Proceso de negocio y Gestión de Procesos de Negocio (BPM), las siguientes:

Los **Procesos de Negocio** son representados manualmente, guiados por el conocimiento del personal de la organización y asistidos por regulaciones organizacionales y procedimientos previamente instalados.

Los **Sistemas de Gestión de Procesos de Negocios (BPMS)** son implementados por un conjunto de programas (software) utilizados como herramientas para representar y coordinar las actividades involucradas en un proceso de negocio.

Las personas son una parte esencial de prácticamente cualquier proceso de negocio, pues son las responsables de aplicar las soluciones y disponer de la visión que hace avanzar a una empresa, por lo que el objetivo debe ser aumentar su capacidad para crear e innovar y ser más productivas, y no “hacer reingeniería” pretendiendo colocar a las personas fuera de los procesos.

Aunque BPM puede considerarse como una entidad al margen de las iniciativas SOA, la capacidad para definir nuevos procesos de negocio de forma flexible y rápida es mucho mayor si los recursos de los sistemas de TIC se exponen en la forma de orientación a servicios.

¹ Weske Mathias, “Business Process Management: Concepts, Languages

3.5 MARCO DE LA ARQUITECTURA EMPRESARIAL (TOGAF)

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

El modelo de arquitectura empresarial esta soportada en 4 arquitecturas, como se muestra en la Figura 3.9, las cuales son:



Figura 3.9 Componentes arquitectura empresarial

Arquitectura de Negocio: En la cual se busca identificar la cadena de valor de la organización desde macro procesos hasta subprocesos. Tras la identificación, pasa entonces a la Definición de la Arquitectura de Procesos de Negocio. Usa BPMN (Business Process Model and Notation) como técnica de modelamiento para proveer una notación estándar fácilmente leíble y entendible para los involucrados en el negocio, como lo son los analistas de negocio, los desarrolladores técnicos y los gerentes y administradores del negocio.

Arquitectura de Datos: Establece el modelo de gestión de todos los aspectos del ciclo de vida de la información, es decir, identifica el modelo de Entidades de Negocio y su relación con los procesos de negocio buscando así la forma de crear, almacenar, mover, utilizar y retirar los datos.

Arquitectura de Aplicación: Identifica esta arquitectura a través del levantamiento oficial del catálogo de aplicaciones actuales, la identificación de iniciativas en ejecución, y un análisis de cubrimiento de estas aplicaciones en los procesos de negocio.

Arquitectura Tecnológica: Se encarga de validar las capacidades de software y hardware que se requieren para apoyar la implementación de servicios de negocio, datos y aplicación. Esto incluye infraestructura de TIC, capa de mediación, redes, comunicaciones, procesamiento y estándares.

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 tal como se muestra en la Figura 3.11. De tal forma que en cada ciclo de ejecución se incremente la madurez de la organización y el valor que aporta al negocio.

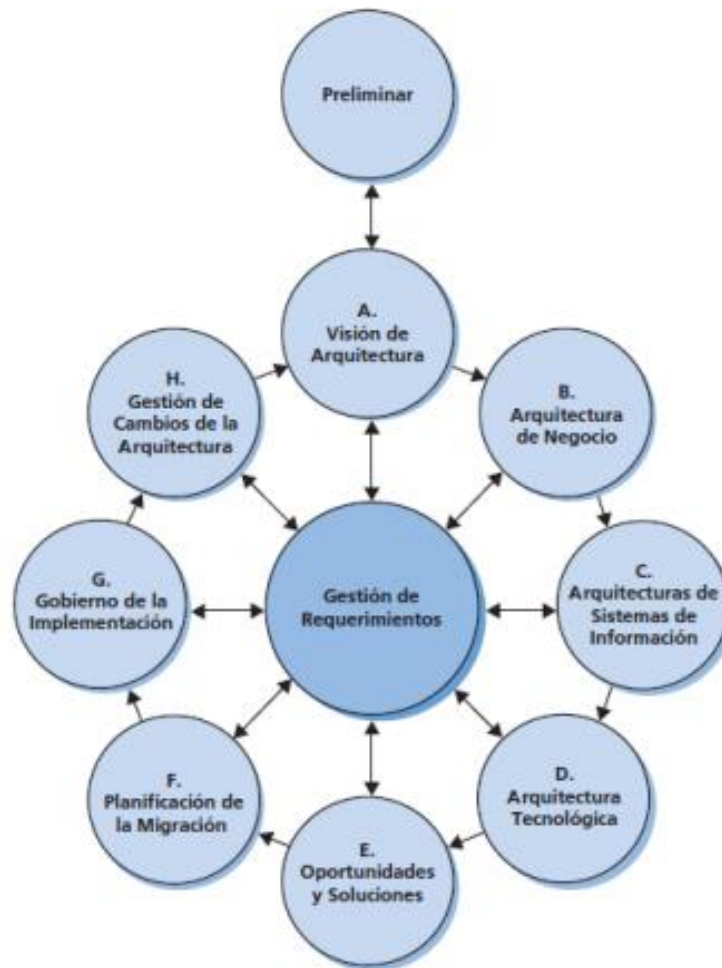


Figura 3.11 Método de desarrollo de Arquitectura, TOGAF 2011

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:** Prepara a una organización para emprender proyectos de Arquitectura Empresarial de manera exitosa
- ✚ **Fase A (Visión de la Arquitectura):** Aborda el establecimiento del proyecto e inicia una iteración del ciclo de desarrollo de la arquitectura, estableciendo el alcance, limitaciones y expectativas de la iteración. Se ejecuta con el objetivo de

validar el contexto del negocio y producir una Declaración de Trabajo de Arquitectura aprobada.

- ✚ **Fase B (Arquitectura de Negocio):** Aborda el desarrollo de negocio que apoye la visión de la arquitectura acordada.
- ✚ **Fase C (Arquitecturas de sistemas de información):** Aborda la documentación de la organización fundamental de los sistemas de TIC de una empresa, representada por los principales tipos de sistemas de información y aplicaciones que los utilizan. En esta fase hay dos pasos que se pueden llevar a cabo secuencialmente o simultáneamente: la Fase C y Fase D.
- ✚ **Fase D (Arquitectura tecnológica):** Aborda la documentación de la organización esencial de sistemas TIC, representada en hardware, software y tecnología de comunicaciones.
- ✚ **Fase E (Oportunidades y soluciones):** Esta fase se refiere a la implementación directamente. Describe el proceso de identificación de los medios de entrega (proyectos, programas o carteras) que proporcionan la Arquitectura de Destino identificada en las fases anteriores.
- ✚ **Fase F (Planificación de la Migración):** Aborda la planificación de la migración, es decir, cómo moverse desde la Arquitectura de la Línea de Base a la Arquitectura de Destino, finalizando un plan de implementación y migración.
- ✚ **Fase G (Gobierno de la Implementación):** La arquitectura delimita los proyectos de implementación, la supervisa al mismo tiempo que se la construye y produce un contrato de arquitectura firmado.
- ✚ **Fase H (Gestión de Cambios de la Arquitectura):** Asegura que los cambios en la arquitectura se gestionen de manera controlada.
- ✚ **Gestión de Requerimientos:** Se aplica a todas las fases del ciclo del ADM, a excepción de la fase preliminar. El proceso de Gestión de Requerimientos es un proceso dinámico que aborda la identificación de los requerimientos de la empresa, almacenándolos y luego gestionándolos al ingreso y egreso de las fases relevantes del ADM. Este proceso es fundamental para conducir el proceso del ADM.

3.6 GESTIÓN ORGANIZACIONAL.

Los cambios tecnológicos que se han producido en la sociedad han ampliado el campo de la gestión. La automatización, la informática, las nuevas tecnologías de la información y las crecientes expectativas de la sociedad han puesto al descubierto muchas carencias de los directivos. La naturaleza de la gestión se ha hecho más compleja para actuar en función de una serie de prioridades, como es la de conseguir beneficios constantes, por encima de todas ellas.

La gestión organizacional, se puede definir como el área cuya acción se encamina a definir la acción, el impacto y el efecto de la integración de los procesos de una organización, permitiendo guiar racionalmente el rumbo de una organización, generando las condiciones para la existencia, supervivencia y prosperidad de esta.

La visión de la gestión organizacional, es entendida como el desarrollo de estrategias de mejoramiento continuo en los procesos administrativos, puesto que las organizaciones tienen en cuenta a su personal en el desarrollo de actividades industriales, comerciales y de servicios que satisfacen necesidades generales y específicas.

Para la gestión de los diferentes procesos de la organización se requiere el establecimiento de tres elementos fundamentales:

1. **La definición de parámetros u objetivos.** - Para esto se debe conocer la situación actual de la empresa, sus fortalezas, debilidades y oportunidades, esto con base en el resultado de establecer estrategias y mecanismos para facilitar la implantación de un proceso de gestión.
2. **La medición de la acción real realizada.** - La medición a través de indicadores de gestión es el componente clave para tomar decisiones acertadas de negocio, así como, un mecanismo efectivo para comunicar las prioridades organizacionales.
3. **El análisis de los resultados mediante herramientas y técnicas de análisis estadístico.** - Es conocer con exactitud cómo se está frente al cumplimiento de cada requisito dado el tiempo y las etapas que han transcurrido, para de esta forma tener un análisis estadístico del avance que ha significado para la organización.

Una buena gestión empresarial está basada en cuatro factores imprescindibles:

- ✚ **Planificación:** Es indispensable tener una buena gestión empresarial desde la perspectiva de la planificación. Establece muy bien las metas y cada una de las actividades y estrategias que vas a seguir para llegar a tus objetivos empresariales, así como también los recursos financieros, humanos y técnicos que necesitas.

"Tener una visión global de la empresa y su entorno, tomando decisiones concretas sobre objetivos concretos"
Luther Gulik (2012)

- ✚ **Organización:** La organización es tan importante que puede marcar el éxito o fracaso de una empresa. Define, asigna y coordina cada una de las actividades que se van a llevar a cabo, quién las realizará, en qué tiempo y cómo lo hará. Una empresa exitosa tiene como característica principal la organización.

"Obtener el mejor aprovechamiento de las personas y de los recursos disponibles para obtener resultados"
Luther Gulik (2012)

- ✚ **Dirección:** No se trata de mandar. Esta función es impulsar y motivar a cada uno de los miembros de la organización. Una buena dirección debe mantener una serie de cualidades necesarias en el trato hacia los demás como la asertividad y la empatía.

"Un elevado nivel de comunicación con su personal y habilidad para crear un ambiente propicio para alcanzar los objetivos de eficacia y rentabilidad de la empresa"
Luther Gulik (2012)

- ✚ **Control:** Coordinar y supervisar el trabajo que se va realizando es de suma importancia. Con todo ello, se van conociendo las fortalezas y debilidades de la empresa, para extraer índices e indicadores de peso que permitan solucionar cualquier futuro inconveniente. El control se convierte al final en la herramienta necesaria para establecer medidas a corto y largo plazo, sin poner en riesgo la situación de la empresa.

"Cuantificar el progreso realizado por el personal en cuanto a los objetivos marcados"
Luther Gulik (2012)

Al terminar este capítulo hemos visto algunas de las buenas prácticas que se deben seguir para contar una visualización que cumpla sus objetivos y sea óptima, desde la selección de métricas para los datos que serán usados en esta, así como las buenas prácticas visuales y de datos en cuestión de accesibilidad y actualización con las que debe de contar una visualización.

Asimismo, vimos las características de la arquitectura SOA, y cuáles son los beneficios que nos trae el implementar un proyecto de software sobre esta, ya que el hacer uso de una arquitectura SOA es complementario al de una empresarial como lo es TOGAF.

De igual forma hemos visto la importancia de contar con y conocer un proceso bien definido que nos indica paso por paso cómo se debe desarrollar la arquitectura empresarial.

TOGAF como hemos visto es un framework para arquitecturas empresariales desarrollado por The Open Group. Es considerada una herramienta para asistir en la aceptación, producción, uso y mantenimiento de arquitecturas empresariales, basándose en un modelo de proceso iterativo soportado por buenas prácticas y un conjunto reusable de activos arquitecturales existentes.

Para TOGAF la arquitectura de TI empresarial no tiene que ver sólo con las aplicaciones de la organización. Pues, propone una arquitectura de TI que funcione para el negocio, es decir, que este alineada y apoye a las estrategias del negocio.

El ADM de TOGAF es una especie de “CMMI para arquitectura”. Nos presenta un proceso iterativo con fases que se deben realizar para desarrollar la arquitectura, pero sin indicarnos como hacer cada uno de los entregables.

Sin embargo, el ADM, como cualquier modelo de procesos, debe adoptarse gradualmente, sin intentar implementarlo de inmediato al 100% y en todos los proyectos.

4. CASO DE ESTUDIO

¿Por qué diseñar una solución basada en una AE (Arquitectura Empresarial)? 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. Asimismo, nos proporciona ventajas como:

- ✚ 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.

Para llevar a cabo este caso de estudio, se ha definido realizar diferentes iteraciones para el desarrollo de elementos, que son importantes para el proyecto. Estas iteraciones son las mostradas en la Figura 4.1.

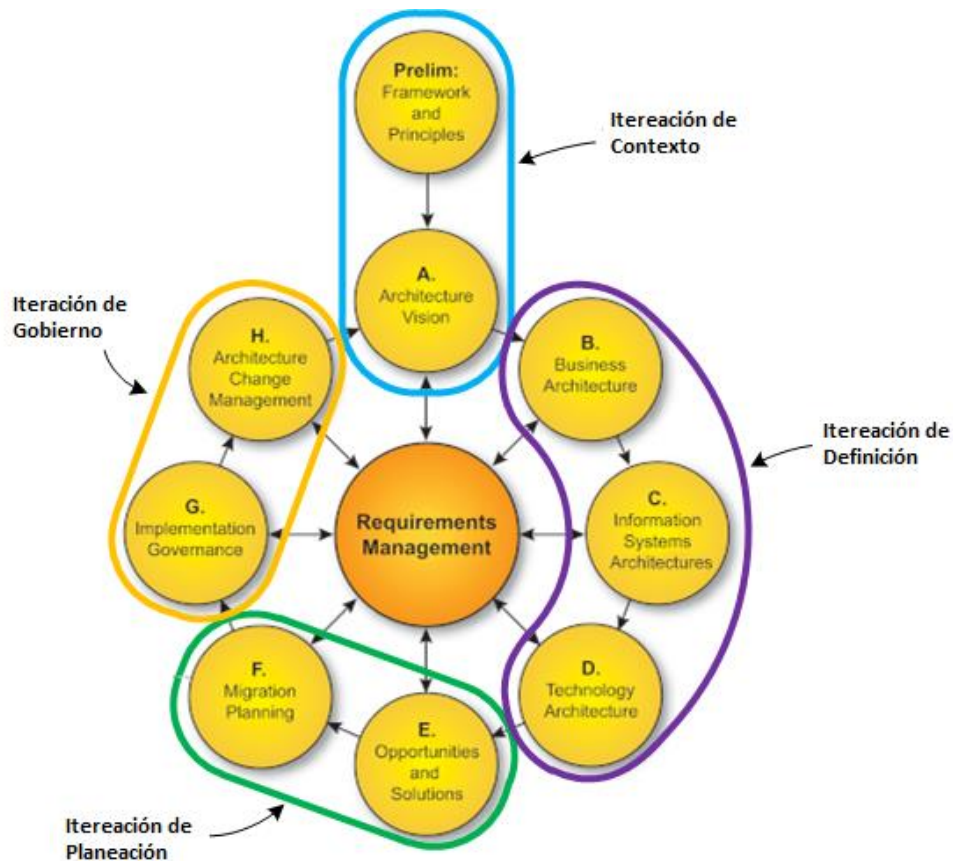


Figura 4.1 Iteraciones del modelo ADM

- ✚ La Iteración de Contexto comprende la fase preliminar y visión de la arquitectura, donde los elementos a desarrollar son la evaluación de madurez, el alcance, la misión y la visión de la capacidad arquitectónica a diseñar.
- ✚ Para la Iteración de Definición se consideran 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 servidores, manejadores, etc.) para soportar las anteriores fases

- ✚ Para la Iteración de Planeación se consideran las fases de Oportunidades y Soluciones, y Planeación de Migración. En estas fases se desarrollan las iniciativas para cumplir e implementar la arquitectura definida en los dominios previamente mencionados, así como se desarrolla un plan de migración.

- ✚ Para la Iteración de Gobierno se consideran 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 caso de estudio involucra Implementar una plataforma para hacer uso de visualizaciones de datos la cual pueda ser accedida a través de la intranet de la institución, con el objetivo que los trabajadores y personas interesadas en estos datos, puedan visualizar y analizar los datos de forma clara y sencilla permitiendo reducir tiempos y mejorando la toma de decisiones dentro de la institución.

Para esto, dividiremos el presente caso de estudio, en dos iteraciones completas sobre la arquitectura, la cual se basa en un modelo en espiral, en la primera iteración, se buscará realizar una limpieza de los datos, de forma que los datos presentados en las visualizaciones, sean los más confiables y actualizados para la toma de decisiones, mientras que para la segunda iteración, se llevará a cabo la creación de las visualizaciones y su presentación a los interesados, para los objetivos perseguidos.

4.1 PRIMERA ITERACIÓN (LIMPIEZA DE LOS DATOS)

4.1.1 ITERACIÓN DE CONTEXTO

4.1.1.1 FASE 0: PRELIMINAR

Institución: Institución Financiera.

Hoy en día, existe dentro de una institución financiera un área la cual provee información financiera, que recibe y procesa a unidades administrativas de la misma institución, 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 genere y provea productos de calidad es indispensable verificar la información recibida.

Estructura organizacional

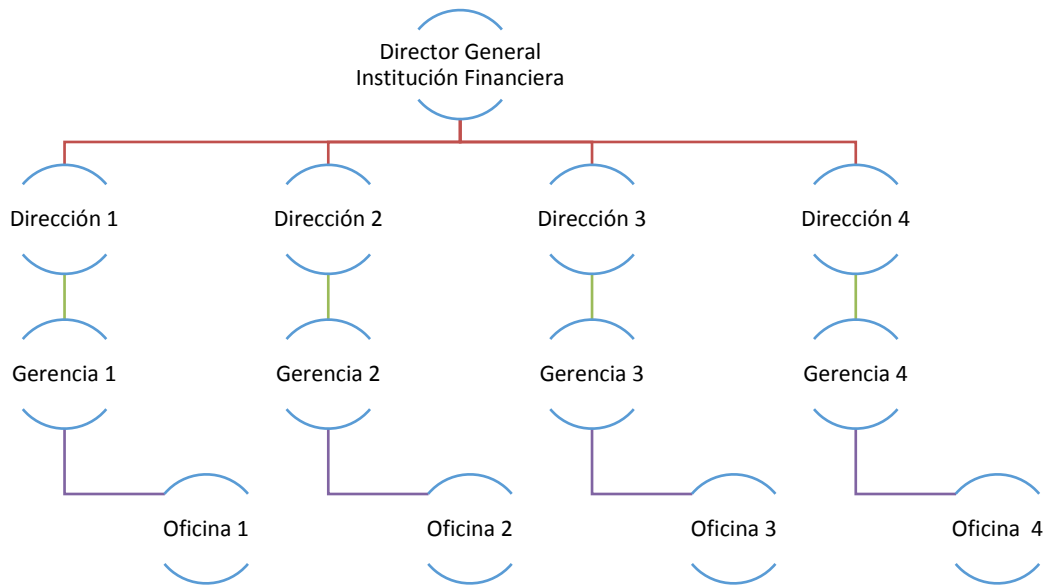


Figura 4.2 Estructura Jerárquica de la Institución

Evaluación de madurez

La evaluación de madurez de la organización para la implementación de se realizará con base en el MDM Maturity Mode, el cual describe un conjunto de conductas, tecnologías y respuestas utilizadas para crear y mantener datos coherentes, completos, contextuales y precisos de todas las partes (usuarios, aplicaciones, almacenes de datos, procesos y socios).

MDM no crea nuevos datos o nuevos entornos aislados de datos, sino que proporciona un método mediante el cual una organización puede gestionar de manera eficaz datos ya presentes en distintos sistemas de información.

Objetivos que se buscan al implementar MDM¹

- ✚ Mejorar la habilidad de una organización para ajustarse rápidamente a los requerimientos cambiantes del negocio
- ✚ Mejorar la eficiencia operacional
- ✚ Impulsar procesos de negocio
- ✚ Aumentar la calidad de datos
- ✚ Mejorar la eficiencia en la administración de la información
- ✚ Habilitar una integración de datos más amplia y compleja
- ✚ Eliminar actividades de administración de datos redundantes
- ✚ Eliminar actividades de integración redundantes
- ✚ Mejorar la toma de decisiones

El modelo define seis componentes, y cada uno de ellos contiene ciertas capacidades (Arquitectura, Gobierno, Administración, Identificación, Integración y Administración de procesos de negocio). Así mismo establece cinco niveles de madurez: inicial, reactivo, gestionado, proactivo y rendimiento estratégico. Al realizar esta evaluación y de acuerdo a los resultados respecto al nivel de madurez de la Arquitectura empresarial en el cual se encuentre la institución es posible identificar a dónde se orientan las inversiones en TI.

Para esto se ponen seis cuestionarios, los cuales buscan evaluar el nivel de madurez de los componentes:

¹ Bentley, J. E. (s.f.). "Master Data Management- What it is and why should care".

Cuestionario 1: Arquitectura

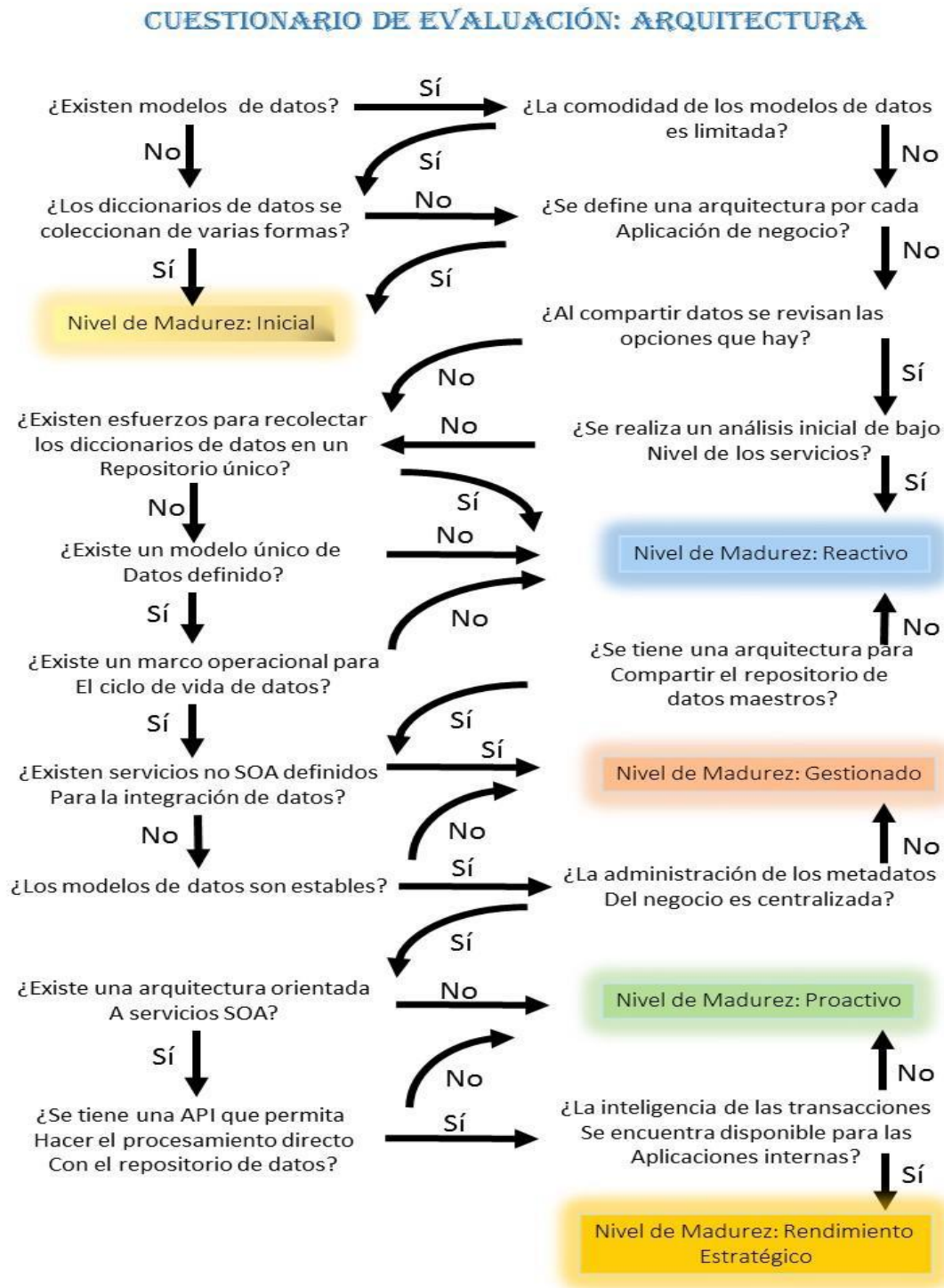


Figura 4.3 Cuestionario de Arquitectura. Adaptado de Loshin, 2009

Cuestionario 2: Gobierno

CUESTIONARIO DE EVALUACIÓN: GOBIERNO

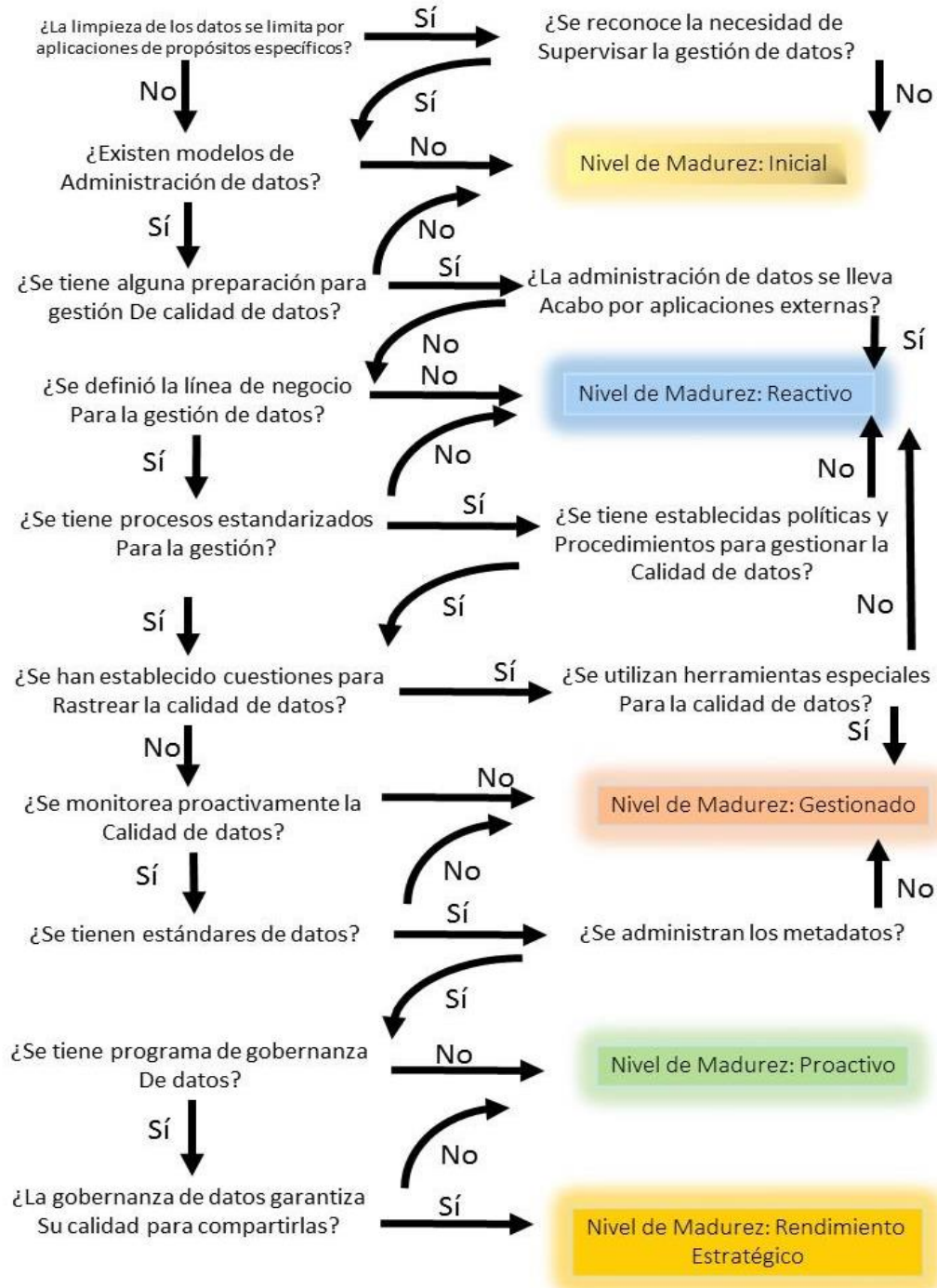


Figura 4.4 Cuestionario de Gobierno. Adaptado de Loshin, 2009.

Cuestionario 3: Administración

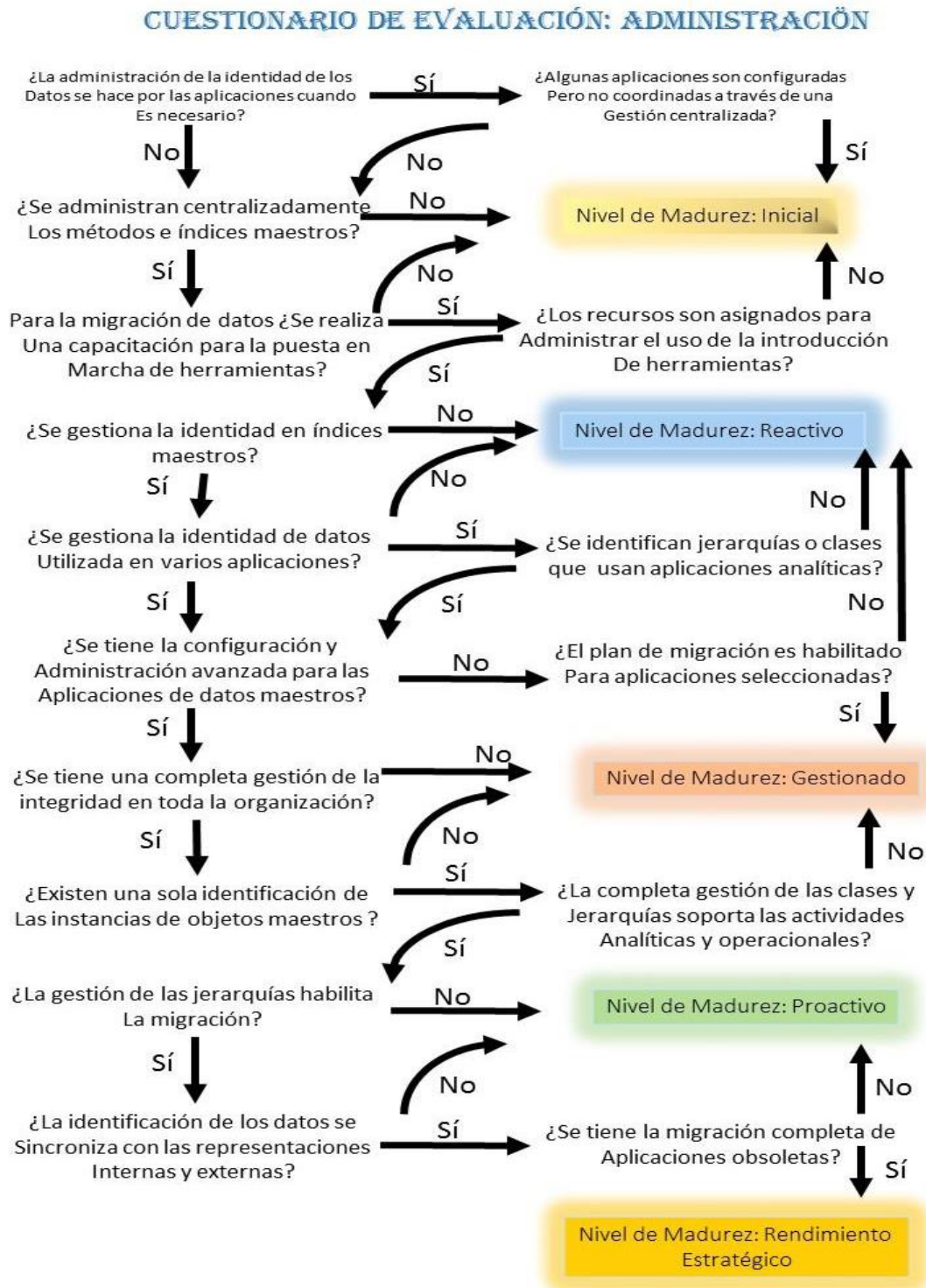


Figura 4.5 Cuestionario de Administración. Adaptado de Loshin, 2009

Cuestionario 4: Identificación

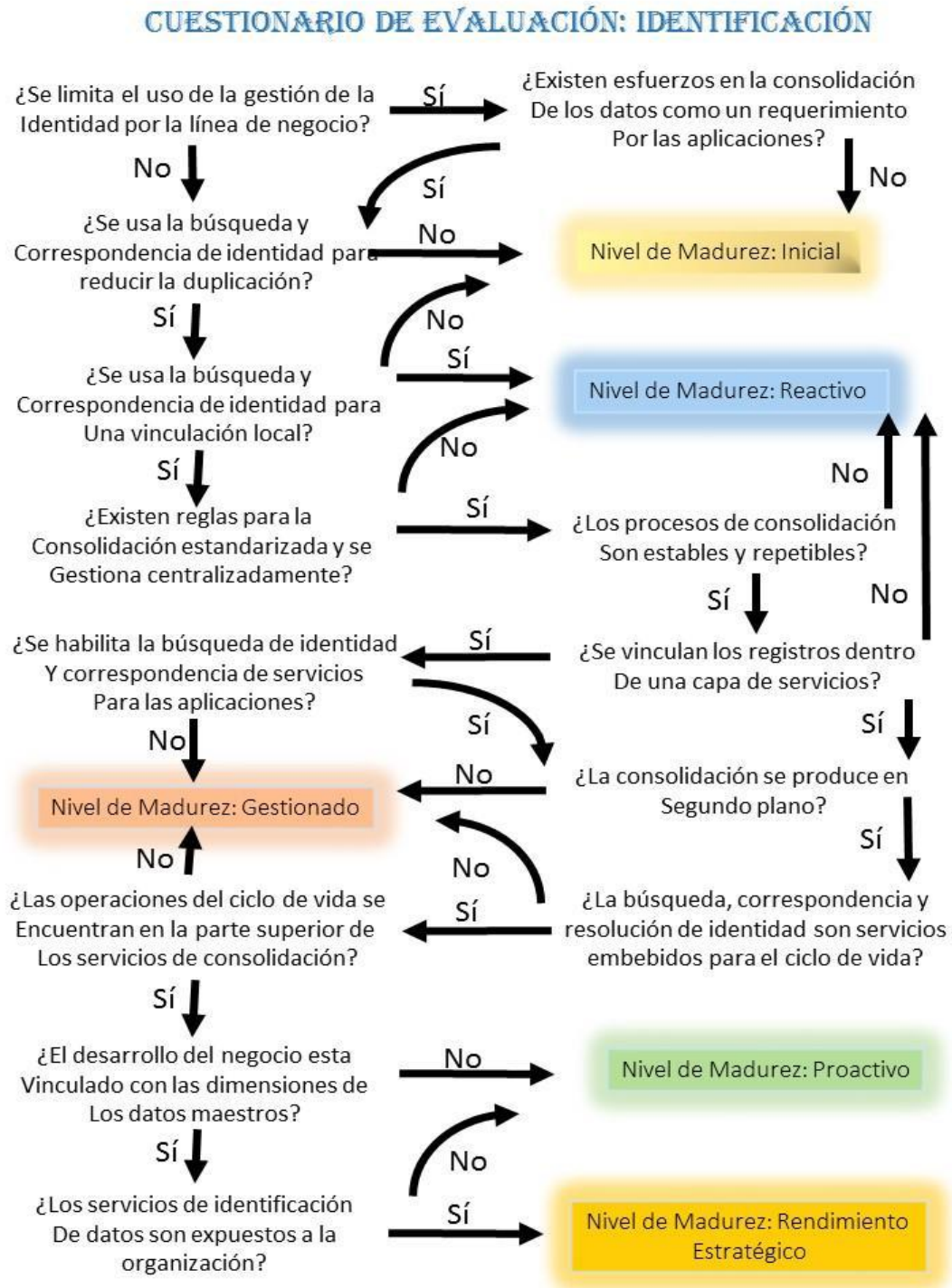


Figura 4.6 Cuestionario de Identificación. Adaptado de Loshin, 2009

Cuestionario 5: Integración

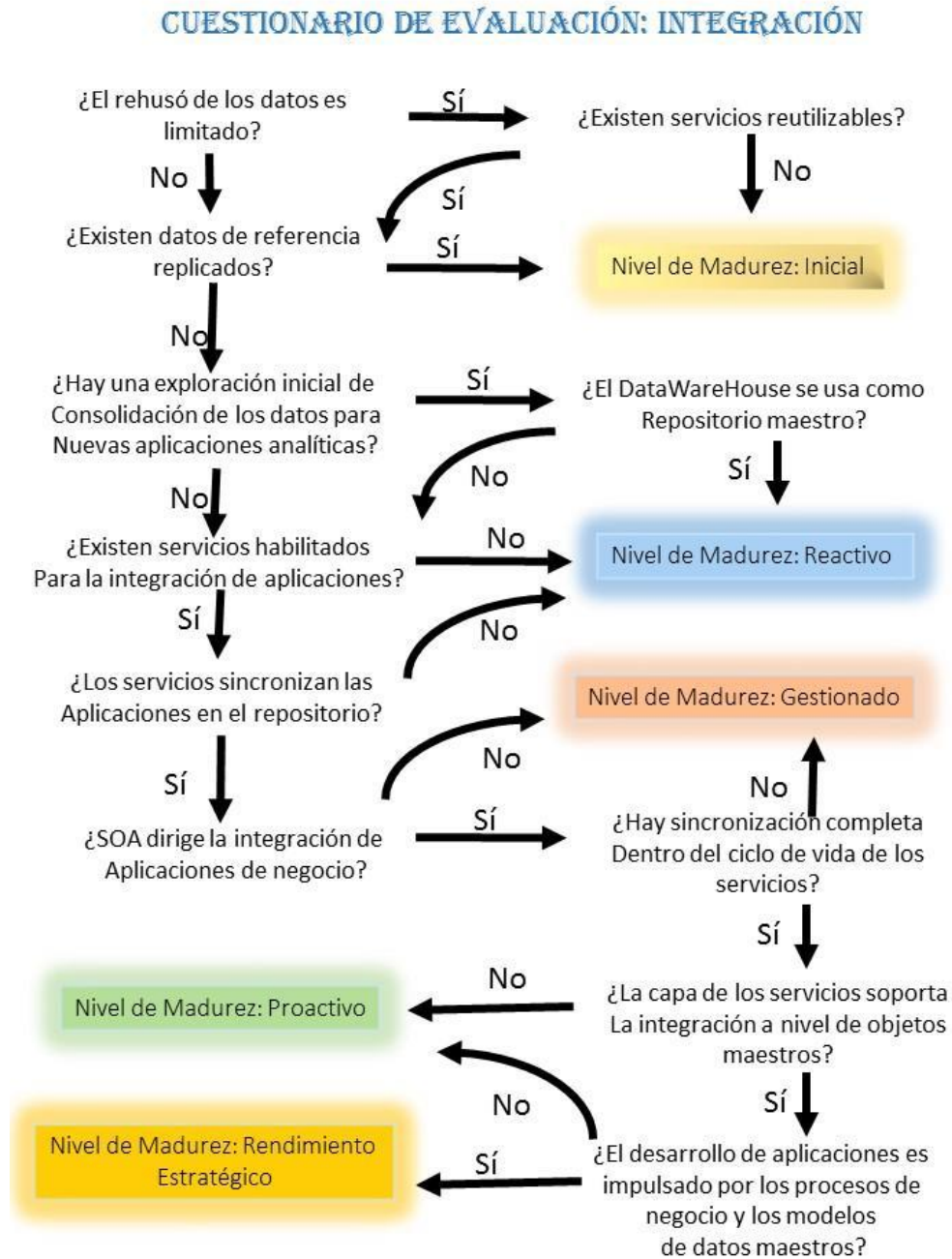


Figura 4.7 Cuestionario de Integración. Adaptado de Loshin, 2009

Cuestionario 6: Administración de procesos de negocio

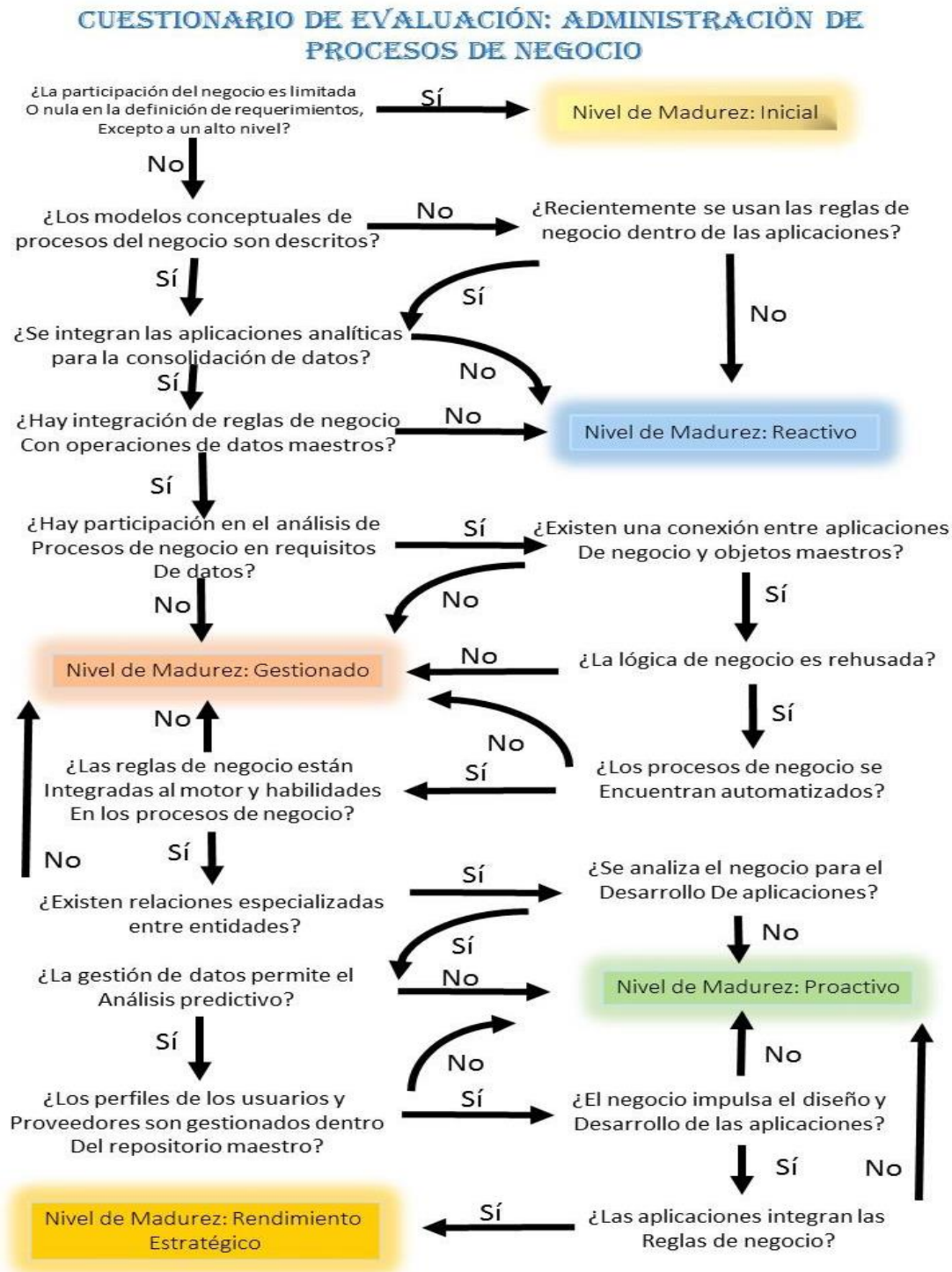


Figura 4.8 Cuestionario de Administración de Procesos de negocio. Adaptado de Loshin, 2009

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, obteniendo los siguientes resultados:

Arquitectura:

Las tres capacidades a evaluar como parte del componente Arquitectura son los modelos de datos, 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, el nivel de madurez para el componente de arquitectura es:

CUESTIONARIO DE EVALUACIÓN: ARQUITECTURA

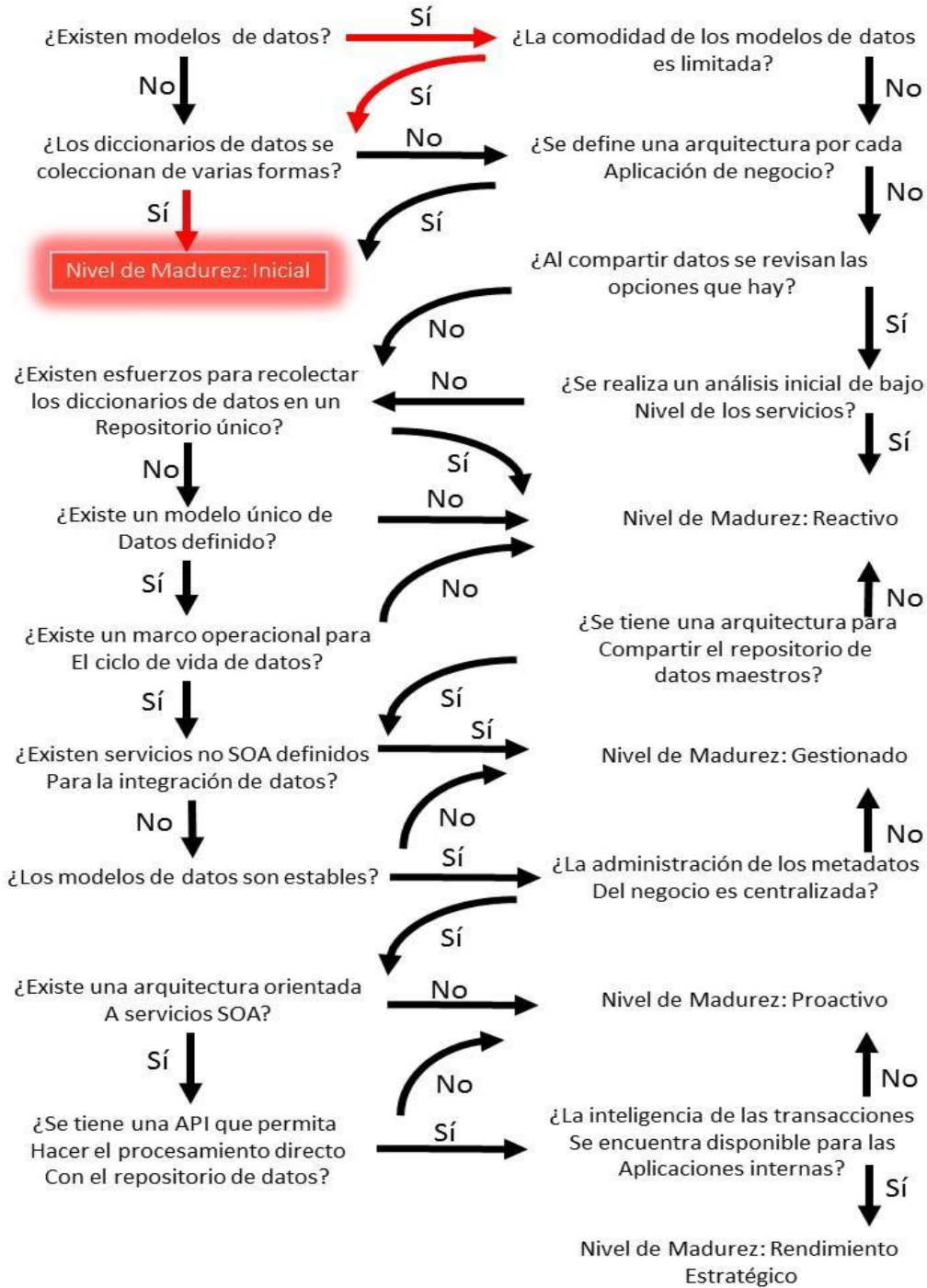


Figura 4.9 Evaluación de arquitectura

Nivel de Madurez: Inicial

- Físicamente existen modelos implementados sin una estandarización de su estructura ni contenido; muchos datos se encuentran duplicados en varias bases de datos.
- 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 (Reactivo) 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.

Gobierno:

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

Para este segundo componente, en el Área Financiera, no se tienen descripciones de usuarios y proveedores, así como para la capacidad de la calidad, no se tienen establecidos indicadores que la midan. Asimismo, tampoco existen lineamientos definidos para la gobernanza de usuarios y proveedores, evidentemente tampoco existen métodos aún definidos para garantizar su cumplimiento. Con base en esto el nivel de madurez que se ha determinado para el componente de gobierno es:

CUESTIONARIO DE EVALUACIÓN: GOBIERNO

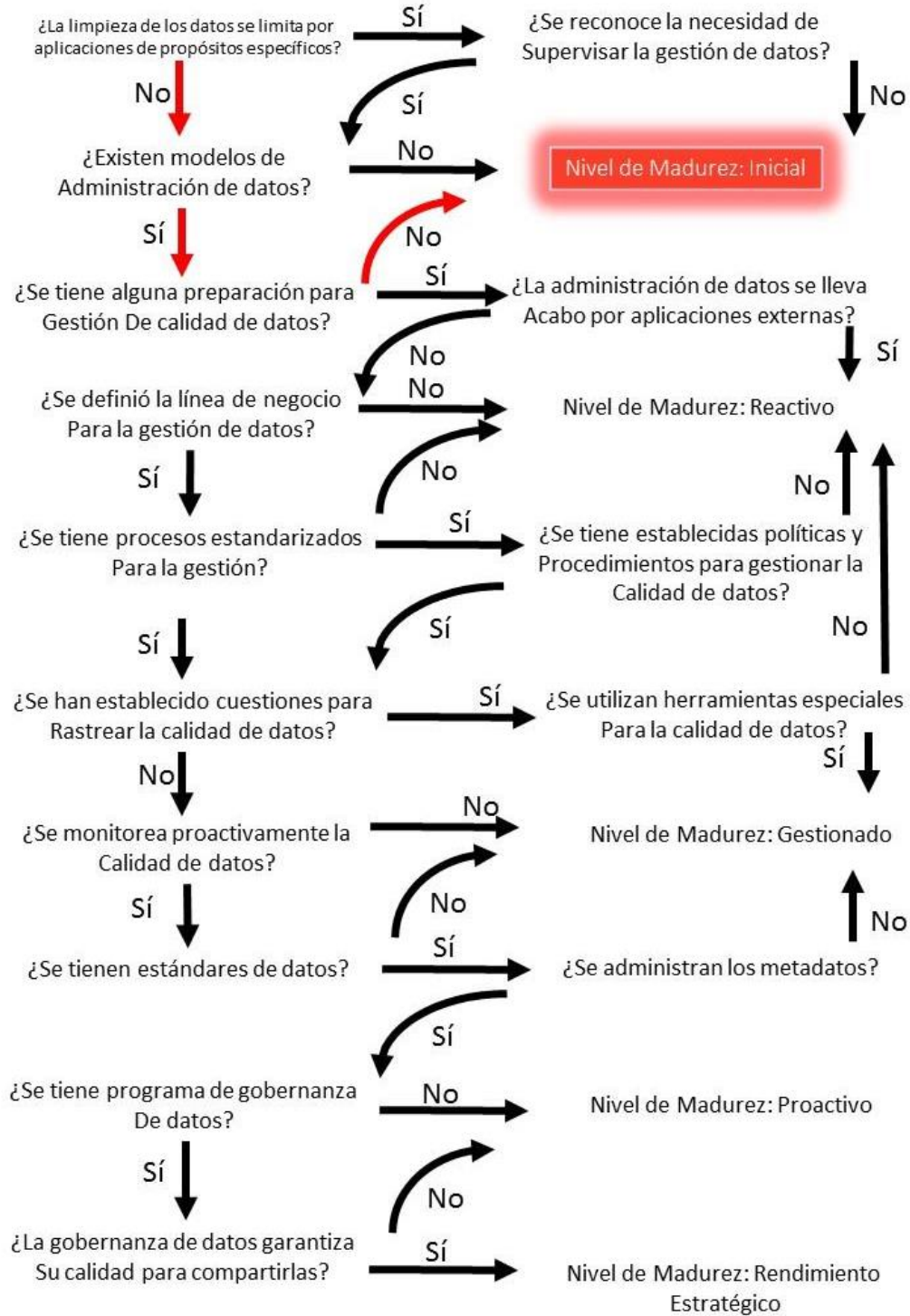


Figura 4.10 Evaluación de Gobierno

Nivel de Madurez: Inicial

- No se tiene una cultura, ni un proceso de realizar la limpieza de datos en un determinado periodo.

En este componente no se considera el nivel de madurez dos (Reactivo), debido a que no se administra calidad de datos para el análisis, estandarización y consolidación.

Administración:

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

Para el área financiera estudiada, 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. Con base en esto el nivel de madurez que se ha determinado para el componente de administración es:

CUESTIONARIO DE EVALUACIÓN: ADMINISTRACIÓN

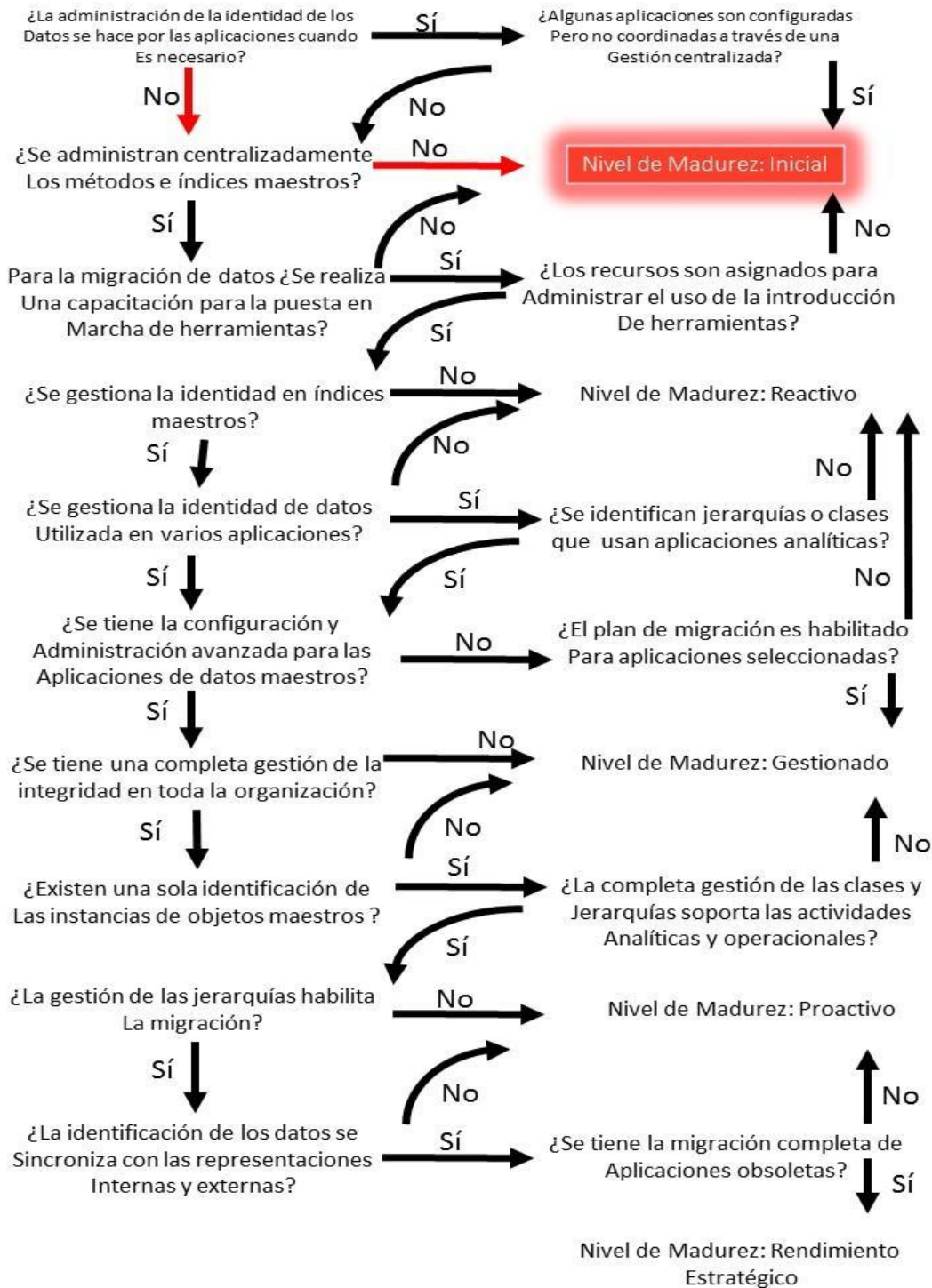


Figura 4.11 Evaluación de Administración

Nivel de Madurez: Inicial

- 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

Identificación:

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

En el área financiera de estudio el ingresar un nuevo dato, se realiza mediante un archivo Excel, lo cual no garantiza la duplicación o la mala inserción de un proveedor o usuario.

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

CUESTIONARIO DE EVALUACIÓN: IDENTIFICACIÓN

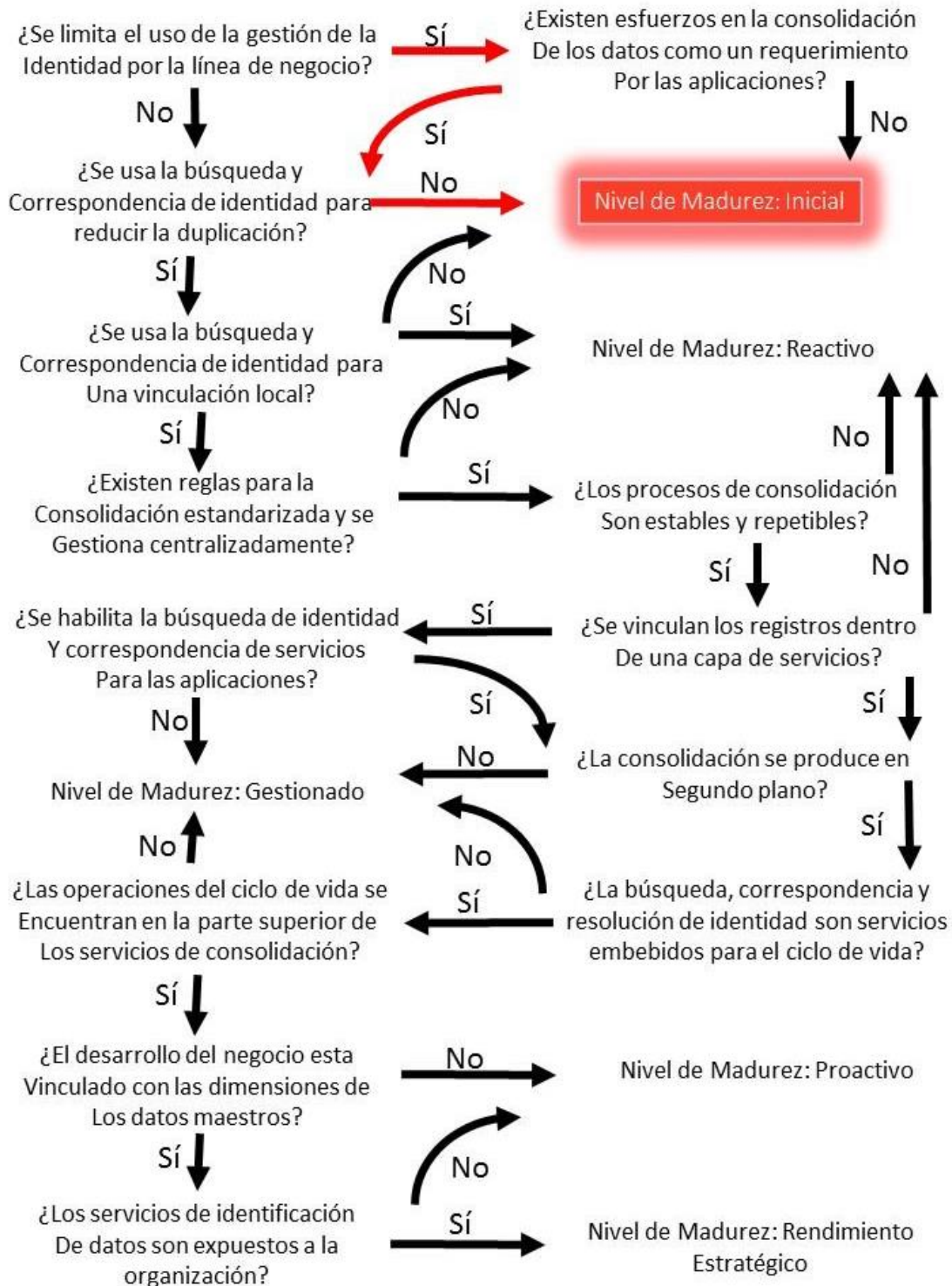


Figura 4.12 Evaluación de Identificación

Nivel de Madurez: Inicial

No se considera el nivel dos (Reactivo) porque no se realiza la búsqueda y correspondencia entre los registros para evitar la duplicación de ellos.

Integración

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

Para el área financiera de estudio, no se ingresa fácilmente al repositorio los datos. Con base en esto el nivel de madurez que se ha determinado para el componente de integración es:

CUESTIONARIO DE EVALUACIÓN: INTEGRACIÓN

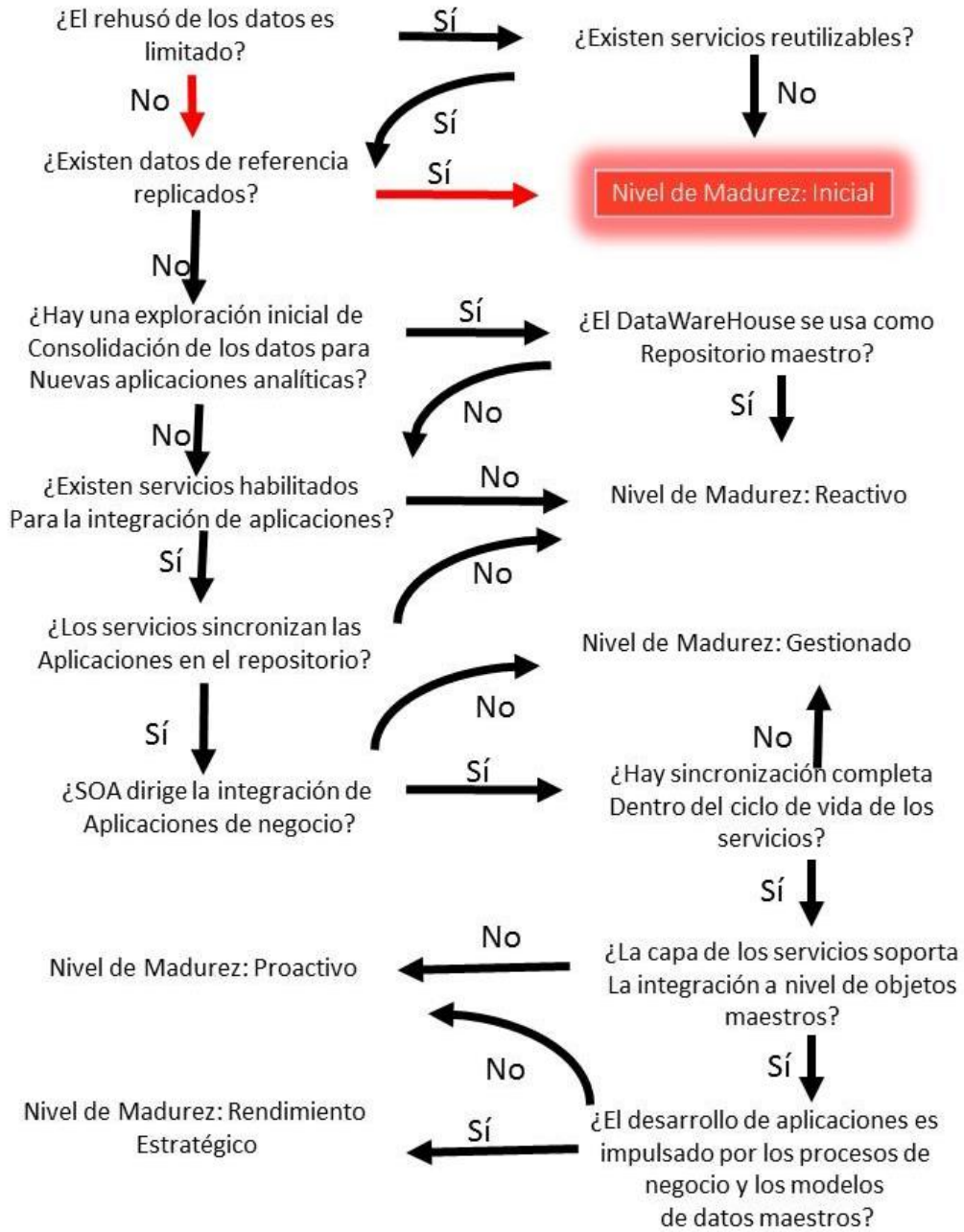


Figura 4.13 Evaluación de Integración

Nivel de Madurez: Inicial

Los catálogos de proveedores y usuarios en cada proceso operativo del área financiera. No se encuentran unificados ni relacionados entre sí.

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, debido a esto no se replican las actualizaciones mediante catálogos de referencia cruzada.

Administración de proceso 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 cuenta con una estandarización de las reglas de negocio utilizadas para sus datos transaccionales. Con base en esto el nivel de madurez que se ha determinado para el componente de administración de proceso de negocio es:

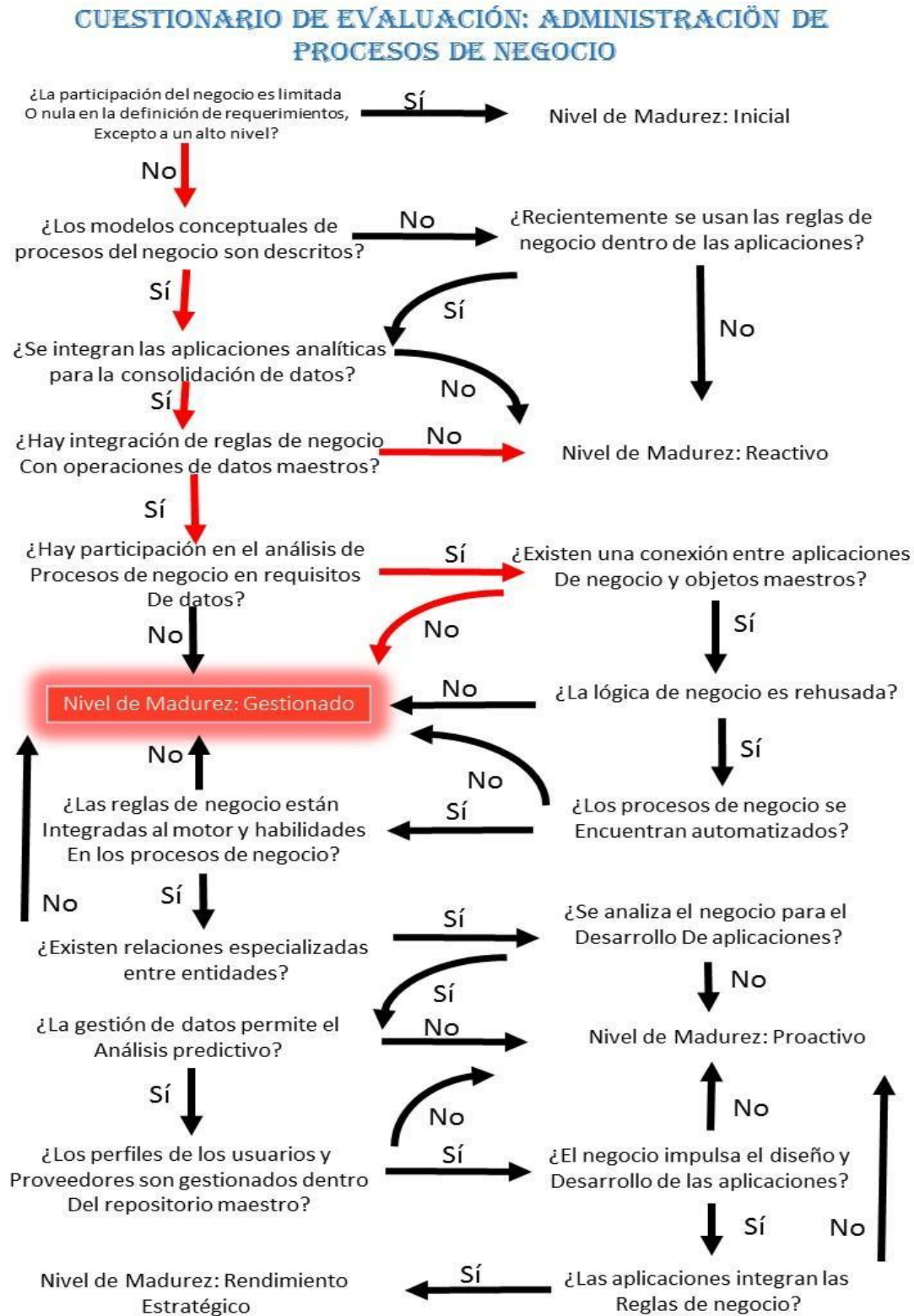


Figura 4.14 Evaluación de Administración de Proceso de Negocio

Nivel de Madurez: Gestionado

Se considera el nivel de madurez tres ya que si se tiene considerado el uso de las reglas de negocio dentro de aplicaciones, así como la participación del negocio no es limitado para los requerimientos.

Evaluación

A continuación, se muestra en la Tabla 4.1 el resumen de la evaluación de madurez realizada por cada componente para usuarios y proveedores de un área financiera.

Componentes del modelo de madurez	Nivel 1: Inicial	Nivel 2: Reactivo	Nivel 3: Gestionado	Nivel 4: Proactivo	Nivel 5: Rendimiento estratégico
Arquitectura	X				
Gobierno	X				
Administración	X				
Identificación	X				
Integración	X				
Administración de procesos de negocio			X		

Tabla 4.1 Resumen de Evaluación de Madurez

Arquitectura: No hay manera sistemática y exhaustiva para asegurar cambios en los datos maestros.

Gobierno: Los procesos no están regulados o documentados durante esta etapa. No hay conciencia de la calidad de la información, por lo tanto, no se hacen intentos para evaluar o mejorar la calidad de la información.

Administración: La organización actúa en respuesta sólo cuando se producen problemas en la calidad de la información

Identificación: Al ingresar un nuevo dato, se realiza por medio de un archivo Excel el cual no garantiza la duplicación o la mala inserción de un proveedor o usuario.

Integración: No se cuenta con un repositorio de datos unificado.

Administración de procesos de negocio: Se cuenta con una mayor participación de las áreas de negocio involucradas en el desarrollo de los requerimientos de las aplicaciones, pero aún la capacidad de análisis está siendo desarrollada.

4.1.1.2 FASE A: VISIÓN DE LA ARQUITECTURA

En esta fase se desarrolla el alcance, la misión, la visión y las expectativas que se tendrá para esta primera iteración.

Misión: Contar una base fiable, segura y fidedigna de los datos críticos para el negocio.

Visión: Administrar los datos del área financiera, de forma que siempre se tenga una confiabilidad y medición de calidad de estos.

Para esta iteración se busca llegar a:

- ✚ Contar con una base con datos sin datos repetidos.
- ✚ Desarrollar un proceso para el aseguramiento de la calidad de los datos.

4.1.2 ITERACIÓN DE DEFINICIÓN

4.1.2.1 FASE B: ARQUITECTURA DE NEGOCIO

Esta arquitectura describe las directrices requeridas para operar y lograr los objetivos de negocio. Sus principales elementos son procesos, funciones, servicios, roles, principios y políticas.

Procesos

La clave para una exitosa transformación de negocios es lograr que todas las personas y grupos de interés comprendan sus roles y los problemas que posiblemente surjan. Uno de los mayores desafíos para la práctica transformación de los negocios es incorporar soluciones nuevas sin interrumpir ni eliminar los componentes de negocio que se encuentran en funcionamiento.

Cuando los datos son recopilados en grandes cantidades, suelen presentarse en una forma incompleta, inconsistente y ruidosa, por lo que tener una limpieza de estos es fundamental para tener un mejor resultado, obteniendo beneficios tales como:

- ✚ Se asegura la calidad de los datos que vamos a procesar.
- ✚ Se evita la información no veraz o errónea.
- ✚ Se ahorra costes de espacio en disco al eliminarse la información duplicada.
- ✚ Se agiliza las consultas por la ausencia de datos repetidos o inservibles.
- ✚ Se ayuda a tomar decisiones estratégicas correctas.

El objetivo de este proceso es medir y controlar la calidad de los datos con los que se cuentan en la institución. Este proceso debe de formar parte de una evaluación periódica, para lo cual se definen como principales entradas a 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.
- ✚ Respaldo histórico para una futura recuperación.
- ✚ Replicar actualización.
- ✚ Notificar cambio realizado.

Se debe seguir el siguiente flujo de proceso asegurar la calidad de los datos de la base de la institución.

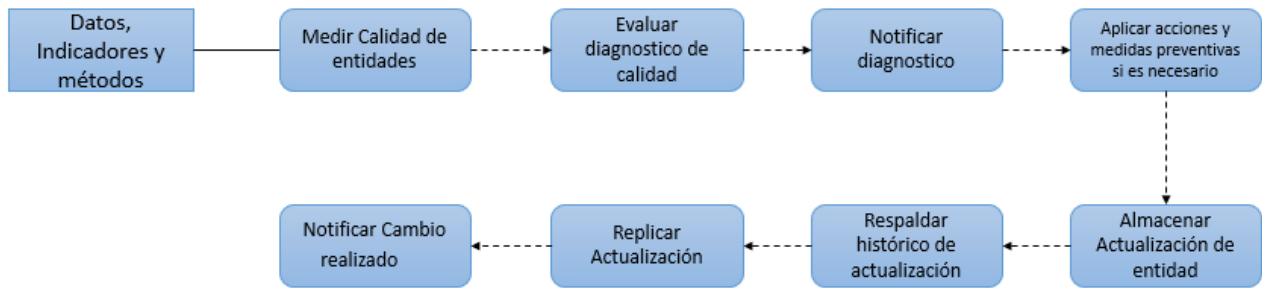


Figura 4.15 Proceso para controlar calidad de datos

Funciones

Las funciones de negocio entregan capacidades, por lo tanto la función principal para esta arquitectura es la de depuración y aumento de la calidad de los datos.

Las funciones para el proceso de calidad de las entidades de datos son las mostradas en la Figura 4.16, las cuales 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:
Aprobar los cambios realizados en la depuración de datos.
- ✚ Integración de datos:
Incorporar los datos nuevos a la base de datos ya existente.
- ✚ Seguridad de datos:
Mantener un respaldo de la información depurada

✚ Sincronización de datos:

Tener una sincronización entre los datos depurados y su incorporación a la base de datos

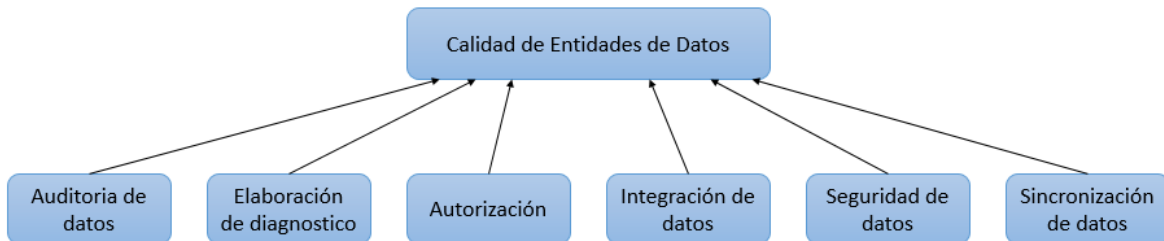


Figura 4.16 Funciones para el proceso: Calidad de entidades

Servicios

Los servicios soportan las capacidades que generan las funciones a través de una interfaz.

Estos servicios son mostrados en la Figura 4.17 y Tabla 4.2, y descritos el servicio de negocio contra la función de negocio en la siguiente tabla

Función de Negocio	Servicio de Negocio
Auditoria de datos	Servicio de medición de calidad a través de indicadores y formulas
Elaboración de diagnostico	Servicio de despliegue de información de diagnóstico
Autorización	Servicio de aprobación o rechazo de cambio
Integración de datos	Servicio de información integrada
Seguridad de datos	Servicio de respaldo de información
Sincronización de datos	Servicio de información sincronizada
	Servicio de proceso finalizado

Tabla 4.2 Servicios de Negocio

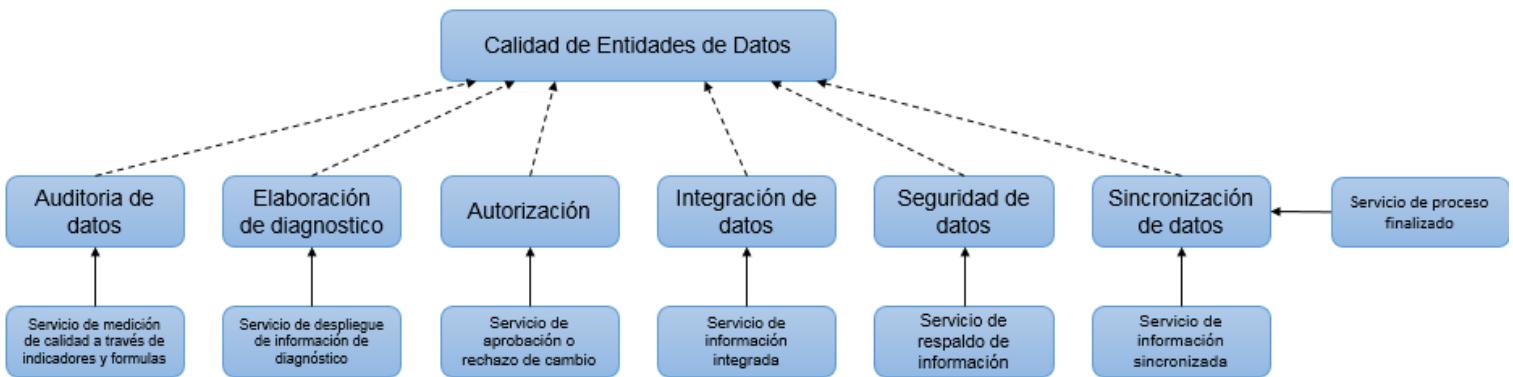


Figura 4.17 Servicios para el proceso de control de calidad de datos

Perfilado

Aunque todavía se considera como una etapa independiente del proceso de limpieza, es altamente recomendable realizar previamente un perfilado de datos donde se decide, a través de un muestreo, qué cambios realizar y de qué forma concreta hacerlos. De esta manera, garantizamos una limpieza posterior óptima y completamente estandarizada.

Conviene remarcar que en el perfilado aún no se realizan modificaciones, puesto que se trata de un análisis previo.

Roles

Esta solución contempla roles para llevar a cabo las actividades de los procesos descritos. El DAMA¹ define tres agrupaciones de roles: responsables, de aprobación y de contribución.

✚ **Roles responsables.** Son aquellas personas expertas que garantizan el control y uso eficaz de los datos. Aprueban, validan y realizan los cambios necesarios para las entidades de datos que contengan valores de oro y se compartan adecuadamente. Los roles de esta agrupación son:

- o Administrador del negocio
- o Custodio de datos
- o Administrador de datos.

¹ **DAMA:** (Data Management Association International) es una asociación internacional sin fines de lucro dedicada a promover conceptos y buenas prácticas en la gestión de recursos de información (IRM “Information Resource Management”) y gestión de recursos de datos (DRM “Data Resource Management”)

- ✚ **Roles de contribución.** Son los participantes que suministran información para la actualización del contenido. Los roles de esta agrupación son:
 - o Usuario de datos
 - o Proveedor de datos
 - o Interesado.

- ✚ **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.
 - o Arquitecto de datos
 - o Arquitecto del negocio
 - o Responsable del negocio
 - o Consumidor

4.1.2.2 FASE C: ARQUITECTURA DE SISTEMAS DE INFORMACIÓN

Esta arquitectura soporta la arquitectura de negocio y 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 4.18.

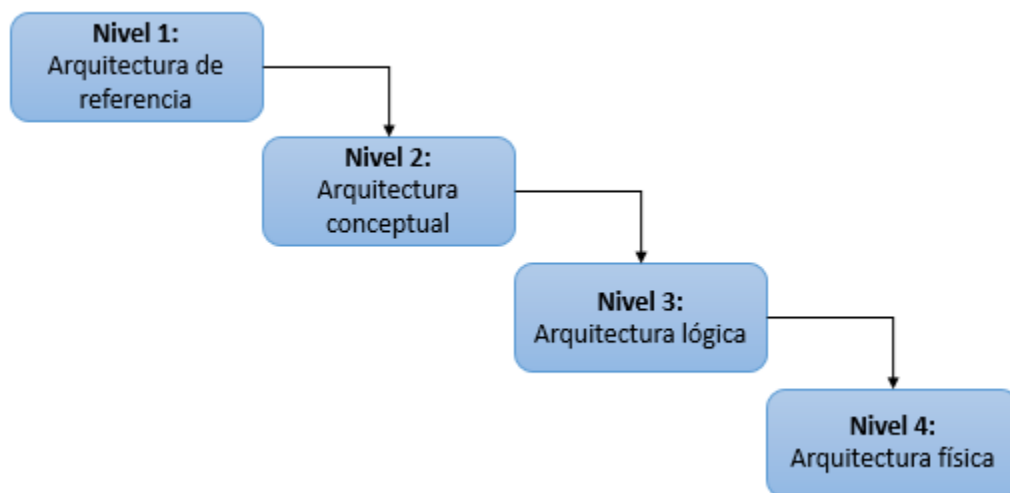


Figura 4.18. Niveles de abstracción de arquitectura de datos

Arquitectura de referencia

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.

Los elementos clave para una Arquitectura de Referencia son:

- ✚ Propósito estratégico:
Explica el contexto, alcance, metas, propósito e intención de uso.
- ✚ Principios:
Son declaraciones de alto nivel, generalmente reglas o lineamientos que dirigen como una organización cumple con sus objetivos
- ✚ Posiciones Técnicas:
Son los requerimientos establecidos para lanzar ingenierías uniformes o criterios técnicos, métodos, procesos y prácticas

Para el caso de estudio, en el Área Financiera se observa:

1. Propósito estratégico:

Para esta iteración se tiene como propósito estratégico de la arquitectura:

- 1) Almacenar centralizadamente los datos de los proveedores en un repositorio único.
- 2) Tener disponibles los registros únicos y más actualizados de los proveedores.
- 3) Tener control de los cambios de estos contactos.

2. Principios.

Los principios para esta iteración son:

- 1) Aplicar reglas de unificación de datos.
- 2) Validaciones de completitud
- 3) Estandarización de datos.

3. Posiciones Técnicas.

Para la posición técnica se debe de analizar varios puntos para que se cumplan con los requerimientos técnicos, necesarios, tales como:

- 1) Cantidad de datos y tipo de datos (texto, binarios, espacial u otros tipos específicos)
- 2) Número simultáneo de usuarios (conurrencia a la base de datos)
- 3) Disponibilidad.
- 4) Escalabilidad.
- 5) Seguridad: cuanto necesitará de características como seguridad e encriptación de datos, administración de usuarios, privilegios.
- 6) Manejo y administración: ¿Qué tan amigable quiere que sea la administración de su base de datos?

Arquitectura Conceptual

La arquitectura conceptual está integrada por las reglas de negocio dado una situación, así como un diagrama el cual representa a estas reglas.

Dentro de la institución financiera, los proveedores se clasifican en primarios, secundarios y operativos, esto en función del tipo de insumos que suministran.

Los proveedores primarios son todas las entidades financieras externas de los cuales se les puede requerir información. Por lo tanto, estos proveedores se pueden identificar por medio de su entidad financiera.

Los proveedores secundarios son todas las entidades o autoridades (externas o internas) que, por acuerdo o contrato, comparten información. Estos proveedores se representan por la entidad financiera y si son internos por la unidad administrativa.

Los proveedores operativos se refieren a otros procesos que conforme a sus responsabilidades y un requerimiento de información suministra información necesaria. Estos proveedores se identificarán por los procesos operativos.

El alcance de este modelo abarca el diseño para almacenar los proveedores primarios y los secundarios. Para ello, en la Figura 4.19 se muestra el escenario respecto a los proveedores a modelar.

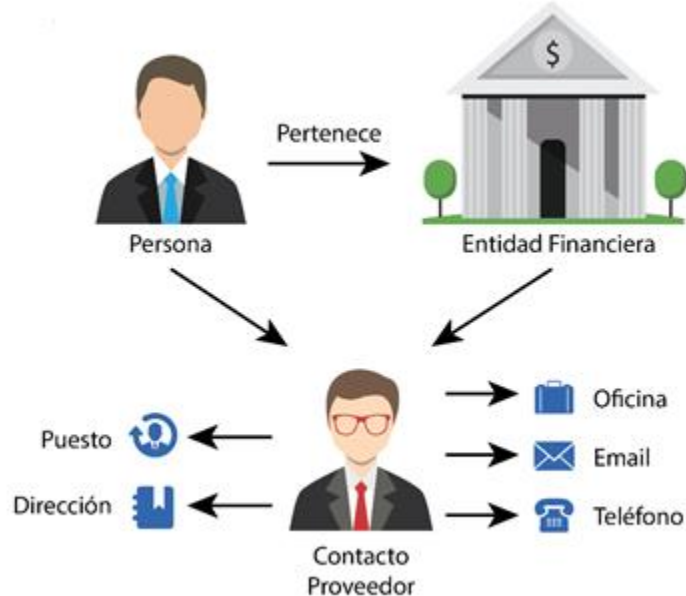


Figura 4.19 Escenario de los proveedores

En el diagrama se puede observar que una persona que pertenece a una entidad financiera, representa a un contacto. Los datos de identificación asociados a un contacto proveedor son: entidad financiera, oficina o unidad administrativa, correo electrónico, teléfono, puesto y dirección de trabajo.

Derivado de esto, se comenzará a enlistar los requerimientos funcionales, los cuales son aquellos que describen la funcionalidad o los servicios que el usuario espera obtener. Los usuarios de esta información son todos los responsables del negocio que necesitan mantener comunicación constante con las instituciones financieras durante el acopio de la información.

Como se mencionó previamente, los propósitos estratégicos de esta primera iteración para la institución son:

- ✚ Almacenar centralizadamente los datos de los proveedores en un repositorio único.
- ✚ Tener disponibles los registros únicos y más actualizados de los proveedores.
- ✚ Tener control de los cambios de estos contactos.

Los requerimientos no funcionales son aquellos que describen características más específicas orientadas al funcionamiento del desarrollo de un sistema. Por ello, los requerimientos no funcionales del escenario de los proveedores son:

- ✚ Normalización del modelo y restricciones de integridad con el fin de que no se repitan contactos duplicados
- ✚ Mecanismos automatizados que controlen las inserciones de nuevos datos y cambios sobre datos ya almacenados.

Teniendo el siguiente modelo conceptual:

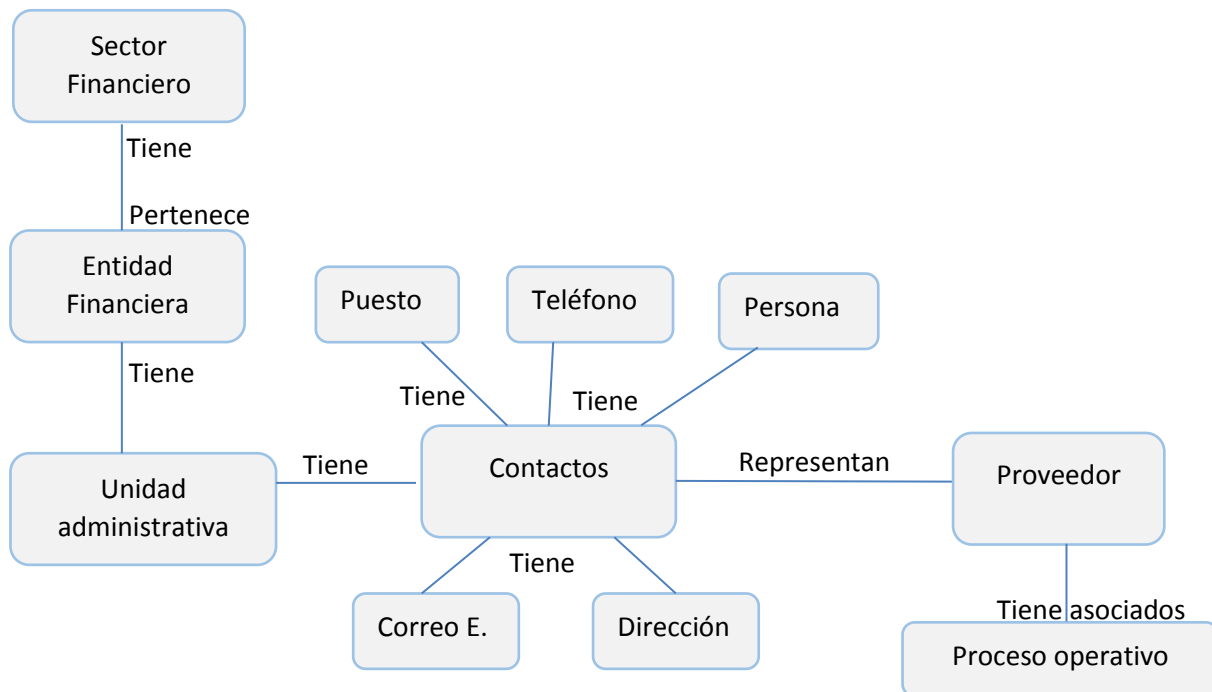


Figura 4.20 Diagrama conceptual de la base de datos

Arquitectura Lógica

La definición de arquitectura lógica está orientada a las consultas y operaciones que puede soportar a través de la definición de sus relaciones y condiciones lógicas. Por lo tanto, posterior al modelado conceptual, en esta sección se describe el modelo lógico usado en la base de datos.

Para ello, con base en el diseño conceptual el primer paso es poder analizar qué información se guardaría en cada entidad de datos para definir sus respectivos atributos y relaciones de cardinalidad entre estas entidades.

Entidad de datos	Descripción
Sector Financiero	Entidad que contiene la clasificación por sector del conjunto de instituciones que componen el sistema financiero. Ejemplo de estos sectores pueden ser, instituciones de banca múltiple, casas de bolsa, arrendadoras financieras, etc.
Entidad Financiera	Contiene los valores de las instituciones financieras que componen el sistema financiero.
Unidad Administrativa	Formadas por los valores de las diferentes unidades administrativas que pertenecen a las entidades financieras.
Contacto	Es la entidad de datos que está representada por una persona y sus correspondientes datos de identificación como puesto, persona, correo, teléfono, unidad administrativa y entidad financiera.
Persona	Contiene los datos de los actores que participan suministrando información. Esta entidad concentra los diferentes nombres y apellidos de las personas que son contactos.
Teléfono	Es la entidad de datos que contiene la información de los números telefónicos que puede tener el contacto. Este catálogo tiene como atributos el número de teléfono y su extensión en caso de aplicar.
Correo Electrónico	Es la entidad de datos que contiene la información de los correos electrónicos que puede tener el contacto. Esta entidad tiene como atributos el identificador del contacto externo al que pertenece y el correo electrónico en sí.
Puesto	Es el catalogo que incluye los valores de los puestos de trabajo para los diferentes contactos. Principalmente este valor depende de la institución financiera y la unidad administrativa a la que sea representada por una persona.

Tabla 4.3 Entidad de datos

Entidad de datos	Descripción
Dirección	Contiene los datos de la dirección de trabajo de los contactos. Este valor depende de la entidad financiera a la que represente una persona, así como de la sucursal en la que labore.
Proveedor	Es la entidad de datos que establece la relación de proveedor entre los contactos de las entidades financieras y el proceso operativo al cual le suministran insumos.
Proceso Operativo	Es el catálogo que contiene el listado de los procesos de información de la institución.

Tabla 4.3 Entidad de datos

Para las relaciones que existen se tiene lo siguiente:

Relación	Condición	Cardinalidad
A	Un proceso operativo puede tener uno o varios proveedores	Proceso Operativo-> Proveedor 1: N
B	Un contacto puede ser proveedor de uno o varios procesos operativos	Contacto-> Proveedor 1:N
C	Una persona puede representar a más de un contacto por medio de diferentes entidades financieras	Persona- >Contacto 1:N
D	Una dirección tiene asociado un asentamiento que representa el código postal. Un asentamiento se relaciona con una o varias direcciones	Asentamiento -> Dirección 1:N
E	Una dirección puede estar asociada a más de un contacto.	Dirección -> Contacto 1:N
F	Un contacto puede tener asociados uno o varios correos electrónicos.	Contacto -> Correo 1:N
G	Un contacto puede tener asociados uno o varios teléfonos	Contacto -> Teléfono 1:N
H	Un puesto puede estar asociado a diferentes contactos	Puesto -> Contacto 1:N

Relación	Condición	Cardinalidad
I	Un contacto puede tener asociados diferentes puestos reportados por diferentes procesos operativos.	Contacto -> ContactoPuesto 1: N
J	Un contacto puede tener asociada más de una dirección.	Contacto -> ContactoDireccion 1:N

Relación	Condición	Cardinalidad
K	Un contacto puede tener asociadas diferentes unidades administrativas.	Contacto -> ContactoUnidad 1:N
L	Una unidad administrativa puede estar asociada a más de un contacto	UnidadAdministrativa -> ContactoUnidad 1: N
M	Una entidad financiera puede estar asociada a más de un contacto, pero un contacto solo puede estar asociado a una entidad financiera	Entidad Unidad-> Contacto 1:N
N	Una entidad financiera pertenece a un sector financiero. Un sector financiero tiene asociadas más de una entidad financiera	SectorFinanciero-> Entidad Financiera 1:N

Tabla 4.4 Relaciones de Tablas

En el siguiente modelo relacional se puede observar a las entidades anteriormente descritas, así como las relaciones entre ellas.

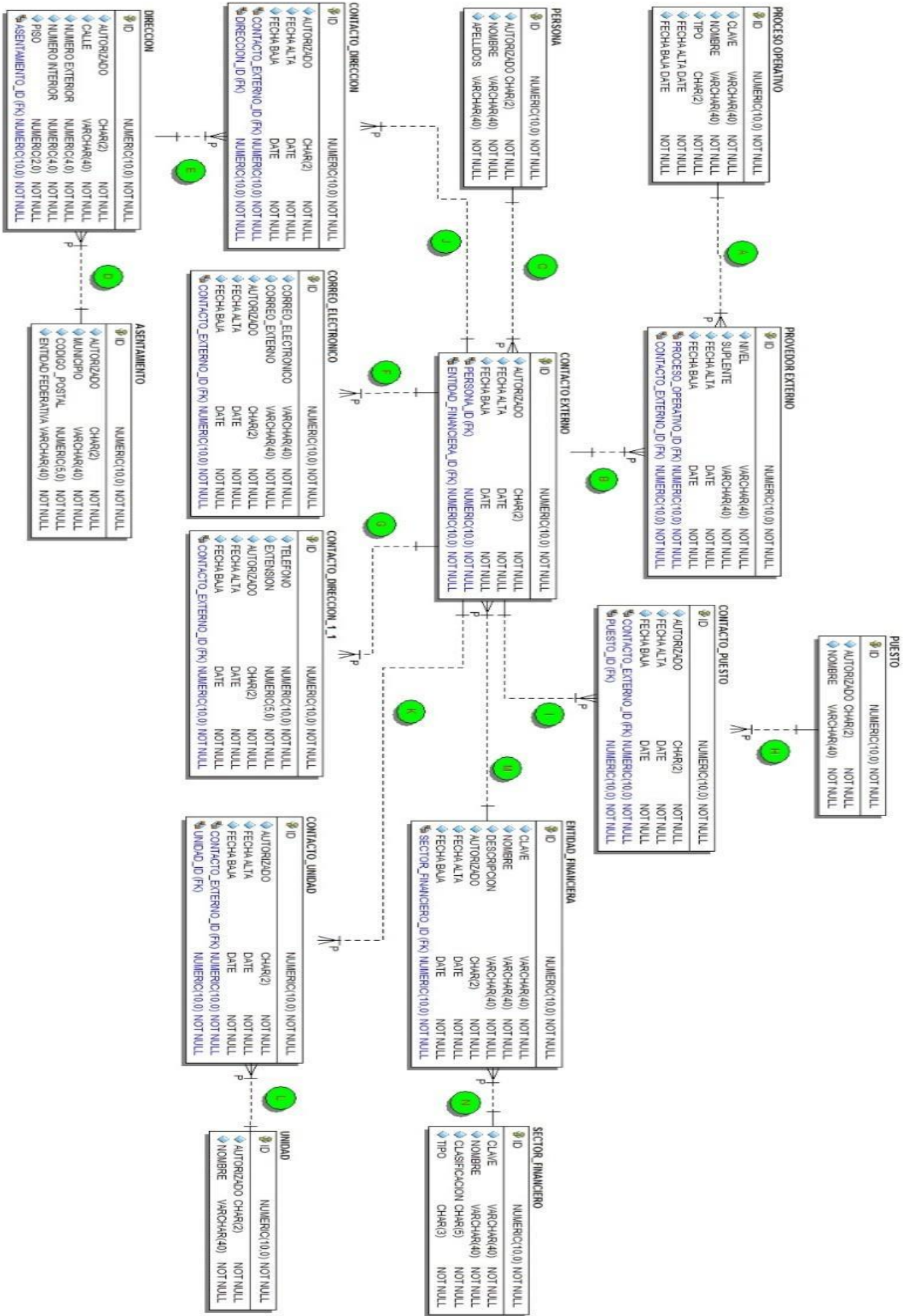


Figura 4.21 Modelo Lógico de la base de datos

Arquitectura de Física

Un modelo de datos físico describe las estructuras de datos implementadas en un manejador de base de datos, también hace referencia al tipo de datos de los atributos de las tablas, así como las restricciones de integridad que corresponden a las condiciones que los datos deben cumplir.

Los responsables del contenido de la base de datos son los responsables de negocio de los procesos operativos. Sin embargo, para cargar nuevos registros en el modelo físico se necesitarán dos perfiles, un arquitecto de datos y un administrador de base de datos. Ellos se encargan de diseñar y configurar todo lo que requiera este esquema propuesto para ejecutar programas informáticos que respondan a la operación del modelo de datos.

4.1.2.3 FASE D: ARQUITECTURA DE TECNOLOGÍA

Si bien un buen programa de calidad de datos es el resultado de una apropiada administración de personas y procesos, las herramientas tecnológicas tienen un papel importante. Muchas empresas realizan esta tarea de limpieza de datos con programas en SQL o herramientas limitadas incluidas en productos de ETL. El mercado de herramientas de calidad de datos es aún pequeño, pero se encuentra en expansión.

La funcionalidad esperable de las herramientas de calidad de datos consiste de:

- **Perfilado de datos:** Consiste en examinar los datos que existen en las fuentes de origen de una organización y recopilar estadísticas e información sobre los mismos con el objetivo de reducir el riesgo al integrar nuevos aplicativos y conseguir métricas de calidad de datos (Curto, 2008)
- **Estandarización o normalización:** Consiste en descartar la repetición de grupos y minimizar la redundancia para optimizar consultas y aumentar la confiabilidad de la información.
- **Verificación:** Consiste en la posibilidad de comparar datos ingresados que van a ser llevados al repositorio central con un dominio específico
- **Matching:** Consiste en la consolidación de registros dentro de un repositorio MDM, buscando registros idénticos (el mismo objeto en diferentes sistemas) y duplicados (el mismo objeto en el mismo sistema).

- **Monitorización:** implementación de controles para asegurar que los datos siguen cumpliendo las reglas de negocio que definen la calidad de datos de la organización.
- **Usabilidad:** provisión de toda la funcionalidad anterior a través de una interfaz orientada a negocio, disponible para usuarios técnicos y de negocio, de una manera que sea apropiada para su rol.

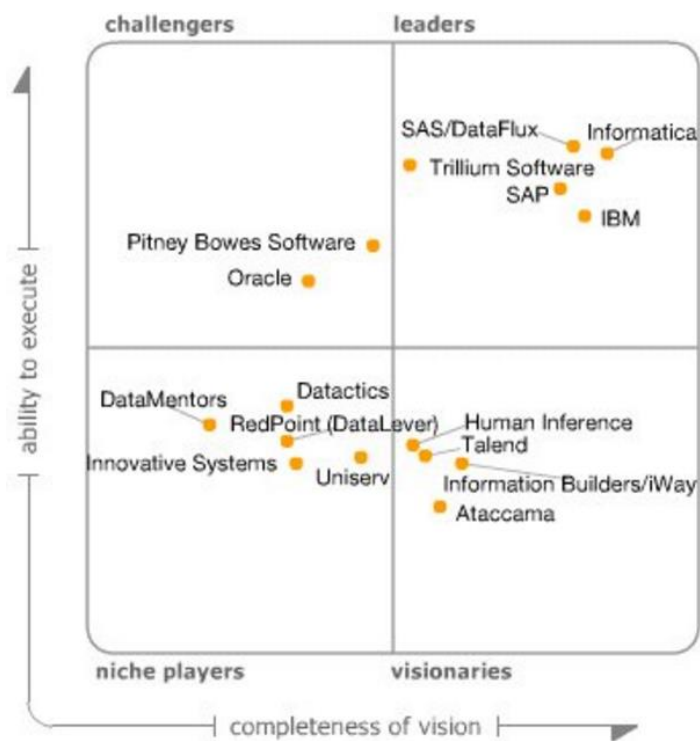


Figura 4.22 Cuadrante Gartner para herramientas de calidad de datos 2012

Gartner pone de relieve en este estudio que la demanda e implementaciones de este tipo de herramientas permanece sólido en el mercado, tal como es mostrado en la Figura 4.22 donde es mostrada una evaluación de las herramientas comerciales existentes en el 2012 para la gestión de la calidad de datos, evaluándolas sobre dos ejes: Habilidad para hacer (Eje vertical) e Integridad de la visión (Eje Horizontal). Por su

parte, los vendedores están expandiendo las funcionalidades de sus soluciones para apoyar roles que no son TIC y diversificando sus modelos de despliegues.

Para este caso de estudio tomaremos una pequeña muestra de la base de datos, en la cual tenemos como datos, los mostrados en la Tabla 4.5.

Nombre	Apellido Paterno	Apellido Materno	Correo
Dolores	Gaspar	Gonzalez	gaspar@hotmail.com
Jose Manuel	Gomez	Saavedra	gomez@hotmail.com
Francisco	Delgado	Zarza	delgado@hotmail.com
Maria Teresa	Perez	Blazquez	perez@hotmail.com
Jesus A.	Costa	Hernandez	costa@hotmail.com
Marta	Jaramillo	Saiz	jaramillo@hotmail.com
Alejandro	Ibanez	Rivas	ibanez@hotmail.com
Alvaro	Padron	Beltran	padron@hotmail.com
Jose Manuel	Asensio	Vega	asensio@hotmail.com
Francisco Javier	Lopez	Reyes	lopez@hotmail.com
Andrea	Ferre	Manzano	ferre@hotmail.com
Angel	Romero	Garcia	romero@hotmail.com
Cristina	Andres	Sanz	andres@hotmail.com

Tabla 4.5 Datos de la base de datos productiva

Se ha recibido un archivo con la actualización de datos que se deben de ingresar a la base, del cual tomamos una pequeña muestra, mostrados en la Tabla 4.6.

Nombre	Apellido Paterno	Apellido Materno	Correo
Jose M.	Gomez	Saavedra	gomez@hotmail.com
Maria	Perez	Blazquez	perez@hotmail.com
Jesus Alejandro	Costa	Hernandez	costa@hotmail.com
Marta	Jaramillo	Saiz	jaramillo@hotmail.com
Alejandro	Revilla	Haro	revilla@hotmail.com
alvaro Daniel	Padron	Beltran	padron@hotmail.com
J. Manuel	Asensio	Vega	asensio@hotmail.com
Francisco J.	Lopez	Reyes	lopez@hotmail.com
Andrea	Ferre	Manzano	ferre@hotmail.com
angel	Romero	Garcia	romero@hotmail.com
Kristina	Andres	Sanz	andres@hotmail.com
Carmen	De Leon	Gonzalez	de leon@hotmail.com

Tabla 4.6 Datos nuevos a ingresar

En este punto se realizará una búsqueda de coincidencia de los datos nuevos a ingresar, respecto a los que ya se encuentran en la base de datos, de forma que evitemos el duplicar registros basura en la base, debido a la existencia de restricciones de tipo unique, no nos asegura una total verificación de dato, ya que en este tipo de restricción debe de existir una coincidencia del 100% para que aplique y en la vida diaria pocas veces se llega a presentar esto.

En este caso de estudio haremos uso de una herramienta de calidad de datos de licencia libre, en la cual a partir de la conexión con una base de datos o un archivo se puede hacer una búsqueda por coincidencias teniendo la oportunidad de limitar el porcentaje de coincidencias que busquemos, así como en los campos que deseemos realizar esta búsqueda.

Lo primero que realizaremos tal como se muestra en la figura 4.23, será cargar ambos archivos en nuestro programa, uno representa los datos cargados en la base de datos productiva y el otro que representa los datos a actualizar

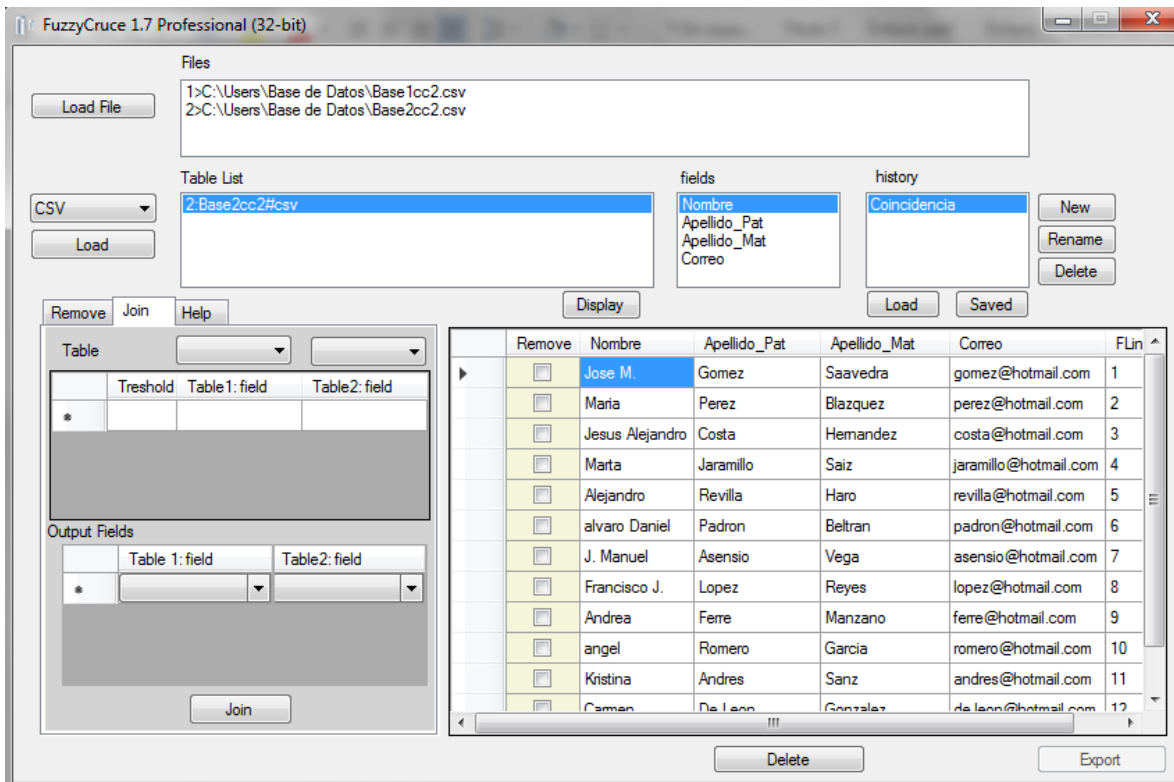


Figura 4.23 Carga de Archivos

Lo siguiente será realizar una consulta donde realicemos un match entre ambos archivos, para poder realizar un filtro por coincidencias entre los registros, tal como se muestra en la Figura 4.24



Figura 4.24 Consulta de coincidencias

Una vez realizada esta consulta tendremos como resultado el mostrado en la Figura 4.25, en el cual podemos observar que existen 3 registros con una coincidencia del 100% y otros con una coincidencia mayor al 60%

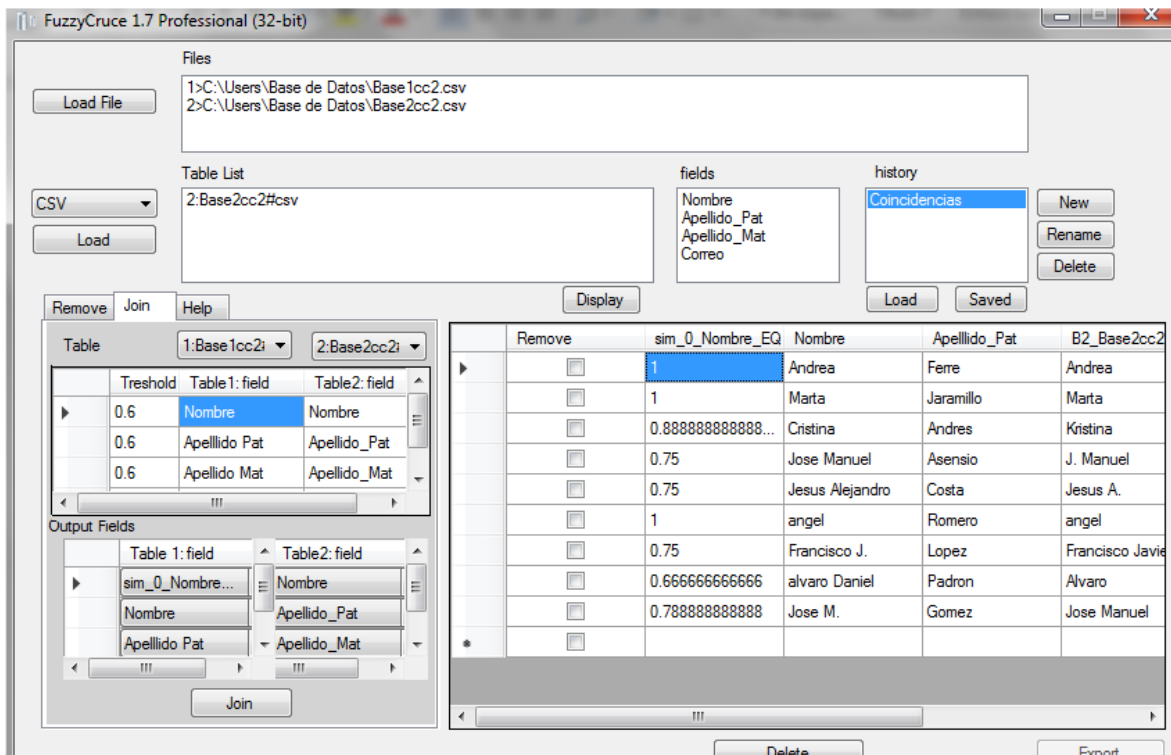


Figura 4.25 Carga de Archivos

A partir de esto podemos eliminar del archivo de actualización, los registros que ya existen en la base de datos, por lo cual del archivo de actualización que teníamos al inicio, únicamente nos quedamos los datos mostrados en la Tabla 4.7.

Nombre	Apellido Paterno	Apellido Materno	Correo
Maria	Perez	Blazquez	perez@hotmail.com
Alejandro	Revilla	Haro	revilla@hotmail.com
Carmen	De Leon	Gonzalez	de leon@hotmail.com

Tabla 4.7 Datos finales

4.1.3 ITERACIÓN DE PLANEACIÓN

4.1.3.1 FASE E: OPORTUNIDADES Y SOLUCIONES

Con base en la metodología de TOGAF la implementación de la arquitectura para el aseguramiento de la calidad de las entidades de datos requiere de una planeación estratégica.

- ✚ Los datos no se deberán duplicar.
- ✚ El intercambio de datos se deberá realizar mediante la implementación de las aplicaciones definidas en la arquitectura de aplicaciones.

La estrategia de migración conduce la implementación para obtener como resultado a la arquitectura objetivo.

Estrategia de migración e implementación

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 Maturity Model y sus capacidades mediante programas y proyectos.

En la siguiente tabla son mostrados los proyectos que se deben de implementar para cubrir cada uno de los componentes del modelo de madurez MDM.

Detección y análisis de datos	
Proyectos	Descripción
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.
Recolección de datos	Este proyecto contempla las actividades para ubicar, extraer a los datos, definir un artefacto para su recolección y validación.
Consolidación del contenido de datos	Este proyecto contempla las actividades para realizar un análisis de los datos recopilados a través de métodos de limpieza de datos.

Tabla 4.8 Proyectos de implementación de limpieza de datos

4.1.3.2 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 del proyecto. Su representación jerárquica permite una identificación más fácil de los paquetes de trabajo que se tienen que generar.

En la Figura 4.26 se muestra la Estructura de Descomposición de Trabajo para el presente trabajo.

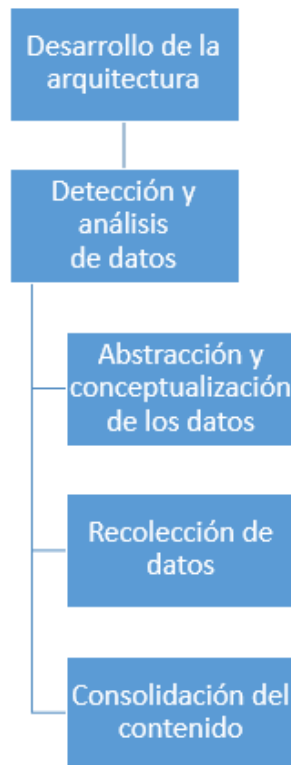


Figura 4.26 Estructura de desglose de trabajo

4.1.4 CONCLUSIÓN DE PRIMERA ITERACIÓN

Como hemos visto, la calidad del dato es esencial tanto para minimizar el riesgo como para informar de las mejores decisiones de negocio con una mayor confiabilidad. Pero gestionar el dato como un activo de negocio requiere de una administración acorde con la política de la empresa, incluyendo las necesidades de los distintos departamentos.

El aseguramiento de la calidad de datos con los que trabaja un sistema, es esencial para el correcto funcionamiento de este, pero para lograrlo debe de existir un alto compromiso y participación de las diferentes áreas de la institución la cual puede implicar un cambio cultural en la organización que puede ser lo que más tiempo requiera en la iniciativa de calidad de datos.

A través de esta primera iteración de arquitectura empresarial, logramos mejorar la calidad de los datos que se buscaban introducir en la base de estudio, logrando tener grandes resultados y en los que pudimos constatar como los datos son de todo tipo y provienen de múltiples fuentes, por lo que no hay que centrarse en un área en concreto para la revisión de la calidad de datos, olvidando a las demás. La calidad de los datos se puede gestionar en cualquier punto del flujo de negocio, he de aquí que se debe buscar un bajo acoplamiento a un punto determinado del flujo.

Asimismo, a pesar de que sí resulta fácilmente adaptable a otros casos de estudio, se recomienda realizar un análisis previo sobre la herramienta a usar para llevar la gestión de la calidad de datos, ya que algunas herramientas de licenciamiento pueden ofrecer funcionalidades más acordes a las necesidades que se tiene como negocio, pero no hay que perder de vista que por muy bien diseñado que esté un programa de gestión de la calidad de datos, si los usuarios de negocio y administradores del programa, no conocen sus responsabilidades o las pasan por alto en sus interacciones con los activos informacionales de la organización, no se podrá avanzar hacia los objetivos de calidad.

4.2 SEGUNDA ITERACIÓN

La capacidad de sacar partido y entender la información está íntimamente ligada a nuestra capacidad para explotarla y transformarla en algo más que solo un dato, de forma que se busque que estos adquieran un significado.

La visualización de datos es la presentación gráfica de información con dos propósitos. Por un lado, la interpretación y construcción de significado a partir de los datos (es decir, el análisis); y, por otro lado, la comunicación.

La visualización es una herramienta muy potente para descubrir y comprender la lógica que se encuentra detrás de un conjunto de datos, así como para compartir esta interpretación con otras personas desde un punto de vista objetivo.

Es importante tener en cuenta que, aunque la visualización de datos se utiliza para representar generalmente variables cuantitativas y relaciones entre ellas, también puede utilizarse para representar relaciones entre entidades de naturaleza cualitativa.

Para esta segunda iteración se realizarán las visualizaciones de datos, a partir de los datos del área Financiera que fueron limpiados en la primera iteración del ADM de TOGAF, mostrando el tipo de visualización más adecuada para el tipo de dato que se presenta, cumpliendo con las características que debe de tener una visualización para que cumpla con su objetivo.

4.2.1 ITERACIÓN DE CONTEXTO

4.2.1.1 FASE 0: PRELIMINAR

El área financiera del banco, busca poder automatizar el reporte de incidentes dirigiéndolos con la persona adecuada para su resolución, de forma que cualquier área de negocio pueda realizar el levantamiento de estos, sin necesidad de mandarlos con un área intermediaria, quien es la responsable de los datos de contacto hoy en día.

El alcance de este caso de estudio busca implementar un sistema web para el levantamiento de los reportes de incidencias, el cual pueda ser accedido a través de la intrared de la institución, con el objetivo que los trabajadores y personas interesadas en estos datos, puedan visualizar y analizar los datos de forma clara y sencilla permitiendo reducir tiempos y mejorando la toma de decisiones.

El objetivo principal es dar a los usuarios (aquella persona con privilegios concedidos por el personal autorizado del proceso), la capacidad de reportar los incidentes (interrupción en algún flujo administrativo), de forma más rápida, reconociendo cuales fueron los eventos (sucesos imprevistos generados por una causa y teniendo una consecuencia) y la causa de estos (ocurrencia identificada para la falla en el estado de un sistema, servicio o red), enviando de forma automática un reporte (informe estructurado y contextualizado sobre problema en el flujo de negocio, alertando a personal de la institución para buscar resolver la necesidad del usuario), para su rápida atención.

4.2.1.2 FASE A: VISIÓN DE LA ARQUITECTURA

Misión de esta arquitectura: Contar con una serie de visualizaciones de datos que cuenten con información veraz, completa y actualizada para la mejor toma de decisiones dentro de la institución.

Visión de esta arquitectura: Gestionar las visualizaciones de datos, contando con datos confiables y actualizados, los cuales puedan ser accedidos y usados por los empleados de la institución, fortaleciendo la gestión organizacional dentro de la empresa.

Para esta iteración se busca llegar a:

- ✚ Compartir la importancia de las visualizaciones de datos.
- ✚ Proporcionar un esquema que muestre cómo son utilizadas las visualizaciones de datos por elementos del negocio.
- ✚ Desarrollar una arquitectura funcional para la implementación de visualizaciones de datos, en una institución.
- ✚ Comunicar dentro de la institución, las visualizaciones para fortalecimiento de la gestión interna.

4.2.2 ITERACIÓN DE DEFINICIÓN

4.2.2.1 FASE B: ARQUITECTURA DE NEGOCIO

Procesos

Se busca poder proporcionar las directrices para la generación de visualizaciones de datos. Para esto, se debe de seguir el siguiente proceso.

Las entradas de este proceso son:

- I. Requerimientos de negocio.
- II. Datos a ser utilizados.

Las salidas de este proceso son:

- I. Nueva Visualización de datos.
- II. Respuesta al requerimiento.

Las etapas de este proceso son:

- ✚ Recibir Requerimiento
- ✚ Confirmar recepción de Requerimiento
- ✚ Atender requerimiento de acuerdo la naturaleza
- ✚ Verificar existencia de la visualización
 - Si existe una visualización que ya cumpla con los requerimientos, entonces notificar el rechazo del requerimiento. Si no existe la visualización:
 - Validar las reglas del requerimiento
 - Analizar tipo de visualización a usar
 - Creación de la visualización
 - Almacenamiento de la visualización
 - Notificar la creación de la visualización

Se debe seguir el flujo mostrado en la figura 4.27, para la creación de la visualización de datos.

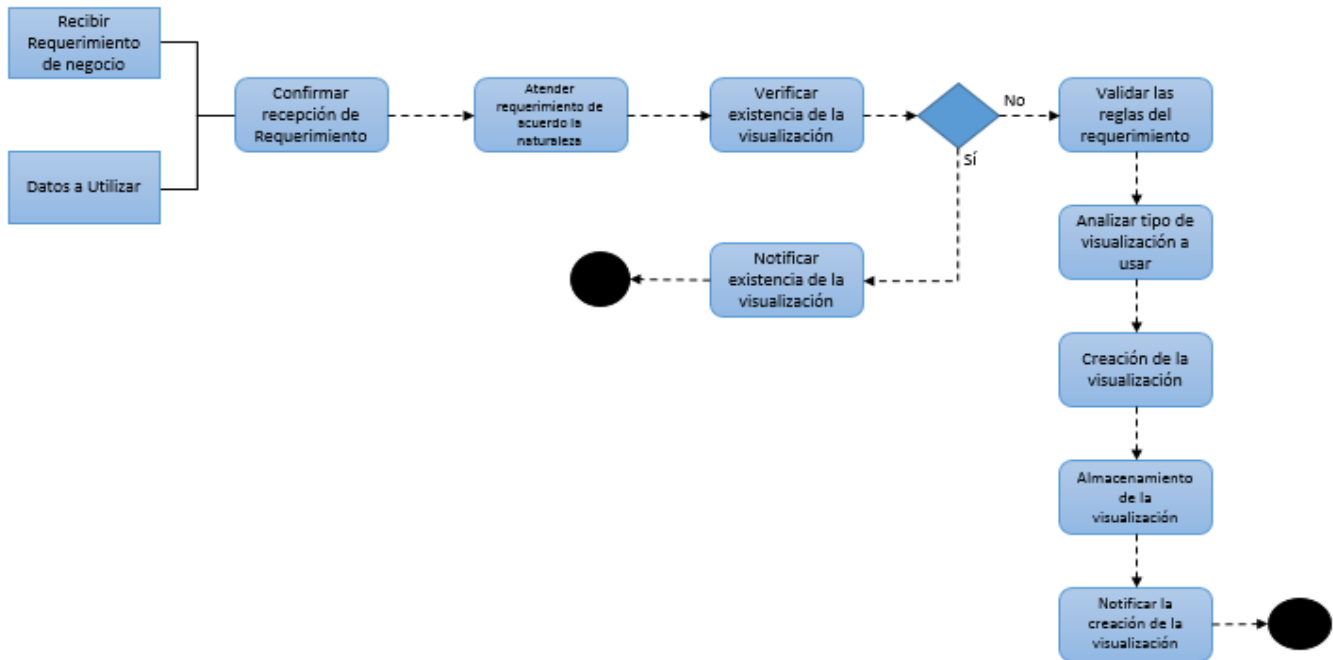


Figura 4.27 Proceso para la creación de nuevas visualizaciones

Funciones

Las funciones para la creación de visualizaciones de datos son las mostradas en la figura 4.28:

- ✚ Importancia de requerimiento
Revisar las prioridades de negocio de los requerimientos de visualizaciones se son ingresadas
- ✚ Identificación de datos
Tiene la funcionalidad para reconocer información que será usada en el requerimiento ingresado
- ✚ Análisis de requerimiento
Tiene la finalidad de realizar un análisis sobre el requerimiento de petición. Este análisis se basa en el resultado de las anteriores funciones para tomar una decisión
- ✚ Búsqueda de visualizaciones

Buscar información sobre las visualizaciones de datos existentes, de forma que se localice si alguna de estas ya cumple con el requerimiento solicitado

✚ Estandarización

En caso de generarse una nueva visualización, esta se deberá de realizar sobre los estándares de la institución

✚ Autorización

Se debe de por un área que funja como governance para el almacenamiento y publicación de la visualización

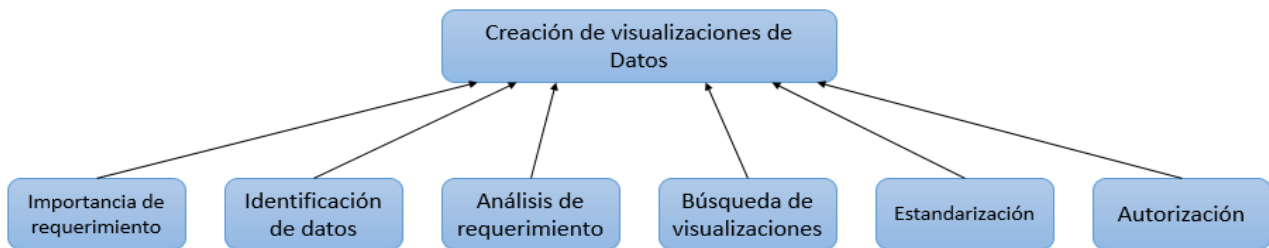


Figura 4.28 Funciones para el proceso: Creación de visualizaciones de datos

Servicios

Los servicios de negocio, son aquellos que dan soporte directo a procesos de negocio, siendo accesibles por los usuarios a través de una acción realizada típicamente en la capa de presentación. Los servicios relacionados a este proceso son los mostrados en la figura 4.29, los cuales son:

Función de Negocio	Servicio de Negocio
Importancia de requerimiento	Servicio de despliegue de solicitud de requerimientos
	Servicio para crear una nueva visualización
Identificación de datos	Servicio de despliegue de información identificada
Análisis de requerimiento	Servicio de despliegue de requerimientos de negocio
Búsqueda de visualizaciones	Servicio de resultados de búsqueda
Estandarización	Servicio de información almacenada en formato definido
Autorización	Servicio de notificación de aprobación o rechazo para su almacenamiento

Tabla 4.9 Servicios de Negocio para visualización de datos

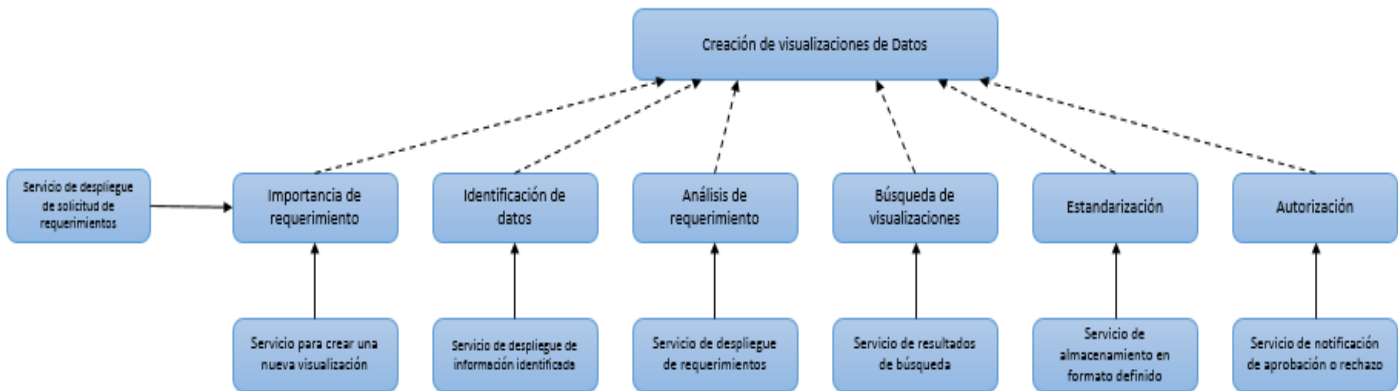


Figura 4.29 Servicios para el proceso de creación de visualizaciones de datos

Roles

Esta solución contempla los mismos roles expuestos para la limpieza de los datos, estos roles son: responsables de aprobación y de contribución.

4.2.2.2 FASE C: ARQUITECTURA DE SISTEMAS DE INFORMACIÓN

Describe el ciclo de vida de las visualizaciones de datos a lo largo de los procesos de negocio.

Para el desarrollo de las visualizaciones debido a la gran adaptación que ofrece este modelo, así como la forma en que facilita la gestión del desarrollo, haciendo que se empiece el desarrollo con mayor rapidez esto gracias a que se deben de contar con los requisitos definidos al inicio, en comparación de otros modelos ágiles de implementación, se hará uso del modelo en espiral, mostrado en la figura 4.30.

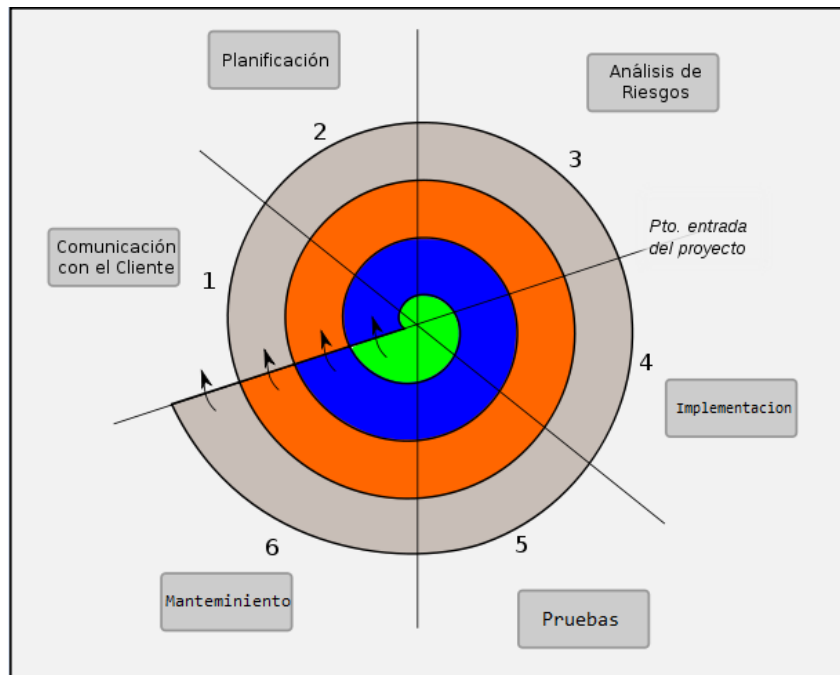


Figura 4.30 Modelo en Cascada para desarrollo de Software

- ✚ **Comunicación con el cliente:** Consiste en reunir las necesidades del producto y casi siempre su salida es texto.
- ✚ **Planificación:** Describe la estructura interna del producto y suele representarse con diagramas y texto.
- ✚ **Análisis de Riesgos:** Las tareas requeridas para evaluar riesgos técnicos y de gestión
- ✚ **Implementación:** Es la parte donde se realiza la programación. El producto de esta etapa es el código en cualquier nivel, incluido el producido por sistemas de generación automática.
- ✚ **Pruebas:** Esta etapa se centra en los procesos lógicos internos del software, asegurando que todas las sentencias se han comprobado, y en la detección de errores.
- ✚ **Mantenimiento:** Debido a que el programa puede tener errores, puede no ser del completo agrado del cliente o puede necesitar, eventualmente acoplarse a los cambios en su entorno. Esto quiere decir que no se rehace el programa, sino que sobre la base de uno ya existente se realizan algunos cambios.

Comunicación con el cliente

Para este caso de estudio, la institución financiera, desea implementar una visualización, donde se muestre el contacto del proveedor quien es el encargado de resolver la incidencia a causa de un evento determinado, de forma que el reporte pueda ser levantado por cualquier área de negocio dentro de la institución para darle su oportuna atención, he de aquí la necesidad de contar con una base de contactos depurada y actualizada.

Para esto, analizaremos el tipo de visualización que se puede presentar, seleccionando la más óptima, de acuerdo a los objetivos presentados anteriormente.

- Tabla

Incidencia	Evento	Causante	Contacto
Incidencia 1	Evento 1	Causante 1	Contacto 1
Incidencia 1	Evento 2	Causante 1	Contacto 1
Incidencia 2	Evento 1	Causante 1	Contacto 1

Las tablas sirven para mostrar los datos prácticamente a secas, por lo cual, para los objetivos perseguidos, este tipo no es una opción por lo que queda descartada.

- Gráfico de Barra

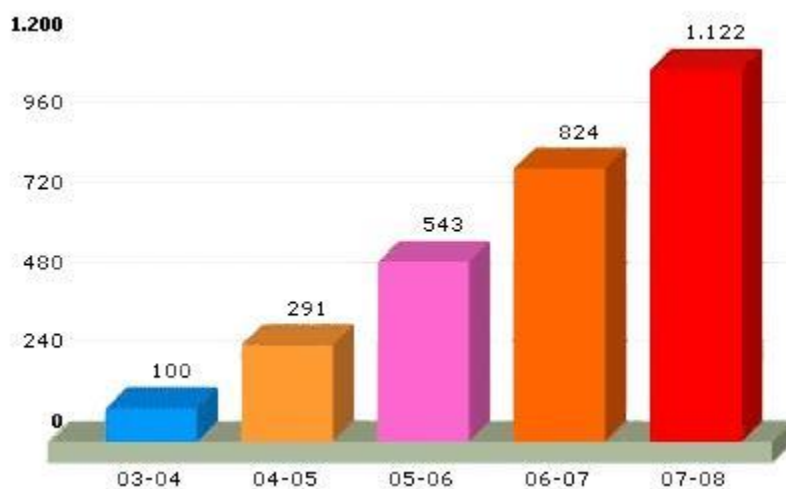


Figura 4.31 Grafica de Barras

Un gráfico de barra o columnas hace énfasis en la comparación entre elementos en un período de tiempo específico, sin embargo, no se presta para una verdadera comprensión de grandes datos, de cómo se ha llegado hasta aquí y hacia dónde vamos, como el ejemplo de la Figura 4.31.

- Gráficos de Línea

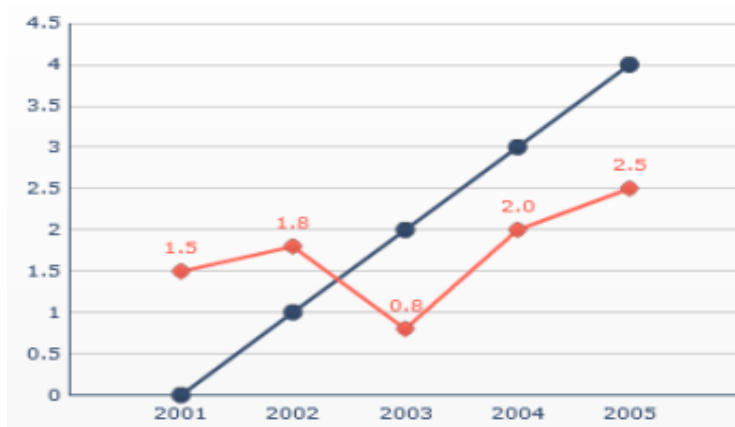


Figura 4.32 Grafica de líneas

Un gráfico de línea muestra las relaciones de los cambios en los datos en un período de tiempo, este tipo es útil para mostrar tendencias, sobre todo para las tendencias al alza, como el ejemplo de la Figura 4.32.

- Gráfico Circular

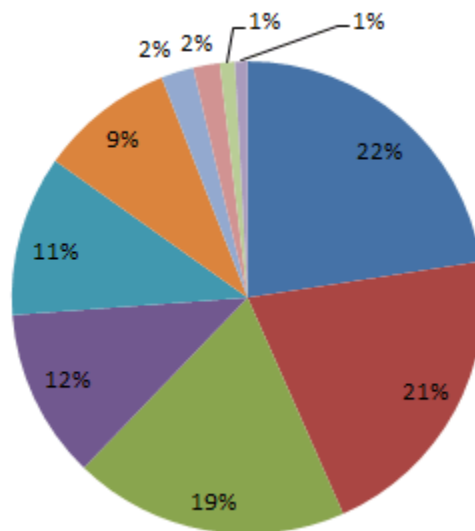


Figura 4.33 Grafica circular

Su función principal es mostrar la división de varios elementos en una cosa. Muy útiles cuando creamos un documento diseñado y pensado para un destinatario que leerá los datos por encima (dirección), como el ejemplo de la Figura 4.33.

- Gráficos de Dispersión y Burbujas

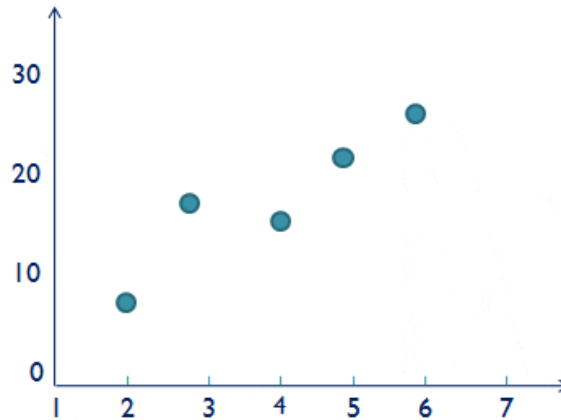


Figura 4.34 Grafica de dispersión o burbujas

Son útiles para mostrar la relación entre diferentes puntos de datos. Este tipo de gráfico utiliza valores numéricos para ambos ejes en lugar de utilizar categorías en alguno de los ejes como en los gráficos anteriores, el tamaño de las burbujas representa una dimensión adicional de los datos, añadiendo otro parámetro de medición dentro de esta.

Para el tipo de requerimiento de negocio, no se busca poder representar datos numéricos o estadísticos, por lo que descartaríamos este tipo de visualizaciones (Gráfico de Barra, Gráfico de Línea, Gráfico de Dispersión, Gráficos de Burbuja y Gráfico circular), como el ejemplo de la Figura 4.34.

- Infografía

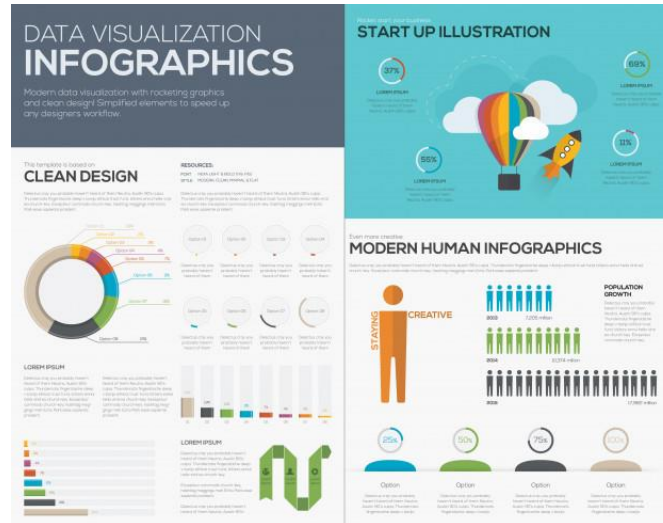


Figura 4.35 Infografía

Es una combinación de imágenes estáticas, explicativas y fáciles de entender y textos con el fin de comunicar información de manera visual para facilitar su transmisión, son usadas principalmente para compartir información, difundirla y generar discusión.

Al tratarse de una visualización estática y debido a las múltiples posibilidades de eventos y contactos que se pueden tener, esta visualización quedaría con un exceso de información, lo cual no entra dentro de las buenas prácticas, como se vio anteriormente, como el ejemplo de la Figura 4.35.

- Tree

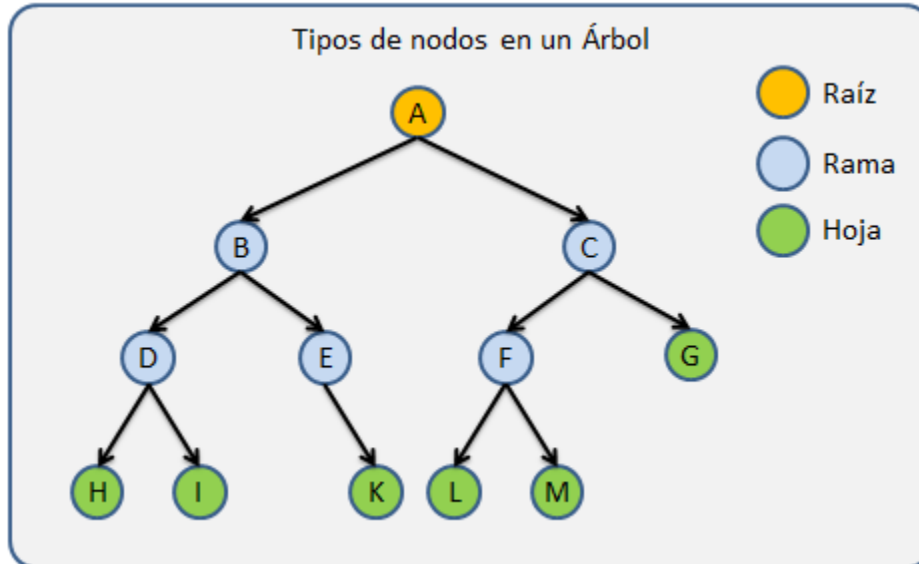


Figura 4.36 Grafico de Árbol

Una visualización de árbol muestra una serie de pasos, en las cuales cada una de las posibilidades, es acompañada de todas sus probabilidades, el resultado final es una visualización jerárquica e intuitiva, como el ejemplo de la Figura 4.36.

Realizando un análisis de cual tipo de grafica se adapta mejor a las necesidades y objetivos que se persiguen con esta visualización, nos damos cuenta que los primeros tipos (De barra, línea, circular, de burbujas y dispersión) son más usados para la representación de datos numéricos, los cuales no son el tipo que buscamos representar en este momento.

Para el tipo de visualización que se busca, de acuerdo a los objetivos planteados al inicio, se llega a que el tipo de grafica más adecuado es el jerárquico, ya que este tipo de gráfico, nos permite tener una mayor visión de los pasos a seguir dependiendo la acción tomada o cuáles son las consecuencias de cada suceso.

Se nos presentan dos opciones, un gráfico de jerarquía circular, como es mostrada en la Figura 4.37, o bien un gráfico de tipo árbol como el que se muestra en la Figura 4.38.

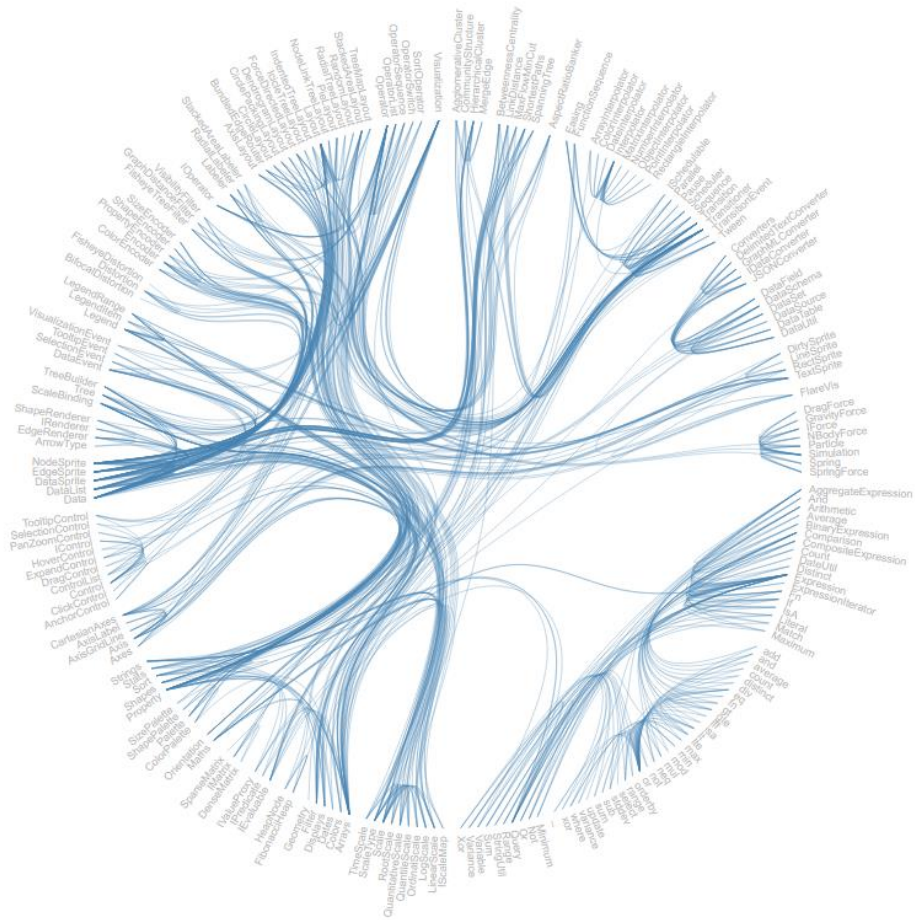


Figura 4.37 Grafico Jerárquico Circular

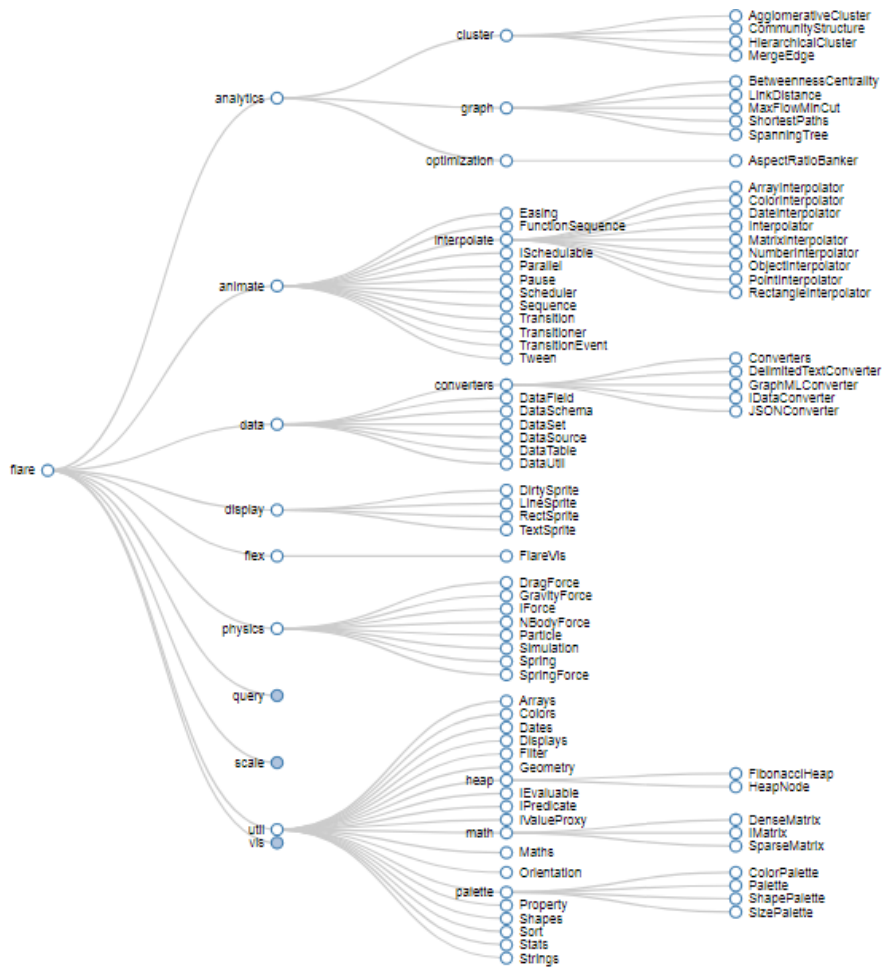


Figura 4.38 Grafico de árbol Horizontal

Del grafico Jerárquico circular tenemos la ventaja que de una sola vista podemos ver cómo se pueden relacionar varios nodos entre sí, pero la desventaja que dependiendo la cantidad de datos puede llegar a tener una sobrecarga de información, la cual la realizaría difícil de analizar y comprender, aun cuando se remarcará de un color diferente los conectores, al seleccionar un nodo, tal como es mostrado en la Figura 4.39.

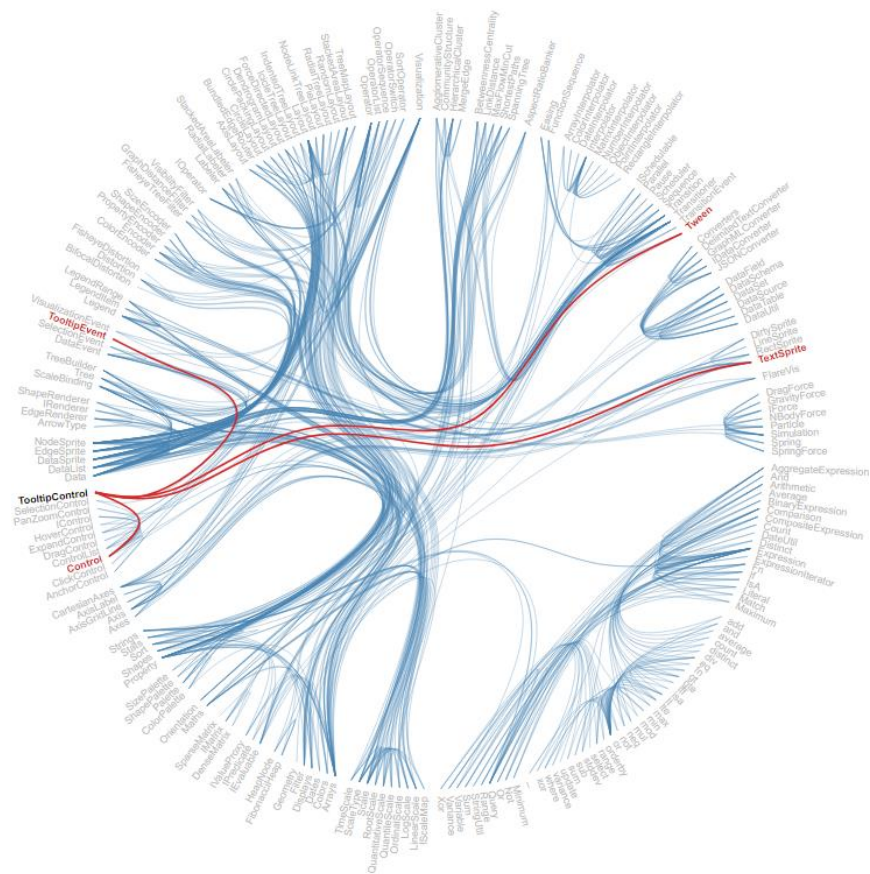


Figura 4.39 Grafico Jerárquico Circular mostrando relación

De igual forma, en la visualización de tipo árbol, podemos observar en una sola vista, todos los eventos y cuáles fueron sus causas, a diferencia que, en esta visualización, no existe cruce entre los conectores de nodo, lo cual facilita más su comprensión a diferencia de la visualización de tipo jerárquico circular.

Por tal motivo, analizando el requerimiento de negocio y las ventajas y desventajas que nos ofrece cada tipo de gráfico, se ha visto que la mejor opción para la implementación es el grafico de tipo árbol, ya que nos será de gran utilidad para mostrar los pasos a seguir para el reporte de incidencias que es el objetivo que busca la visualización.

Planificación y Análisis de riesgos

Al realizar el diseño de la aplicación, se planea realizarlo sobre una arquitectura de 3 capas, de forma que se tenga una separación entre la interfaz de usuario y la lógica

de la aplicación facilitando la reutilización de componentes o servicios para aplicaciones futuras, teniendo como modelo de arquitectura el presentado en la Figura 4.40

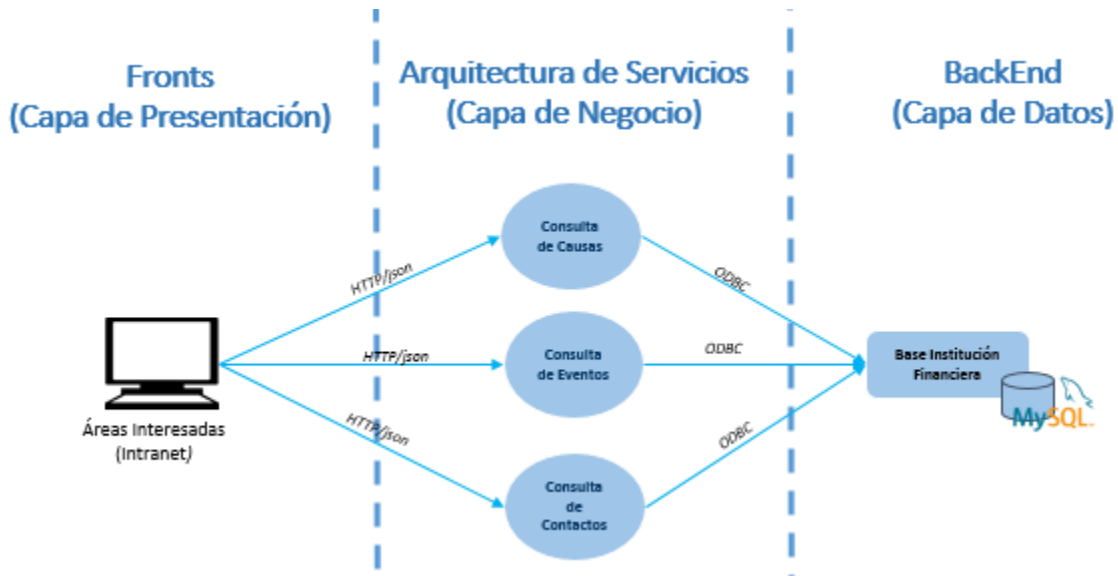


Figura 4.40 Modelo de Arquitectura de la Aplicación

Capa de Presentación

Para esta capa, se tiene que diseñar la parte del aspecto visual de la página que será mostrada a los usuarios finales, para lograrlo de manera efectiva, se debe tener en mente las siguientes preguntas que son de gran importancia

- ✚ ¿Hacia dónde se dirige la mirada primero?
- ✚ ¿Qué elementos llaman la atención del usuario inicialmente?
- ✚ ¿Son ellos diseñados para los puntos importantes en los objetivos estratégicos?

Debido a que uno de los puntos es que esta visualización sea colocada dentro de la intranet de la institución, ya se tiene un formato en cuando a cabeceras y menú definido para esto, lo cual se debe de respetar para mantener el estándar de formato de la institución.

Quedando el diseño preliminar del Front de la visualización tal como se muestra en la Figura 4.41.

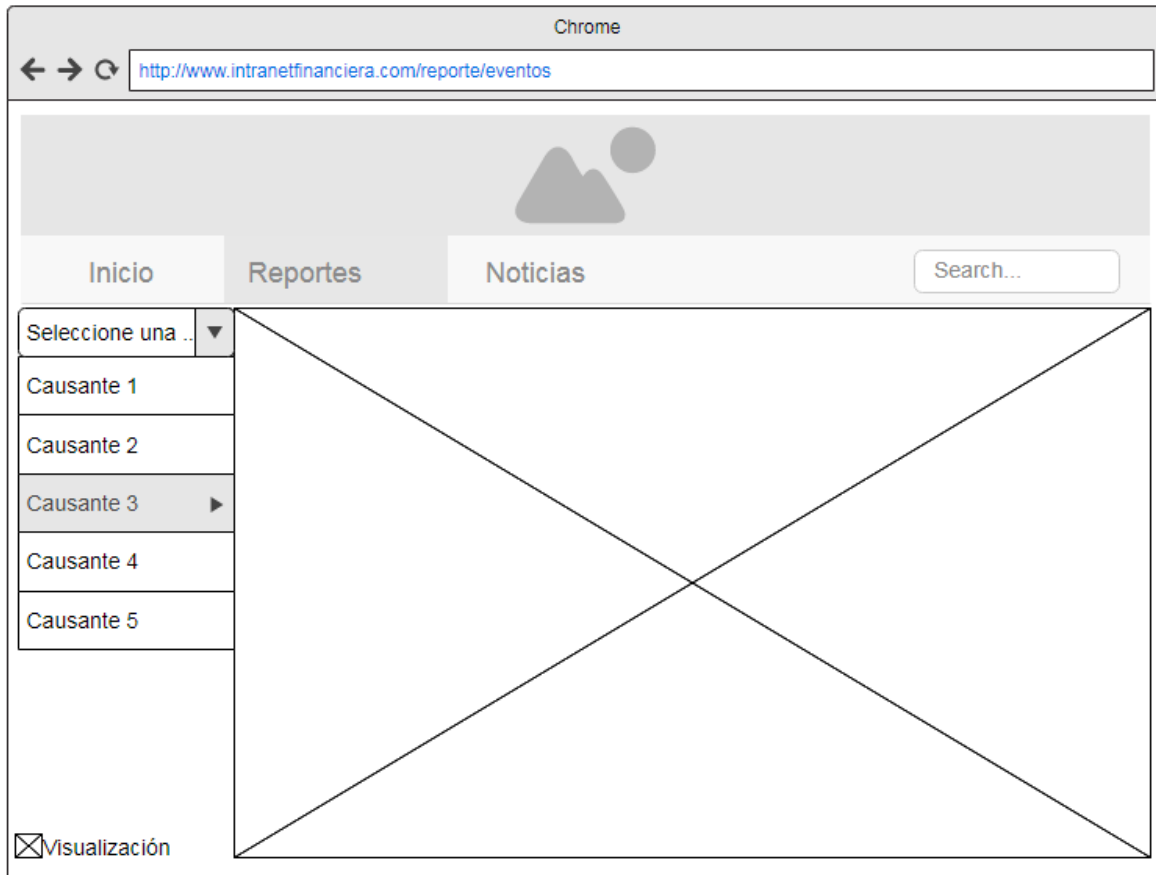


Figura 4.41 Diseño Front de la visualización

Una vez seleccionado el contacto, del responsable de resolver la incidencia, se pasará a un formulario, cuyo diseño se muestra en la Figura 4.42, donde se tendrán los datos pre llenados, para facilitar el uso de la aplicación, se podrá añadir una pequeña descripción de la incidencia para facilitar su entendimiento y su pronta resolución.

Este formulario será enviado vía correo electrónico, el cual será recuperado de la base de datos de la institución, para ser presentada al contacto responsable.

The image shows a web browser window with the address bar containing the URL <http://www.intranetfinanciera.com/reportes/eventos/formulario>. The browser title is 'Chrome'. The page layout includes a header with a logo and a navigation menu with 'Inicio', 'Reportes', and 'Noticias'. A search box is located in the top right corner. The main content area contains a form with four input fields: 'Incidencia', 'Evento', 'Causante', and 'Descripción'. Below the form is a blue 'Enviar' button.

Figura 4.42 Diseño Front del formulario.

Capa de Negocio

Se busca que los servicios de negocio sean atómicos, para su baja asociación a la aplicación y su futura reutilización.

Por lo tanto, el diseño de los servicios de negocio en cuanto a sus entradas y salidas queda representado en las siguientes tablas:

Servicio Consulta de Causa	
Entrada	Salida
/Get/'Incidente'	Json con las causas resultado del incidente seleccionado

Tabla 4.10 Servicios SOA

Servicio Consulta de Eventos	
Entrada	Salida
<i>/Get/’Causa’</i>	Json con los eventos resultado de la causa seleccionada

Servicio Consulta de Proveedores	
Entrada	Salida
<i>/Get/’Evento’</i>	Json con los proveedores a quien se puede levantar la incidencia resultado del evento seleccionada

Tabla 4.10 Servicios SOA

Capa de Datos

Para la capa de Datos, ya se tiene una base de datos depurada y actualizado resultado de lo realizado previamente,

Implementación

Uno de los problemas más importantes a la hora de hacer posible el diseño y ejecución, es la selección de lenguaje de desarrollo o implementación de la parte frontal o visual de la página.

La evaluación de los lenguajes para el diseño visual es un problema difícil de resolver. Esto ha sido debido, no solo a la dificultad del problema sino también a su subjetividad. El amplio ámbito de aplicación de los lenguajes hace que unos modelos puedan ser válidos ante algunas necesidades e inválidos en otros.

Arbor

Es una biblioteca adecuada que gestiona de manera eficaz abstracciones gráficas de organización y manejo de actualizaciones.

Cubism

Crean interacciones flexibles, apoyadas en el formato vector SVG para su visualización.

Envision

Es muy popular, ya que crea visualización rápida y dinámica de datos en tiempo real.

Polymaps

Es útil en la elaboración de mapas dinámicos e interactivos.

Circos






Permite crear visualizaciones de datos de manera circular, mostrando la interrelación que existe entre datos.

D3

Data-Driven Documents, Permite crear infogramas dinámicos e interactivos en navegadores web, haciendo uso de tecnologías como SVG, HTML5, y CS.

Debido al tipo de visualización que deseamos obtener, haremos uso de la librería de D3, ya que nos permite tener control completo sobre el resultado visual final, así como grandes grupos de datos pueden ser fácilmente asociados a objetos SVG de forma sencilla con funciones de D3 para generar textos extendidos o infogramas elaborados.

Para poder realizar la visualización es indispensable tener conocimientos previos tales como:

-  HTML
-  CSS
-  DOM (Modelo de Objetos del Documento)
-  Javascript
-  SVG (Scalable Vector Graphics)

Uno de los primeros pasos para la realización de la visualización consiste en la descarga de la librería d3, poniéndolo disponible dentro de la carpeta del proyecto a crear. Asimismo, es necesario crear un nuevo elemento en el DOM para el inicio de la visualización.

La visualización de datos es un proceso de mapeo de datos a despliegues gráficos. Los datos ingresan y su representación en formato gráfico constituye el resultado final.

Con D3, se asocian los datos de entrada a los elementos del DOM. Asociar en este sentido es similar a pegar a elementos específicos, de tal forma que posteriormente se puedan referenciar esos valores y aplicar las reglas de mapeo. Si no se lleva a cabo el paso de asociación, lo único que se obtiene es un listado de elementos del DOM sin datos que no pueden ser mapeados.

En D3 se utiliza el método `selection.data()` para asociar datos a los elementos del DOM. Sin embargo, es necesario tener dos componentes organizados antes de poder asociar datos:

- Los datos en sí mismos
- Una selección de elementos del DOM

Los datos

Para la obtención de los datos se realizará una consulta a una base de datos SQL, obteniendo con ella los datos mostrados en la figura 4.43, los cuales son los datos principales que serán usados para nuestra visualización

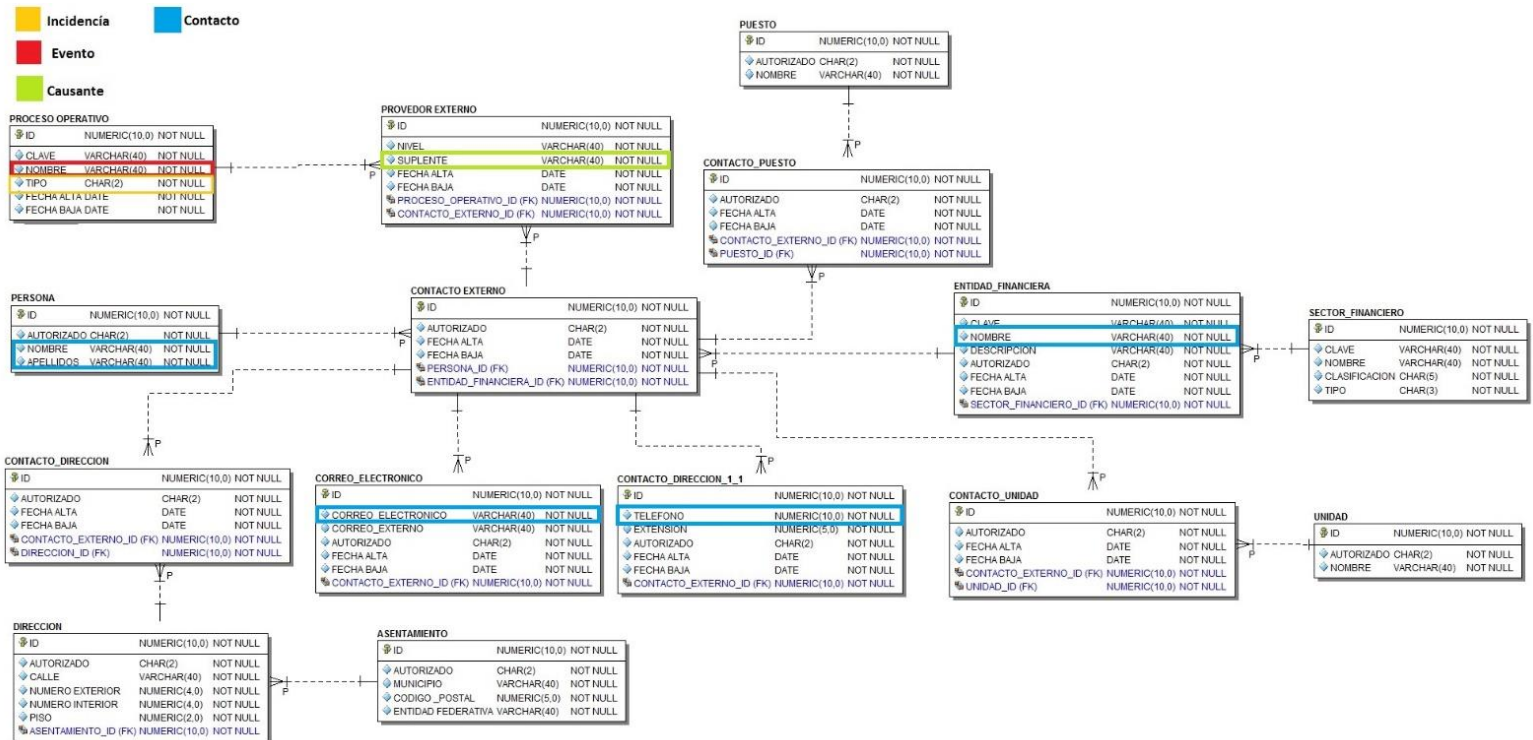


Figura 4.43 Modelo de datos necesario para la visualización de datos

A través de la realización de la consulta se buscará que nos regrese la información codificada jerárquicamente en JavaScript Object Notation (JSON), cada nodo debe tener un nombre y un nodo principal o secundario, o ambos, esto para que pueda ser representada por las librerías de D3 en una visualización.

El formato Json que se debe buscar es el basándonos en el siguiente:

```
{
  "name": "Incidencia", "subtipo": "I",
  "children": [{
    "name": "Evento", "subtipo": "E",
    "children": [{
      "name": "Causa 1", "subtipo": "C",
      "children": [
        {"name": "Contacto Entidad 1", "subtipo": "Cont"}
      ]
    }
  ]
}]
}
```

En este ejemplo podemos ver que esta ordenado de manera jerárquica de acuerdo al tipo de incidencia que se realiza de búsqueda.

Selección de elementos DOM

Se contará con una sección css donde se establecerá el estilo de los círculos que representas los nodos, el texto junto a ellos y los enlaces entre estos.

```
.node {
  cursor: pointer;
}
.node circle {
  fill: #fff;
  stroke-width: 1.5px;
}
.node text {
  font: 10px sans-serif;
}
.link {
  fill: none;
  stroke: #ccc;
  stroke-width: 1.5px;
}
```

La siguiente sección declara algunas de las funciones estándar para la visualización, como el tamaño y la forma del contenedor svg con márgenes incluidos.

```
var margin = {top: 20, right: 120, bottom: 20, left: 120},
    width = 960 - margin.right - margin.left,
    height = 600 - margin.top - margin.bottom;
```


La función `tree` se utiliza para asignar y calcular los nodos y enlaces necesarios para el siguiente nivel de nuestra visualización.

```
var tree = d3.layout.tree()  
    .size([height, width]);
```

La función `diagonal` se usará para dibujar los enlaces entre los nodos, ayudando a dibujar un camino entre dos puntos, haciendo que la línea muestre algunas curvas para que la visualización sea más atractiva.

```
var diagonal = d3.svg.diagonal()  
    .projection(function(d) { return [d.y, d.x]; });
```

El siguiente bloque de código agrega el área de trabajo SVG al cuerpo de la página web y creando un grupo de elementos (`<g>`) que contendrán los objetos svg (nodos, texto y enlaces).

```
var svg = d3.select("body").append("svg")  
    .attr("width", width + margin.right + margin.left)  
    .attr("height", height + margin.top + margin.bottom)  
    .append("g")  
    .attr("transform", "translate(" + margin.left + "," +  
margin.top + ")");
```

Una de las funciones más importantes es la de `update()`, pues es la parte del código que reúne las funciones y datos que hemos declarado y realiza la impresión del árbol.

El primer paso en ese proceso es asignar nuestros nodos y enlaces.

```
var nodes = tree.nodes(root).reverse(),  
    links = tree.links(nodes);
```

La función `nodeEnter` ayuda a agregar un nodo a una posición en particular.

```
var nodeEnter = node.enter().append("g")  
    .attr("class", "node")  
    .attr("transform", function(d) { return "translate(" +  
source.y0 + "," + source.x0 + ")"; })  
    .on("click", click);
```

A partir de estas funciones es que obtendremos una visualización como la que es mostrada en la Figura 4.44, la cual se buscará ir adaptando de acuerdo a las buenas prácticas de visualizaciones de datos que se han repasado anteriormente.

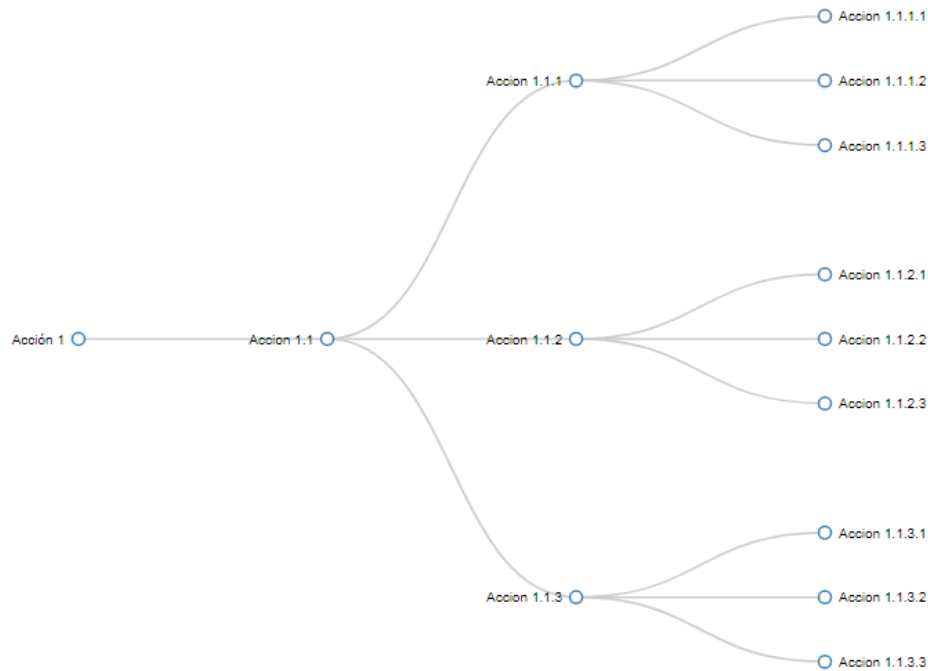


Figura 4.44 Visualización de árbol estándar.

D3 es de mayor utilidad cuando se usa para generar información gráfica en formato SVG. El estilo por defecto de los SVGs es llenado en negro y sin línea. Si se necesita algo diferente, se tienen que aplicar estilos a los elementos.

Para el caso de la visualización, hemos visto la importancia del uso de colores para poder destacar algunos elementos, así como del uso de los contrastes para poder anteponer la facilidad de la lectura, por lo mismo se buscará que los nodos del árbol puedan ser diferenciados a través del color, de forma que se tenga una apreciación visual más rápida del conocimiento e información que se está tratando de compartir. Se usará la siguiente función para dar estilo a los nodos:

```
nodeUpdate.select("circle")
.attr("r", 8.5)
.style("fill", function(d) {
  if(d.subtipo == "Nodo1"){ return d._children ?
    "Red" : "#EC8282"; }
  else if(d.subtipo == "Nodo2"){
    return d._children ? "Orange" : "#F6BF52"; }
})
```

```
else if(d.subtipo == "Nodo3"){  
    return d._children ? "Yellow" : "#F9D806"; }  
else if(d.subtipo == "Nodo 4"){  
    return d._children ? "LightGreen" : "#C1F210"; }  
});
```

De forma que obtenga una visualización más personalizada como la mostrada en la figura 4.45

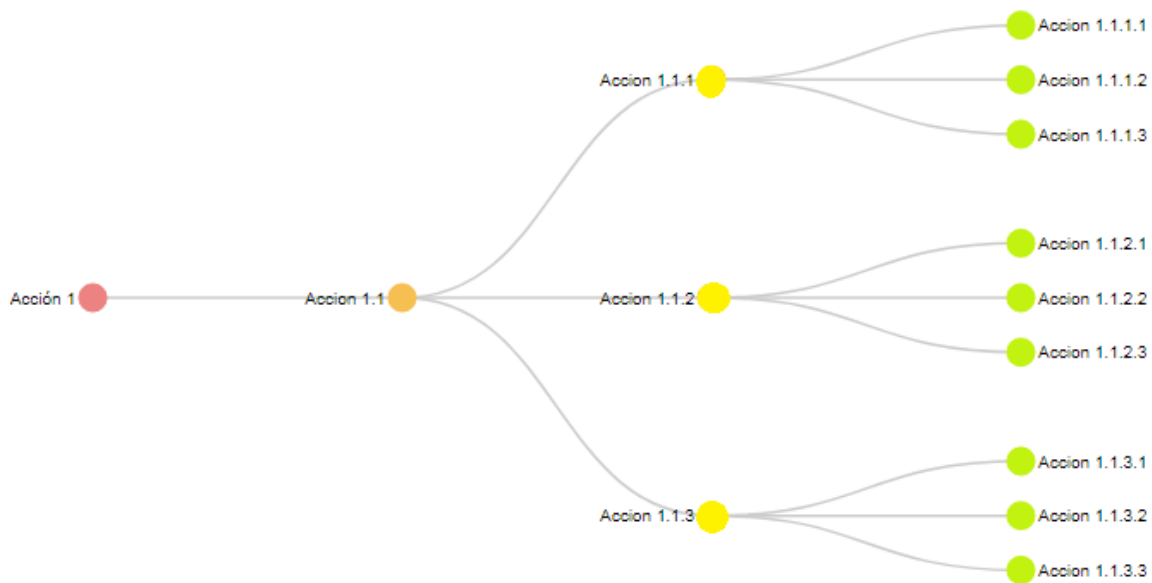


Figura 4.45 Visualización personalizada a tipo de nodos

Como hemos visto el desarrollar visualizaciones de datos haciendo uso de la biblioteca de D3, es muy sencillo y se puede adaptar a las necesidades que se tengan en cada negocio. Por ejemplo, una de visualizaciones adicionales, que fueron desarrolladas con el uso de la misma biblioteca, para la presentación de datos es la mostrada en la Figura 4.46

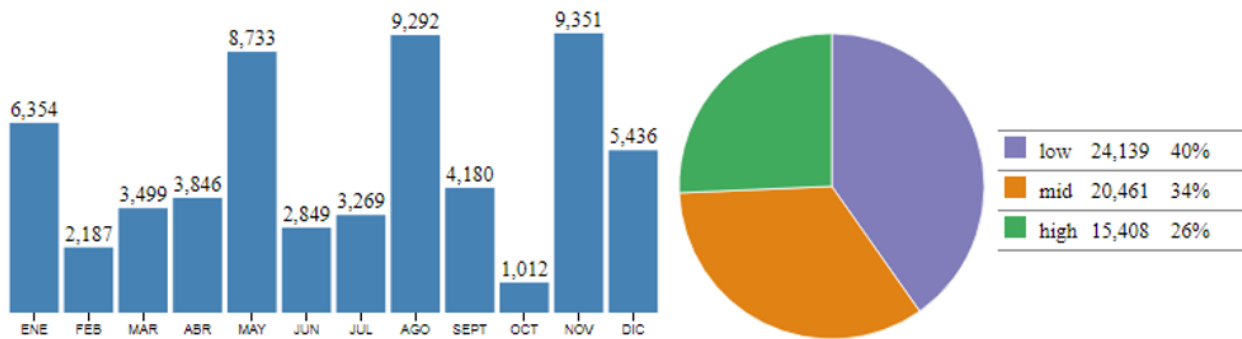


Figura 4.46 Visualización adicional creada con D3

Esta es una visualización que nos presenta tres de los tipos de grafico que analizamos previamente, tal como grafico de barras, grafico circular y tabla, en el cual cada que pasemos el mouse por alguno de los tres sectores de la gráfica circular o la tabla, se verá reflejado el cambio en la gráfica de barras para observar su comportamiento a lo largo del tiempo, tal como se muestra en la Figura 4.47.

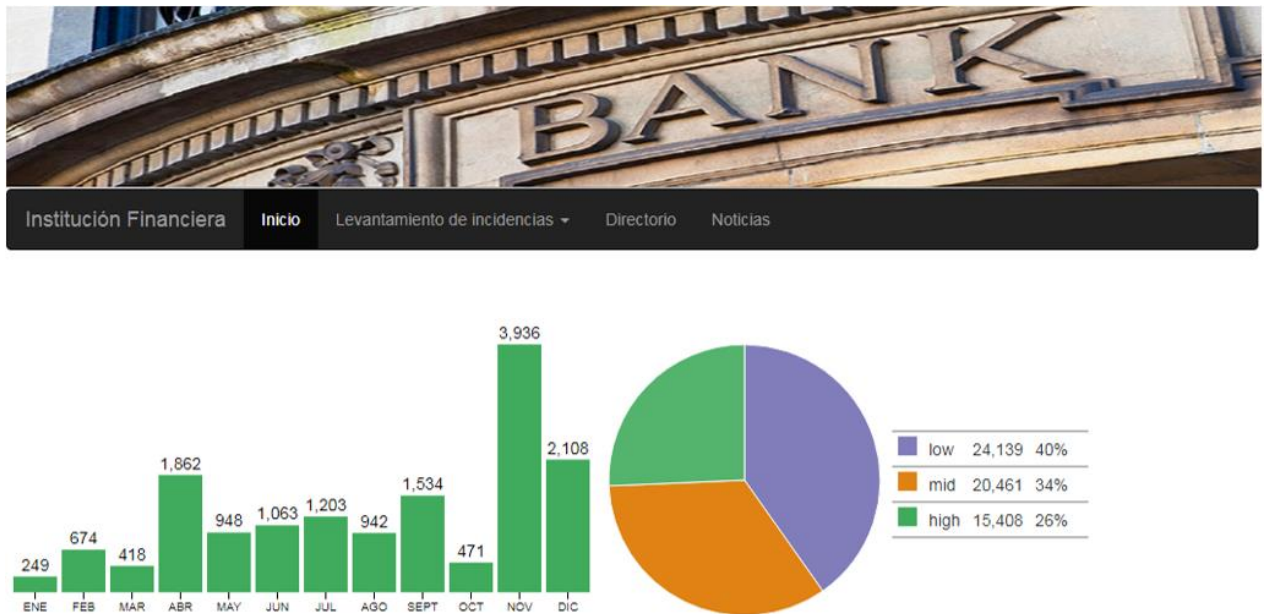


Figura 4.47 Comportamiento de la Visualización adicional creada con D3

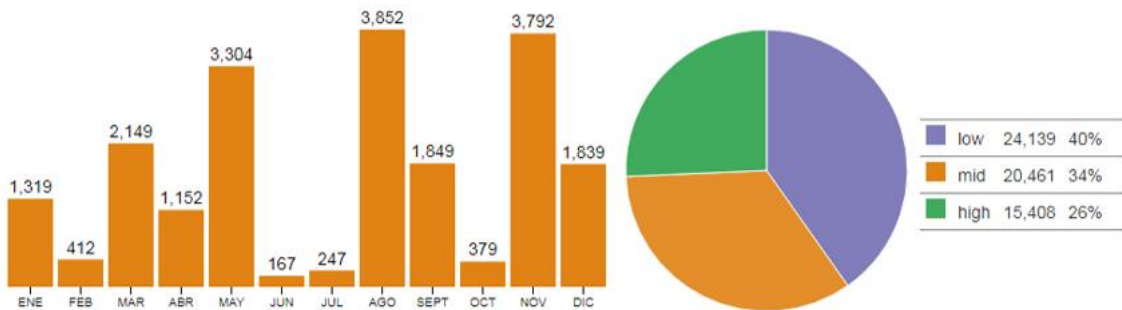
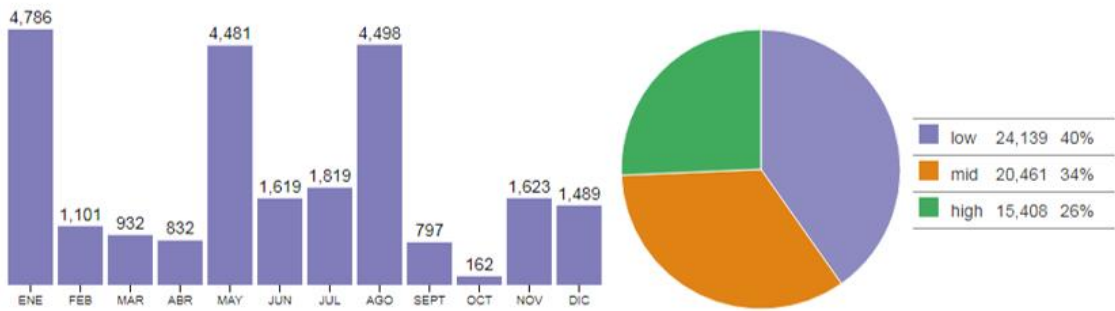


Figura 4.47 Comportamiento de la Visualización adicional creada con D3

Pruebas

El plan de pruebas se debe de elaborar para atender los objetivos de calidad de la visualización, para lo cual se propone seguir lo siguientes pasos:

1. Analizar los requerimientos

Para elaborar un plan de pruebas de software lo primero que se debe realizar es entender los requerimientos de usuario que componen la iteración o proyecto, que son el sujeto de la verificación de calidad que se va a realizar.

2. Identificar las funcionalidades nuevas a probar

Se debe identificar e incluir la lista de las funcionalidades totalmente nuevas.

3. Identificar las funcionalidades de sistemas existentes que deben probarse

Se debe identificar las funcionalidades existentes que estén siendo impactadas por el desarrollo de alguna forma, considerando todos los componentes afectados en todas las capas de la arquitectura de software.

Existen dos situaciones que se puede encontrar al identificar estas funcionalidades:

- Funcionalidades modificadas de cara al usuario
- Funcionalidades modificadas en sus componentes internos

4. Definir la estrategia de pruebas

Consiste básicamente en seleccionar cuáles son los tipos de pruebas de software que se deben realizar.

Se determinan los conjuntos de pruebas a realizar, correspondiente con cada funcionalidad nueva o existente que se esté modificando.

5. Identificar los entornos (ambientes) requeridos

Documentar las características de los entornos de Hardware y Software necesarios para realizar la ejecución de las pruebas de software.

Como mejor práctica, el ambiente de pruebas de software debería ser lo más similar posible al ambiente de producción.

6. Identificar los riesgos y definir planes de respuesta

Los riesgos por lo general están vinculados con factores como:

- Posibles dificultades en la disponibilidad de entornos.
- Pruebas que dependen de factores externos al proyecto y la organización.

- Dependencias con otros proyectos.

Al tratarse de un sistema web, se recomienda realizar pruebas de Interfaz del Usuario, y de compatibilidad.

En la prueba de Interfaz de usuario, las características de la interfaz se prueban para asegurar que las reglas del diseño, la estética y el contenido visual relacionado están a disposición del usuario sin error alguno, los mecanismos individuales de la interfaz se prueban en forma unitaria, asimismo, la interfaz se prueba frente cada caso de uso (USN) para descubrir errores de semántica, y facilidad de uso.

Mientras que en la prueba de compatibilidad se realiza para descubrir errores asociados con un ambiente específico. Para esto se debe definir un conjunto de configuraciones de computadoras cliente más comunes y en función de estas armar las pruebas.

Mantenimiento

Con el mantenimiento se busca poder encontrar y corregir los problemas menores antes de que estos provoquen fallas.

El tiempo con el que se dé un mantenimiento dependerá de las necesidades del usuario, pero lo ideal es tener un mantenimiento planificado para poder evitar errores antes de que estos sucedan.

El mantenimiento puede ser definido como una lista completa de actividades, todas ellas realizadas por los usuarios y operadores de la visualización.

En conclusión vemos como hacer uso de este modelo nos ha permitido tener un sistema más adaptable al ser sobre un enfoque evolutivo permitiendo entender y reaccionar a los riesgos en cada nivel, pues el sistema se evalúa y se prueba con todos los elementos necesarios y el mantenimiento se realiza sobre una base continua para prevenir grandes fallas y reducir al mínimo el tiempo perdido.

4.2.2.3 FASE D: ARQUITECTURA DE TECNOLOGÍA

Siguiendo el modelo de arquitectura de la aplicación mostrado en la Figura 4.40 se muestra el diseño que se deberá seguir, buscando en esta sección profundizar más a nivel técnico el desarrollo de esta, el cual se basará en el modelo mostrado en la Figura 4.48, donde se observa la interacción entre cada una de las capas mostradas previamente.

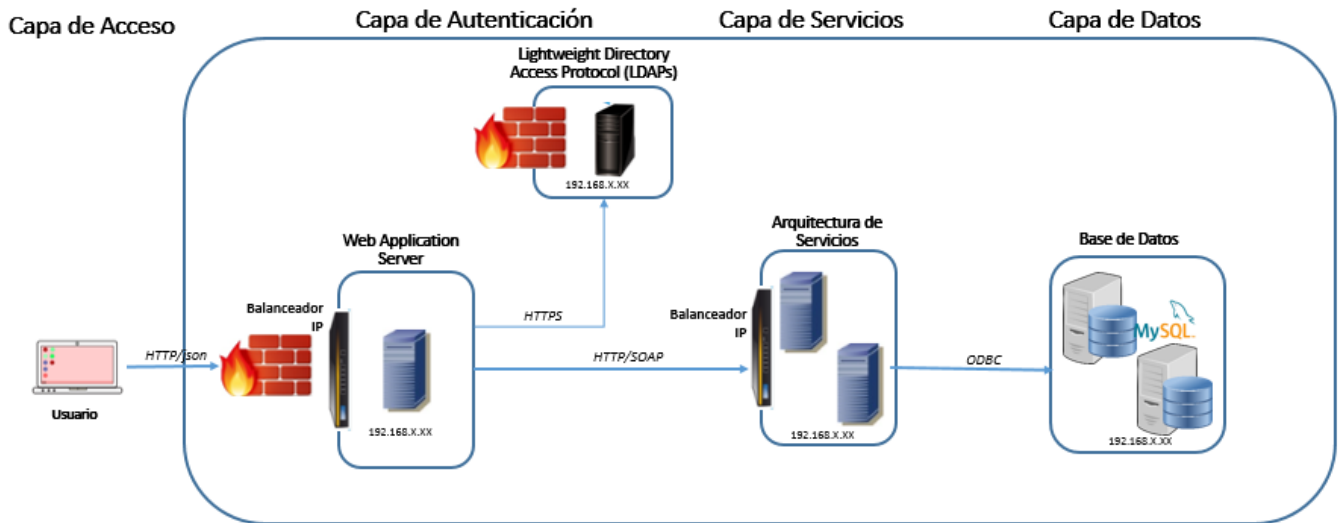


Figura 4.48 Modelo de Técnico para el desarrollo de la aplicación

Se debe considerar una comunicación LDAPs¹ desde el Web Application Server, para la validación de credenciales de los usuarios, así como la infraestructura, recursos de espacio en disco y memoria necesaria para el correcto funcionamiento de la aplicación

4.2.3 ITERACIÓN DE PLANEACIÓN

4.2.3.1 FASE E: OPORTUNIDADES Y SOLUCIONES

La estrategia de implementación conduce la implementación para obtener como resultado a la arquitectura objetivo.

Estrategia de implementación

Creación de visualizaciones	
Proyectos	Descripción
Definición de los requerimientos	Este proyecto contempla las actividades para definir conceptualmente las visualizaciones y el contenido que se tendrán en ellas.
Diseño de visualización	Este proyecto contempla las actividades para el diseño de la visualización de datos

Tabla 4.11 Proyectos para la implementación de las visualizaciones

¹ LDAP son las siglas de Lightweight Directory Access Protocol es un protocolo a nivel de aplicación que permite el acceso a un servicio de directorio, donde se almacena la información de autenticación (usuario y contraseña)

4.2.3.2 FASE F: PLANEACIÓN DE MIGRACIÓN

Basándonos en la estructura WBS, se muestra en la Figura 4.46 la estructura para este alcance.

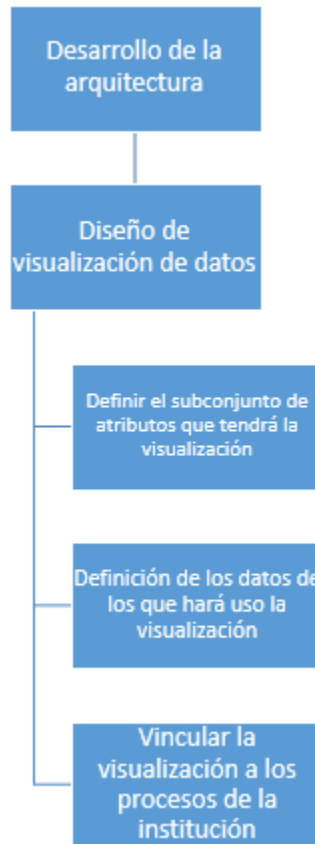


Figura 4.49 Estructura de desglose de trabajo para la creación de la visualización

Debido a que este proyecto fue iniciado como una necesidad nueva, y no como una actualización a una arquitectura ya existente, los resultados finales de la planeación e implementación de la arquitectura son el desarrollo de estas visualizaciones, las cuales serán analizadas más a detalle en el siguiente punto.

4.2.4 ANALISIS DE RESULTADOS

Una vez concluida esta iteración, pudimos obtener como resultado que una visualización que tiene como objetivo el simplificar el proceso de levantamiento de reportes por diversas incidencias, haciendo llegar este a la persona adecuada de forma sencilla para cualquier área de negocio.

Obtuvimos como resultado una visualización de tipo Tree, la cual es mostrada en la Figura 4.50.

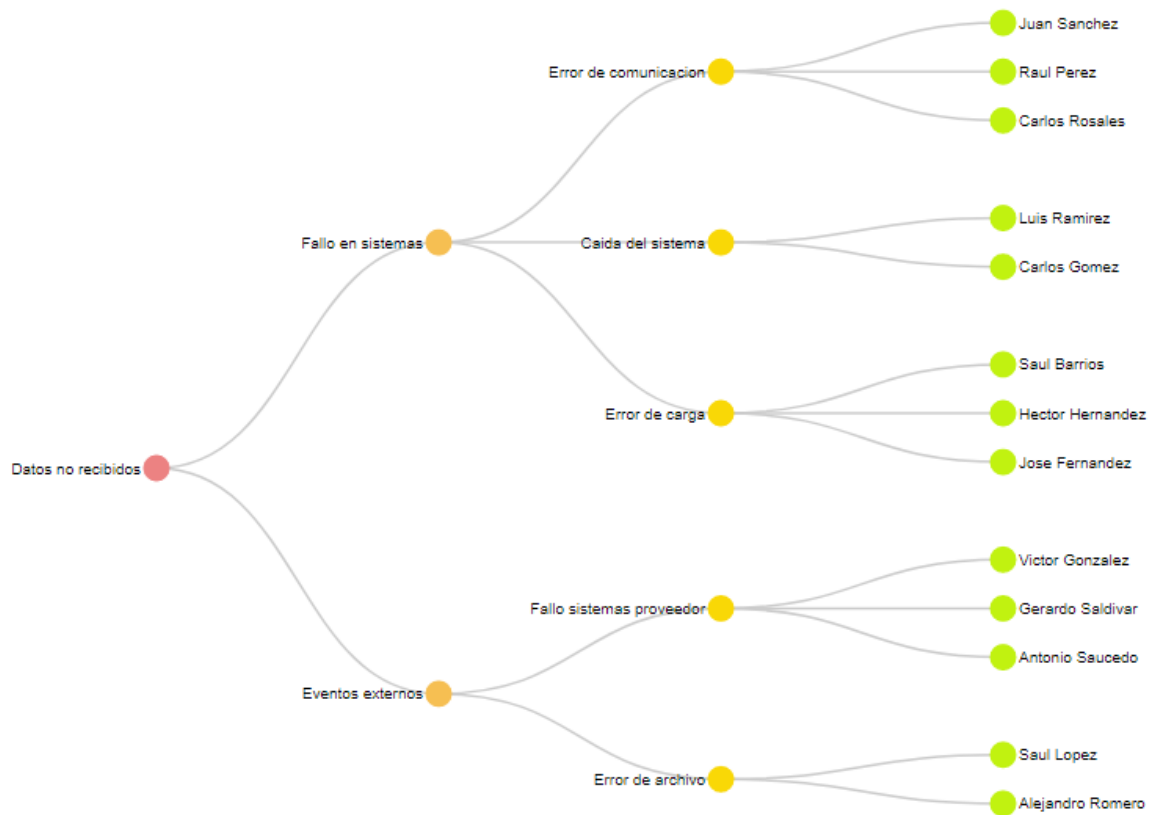


Figura 4.50 Visualización tipo Tree obtenida.

Para poder hacer una visualización de forma más interactiva, se requiere de una solución dinámica, de forma que al inicio solo nos mostrará todos los posibles eventos (nodos de segundo nivel) que llevaron al incidente, tal como es mostrado en la Figura 4.51

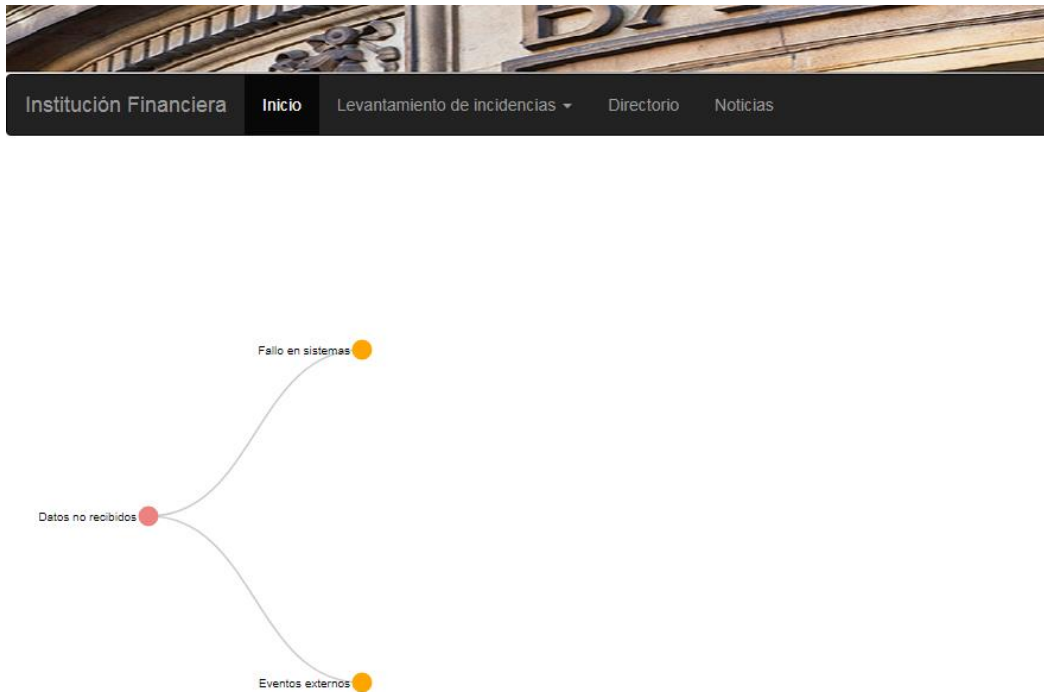


Figura 4.51 Visualización mostrando los nodos de segundo nivel

A partir del evento que seleccionará el personal de la institución que este levantando la incidencia, dará clic sobre el nodo correspondiente, mostrando los nodos del tercer nivel (Causas) del incidente, tal como es mostrado en la Figura 4.52

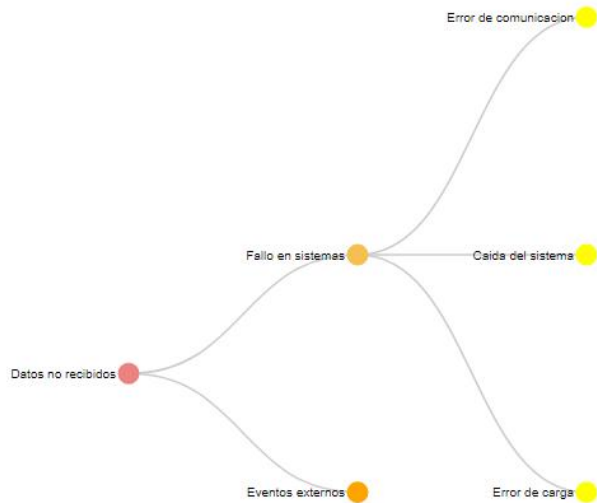


Figura 4.52 Visualización mostrando los nodos de tercer nivel

De igual manera una vez seleccionada la posible causa el incidente, dará clic sobre el nodo correspondiente, mostrando los nodos del cuarto nivel, correspondiente a los contactos a los cuales se les puede levantar la incidencia para una correcta atención y solución a este, tal como es mostrado en la Figura 4.53

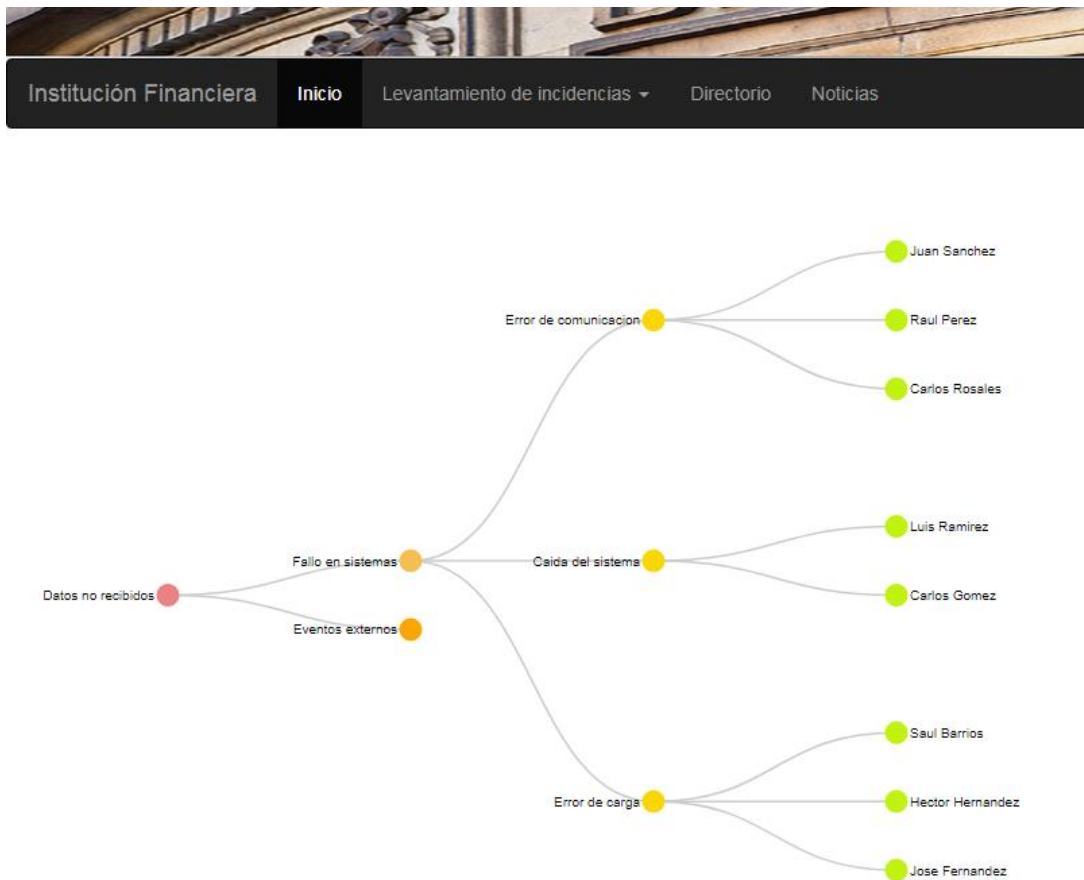
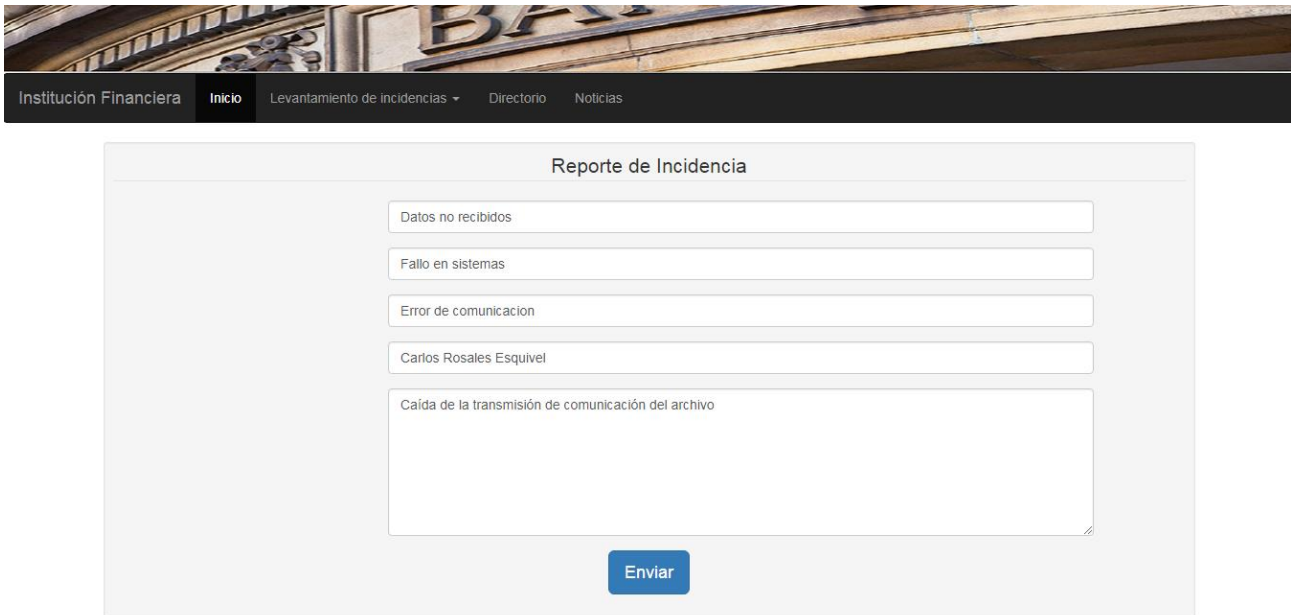


Figura 4.53 Visualización mostrando los contactos a atender la visualización

Finalmente, al seleccionar el contacto al cual se quiere reportar la incidencia, serán pasados los datos a través de la sesión de la página, para su llenado de forma automática en un formulario, donde el empleado tendrá la posibilidad de agregar una descripción más amplia del problema suscitado, tal como es mostrado en la Figura 4.54. De forma que se envíe vía correo electrónico a la dirección del proveedor, cuyo correo electrónico fue recuperado a través del consumo de un servicio, el cual no es visible en pantalla, solo almacenado en una variable de sesión de la página, mostrando si el mensaje fue enviado de manera correcta o no, como es mostrado en la Figura 4.55.



The screenshot displays a web application interface for reporting incidents. At the top, there is a navigation bar with the following items: "Institución Financiera", "Inicio", "Levantamiento de incidencias" (with a dropdown arrow), "Directorio", and "Noticias". Below the navigation bar is a form titled "Reporte de Incidencia". The form contains five input fields: "Datos no recibidos", "Fallo en sistemas", "Error de comunicación", "Carlos Rosales Esquivel", and "Caída de la transmisión de comunicación del archivo". A blue "Enviar" button is positioned at the bottom center of the form.

Figura 4.54 Formulario para envío de reporte



Figura 4.55 Mensaje de envío correcto de reporte

4.2.3 CONCLUSIONES DE SEGUNDA ITERACIÓN

Recoger, registrar y actualizar información son acciones importantes que optimizan la gestión y dan paso a un mayor desarrollo, pero exigen que se le dedique una parte importante de tiempo para cuidar todo este proceso.

A lo largo de esta iteración, vimos la importancia de integrar las nuevas tecnologías a favor de cualquier empresa o institución, de forma que se puedan crear herramientas más accesibles, que supongan un ahorro de tiempo y con resultados satisfactorios, de forma que los colaboradores de la institución tengan una mayor facilidad para la realización de sus tareas.

Para este caso, apoyándonos de las nuevas tecnologías, se planteó el desarrollo de una visualización de datos, que ayudará y facilitará las tareas diarias del personal, la cual a partir de los requerimientos dados por negocio, se analizó la mejor opción en cuanto a tipo de visualización para que cumpliera su objetivo y presentará los mejores resultados obteniendo una visualización de árbol, en la cual nos ayudó a tener una representación de los aspectos de una decisión, de forma que se conocía el proceso de inicio a final, permitiendo tomar decisiones adecuadamente. .

5. CONCLUSIONES

La presente tesis tuvo como objetivo demostrar la importancia de las visualizaciones de datos, para la representación de la información, y cómo, es que estas, contribuyen al fortalecimiento de la gestión organizacional en una empresa.

Para demostrar esto, en la aplicación del caso de estudio, se realizaron dos iteraciones donde como primer punto pudimos comprobar cómo TOGAF facilita la definición del contenido, debido a que provee fases para el diseño, planificación e implementación de una arquitectura empresarial, comprobando la importancia cada una de las fases que conforman esta arquitectura, para lograr los objetivos deseados.

En el desarrollo de este caso de Estudio, en los dos ciclos realizados (uno para la limpieza de datos y otro para la construcción de la visualización), se realizaron bajo tres iteraciones, las cuales son:

- Iteración de Arquitectura de Contexto
- Iteración de Arquitectura de Definición
- Iteración de Planeación

Para la iteración de contexto la cual abarca la fase preliminar y la fase de visión de arquitectura, se confirmó la importancia de conocer la situación de la empresa o institución sobre la cual se implementará todo el desarrollo, conocer bien los objetivos y la visión que se debe perseguir, de forma que todos los trabajos realizados estén orientados a su cumplimiento.

Como se observó en el caso de estudio, en ocasiones la institución no conoce el estado de madurez en el cual se encuentra, por lo que es necesario realizar una serie de ejercicios, para conocer el estado de la institución en el ámbito de arquitectura, gobierno, administración, identificación, integración y administración de negocio, de forma que conociendo las capacidades que se tienen se pueda realizar un plan de trabajo más adecuado.

De la misma forma para la iteración de definición la cual abarca las fases de Arquitectura de Negocio, Arquitectura de sistemas de información y Arquitectura Tecnológica, a lo largo de esta comprobamos la importancia de contar con los procesos bien definidos, donde conozcamos cuáles son las funciones, servicios y roles que se tendrán a lo largo de este pues esto nos ayudará a tener un panorama más claro de las capacidades del personal así como el papel que tendrá a lo largo del desarrollo del sistema; asimismo fue la iteración de definición donde se desarrolló el sistema en cada

uno de sus componentes (conceptual, lógico y físico), fue aquí donde llevamos los requerimientos de negocio a un desarrollo más técnico.

Para la iteración de Planeación que considera las fases de Oportunidades y Soluciones, así como la de Planeación y Migración se comprobó que el seguir una estrategia para la implementación de la arquitectura ayuda a no perder el principal objetivo y desgastarse en otras tareas que no son primordiales, asimismo esta nos ayuda a tener una mejor definición de los trabajos y estimación de costos y tiempos.

Se observó que la mayoría del personal, tanto de negocio como de TI, de la institución no tiene conocimiento de la importancia que presentan los datos y la ventajas que se pueden obtener al presentarlos a los interesados de una forma sencilla; los beneficios de realizar esto se ve reflejado en la mejora de la gestión empresarial pues estos contribuyen a la mejor toma de decisiones y permiten llevar una mejor planeación, organización, dirección y control de la institución con la intención de cumplir todos los objetivos y visión de ésta. Realizar la implementación de visualizaciones, traerán grandes ventajas a la institución, siendo en la gestión empresarial de los puntos más importantes, pues permite generar las condiciones para la existencia, supervivencia y prosperidad de la empresa.

Como conclusión vimos cómo la visualización generada, en la cual se presentaba una serie de sucesos-eventos, ayudó a automatizar el levantamiento de reportes, pues el mismo usuario de negocio, puede realizarlos, sin la necesidad de involucrar a un área intermedia de sistemas, mejorando la satisfacción de los usuarios al disminuir el tiempo de respuesta, a los diversos incidentes que se presentaban, así como tener la información necesaria a la mano, presentada de una forma sencilla y dinámica.

6. BIBLIOGRAFÍA

Oracle, 2008, Gestión de Procesos de Negocio, Arquitectura Orientada a Servicios y Web 2.0: ¿Transformación de Negocios o Problemática Global?
<<http://www.oracle.com/technetwork/es/middleware/fusion-middleware/documentation/gestion-proceso-negocio-soa-web-450487-esa.pdf>>

Panduit, 2012, Arquitectura de Referencia: La construcción de bloques de una base ágil, - exible, escalable y sustentable
<<http://www.panduit.com/ccurl/624/431/CPAT08--WW-SPA-ArquitecturaReferencias-WEB,0.pdf>>

PowerData, 2013, La Limpieza de Datos: la etapa previa a los Procesos ETL
<<https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/312597/la-limpieza-de-datos-la-etapa-previa-a-los-procesos-etl>>

Gobierno de España, 2016, Visualización de datos: definición, tecnologías y herramientas
<http://datos.gob.es/sites/default/files/doc/file/informe_herramientas_visualizacion.pdf>

Loshin, 2009, Master Data Management, U.S.A., Morgan Kaufmann, pág. 43-65

Bruno Munari, 2016, Diseño y Comunicación Visual: Contribución a una Metodología Didáctica, España, Gustavo Gili

Ignasi Alcalde, 2015, Visualización de la información de los datos al conocimiento, España, GUSTAVO GILI

Weske Mathias, 2012, Business Process Management: Concepts Languages, USA, Springer-Verlag

The Open Group, 2013, TOGAF Version 9.1, USA, Van Haren Publishing