



FACULTAD DE POSGRADOS

FORMULACIÓN DE UN MARCO METODOLÓGICO PARA EL DESARROLLO
DE SOLUCIONES DE INTELIGENCIA DE NEGOCIOS, EMPLEANDO
METODOLOGÍAS AGILES.
CASO: ÁREA DE DATAWAREHOUSE DEL SERVICIO DE RENTAS
INTERNAS

Trabajo de Titulación presentado en conformidad con los requisitos
establecidos para optar por el título de Magister en Gerencia de Sistemas y
Tecnologías de la Información.

Profesor Guía
MSc. Freddy Mauricio Tapia León

Autor
Juan Fernando Analuisa Barona

Año
2016

DECLARACIÓN DEL PROFESOR GUÍA

“Declaro haber dirigido este trabajo a través de reuniones periódicas con el estudiante Juan Fernando Analuisa Barona, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación”.

Freddy Mauricio Tapia León

Master universitario en investigación e innovación en tecnologías de la
información y las comunicaciones

CI: 1714745690

DECLARACIÓN DE AUDITORIA DEL ESTUDIANTE

“Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes”

Juan Fernando Analuisa Barona

CI: 1711694180

AGRADECIMIENTOS

A mi esposa por su apoyo y ánimo brindado para seguir esta maestría, y sobre todo por ser esa persona tan especial en mi vida.

A mis hijos Sofía y Juan Sebastián por su paciencia durante este tiempo de estudios.

A mis padres y hermanos por ser el soporte en esos momentos difíciles y ser parte de este logro.

A la UDLA por abrirme las puertas a tan prestigiosa universidad.

A los profesores de la UDLA que me brindaron su conocimiento y su experiencia durante las horas académicas.

Al SRI por darme el apoyo para realizar este trabajo de titulación.

DEDICATORIA

Uno entiende ese amor tan inmenso que los padres tienen a sus hijos, cuando se es padre. Los hijos son esas personitas que nos dan el valor y las fuerzas para cumplir todos nuestros objetivos en bienestar de ellos. Es por eso que este trabaja va para ustedes, Sofía y Juan Sebastián que son la razón de mi vivir.

Quiero dedicar este trabajo también a mi esposa Kary por la paciencia y el cariño que me ofreció, a mi papito Juani, a mis hermanos Napo, Sandy y Pachi, a todos mis sobrinos, y especialmente a mis dos mamis Glorita y Eloysita.

RESUMEN

Las unidades de inteligencia de negocios son poco comunes y solo se las halla en grandes organizaciones como un área de TI y en algunos casos son parte de la unidad de planificación a lado del negocio. En el caso del SRI existe el área de Datawarehouse el cual inicialmente fue creado como área de Business Intelligence desde el año 2008 como parte del departamento de Desarrollo de la Dirección Nacional de Tecnología.

Al ser un área relativamente nueva se empezaron a establecer los métodos de trabajo y los estándares de desarrollo. Así también se inició la adquisición de los elementos que conforman la arquitectura actual de Data Warehouse - DWH, sin embargo no se cuenta con un procedimiento estandarizado y formalizado que establezca los pasos a seguir en la construcción de soluciones de Data Warehouse.

Adicionalmente se ha visto la necesidad de aplicar prácticas ágiles en el desarrollo de las soluciones de Data Warehouse que permitan entregar productos de manera rápida, oportuna y eficiente al negocio para la toma de decisiones adecuada en función de cumplir los objetivos estratégicos de la institución.

Este trabajo de tesis tiene como finalidad analizar la situación actual del área de Datawarehouse del SRI, entender las metodologías tradicionales de construcción de soluciones de Data Warehouse, entender los marcos de trabajo ágiles y desarrollar un marco metodológico cuyo entregable sea un procedimiento de desarrollo de productos de Data Warehouse. El mismo deberá incluir las mejores prácticas de metodologías tradicionales de desarrollo de soluciones de Data Warehouse, así como el uso de las prácticas de marcos de trabajo ágiles, permitiendo entregar un procedimiento al área de Datawarehouse que sirva de guía para la construcción de los productos a desarrollar de acuerdo a las necesidades de los usuarios de la institución.

ABSTRACT

The Business Intelligence units are rare and only found in the high organizations as an IT area and in some cases are part of the planning unit next to the business. In the case of SRI Data Warehouse area exists which initially was created as Business Intelligence area since 2008 as part of the Department of Development of the National Directorate of Technology.

Being a relatively new area they began to establish the working methods and standards development. Thus the acquisition of the elements of the current architecture of Data Warehouse was also launched - DWH, however there is not a standardized and formalized setting out the next steps in building Data Warehouse solutions procedure.

Additionally it has seen the need to implement agile practices in developing solutions that enable Data Warehouse deliver products fast, timely and efficient manner to the business for making appropriate decisions based on fulfilling the strategic objectives of the institution.

This thesis aims to analyze the current situation in the area of Datawarehouse SRI understand traditional construction methodologies solutions Data Warehouse, understand frameworks agile work and develop a methodological framework whose deliverable is a process of development of Data Warehouse products. It should include the best practices of traditional development methodologies solutions Data Warehouse apply to the practices of frames agile work, allowing deliver a procedure to the area of data warehouse as a guide for building products developed according to the needs of business users.

ÍNDICE

INTRODUCCIÓN	1
1. CAPITULO I DESCRIPCIÓN Y OBJETIVOS	2
1.1 Antecedentes del SRI.....	2
1.2 Justificación de la investigación.....	5
1.3 Objetivos de la investigación	7
1.3.1 Objetivo general.....	7
1.3.2 Objetivos específicos	8
1.4 Alcance de la investigación.....	8
1.5 Metodología de investigación	9
2. CAPITULO II MARCO TEÓRICO	11
2.1 Inteligencia de Negocios	11
2.2 Data Warehouse.....	13
2.3 Definición de Marco Metodológico	18
2.4 Metodologías de desarrollo Data Warehouse	20
2.4.1 Metodología de Bill Inmon.....	22
2.4.1.1Ciclo de Vida de Inmon	23
2.4.1.2 Estructura de un DWH	24
2.4.1.3 Proceso de construcción de un DWH.....	25
2.4.1.4 Granularidad	26
2.4.1.5 Niveles de la Arquitectura de DWH.....	27
2.4.1.6 Usuarios del DWH.....	28
2.4.1.7 Modelo de Datos del DWH.....	28
2.4.1.8 Varios conceptos adicionales relevantes	29
2.4.2 Metodología de Ralph Kimball	30
2.4.3.1 Esquema Estrella vs Cubo OLAP	32
2.4.3.2 Tabla de Hechos y Medidas.....	33
2.4.3.3 Tabla de Dimensiones.....	35
2.4.3.4 Hechos y Dimensiones en un esquema Estrella	36
2.4.3.5 Arquitectura de DWH/BI de Kimball	37

2.4.3.6 Ciclo de vida de Kimball	39
2.5 Marcos de Trabajo Ágiles.....	41
2.5.1 Scrum.....	42
2.5.2 Kanban.....	47
3. CAPITULO III SITUACIÓN ACTUAL DEL ÁREA DE	
DATAWAREHOUSE	51
3.1 Introducción	51
3.2 Arquitectura de Data Warehouse DWH	52
3.3 Herramientas.....	56
3.3.1 Herramientas de base de datos	56
3.3.1.1 Base de datos DWH.....	56
3.3.1.2 Base de datos STAR.....	57
3.3.1.3 Base de datos Frontera.....	58
3.3.2 Herramientas de ETL.....	58
3.3.3 Herramientas de Publicación y Explotación de Información	59
3.4 Roles	61
3.5 Atención de Requerimientos del Negocio	62
3.6 Inconvenientes proceso actual Desarrollo de	
Productos DWH.....	71
3.7 Encuesta situación actual.....	72
4. CAPITULO IV PROPUESTA METODOLÓGICA	76
4.1 Introducción	76
4.2 Estructura del Marco Metodológico.....	76
4.3 Selección de la Metodología de DWH.....	76
4.4 Prácticas de Metodologías Ágiles a considerar	78
4.5 Procedimiento de Desarrollo de Productos de DWH	79
4.5.1 Identificación del Proceso	80
4.5.2 Objetivo.....	81
4.5.3 Alcance	81
4.5.4 Normas de Operación	82

4.5.4.1 De la definición de Requerimientos del Negocio	83
4.5.4.2 Del Diseño de la Arquitectura Técnica y Selección de Productos.	87
4.5.4.3 Del Modelo Dimensional	88
4.5.4.4 Del Diseño Físico	89
4.5.4.5 Del Diseño y Construcción del ETL.....	91
4.5.4.6 De las Especificaciones y Desarrollo de Aplicaciones de BI	93
4.5.4.7 De la Implementación de la Solución	94
4.5.5 Perfiles y Niveles de Responsabilidad	96
4.5.6 Flujo del Proceso	99
4.5.7 Entregables del Proceso	100
4.6 Sustentación del Procedimiento	101
5. CAPITULO V CONCLUSIONES Y RECOMENDACIONES.....	106
5.1 Conclusiones	106
5.2 Recomendaciones	107
6. REFERENCIAS	109
7. ANEXOS	111

ÍNDICE DE FIGURAS

Figura 1. Reestructura Organizacional	3
Figura 2. Pirámide Organizacional	12
Figura 3. Elementos de un Data Warehouse	17
Figura 4. Elementos de una Metodología de Desarrollo de Software	19
Figura 5. Ciclo de vida de desarrollo tradicional vs ciclo de vida DWH	24
Figura 6. Estructura del DWH.....	25
Figura 7. Niveles de la Arquitectura	27
Figura 8. Esquema Estrella vs cubo OLAP	33
Figura 9. Tablas de hechos y dimensiones en un modelo dimensional	36
Figura 10. Elementos de la Arquitectura de DWH y BI de Kimball	37
Figura 11. Ciclo de Vida de Kimball	39
Figura 12. Fases de SCRUM	46
Figura 13. Flujo Kanban	49
Figura 14. Arquitectura DWH del SRI.....	54
Figura 15. Servidores que intervienen en la arquitectura de DWH.....	61
Figura 16. Incidentes atendidos vs tiempos de atención establecidos	66
Figura 17. Requerimientos Atendidos 2014 - 2015	68
Figura 18. Cambios en producción 2014 - 2015.....	70
Figura 19. Estructura Procedimiento de Desarrollo de Productos DWH	80
Figura 20. Flujo del Proceso Desarrollo DWH.....	99
Figura 21. Encuesta situación futura	103

ÍNDICE DE TABLAS

Tabla 1. Datos Estructurados	14
Tabla 2. Estadísticas incidentes cerrados últimos 2 años	65
Tabla 3. Tiempos resolución de incidentes	66
Tabla 4. Número requerimientos atendidos años 2014 - 2015.....	67
Tabla 5. Número cambios atendidos años 2014 - 2015	69
Tabla 6. Identificación Proceso DWH.....	80
Tabla 7. Perfiles y Niveles de Responsabilidad.....	96
Tabla 8. Entregables Proceso de Desarrollo de Productos de DWH	100
Tabla 9. Evaluación beneficios procedimiento propuesto situación actual.....	101
Tabla 10. Universo de encuestados Situación Futura	103
Tabla 11. Evaluación beneficios procedimiento propuesto situación futura ...	104

INTRODUCCIÓN

El Servicio de Rentas Internas (SRI) se ha convertido en una de las principales entidades del Estado que se ha destacado por desarrollar aplicaciones tecnológicas innovadoras y de alta calidad, adoptando para ello tecnologías de punta y metodologías de desarrollo, que permitan generar productos que cumplan con los requerimientos de las distintas áreas de la Institución, los cuales se han enfocado en resolver las necesidades cambiantes del negocio en aspectos tributarios mediante la mejora continua a los procesos de la organización.

Debido a los cambios acelerados en el entorno global, se observa un crecimiento en los requerimientos tecnológicos, a los cuales se debe dar frente de una manera oportuna y eficaz de tal forma que genere resultados con valor agregado para la organización.

Es por tal motivo que se propone el desarrollo de un marco metodológico que agilice el desarrollo y construcción de soluciones de Inteligencia de Negocios enfocado al Área de Datawarehouse del SRI, permitiendo alinear las necesidades del negocio a los objetivos estratégicos de la institución.

1. CAPITULO I DESCRIPCIÓN Y OBJETIVOS

1.1 Antecedentes del SRI

De acuerdo a la página institucional del SRI (2016):

“El Servicio de Rentas Internas (SRI) es una entidad técnica y autónoma que tiene la responsabilidad de recaudar los tributos internos establecidos por Ley, mediante la aplicación de la normativa vigente. Su finalidad es la de consolidar la cultura tributaria en el país a efectos de incrementar sostenidamente el cumplimiento voluntario de las obligaciones tributarias por parte de los contribuyentes.”

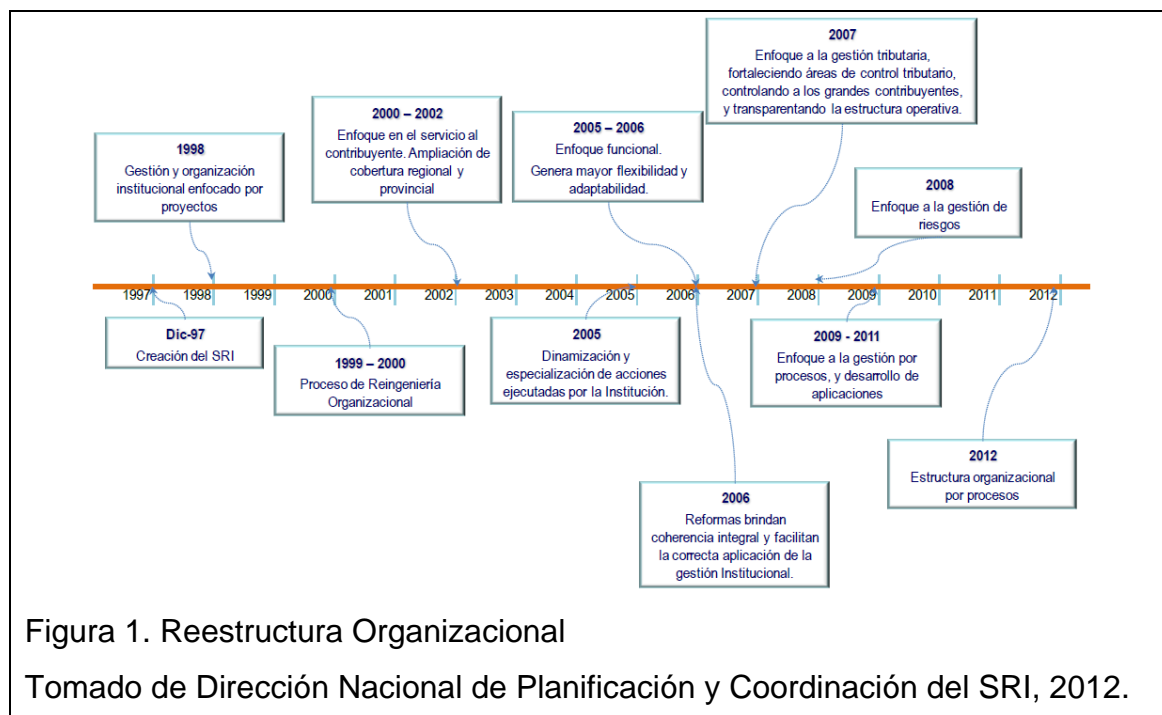
Los objetivos institucionales del SRI son:

- Incrementar anualmente la recaudación de impuestos con relación al crecimiento de la economía.
- Diseñar propuestas de política tributaria orientadas a obtener mayor equidad, fortalecer la capacidad de gestión institucional y reducir el fraude fiscal.
- Lograr altos niveles de satisfacción en los servicios que presta al contribuyente.
- Reducir los índices de evasión tributaria y procurar la disminución en la aplicación de mecanismos de elusión de impuestos.

El SRI ha consolidado su estructura organizacional en áreas funcionales, siendo una de estas, la encargada del establecimiento de políticas a nivel nacional; otras a nivel regional se encuentran alineadas a la operación, dando énfasis a la implementación de servicios y transacciones internas y externas que faciliten el cumplimiento de las obligaciones fiscales por parte del contribuyente.

En la figura 1, se puede observar la reestructuración organizacional que la institución ha ejecutado en los últimos 12 años. Tal es así, que para el año 2012

la administración del SRI impulsó la formulación del Plan Estratégico 2012 -2015, bajo los lineamientos del Plan Nacional para el Buen Vivir, siendo el aspecto tecnológico uno de los factores estratégicos en la gestión del SRI.



Adicionalmente se puede destacar que en el año 2008, mediante resolución No.

DSRI-028-2008, se resuelve reformar el Reglamento Orgánico Funcional del SRI señalándose en el artículo 1, literal j, párrafo 4.2, la nueva estructura vigente a esa fecha, en la cual se incluye el Área de Business Intelligence (BI) y Soluciones Regionales, como parte del Departamento de Desarrollo e Implementación de la Dirección Nacional de Desarrollo Tecnológico cuya principal función consiste en “Planificar, coordinar e implementar soluciones que permitan la explotación de la información que requiera el negocio para el análisis y toma de decisiones” (Servicio de Rentas Internas, 2008).

El Área en mención ejecuta tareas de atención a los requerimientos de información operativa y analítica de las distintas áreas del SRI, mediante la construcción de modelos de información multidimensionales, así como la atención de convenios de intercambio con instituciones externas como por

ejemplo: Instituto Ecuatoriano de Seguridad Social (IESS), Dirección Nacional del Registro Civil, Servicio Nacional de Aduana del Ecuador (SENAE), las cuales replican su información hacia el SRI, lugar que centraliza las Bases de Datos (BD) replicadas, para su uso posterior en procesos de control en la Dirección Nacional del SRI.

Para ejecutar las tareas anteriormente descritas, el Área de BI dispone de las siguientes herramientas de software:

- Repositorio de Data Warehouse en una instancia de Base de Datos Oracle.
- Oracle Warehouse Builder para construcción de ETL.
- Oracle Discoverer para la publicación de la información.

Como parte de la reestructuración orgánica efectuada, una de las acciones fue adoptar una metodología de desarrollo de software - MSF (Microsoft Solutions Framework), la cual no se llegó a implementar en su totalidad en el Área de BI, y al poco tiempo fue desechada por el Departamento de Desarrollo Tecnológico debido a su complejidad y gran cantidad de documentos que esta requería. Posteriormente, se contrató una consultoría para la reestructuración de la arquitectura del Data Warehouse Institucional cuyo objetivo fue "...mostrar el diseño de una arquitectura de BI particularizada y acoplada a las necesidades de la organización en todos los posibles escenarios de manejo de información" (Innovasys, 2010).

Entre los resultados que presentó la consultoría, y específicamente en el Documento de Diseño de Arquitectura de *Business Intelligence* (BI) del SRI (Innovasys, 2010), se refieren entre sus conclusiones a una arquitectura "IDEAL" de BI, la cual no contrasta con las realidades de las ofertas en el mercado debido a su complejidad, lo que obliga a estudiar cada uno de los elementos que conforman la arquitectura de BI para la implementación de este diseño. Otra conclusión señala que no existen herramientas adecuadas de BI como son herramientas ETL y de publicación, por lo que se adquirió la Suite de BI de la

casa comercial SAP, y la herramienta de minería de datos SPSS de la casa comercial de IBM. Sin embargo pese la adquisición de las mencionadas herramientas no se logró migrar los modelos multidimensionales que se tenía hasta ese momento.

Actualmente, el SRI ha adoptado una estructura organizacional por procesos donde el Área de BI pasa a nombrarse como Área de Datawarehouse, pero mantiene los mismos desafíos en la construcción de soluciones de información en beneficio del SRI.

1.2 Justificación de la investigación

Dentro de la Planificación Gubernamental ejecutada por medio del Plan Nacional del Buen Vivir, se ha establecido el uso de metodologías y herramientas que permitan planear, monitorear y gestionar objetivos, indicadores, procesos y proyectos denominado Gobierno por Resultados (GPR), por lo cual el SRI, entre sus objetivos institucionales ha desarrollado un plan estratégico el cual especifica las políticas y acciones para lograr alcanzar los objetivos establecidos a nivel nacional.

Bajo este contexto, el SRI ha ejecutado numerosos proyectos que han permitido construir aplicaciones tecnológicas para automatizar los procesos del negocio que permitan apoyar a la consecución de los objetivos estratégicos de la institución. Dichas aplicaciones han registrado gran cantidad de datos lo que ha originado que las BD crezcan constantemente así como las necesidades de información.

Los usuarios de la Institución (usuarios finales), quienes realizan análisis de información para la toma de decisiones, no tienen claro sus necesidades de información analítica y estratégica, y en un intento por cubrir estas necesidades, solicitaban al Área de BI, actualmente Área de Datawarehouse, la publicación de fuentes de datos transaccionales, para posteriormente poder determinar sus

requerimientos de información. Los requerimientos se vuelven extensos al igual que los tiempos de desarrollo, lo que ocasiona cambios en el requerimiento debido a funcionalidades o necesidades inicialmente no consideradas incrementando los tiempos del proyecto.

Para los usuarios de la Institución se vuelve complejo entender y transmitir sus necesidades de información, lo que ocasiona un mayor tiempo en la elaboración del documento de requerimientos de información, y como consecuencia no se desarrolla un producto visible a corto plazo hacia los interesados del proyecto.

Entre los problemas que se evidencia en el Área de Datawarehouse al momento de desarrollar soluciones de BI están:

- No existe un procedimiento adecuado para el levantamiento de necesidades de información, así como el uso de formatos estandarizados que faciliten el levantamiento del requerimiento.
- No existe un procedimiento definido para el desarrollo de soluciones BI en el cual se especifique las fases de: requerimientos, análisis, diseño, construcción, pruebas, certificación y puesta en producción de las mismas.
- Los tiempos de implementación de los productos de BI son altos por lo que los usuarios ven un resultado al final del mismo cuando el interés o la necesidad de la información ya no es requerida, es decir la información no está disponible de manera oportuna.
- La arquitectura correspondiente a la infraestructura técnica de BI no está levantada, lo que ocasiona que los desarrolladores de soluciones de información no conozcan elementos técnicos sobre los cuales se deben construir dichas soluciones.
- No se maneja un modelo de gestión del conocimiento con respecto a la documentación que se genera de los proyectos o requerimientos atendidos.

- Las fases de desarrollo de algunos de los proyectos de BI se cumplen parcialmente.
- Los estándares para la construcción de objetos o elementos de base de datos no se aplican de la manera más adecuada y tampoco se controla la utilización de los mismos.
- La construcción como tal no está sujeta a un proceso de control de calidad adecuado.

En las condiciones descritas se vuelve complicado determinar la capacidad operativa del Área de Datawarehouse, debido a que no existen parámetros de medición, ni un control adecuado sobre el seguimiento de los requerimientos que son atendidos.

Esta Área atiende en su mayoría requerimientos que se enfocan a publicación de datos transaccionales, dejando escapar la posibilidad de atacar aquellos requerimientos de información como por ejemplo construcción de modelos de información multidimensionales que permitan solventar los indicadores de los procesos estratégicos de la institución.

En el capítulo tres de este trabajo se realiza un levantamiento de la situación actual, en el cual se sustentan los puntos anteriores que corresponden a la justificación de esta tesis.

1.3 Objetivos de la investigación

1.3.1 Objetivo general

Formular una metodología ágil para optimizar el desarrollo de soluciones de inteligencia de negocios mediante el uso adecuado de herramientas, técnicas y procesos que permita atender las necesidades de información analítica y estratégica para la toma adecuada de decisiones de los niveles directivos y operativos del SRI.

1.3.2 Objetivos específicos

- Analizar la situación actual del Área de Datawarehouse, a través del cual se identifiquen los procesos orientados a la atención de necesidades de información del SRI, elementos de arquitectura utilizados y las herramientas tecnológicas que apoyan a la ejecución de las tareas del área.
- Identificar y analizar las principales características de las metodologías más importantes de desarrollo de soluciones de inteligencia de negocios y las metodologías y marcos de trabajo ágiles de desarrollo de software que puedan ser aplicadas al Área de Datawarehouse.
- Proponer una metodología ágil alineada al contexto del Área de Datawarehouse en base a las principales características de las metodologías de desarrollo de soluciones de inteligencia de negocios y ágiles de desarrollo de software identificadas.
- Identificar opciones de mejora al proceso actual de atención de necesidades de información del SRI en base a la metodología ágil propuesta.

1.4 Alcance de la investigación

El presente trabajo de tesis incluye la identificación de las principales características de las metodologías tradicionales de Data Warehouse y marcos de trabajo ágiles, el levantamiento de la situación actual del proceso de desarrollo de soluciones de información aplicando técnicas de encuestas e investigación descriptiva, y la elaboración de un procedimiento en base a las fases de desarrollo de soluciones de Data Warehouse aplicando las características de los marcos de trabajo ágiles.

Esta fuera del alcance el implementar la propuesta metodológica que se genere de este trabajo de tesis en el Área de Datawarehouse del SRI.

1.5 Metodología de investigación

El presente trabajo de tesis propone una serie de pasos para el mejoramiento del proceso de desarrollo de soluciones de inteligencia de negocios de una manera ágil asegurando una atención adecuada a las necesidades de información analítica y estratégica del SRI.

Para lo cual se realizará un estudio descriptivo de los procesos actuales de desarrollo de soluciones de inteligencia de negocios mismos que permitan medir y evaluar la atención de requerimientos de información, elementos de arquitectura utilizados y herramientas tecnológicas que apoyan la ejecución de las tareas del Área de Datawarehouse.

Como parte inicial de este trabajo de tesis se describe la estructura que debe contener un marco metodológico, así como también los principales conceptos que son nombrados en la teoría de Data Warehouse y de Inteligencia de Negocios. También se describen las características de las metodologías de desarrollo de Data Warehouse de Bill Inmon y Ralph Kimbal, así como las principales características de los marcos ágiles: Scrum y Kanban.

Para poder sustentar la justificación del presente trabajo, se realiza un levantamiento de la situación actual del Área de Datawarehouse, para lo cual, mediante investigación descriptiva, se identifican y describen la arquitectura actual de Data Warehouse del SRI, las herramientas tecnológicas utilizadas y los roles que intervienen en las fases de desarrollo, así como un levantamiento de los principales inconvenientes detectados en el procedimiento actual de desarrollo de soluciones de información. También se obtienen estadísticas referentes a la atención de requerimientos de información, y se realizan encuestas que permiten identificar los inconvenientes al actual procedimiento y sustentar este trabajo de titulación.

Posteriormente se plantea el procedimiento de desarrollo de productos de Data Warehouse el cual se base en las fases del ciclo de vida propuesto en la metodología de Ralph Kimball, así como también en las prácticas ágiles propuestas por Scrum.

Finalmente se establecen las conclusiones y recomendaciones en cuanto a la propuesta definida en este trabajo de titulación.

2. CAPITULO II MARCO TEÓRICO

2.1 Inteligencia de Negocios

La Inteligencia de Negocios pretende describir un conjunto de conceptos y métodos que buscan mejorar el proceso de toma de decisiones, para lo cual se apoya en aplicaciones tecnológicas orientadas a integrar datos para generar información, la cual sea accesible a las personas adecuadas, ejecutoras de decisiones, en un tiempo oportuno, que mediante análisis, permitirán convertir esa información en conocimiento.

Un ejemplo de este tipo de aplicaciones se puede observar en cadenas de supermercados con el objetivo de identificar el perfil de sus clientes más rentables para aplicar políticas de fidelización, para lo cual utilizan los datos de las ventas para extraer la información relevante que permita analizar y determinar los clientes con mayores compras. Esta información es analizada desde varias perspectivas del negocio.

Otro de los objetivos que busca la Inteligencia de Negocios es analizar la información de diferentes perspectivas o aristas de análisis de las organizaciones, de tal forma que a más de responder al estado actual de las empresas, también pueda encontrar patrones o tendencias que permitan predecir escenarios futuros para una mejor toma de decisiones.

Para lo cual la Inteligencia de Negocios se enfoca a los tres principales niveles de las organizaciones: operacional, táctico o analítica y estratégico. En la siguiente figura se esquematiza lo mencionado:



En el nivel operacional la información es utilizada por el personal operativo de la organización, el cual ejecuta las tareas de los procesos de la cadena de valor de las empresas.

En el nivel táctico o analítico la información es utilizada por los supervisores o directores operativos para tomar decisiones respecto a las actividades operativas en mejora de los procesos de las empresas.

En cuanto al nivel estratégico la información es utilizada por los altos directivos para la toma de decisiones gerenciales, permitiendo siempre alinear a la consecución de los objetivos gerenciales.

Entre los beneficios que proporciona la Inteligencia de Negocios a las organizaciones están: (Cano, 2007)

- Beneficios tangibles como: aumento de ingresos, disminución de costos, mejora de los procesos internos mediante el análisis de los mismos.
- Beneficios intangibles como: información disponible a los usuarios claves para la toma de decisiones para mejorar la posición competitiva de las empresas.

- Beneficios estratégicos ya que apoya a la ejecución de los objetivos estratégicos de las organizaciones mediante la creación de indicadores claves.

Como conclusión se puede mencionar que: Inteligencia de Negocios consolida los datos de la organización en un solo repositorio central denominado bodega de datos, comúnmente conocida como Data Warehouse, de una manera organizada por temas específicos, que permita realizar análisis de tiempo real que genere información para la toma de decisiones adecuadas en los 3 niveles de la organización.

2.2 Data Warehouse

De acuerdo a lo descrito en el literal 2.1, la Inteligencia de Negocios apoya a la toma de decisiones, para lo cual se requiere que la información se encuentre disponible y organizada en un solo lugar denominado Data Warehouse (DWH), o su traducción al lenguaje español como bodega o almacén de datos, el cual es un recurso esencial al momento de realizar el proceso de Inteligencia de Negocios.

Se puede definir a un DWH en los siguientes puntos:

- Es una base que agrupa todos los datos que una organización utiliza como insumo para la toma de decisiones.
- Los datos son almacenados de una manera ordenada e integrada.
- Almacena grandes volúmenes de datos por lo que maneja información histórica.
- Se ejecutan procesos de limpieza y análisis de calidad sobre los datos procesados.
- Los usuarios de la organización acceden fácilmente hacia el DWH de forma rápida y eficiente.

- Apoya a los Sistemas de Soporte de Decisiones (DSS) y Sistemas de Información para Ejecutivos (EIS).
- Almacena los metadatos, es decir la información específica de los datos almacenados.

Los datos son extraídos de diferentes fuentes de información como: sistemas transaccionales, archivos de Excel, correo electrónico, sistemas documentales, etc., implementadas en las organizaciones.

En una organización se pueden identificar varios tipos de datos a los cuales se los puede clasificar en datos estructurados, no estructurados y semiestructurados.

Los datos estructurados son aquellos que son definidos en base a sus atributos, los cuales poseen tipo de datos y son almacenados en BD relacionales, como por ejemplo: SQL Server, Oracle, MS Access. En la siguiente tabla se muestra un ejemplo de datos estructurados:

Tabla 1. Datos Estructurados

Numero de RUC	Razón Social	Fecha Inscripción	Ingresos
Texto de 13 caracteres.	Texto de 500 caracteres	Fecha en formato dd/mm/yyyy	Entero de 2 decimales
0961117778001	Empresa XYZ	15/05/2014	1800.30

Interpretando la tabla 1 se concluye que “La Empresa XYZ con número de RUC 0961117778001 fue inscrita el 15 de mayo del 2014 y reportó ingresos por \$1800.30”.

Los datos no estructurados no tienen atributos ni definición de tipos de datos, y por lo general corresponden a narraciones textuales sin ninguna estructura, almacenados en archivos de diferentes formatos como por ejemplo: HTML, DOC, Correos Electrónicos, entre otros.

Un dato semiestructurado es la combinación entre un dato estructurado y no estructurado. Para lo cual manejan una estructura tipo árbol donde en cada nodo del árbol se encuentra un texto que corresponde al dato como tal. Ejemplo de este tipo de datos son los archivos XML.

Un DWH debe tener la capacidad de integrar y almacenar de una forma homogénea los tres diferentes tipos de datos descritos anteriormente, para lo cual existen herramientas especializadas en realizar estas tareas.

El proceso encargado de alimentar de datos al DWH se denomina ETL y tiene las siguientes principales tareas:

- Conectarse a las diferentes fuentes de datos heterogéneas y multiplataforma.
- Extracción de los datos en donde se realizan procesos de calidad y limpieza de los datos.
- Transformación de los datos donde se realizan cálculos complejos y operaciones agregadas.
- Almacenamiento de la data transformada en estructuras físicas de la base de datos del DWH que soportan cargas masivas de datos así como lecturas rápidas.

Al proceso de ETL, cuyas siglas significan proceso de Extracción, Transformación y Carga, también se le conoce como el proceso de *Datawarehousing*.

Los modelos de datos que se utilizan en la construcción de un DWH se los denomina modelos multidimensionales. Existen dos modelos multidimensionales: modelos estrella y copo de nieve. Sin embargo de acuerdo a la metodología de construcción de DWH no es siempre mandatorio la construcción de modelos multidimensionales ya que se puede construir un DWH a partir de base de datos transaccionales. Este tema se tratará a detalle en el literal 2.3 de Metodologías de Desarrollo de Data Warehouse.

Hay que señalar que a diferencia del modelo relacional donde se tiende a normalizar las estructuras de datos, en un DWH se tiende a desnormalizar las estructuras de datos en tablas planas o estructuras que relacionan varias tablas de un mismo tema en una sola estructura.

Otros de los elementos que se consideran en un DWH mencionados por varios autores como Inmon (2002), Kimball y Ross (2013) son:

- Área de *Staging*, o área temporal donde los procesos ETL realizan tareas de calidad de datos que son únicamente requeridos para ese proceso en específico por lo que luego son eliminados.
- Área de ODS, *Operational Data Store* o almacenamiento de Datos Operacionales, donde se trasladan los datos transaccionales hacia el DWH para realizar ahí las transformaciones y cálculos complejos, sin afectar el desempeño de las aplicaciones transaccionales.
- Área de modelos multidimensionales, que es donde residen las estructuras, modelos estrella o copo de nieve. A este elemento también se le conoce como *data mart*.
- Cubos, los cuales son construidos a partir de los *data marts*.
- Metadatos corresponde a la información de los datos desde su origen, transformación y explotación.
- Publicación, que se refiere al uso de los cubos y *data marts* construidos en los sistemas de toma de decisión, tableros de control, herramientas de minería de datos, etc.

De acuerdo a los elementos especificados en los puntos anteriores, se resumen en la siguiente figura los elementos principales que intervienen en un DWH:

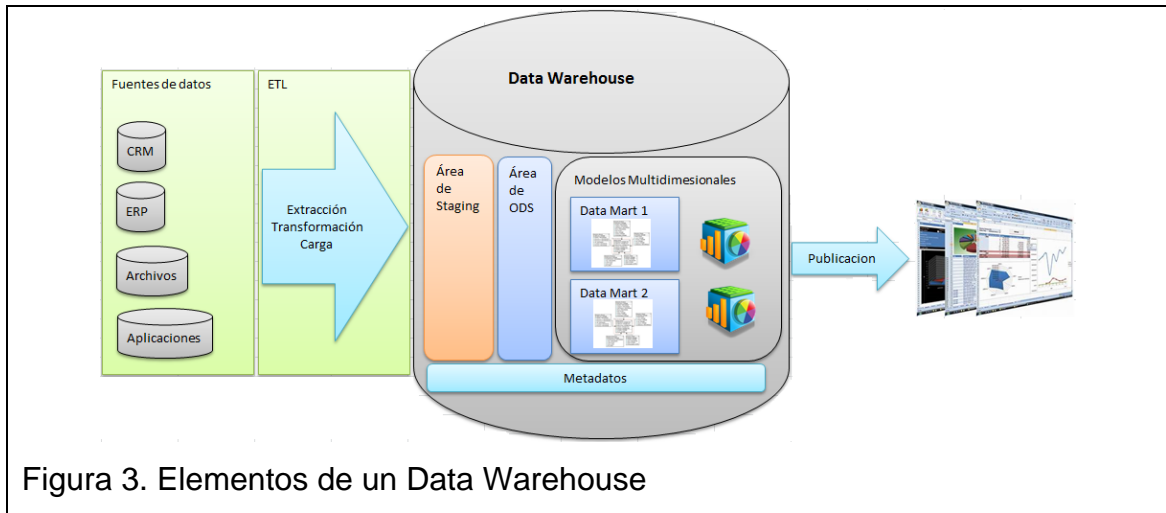


Figura 3. Elementos de un Data Warehouse

Como se observa en la figura, existen 4 elementos principales:

1. Fuentes de datos que pueden ser generados por las aplicaciones transaccionales, CRM, ERP o archivos de varios tipos.
2. El ETL encargado de extraer los datos de las fuentes, integrarlas, transformarlas y cargarlas en el DWH.
3. El DWH como tal, el cual está conformado de tres principales áreas: Área de *Staging* donde se almacena datos temporales para realizar tareas de calidad y procesamiento de datos. Área de ODS donde se almacenan estructuras con datos operacionales con cierto nivel de procesamiento. Y el área de modelos multidimensionales donde se almacenan físicamente los modelos estrellas.
4. La publicación de ODS y modelos multidimensionales en herramientas de explotación y análisis de información que están a disposición de los usuarios finales.

El elemento fundamental en una arquitectura de DWH es el componente de base de datos utilizado para almacenar la data extraída de las fuentes de datos de la organización. A diferencia de las bases de datos transaccionales OLTP, un DWH utiliza base de datos multidimensionales OLAP. En el caso de implementar un DWH la opción más adecuada es la utilización de base de datos OLAP, sin

embargo en el nivel de madurez de un BI por lo general las organizaciones inician construyendo su DWH en bases de datos OLTP.

Existen tres tipos de bases de datos OLAP:

1. ROLAP, Relacional OLAP, las cuales trabajan sobre base de datos relacionales, se puede construir DWH sobre estas bases pero no son especializadas para lecturas y escrituras masivas y eficientes.
2. MOLAP, Multidimensional OLAP, trabajan sobre bases de datos multidimensionales en que la información es almacenada multidimensionalmente proveyendo mejor rendimiento al momento de realizar consultas por las distintas dimensiones de análisis. El inconveniente que se presenta en este tipo de base de datos es que se requiere almacenar grandes volúmenes de información.
3. HOLAP, Híbrido OLAP, es un híbrido entre ROLAP y MOLAP ya que se puede trabajar en base de datos relacional y agregar los datos de únicamente las principales dimensiones de análisis que se requiera.

Existen varias metodologías para el diseño y construcción de Data Warehouse, en el siguiente capítulo se detallan las principales metodologías que son reconocidas en este tema.

2.3 Definición de Marco Metodológico

De acuerdo a la definición propuesta por Piattini (1996), menciona que una metodología de desarrollo de software es un conjunto de procedimientos, técnicas, herramientas y soporte documental que ayuda a los desarrolladores a construir software.

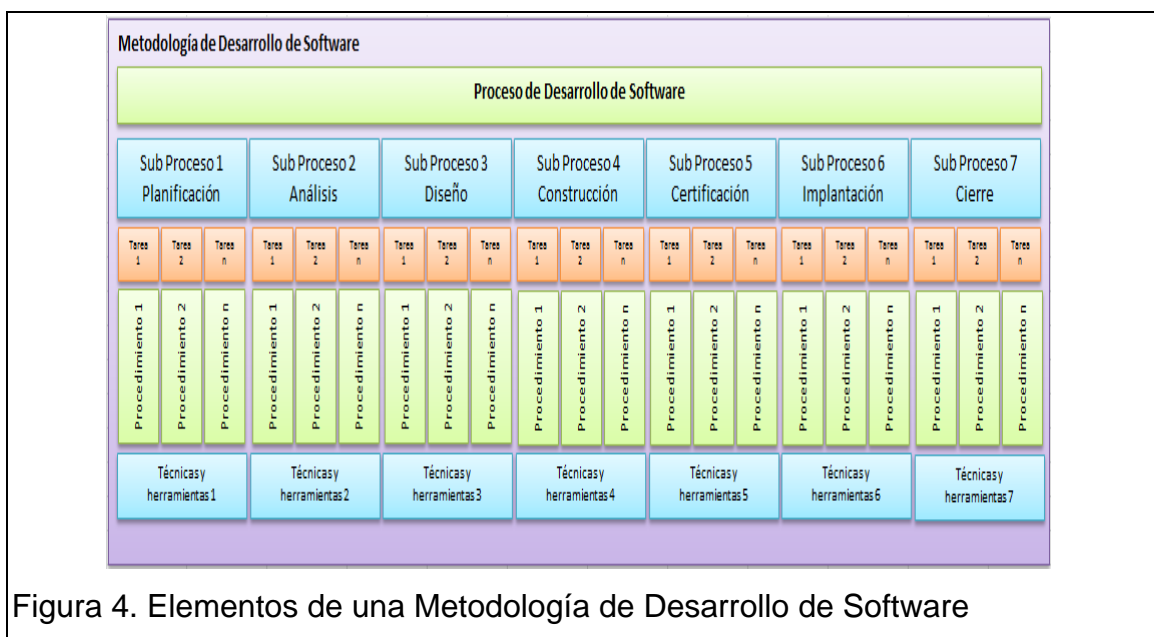
En ese sentido una metodología establece paso a paso las actividades que se deben ejecutar para el desarrollo de software de una manera integral. También debe establecer los roles que participan en el desarrollo de dichas actividades, y

detallar la información que debe producir como resultado de una actividad así como la información necesaria para comenzarla.

Según Paittini (1996) una metodología persigue cubrir los siguientes objetivos:

- Construir aplicaciones informáticas en lo posible sin errores, con un alto grado de calidad, que respondan a las necesidades de los usuarios.
- Generar la documentación de cada una de las fases de desarrollo lo que posteriormente permita realizar las tareas de mantenimiento eficientemente.
- Especificar los requerimientos de un software de manera adecuada.
- Contar con un método sistemático de desarrollo donde se pueda controlar el progreso de cada una de sus fases, asegurando el uso de recursos apropiados y costo adecuado.
- Contar con un proceso estandarizado a toda la organización, independiente del personal.

En la figura 4 se resumen los elementos principales que debe abordar una metodología de desarrollo de software:



Como se observa en la figura 4, una metodología consiste en la descomposición del proceso de desarrollo en subprocesos o fases los cuales son apoyados por técnicas que indican cómo debe realizarse cada actividad del subproceso. Esta descomposición facilita la planificación, gestión, control y evaluación de los proyectos de desarrollo de software durante todo su ciclo.

Cada subproceso o fase se debe descomponer en tareas, que son las actividades elementales en las que se dividen los procesos.

Para cada tarea se define un procedimiento en donde se especifica la manera de ejecutar la tarea. También se establece la manera de comunicación entre los usuarios y los desarrolladores.

Del resultado de la ejecución del procedimiento se obtiene un producto el cual sirve de entrada al siguiente proceso. El producto resultante del último proceso corresponderá al resultado final esperado por el usuario.

Los procedimientos establecen las técnicas que permiten llevar a cabo las tareas del subproceso o fase. Estas técnicas pueden ser apoyadas por herramientas como por ejemplo formatos de levantamiento de requerimientos, software de diseño de modelos conceptuales, software de construcción y test de aplicaciones, entre otros.

2.4 Metodologías de desarrollo Data Warehouse

Como se describió en el literal anterior 2.3, una metodología persigue establecer los pasos y actividades que se requieren ejecutar, guiados por procedimientos y técnicas para la construcción de software. Este mismo concepto puede ser utilizado al momento de hablar sobre las metodologías para la construcción de un DWH.

En la actualidad se puede obtener gran cantidad de información sobre varias metodologías tradicionales que permiten la construcción de un DWH. Por nombrar algunas están las metodologías de Ralph Kimball, Bill Inmon, Data Warehouse Engineering Process (DWEPE), SAS Rapid Warehousing, HEFESTO, CRISP-DM, entre otros.

Para elaborar la propuesta metodológica con un enfoque agile, aplicada al contexto de desarrollo de productos de DWH del SRI, primero se debe seleccionar una metodología tradicional de construcción de DWH que se adapte a las características de desarrollo de productos de DWH del SRI, y el mismo sirva como referencia al momento de definir el procedimiento para cada una de las fases de la propuesta metodológica.

En este sentido se identifican dos principales metodologías que se adaptan a las fases de desarrollo de productos de DWH del SRI, estas son las metodologías de Bill Inmon y Ralph Kimball. La primera hace mención a un esquema *top-down*, es decir crear un DWH con todos los datos de la organización al mismo nivel de detalle que se encuentran en las fuentes transaccionales para posteriormente crear los *data marts* específicos de cada unidad o proceso de negocio. Y la segunda cuyo enfoque es *botton-up*, es decir primero construye los *data marts* específicos de un proceso de negocio, los cuales se relacionan con otros *data marts* mediante dimensiones comunes, como por ejemplo la dimensión cliente, conformando un DWH empresarial, basándose en modelos multidimensionales (no normalizados).

Tanto la metodología de Bill Inmon como la de Ralph Kimball son aplicables para la construcción de productos de DWH del SRI, ya que se podría considerar construir un DWH que contenga todos los datos estructurados almacenados en las bases de datos del SRI para posteriormente ir creando los *data marts* específicos de cada proceso de negocio, o primero crear *data marts* basados en modelos multidimensionales por cada necesidad de negocio. En ese sentido las dos metodologías plantean el uso de un único repositorio de información donde integra y explota la información de diversas fuentes.

Finalmente hay que considerar que las dos metodologías se ajustan a la estructura actual de la arquitectura de DWH así como al equipo de desarrollo del Área de Datawarehouse del SRI por lo que son aplicables.

Adicionalmente se puede mencionar que las metodologías de Bill Inmon y Ralph Kimball son pioneras en establecer los conceptos arquitectónicos de DWH influyentes (Russom, 2014). Así también como se menciona en el artículo Data Warehousing Battle of Giant (Breslin, 2004), donde se considera a Bill Inmon como el padre del Data Warehousing, ya que al publicar su libro *Building The Data Warehouse* en el año de 1990, las empresas empezaron a implementar su visión. Y por otro lado Ralph Kimball quien desarrolló un nuevo modelo que compite con el de Inmon en su libro *The Data Warehouse Toolkit* del año 1996 haciendo hincapié a la construcción de bases de datos específicas de análisis de negocio denominadas *data marts* las cuales están enfocadas a los procesos de negocio de las empresas.

A continuación se describe las principales características de las metodologías de Bill Inmon y Ralph Kimball, las cuales son usadas como referencia al momento de definir el marco metodológico ágil enfocado al área de Datawarehouse del SRI.

2.4.1 Metodología de Bill Inmon

Inmon (2002, p. 31) define a un DWH como "...una colección de datos orientado a temas, integrados, no volátiles y variables en el tiempo, organizados para soportar necesidades empresariales...". De esta definición se puede establecer las 4 características principales de un DWH:

1. Orientado a temas, son aquellos datos relacionados que se generan a partir de un proceso de la operación de la organización los cuales son susceptibles a análisis, por ejemplo en la presentación de declaraciones

- se identifica al contribuyente, impuesto, periodo fiscal, la declaración, entre otros.
2. Integrados, los datos que almacena el DWH provienen de varias fuentes heterogéneas, cuyas estructuras tienen atributos diferentes como tipos de datos o precisión de los campos numéricos, almacenamiento diferente pero con el mismo significado. La integración permite resolver estos inconvenientes al formatear o convertir los datos de las distintas fuentes en un dato entendible para toda la organización la cual reside en el DWH.
 3. No volátil, los datos no son actualizables en el DWH, se mantiene la información histórica sobre los cambios que se realizan en las fuentes transaccionales.
 4. Variante en el tiempo, la información almacenada en el DWH al ser histórica es correcta hasta un momento en el tiempo.

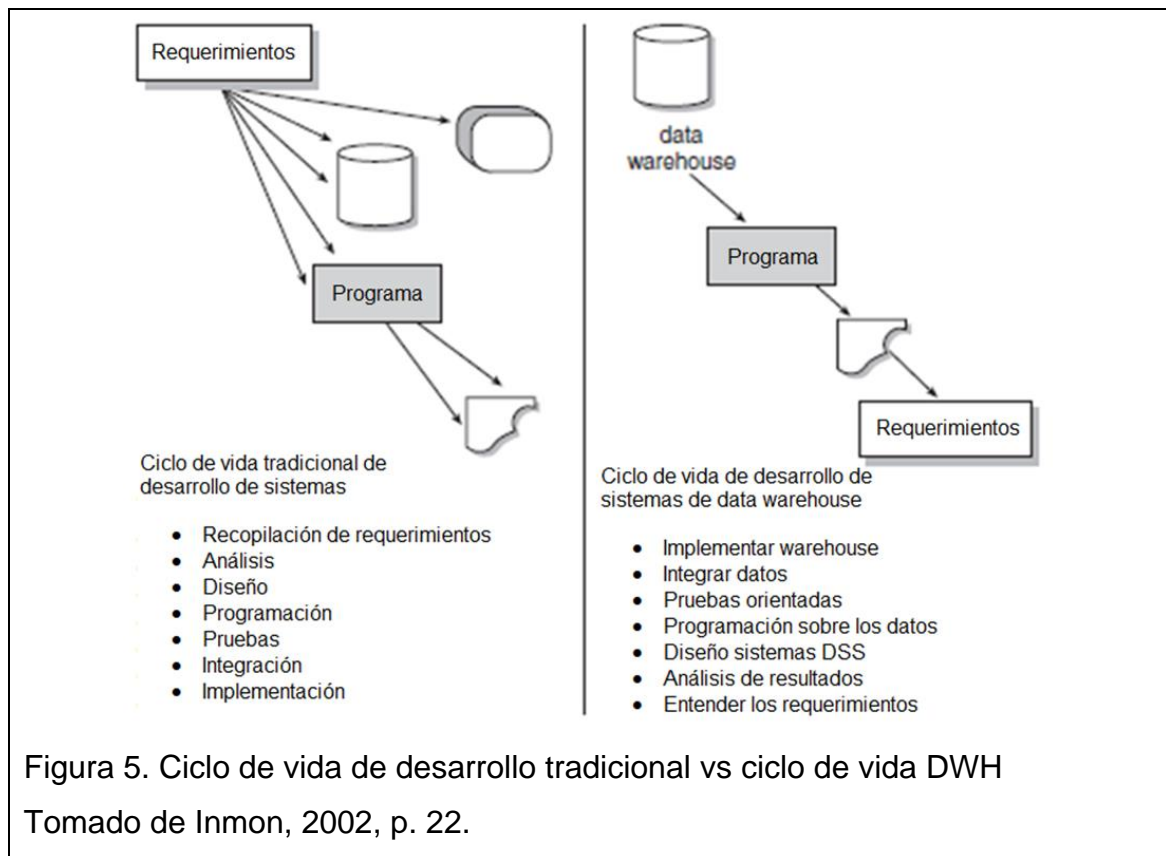
Básicamente el enfoque de esta metodología es trasladar el máximo nivel de detalle de las fuentes transaccionales hacia el DWH, al mismo nivel de normalización que por lo general está en tercera forma normal, es decir prácticamente se tiene una copia de los datos transaccionales en un área específica del DWH denominada *Staging*.

A continuación se describen los principales conceptos que menciona Inmon (2002) en referencia a la construcción de un DWH.

2.4.1.1 Ciclo de Vida de Inmon

El ciclo de vida de desarrollo de aplicaciones transaccionales siempre inicia con los requerimientos funcionales, mientras que el ciclo de vida del desarrollo de un DWH es lo contrario: inicia con los datos de las fuentes transaccionales, los cuales son integrados y probados, luego se desarrollan programas que leen dichos datos, a lo que se podría llamar publicación de datos para que sean analizados, y finalmente se establecen y definen los requisitos de los DSS. En la

figura 5 se esquematiza lo anteriormente señalado. Este ciclo de vida calza perfectamente en una metodología tipo espiral.



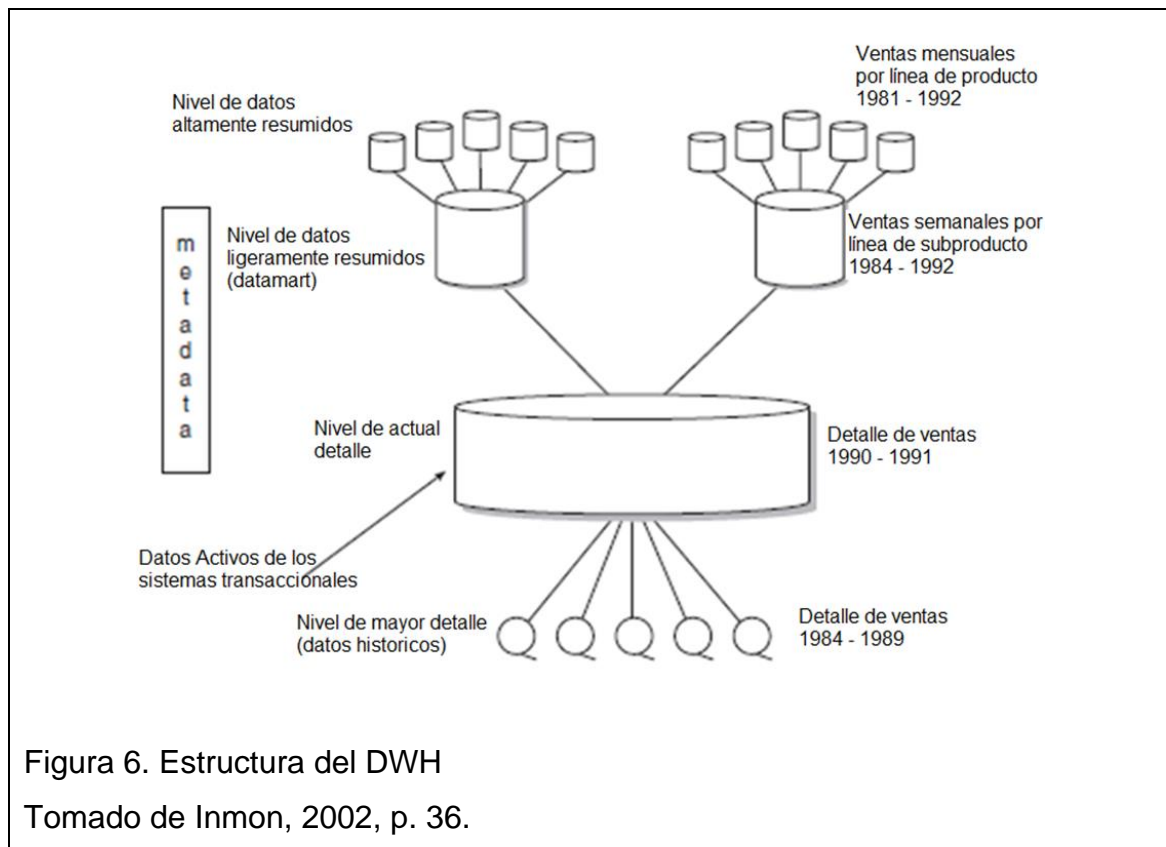
2.4.1.2 Estructura de un DWH

Inmon (2002) establece que la estructura básica de un DWH está dividida en cuatro niveles:

1. Nivel de mayor detalle, que corresponde a los datos históricos de los sistemas transaccionales
2. Nivel de detalle actual, el cual corresponde a los datos activos de los sistemas transaccionales
3. Nivel de datos ligeramente resumidos, el cual corresponde a data que tiene una primera transformación.

4. Nivel de datos altamente resumidos, el cual corresponde a data ya agregada o resumida específica para una unidad de negocio de la organización.

En la siguiente figura se esquematiza los cuatro niveles mencionados:



2.4.1.3 Proceso de construcción de un DWH

A continuación se enumeran los pasos de construcción de un DWH que propone Inmon (2002):

1. Identificar los sistemas transaccionales de la organización que sirven de fuentes de datos del DWH.
2. Las primeras tablas de la primera unidad de negocio son pobladas en el DWH. Inician los usuarios a acceder a los datos integrados.

3. Se cargan más tablas al DWH y por ende existen más usuarios que acceden a la data almacenada al DWH.
4. Los datos operativos son cargados correctamente al DWH y aparecen varios sistemas DSS que requieren consumir dichos datos.
5. Aparecen los *data marts* específicos de cada unidad de negocio. La ventaja de crear un *data mart* es el acceso más eficiente a los datos requeridos por la unidad del negocio.
6. Aparecen más necesidades que requieren la construcción de *data marts* para cada una de las unidades de negocio.
7. Finalmente la arquitectura está desarrollada. Hay pocos usuarios que acceden a los datos del DWH y muchos usuarios que prefieren acceder a los *data marts* donde se realizan las actividades de procesamiento analítico de los sistemas DSS

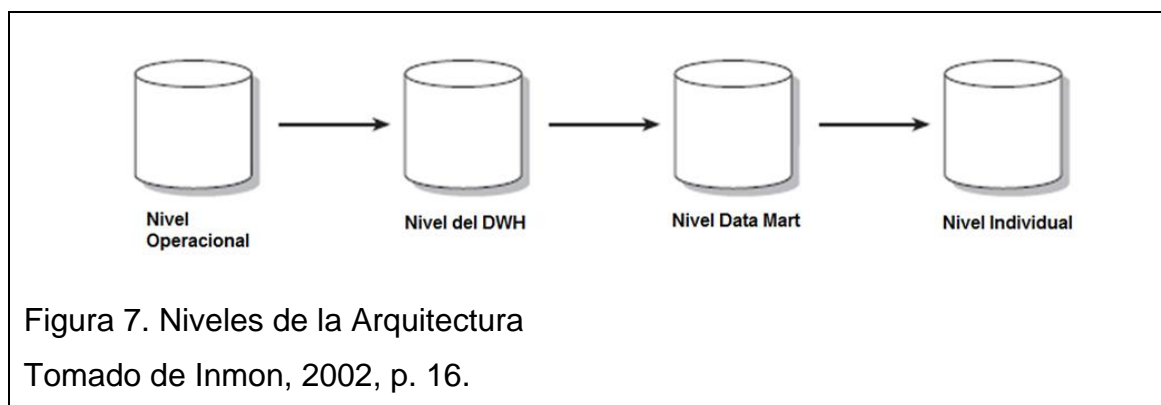
La evolución del punto 1 al 7 toma largo tiempo, por lo general varios años.

2.4.1.4 Granularidad

Según Inmon (2002) se refiere al nivel de detalle o de resumen de las unidades de datos en el DWH. Cuanto más detalle existe, menor es el nivel de granularidad. A menor detalle que exista, es mayor el nivel de granularidad.

2.4.1.5 Niveles de la Arquitectura de DWH

En la figura 7 se esquematizan los niveles de la arquitectura propuesta por Inmon (2002):



- Nivel operacional. Proceso que registra los datos de las aplicaciones transaccionales, corresponde a los datos que se generan en el ambiente transaccional.
- Nivel del DWH. Datos transaccionales integrados, históricos que no son actualizables, con ciertos datos derivados requeridos por la unidad de negocio así como nuevos campos calculados o marcas. En este nivel los datos son usados para explotar información voluminosa.
- Nivel *Data Mart*. Datos exclusivos de una unidad específica de negocio en base a los requerimientos de los usuarios finales adaptados a las necesidades de la unidad de negocio. Por ejemplo *data mart* de marketing, ventas, contabilidad, etc. Si bien los datos se encuentran a nivel operativo en el *data mart*, estos se encuentran desnormalizados y resumidos para atender a los requerimientos de la unidad de negocio.
- Nivel Individual. Datos específicos para hacer análisis heurístico, suelen ser temporales y enfocados sobre un grupo de datos no voluminosos.

Los datos son pasados desde el entorno transaccional u operativo hacia el ambiente de DWH. Deben ser almacenados en el DWH ya integrados (dato entendible para la organización), ya que si los datos son almacenados de la

misma forma que se encuentran en el entorno operacional, no podrán ser utilizados para apoyar una visión global de la información de toda la organización, este punto es esencial en la arquitectura.

2.4.1.6 Usuarios del DWH

Inmon (2002) los llama también como analistas de los sistemas DSS. Tienen las siguientes características:

1. Definir y descubrir la información utilizada en la toma de decisiones corporativas.
2. Requieren acceso a todos los datos para observar todas las posibilidades de análisis existentes, y posteriormente descubrir y definir los criterios de análisis que cubran sus necesidades de información.

2.4.1.7 Modelo de Datos del DWH

Inmon menciona (2002) que el modelo de procesos utilizado en los ambientes operativos, podría ser usado como referencia para generar los modelos de datos del DWH, de tal forma que se genere un modelo de datos corporativo el cual podría compararse al mapa general de procesos de la organización. Existen 3 niveles de modelamiento:

1. Modelado de alto nivel (ERD, *entity relationship level*), diagrama o modelos de datos a un alto nivel entendible para los usuarios finales o de negocio. Para cada área o entidad de negocio se crea un ERD.
2. Modelado de nivel medio (DIS, *data item set*), se crean cuatro construcciones de modelado básicos usados para identificar los atributos de los datos y sus relaciones:
 - Agrupación primaria de datos: Contienen atributos que existen una sola vez para cada área principal.

- Agrupación secundaria de datos: Son atributos de datos que pueden existir varias veces para cada área principal.
 - Relación entre grupos.
 - Tipos de datos.
3. Modelado de bajo nivel (Modelo físico), es importante especificar la granularidad del dato, ya que influye directamente con respecto a las características de rendimiento del DWH, específicamente sobre el almacenamiento. El diseño físico de los datos debe estar de tal forma que se optimicen las consultas que se realicen sobre el DWH.

A partir del modelo de datos se generan las iteraciones de desarrollo, ya que de no ser así puede haber superposición de esfuerzos y desarrollos separados.

2.4.1.8 Varios conceptos adicionales relevantes

Inmon (2002) menciona varios conceptos adicionales a los ya nombrados y que son necesarios considerarlos al momento de la construcción de un DWH. Entre los que se debe destacar están:

- Normalización y des normalización, cuyo objetivo es combinar las tablas de los datos en una única estructura des normalizada, donde se aprecian datos redundantes, datos derivados o nuevos cálculos, paso de filas a columnas, entre otros, lo que permite un acceso más eficiente a los datos almacenados en el DWH.
- El almacenamiento de eventos como el registro de llamadas telefónicas, la compra o venta de productos para realizar análisis en el tiempo sobre dichos eventos.
- El concepto de meta data, que permite a los usuarios acceder a la información del DWH en un modo que es fácilmente interpretado por los usuarios.

- El concepto de ciclicidad del dato donde Inmon se refiere a la frecuencia en que debe actualizarse los cambios del transaccional hacia el DWH. Por regla general se establece cada 24 horas.

2.4.2 Metodología de Ralph Kimball

Ralph Kimball junto con Margy Ross (2013) son otros de los autores de los principales exponentes de las metodologías de construcción de DWH. En su modelo difieren en varios aspectos importantes al enfoque tradicional de base de datos relacionales. Establecen a los modelos multidimensionales como base del DWH.

Kimball y Ross (2013) definen una clara diferencia entre datos que son capturados de los sistemas transaccionales, y los datos para uso de análisis para la toma de decisiones. El primero nace de la automatización de los procesos de la operación de las empresas los cuales son almacenados en bases de datos relacionales, y el segundo utiliza dichos datos para ser analizados lo que permite generar información y conocimiento para la toma de decisiones de las organizaciones.

Los objetivos de un sistema DWH y BI según Kimball y Ross (2013) se resume en los siguientes puntos:

- Fácil acceso a la información, el contenido del DWH debe ser comprensible, intuitivo y evidente para los usuarios de negocio. Las herramientas de BI que acceden a los datos del DWH deben ser simples y fáciles de usar retornando la información requerida por el usuario en tiempos óptimos.
- Almacenar información coherente, los datos en el DWH deben ser creíbles, así como cuidadosamente ensamblados a partir de una variedad de fuentes. Deben pasar por un proceso de limpieza y calidad garantizada. Deben ser liberados sólo cuando es apta para el consumo

del usuario. La coherencia también implica definiciones comunes para los contenidos que alberga el DWH.

- Adaptación al cambio, el DWH y las aplicaciones de BI deben estar diseñados para manejar los cambios que surjan de las nuevas necesidades de los usuarios, así como a las condiciones del negocio y hasta considerar los cambios de tecnología.
- Presentación de información de manera oportuna, los usuarios del negocio necesitan tener expectativas realistas al momento de realizar toma de decisiones.
- Seguridad de la información, se debe administrar de manera eficaz el acceso a la información confidencial que alberga el DWH.
- Apoyar a la toma de decisiones, el DWH debe proveer información fidedigna y confiable para una adecuada toma de decisiones. El DWH será evaluado por la certeza de sus resultados al momento que los usuarios realicen el proceso de toma de decisiones.
- Aceptar un sistema de DWH por la organización, no por el hecho de tener la mejor infraestructura de DWH se puede considerar un éxito. La organización debe proponer proyectos que permitan explotar dicha infraestructura.

La metodología propuesta por Kimball y Ross (2013) se basa en el modelamiento multidimensional para la presentación de datos analíticos. Este modelo pretende de forma general atender los siguientes puntos:

- Entregar datos que sean comprensibles para los usuarios de negocios.
- Entregar rendimiento a consultas rápidas.

Según Kimball y Ross (2013) el modelo dimensional es una técnica para generar bases de datos simples, estructuras multidimensionales sencillas entendibles para las personas. La simplicidad es fundamental, ya que garantiza que los usuarios puedan entender fácilmente los datos. Así también el software de BI

que se utilice, debe permitir navegar y obtener resultados de forma rápida y eficiente.

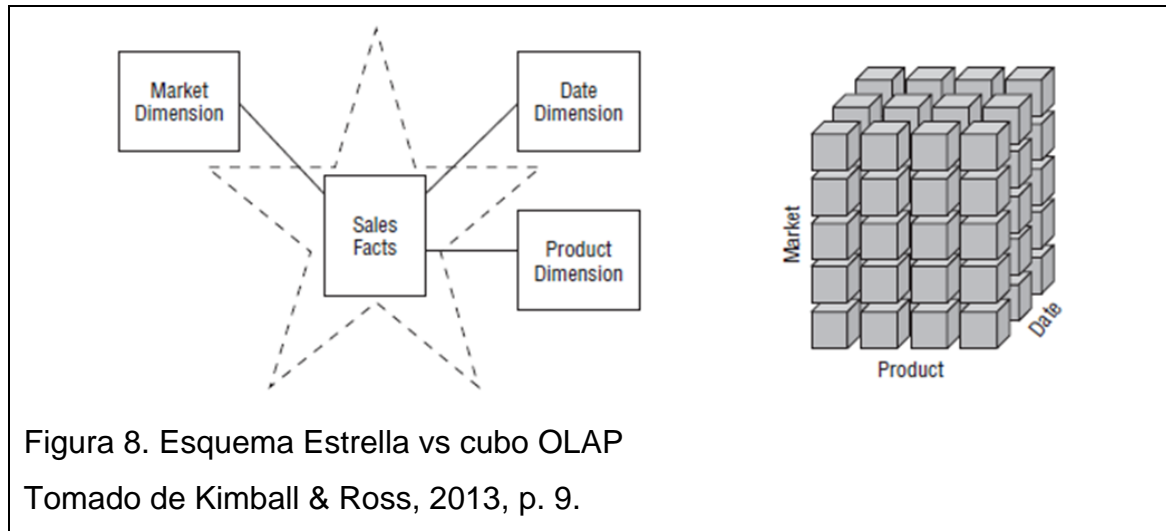
A diferencia de los modelos de datos utilizados en los sistemas transaccionales, los cuales están en tercera forma normal (normalizados) o conocidos también como modelos entidad relación, que son útiles en los procesamientos de inserción y actualización de datos, los modelos multidimensionales están en primera forma normal lo que permite consultas más rápidas y eficientes.

Un modelo dimensional contiene la misma información que un modelo normalizado, pero empaqueta los datos en un formato que ofrece comprensibilidad al usuario, rendimiento en las consultas, y adaptación al cambio.

A continuación se describen los principales conceptos mencionados en la metodología de Ralph Kimball y Margy Ross en referencia a la construcción de un DWH.

2.4.3.1 Esquema Estrella vs Cubo OLAP

Los modelos multidimensionales implementados en bases relacionales se denominan esquemas estrellas debido a su parecido con una estructura en forma de estrella. Los modelos multidimensionales implementados en entornos de bases de datos multidimensionales se denominan cubos OLAP (Kimball & Ross, 2013). En la siguiente figura se ilustra lo antes mencionado:



Como recomiendan Kimball y Ross (2013) la información detallada debe ser cargada en un esquema estrella, y los cubos OLAP poblados a partir de este esquema.

Un cubo OLAP provee capacidades de análisis más eficientes frente a las implementaciones de base de datos relacionales, sin embargo esa distinción ha perdido importancia con los avances en hardware, tales como equipos especializados en procesamientos en paralelo, bases de datos columnares y en memoria.

La metodología de Kimball y Ross (2013) enfatiza el estudio del esquema estrella bajo una plataforma de base de datos relacional enfocándose a los dos componentes principales del modelo dimensional: tablas de hechos y dimensiones.

2.4.3.2 Tabla de Hechos y Medidas

Entre las características identificadas de las tablas de hechos, y que son mencionadas por Kimball y Ross (2013), se pueden nombrar las siguientes:

- Una tabla de hechos, en un modelo multidimensional, almacena las mediciones de desempeño o rendimiento de los eventos de los procesos de negocio de la organización.
- Cada fila de una tabla de hechos corresponde a un evento de medición. Los datos de cada fila están en un nivel específico de granularidad.
- Uno de los principios básicos del modelado dimensional es que todas las filas de medición en una tabla de hechos deben estar en el mismo nivel de granularidad.
- Los hechos por lo general son representaciones que permiten realizar operaciones de suma tales como la cantidad de ventas en dólares. Es posible que existan hechos textuales, pero esta condición rara vez se presenta.
- No se debe almacenar información textual redundante en tablas de hechos, a menos que el texto sea único para cada fila de la tabla de hechos.
- Las tablas de hechos por lo general constituyen el 90 por ciento o más del total de espacio consumido por un modelo dimensional. Tienden a ser profundas en términos del número de filas, pero estrechas en términos del número de columnas. Dado su tamaño, se debe considerar el espacio requerido de almacenamiento que va a ocupar en el DWH.
- Todas las tablas de hechos tienen dos o más claves externas que se conectan a las claves principales de las tablas de dimensiones. La combinación de todas las claves externas conforma la clave compuesta de la tabla de hechos.

Ejemplo de tablas de hechos se puede nombrar los eventos de ventas, compras, registro contable, movimiento de inventarios, por nombrar algunos.

Por lo general los hechos son todos aquellos eventos en los procesos de negocio donde se identifica un valor de medida que sirve para analizar a dichos procesos (Kimball & Ross, 2013).

2.4.3.3 Tabla de Dimensiones

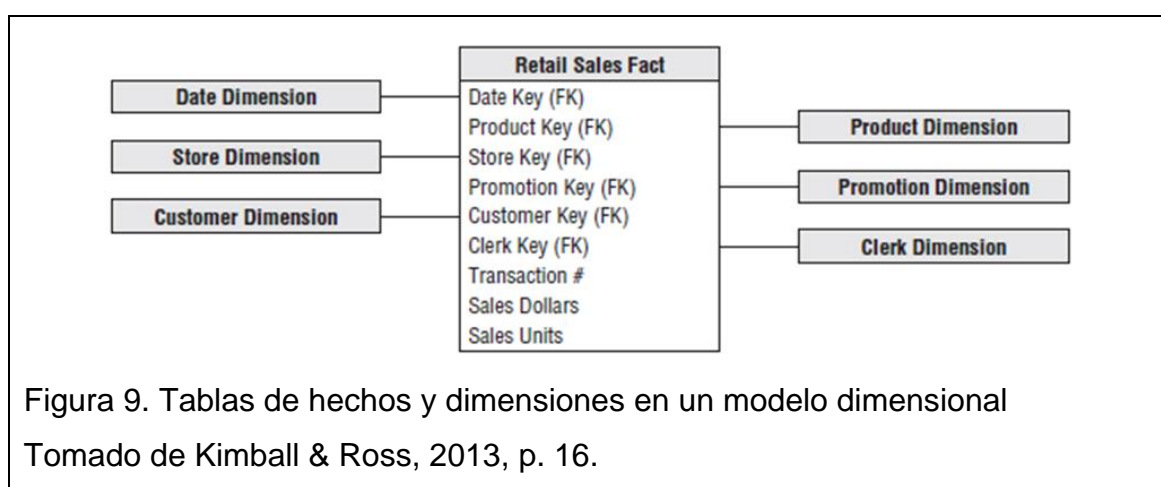
Según Kimball y Ross (2013), sostienen que las tablas de dimensiones están asociados a la tabla de hechos y contienen el contexto textual asociada a un evento de la medición de un proceso de negocio. Describen el quién, qué, dónde, cuándo, cómo y por qué asociado con el evento. Mencionan también las particularidades que definen a las tablas de dimensiones, las cuales se pueden resumir en las siguientes:

- Tienen pocos registros pero pueden tener muchas columnas o atributos. Cada dimensión posee una sola clave primaria que permite relacionarse con las tablas de hechos.
- Los atributos de las dimensiones son utilizados en los filtros de las consultas, así como también son parte de los campos agrupados al momento de realizar operaciones agregadas sobre los campos de medición de las tablas de hechos.
- Los atributos deben ser de tipo texto evitando la utilización de códigos ya que son parte de los reporte por lo que se requiere que sean entendibles por los usuarios, es decir deben enmarcarse a la terminología del negocio. En el caso de que los usuarios requieran visualizar códigos operacionales, estos deberán ser parte de los atributos de las dimensiones.
- La estructura de una dimensión pretende facilitar a los usuarios el análisis de los eventos que son representados por los campos de medidas de las tablas de hechos.
- Poseen niveles y jerarquías que proveen al usuario la funcionalidad de navegar por niveles de la información y sobre los cuales se puede realizar operaciones de agregación.

Ejemplos de dimensiones se puede nombrar el tiempo, producto, mercados, personas, ubicación geográfica, sectores, por nombrar algunas.

2.4.3.4 Hechos y Dimensiones en un esquema Estrella

Cada proceso de negocio está representado por un modelo multidimensional que consiste en una tabla de hechos la cual contiene mediciones numéricas del evento, rodeado de varias tablas de dimensiones que contienen el contexto textual al momento de que se originó el evento. Esta estructura se la denomina esquema estrella (Kimball & Ross, 2013). En la siguiente figura se esquematiza lo antes mencionado:



Entre las características que se identifican en este esquema se pueden nombrar las siguientes:

- En este esquema multidimensional prevalece la sencillez y simetría, permitiendo a los usuarios de negocios beneficiarse de la simplicidad porque los datos son más fáciles de entender y navegar.
- Proporciona ventajas de rendimiento ya que las bases de datos procesan más eficientemente las consultas a estos modelos.
- Su diseño es altamente reconocible para los usuarios de negocios por lo que proporciona una funcionalidad más intuitiva al usuario al momento de generar los reportes analíticos.
- La estructura de este diseño se basa en el nivel de granularidad, o también llamado nivel atómico de detalle de los datos almacenados en las tablas de hechos, ya que a mayor nivel de detalle el modelo puede

responder a las preguntas inesperadas de los usuarios de negocio, y resolver los reportes *ad hoc*.

2.4.3.5 Arquitectura de DWH/BI de Kimball

Kimball & Ross (2013) proponen una arquitectura de cuatro componentes principales: fuentes de sistemas transaccionales, sistemas de ETL, área de presentación de datos y aplicaciones de BI, como se observa en la siguiente figura.

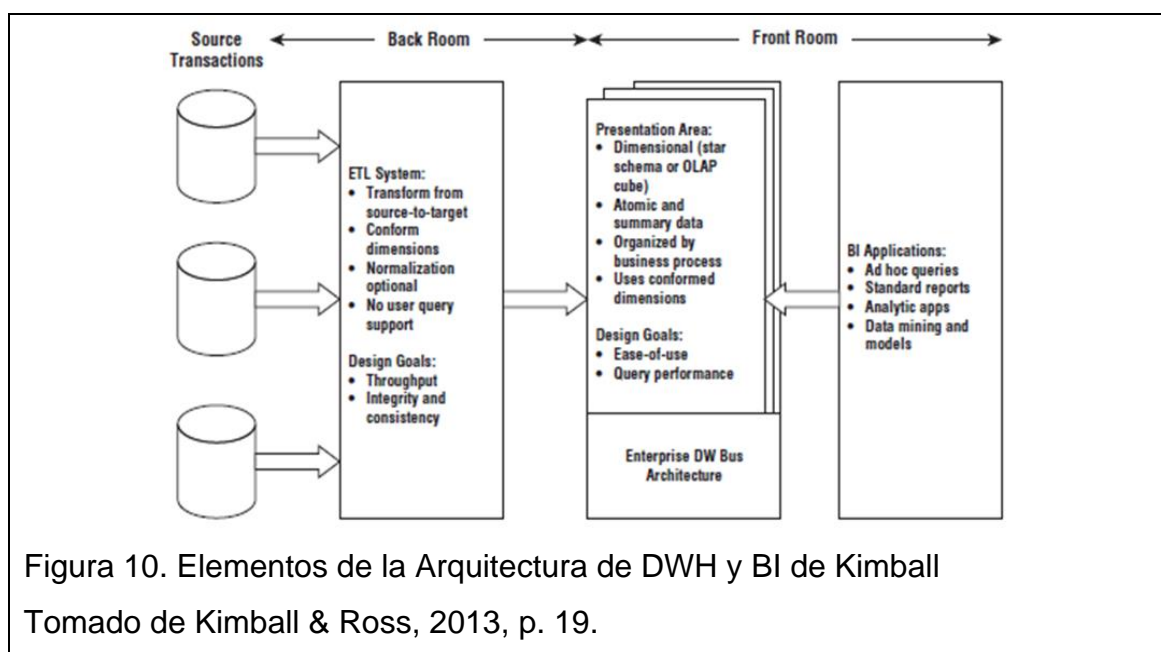


Figura 10. Elementos de la Arquitectura de DWH y BI de Kimball

Tomado de Kimball & Ross, 2013, p. 19.

A continuación se describen cada uno de los componentes:

Fuentes transaccionales.- Aplicaciones que capturan las transacciones de los procesos de las organizaciones. Este elemento no es parte del DWH pero sirve de insumo para la carga de datos hacia el DWH. El principal objetivo de este elemento es alta disponibilidad de las aplicaciones y alto rendimiento en el procesamiento transaccional.

Las consultas que se ejecutan bajo este ambiente corresponden a reportes operativos, en tiempo real y sobre datos específicos que corresponden al flujo

normal de las transacciones. En este ambiente transaccional no se deben ejecutar consultas masivas y de rangos amplios que normalmente son competencia de los sistemas de DWH y BI.

Fuentes transaccionales mantienen pocos datos históricos, en este ambiente se ejecutan aplicaciones transaccionales como son los ERP.

Sistema de ETL.- El primer paso es la obtención de los datos. El segundo paso se realizan operaciones de transformación, tales como la limpieza y estandarización de los datos, combinación de datos de múltiples fuentes y calidad de datos como revisión de datos duplicados. El objetivo es mejorar la calidad de la data. El paso final es la carga de los datos en una estructura física en un modelo dimensional.

Kimball & Ross (2013, pág. 20) ponen a consideración de cada organización la necesidad de cargar los datos normalizados hacia el DWH antes de generar los modelos multidimensionales para la presentación de consultas e informes. Sin embargo este proceso de dos etapas requiere más tiempo e inversión para el desarrollo, más tiempo para la carga periódica o actualización de los datos, y más capacidad para almacenar las múltiples copias de los datos.

Área de Presentación.- El Área de presentación es el lugar donde se organizan los datos almacenados en modelos multidimensionales para ponerlos a disposición de los usuarios del negocio, con el objetivo de realizar consultas directas o *ad hoc*, generación de informes y otras aplicaciones de BI y analíticas.

En esta área se establecen las relaciones entre las dimensiones y hechos que conforman un esquema estrella, y las relaciones entre varios esquemas estrella mediante dimensiones en común conformando la arquitectura de bus de almacenamiento de datos de la empresa.

Aplicaciones de BI.- Corresponde al componente final de la arquitectura y se refiere a la gama de funcionalidades proporcionadas a los usuarios de negocio

para aprovechar el Área de presentación para la toma de decisiones analíticas. Pueden ser herramientas tan simples como de consultas *ad hoc* o tan complejas como de minería de datos.

2.4.3.6 Ciclo de vida de Kimball

En la siguiente figura se describe las fases que proponen Kimball & Ross (2013) durante el ciclo de vida de desarrollo de productos DWH/BI:

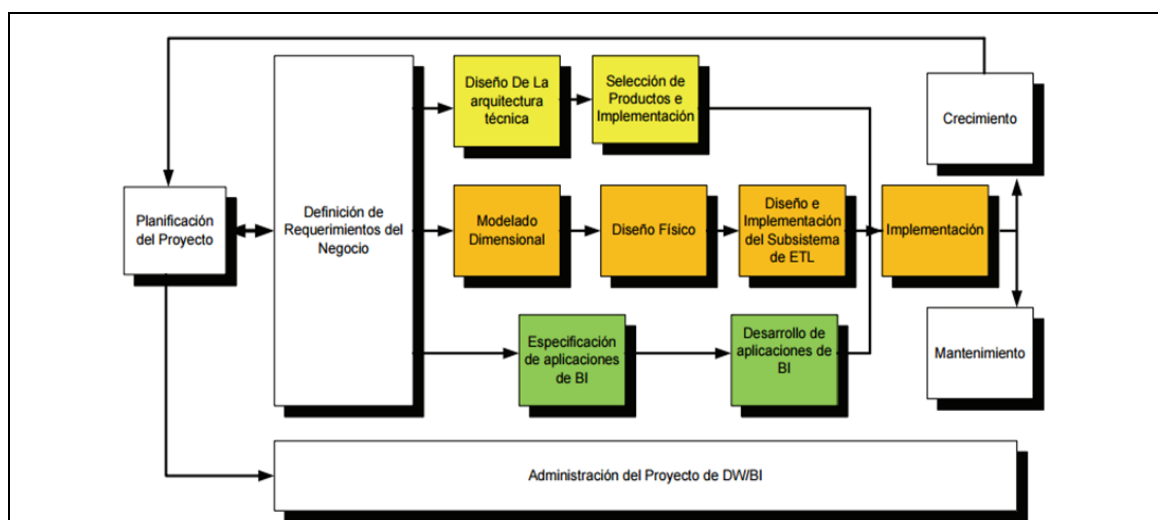


Figura 11. Ciclo de Vida de Kimball

Tomado de Kimball & Ross, 2013, p. 404.

El enfoque del ciclo de vida de Kimball y las metodologías ágiles comparten algunas doctrinas comunes: un enfoque en generar valor al negocio mediante desarrollos incrementales la cual tiene que ser construida sobre una sólida arquitectura y gobernanza de los datos impulsada por la arquitectura de bus (Kimball & Ross, 2013). A continuación se describen cada una de sus fases:

El ciclo de vida inicia con la planificación del proyecto donde se establece el alcance y justificación así como los recursos necesarios del proyecto. Establece los lineamientos necesarios para gestionar el resto de actividades del proyecto.

La siguiente fase corresponde a la definición de los requerimientos del negocio la cual está asociada a la planificación del proyecto. Los requerimientos deben estar alineados a los objetivos estratégicos de la organización y constituyen una entrada principal al resto de tareas en el ciclo de vida. Parte de esta fase se encuentra la elaboración de la Matriz del Bus Empresarial de Data Warehouse que es una herramienta de diseño multidimensional y de arquitectura de datos empresarial, en la cual se representan los principales procesos de la organización asociados a los hechos o medidas de análisis de los mismos, y la relación entre las dimensiones de análisis comunes de la organización por cada proceso de negocio.

La Matriz es un elemento esencial en la metodología propuesta por Kimball y Ross, ya que permite visualizar por cada proceso de negocio los hechos construidos y la relación con las dimensiones comunes que son las que permiten establecer la relación con el resto de los procesos de la organización.

Las fases que se encuentran en los cuadros de color amarillo se ocupan de la solución tecnológica: El diseño de la arquitectura técnica establece el marco general para apoyar la integración de múltiples tecnologías. Utilizando las capacidades identificadas en el diseño de la arquitectura, se establecen las características que debe cumplir la solución tecnológica, y se evalúan y seleccionan los productos específicos acorde a las necesidades establecidas en el diseño de arquitectura.

Las fases que se encuentran en los cuadros de color tomate se centran en los datos. Se inicia mediante la traducción de los requisitos en un modelo multidimensional, que posteriormente se transforma en una estructura física. El diseño físico se centra en las estrategias de optimización de rendimiento, tales como la agregación, la indexación y particionamiento. Por último, corresponde al diseño y desarrollo del ETL.

El conjunto final de tareas que se encuentran en los cuadros de color verde corresponden a las fases de diseño y desarrollo de las aplicaciones de BI. Y por

último, una gestión del crecimiento y mantenimiento de los proyectos implementados, iniciando nuevamente el ciclo de vida de desarrollo de DWH.

2.5 Marcos de Trabajo Ágiles

Antes de describir una definición sobre marcos de trabajo ágiles, se identifica que en la bibliografía revisada sobre este tema, se denota que varios autores denominan al proceso de desarrollo ágil como metodologías, marcos metodológicos o marcos de trabajo ágiles. En tal sentido se propone considerar al proceso de desarrollo ágil como un marco de trabajo agile considerando lo expuesto por Kniberg (2007) donde menciona que Scrum no es una metodología sino más bien un marco de trabajo.

Un marco de trabajo *agile* no establece los lineamientos de construcción o de desarrollo de software, más bien propone simplificar el proceso de ciclo de vida de desarrollo de software, a través de la colaboración de equipos multidisciplinarios en los cuales está incluido al cliente como parte del equipo, entregando productos con valor en intervalos cortos de tiempo lo que permite proveer a los interesados funcionalidades para un retorno de inversión a corto plazo (SCRUMstudy, 2013).

Existe una gran variedad de marcos de trabajo que promueven el enfoque ágil, entre los principales marcos más populares para el desarrollo ágil de proyectos se puede nombrar a Scrum y XP (Programación Extrema) (INTECO, 2009). Otro marco de trabajo ágil es Kanban el cual no se enfoca necesariamente a proyectos que tienen un inicio y fin, más bien se orienta a un flujo de trabajo continuo o trabajo en curso (Bermejo, 2010).

Este proyecto de investigación se enfoca al estudio del marco ágil Scrum ya que ha sido utilizado en proyectos de desarrollo de software y es considerada como el marco de trabajo para desarrollo de proyectos de manera ágil preferida a nivel mundial (SCRUMstudy, 2013). También se considera el marco ágil de Kanban

ya que se enfoca a entornos de desarrollo de mantenimiento de aplicaciones, así como también permite controlar el avance en cada una de las fases de desarrollo (Bermejo, 2010). A continuación se describen las características principales de cada uno de ellos.

2.5.1 Scrum

En el documento de Hundermark (2009, p. 2) define a Scrum como "... un framework para el manejo de proyectos que tienen como fin el desarrollo de productos complejos...".

Según el SBOK (SCRUMstudy, 2013, p. 2) Scrum "es una metodología de adaptación, iterativa, rápida, flexible y eficaz, diseñada para ofrecer un valor significativo de forma rápida en todo el proyecto." y que "...está estructurado de tal manera que es compatible con los productos y el desarrollo de servicio en todo tipo de industrias y en cualquier tipo de proyecto, independientemente de su complejidad".

El marco de trabajo de Scrum es adaptable a los proyectos de desarrollo de software el cual tiene como objetivo reducir las restricciones de tiempo y simplificar los procesos del ciclo de vida de construcción de aplicaciones permitiendo adaptarse a los entornos cambiantes de la organización ofreciendo productos de calidad. Los expertos Ken Schwaber y Jeff Sutherland son los fundadores del concepto de Scrum y su aplicabilidad al desarrollo de software. Ellos fueron parte de los 17 expertos autodenominados la Alianza Ágil, que se reunieron para elaborar el Manifiesto de Desarrollo Ágil de Software en el cual constan los valores y principios que respaldan este manifiesto (SCRUMstudy, 2013).

Los valores que se establecen en el Manifiesto Ágil son:

- Individuos e interacciones sobre procesos y herramientas, es decir crear un buen equipo de trabajo antes que construir un entorno.
- Desarrolla software funcionando sobre documentación extensa, únicamente generar la documentación que realmente sea necesaria de forma inmediata.
- Colaboración con el cliente sobre negociación de un contrato, interacción constante entre el cliente y el equipo de trabajo que asegure el éxito del proyecto.
- Responder ante el cambio sobre seguir un plan, la planificación no es estricta, y se valora la habilidad de responder a los cambios.

De los valores anteriores se generan los doce principios definidos por el Manifiesto Ágil. De acuerdo a lo mencionado por Hundermark (2009), se resumen los siguientes doce principios:

1. La prioridad es satisfacer al cliente mediante la entrega temprana y continua de software que aporte valor a la organización.
2. Aceptar que los requisitos cambien, incluso en etapas tardías del desarrollo. Los procesos ágiles aprovechan el cambio para proporcionar ventajas competitivas a la organización.
3. Entregar software funcional en intervalos cortos de tiempo.
4. Los responsables de negocio y los desarrolladores trabajan juntos de forma cotidiana durante todo el proyecto.
5. Los proyectos se desarrollan en torno a individuos motivados.
6. El método más eficiente y efectivo de comunicar información al equipo es la conversación cara a cara.
7. El software funcionando es la medida principal de progreso.
8. Los procesos ágiles promueven el desarrollo sostenible. Los promotores, desarrolladores y usuarios deben ser capaces de mantener un ritmo constante de forma indefinida.

9. La atención continua a la excelencia técnica y al buen diseño mejora la agilidad.
10. La simplicidad, o el arte de maximizar la cantidad de trabajo no realizado, es esencial.
11. Las mejores arquitecturas, requisitos y diseños emergen de equipos auto organizados.
12. A intervalos regulares el equipo reflexiona sobre cómo ser más efectivo ajustando y perfeccionando su método de desarrollo.

De acuerdo a lo que se menciona por Hundermark (2009), los roles que conforman el equipo Scrum son:

Dueño del producto o *Product Owner* que entre sus responsabilidades están:

- Posee una visión general de la problemática a resolver.
- Identifica las necesidades funcionales del producto.
- Administra y prioriza la lista de necesidades (*Product Backlog*).
- Acepta el software construido al final de cada iteración.
- Administra el plan de liberación.
- Maximiza el retorno de inversión (ROI) del proyecto.

Facilitador o *Scrum Master*, hace todo lo que esté en sus manos para ayudar al equipo a cumplir sus objetivos y cumplir con éxito los entregables del proyecto:

- Crea un entorno adecuado de trabajo para el equipo.
- Elimina impedimentos.
- Asegura que se usen las técnicas de Scrum dentro del proyecto.

El Equipo que está conformado por miembros multidisciplinarios con las habilidades requeridas para desarrollar el trabajo encomendado. Entre sus responsabilidades están:

- Estimar el esfuerzo de desarrollo de la lista de necesidades.
- Compromiso de entregar productos con la funcionalidad requerida en producción.
- Auto gestión de su propio avance.
- Auto organización, con la responsabilidad ante el *Product Owner* de entregar aquello que fue comprometido.
- Interesados del proyecto, usuarios finales del producto, personas a quienes el proyecto producirá un beneficio y administradores que toman las decisiones respecto al proyecto.

Una de las principales bases de Scrum es la creación de ciclos frecuentes donde se realizan las actividades de desarrollo llamadas iteraciones y que Scrum los denomina *Sprints*.

De acuerdo a lo mencionado por Gallego (), el proceso de desarrollo de Scrum se divide en 3 grandes fases o también consideradas como reuniones, a continuación se describe cada una de las tres fases:

1. Planificación del *Product Backlog*, detalla las necesidades funcionales del producto por prioridades, se establece los objetivos y las actividades que se realizan en la iteración.
2. Seguimiento del *Sprint*.- Son reuniones diarias durante el tiempo que dure el *Sprint* en la que se evalúa el trabajo realizado desde la reunión anterior, el trabajo que se realiza hasta una nueva reunión, e identifica los inconvenientes que han surgido y cómo solucionarlos.
3. Revisión del *Sprint*.- Es la reunión de finalización de la iteración donde se evalúan los resultados finales con la demo de la funcionalidad construida la cual se debe desplegar en producción.

Los elementos que forman parte de Scrum según Gallego () son:

1. *Product Backlog*, lista de necesidades del proyecto priorizadas, con su estimación de esfuerzo que incluirán las condiciones o criterios de

aceptación del entregable final. Las necesidades son las descripciones de las funcionalidades que va a tener el software, a eso se le denomina historia de usuario.

2. *Sprint Backlog*, lista de tareas ordenadas que se realizan en un *Sprint*, detallando los responsables de desarrollar las tareas, el tiempo requerido para terminarlas y el estado en que se encuentran las tareas.
3. Incremento, funcionalidad terminada lista para desplegarse en un ambiente productivo.

Dentro del marco de trabajo de Scrum se desarrollan las fases anteriormente mencionadas mediante prácticas, artefactos y roles por cada iteración o *Sprint*. En la siguiente figura se esquematiza a detalle las fases que se ejecutan en Scrum:

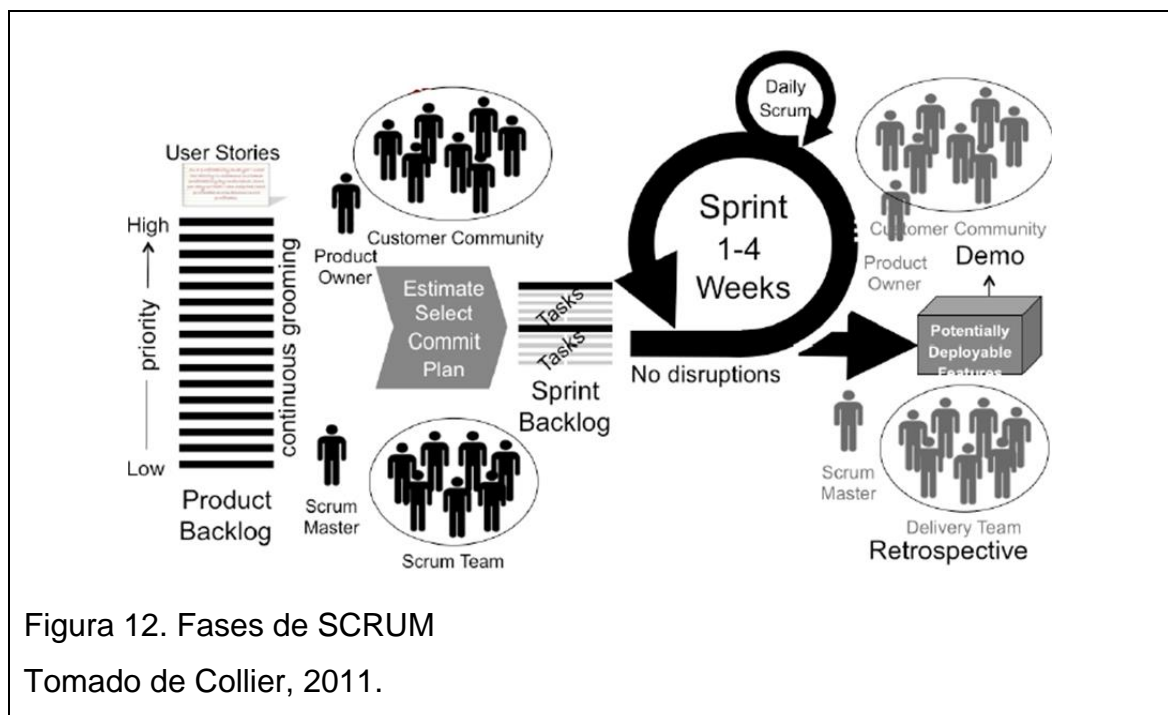


Figura 12. Fases de SCRUM

Tomado de Collier, 2011.

Collier (2011) detalla los pasos que intervienen en el marco de trabajo de Scrum, los cuales se los puede resumir en los siguientes puntos:

1. Un proyecto inicia con los clientes, interesados y el dueño del producto creando el *product backlog* o funcionalidades deseadas.

2. El *product backlog* consiste de historias de usuario, que representan fragmentos de funcionalidad requeridas por el cliente o el usuario.
3. El equipo junto con el *scrum master* estiman las historias utilizando varias técnicas como el de tamaño relativo.
4. Al inicio de cada *Sprint* el equipo selecciona pocas historias para desarrollar en base a su capacidad.
5. El equipo analiza cada historia y desarrolla una lista detallada de tareas que conforman el *Sprint Backlog*.
6. Una vez que el *Sprint* está en marcha el equipo debe ser protegido de cambios externos e interrupciones.
7. Cada día el equipo realiza la reunión de seguimiento de *Sprint* la cual no debe durar más de 15 minutos.
8. Cada resultado del *Sprint* debe tener una funcionalidad lista que pueda ser desplegada en producción.
9. El *product owner* con la ayuda del *scrum master* y del equipo muestra las características del producto obtenido del *Sprint* a los clientes e interesados.
10. Finalmente el equipo junto al *scrum master* y *product owner* participan en una retrospectiva para revisar los resultados y evaluar lo que se realizó mal para mejorar en el siguiente sprint.

En este apartado se resumen las características principales del marco de trabajo Scrum que serán consideradas para establecer el marco de referencia propuesto en este trabajo.

2.5.2 Kanban

Es otro marco de trabajo ágil que se basa en los mismos valores y principios del Manifiesto Ágil y que pretende "... producir exactamente aquella cantidad de trabajo que el sistema es capaz de asumir...". (Bermejo, 2010, p. 8)

Se adapta al proceso actual de desarrollo de software y es utilizado especialmente en entornos de mantenimientos. Utiliza herramientas para la visualización del proceso lo que permiten gestionar el esfuerzo de un equipo de trabajo.

Una de las ventajas a destacar de este marco de trabajo es que permite entender el actual proceso de desarrollo de software, identificar la capacidad de cada parte del proceso, identificar los cuellos de botella y determinar o establecer la mejora continua al proceso, apoyado de un entorno de visualización para el equipo y para los interesados, lo que genera una mayor colaboración de trabajo.

Kanban establece que se debe limitar el trabajo a realizar WIP en cada fase del proceso, y solo se empieza una nueva tarea cuando el trabajo anterior haya sido entregado al siguiente paso de la cadena de producción (Kniberg, Skarin, Poppendieck, & Anderson, 2010). A continuación se resumen las principales características de este marco de trabajo:

1. Visualiza el flujo de trabajo, divide al proceso de trabajo en bloques que representan las fases del proceso. Por cada fase se debe mostrar las tareas que se realizan las cuales deben estar claramente identificadas.
2. Limita el WIP, especifica el número de tareas que el equipo puede atender o estar en progreso en cada fase del proceso.
3. Establece el tiempo de ciclo, determinar el tiempo promedio de atención por cada tarea en cada fase del proceso. Se pretende optimizar el proceso de tal forma que el tiempo de ejecución de una tarea sea el mínimo requerido.

Al igual que en Scrum, Kanban también establece un *sprint backlog* donde se especifican las historias de usuario las cuales son efectuadas. En la siguiente figura se esquematiza un ejemplo del flujo de trabajo de este marco de trabajo:

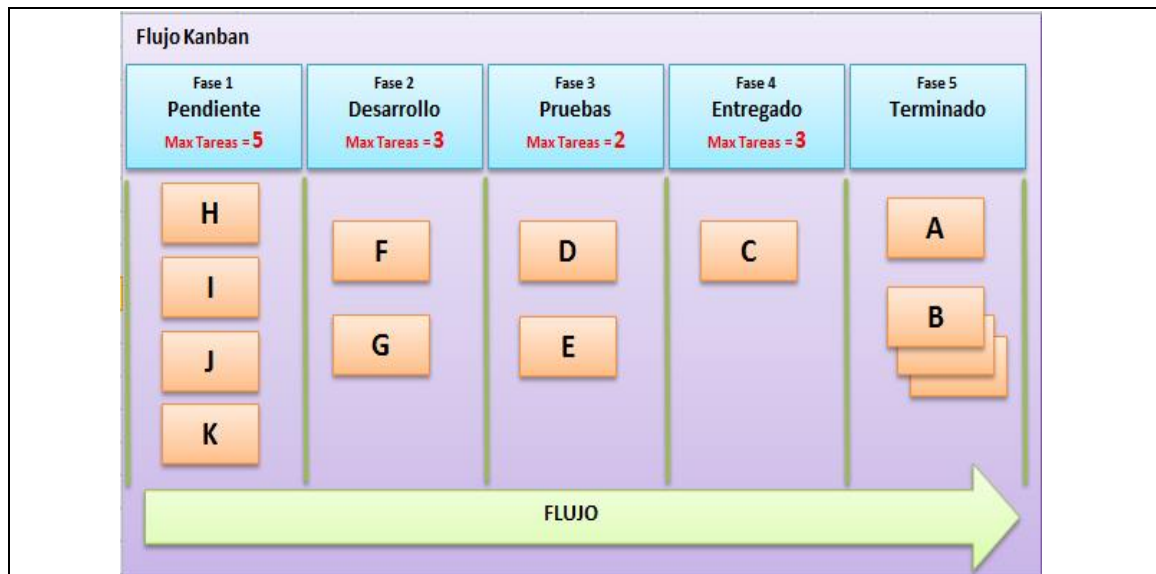


Figura 13. Flujo Kanban

Adaptado de Kniberg, Skarin, Poppendieck, y Anderson, 2010, p. 5.

Como se observa en la figura, el flujo está descompuesto en 5 fases en donde cada una de ellas se establece el máximo de tareas que pueden ser atendidas, a excepción de la última fase que corresponde a las tareas culminadas. El flujo va de izquierda a derecha empezando con las tareas que están pendientes por realizar, continúa con las tareas que deben ser desarrolladas, posteriormente probadas, entregadas y finalmente terminadas.

En cada fase se deben establecer los criterios que deben cumplir las tareas para cambiar de estado o de fase, los que deben estar en conocimiento de todo el equipo.

Bermejo (2010) detalla los pasos que se deben realizar para utilizar Kanban en un proceso de desarrollo de software. A continuación se resumen los mismos:

1. Visualizar en un tablero las fases generales del proceso y las tareas del equipo, especificando el número promedio de tareas que pueden ser atendidas en cada fase. Esta será la primera aproximación del esfuerzo que el equipo podrá realizar, pero en cada iteración se debe establecer un valor más real de tareas que puedan ser atendidas. Al separar las fases

se puede identificar cuellos de botella en el proceso, lo que permite proponer mejoras y optimizaciones sobre el mismo.

2. Analizar las fases del tablero en cada iteración por el equipo para detectar fases que falten o sobren.
3. El equipo debe actualizar el tablero de forma constante con el objetivo de visualizar el estado actual de las tareas en cada fase.
4. Se debe asegurar que el equipo realice únicamente el trabajo que puede atender por cada fase. No es recomendable que los miembros del equipo realicen multitareas, ya que puede provocar desperdicios cada vez que el programador cambia de tarea. Hay un tiempo de entendimiento o asimilación del contexto de la tarea considerado como desperdicio.
5. Identificar los cuellos de botella lo que permite al equipo proponer mejoras en cada iteración.
6. Finalmente, en el tablero se deben reflejar las tareas que van culminando las cuales deben haber pasado por los criterios de aceptación del cliente.

Kanban permite a un equipo de desarrollo continuo (mantenimiento de software) tomar las tareas o requerimientos de los clientes en la medida que el equipo pueda atenderlos. Visualiza el flujo de trabajo lo que proporciona al jefe o gerente del proyecto medir y gestionar el flujo. Optimiza el proceso en cada ciclo y mejora los entregables finales más eficientemente permitiendo obtener productos a satisfacción del cliente.

A diferencia con Scrum, en Kanban no existen iteraciones, el tablero nunca se limpia, no se establecen roles o no son obligatorios y se establece el número máximo de tareas a realizar en cada fase del proceso de desarrollo.

3. CAPITULO III SITUACIÓN ACTUAL DEL ÁREA DE DATAWAREHOUSE

3.1 Introducción

En este capítulo se explica brevemente la situación actual del Área de Datawarehouse del SRI, con el objetivo de describir la arquitectura de información de alto nivel sobre la cual se realizan los procesos de explotación de datos, roles y responsabilidades que se identifican en el área, herramientas que apoyan al desarrollo de aplicaciones de BI e identificación de tipos de requerimientos de necesidades de información que son atendidas por el Área de Datawarehouse, así como estadísticas de tiempos de implementación y número de requerimientos atendidos.

También se presentan los resultados de las encuestas realizadas, así como la identificación de los principales inconvenientes en el proceso actual de desarrollo de productos de DWH, con la finalidad de sustentar las justificaciones de este proyecto de tesis.

El análisis de la situación actual permite establecer una línea base sobre la cual se especifiquen actividades que permitan desarrollar la propuesta de un marco metodológico para el desarrollo de soluciones de inteligencia de negocios empleando metodologías ágiles.

Mediante resolución No. DSRI-028-2008 (Servicio de Rentas Internas, 2008) el SRI reforma el Reglamento Orgánico Funcional, en el cual se establece la creación del Área de Business Intelligence (BI) y Soluciones Regionales como parte del Departamento de Desarrollo e Implementación de la Dirección Nacional de Desarrollo Tecnológico, cuya principal función consiste en planificar, coordinar e implementar soluciones que permitan la explotación de la información que requiera el negocio para el análisis y toma de decisiones.

En el año 2014, mediante el Registro Oficial No. 134 (2014), la institución ejecuta una nueva reestructuración orgánica, de un enfoque funcional a una gestión organizacional por procesos. En esta evolución, el Área de BI cambia de nombre a Área de Datawarehouse y se establecen los servicios y productos que deben generar, lo cuales son los siguientes:

- Análisis, diseño e implementación de modelos físicos para explotación de información.
- Extracción, transformación y carga de datos internos y externos.
- Mantenimiento y actualización de modelos físicos para explotación de información.

3.2 Arquitectura de Data Warehouse DWH

Antes de exponer la arquitectura de DWH del SRI se describen algunos conceptos de la arquitectura de datos definidos en el marco de referencia TOGAF, que tienen relación sobre la gestión de los datos y pueden aportar al análisis del diseño actual de la arquitectura de DWH.

Según Monsalve (2014), TOGAF define a la arquitectura de datos como la estructura de datos lógicos y físicos que posee una organización y los recursos de gestión sobre los mismos. Sobre esta arquitectura se identifican los flujos de información, los modelos analíticos y gerenciales de la información para la toma de decisiones, sistemas de gestión de datos maestros y elementos de seguridad de la información. Para el efecto, se listan las tareas que se deben considerar para elaborar una arquitectura de datos:

1. Identificar la situación actual de la arquitectura de datos.
2. Definir la situación futura o el referente al cual se desea alcanzar.
3. Identificar las brechas entre la situación actual y la situación futura.
4. Priorizar las actividades que se deben realizar para alcanzar la situación futura.

5. Elaboración del plan que permita alcanzar el objetivo deseado.

El objetivo de la arquitectura de datos es analizar y evaluar los modelos de información analíticos y gerenciales requeridos por una organización para la toma de decisiones efectivas.

El alcance de la arquitectura de datos (Monsalve, 2014) se resume en los siguientes puntos:

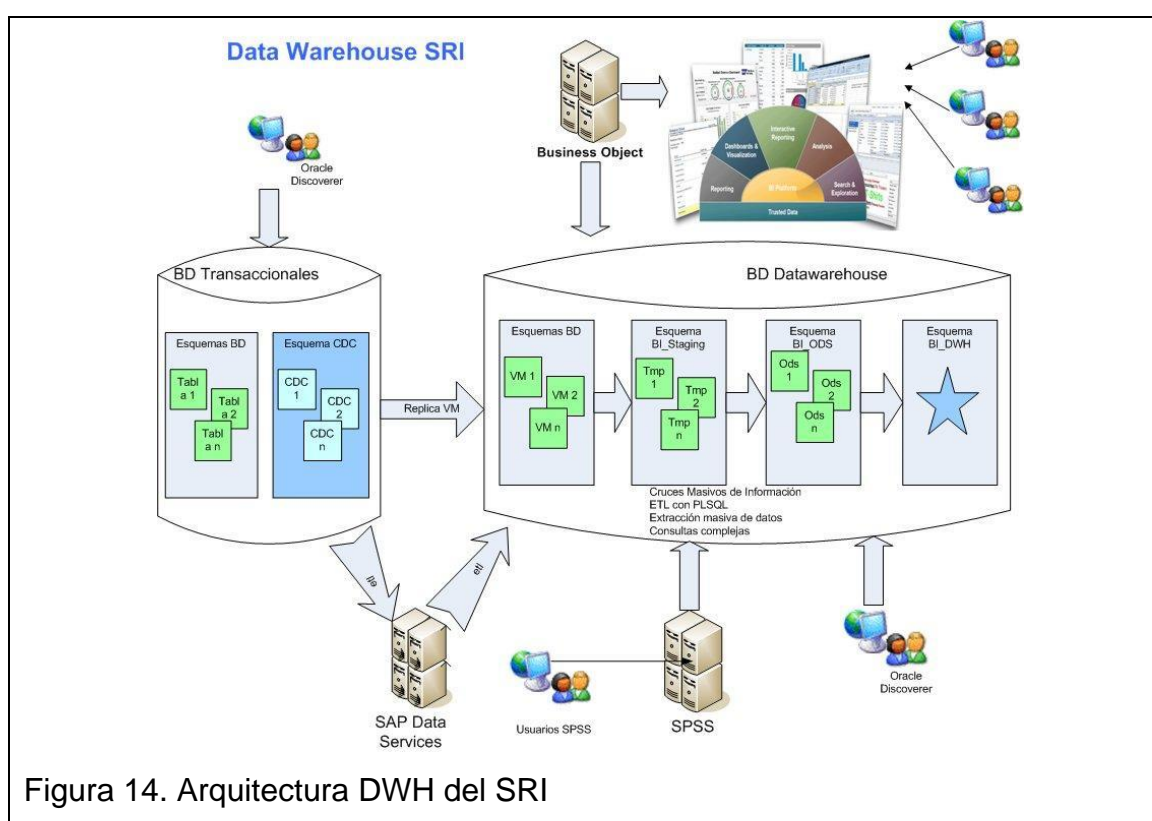
- Identificar los repositorios de las aplicaciones.
- Revisar las necesidades de información gerencial.
- Análisis de contenido de los repositorios y la consistencia de la información.
- Identificar fuentes de información, análisis de duplicidad, vulnerabilidad y consistencia.
- Garantizar la seguridad, confidencialidad, integridad y disponibilidad de la información.
- Establecer el gobierno del dato, niveles de sensibilidad de la información, administradores, ciclo de vida, almacenamiento de la información.

Como parte de la arquitectura tecnológica que posee el SRI, se encuentra desarrollada la de DWH, cuyo objetivo es integrar los datos resultantes de numerosas aplicaciones transaccionales alojados en varios repositorios de base de datos Oracle hacia un único repositorio, en el cual se realicen tareas de procesamiento masivo, de análisis y publicación de la información, que permita realizar procesos de toma de decisiones a nivel operativo, analítico y estratégico de la organización.

Cabe mencionar que a la fecha de elaboración de esta tesis, no existe una arquitectura de DWH formalizada para la institución, por lo que la coordinación de DWH ha solicitado al departamento de Planificación de TI del SRI, encargado de plantear los lineamientos arquitectónicos para la construcción de soluciones

informáticas, que gestione el desarrollo, formalización y socialización del estándar de arquitectura de DWH, en base a los elementos que actualmente son utilizados en la implementación de soluciones de DWH y que son identificados y analizados a continuación.

En la siguiente figura se esquematiza la arquitectura de DWH que maneja informalmente el SRI, la cual muestra los siguientes elementos: gestores de base de datos, servidores de ETL y servidores de publicación que conforman dicha arquitectura.



Las aplicaciones transaccionales registran los datos de sus respectivos procesos los cuales son almacenados en bases de datos Oracle configuradas para este propósito.

Dependiendo de las necesidades de análisis de los usuarios del negocio, existen herramientas de publicación que acceden directamente a las bases transaccionales, lo que en algunas ocasiones provoca impacto sobre dichas

bases, debido a consultas complejas efectuadas por los mismos. En los casos en que los usuarios de negocio tienen definido sus necesidades de información, se replican únicamente aquellos datos transaccionales requeridos para el análisis, y dependiendo del tipo de análisis, los datos son transformados y clasificados en esquemas de la base de datos mediante las herramientas de ETL.

A continuación se detallan los esquemas que se identifican en la arquitectura actual de DWH en el SRI:

- Esquema de base de datos: Aloja la réplica de los datos transaccionales sin ningún proceso de limpieza o transformación, y únicamente aquellos atributos requeridos para el análisis posterior.
- Esquema BI_staging: Área temporal donde se realizan procesos de limpieza y perfilamiento sobre los datos como: análisis de duplicidad, integridad y consistencia de los datos. Es utilizado como parte de los procesos ETL para la transformación de los datos antes de cargarlos a las estructuras de almacenamiento de destino.
- Esquema BI_ods: Aloja un primer nivel de información destinada al análisis operativo de las aplicaciones transaccionales.
- Esquema BI_dwh: Almacena los *data marts* y cubos de información utilizados para la toma de decisiones a nivel analítico y estratégico del negocio.

La información almacenada en estos esquemas es publicada en diferentes herramientas de análisis que son utilizados por los usuarios del negocio para realizar la toma de decisiones de sus procesos. En el siguiente apartado se detalla cada una de estas herramientas.

De acuerdo a lo descrito anteriormente, se puede concluir que la arquitectura actual de DWH del SRI cumple con la línea base establecida en la arquitectura de datos propuesta por TOGAF, lo cual habilita la generación de una metodología para desarrollo de productos de Data Warehouse.

3.3 Herramientas

El Área de Datawarehouse posee tres grupos de herramientas utilizadas para el desarrollo de soluciones de BI, las cuales se clasifican de la siguiente forma:

- Herramientas de base de datos.
- Herramientas de ETL.
- Herramientas de publicación y explotación de información.

A continuación se detalla cada uno de los grupos de herramientas.

3.3.1 Herramientas de base de datos

Corresponde a las bases de datos encargadas de almacenar la información procesada por las herramientas ETL construidas por el Área de Datawarehouse, dichas bases están configuradas con el software de gestor de base de datos Oracle, las cuales se encuentran a la fecha con la versión 11g, con proyección a migración a la versión 12c. El Área de Datawarehouse realiza sus actividades de construcción sobre las siguientes bases de datos:

3.3.1.1 Base de datos DWH

Es la primera base de datos del SRI destinada para consolidar varias fuentes de información y albergar modelos multidimensionales con el objetivo de proveer información para el análisis de los usuarios del negocio.

Entre los principales modelos se encuentran los ODS de Declaraciones, los cuales integran de manera adecuada la información de declaraciones de formularios presentadas por los contribuyentes. Estos ODS son utilizados por varios procesos de la organización como control de inconsistencias, diferencias, omisidad, entre otros.

Otro modelo a destacar es el cubo de Recaudación, el cual consolida en un modelo estrella los hechos de recaudación registrados por la Administración Tributaria, ofreciendo la opción de análisis desde varias dimensiones o perspectivas de la organización.

También se puede nombrar el ODS de Transaccionalidad, el cual integra varias fuentes de datos referentes a la información del contribuyente, como por ejemplo el ranking de ingresos y egresos de los mismos. Este ODS es utilizado en procesos de análisis de control y riesgos tributarios.

La base de datos DWH del SRI, pese a ser una base relacional con configuraciones específicas para proveer las funcionalidades propias de Data Warehouse, no cumple todas las características necesarias para su objetivo principal, debido a que aloja modelos multidimensionales y modelos relacionales, convirtiéndose en una base de datos híbrida.

3.3.1.2 Base de datos STAR

Esta base de datos fue creada con el objetivo de atender dos principales necesidades:

1. Replicar los datos transaccionales a un repositorio (BD STAR) que sirva como fuente de datos para los reportes *ad hoc*, los cuales afectan el desempeño de las bases de datos transaccionales.
2. Poseer una base de datos con características potentes de procesamientos masivos de cruces de información transaccional para la ejecución de los procesos ETL y generación de reportes.

Los datos que se almacenan en esta base sirven de fuente para la construcción de los modelos ODS y modelos estrella, así como para las herramientas de explotación de información.

3.3.1.3 Base de datos Frontera

El SRI es una institución del Estado que cuenta con datos útiles para otras organizaciones gubernamentales. A la fecha, estos datos son requeridos y compartidos directamente desde sus bases de datos transaccionales. En tal sentido, el SRI está implementando una base de datos denominada “Frontera” a la cual se replican los datos a ser compartidos a las instituciones externas, con el objetivo de mitigar problemas de seguridad.

3.3.2 Herramientas de ETL

La adopción de herramientas ETL ha sido un proceso de evolución sobre las necesidades de explotación de los datos transaccionales por parte de los usuarios del negocio. A continuación se detallan los pasos de evolución y las herramientas que se han adoptado en los últimos años:

1. Inicialmente no se construían procesos de ETL, simplemente se publicaban los datos transaccionales en la herramienta Oracle Discoverer, el cual es utilizado por los usuarios para la generación de los reportes *ad hoc* requeridos.
2. Al incrementarse los datos transaccionales, los usuarios del negocio evidencian complicaciones para explotar la información directamente desde las fuentes transaccionales. Un claro ejemplo es la cantidad de declaraciones receptadas por la Administración Tributaria, específicamente para la declaración de IVA, la cual posee más de 400 campos que son registrados a detalle. Se construyen los primeros ETL con código de programación estructurada PL SQL propio de la base de datos Oracle, permitiendo extraer el detalle de las declaraciones y trasponer dicho detalle en una estructura física de la base de datos denominada ODS, facilitando un acceso más rápido a las declaraciones del contribuyente.

3. En el licenciamiento de la base de datos Oracle se provee de la herramienta Oracle Warehouse Builder, la cual contiene funciones gráficas para la construcción de ETL, por lo que se empieza a utilizar dicha herramienta para construir algunos modelos estrella. Sin embargo al realizar una comparación entre un ETL construido en la herramienta de Oracle Warehouse Builder y PL SQL se observa que en este último, el desempeño en cuanto a procesamiento y velocidad son mejores. En tal sentido se descarta dicha herramienta y se inicia el proceso de adquisición de herramientas más eficientes para las tareas de procesamiento de información requeridas por los usuarios del negocio del SRI.
4. Finalmente se adquiere una solución para la construcción de ETL proporcionada por la casa comercial SAP, llamada Data Services, la cual proporciona varias funcionalidades de calidad de datos, perfilamiento y transformadores que permiten procesar los datos en modelos ODS o estrella de una manera eficiente. Actualmente las herramientas de SAP son institucionalmente autorizadas para realizar las tareas de extracción, transformación y carga.

3.3.3 Herramientas de Publicación y Explotación de Información

Al igual que las herramientas de ETL, las herramientas de publicación han pasado por un ciclo de madurez. A continuación se detalla su evolución:

1. Inicialmente dentro de la construcción de los aplicativos transaccionales se contemplaba la construcción de reportes fijos, sin embargo el costo de mantenimiento de los reportes era alto debido a los cambios frecuentes sobre los mismos.
2. Al ver la necesidad de reportes *ad hoc* se opta por contratar el licenciamiento de la herramienta Oracle Discoverer, la cual provee una funcionalidad para crear reportes a partir de las fuentes transaccionales. En ese momento no existía problemática en la publicación de fuentes correspondientes a las tablas de los modelos transaccionales, por lo que

se utilizó de manera masiva: para cada nueva aplicación se creó un área de negocio, lo que desencadenó problemas de rendimiento en las bases transaccionales.

3. La herramienta Oracle Discoverer es una aplicación cliente – servidor que corresponde a una tecnología obsoleta, por lo que se realiza un proceso de contratación donde se adquieren herramientas de BI para la publicación de información. En este proceso se adquieren las herramientas de Business Object de la casa comercial SAP. Actualmente conviven las herramientas de Oracle Discoverer y Business Object, el primero para la publicación de datos transaccionales, y el segundo para la publicación de modelos multidimensionales.
4. Debido a las necesidades de análisis predictivo de información, se han ejecutado algunos proyectos de minería de datos por lo que la Administración Tributaria identificó la necesidad de contratar el licenciamiento de la herramienta SPSS de IBM, entre cuyas funcionalidades está realizar predicciones sobre los datos transaccionales y modelos multidimensionales mediante modelos y algoritmos matemáticos.

La figura siguiente muestra los servidores en los cuales se alojan las herramientas de base de datos, herramientas de ETL y herramientas de publicación y explotación de información:

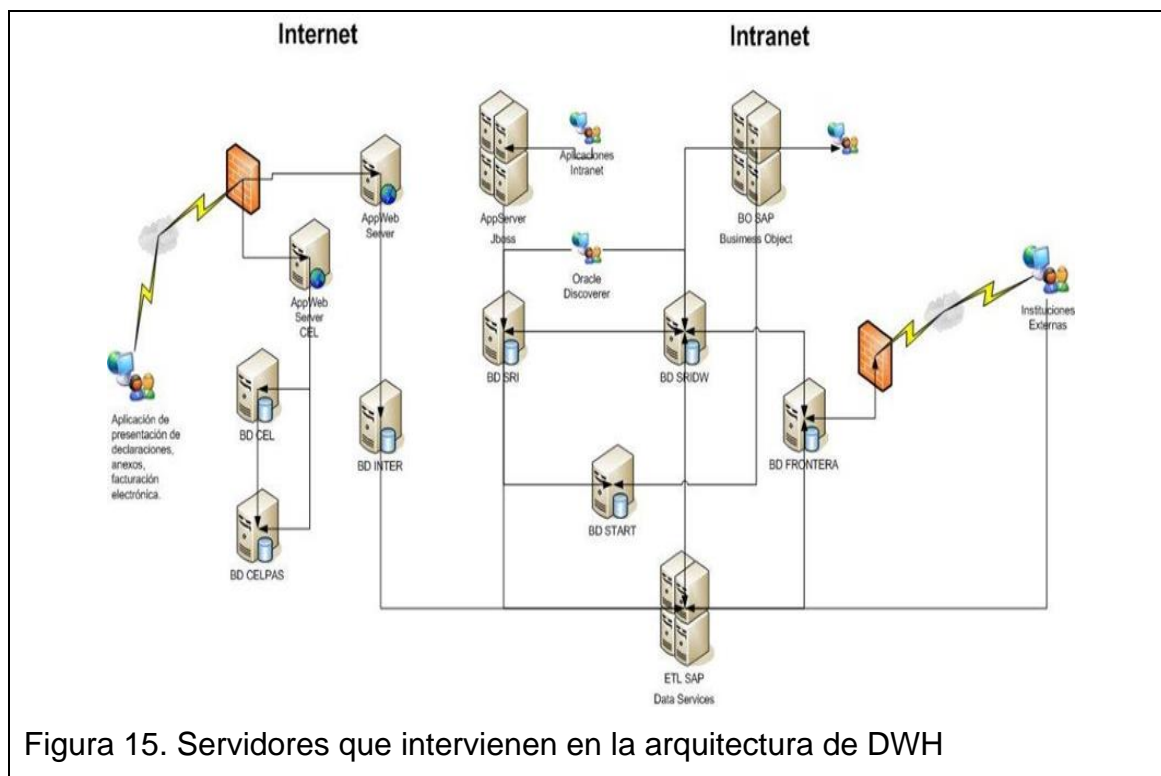


Figura 15. Servidores que intervienen en la arquitectura de DWH

3.4 Roles

El Área de Datawarehouse es parte del Departamento de Desarrollo de la Dirección Nacional de Tecnología del SRI, en la cual se identifican los siguientes roles:

- **Coordinador.-** Corresponde al jefe del Área. Entre sus actividades se pueden mencionar: distribuir las tareas a cada miembro del equipo, dar seguimiento a las actividades que debe cumplir el Área, eliminar los impedimentos que surjan en el desarrollo de las actividades, entre otras.
- **Administrador de las herramientas SAP.-** Responsable de monitorear la ejecución de los procesos ETL, velar por el cumplimiento de los estándares de desarrollo de ETL, asignación de permisos a los usuarios de negocio que requieran acceso a las herramientas de publicación.
- **Administrador de la herramienta Oracle Discoverer.-** Encargado de mantener las actuales Áreas de Negocio existentes y crear nuevas Áreas de Negocio.

- Construcción de ETL.- Personal especializado en el procesamiento de datos mediante la utilización de herramientas ETL, ya sean desarrolladas mediante código PL SQL o mediante las funcionalidades de la herramienta Data Services de SAP.
- Construcción de reportes y tableros de control.- Personal especializado en el desarrollo de la capa conceptual de los datos que para el caso de la herramienta de Business Object se denomina Universos y que para la herramienta de Oracle Discoverer se llama Áreas de Negocio. A partir de esta capa se crean los reportes y tableros de control requeridos por los usuarios de negocio.
- Proyectos de BI.- Personal especializado en gestionar los proyectos de BI, para lo cual se requiere la participación de los roles descritos en los puntos anteriores como parte del equipo encargado de desarrollar los proyectos.

3.5 Atención de Requerimientos del Negocio

El Área de Datawarehouse atiende diversos requerimientos de información solicitados por usuarios de la Administración Tributaria respecto al pedido, administración o procesamiento de información. A continuación se detalla los tipos de requerimientos identificados en los últimos años:

- Descarga de información.
- Creación de nuevas Áreas de Negocio en Oracle Discoverer.
- Creación de nuevos Universos en Business Object SAP.
- Construcción de Indicadores.
- Automatización de cruces de información.
- Construcción de modelos multidimensionales.
- Convenios con instituciones externas.

Inicialmente los requerimientos eran solicitados a la Dirección de Tecnología por medio de memorándums donde los usuarios del negocio describían las

necesidades de información de manera general. Dichos requerimientos no estaban enmarcados en un formato de levantamiento de necesidades de información, por lo que muchas veces eran necesarias varias reuniones para detallar y entender el requerimiento solicitado.

La manera como se gestionaban los requerimientos no permitía generar estadísticas sobre los mismo como: requerimientos atendidos, pendientes de atención, tiempo estimado de atención, entre otros aspectos que permitan establecer y parametrizar métricas de seguimiento adecuadas.

La Dirección de Tecnología, siguiendo las buenas prácticas del marco de referencia ITIL, adoptó y socializó este método para garantizar una entrega eficaz de los servicios que ofrece a la organización. Para lo cual adquirió la herramienta HP Service Manager (HPSM) con el objetivo de organizar de manera adecuada los servicios a cargo de TI en base al estándar de ITIL.

La herramienta HPSM ha permitido facilitar el registro y monitoreo de forma más organizada de los requerimientos que el negocio ha solicitado al Área de Datawarehouse. Por lo tanto se ha clasificado a los requerimientos en los siguientes grupos:

- Incidentes, corresponde a la interrupción no planeada de una aplicación o servicio en ambiente de producción que está siendo utilizado por el negocio.
- Requerimientos estándar, corresponden a aquellos requerimientos que son considerados como tareas específicas, cuya complejidad no es alta y no demanda mayor tiempo para su atención, por lo general la atención de este tipo de requerimientos no supera cinco días laborables de esfuerzo. Las actividades que se realizan en este tipo de requerimientos no afectan a los ítems o elementos de configuración existentes en ambientes de producción. En este grupo se encuentra los siguientes requerimientos:
- Descarga de información, solicitados por los usuarios del negocio cuando no pueden descargar grandes volúmenes de información.

- Cambios en la publicación de Áreas de Negocio de Oracle Discoverer o Universos de Business Object que no conllevan mayor complejidad.
 - Réplicas de información en las bases de datos Oracle.
 - Otras tareas adicionales que se le delegue al Área de Datawarehouse como generación de documentación, reuniones de trabajo, entre otros.
-
- Cambios, corresponde a la creación de una nueva aplicación o un servicio, o modificaciones a aplicaciones o servicios ya existentes. La complejidad de este tipo de requerimientos es mayor por lo que debe pasar por un comité de cambios encargado de autorizar y priorizar el desarrollo de los mismos. El tiempo de implementación de un cambio no puede ser mayor a 6 semanas laborables, caso contrario es considerado como proyecto (Dirección Nacional de Desarrollo Tecnológico del SRI, 2014). En este grupo se pueden considerar los siguientes requerimientos:
 - Creación y mantenimiento de Áreas de Negocio de Oracle Discoverer.
 - Creación y mantenimiento de Universos de Business Object SAP.
 - Automatización de cruces de información.
 - Construcción de indicadores de un proceso específico de complejidad baja.
 - Implementación de convenios con instituciones externas.
 - Creación de modelos multidimensionales específicos.
-
- Proyecto, conlleva un mayor esfuerzo para construir el requerimiento del usuario. Por lo general intervienen recursos de otras áreas que apoyan a la construcción del mismo. Se requieren recursos especialistas de base de datos para configurar adecuadamente los procesos de carga, recursos de control de calidad para certificar la nueva aplicación o servicio, recursos de desarrollo para construir los ETL, publicar las Áreas de

Negocio y Universos, y recursos para desplegar la solución en ambientes de producción. En este grupo se pueden considerar los siguientes requerimientos:

- Creación de modelos multidimensionales, construcción de ODS y modelos estrella.
- Automatización de cruces de información.
- Creación de nuevas Áreas de Negocio y Universos con complejidad alta.
- Construcción de indicadores de complejidad alta.

A continuación se muestran estadísticas de atención de acuerdo al grupo de requerimientos identificados correspondiente a los últimos dos años, las cuales son obtenidas de la herramienta HPSM:

Tabla 2. Estadísticas incidentes cerrados últimos 2 años

Año Registro	Prioridad	Tiempo Promedio	Número de Incidentes
2014	P2	11 días	2
2014	P3	5 días	33
2014	P4	14 días	28
2015	P1	19 minutos	1
2015	P3	5 días	1
2015	P4	6 días	12

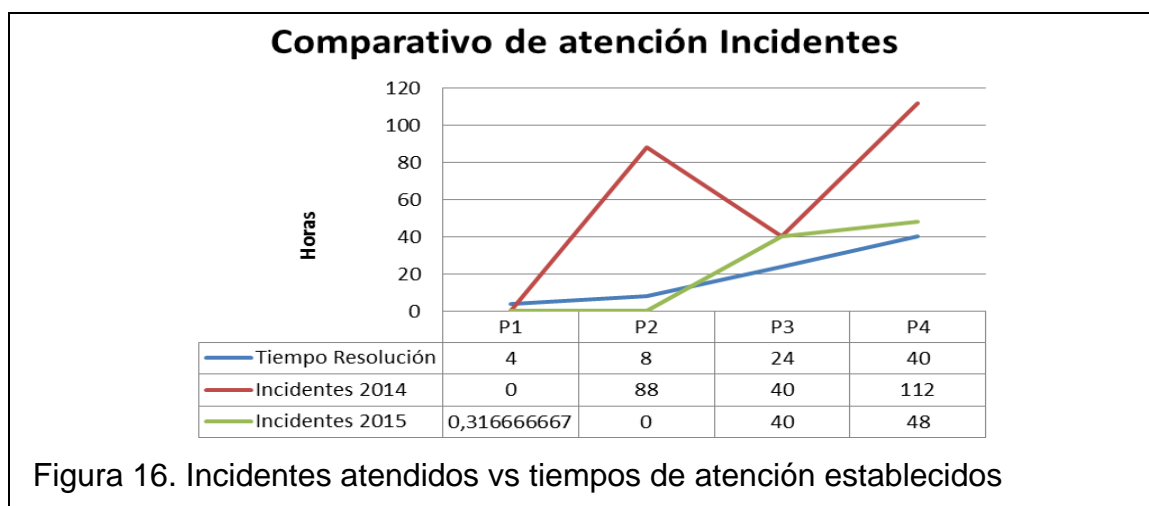
La Dirección de Tecnología ha establecido los niveles de atención de los servicios tecnológicos ofrecidos, los cuales rigen desde el año 2015. En la siguiente tabla se muestra los tiempos mínimos de resolución de incidentes de acuerdo a su prioridad:

Tabla 3. Tiempos resolución de incidentes

Prioridad	Horario	Tiempo de Resolución Desarrollo
P1 – Crítica	Horario Normal	4 horas
P2 – Alta	Horario Normal	8 horas
P3 – Media	Horario Normal	24 horas
P4 – Baja	Horario Normal	40 horas

Como se observa en la tabla 2, los tiempos promedios de resolución de incidentes para el año 2014 sobrepasan los actuales tiempos comprometidos por la Dirección de Tecnología. Esto se debe a que en el año 2014 el Área de Datawarehouse no había sido parte de los acuerdos de nivel de servicio. Sin embargo para el año 2015, los incidentes de prioridad crítica fueron resueltos dentro de los tiempos comprometidos.

En la siguiente figura se muestra el comparativo entre los tiempos de atención de incidentes atendidos en los dos últimos años frente a los tiempos de atención acordados por la Dirección de Tecnología:



Por ejemplo, observando el anterior gráfico, un incidente P4 catalogado con prioridad baja, tiene un tiempo establecido de resolución de 40 horas (línea azul). Para el año 2014 el promedio de atención de incidentes fue de 112 horas (línea roja). Y para el año 2015 el promedio de atención de incidentes fue de 48 horas

(línea verde), evidenciándose una disminución en los tiempos de atención de incidentes.

En la siguiente tabla se muestra las estadísticas de atención sobre los requerimientos atendidos por el Área de Datawarehouse, las mismas son obtenidas de la herramienta HPSM:

Tabla 4. Número requerimientos atendidos años 2014 - 2015

Mes Registro	Número de Requerimientos 2014	Tiempo estimado de cierre 2014 (horas)	Número de Requerimientos 2015	Tiempo estimado de cierre 2015 (horas)
1	15	28,27	40	5,58
2	16	32,26	39	21,71
3	19	8,34	58	10,81
4	31	23,02	41	3,57
5	18	32,87	45	13,91
6	25	27,25	42	12,14
7	28	21,41	64	6,39
8	22	30,68	45	6,10
9	28	13,42	52	5,67
10	38	12,36	67	4,05
11	50	10,93	38	3,70
12	38	9,98	41	2,72

Como se observa en la tabla anterior el número de requerimientos en el año 2015 ha duplicado el número de requerimientos atendidos en comparación del año 2014. Esto se debe a que a partir del segundo semestre del año 2014 se regularizó el registro de requerimientos atendidos por el Área de Datawarehouse, y en el año 2015 ya es una política institucional en la que todo requerimiento de usuario sea registrado en la herramienta HPSM.

En la siguiente figura se visualiza el aumento del número de requerimientos atendidos y la disminución de tiempos de atención de los mismos en comparación al año anterior.

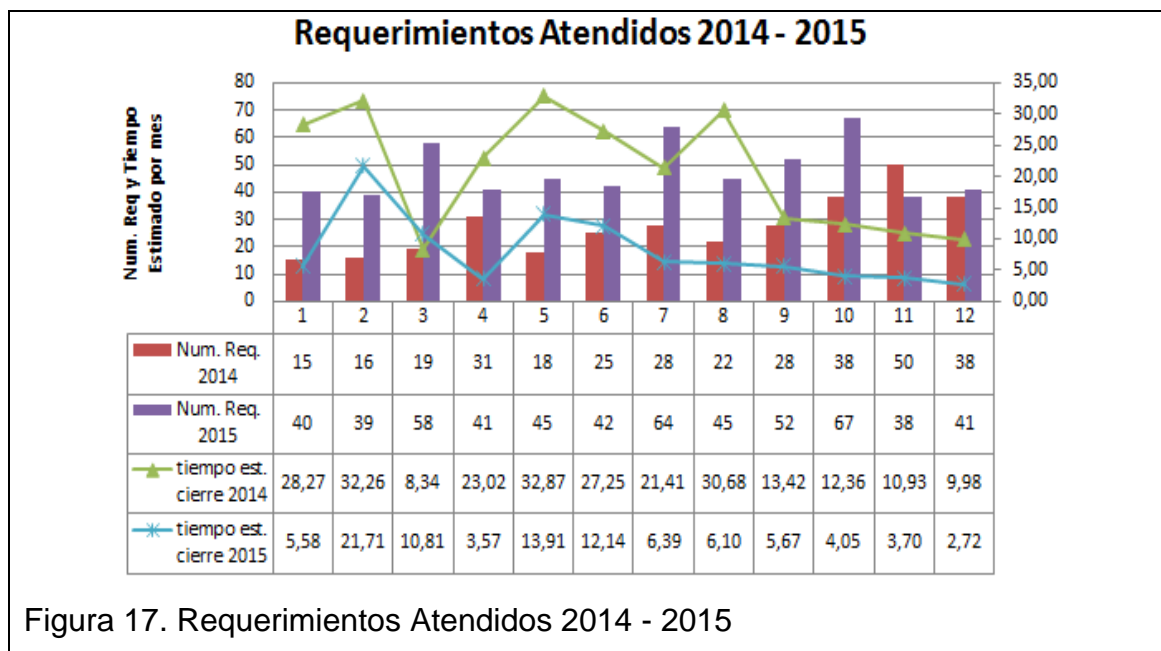


Figura 17. Requerimientos Atendidos 2014 - 2015

Por ejemplo, se observa que en el mes de julio el número de requerimientos para el año 2014 fue de 28 (barra roja), y para el año 2015 fue de 64 (barra morada). Así también se observa que para ese mes el tiempo promedio de cierre de los 28 requerimientos atendidos en el año 2014 fue de 21,41 horas (línea verde), y el tiempo promedio de cierre de los 64 requerimientos atendidos en el año 2015 fue de 6,39 horas (línea celeste), evidenciándose una disminución en los tiempos de resolución de los requerimientos.

Debido a las actuales limitaciones de la herramienta HPSM, no se dispone del tiempo real de atención de los requerimientos, sin embargo se dispone de la fecha de registro y cierre del requerimiento los cuales son utilizados para obtener una estimación sobre el tiempo aproximado de atención a los mismos.

En la siguiente tabla se muestra las estadísticas de atención sobre los cambios atendidos por el Área de Datawarehouse, las mismas son obtenidas del sistema HPSM:

Tabla 5. Número cambios atendidos años 2014 - 2015

Mes Registro Producción	Número Cambios 2014	Tiempo promedio atención 2014 (días)	Número Cambios 2015	Tiempo promedio atención 2015 (días)
1	4	5,55	1	1,00
2	2	11,50	2	63,00
3	8	24,49	0	0,00
4	2	81,50	2	5,35
5	6	12,68	3	59,00
6	7	14,33	3	10,33
7	4	15,58	2	1,00
8	6	43,06	2	28,50
9	4	7,10	1	15,00
10	5	41,14	1	1,00
11	0	0,00	5	25,88
12	1	29,00	3	15,24

Al comparar el número de cambios desplegados en producción entre los años 2014 y 2015 se observa una disminución en el número de cambios atendidos. En cuando al tiempo promedio de atención, en algunos meses disminuye pero en otros aumenta, sin embargo se observa que existen tiempos mayores a 30 días tanto en el año 2015 como en el 2014.

En la siguiente figura se visualiza gráficamente la tendencia de los cambios en número y en tiempo de atención:

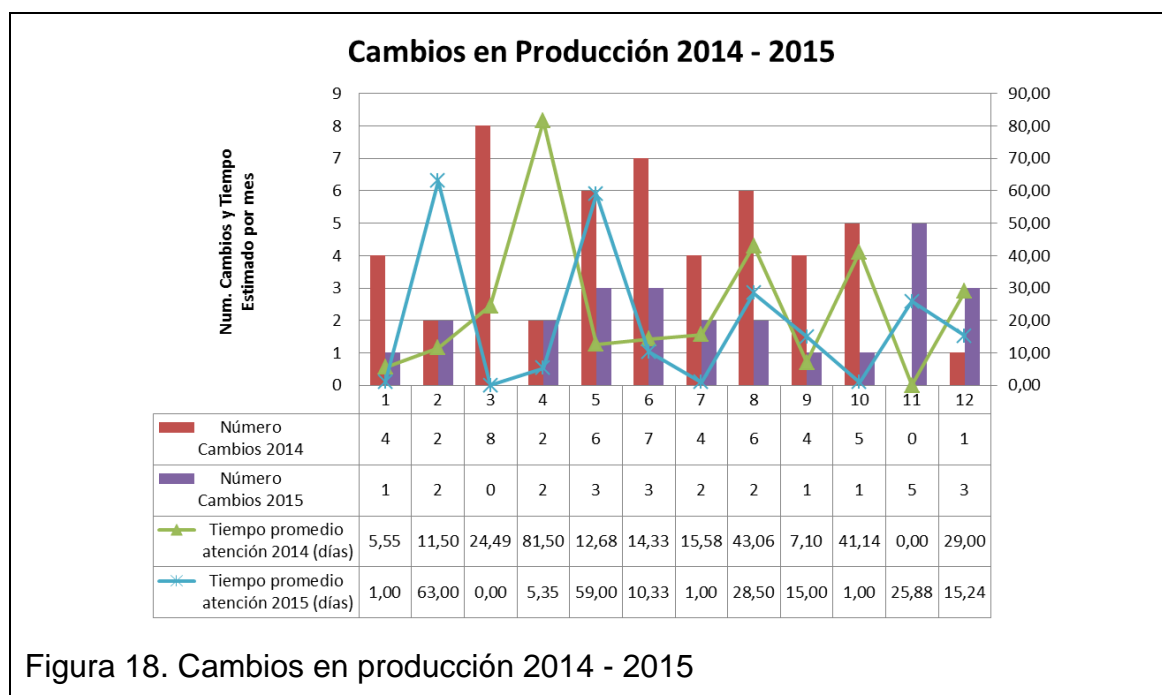


Figura 18. Cambios en producción 2014 - 2015

Como se observa en los meses de agosto y octubre del 2014, en promedio la atención de una solicitud de cambio está entre 43 y 41 días respectivamente, es decir los usuarios del negocio esperaron ese tiempo para tener a disposición información y poder realizar sus actividades de análisis sobre la información generada. Lo mismo sucede en el año 2015 donde en el mes de febrero y mayo la atención de los cambios sobrepasa los 30 días de atención. Esto podría ocasionar que la información no esté disponible a los usuarios en el momento adecuado para la toma de decisiones.

En cuanto a las estadísticas de proyectos de BI se puede destacar cuatro proyectos registrados en la Oficina de Proyectos del SRI, cada uno de los cuales por lo general consta de dos fases de implementación. En promedio cada fase ha tenido un tiempo estimado de implementación de un año. Entre los proyectos a destacar están:

- Proyecto de Controles Extensivos, Fase I y II.
- Proyecto de Contribuyentes Inexistentes, Fase I y II
- Proyectos de Seguridad de la Información, Fase I y II
- Proyecto de construcción de Indicadores.

Para los proyectos mencionados se ha observado que los usuarios han obtenido resultados al final de cada fase de los mismos, esto quiere decir que los usuarios han esperado alrededor de un año para poder disponer de información que les permita realizar análisis sobre sus procesos de control.

En el siguiente capítulo se propone una metodología que permita atender los tipos de requerimientos identificados de una manera más ágil permitiendo realizar entregas a los usuarios finales en tiempos más cortos.

Otra problemática identificada es la poca documentación técnica generada, en donde se detallen los diseños lógicos, físicos, de ETL y de las aplicaciones de BI implementadas, lo que causa demoras al momento de realizar las tareas de mantenimiento sobre los productos de DWH. Esto se debe principalmente a que no existe un procedimiento que norme la generación de la documentación de los productos implementados.

3.6 Inconvenientes proceso actual Desarrollo de Productos DWH

Del análisis descriptivo realizado al proceso actual que ejecuta el Área de Datawarehouse para el desarrollo de productos de DWH se identifican los principales inconvenientes:

- Existen productos generados directamente en el ambiente de producción, que a pesar de utilizar los elementos de arquitectura disponibles, no son correctamente probados, corriendo el riesgo de generar información no adecuada la cual es posteriormente publicada a los usuarios del negocio.
- Al desarrollar directamente en ambientes de producción existe el riesgo de generar indisponibilidad a otros productos relacionados.
- El desarrollo se lo realiza sin utilizar adecuadamente los estándares de nomenclaturas de programación de DWH y tampoco son controlados por el Área de Control de Calidad.

- No existe un control adecuado de calidad sobre los procesos de carga o ETL ya que el Área de Control de Calidad desconoce cómo realizarlos y no posee un ambiente de certificación en el cual realizarlos.
- La implementación de un cambio tecnológico con componente de DWH es realizado sin seguir un procedimiento formal donde se especifiquen los pasos a realizar para la implementación de la solución.
- El tiempo requerido en realizar el análisis de la información disponible hasta redactar el requerimiento de la necesidad de información suele ser extenso, esto se debe a que los usuarios no tienen claro sus necesidades de información.
- Para el levantamiento de requerimientos no existen formatos donde los usuarios describan sus necesidades de información.
- En algunos casos no se realizan diseños de los modelos lógicos y en los casos que sí se generan, no son documentados y solo quedan en conocimiento de las personas que realizan el desarrollo.

De los puntos mencionados se observa la clara necesidad de especificar un procedimiento que permita normar el proceso de desarrollo de productos de DWH.

3.7 Encuesta situación actual

Con el objetivo de evaluar el nivel de servicio que ofrece el Área de Datawarehouse y justificar este proyecto de tesis, se ha elaborado una encuesta la cual ha sido dirigida hacia 18 funcionarios responsables del negocio, los cuales principalmente han generado algún tipo de requerimiento al Área de Datawarehouse en el último año. En el anexo uno, se detallan las preguntas formuladas así como las opciones de respuesta para cada una de las mismas.

Las tabulaciones de la encuesta se encuentran detalladas en el anexo dos en donde se puede observar los resultados obtenidos.

A continuación se detallan las conclusiones obtenidas de la tabulación de la encuesta:

- En base a las preguntas 2, 3, 4, 5, 6 y 7 se concluye que: el 100% de los encuestados conoce del Área de Datawarehouse, conoce de manera general cuál es su finalidad en la Institución y ha generado algún requerimiento al Área, de los cuales 1 usuario contestó que el requerimiento generado no fue atendido por falta de recurso humano, cambios de prioridad en la Institución, cambio de personas asignadas, lo que representa un 6% de la muestra.
- De acuerdo a la pregunta 9, el 50% de los encuestados señalan que los requerimientos enviados al Área de Datawarehouse son atendidos con demora y aun así el requerimiento era necesario para el negocio, lo cual sustenta la justificación de esta tesis respecto a los tiempos de implementación altos y la información no disponible de manera oportuna a los interesados de las soluciones de BI.
- De acuerdo a las preguntas 9 y 10, el 33% de los encuestados señalan que los requerimientos son atendidos dentro del tiempo adecuado y esperado, y el 11% restante responde que dependiendo del funcionario que atienda el requerimiento es atendido a tiempo, así como también puede depender de la complejidad del requerimiento. El 6% restante no aplica a esta pregunta.
- La pregunta 11 trata de medir la complejidad que tienen los usuarios en entender las problemáticas de la gestión de información, y transmitir las mediante requerimiento al Área de Datawarehouse. Se observa que para el 53% de los encuestados es complejo y medianamente complejo, lo que evidencia la necesidad de establecer un procedimiento para el levantamiento de requerimientos.
- En base a la pregunta 12, en lo que se refiere a la información disponible y actualizable de los cubos de información, el 61% de los encuestados señalan que es adecuada y un 33% señalan que no es la adecuada pero es útil para las actividades. De esta manera también se sustenta que la

información no es oportuna a los interesados. El 6% restante no aplica a esta pregunta.

- La pregunta 13 hace referencia sobre, sí los productos proporcionados por el Área de Datawarehouse cumplieron con las expectativas, el 17% de los encuestados están muy satisfechos, un 67% están satisfechos y el resto poco satisfechos. Al elaborar un marco metodológico se esperaría mejorar este indicador y subir a más de un 60% los usuarios muy satisfechos.
- En base a las preguntas 15, 16 y 17, sobre la atención a incidentes en el último año, el 56% de los encuestados han experimentado la caída de un servicio provisto por el Área de Datawarehouse, sobre los cuales el 90% menciona que la atención para la solución del mismo fue adecuada, de los cuales el 22% menciona que fue solventado en más de 8 horas laborables. A lo que se concluye que de manera general el Área está atendiendo adecuadamente los incidentes reportados.
- En la pregunta 18 sobre la inconsistencia de la información presentada en los cubos de información, el 61% de los encuestados menciona que han detectado inconsistencias por lo que el marco metodológico debe contemplar realizar control de calidad de los productos a implementar.
- En base a la pregunta 19 se concluye que los requerimientos más comunes y con mayor prioridad son:
 1. Creación de reportes con complejidad alta.
 2. Creación y publicación de información para la toma de decisiones.
 3. Descargas de información masiva.
 4. Cruces de información masiva.
- De acuerdo a la pregunta 20, se evidencia que para el 100% de los encuestados es muy importante los servicios o productos que ofrece el Área de Datawarehouse por lo que el marco metodológico tiene como objetivo generar productos que apoyen a las necesidades del resto de áreas de la Institución.

- En base a las preguntas 21, 22 y 23 referente a la información entregada por el Área de Datawarehouse, es valorada por los encuestados como excelente un 44% y bueno un 56%. La información que se genere a partir de la aplicación de un marco metodológico deberían mejorar estos indicadores. Se debe considerar las oportunidades de mejora sobre los datos inconsistentes, desempeño de consultas, accesos a la información, entre otros.
- La pregunta 24 identifica que el 72% de los encuestados se consideran usuarios de nivel avanzado por lo que las soluciones de DWH que se construyan deben satisfacer las necesidades de este grupo de usuario.
- La pregunta 25 identifica que el 72% de los encuestados usan diariamente las herramientas de BI por lo que se vuelve importante proveer de soluciones de BI adecuadas para las actividades de los usuarios.
- En base a las preguntas 26, 27 y 28, el 50% de los encuestados señalan que las capacidades de las herramientas de BI no satisfacen las necesidades requeridas para su trabajo, debido a que las mismas no ofrecen interfaces más amigables, no son tan fáciles de usar, así como tiempos de procesamientos altos, entre otros.

4. CAPITULO IV PROPUESTA METODOLÓGICA

4.1 Introducción

El desarrollo de la propuesta metodológica tiene como objetivo establecer el marco de trabajo para cada una de las fases del ciclo de vida de desarrollo de soluciones de DWH, para ello se han planteado los siguientes hitos:

1. Considerar la estructura de un marco metodológico.
2. Seleccionar la metodología de DWH para este proyecto de tesis.
3. Establecer las prácticas de metodologías ágiles a considerar.
4. Definir el procedimiento de desarrollo de productos de DWH.
5. Sustentación del procedimiento en un escenario futuro.

4.2 Estructura del Marco Metodológico

Como se mencionó en el literal 2.3 (Definición de Marco Metodológico), la propuesta metodológica a desarrollar, consiste en descomponer el proceso de desarrollo de DWH en sub procesos o fases, identificar las tareas a ejecutar en cada fase y establecer los procedimientos para su ejecución, para lo cual se cuenta con herramientas y técnicas pre establecidas. (Ver figura 4 para revisar los elementos de un marco metodológico).

La propuesta metodológica tiene como entregable el Procedimiento de Desarrollo de Productos de DWH el cual está desarrollado en base a la estructura metodológica mencionada en el párrafo anterior.

4.3 Selección de la Metodología de DWH

El marco metodológico propuesto en este trabajo de tesis está basado en el Ciclo de Vida de DWH de Kimball y Ross (2013). A continuación se establece los

motivos por los cuales se ha seleccionado esta metodología (ver lit. 2.4.2 Metodología de Ralph Kimball):

- Se basa en modelos multidimensionales permitiendo a los usuarios un acceso más fácil y comprensible a la información.
- Los tiempos de respuesta de las consultas sobre este tipo de modelos son óptimos.
- Esta metodología plantea entregables a corto plazo e iterativos, lo que se acopla a los marcos de trabajo ágiles.
- Es utilizada en la mayoría de industrias desde la década de los noventa por lo que se considera una metodología madura.
- Establece claras fases para la construcción de un DWH las cuales se acoplan a las actividades de desarrollo de DWH que se ejecutan en el SRI.

El proceso de desarrollo de DWH según Kimball y Ross (2013) se descompone en las siguientes fases las cuales son consideradas en el marco metodológico propuesto:

- Definición de Requerimientos del Negocio.
- Diseño de la Arquitectura Técnica.
- Selección de Productos.
- Modelo Dimensional.
- Diseño Físico.
- Diseño y Construcción del ETL.
- Especificaciones y Desarrollo de Aplicaciones BI.
- Implementación de la Solución.

Para cada una de estas fases se establecen las actividades y tareas que se deben realizar al momento de construir la solución de DWH, las cuales se detallan en el Procedimiento de Desarrollo de Productos de DWH, así como también se establecen técnicas para el levantamiento de requerimientos, y

herramientas informáticas que apoyan a la ejecución de las tareas establecidas en el procedimiento.

4.4 Prácticas de Metodologías Ágiles a considerar

Para acoplar el marco metodológico que se plantea, se consideran las buenas prácticas propuestas por el marco de trabajo Scrum. Entre los principales puntos a considerar están:

- Redactar las historias de usuario entre los requirentes del negocio y los técnicos que desarrollan la solución. El usuario requirente deberá especificar los criterios de aceptación de cada historia de usuario.
- Las historias de usuario deben estar priorizadas y estimadas. La priorización la establece el negocio y la estimación la estable los técnicos.
- La estimación de las historias de usuario se la realiza con puntos de historias mediante la escala Fibonacci donde el valor 1 es la historia de menor complejidad y el valor 13 es de mayor complejidad. Se utiliza la técnica del tamaño relativo.
- El grupo de historias de usuario priorizadas y estimadas se denomina *backlog*, lo que permite tener una visión general de lo que solicita el negocio.
- Se establece el tamaño de cada iteración la cual puede durar entre 1 a 4 semanas.
- De acuerdo a datos históricos referente a la atención de requerimientos similares, se establece la velocidad media del equipo, y de acuerdo al total de puntos estimados del total de las historias de usuario del *backlog* se establece el número de iteraciones y por ende el total de semanas requeridas.
- Se planifica cada una de las iteraciones, especificando las historias de usuario que se desarrollarán en cada iteración.
- Se ejecutan cada una de las iteraciones en el orden establecido en el *backlog*. En cada iteración se establecen las tareas técnicas que se deben

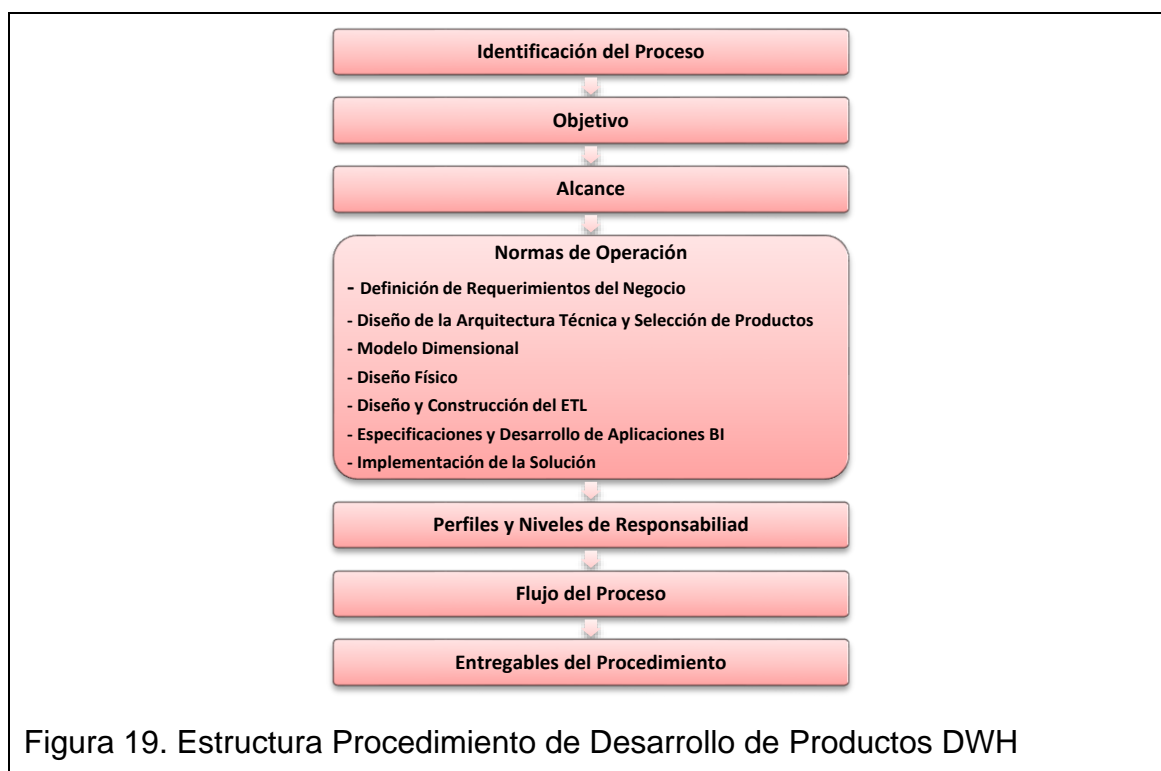
ejecutar y se estiman las horas que se requiere para su ejecución. Esto permitirá dar un seguimiento al avance de la iteración.

- Al final de cada iteración se revisa junto con el requirente del negocio el cumplimiento de los criterios de aceptación que fueron definidos previamente, lo que permitirá desplegar en producción las funcionalidades aceptadas.
- Finalmente se realiza una retrospectiva de la iteración con los miembros que conforman el equipo para revisar los resultados y evaluar lo que se realizó mal y mejorar en la siguiente iteración.

Los puntos mencionados son actividades que permiten aplicar las prácticas de agilidad las cuales son consideradas para elaborar el procedimiento de DWH.

4.5 Procedimiento de Desarrollo de Productos de DWH

El Procedimiento de Desarrollo de Productos de DWH junta las fases del ciclo de vida de desarrollo de DWH propuesta por Ralph Kimball y Margy Ross, así como las características ágiles planteadas por Scrum. En la siguiente figura se especifican los puntos a considerar en la elaboración del procedimiento:



A continuación se realiza una descripción detallada de cada uno de los puntos mencionados en la figura anterior, describiendo las actividades que se deben considerar para el desarrollo de productos de DWH.

4.5.1 Identificación del Proceso

En la siguiente tabla se identifican los procesos de apoyo del SRI, hasta tercer nivel, en donde se aplicará el uso del procedimiento propuesto.

Tabla 6. Identificación Proceso DWH

Proceso de Primer Nivel:	Gestión Tecnológica e Innovación		
Proceso de Segundo Nivel:	Transición de Servicios Tecnológicos		
Proceso de Tercer Nivel:	Gestión de Desarrollo		
Versión del Documento:	1.0	Frecuencia de Ejecución:	Bajo demanda
Responsable:	Departamento de Desarrollo		

4.5.2 Objetivo

Establecer el procedimiento que defina los lineamientos a seguir en cada una de las fases de desarrollo de productos de DWH, para crear de manera iterativa e incremental los entregables necesarios para la implementación de los requerimientos solicitados por los usuarios del negocio de una manera óptima, utilizando los estándares de desarrollo definidos por la Institución.

4.5.3 Alcance

El presente procedimiento involucra las fases de Definición de Requerimientos del Negocio, Diseño de la Arquitectura Técnica, Selección de Productos, Modelo Dimensional, Diseño Físico, Diseño y Construcción del ETL, Especificaciones y Desarrollo de Aplicaciones BI e Implementación de la Solución de acuerdo al modelo propuesto en el ciclo de vida de construcción DWH de Kimball y Ross (2013).

Cualquier requerimiento solicitado por la Institución respecto al pedido, administración o procesamiento de información se considera como un cambio tecnológico y serán gestionados de acuerdo al Procedimiento de Gestión de Cambios Tecnológicos (Dirección Nacional de Desarrollo Tecnológico del SRI, 2014). El resultado de la implementación del cambio es llamado también solución de información.

Este procedimiento aplica a cambios tecnológicos normales cuyo esfuerzo no supera a 30 días/hombre (días laborables), y cambios que estén relacionados a proyectos con componente de DWH. En cualquiera de los dos casos, la implementación de los cambios se realizará con desarrollo interno.

4.5.4 Normas de Operación

- a) Las tareas que se deben realizar en cada una de las fases deben ser gestionadas por el Coordinador del Área de Datawarehouse y en el caso de proyectos por el gerente del proyecto.
- b) La fase de Definición de Requerimientos así como el resto de fases de implementación de las soluciones de información se debe ejecutar en iteraciones.
- c) Las iteraciones deben tener una duración de tiempo fija en las cuales se desarrollan las actividades para la implementación de la solución.
- d) El recurso del Área de Datawarehouse responsable de la construcción de la solución, el cual en adelante se denominará técnico de DWH, debe definir la duración de cada iteración la cual podrá ser entre 1 a 4 semanas dependiendo del esfuerzo días/hombre requeridas para la implementación de la solución o cambio.
- e) Si un cambio tiene una estimación de implementación de 4 semanas, se debe definir iteraciones de 1 semana y planificar los entregables de cada iteración.
- f) En el caso de proyectos donde existe mayor número de entregables y los tiempos estimados de implementación son mayores a 6 semanas, las iteraciones deben tener una duración entre 2 a 4 semanas.
- g) Todo requerimiento o cambio con componente de DWH será registrado, analizado, autorizado, programado, implementado, evaluado y cerrado de acuerdo a lo establecido en el Procedimiento de Gestión de Cambios Tecnológicos.
- h) El usuario experto de la unidad requirente, el que aquí en adelante se le denominará Requirente, debe estar disponible y presto para aclarar cualquier duda o inquietud durante todas las fases de la implementación de la solución.
- i) Toda la documentación que genere la implementación del cambio o proyecto debe ser registrada en la herramienta de gestión documental.

4.5.4.1 De la definición de Requerimientos del Negocio

El objetivo de definir los requerimientos del negocio es levantar y comprender las necesidades de información requeridas, analizar y estimar el esfuerzo y tiempos de desarrollo e identificar y evaluar los riesgos de implementación.

Otro objetivo de esta fase es definir un documento donde se detallen los criterios de aceptación que debe cumplir la solución y sobre los cuales se pueda verificar su cumplimiento al momento de realizar las pruebas de certificación de la solución.

También es objetivo de esta fase obtener el entendimiento necesario que sirva de base para generar los diseños de solución que se realizan en las fases posteriores.

- a) El proceso de Definición de Requerimientos inicia con reuniones previas entre el Requirente, interesados de la unidad requirente, el técnico de DWH y el Experto del Área de Inteligencia de Información, con el objetivo de justificar o no el desarrollo del cambio tecnológico, entender la necesidad de información requerida, levantar y redactar las historias de usuario.
- b) En el caso de proyectos, la Oficina de Proyectos (ODP) deberá haber priorizado y aprobado la ejecución del mismo. En el caso de cambios tecnológicos, deberá existir un pedido de la necesidad de información por parte del Jefe de Área de la unidad requirente hacia el Coordinador del Área de Datawarehouse. El pedido inicialmente podrá ser un correo electrónico y posteriormente deberá ser formalizado con el ingreso de una solicitud de cambio (SDC).
- c) La primera reunión a realizarse debe ser convocada por el Requirente, en donde el técnico de DWH debe solicitar exponer los siguientes puntos los cuales serán útiles para entender la problemática a resolver:

- Realizar una explicación general sobre las necesidades de información requeridas.
 - Especificar la situación actual así como la problemática de no contar la funcionalidad que se solicita.
 - Describir los objetivos que persigue obtener.
 - Detallar los roles y responsabilidades de los usuarios que harán uso de la solución.
 - Explicar el proceso del negocio relacionado a la solución.
- d) Esta primera reunión no debe tomar más de 2 horas de la cual el Requirente realizará un acta de reunión donde se describan los principales puntos.
- e) En el caso de que la implementación de la necesidad de información sea viable, se establecerán las fechas de las próximas reuniones para elaborar el documento de Levantamiento de Requerimientos de Información.
- f) El Requirente será el encargado de registrar el documento de Levantamiento de Requerimientos de Información (Anexo 3), el cual debe contemplar toda la funcionalidad referente a procesamiento y publicación de información requerida en el formato de historias de usuario con la debida priorización correspondiente.
- g) Las historias de usuario deben detallar y especificar los reportes operativos, analíticos o estratégicos que el Requirente desea obtener como resultado final de la solución de DWH, incluyendo las funcionalidades que requiera como parámetros de entrada, navegación drill down, drill up, gráficos estadísticos, ente otros.
- h) Para elaborar el documento de Levantamiento de Requerimientos, el Requirente debe contar con el apoyo del técnico de DWH y el Experto de Inteligencia de Información.
- i) Es responsabilidad del técnico de DWH y el Experto de Inteligencia de Información guiar al Requirente en la redacción de las historias de usuario de tal forma que si las mismas son complejas y requieren de un esfuerzo

mayor al tiempo estipulado de la iteración, se debe descomponer en historias de menor tamaño.

- j) La estimación del esfuerzo para la construcción de las historias de usuario estará a cargo del técnico de DWH el cual podrá solicitar la colaboración de un recurso Experto técnico del Área de Datawarehouse que permita realizar una estimación más real.
- k) La estimación se la realiza con la escala Fibonacci donde la historia de usuario con menor complejidad es 1 y la de mayor complejidad es 13. Si se identifica que una historia de usuario va ser mayor a 13 puntos, se debe descomponer a la misma en entregables más pequeños. Se empieza estimando por una historia de usuario que sirva como referencia.
- l) La prioridad de cada historia de usuario es estimada por el Requirente de acuerdo a la clasificación especificada en la plantilla de Levantamiento de Requerimientos de Información (Anexo 3).
- m) El técnico de DWH debe especificar la velocidad media o número de puntos que se prevé ejecutar en una iteración. Este es un valor que se obtiene en base a la experiencia previa sobre otras historias de usuario ya realizadas y corresponde al promedio de puntos realizados por cada iteración en otros proyectos o cambios de similar esfuerzo. En caso de que no se tenga la certeza del valor de la velocidad media que se deba ingresar, se debe solicitar el criterio experto de algún recurso del Área de Datawarehouse que ya haya desarrollado similares historias.
- n) De acuerdo al total de puntos de historia y la velocidad media estimada se establece el número de iteraciones requeridas.
- o) El Requirente debe establecer la prioridad de atención de las historias de usuario y especificar claramente los criterios de aceptación de cada una de ellas las cuales servirán de insumo al momento de realizar el proceso de certificación de la solución.
- p) Se planifica cada una de las iteraciones especificando las historias de usuario que se desarrollarán en cada iteración.
- q) Se ejecutan cada una de las iteraciones en el orden establecido en el *backlog*. En cada iteración se establecen las tareas técnicas que se deben

- ejecutar y se estiman las horas que se requiere para su ejecución, esto permitirá dar un seguimiento al avance de la iteración.
- r) Al final de cada iteración se revisa junto con el requirente del negocio el cumplimiento de los criterios de aceptación que fueron definidos previamente lo que permitirá desplegar en producción las funcionalidades aceptadas.
 - s) Se realiza una retrospectiva de la iteración con los miembros que conforman el equipo para revisar los resultados y evaluar las actividades realizadas de manera errónea, de tal forma que se planifique mejorar en la siguiente iteración. Considerar los formatos que incluye este procedimiento para el registro del *backlog*, seguimiento de las iteraciones así como la revisión y retrospectiva de cada iteración especificada en las plantillas de Levantamiento de Requerimientos de Información (Anexo 3).
 - t) El tiempo requerido para elaborar el documento de Levantamiento de Requerimientos de Información, estimar y priorizar las historias de usuario y establecer los criterios de aceptación no debería ser mayor a 5 días laborables.
 - u) El documento debe ser revisado y aprobado por el Jefe Departamental de la unidad requirente, el usuario Requirente, el técnico de DWH y el Experto de Inteligencia de Información.
 - v) El Requirente debe registrar la solicitud de cambio en la herramienta de Gestión de Cambios Tecnológicos adjuntando el documento de Levantamiento de Requerimientos de Información con las debidas firmas de aceptación.
 - w) El técnico de DWH debe realizar un análisis de todas las historias de usuario registradas en el documento de Requerimientos de Información el cual tiene como resultado la identificación de los hechos o medidas y las dimensiones de análisis de las mismas.
 - x) Se debe verificar en la Matriz de Data Warehouse del SRI si existe o no las dimensiones y hechos identificados con la finalidad de evitar redundancia.

- y) La Matriz de Data Warehouse debe contener las dimensiones comunes utilizadas en el SRI y los hechos construidos por cada proceso de negocio, señalando la relación entre dimensiones y hechos. (Anexo 4).
- z) La Matriz debe estar disponible en la herramienta de gestión documental, la misma debe ser actualizada en la etapa del análisis del requerimiento por el técnico de DWH, ingresando los nuevos hechos relacionados a un proceso de segundo nivel, y las dimensiones identificadas.

4.5.4.2 Del Diseño de la Arquitectura Técnica y Selección de Productos.

Esta fase permite identificar los elementos de la arquitectura técnica requeridos para cumplir con la solución tecnológica, así como la selección y adquisición de los productos específicos acorde a las necesidades establecidas en el diseño de arquitectura.

- a) Luego de haber identificado las tareas técnicas de cada una de las historias de usuario previamente priorizadas y estimadas, el técnico de DWH junto con el responsable de la gestión de la capacidad tecnológica deben contrastar con la arquitectura de DWH y especificar si se cumplen los siguientes puntos:
 - ¿Se dispone de almacenamiento suficiente para procesar los datos requeridos?
 - ¿Las herramientas de ETL son suficientes para construir los procesos de carga?
 - ¿Las configuraciones o recursos de los servidores de base de datos y ETL son suficientes para procesar la cantidad de datos requeridos?
 - ¿Las herramientas de publicación para generar reportes analíticos y gerenciales son suficientes para construir las soluciones de BI requeridas?
 - ¿No se requiere de algún otro elemento técnico adicional para construir la solución?

- b) En caso de no cumplir con los puntos anteriores, se debe informar al Requirente los motivos por los cuales no se cumple la arquitectura y determinar si el proyecto tiene recursos monetarios para adquirir los elementos técnicos que permitan cumplir con el desarrollo de la solución.
- c) Si el proyecto dispone de recursos, el gerente del proyecto debe ejecutar el Procedimiento de Adquisiciones y solicitar al departamento de Planificación TI elaborar el documento de capacidad de infraestructura.

4.5.4.3 Del Modelo Dimensional

En esta fase se identifican los hechos con sus medidas y las dimensiones de análisis de los mismos.

- a) Los documentos de entrada de este proceso son: el documento de Levantamiento de Requerimientos de Información y la Matriz de DWH del SRI.
- b) Se debe identificar claramente el proceso o procesos de negocio a modelar en base al análisis de las historias de usuario.
- c) Se debe definir la granularidad esperada de las dimensiones y hechos, la cual corresponde al nivel de detalle que se espera almacenar la información en las estructuras de tipo estrella para ser analizada por los usuarios de negocio.
- d) Se debe identificar las dimensiones, sus atributos, jerarquías y orígenes o fuentes de datos. Una dimensión debe proporcionar el "quién, qué, dónde, cuándo, por qué y cómo" que rodean al evento del proceso de negocio.
- e) Se debe identificar los hechos y orígenes o fuentes de datos de los mismos. Un hecho debe proporcionar las medidas que resultan del evento del proceso de negocio.
- f) A partir de la identificación de los hechos, dimensiones y granularidad, se debe generar el modelo dimensional el cual debe constar de una tabla

- central de hechos con sus medidas, relacionada a las tablas de dimensiones. A este modelo también se le conoce como modelo estrella.
- g) El modelo dimensional debe ser generado en la herramienta CASE destinada para el caso.
 - h) Todas las tablas (hechos y dimensiones) deben contener los campos o atributos de las mismas. Se debe especificar por cada atributo el tipo de dato, longitud, clave primaria, clave foránea, los valores permitidos de ser el caso y el comentario que describa su uso.
 - i) Todas las tablas (hechos y dimensiones), deben contener en sus comentarios la fuente de donde proviene: instancia de base de datos, nombre del esquema, tabla y aplicativo transaccional asociado a la fuente.
 - j) Se debe realizar un proceso de perfilamiento de datos mediante ejecución de sentencias de consulta a las tablas transaccionales, que permita obtener estadísticas de los mismos, con el objetivo de entender los problemas derivados de los datos y que puedan modificar al modelo dimensional propuesto.
 - k) Para probar que el modelo dimensional pueda solventar las necesidades de los usuarios del negocio, se debe tomar cada una de las historias de usuario del requerimiento de información y verificar que el modelo pueda contestar o resolver cada una de ellas. Si existiera algún caso no contemplado se debe realizar la debida actualización al modelo.
 - l) La salida de este proceso es la Matriz de DWH del SRI actualizada con la descripción detallada de los hechos y dimensiones identificadas, el modelo dimensional lógico generado en la herramienta CASE y el análisis de perfilamiento a los datos fuentes.

4.5.4.4 Del Diseño Físico

En esta fase se identifican las tareas para crear el modelo dimensional en estructuras de almacenamiento, considerando las prácticas de indexación y particionamiento.

- a) La entrada a este proceso es la Matriz de DWH del SRI, el documento de estándares de base de datos de BI y el modelo dimensional lógico en la herramienta CASE.
- b) A partir del diseño lógico se debe construir las estructuras físicas que serán creadas posteriormente en la base de datos, para lo cual se debe tomar como referencia el documento de estándares de base de datos de BI, el cual especifica los lineamientos de nomenclatura, documentación y programación PL SQL.
- c) El modelo inicialmente debe ser construido en la base de datos de desarrollo de DWH, el cual será parte de la construcción y pruebas de los ETL que procesen los datos hacia el modelo.
- d) Se debe identificar los campos de indexación y particionamiento de las tablas de hechos que permitan optimizar los criterios de búsqueda, consultas o análisis sobre los datos almacenados.
- e) Las estructuras agregadas que se creen a partir de las tablas de hechos también deben considerar una estrategia de indexación y particionamiento.
- f) Por cada una de las estructuras especificadas en el modelo físico se debe estimar los espacios iniciales de almacenamiento tanto en datos como en índices, así como el crecimiento anual de las mismas.
- g) Se debe establecer los rangos o periodos de tiempo que la información estará almacenada en el DWH (información activa). Posteriormente la información deberá ser respaldada en dispositivos de almacenamiento como discos lentos o cintas magnéticas (información pasiva).
- h) Todos los objetos a construir deberán ser implementados en la base de datos STAR bajo los siguientes esquemas:
 - BI_STAGING, utilizado para tablas temporales.
 - BI_ODS, utilizado para tablas donde se almacena datos operacionales como cruces de información y ODS.
 - BI_DWH, utilizado para tablas de hechos, dimensiones, cubos de información e indicadores.
 - Para vistas materializadas se utilizará el mismo esquema dueño de las tablas fuentes.

- i) Las estructuras físicas deben ser registradas, documentadas y versionadas en la herramienta CASE.
- j) Los scripts de creación de objetos de base de datos deben ser probados y ejecutados previamente en el ambiente de desarrollo de DWH.
- k) Será responsabilidad del Área de Control de Calidad velar por el cumplimiento de los estándares de base de datos de BI que sean enviados a certificar.
- l) Para la implementación de los objetos, el responsable del Área de Centro de Cómputo que ejecuta los pasos a producción, debe tomar los scripts que genere la herramienta CASE.

4.5.4.5 Del Diseño y Construcción del ETL

En esta fase se detallan las tareas para el diseño y construcción de los procesos de carga hacia los modelos multidimensionales creados físicamente en la base de datos.

- a) Los documentos de entrada de este proceso son: el documento de Levantamiento de Requerimientos de Información, la Matriz de DWH del SRI, Diseño Físico implementado y perfilamiento de los datos fuentes.
- b) En base al estudio de las fuentes de datos, se debe determinar si existe algún atributo como una fecha de transacción relacionada al evento que pueda ser usado para las cargas incrementales, caso contrario se deberá optar por otra estrategia como el uso de CDC (*Change Data Capture*).
- c) Las fuentes transaccionales deben ser replicadas a la instancia de base de datos de DWH haciendo uso de herramientas ETL o soluciones específicas de la base de datos. En el caso de Oracle con el uso de vistas materializadas.
- d) Los ETL deben contener procesos de calidad y limpieza de datos que permitan obtener resultados adecuados.
- e) El diseño y construcción del ETL son actividades que se realizan a la par, ya que mientras se construye el ETL se genera el diseño del mismo.

- f) La construcción del ETL debe realizarse inicialmente bajo el ambiente de desarrollo de DWH y debe basarse en el documento de estándares de objetos de ETL.
- g) Como regla general el flujo de un ETL debe seguir el siguiente orden:
- Réplica de las fuentes transaccionales las cuales podrán ser vistas materializadas en el caso de Oracle.
 - Extracción y carga de datos transaccionales a tablas temporales. En este punto el ETL podrá contemplar tareas de limpieza y calidad de los datos fuentes.
 - A partir de datos temporales se realizan las transformaciones del caso y se carga la información procesada hacia los ODS o modelos estrella.
 - Para los procesos incrementales se recomienda tener una bandera en las tablas fuentes que permita identificar los registros cargados por día. En caso de no existir dicha bandera, se podrá gestionar al Área de Desarrollo la inclusión de la misma o hacer uso de CDC para registrar los cambios sobre las tablas fuentes, permitiendo al ETL realizar procesos de cargas incrementales.
- h) Los tiempos de ejecución de los procesos de carga incrementales no deberán ser mayores a una hora por día considerando una carga no mayor a un millón de registros, y máximo 4 horas considerando una carga no mayor a cuatro millones de registros. Como regla general se establece una hora de procesamiento por un millón de registros.
- i) En el ETL también se debe determinar la frecuencia de actualización de la información. Todos los ETL incrementales deben ejecutarse diariamente salvo algún tipo de inconvenientes en los datos que no permita dicha actualización.
- j) Se debe registrar en el sistema documental los horarios de ejecución, tiempos estimados de duración, nombre del producto con la finalidad de visualizar las ventanas de tiempo de los procesos de carga.

4.5.4.6 De las Especificaciones y Desarrollo de Aplicaciones de BI

En esta fase se analizan los reportes operativos, analíticos y estratégicos definidos en el Requerimiento de Información y se procede con el desarrollo de los mismos.

- a) Los documentos de entrada de este proceso son: el documento de Levantamiento de Requerimientos de Información, la Matriz de DWH del SRI, Diseño Físico implementado y el documento de estándares de aplicaciones de BI.
- b) En base al Requerimiento de Información se debe identificar una lista de los principales reportes que requiere el usuario de negocio para cubrir las necesidades especificadas en las historias de usuario, los cuales deben estar priorizados en el *backlog*.
- c) En la plataforma de BI del SRI se debe especificar una estructura de contenidos que permita agrupar y organizar los reportes que son construidos, así como los futuros reportes que generen y compartan los usuarios.
- d) Mediante las herramientas de BI institucionales se debe estructurar y construir la capa semántica, la cual relaciona el modelo físico implementado y la interfaz de usuario, permitiendo publicar el modelo multidimensional. A partir de esta estructura se deben construir los reportes y dashboard necesarios para el negocio.
- e) En cada uno de los reportes a construir debe existir una descripción detallada del objetivo del reporte, los parámetros de ingreso y los resultados o medidas que entregará el mismo. Esta descripción debe estar registrada en la opción de metadatos de la plataforma de BI.
- f) Los nombres de los reportes construidos deben manejar un estándar de nomenclatura de negocio que serán consensuados con el Requirente.
- g) Todos los reportes deben ser diseñados en base al documento de estándares de aplicaciones de BI, en la parte correspondiente a estándares de visualización.

- h) Para la construcción de una aplicación de BI se debe utilizar las herramientas institucionales disponibles para el efecto, así como hacer uso de las bondades de registro de metadatos para cada uno de los objetos que se construyan.
- i) Todos los reportes construidos deben ser probados, verificando su funcionalidad por parte del usuario, tiempos de respuesta en desplegar datos los cuales deben ser consistentes y coherentes.
- j) El Requirente debe validar y certificar que se han cumplido los criterios de aceptación, para lo cual se firmará el acta de certificación que permitirá dar paso a la siguiente fase de implementación.
- k) Se debe nombrar un responsable del negocio que será el encargado de la autorización de accesos a los reportes así como cambios posteriores que se requiera sobre los mismos.

4.5.4.7 De la Implementación de la Solución

En esta fase se detallan las tareas necesarias para implementar los diseños físicos de los modelos multidimensionales, los ETL y las aplicaciones de BI en un entorno de producción.

- a) La entrada de este proceso es el documento o acta de certificación aprobado por el Requirente.
- b) La implementación corresponde a los pasos a producción que deben registrarse, revisarse por el responsable del Área de Control de Calidad y ejecutarse en el ambiente de producción por el responsable del Área del Centro de Cómputo.
- c) Los pasos a producción deben ser ejecutados cuando se cuente con la certificación del usuario, los mismos se realizarán al final de cada iteración.
- d) Los pasos a producción de creación de objetos de base de datos así como de ETL deben ser revisados por el Área de Control de la Calidad. Los pasos a producción de las aplicaciones de BI deben ser revisados y

- ejecutados en producción por los recursos del Área de Datawarehouse encargados de administrar la plataforma de BI.
- e) Todos los pasos que requieran revisión del Área de Control de la Calidad deben ser registrados en el Sistema de Gestión TI donde se asociará a la solicitud de cambio.
 - f) Se debe dar seguimiento a la ejecución del paso por los responsables del Área de Centro de Cómputo, y verificar que se hayan ejecutado todos los elementos que conforman la solución de DWH.
 - g) El Requirente debe validar que la solución implementada se encuentre en correcto funcionamiento y cubra todos los puntos definidos en las historias de usuario del documento de Levantamiento de Requerimientos de Información.
 - h) Existirá un tiempo de estabilización de la solución, donde se monitoreará el correcto funcionamiento de los procesos de carga, así como las aplicaciones de BI implementadas.
 - i) Las tareas especificadas por cada una de las fases del ciclo de vida serán realizadas en cada una de las iteraciones definidas que contienen las historias de usuario priorizadas y estimadas.
 - j) Cuando se hayan culminado todas las iteraciones se realizarán las actas de cierre del cambio o proyecto, así como la encuesta de satisfacción por parte del Requirente sobre la calidad del servicio ofrecido por el o los Técnicos de DWH que intervinieron en todo el ciclo de vida de la construcción de la solución de DWH.
 - k) En caso de que haya existido cambios en la arquitectura de DWH, el responsable de la capacidad será el encargado de actualizar los documentos correspondientes a la arquitectura.
 - l) Toda la documentación del cambio o proyecto debe estar disponible en el sistema de gestión documental ya sea para procesos de auditoria como para que las personas encargadas de dar soporte y mantenimiento a la solución puedan ejecutar estas tareas.

4.5.5 Perfiles y Niveles de Responsabilidad

En la siguiente tabla se detalla los roles identificados que intervienen en el proceso de desarrollo de una solución de DWH, así como sus responsabilidades de acuerdo al perfil del puesto y sus competencias establecidas por la unidad de Talento Humano del SRI:

Tabla 7. Perfiles y Niveles de Responsabilidad

Rol o Cargo del Responsable	Nivel de Responsabilidad / Funciones
Coordinador del Área de Datawarehouse	<ul style="list-style-type: none"> • Gestiona la ejecución de proyectos y cambios con componente tecnológico. • Coordina la entrega de los servicios tecnológicos existentes referentes a aplicaciones o DWH de acuerdo a las necesidades Institucionales. • Supervisa la ejecución de las actividades de los procesos asignados a su cargo. • Planifica las actividades de los procesos asignados a su cargo. • Coordina la gestión de adquisiciones de bienes o servicios tecnológicos de proyectos a su cargo. • Controla la atención de observaciones, quejas y sugerencias relacionadas con los procesos a su cargo. • Revisa y define estándares, políticas y procedimientos. • Coordina y participa en la investigación de mejoras e innovación tecnológica. • Controla el cumplimiento de las políticas, procedimientos y estándares.
Gerente de Proyectos	<ul style="list-style-type: none"> • Las establecidas en los Procedimientos de Gestión de Proyectos.
Técnico de DWH	<ul style="list-style-type: none"> • Ejecuta actividades vinculadas a la implementación y mantenimiento de aplicaciones o de soluciones de DWH. • Resuelve y colabora en la solución de incidentes y problemas que se presentan en la funcionalidad de las aplicaciones o del DWH.

	<ul style="list-style-type: none"> • Procesa y entrega requerimientos de información al usuario cuya complejidad no puede ser resuelta por las herramientas de explotación de información. • Resuelve incidentes y problemas técnicos referentes a las herramientas de inteligencia de negocios que administra. • Implementa técnicamente la información externa que se genera a través de convenios del SRI.
<p>Experto técnico de DWH</p>	<ul style="list-style-type: none"> • Supervisa, diseña y ejecuta la administración técnica de los procesos ETL. • Supervisa, diseña e implementa modelos: dimensionales, ODSs, starting, staging y proyectos de explotación de información. • Asesora e implementa técnicamente la información externa que se genera a través de convenios de intercambio de información, acuerdos u otros instrumentos que el SRI gestiona para la obtención de la mencionada información. • Gestiona proyectos auspiciados por la Dirección Nacional de Tecnología e Innovación. • Propone metodologías y estándares de construcción de DWH. • Supervisa el registro de metadatos referente al contenido del DWH institucional. • Elabora informes de actividades efectuadas según se le solicite.
<p>Experto del Área de Inteligencia de Información</p>	<ul style="list-style-type: none"> • Actualiza los activos de información para ponerlos a disposición de los usuarios de la Institución. • Colabora en el levantamiento de Requerimientos de Información.
<p>Requirente</p>	<ul style="list-style-type: none"> • Las establecidas en los procedimientos de Gestión de Proyectos • Proporciona una descripción general del proceso de negocio que requiere de una solución de información. • Elabora y aprueba el documento de Levantamiento de Requerimientos de Información. • Establece los criterios de aceptación que debe cumplir la solución de DWH. • Participa en todas las fases del ciclo de vida de construcción de la solución junto con el técnico de DWH.

	<ul style="list-style-type: none"> • Certifica que la solución de DWH cumple con todos los criterios de aceptación. • Verifica que la solución de DWH se encuentra disponible y estable en ambientes de producción. • Establece los responsables de la solución de DWH que serán los encargados de dar accesos a la información.
Jefe de Área de la unidad requirente	<ul style="list-style-type: none"> • Las establecidas en los procedimientos de Gestión de Proyectos • Solicita al Coordinador del Área de Datawarehouse el desarrollo de una solución de DWH. • Revisa el documento de levantamiento de requerimiento de información.
Jefe Departamental de la unidad requirente	<ul style="list-style-type: none"> • Las establecidas en los procedimientos de Gestión de Proyectos • Revisa y aprueba el documento de Levantamiento de Requerimientos de Información. • Revisa y aprueba el documento de Solicitud de Cambios.
Responsable Área de Control de Calidad	<ul style="list-style-type: none"> • Verifica el cumplimiento de los estándares de objetos de base de datos y la ejecución sin errores de los procesos ETL contruidos en la herramienta Data Services. • Aprueba los pasos a producción enviados por la herramienta Gestión TI para que sean ejecutados por el responsable del Área de Centro de Cómputo.
Responsable Área de Centro de Cómputo	<ul style="list-style-type: none"> • Ejecuta en producción los scripts señalados en el paso a producción desde la herramienta CASE Oracle Designer.
Responsable de Planificación de TI	<ul style="list-style-type: none"> • Realiza el análisis de capacidad tecnológica requerida por la solución de DWH a desarrollar.

4.5.6 Flujo del Proceso

En la siguiente figura se muestra el flujo general de las fases del procedimiento propuesto para el Desarrollo de Productos de DWH con los responsables que intervienen en cada una de las fases.

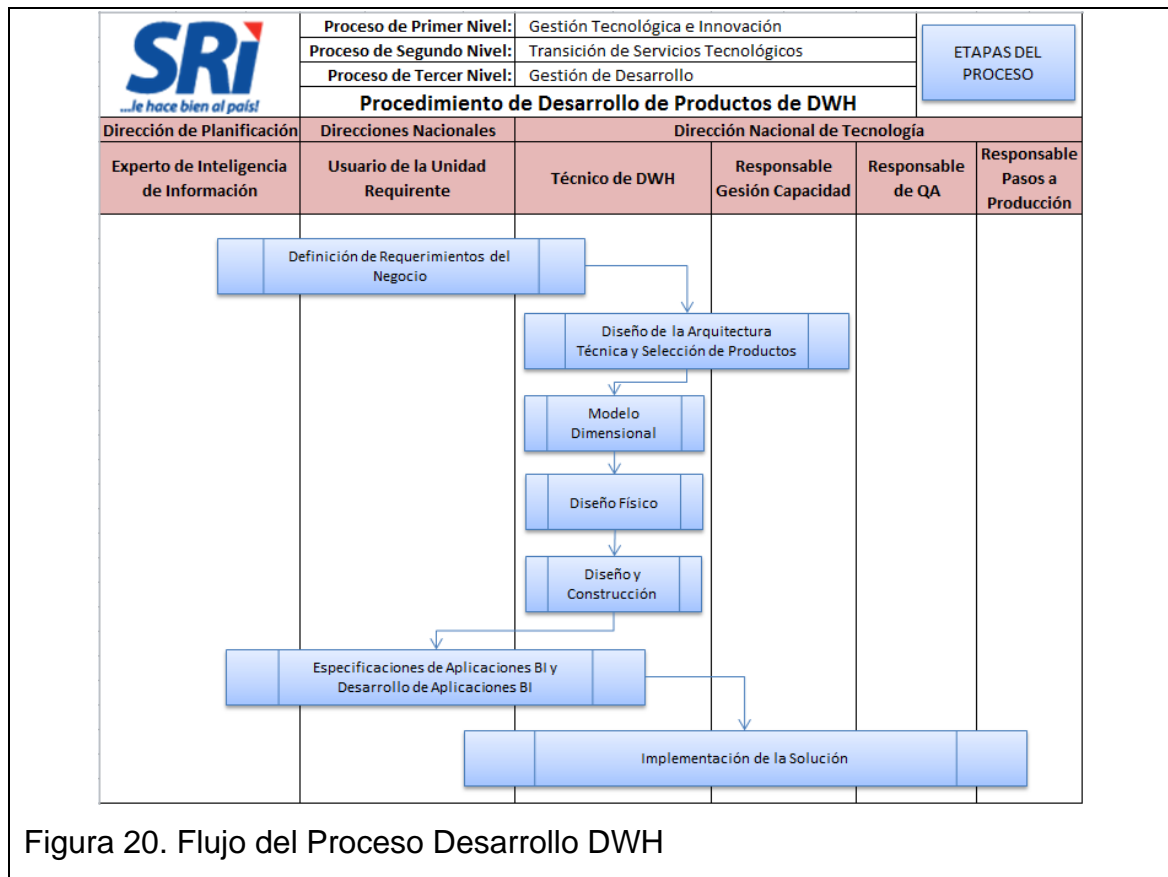


Figura 20. Flujo del Proceso Desarrollo DWH

4.5.7 Entregables del Proceso

A continuación se muestra la tabla que describe los entregables finales por cada una de las fases del proceso de Desarrollo de Productos de DWH propuesto misma que fue generada considerando las características del ciclo de vida de Kimball y Ross (2013) y del marco de trabajo ágil SCRUM.

Tabla 8. Entregables Proceso de Desarrollo de Productos de DWH

Fase 1	Definición de Requerimientos del Negocio
	Documento de Levantamiento de Requerimientos de Información
	Registro de las historias de usuario con sus criterios de aceptación
	Estimación y priorización de las historias de usuario
	<i>Backlog</i> de historias de usuario por iteración
Fase 2	Actualización de la Matriz de DWH del SRI
Fase 2	Diseño de la Arquitectura Técnica y Selección de Productos.
	Informe de arquitectura
	Documento de capacidad de infraestructura
Fase 3	Procedimiento de Adquisiciones
	Modelo Dimensional
	Actualización de la Matriz de DWH del SRI
Fase 3	Modelo dimensional lógico
	Análisis de perfilamiento datos fuentes
	Diseño Físico
Fase 4	Modelo dimensional físico
	Diseño y Construcción del ETL
Fase 5	Diseño y construcción de los procesos de carga en la herramienta de ETL
	Especificación y Desarrollo de Aplicaciones de BI
Fase 6	Reportes y dashboard construidos, publicados y probados en ambiente de desarrollo
	Documentos de certificación aprobado
	Implementación de la Solución
Fase 7	Registro de los pasos a producción en Sistema de Gestión TI
	Certificación pasos a producción.
	Pasos a producción ejecutados.
	Documentación del cambio actualizada en la herramienta documental

4.6 Sustentación del Procedimiento

Con la elaboración del procedimiento propuesto se pretende solventar los problemas que surgen durante la construcción de soluciones de información, que ocasionan proveer información inoportuna e inconsistente a los usuarios del negocio, además de otros inconvenientes identificados en la justificación de esta tesis.

De acuerdo a las conclusiones de las encuestas realizada sobre el nivel de servicios del Área de Datawarehouse y del proceso actual de levantamiento de requerimientos (ver lit. 3.7), se ha podido identificar el escenario actual en cuanto el nivel de satisfacción de los usuarios de negocio con respecto a las soluciones de información implementadas. Así mismo en base al análisis del proceso actual de desarrollo de DWH se identificaron varios inconvenientes (ver lit. 3.6).

En la siguiente tabla se detallan los beneficios que provee la implementación del procedimiento propuesto. Así también se determina si en el actual escenario se disponen o no de los beneficios del procedimiento propuesto:

Tabla 9. Evaluación beneficios procedimiento propuesto situación actual

Situación Actual			
No	Beneficios del Procedimiento	Cumple	No Cumple
1	Obtener entregables parciales durante la implementación de los requerimientos solicitados por los usuarios funcionales.		X
2	Contar con un proceso de levantamiento de requerimiento apoyado en formatos que permita describir los requerimientos de información en un lenguaje natural así como los criterios de aceptación que permitan verificar el cumplimiento de dichos requerimiento previa implementación.		X
3	Contar con un procedimiento de desarrollo de productos de DWH en el cual se establezcan tareas de perfilamiento de datos así como de diseño, construcción y control de calidad sobre los productos construidos que permitan generar información		X

	adecuada y precisa de manera oportuna para poner a disposición de los usuarios del negocio.		
4	Al seguir los lineamientos de un procedimiento de desarrollo de productos de DWH se generan productos de alta calidad que cumplan satisfactoriamente con las expectativas de los usuarios de negocio.		X
5	Establece el uso obligatorio de los estándares para la construcción de los productos de DWH.		X
6	Contar con un procedimiento formal de desarrollo de productos de DWH que defina los lineamientos a considerar en cada una de las etapas del ciclo de vida del desarrollo.		X
7	Provee herramientas para el seguimiento de las tareas que se realizan en las iteraciones de manera regular, permitiendo conocer el progreso real del cambio o proyecto.		X
8	Contar con documentación de cada una de las fases del desarrollo registrada en la herramienta de gestión documental compartida a todos los interesados del cambio o proyecto.		X

La Tabla 9 se genera del análisis realizado al proceso actual de desarrollo de productos de DWH y a las conclusiones de la tabulación de la encuesta de la situación actual (ver Lit. 3.6 y 3.7)

Se concluye que el proceso actual de desarrollo de productos de DWH no cumple ninguno de los beneficios que se tendría si se implementa el procedimiento de desarrollo planteado.

Con la finalidad de proyectar la situación futura sobre la aceptación de la implementación del procedimiento propuesto, se realiza una encuesta a 10 funcionarios del negocio que inicialmente también participaron en el registro de la primera encuesta, considerados como usuarios analíticos de nivel avanzado. En la siguiente tabla se muestra el universo de los encuestados por áreas del SRI:

Tabla 10. Universo de encuestados Situación Futura

Áreas del SRI	Número de encuestados
Planificación Institucional	1
Inteligencia de la Información	3
Riesgos Tributarios	3
Grandes Contribuyentes	1
Centro de Estudios Fiscales	1
Control Tributario	1
Total:	10

Las preguntas realizadas en la encuesta así como el resultado tabulado se detallan en el anexo cinco. Cada una de las preguntas hace referencia a los beneficios de contar con un procedimiento para el desarrollo de productos de DWH. Las preguntas se enfocan a determinar la aceptación, uso y cumplimiento del procedimiento propuesto por parte de los encuestados.

En la siguiente figura se muestran los resultados tabulados de la encuesta por cada una de las preguntas formuladas:

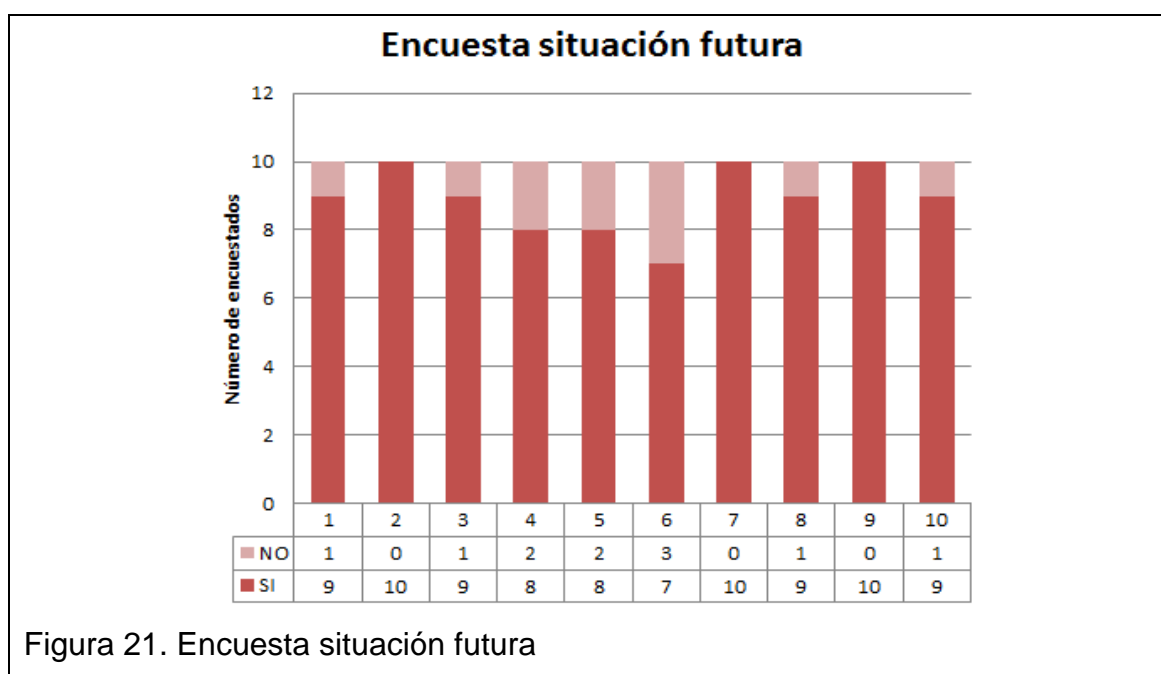


Figura 21. Encuesta situación futura



De acuerdo al análisis de la tabulación de esta segunda encuesta se observa que la mayoría de los encuestados estarían de acuerdo con los beneficios que

aportaría la implantación del Procedimiento de DWH propuesto. En el anexo cinco se muestra la tabulación de la encuesta por áreas del SRI.

En la siguiente tabla se muestra la evaluación sobre los beneficios proporcionados por el procedimiento propuesto:

Tabla 11. Evaluación beneficios procedimiento propuesto situación futura

Situación Futura			
No	Beneficios del Procedimiento	Si Cumpliría	No Cumpliría
1	Obtener entregables parciales durante la implementación de los requerimientos solicitados por los usuarios funcionales.	✓	
2	Contar con un proceso de levantamiento de requerimiento apoyado en formatos que permita describir los requerimientos de información en un lenguaje natural así como los criterios de aceptación que permitan verificar el cumplimiento de dichos requerimiento previa implementación.	✓	
3	Contar con un procedimiento de desarrollo de productos de DWH en el cual se establezcan tareas de perfilamiento de datos así como de diseño, construcción y control de calidad sobre los productos construidos que permitan generar información adecuada y precisa de manera oportuna para poner a disposición de los usuarios del negocio.	✓	
4	Al seguir los lineamientos de un procedimiento de desarrollo de productos de DWH se generan productos de alta calidad que cumplan satisfactoriamente con las expectativas de los usuarios de negocio.	✓	
5	Establece el uso obligatorio de los estándares para la construcción de los productos de DWH.	✓	
6	Contar con un procedimiento formal de desarrollo de productos de DWH que defina los lineamientos a considerar en cada una de las etapas del ciclo de vida del desarrollo.	✓	

7	Provee herramientas para el seguimiento de las tareas que se realizan en las iteraciones de manera regular, permitiendo conocer el progreso real del cambio o proyecto.		
8	Contar con documentación de cada una de las fases del desarrollo registrada en la herramienta de gestión documental compartida a todos los interesados del cambio o proyecto.		

Con los datos recolectados y validados se concluye que el Procedimiento de Desarrollo de Productos de DWH propuesto es factible y viable su implementación y aplicabilidad dentro del SRI.

5. CAPITULO V CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

El resultado del marco metodológico propuesto es la elaboración del Procedimiento de Desarrollo de Productos de DWH, en el cual se establecen las tareas que deben realizarse por medio de prácticas ágiles, mismas que constan descritas en el marco de trabajo SCRUM.

El Procedimiento de Desarrollo de Productos de DWH propuesto, en su parte inicial hace énfasis al levantamiento de requerimientos de información, el cual plantea como objetivo justificar la necesidad generada por el requirente a través del uso de plantillas, las cuales permitirían un levantamiento de requerimientos de información formal y consensuada.

El Procedimiento de Desarrollo de Productos de DWH propuesto, define fases para la construcción de soluciones de DWH, las cuales se basan en el ciclo de vida propuesto por Kimball y Ross (2013), donde cada una de las mismas debe ser realizada como parte de las iteraciones planificadas.

En cada iteración se planifica las historias de usuario a desarrollar y se establecen las actividades técnicas, mismas que son parte del ciclo de vida de DWH. La iteración tiene como objetivo entregar un producto funcional al usuario de negocio por lo que al final de cada iteración el Requirente deberá verificar los criterios de aceptación y aprobar o no el despliegue a producción de la funcionalidad probada. Esto permitirá mejorar el desempeño del Área de Datawarehouse al entregar productos funcionales en menor tiempo y de manera oportuna.

La entrega al usuario del negocio de productos funcionales en menor tiempo permitirá mejorar la eficiencia empresarial ya que se proveerán herramientas tecnológicas para acceder a la información y apoyar a la toma de decisiones

efectivas en menor tiempo, permitiendo alcanzar los objetivos estratégicos de la Institución como son: mejorar la recaudación, reducir el fraude fiscal y reducir los índices de evasión tributaria.

Al final de cada iteración se realiza una evaluación denominada retrospectiva, con el objetivo de identificar los aspectos de mejora que se pueden adoptar para las siguientes iteraciones. De esa forma se pretende realizar un proceso de mejoramiento continuo al procedimiento de construcción de soluciones de DWH.

La arquitectura de Hardware y Software de DWH que actualmente se encuentra disponible en el SRI, permite implementar la mayoría de soluciones de información requeridas por el negocio, ya que cuenta con una infraestructura adecuada de servidores sobre los cuales se ejecuta un motor de base de datos de alto desempeño así como también las aplicaciones de ETL y BI.

La base de datos del DWH actual es de tipo relacional por lo que la mayoría de soluciones de información que se desarrollen deben tener como resultado la implementación de un modelo multidimensional, que provea a los usuarios del negocio un producto de fácil uso y sencillo al momento de realizar y ejecutar los reportes de análisis.

El Procedimiento de Desarrollo de Productos de DWH propuesto hace énfasis en la generación de la documentación en cada una de las fases del procedimiento la cual debe estar registrada en la herramienta de gestión documental del SRI.

5.2 Recomendaciones

Documentar la arquitectura del Data Warehouse del SRI, la cual refleje los elementos principales de la misma, así como una descripción de cada uno de ellos. Este documento deberá estar publicado en las herramientas documentales oficiales que están a disposición de los funcionarios de la Institución, así como

también se deberá socializar dicho documento entre las jefaturas de áreas y departamentales, además del Director de Tecnología del SRI.

Inicialmente se plantea el uso de formatos para las fases de levantamiento de requerimientos, seguimiento y avance de las iteraciones. Posteriormente se sugiere hacer uso de alguna herramienta informática que permita automatizar dichas tareas, entre las cuales se sugiere la herramienta Jira Agile.

Toda documentación que se genere de la implementación de productos de DWH deberá ser registrada en la herramienta documental institucional, la cual debe estar organizada adecuadamente para que sea de fácil acceso a las personas que requieran utilizarla.

En el caso de implementar el Procedimiento de Desarrollo de Productos de DWH, debe ser revisado y aprobado por todos los involucrados así como por la máxima autoridad de la institución, en este caso por el Director General del SRI. Posterior a su aprobación el procedimiento debe entrar en vigencia y cada uno de los roles que intervienen en el proceso deberán acatar lo estipulado en el mismo.

Para una efectiva aplicación del Procedimiento de Desarrollo de Productos de DWH, se recomienda capacitar a los roles involucrados en las técnicas y prácticas ágiles propuestas en el marco de trabajo Scrum.

Se recomienda el uso del marco Kanban para la gestión de tareas asignadas al personal del Área de Datawarehouse, lo que permitirá tener un mejor control sobre las actividades realizadas por el Área.

6. REFERENCIAS

- Bermejo, M. (2010). *El Kanban*. Recuperado el 23 de noviembre de 2014, de FUOC. Fundación para la Universitat Oberta de Catalunya: [https://www.exabyteinformatica.com/uoc/Audiovisual/Produccion_multimedia/Produccion_multimedia_\(Modulo_4\).pdf](https://www.exabyteinformatica.com/uoc/Audiovisual/Produccion_multimedia/Produccion_multimedia_(Modulo_4).pdf)
- Breslin, M. (2004). *Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models*. Recuperado el 29 de mayo de 2015, de Business Intelligence Journal: https://cours.etsmtl.ca/mti820/public_docs/lectures/DWBattleOfTheGiants.pdf
- Cano, J. L. (2007). *Business Intelligence: Competir con Información*. ESADE.
- Collier, K. (2011). *Agile Analytics*. Boston: Pearson Education, Inc.
- Dirección de Tecnologías de Información de la Universidad de Concepción de Chile. (s.f.). *Pirámide Organizacional*. Recuperado el 15 de marzo de 2015, de Dirección de Tecnologías de Información: <http://www.udec.cl/dti/node/108>
- Dirección Nacional de Desarrollo Tecnológico del SRI. (2014). *Procedimiento de Gestión de Cambios Tecnológicos*. Quito.
- Dirección Nacional de Planificación y Coordinación del SRI. (2012). *Plan Estratégico Institucional 2012-2015*. Quito.
- Gallego, M. (s.f.). *Metodología Scrum*. Recuperado el 7 de junio de 2015, de Universidad Oberta de Catalunya Repositorio institucional: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/17885/1/mtrigas_TFC0612memoria.pdf
- Hundermark, P. (Noviembre de 2009). *Un Mejor Scrum*. Recuperado el 6 de junio de 2015, de ScrumSense: <http://www.scrumsense.com/wp-content/uploads/2012/03/Un-mejor-Scrum-2.pdf>
- Inmon, W. H. (2002). *Building the Data Warehouse*. John Wiley & Sons, Inc.
- Innovasys. (2010). *Documento de Diseño de Arquitectura de Business Intelligence del SRI*. Quito.
- INTECO. (2009). *Ingeniería del Software: Metodologías y Ciclos de Vida*. Recuperado el 16 de noviembre de 2015, de

http://datateca.unad.edu.co/contenidos/301569/guia_de_ingenieria_del_software.pdf

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit Third Edition*. Indianapolis: John Wiley & Sons, Inc.

Kniberg, H. (2007). *Scrum y XP desde las trincheras*. C4Media Inc. InfoQ.

Kniberg, H., Skarin, M., Poppendieck, M., & Anderson, D. (2010). *Kanban y Scrum – obteniendo lo mejor de ambos. Prólogo de Mary Poppendieck & David Anderson*. Estados Unidos: C4Media Inc.

M.G. Piattini, J.A. Calvo-Manzano, J.Cervera, & L.Fernández. (1996). *Análisis y Diseño detallado de Aplicación Informáticas de Gestión*. Recuperado el 4 de mayo de 2015, de SEDICI: http://sedici.unlp.edu.ar/bitstream/handle/10915/4055/3_-_La_ingenier%C3%ADa_de_software.pdf?sequence=7

Monsalve, M. (2014). *Arquitectura Empresarial*. 189. Quito, Pichincha, Ecuador.

Registro Oficial. (2014). Servicio De Rentas Internas Estatuto Orgánico de Gestión Organizacional por Procesos. *Registro Oficial No 134*, 34.

Russom, P. (2014). *Evolving Data Warehouse Architectures In the Age of Big data*. The Data Warehousing Institute.

SCRUMstudy. (2013). *Una guía para el conocimiento de Scrum (Guía SBOK)*. Phoenix, Arizona: SCRUMstudy.

Servicio de Rentas Internas. (2008). *Texto Sustitutivo del Reglamento Orgánico Funcional Codificado del SRI*. Quito.

Servicio de Rentas Internas. (2016). *¿Qué es el SRI?* Recuperado el 15 de enero de 2016, de Servicio de Rentas Internas: <http://www.sri.gob.ec/web/guest/67>

ANEXOS

Anexo 1 - Preguntas Encuesta – Situación Actual



La siguiente encuesta tiene como objetivo evaluar el nivel de servicio que ofrece el Área de Datawarehouse a los usuarios del SRI que han solicitado la implementación de una solución de información al área, identificar necesidades de información actuales y valorar el uso de las herramientas de visualización de información. También pretende determinar la situación actual del proceso de identificación y levantamiento de las necesidades de información para proponer alternativas de mejora al actual proceso.

Preguntas Encuesta DWH

No	Pregunta
1	Ingrese el ÁREA / DEPARTAMENTO / DIRECCIÓN a la que pertenece.
2	¿Conoce usted sobre la existencia del área DWH?
3	En caso de que su respuesta es afirmativa: ¿Cuál cree que es la finalidad del área DWH?, escoja una o varias opciones: <ul style="list-style-type: none">• Administración de Bases de Datos del SRI.• Desarrollo de aplicaciones para registro de información de contribuyentes.• Apoyo al usuario en la generación de reportes.• Proveer al usuario herramientas y bases de datos que le permitan en forma sencilla y ágil acceder a la información del negocio con el objetivo de brindar soporte en sus procesos de toma de decisiones.• No conozco.• Otros.
4	Si en la pregunta anterior, su respuesta fue la opción "Otros", especifique cuál es la finalidad del Área de Datawarehouse a su criterio:
5	¿Ha realizado algún requerimiento al Área de Datawarehouse?
6	¿Fueron atendidos los requerimientos enviados al Área de Datawarehouse?
7	Si su respuesta anterior fue negativa, ¿por qué razón no fueron atendidos?, escoja una o varias opciones: <ul style="list-style-type: none">• Falta de recurso humano.• Cambio de prioridades en la Institución.• Falta de capacidad o infraestructura.• Otros.

8	Si la selección a la pregunta anterior fue "Otros", especifique otras razones por las que no fueron atendidos los requerimientos enviados:
9	<p>¿Cómo fue la atención de los requerimientos enviados al Área de Datawarehouse? escoja una opción:</p> <ul style="list-style-type: none"> • Dentro del tiempo adecuado y esperado. • Con demora, el requerimiento atendido aún era necesario. • Demasiado tarde, el requerimiento ya no era necesario. • Otros. • No aplica.
10	Si en la pregunta anterior su respuesta fue la opción "Otros", especifique cómo fueron atendidos los requerimientos enviados al Área de Datawarehouse a su criterio:
11	<p>¿Qué tan complejo es para Usted entender las problemáticas de gestión de información de su área y trasmitirlas mediante requerimiento al equipo de Datawarehouse para que resuelvan su necesidad?</p> <ul style="list-style-type: none"> • Totalmente complejo. • Medianamente complejo. • Normal. • Medianamente fácil. • Fácil. • No aplica.
12	<p>La frecuencia con que se actualiza los cubos de información que están a su disposición para las actividades que realiza en su área es:</p> <ul style="list-style-type: none"> • La frecuencia de actualización de los cubos es la adecuada. • La frecuencia de actualización de los cubos no es la adecuada, pero es útil para las actividades del área. • La frecuencia de actualización de los cubos no es la adecuada, y no es útil para las actividades del área. • No aplica.
13	<p>¿Los servicio o productos proporcionado por el Área de Datawarehouse cumplieron con sus expectativas?, escoja una opción:</p> <ul style="list-style-type: none"> • Muy Satisfactorio. • Satisfactorio. • Poco satisfactorio. • Nada satisfactorio. • No aplica.
14	¿Explique el por qué selecciono la opción de la pregunta anterior?
15	¿Durante el último año ha experimentado caída de alguno de los servicios provistos por el Área de DWH?

16	<p>Si su respuesta fue positiva, indique el tiempo en horas que duró esta caída, escoja una opción:</p> <ul style="list-style-type: none"> • Menos de una hora. • 1 a 2 horas. • 2 a 4 horas. • a 6 horas. • 6 a 8 horas. • Más de 8 horas.
17	<p>¿Qué tipo de atención recibió del Área de Datawarehouse cuando usted reportó la caída del servicio?, escoja una opción:</p> <ul style="list-style-type: none"> • No hubo atención. • Atención insuficiente. No se comunica el tiempo para restablecer el servicio. • Atención adecuada. • No aplica.
18	<p>¿Durante el último año ha detectado inconsistencia de información sobre los cubos implementados por el Área de Datawarehouse y que están a su disposición?</p>
19	<p>¿Qué tipo de requerimiento son los más comunes que solicita al Área de Datawarehouse? Ingrese la prioridad de cada uno, siendo 1 la menor prioridad y 10 la mayor:</p> <ul style="list-style-type: none"> • Descargas de información masiva. • Creación y publicación de información para la toma de decisiones. • Creación de reportes con complejidad alta. • Recibir y/o compartir datos con Instituciones Externas. • Actualización de las áreas de negocio y/o universos existentes. • Cruces de información masiva. • Reprocesos de información. • Permisos de acceso a las herramientas de BI. • Construcción de indicadores. • Ninguno.
20	<p>¿Cuál es la importancia que tiene para su trabajo los servicios o productos que el Área de Datawarehouse ha implementado y pone a su disposición?</p> <ul style="list-style-type: none"> • Muy importante. • Importante. • Nada importante.
21	<p>Califique la calidad de información entregada por el Área de Datawarehouse:</p> <ul style="list-style-type: none"> • Excelente. • Bueno. • Regular.

	<ul style="list-style-type: none"> • Malo. • No aplica.
22	<p>En caso de que su respuesta sea diferente a Excelente especifique las novedades, observaciones y errores que ha identificado usted, escoja una o varias opciones:</p> <ul style="list-style-type: none"> • Datos inconsistentes. • No refleja cambios históricos. • Problemas de desempeño de consultas. • No tiene privilegios. • Otros.
23	<p>Si en la pregunta anterior seleccionó la opción "Otros", especifique las novedades, observaciones y errores que ha identificado usted:</p>
24	<p>¿Qué tipo de usuario se considera usted?</p> <ul style="list-style-type: none"> • Nivel Básico (ejecuta consultas prediseñadas) • Nivel Intermedio (utiliza filtros y cálculos complejos) • Nivel Avanzado (realiza sus propias consultas específicas que resuelven un problema puntual)
25	<p>Seleccione la frecuencia con que utiliza las herramientas de BI para sus actividades:</p> <ul style="list-style-type: none"> • Mensual. • Semanal. • Diaria. • Nunca.
26	<p>¿Las capacidades de las herramientas de visualización satisfacen las necesidades que usted requiere para su trabajo?</p>
27	<p>En caso de que la respuesta anterior es negativa, indique las capacidades que debe tener la herramienta:</p>
28	<p>¿Las herramientas de visualización son fáciles de usar?</p> <ul style="list-style-type: none"> • Fácil de usar. • No tan fácil de usar. • Difícil de usar.

Anexo 2 - Tabulación Encuesta – Situación Actual

1. Ingrese el ÁREA / DEPARTAMENTO / DIRECCIÓN a la que pertenece.

- Inteligencia de la Información.
- Planificación Institucional.
- Grandes Contribuyentes.
- Riesgos Tributarios.
- Desarrollo Tecnológico.
- Control Tributario.
- Centro de Estudios Fiscales.

2. ¿Conoce usted sobre la existencia del área DWH?

Respuesta	Media	Total
Sí	100%	18
No		0
Total	100%	18/18

Tabulación pregunta 2

3. En caso de que su respuesta es afirmativa: ¿Cuál cree que es la finalidad del área DWH?, escoja una o varias opciones:

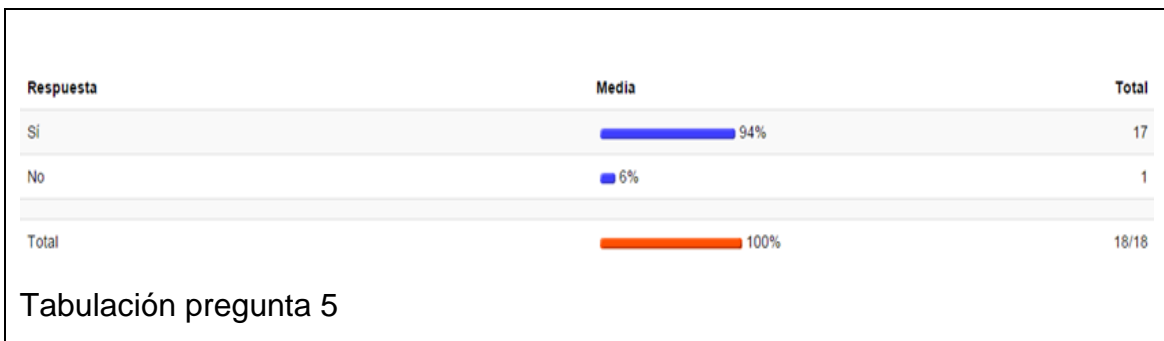
Respuesta	Media	Total
Administración de Bases de Datos del SRI.	11%	4
Desarrollo de aplicaciones para registro de información de contribuyentes.	5%	2
Apoyo al usuario en la generación de reportes.	24%	9
Proveer al usuario herramientas y bases de datos que le permitan en forma sencilla y ágil acceder a la información del negocio con el objetivo de brindar soporte en sus procesos de toma de decisiones.	46%	17
Otros.	14%	5

Tabulación pregunta 3

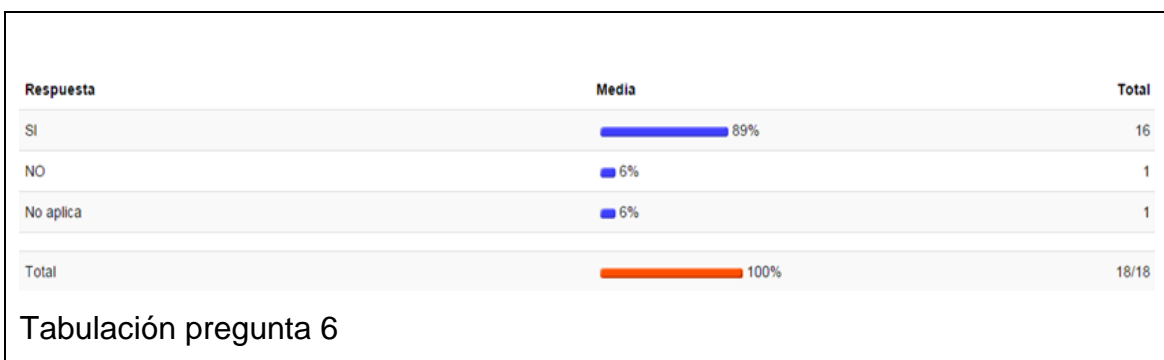
4. Si en la pregunta anterior, su respuesta fue la opción "Otros", especifique cuál es la finalidad del Área de Datawarehouse a su criterio:

- Tener conocimiento y dar soporte en temas de información de las bases de datos del SRI.
- Proveer al usuario de soluciones de datos masivos que permitan agilizar el análisis de la información y poder actuar de manera más ágil.
- Permitir visualizar la información.
- Apoyo en mejores soluciones tecnológicas.
- Apoyo al usuario en generación de información para toma de decisiones a través de visores de gestión y análisis de información.

5. ¿Ha realizado algún requerimiento al Área de Datawarehouse?



6. ¿Fueron atendidos los requerimientos enviados al Área de Datawarehouse?



7. Si su respuesta anterior fue negativa, ¿por qué razón no fueron atendidos?, escoja una o varias opciones:

Respuesta	Media	Total
Falta de recurso humano.	20%	1
Cambio de prioridades en la institución.	20%	1
Otros.	60%	3

Tabulación pregunta 7

8. Si la selección a la pregunta anterior fue "Otros", especifique otras razones por las que no fueron atendidos los requerimientos enviados:

- Cambio de personas asignadas que tenían el conocimiento.
- Demoras por prioridades.
- Existen pedidos que se requieren de forma periódica pero DWH solo procesa una vez y para otro requerimiento se debe hacer otra vez el mismo pedido.

9. ¿Cómo fue la atención de los requerimientos enviados al Área de Datawarehouse? escoja una opción:

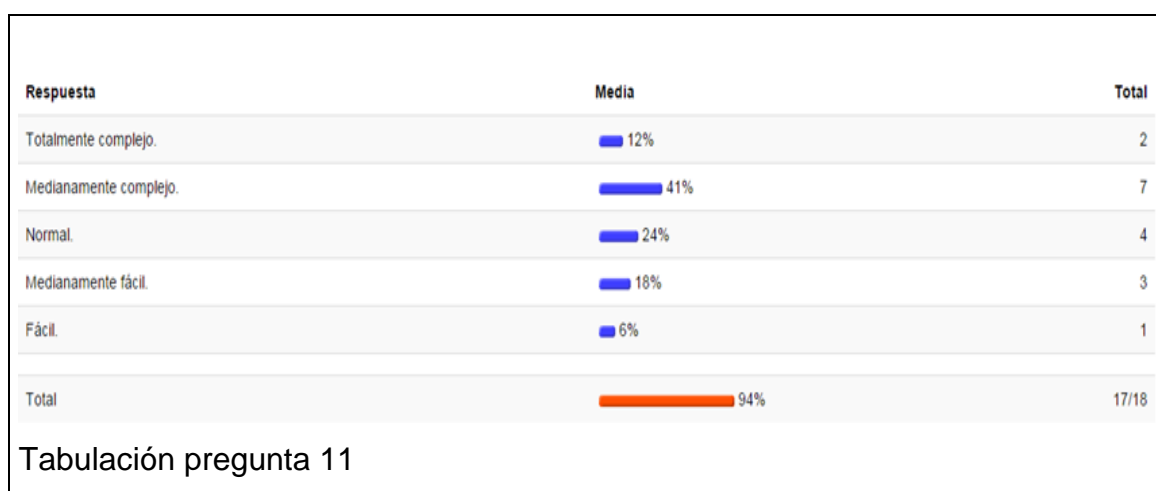
Respuesta	Media	Total
Dentro del tiempo adecuado y esperado.	33%	6
Con demora, el requerimiento atendido aún era necesario.	50%	9
Otros.	11%	2
No aplica.	6%	1
Total	100%	18/18

Tabulación pregunta 9

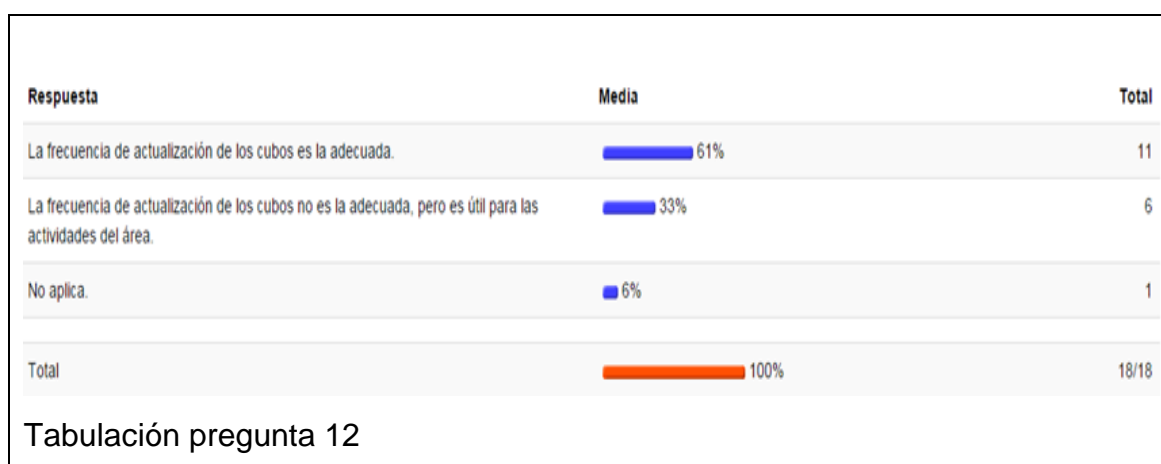
10. Si en la pregunta anterior su respuesta fue la opción "Otros", especifique cómo fueron atendidos los requerimientos enviados al Área de Datawarehouse a su criterio:

- Depende de la persona que atiende el requerimiento.
- Depende del recurso de DWH que atiende el requerimiento y la complejidad del mismo.

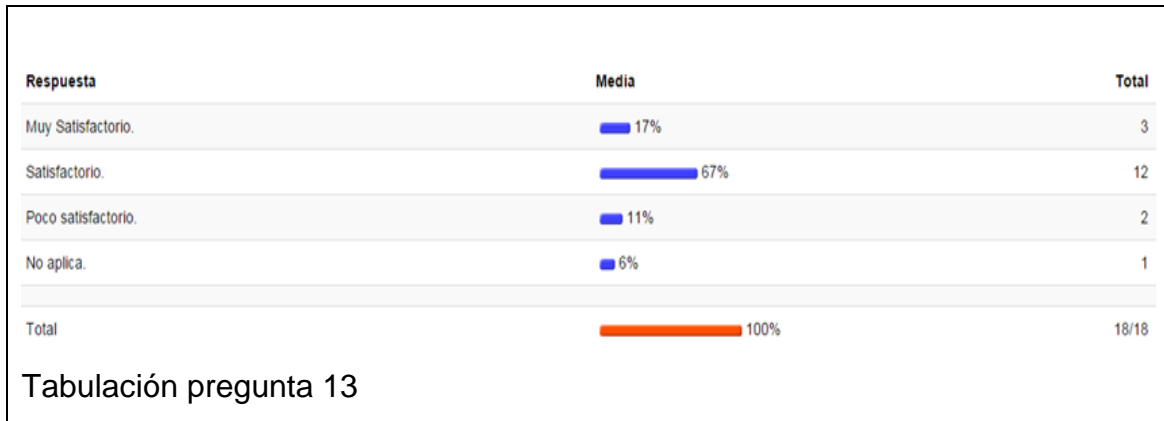
11. ¿Qué tan complejo es para Usted entender las problemáticas de gestión de información de su área y transmitir las mediante requerimiento al equipo de Datawarehouse para que resuelvan su necesidad?



12. La frecuencia con que se actualiza los cubos de información que están a su disposición para las actividades que realiza en su área es:



13. ¿Los servicio o productos proporcionado por el Área de Datawarehouse cumplieron con sus expectativas?, escoja una opción:



14. ¿Explique el por qué selecciono la opción de la pregunta anterior?

- Porque cuando se ha necesitado la ayuda de DWH ha sido en el tiempo establecido.
- Problemas: limitada cantidad de recursos ofrecidos para la solución de muchos usuarios y desenfoco en la atención de necesidades de la organización.
- La atención en los servicios brindados varía dependiendo de la persona que los atiende.
- Solo dan prioridad en la atención a los procesos de Gestión Tributaria.
- Debido a que deben ser consensuadas con más áreas para tener una visión más global.
- La necesidad fue solucionada, sin embargo podía haberse mejorado las relaciones con otras bases
- Debido a que cumplen con el requerimiento solicitado
- Porque no se atendieron todos en el momento requerido.
- Porque tienen un excelente equipo de personas pero las necesidades son muchas que no se pueden priorizar a todas.
- Es bueno pero a veces demora demasiado.
- Depende de quien realiza el requerimiento y muchas veces suelen ser burocráticos en los pedidos.
- El trabajo es eficiente, pero siempre se podría agilizar los procesos.

- Cumple con las expectativas, pero se requiere actualizar las herramientas de visualización de datos.
- Tiempo adecuado de respuesta.
- Son productos muy buenos aunque no adaptados 100% a la necesidad por la complejidad de los productos solicitados.
- Porque el servicio es bueno.
- Cumplen con los parámetros establecidos y superan las expectativas de calidad.
- Se cumple con el requerimiento exacto, aunque el proceso es un poco largo.

15. ¿Durante el último año ha experimentado caída de alguno de los servicios provistos por el Área de DWH?

Respuesta	Media	Total
Sí	56%	10
No	44%	8
Total	100%	18/18

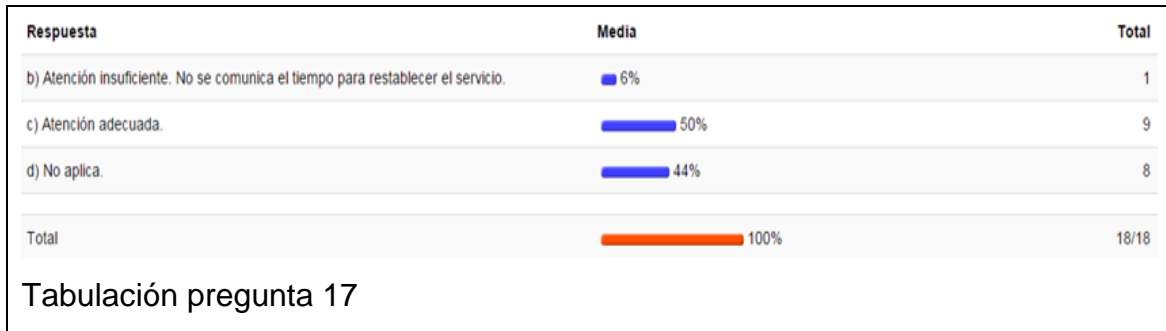
Tabulación pregunta 15

16. Si su respuesta fue positiva, indique el tiempo en horas que duró esta caída, escoja una opción:

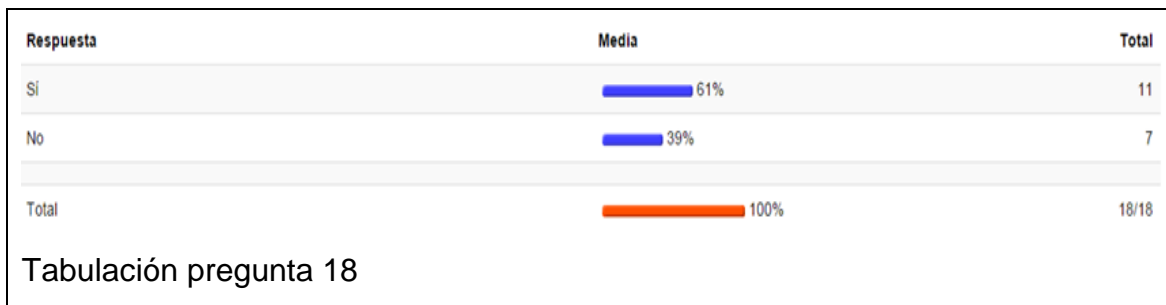
Respuesta	Media	Total
menos de una hora.	11%	1
1 a 2 horas.	11%	1
2 a 4 horas.	11%	1
4 a 6 horas.	33%	3
6 a 8 horas.	11%	1
Más de 8 horas.	22%	2
Total	50%	9/18

Tabulación pregunta 16

17. ¿Qué tipo de atención recibió del Área de Datawarehouse cuando usted reportó la caída del servicio?, escoja una opción:



18. ¿Durante el último año ha detectado inconsistencia de información sobre los cubos implementados por el Área de Datawarehouse y que están a su disposición?



19. ¿Qué tipo de requerimiento son los más comunes que solicita al Área de Datawarehouse? Ingrese la prioridad de cada uno, siendo 1 la menor prioridad y 10 la mayor:

Ranking de la media											↓
	1	2	3	4	5	6	7	8	9	10	
Descargas de información masiva.											6.3
Creación y publicación de información para la toma de decisiones.											6.3
Creación de reportes con complejidad alta.											7.7
Recibir y/o compartir datos con Instituciones Externas.											5.6
Actualización de las áreas de negocio y/o universos existentes.											5.4
Cruces de información masiva.											6.1
Reprocesos de información.											4.9
Permisos de acceso a las herramientas de BI.											5.9
Construcción de indicadores.											4.8
Ninguno.											1.9
Respuestas	1	2	3	4	5	6	7	8	9	10	Total
Descargas de información masiva.	2 (11%)	0	3 (17%)	2 (11%)	2 (11%)	1 (6%)	0	0	2 (11%)	6 (33%)	18
Creación y publicación de información para la toma de decisiones.	0	3 (17%)	0	3 (17%)	2 (11%)	1 (6%)	1 (6%)	2 (11%)	3 (17%)	3 (17%)	18
Creación de reportes con complejidad alta.	0	0	0	1 (6%)	0	1 (6%)	4 (22%)	8 (44%)	3 (17%)	1 (6%)	18
Recibir y/o compartir datos con Instituciones Externas.	0	6 (33%)	2 (11%)	1 (6%)	0	1 (6%)	1 (6%)	2 (11%)	1 (6%)	4 (22%)	18
Actualización de las áreas de negocio y/o universos existentes.	2 (11%)	1 (6%)	2 (11%)	1 (6%)	1 (6%)	2 (11%)	6 (33%)	2 (11%)	1 (6%)	0	18
Cruces de información masiva.	0	2 (11%)	2 (11%)	1 (6%)	3 (17%)	3 (17%)	1 (6%)	1 (6%)	3 (17%)	2 (11%)	18
Reprocesos de información.	0	1 (6%)	2 (11%)	6 (33%)	3 (17%)	2 (11%)	2 (11%)	2 (11%)	0	0	18
Permisos de acceso a las herramientas de BI.	0	0	4 (22%)	1 (6%)	3 (17%)	4 (22%)	2 (11%)	0	3 (17%)	1 (6%)	18
Construcción de indicadores.	1 (6%)	4 (22%)	2 (11%)	1 (6%)	4 (22%)	2 (11%)	0	1 (6%)	2 (11%)	1 (6%)	18




Tabulación pregunta 19

20. ¿Cuál es la importancia que tiene para su trabajo los servicios o productos que el Área de Datawarehouse ha implementado y pone a su disposición?

Respuesta	Media	Total
Muy importante.	100%	18
Total	100%	18/18






Tabulación pregunta 20

21. Califique la calidad de información entregada por el Área de Datawarehouse:

Respuesta	Media	Total
Excelente.	 44%	8
Bueno.	 56%	10
Total	 100%	18/18

Tabulación pregunta 21

22. En caso de que su respuesta sea diferente a Excelente especifique las novedades, observaciones y errores que ha identificado usted, escoja una o varias opciones:

Respuesta	Media	Total
Datos inconsistente.	 25%	6
No refleja cambios históricos.	 4%	1
Problemas de desempeño de consultas.	 21%	5
No tiene privilegios.	 25%	6
Otros.	 25%	6

Tabulación pregunta 22

23. Si en la pregunta anterior seleccionó la opción "Otros", especifique las novedades, observaciones y errores que ha identificado usted:

- Información no actualizada.
- Podría ser mejor.
- Siempre existen observaciones a la data en producción y dependiendo del usuario estas observaciones se corrigen rápidamente o no.
- Demoras en la ejecución de reprocesos.

24. ¿Qué tipo de usuario se considera usted?

Respuesta	Media	Total
Nivel Básico (ejecuta consultas prediseñadas)	17%	3
Nivel Intermedio (utiliza filtros y cálculos complejos)	11%	2
Nivel Avanzado (realiza sus propias consultas específicas que resuelven un problema puntual)	72%	13
Total	100%	18/18

Tabulación pregunta 24

25. Seleccione la frecuencia con que utiliza las herramientas de BI para sus actividades:

Respuesta	Media	Total
Mensual.	17%	3
Semanal.	11%	2
Diaria.	72%	13
Total	100%	18/18

Tabulación pregunta 25

26. ¿Las capacidades de las herramientas de visualización satisfacen las necesidades que usted requiere para su trabajo?

Respuesta	Media	Total
Sí	50%	8
No	50%	8
Total	89%	16/18




Tabulación pregunta 26

27. En caso de que la respuesta anterior es negativa, indique las capacidades que debe tener la herramienta:

- Respuesta ágil, interfaz amigable.
- Más flexibles y fáciles de usar.
- Más visual.

- Se requieren más áreas de negocio con información manual y mejores procesamientos.
- Debería tener una mayor capacidad de descarga masiva, y mayor velocidad en el procesamiento de la información.
- Gran cantidad de visualización, la herramienta proporcionada tiene límites en la visualización y exportación de reporte.
- Debe tener la capacidad de hacer cálculos complejos, filtros, desagregaciones, navegación de maestros a detalles, etc.
- Debe tener una interfaz más grafica para usuario final.

28. ¿Las herramientas de visualización son fáciles de usar?

Respuesta	Media	Total
Fácil de usar.	 31%	4
No tan fácil de usar	 69%	9
Total	 72%	13/18

Tabulación pregunta 28

Anexo 3 – Plantilla de Levantamiento de Requerimientos

FORMATO DE LEVANTAMIENTO DE REQUERIMIENTOS DE INFORMACIÓN DEPARTAMENTO DE DESARROLLO TI ÁREA DE DATAWAREHOUSE			
1. DATOS GENERALES SOLICITANTE			
FECHA		CIUDAD	
NOMBRES Y APELLIDOS			
CARGO			
DIRECCIÓN	<input checked="" type="checkbox"/> Nacional	<input type="checkbox"/> Zonal	<input type="checkbox"/> Provincial
DEPARTAMENTO		AREA	
2. ROLES DE NEGOCIO			
Detalle los roles de negocio que harán uso de la solución.			
ROL	DESCRIPCIÓN		
3. DESCRIPCIÓN DEL CAMBIO O PROYECTO			
Descripción breve del cambio o proyecto			
NOMBRE			
DESCRIPCIÓN			
4. OBJETIVOS DE LA SOLUCIÓN			
Describa los objetivos de negocio que esperan lograr con la solución. Pueden existir objetivos por cada rol.			
NO	OBJETIVO		
1			
2			
3			

Levantamiento de Requerimientos de Información - Cabecera

5. HISTORIAS DE USUARIO

Detalla las historias de usuario que se requiere desarrollar de acuerdo al siguiente formato:
Como <ROL> Necesito la capacidad de <hacer algo> para que pueda <declaración de la meta>

CÓDIGO	HISTORIAS DE USUARIO	CRITERIOS DE ACEPTACIÓN
HU1		
HU2		
HU3		
HU4		
HU5		

6. REVISIONES Y APROBACIONES

Elaborado por:

Nombre y Apellido	Cargo	Firma
-------------------	-------	-------

Revisado por:

Nombre y Apellido	Cargo	Firma
-------------------	-------	-------

Nombre y Apellido	Cargo	Firma
-------------------	-------	-------

Aprobado por:

Nombre y Apellido	Cargo	Firma
-------------------	-------	-------

Nombre y Apellido	Cargo	Firma
-------------------	-------	-------

Levantamiento de Requerimientos de Información - Detalle

ESTIMACIÓN Y PRIORIZACIÓN DE HISTORIAS DE USUARIO

DEPARTAMENTO DE DESARROLLO TI
ÁREA DE DATAWAREHOUSE

1. ESTIMACIÓN DE HISTORIAS DE USUARIO

Estimar el esfuerzo de desarrollo en puntos de historias considerando escala de la serie de Fibonacci

COD	HISTORIA DE USUARIOS	PUNTOS	PRIORIDAD
HU1		3	Alta
HU2		5	Media
HU3		3	Baja
HU4		2	Alta
HU5		8	Media
		TOTAL PUNTOS	21

2. ITERACIÓN

Especifique el tamaño en semanas de cada iteración

Tamaño de la iteración: 2 Semana(s)

Especifique la velocidad media por iteración

Velocidad Media: 21 Puntos de usuario

Número de iteraciones requeridas: 1,00

Semanas requeridas: 2,00

Estimación y priorización de historias de usuario

BACKLOG HISTORIAS DE USUARIO

DEPARTAMENTO DE DESARROLLO TI
ÁREA DE DATAWAREHOUSE

1. Backlog y criterios de aceptación

Ordenar las HU en base a la prioridad y estimación.

PRIORIDAD	ORDEN	COD HU	PUNTOS	HISTORIA DE USUARIO	ITERACIÓN	RELEASE
Alta	1	HU1	8	0	1	R1
	2			0		
	3			0		
	4			0		
	5			0		
Media	6			0		
	7			0		
	8			0		
	9			0		
	10			0		
Baja	11			0		
	12			0		
	13			0		
	14			0		
	15			0		

Backlog historias de usuario

SEGUIMIENTO DE LA ITERACIÓN N_
DEPARTAMENTO DE DESARROLLO TI
ÁREA DE DATAWAREHOUSE

1. Datos de la Iteración

Fecha Inicio:		Fecha Finalización:	
Responsable:			
Objetivo:			

2. Listado de historias de usuario

Ordenar las HU en base a la prioridad y estimación. Seguimiento del estado de las tareas técnicas.

COD HU	PUNTOS	EST HU	TAREA	DESCRIPCIÓN TAREAS TÉCNICAS	HORAS	ESTADO	
HU5	8	Pendiente	1	T1		8	Pendiente
			2	T2		8	Enprogrsa
			3	T3		3	Enprogrsa
			4	T4		5	Concluida
			5	T5		4	Concluida
				AVANCE DE LA HISTORIA DE USUARIO	TOTAL HORAS ESTIMADAS:	28	
					HORAS CONCLUIDAS:	9	
					% DE AVANCE:	32%	
HU2	5	Concluida	6	T1		8	Concluida
			7	T2		8	Concluida
			8	T3		3	Concluida
			9	T4		5	Concluida
			10	T5		4	Concluida
				AVANCE DE LA HISTORIA DE USUARIO	TOTAL HORAS ESTIMADAS:	28	
					HORAS CONCLUIDAS:	28	
					% DE AVANCE:	100%	
		Pendiente		T1			
				T2			
				T3			
				T4			
				T5			
				AVANCE DE LA HISTORIA DE USUARIO	TOTAL HORAS ESTIMADAS:	0	
					HORAS CONCLUIDAS:	0	
					% DE AVANCE:	#DIV/0!	
		Pendiente		T1			
				T2			
				T3			
				T4			
				T5			
				AVANCE DE LA HISTORIA DE USUARIO	TOTAL HORAS ESTIMADAS:	0	
					HORAS CONCLUIDAS:	0	
					% DE AVANCE:	#DIV/0!	
		Pendiente		T1			
				T2			
				T3			
				T4			
				T5			
				AVANCE DE LA HISTORIA DE USUARIO	TOTAL HORAS ESTIMADAS:	0	
					HORAS CONCLUIDAS:	0	
					% DE AVANCE:	#DIV/0!	
Total puntos	13						
Puntos reales	5						
% avance	38%						

Seguimiento de la Iteración

Reporte de avance de la Iteración

Avance 38% No__
Fecha inicio: 00/01/1900
Fecha finalización: 00/01/1900
Responsable: 0

Estado Tareas	Estados			
Tarea	En progreso	Pendiente	Concluido	Total:
1		1		1
2	1			1
3	1			1
4			1	1
5			1	1
6			1	1
7			1	1
8			1	1
9			1	1
10			1	1
Total:	2	1	7	10

Reporte de avance de la iteración

REVISIÓN Y RETROSPECTIVA DE LA ITERACIÓN N_

DEPARTAMENTO DE DESARROLLO TI
ÁREA DE DATAWAREHOUSE

1. Revisión y aceptación de las historias de usuario

Fecha de revisión:

Revisión con el Solicitante el cumplimiento de los criterios de aceptación

COD HU	PUNTOS	DESCRIPCIÓN DE LA HISTORIA DE USUARIO	CRITERIOS DE ACEPTACIÓN	ESTADO
HU5	8			

2. Retrospectiva de la iteración

Detallar los aspectos buenos, aspectos malos y posibles mejoras que se pueden realizar

3. Firmas de Aceptación

Se certifica o no que se cumplieron los objetivos de la iteración

Las historias de usuarios ____ (SI o No) cumplen con todos los criterios de aceptación y se da el visto bueno para que sea enviado a producción.

Elaborado por:

Nombre y Apellido		Cargo		Firma	
-------------------	--	-------	--	-------	--

Revisado por:

Nombre y Apellido		Cargo		Firma	
-------------------	--	-------	--	-------	--

Nombre y Apellido		Cargo		Firma	
-------------------	--	-------	--	-------	--

Aprobado por:

Nombre y Apellido		Cargo		Firma	
-------------------	--	-------	--	-------	--

Nombre y Apellido		Cargo		Firma	
-------------------	--	-------	--	-------	--

Revisión y retrospectiva de iteración

Anexo 4 – Matriz Data Warehouse Del SRI

Tipo de Procesos	Procesos de primer nivel	Procesos de segundo nivel	Hechos	Actividad Económica	Clase Contribuyente	Contribuyente	Distrito Aduanero	Estado Contribuyente	Estructura Organizacional	Fecha	Grupo Contribuyente	Impuesto	Institución Financiera	Moneda	País	Partida Presupuestaria	Periodo Fiscal	Prelación	Profesión	Régimen	Tipo Artesano	Tipo Contribuyente	Tipo Declaración	Tipo Documento	Tipo Formulario	Tipo Liquidación	Tipo Pago	Tipo Presentación	Tipo Recaudación	Ubicación Geográfica	Usuario
Cadena de Valor	Asistencia al Cumplimiento	Información y Educación Tributaria																													
Cadena de Valor	Asistencia al Cumplimiento	Atención de Peticiones y Consultas																													
Cadena de Valor	Asistencia al Cumplimiento	Gestión de Trámites																													
Cadena de Valor	Asistencia al Cumplimiento	Gestión de Notificaciones																													
Cadena de Valor	Asistencia al Cumplimiento	Derechos del Contribuyente																													
Cadena de Valor	Gestión del Cumplimiento Tributario	Catastro Tributario																													
Cadena de Valor	Gestión del Cumplimiento Tributario	Control del Catastro Tributario																													
Cadena de Valor	Gestión del Cumplimiento Tributario	Control de Transacciones y Mercancías																													
Cadena de Valor	Gestión del Cumplimiento Tributario	Control de Inconsistencias	Extensivos			X		X		X		X																			X
Cadena de Valor	Gestión del Cumplimiento Tributario	Control de la Veracidad																													
Cadena de Valor	Gestión del Cumplimiento Tributario	Gestión de Sanciones																													
Cadena de Valor	Gestión del Cumplimiento Tributario	Gestión de Cobranzas	Cartera			X		X		X	X				X	X	X					X									
Cadena de Valor	Gestión de los procesos	Gestión de Reclamos																													
Cadena de Valor	Gestión de los procesos Administrativos y Judiciales Tributarios	Gestión de Recursos de Revisión																													
Cadena de Valor	Gestión de los procesos Administrativos y Judiciales Tributarios	Gestión de los Procesos Judiciales Tributarios																													
Cadena de Valor	Gestión de Recaudación	Conciliación Tributaria	Recaudación			X	X		X	X	X	X	X			X	X			X		X	X		X	X	X	X	X	X	X
Cadena de Valor	Gestión de Recaudación	Acreditación a Terceros																													

Matriz Data Warehouse del SRI

Anexo 5 – Tabulación Encuesta - Situación Futura

Tabulación encuesta aceptación procedimiento propuesto

No	Pregunta	SI	NO
1	¿Considera efectivo contar con un formato que le permita describir sus necesidades de información en un lenguaje natural así como los criterios de aceptación que le permitan verificar el cumplimiento de su implementación?	9	1
2	¿Considera adecuado obtener resultados prioritarios y tangibles en intervalos cortos de tiempo durante la implementación de sus necesidades de información?	10	0
3	¿Considera efectivo que el Área de Datawarehouse implemente sus requerimientos de información siguiendo un lineamiento metodológico que garantice entregar información adecuada, precisa y oportuna?	9	1
4	¿Considera adecuado contar con una herramienta que permita visualizar los requerimientos de los usuarios de una manera priorizada y que permita estimar su esfuerzo así como el seguimiento de los entregables?	8	2
5	¿Considera adecuado que se norme el uso de estándares de desarrollo por medio de un procedimiento de desarrollo de productos de DWH que conlleve generar productos de alta calidad a beneficio de los usuarios de negocio?	8	2
6	¿Considera adecuado que el usuario funcional sea parte del equipo de desarrollo para solventar dudas durante la implementación de las soluciones de DWH?	7	3
7	¿Considera adecuado realizar reuniones de retrospectivas en la finalización de cada iteración junto con el usuario para describir las situaciones que se hicieron de manera inadecuada, resaltar las realizadas correctamente y proponer cambios en las siguientes iteraciones?	10	0
8	¿Considera adecuado contar con un procedimiento en el cual se especifiquen los pasos que se deben realizar para el diseño de la arquitectura, diseños lógicos y físicos, perfilamiento de datos, diseños y construcción de ETL, especificación y desarrollo de aplicaciones de BI, control de calidad así como los pasos para la implantación de la solución de DWH	9	1

9	¿Considera adecuado contar con la documentación referente a la implementación de los productos de DWH?	10	0
10	¿Considera adecuado contar con herramientas para visualizar el avance de las tareas que se requieren ejecutar para construir el producto de información requerido?	9	1

Tabulación encuesta situación futura por área

Área/Pregunta		1	2	3	4	5	6	7	8	9	10	Total
Centro Estudios Fiscales	SI	0	1	0	0	0	0	1	0	1	0	3
	NO	1	0	1	1	1	1	0	1	0	1	7
Control Tributario	SI	1	1	1	0	0	0	1	0	1	1	6
	NO	0	0	0	1	1	1	0	1	0	0	4
Grandes Contribuyentes	SI	1	1	1	1	1	1	1	1	1	1	10
	NO	0	0	0	0	0	0	0	0	0	0	0
Inteligencia Información	SI	3	3	3	3	3	3	3	3	3	3	30
	NO	0	0	0	0	0	0	0	0	0	0	0
Planificación Institucional	SI	1	1	1	1	1	0	1	1	1	1	9
	NO	0	0	0	0	0	1	0	0	0	0	1
Riesgos Tributarios	SI	3	3	3	3	3	3	3	3	3	3	30
	NO	0	0	0	0	0	0	0	0	0	0	0
Total	SI	9	10	9	8	8	7	10	8	10	9	88
	NO	1	0	1	2	2	3	0	2	0	1	12

Anexo 6 – Glosario de Términos

Área de Centro de Cómputo: Unidad técnica del Departamento de Operaciones de TI encargada de garantizar el funcionamiento y disponibilidad de las aplicaciones de negocio así como instalar los nuevos productos y servicios requeridos.

Área de Control de Calidad: Unidad técnica del Departamento de Redes de TI encargada de certificar las aplicaciones informáticas antes de ser desplegadas en ambientes productivos.

Área de Datawarehouse: Unidad técnica del Departamento de Desarrollo de TI encargada en la construcción de soluciones de información.

Área de Inteligencia de la Información: Unidad de negocio del Departamento de Planificación Institucional encargada de la gestión de la información del SRI.

Áreas de Negocio: Capa conceptual de Oracle Discoverer que contiene la información relacionada al cual tienen acceso los usuarios para la generación de reportes.

Ad hoc: Reporte dinámico o personalizados en tiempo real construidos por los usuarios del negocio.

Base de Datos Oracle: Base relacional encargada de almacenar grandes cantidades de datos y que posee varias funcionalidades asociadas a la gestión de base de datos relacionales, desarrollado por la Corporación Oracle.

Cadena de valor: Forma de análisis de la actividad empresarial que descompone a una empresa en sus partes constitutivas, buscando identificar fuentes de ventaja competitiva en aquellas actividades generadoras de valor.

Cambios tecnológicos: Cualquier modificación y/o creación de servicios o elementos de configuración (CI).

Capa Empresarial: Organización del modelo físico en una estructura de fácil acceso por los usuarios de negocio.

CASE: (Ingeniería de Software Asistida por Computadora), herramientas informáticas que apoyan a las tareas de desarrollo de software.

Ciclo de Vida de Ralph Kimball: Estable las fases de construcción de un Data Warehouse iniciando por el levantamiento de requerimientos, estableciendo la arquitectura requerida, definiendo los modelos lógicos y físicos, diseñando y construyendo los ETL, construyendo las aplicaciones de BI hasta la implementación de la solución desde la perspectiva dimensional.

CRM: Customer relationship management: Sistemas informáticos para la gestión de las relaciones con los clientes.

DOC: Formatos de archivos del paquete procesador de texto Microsoft Word.

Documento de estándares de base de datos de BI: Documento donde se definen los estándares de nomenclatura, documentación y programación PL SQL que deben cumplir los objetos de base de datos.

Documento de estándares de objetos de Data Services: Documento donde se establece la nomenclatura de los elementos usados en la construcción de un ETL así como la organización de dichos elementos.

Documento de Estándares de Universos: Documento donde se establece la nomenclatura de los elementos usados en el diseño y construcción de la infraestructura de datos y la carga empresaria que conforman un universo así como la organización de las carpetas y campos que son publicados.

Elemento de Configuración (CI): Un Elemento de Configuración es un activo tecnológico o componente de servicio.

Escala Fibonacci: Sucesión de números naturales 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89.

ERP: Enterprise resource planning: Sistemas de información gerenciales que automatizan los principales procesos de las organizaciones.

ETL: Proceso de extracción de datos desde las fuentes de información internas o externas, procesamiento de los datos para transformarlos en información y conocimientos.

Historias de usuario: Son las descripciones de las funcionalidades que va a tener la solución de información.

HTML: HyperText Markup Language, lenguaje de marcas de hipertexto, lenguaje de programación utilizado para la construcción de páginas web.

Infraestructura de Datos: Corresponde al modelo físico dimensional.

Instituciones Externas: Instituciones del Estado con las cuales el SRI tenga firmado convenios de intercambio de información.

ITIL: (Biblioteca de Infraestructura de Tecnologías de Información), marco de referencia para la gestión de servicios de tecnología.

MS Access: Sistema de gestión de bases de datos parte de las herramientas de Microsoft Office.

Modelos de información multidimensionales: Formas de almacenar eficientemente grandes volúmenes de información, corresponde a un conjunto de medidas descritas por dimensiones que permiten resumir y agrupar datos.

Modelo Estrella: Modelo dimensional que consta de una tabla de hechos la cual está relacionada a las dimensiones de análisis.

Obligaciones fiscales: Vínculo jurídico personal existente entre Estado y contribuyente para satisfacer una prestación en dinero, especies o servicios apreciables en dinero, al verificarse el hecho generador previsto por la Ley.

ODS: De las siglas en inglés Operational Data Store o almacenamiento de Datos Operacionales.

Oficina de Proyectos (ODP): Unidad de la Dirección de Planificación del SRI encargada de dar seguimiento y control a los planes, programas y proyectos institucionales.

OLAP: On Line Analytical Processing, orientado a sistemas estratégicos o de soporte de decisiones, encargados de almacenar datos históricos, parametrizadas para lecturas y escrituras masivas.

OLTP: On Line Transaction Processing, orientado sistemas transaccionales encargado de almacenar los datos recientes de la operación de las organizaciones, para inserciones y actualizaciones rápidas no masivas.

Oracle Discoverer: Herramienta de publicación de modelos multidimensionales de la empresa Oracle

Oracle Warehouse Builder: Herramienta de construcción de ETL de la empresa Oracle.

PL SQL: (*Procedural Language/Structured Query Language*) Lenguaje de programación estructura de Oracle.

Plan Nacional para el Buen Vivir: Plan nacional de desarrollo, mediante el cual el Gobierno Nacional articula las políticas públicas con la gestión y la inversión pública.

Procedimiento de Adquisiciones: Norma y estandariza el proceso de adquisición de bienes, ejecución de obras y prestación de servicios incluidos los de consultoría, tanto en régimen general como especial.

Procedimiento de Gestión de Cambios Tecnológicos: Asegurar el uso de procedimientos estandarizados que permitan manejar de manera eficaz y oportuna los cambios tecnológicos, minimizando el impacto que producen estos en el ambiente de producción.

SAP Business Object: Herramientas para la construcción de la infraestructura de datos y capa empresarial denominada Universos así como para la publicación gestión de permisos a las mismas.

SAP Data Services: Herramienta para la construcción de ETL que integra fuentes heterogéneas.

Servidor de Frontera: Servidor que se encuentra fuera de la DMZ y permite poner a disposición de las Instituciones Externas los datos que el SRI comparte.

Sistema Gestión TI: Aplicación informática para la gestión de la implementación de los elementos tecnológicos.

Solicitud de Cambio (SDC): Es una petición formal para cambiar uno o más elementos de configuración.

SQL Server: Sistema de administración de bases de datos relacionales desarrollado por Microsoft.

Sub-comité de Gestión Tecnológica: Es un grupo interdisciplinario que evalúa las solicitudes de cambio, de acuerdo al riesgo, impacto, planificación y calendarización requerida, toma decisiones sobre los cambios emergentes. Cualquier modificación y/o creación de servicios o elementos de configuración (CI). Se consideran también los cambios a la infraestructura física, suministro eléctrico y cableado sobre los Centros de Cómputo por su afectación a la disponibilidad de los servicios tecnológicos.

TOGAF: (*The Open Group Architecture Framework*), Marco de trabajo para el desarrollo de arquitecturas empresariales.

Universos: Conformado por la infraestructura de datos y la capa empresarial sobre la cual se dan accesos de uso y se construyen los reportes a los niveles operativo, analítico y estratégico.

Vistas Materializadas: Estructura de la base de datos Oracle en la cual se almacena los datos que retorna una sentencia de consulta y que se refresca cada cierto tiempo.

XML: Extensible Markup Language, lenguaje estándar de marcas utilizado para almacenar datos de una manera semiestructurada y es utilizado principalmente para intercambio de información.