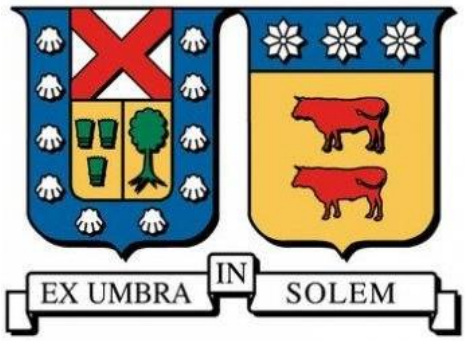


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE INDUSTRIAS

ANÁLISIS DEL FLUJO DE INFORMACIÓN B2B DE ABASTECIMIENTO ENTRE PROVEEDOR Y RETAILER; Y PROPUESTA DE DISEÑO DE UN MODELO DE INTEGRACIÓN Y GOBERNANZA DE DATOS B2B: CASO APLICADO TRESMONTES LUCCHETTI - WALMART

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERIA CIVIL INDUSTRIAL

AUTOR

CARLOS NARVÁEZ HIDALGO

PROFESOR GUÍA

GONZALO AMÉSTICA

SANTIAGO DE CHILE, 30 DE OCTUBRE DE 2026

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: ANÁLISIS DEL FLUJO DE INFORMACIÓN B2B DE ABASTECIMIENTO ENTRE PORVEEDOR Y RETAILER; Y PROPUESTA DE DISEÑO DE UN MODELO DE INTREGRACIÓN Y GOBERNANZA DE DATOS B2B: CASO APLICADO ENTRE TRESMONTES LUCCHETTI-WALMART

Nombre del candidato(a): Carlos Ruben Emilio Narváez Hidalgo

Carrera / Grado: Ingeniería Civil Industrial

Campus: Vitacura **Departamento:** Industrias

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Gonzalo Améstica Hernández, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad **(embargo)** por **(marcar una opción)**:

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

_____

_____

_____

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 30/10/2026 **Firma:** _____

Estudiante o Candidato(a):

Fecha: 30/10/2026 **Firma:** _____

Escaneado con CamScanner

1. Agradecimientos

Agradezco de todo corazón a mi mamá, a mi papá, a mi hermana y Natalia, por su paciencia y apoyo incondicional durante esta etapa de mi vida. Gracias por creer en mí, gracias por estar presentes en todo momento. Este no es solo un logro personal, sino que es un triunfo que jamás pude haber conseguido sin todo el apoyo que he tenido.

Quiero agradecer a todos esos amigos que he tenido y que logré formar durante esta etapa, en especial a los prestigio. Gracias por esas tardes interminables de estudio y de trabajo.

Finalmente quiero agradecer a todos los profesores por entregar su conocimiento y dedicación durante esta etapa, guiándome en mi formación profesional y personal.

Índice

1. Agradecimientos 2
2. Resumen Ejecutivo 7
3. Introducción 9
4. Problema de investigación 11
5. Objetivos 14
5.1 Objetivo General 14
5.2 Objetivos Específicos 14
6. Marco Teórico 15
6.1 Fundamento de la Cadena de Suministro Colaborativa y Logística B2B 15
6.1.1 El Paradigma de la Integración en las Cadenas de Suministro 15
6.1.2 Asimetría de la Información y Efecto Látigo (Bullwhip Effect) 16
6.2 Inteligencia de Negocios (BI) en la Gestión de la Cadena de Suministro Proveedor-Retailer 16
6.2.1 Evolución Histórica y Fundamentos de BI para SCM 16
6.3 Indicadores Abastecimiento y de Disponibilidad en Retail 19
6.3.1 Días de Inventario (DOH) 19
6.3.2 Instock Operacional 19
6.3.3 Fill Rate 20
6.4 Tratamiento de Datos, Procesos ETL. 20
6.4.1 Tipología de Datos en Entornos Corporativos 20
6.4.2 Procesos de Extracción, Transformación y Carga (ETL) 21
6.5 Gobernanza de la Información: Arquitectura y Modelamiento de Almacenes de Datos. 23
6.5.1 Data Lake: Ingestión y Flexibilidad de Datos Crudos 23
6.5.2 Data Warehouse: Consolidación e Integración Corporativa 24
6.5.3 Data Mart y Modelamiento Dimensional 25
6.5.4 Arquitectura de Medallón: Integración de las Etapas de Refinamiento del Dato 27
6.5.5 Motores de Bases de Datos Analíticas Embebidas 28
6.6 Metodología Hefesto para el Diseño de Almacenes de Datos 29
6.6.1 Lean Six Sigma (Metodología DMAIC) 29
6.6.2 Modelo en Cascada (Waterfall) 32
6.6.3 Modelo en V (Verificación y Validación) 33
6.6.4 Metodologías Ágiles (Framework Scrum) 35
6.7 Metodologías y Herramientas para la Evaluación y Modelamiento de Procesos 36
6.7.1 Modelamiento de Procesos de Negocio (BPMN) 37
6.7.2 Mapeo de la Cadena de Valor (Value Stream Mapping - VSM) 38
6.7.3 Capacidad Tecnológica del Repositorio de Datos 40

6.7.4 Modelo de Madurez de Inteligencia de Negocios (BIMM) - Modelo de Wayne Eckerson 41
6.7.5 El Modelo de Madurez de Gartner (Gartner Maturity Model for BI and Performance Management) 44
6.7.6 El Modelo de Desarrollo de BI (Business Intelligence Development Model - BIDM) 45
7. Antecedentes 48
7.1 Estructura Organizacional y Gobernanza Corporativa. 49
7.2 Modelo de Negocio y Diversificación de Canales 50
7.3 Canal Moderno y la Relevancia Estratégica de Walmart 51
7.4 El Ecosistema Tecnológico Actual 52
7.4.1 Sistema de Planificación de Recursos Empresariales (ERP) 52
7.4.2 Plataforma Externa de Inteligencia Operacional (Teamcore) 52
7.4.3 Portal Retail Link (WMC): Decision Support System 53
7.4.4 Herramienta de consolidación y preparación de reportes 54
7.5 Proceso General de Gestión y Flujo de Información para Indicadores de Abastecimiento 54
7.6 Indicadores Claves de Desempeño (KPIs) en la relación B2B TMLUC-Walmart 55
7.6.1 Fill Rate (Nivel de Servicio del Pedido) 56
7.6.2 Instock Operacional (Disponibilidad en Góndola) 56
7.6.3 Días de Inventario (DOH – Days On Hand) 57
8. Metodología 58
8.1 Enfoque del Diseño y Alcance del Modelo 58
8.1.1 Componente de Ingeniería de Datos 59
8.1.2 Alcance de Diagnóstico y Visibilidad 60
8.1.3 Criterio de no Intervención en la Fuente 60
8.1.4 Enfoque del Ciclo de Vida del Proyecto 61
8.2 Fase de Diagnóstico y Levantamiento de la Situación Actual 62
8.2.1 Levantamiento y Modelamiento del Proceso Mediante BPMN 63
8.2.2 Mapeo de la Cadena de Valor (Value Stream Mapping - VSM) 63
8.2.3 Capacidad Tecnológica del Repositorio de Datos 64
8.2.4 Análisis Causa Raíz de Problemas 65
8.2.5 Diagnóstico de Madurez de Inteligencia de Negocios 66
8.3 Procedimiento para el Diseño del Modelo B2B (Basado en el Modelo en V) 67
8.3.1 Procedimiento de Especificación de Requerimientos e Ingesta Híbrida 67
8.3.2 Procedimiento de Diseño Dimensional (Constelación de Hechos) 67
8.3.3 Procedimiento para el Diseño de la Arquitectura de Datos (Capa Medallón) 68
8.3.4 Procedimiento de Especificación del Detalle Tecnológico 68
8.4 Protocolo de Verificación y Pruebas (Rama Derecha del Modelo en V) 69
8.4.1 Pruebas Unitarias (Compuerta 1) 69

8.4.2 Pruebas de Integración (Compuerta 2) 70
8.4.3 Pruebas del Sistema: Nivel de Especificación Funcional (Compuerta 3) 70
8.4.4 Pruebas de Aceptación del Usuario (Compuerta 4) 71
8.5 Procedimiento de Evaluación Económica 71
9 Análisis de la Situación Actual y Resultado de la Propuesta de Solución 73
9.1 Análisis de la Situación Actual 73
9.1.1 Fase A: Extracción Manual y Desincronizada de Datos (Latencia Inicial) 73
9.1.2 Fase B: Consolidación y Cruce en Excel (Fase Crítica) 74
9.1.3 Fase C: Bifurcación de Visualización 75
9.1.4 Fase D: Distribución y Consultas Ad-hoc (Bucle de Pedidos) 76
9.1.5 Fase E: Verificación y Ciclo de Retrabajo (Muda de Error) 76
9.1.6 Desglose Actividades en Diagrama BPMN de la Situación Actual de Flujo de Información TMLUC-WMC 77
9.2 Cuantificación de Ineficiencias y Análisis de Criticidad 82
9.2.1 Indicador 1: Fuga de Valor de Ingeniería 83
9.2.2 Indicador 2: Eficiencia del Ciclo del Proceso (PCE) 83
9.2.3 Indicador 3: Capacidad y Límite Tecnológico 84
9.3 Análisis del Potencial Desaprovechado y Ceguera Operativa 87
9.3.1 Problema 1: El Sistema de Información Artesanal 87
9.3.2 Problema 2: Ceguera Operativa Longitudinal 88
9.3.3 Problema 3: Fragmentación de la Única Verdad 89
9.3.4 Problema 4: Ciclo de Retrabajo Crítico 89
9.3.5 Problema 5: Pérdida de Información por Fin de Semana, Vacaciones, Motivos Exógenos 89
9.4 Análisis Causa Raíz y Oportunidad de Mejora 90
9.4.1 Problema 1: El Sistema de Información Artesanal 90
9.4.2 Problema 2: Ceguera Operativa Longitudinal 91
9.4.3 Problema 3: Fragmentación de la Única Verdad 92
9.4.4 Problema 4: Ciclo de Retrabajo Crítico 93
9.4.5 Problema 5: Pérdida de Información por Fin de Semana, Vacaciones, Motivos Exógenos. 93
9.4.6 Síntesis de Causa Raíz 94
9.5 Diagnóstico de Madurez de Business Intelligence según el Modelo de Wayne Eckerson 95
9.5.1 Clasificación en la Etapa 1: Prenatal / Infant (La "Infancia" Analítica) 95
9.5.2 Justificación de la Exclusión de Estados Superiores 96
9.5.3 El Desafío del Punto de Fricción Gulf (El Golfo) 97
9.6 Resultado del Desarrollo de la Solución 98
9.6.1 Nivel 1 del Modelo en V: Análisis de Requisitos del Negocio y Catálogo de Fuentes 99
9.6.2 Nivel 2 del Modelo en V: Especificación Funcional y Diseño Dimensional 103

9.6.3 Nivel 3 del Modelo en V: Arquitectura del Almacén de Datos (Capa Lógica) 106
9.6.4 Nivel 4 del Modelo en V: Especificación del Detalle Tecnológico 110
9.6.5 Resultado del Tablero de Control Implementado 115
9.6.6 Resultado del Protocolo de Verificación y Pruebas 116
9.6.7 Síntesis del Proceso B2B Rediseñado 118
9.7 Evaluación Económica de la Solución 119
9.7.1 Costo Total de Propiedad (TCO) 119
9.7.2 Beneficio Económico 121
9.7.3 Indicadores de Rentabilidad 121
9.7.4 Análisis de Sensibilidad 123
9.7.5 Síntesis de la Evaluación Económica 124
10 Conclusiones y Recomendaciones 125
10.1 Conclusiones 125
10.2 Recomendaciones 127
11 Discusión 130
12 Limitaciones 134
13 Referencias 137

2. Resumen Ejecutivo

TMLUC gestiona el flujo de información de abastecimiento (Ventas, Stock y Forecast) con Walmart Chile mediante un proceso manual, el Analista CPFR descarga diariamente reportes planos desde el portal Retail Link y los consolida manualmente en planillas de cálculo, un procedimiento que consume 8 horas semanales y expone a la organización a un riesgo de error estructural de entre el 20% y el 40% propio del uso de hojas de cálculo para volúmenes masivos de datos (Davenport, 2006), además de una ceguera operativa longitudinal que impide sostener más de 1 a 3 días de información histórica analizable.

Frente a este problema, el presente estudio tuvo como objetivo diseñar un modelo de integración y gobernanza de datos B2B que automatizara la captura, transformación y consolidación de la información de Ventas, Stock y Forecast proveniente de Walmart, con el fin de mejorar la integridad, disponibilidad y oportunidad del dato granular Producto-Sala para la operación de TMLUC.

Metodológicamente, se diagnosticó la situación actual mediante el modelado de procesos (BPMN), el Mapeo de Cadena de Valor (VSM) y el Análisis Causa Raíz (5 Porqué e Ishikawa), para luego diseñar e implementar la solución siguiendo los cuatro niveles del Modelo en V, especificación de requerimientos, diseño dimensional, arquitectura de datos y detalle tecnológico, validando cada nivel mediante un protocolo de verificación de cuatro compuertas secuenciales.

El diagnóstico reveló que el 77% del tiempo del proceso se destinaba a tareas

manuales de normalización de datos. Esto equivale a una Fuga de Valor de Ingeniería de $272.000 mensuales y a una Eficiencia del Ciclo del Proceso de apenas 23%, y ubica a TMLUC en la Etapa 1 (Prenatal/Infant) del Modelo de Madurez de Inteligencia de Negocios de Eckerson. La solución implementada, una Arquitectura de Medallón sobre un mecanismo híbrido de ingesta vía API y RPA, fue validada mediante un protocolo de verificación que demostró una coincidencia exacta y sostenida, durante cinco días consecutivos y sin registrar diferencias, entre el modelo dimensional construido y la fuente nativa de Walmart. La evaluación económica valorizó el desarrollo con el ingreso promedio de mercado de la carrera y estimó un Costo Total de Propiedad de $8.460.604 a 36 meses, con un período de recuperación de 30,2 meses y un retorno sobre la inversión de 16%, considerando únicamente la capacidad liberada del Analista CPFR. Bastaría con una mejora del Instock de 0,0024 puntos porcentuales para cubrir el costo anual de la solución.

El aporte de esta investigación consiste en resolver el prerrequisito técnico que permite a TMLUC avanzar hacia etapas superiores de madurez de Inteligencia de Negocios, liberando al Analista CPFR de la carpintería de datos para habilitar un rol de análisis estratégico del abastecimiento, y sentando una base arquitectónica replicable hacia otras cuentas del canal moderno.

Palabras Clave: Integración de Datos B2B, Gobernanza de Datos, Arquitectura de Medallón, Modelo en V, Inteligencia de Negocios, Cadena de Suministro.

3. Introducción

La gestión eficiente de la cadena de suministro depende, en gran medida, de la calidad y oportunidad de la información que las organizaciones comparten con sus socios comerciales. En el contexto del canal moderno de retail, esta dependencia se vuelve crítica para los proveedores que, como Tresmontes Lucchetti (TMLUC), deben sincronizar su operación de abastecimiento como los sistemas y plataformas de cadenas de Walmart Chile, bajo relaciones de colaboración que exigen altos estándares de trazabilidad y disponibilidad del dato.

El presente documento aborda el diseño e implementación de un modelo de integración y gobernanza de datos B2B para la relación comercial entre TMLUC y Walmart, orientado a sustituir un proceso de tratamiento manual de la información por una arquitectura de datos automatizada, trazable y verificable. El desarrollo de esta propuesta responde a un diagnóstico previo de las ineficiencias operativas y los riesgos estructurales que el proceso manual actual genera para la organización, los cuales se describen en detalle en el cuerpo de este informe.

Para ello, el presente documento se estructura en los siguientes capítulos. El capítulo 4 presenta el Problema de Investigación que motiva este estudio. El capítulo 5 define los Objetivos Específicos que orienten el desarrollo del proyecto. El capítulo 6 expone el Marco Teórico que sustenta las herramientas y modelo conceptuales utilizados. El capítulo 7 describe los Antecedentes de la organización y de la relación comercial estudiada. El capítulo 8 detalla la Metodología aplicada para le diagnóstico y el diseño de la solución. El capítulo 9 presenta los Resultados obtenidos, tanto el diagnóstico de la situación actual como

de la implementación y verificación de la solución propuesta. Finalmente, los capítulos 10, 11 y 12 exponen, respectivamente, las Conclusiones, Recomendaciones, la Discusión de los hallazgos a la luz del Marco Teórico, y las Limitaciones reconocidas en el desarrollo de esta investigación.

4. Problema de investigación

En el marco de la industria de productos de consumo masivo, la integración digital del flujo de información entre proveedores y grandes retailers ha dejado de ser una ventaja competitiva para convertirse en un requisito de subsistencia dentro de la industria. Para Tresmontes Luchetti (TMLUC), la gestión del abastecimiento de las salas de la cadena Walmart (Walmart Chile) depende críticamente de la capacidad de transformar datos brutos en decisiones precisas. El núcleo del problema radica principalmente en la ineficiencia del proceso de captura, integración y gobernanza de datos B2B, la cual, actualmente depende de una ejecución manual.

Actualmente, Walmart genera a diario un reporte de su Stock MultiEstado, Venta Sell-Out, Forecast, entre otros, dicha información es cargada en el portal Retail Link de Walmart para que TMLUC pueda procesarla. Esta información es la única fuente de visibilidad que tiene TMLUC sobre el flujo de sus productos desde que se entregan en los Centros de Distribución (CD) de Walmart, abarcando las ventas, forecast y el stock en los distintos estados en los cuales se pueden encontrar sus productos: Stock Pedido por cada Sala (On Order), Stock en Centro de Distribución en preparación para despacho a Salas (In Whse), Stock en Tránsito hacia las Salas (In Transit) y Stock en Sala (On Hand).

A pesar de que la información es generada diariamente por Walmart el proceso de bajada y consolidación de estos reportes a nivel de Producto-Sala es realizado de forma manual por el Analista CPFR. Esta metodología de trabajo genera una latencia de información que degrada el valor del dato. Según Christopher (2016), la visibilidad en tiempo real es el sustituto moderno de inventario físico (p. 178), cuando el acceso al dato depende de procesos analógicos, se genera una brecha temporal que impide a la organización actuar con la velocidad que el mercado exige.

Académicamente, según Lee et al., (1997), esta deficiencia en el flujo de datos alimenta

el efecto látigo (p. 546-547), además la distorsión y el retraso en la transmisión de información de ventas y existencias provocan fluctuaciones innecesarias en los pedidos, derivando quiebres de stock en góndola o sobrestock inmovilizado.

El impacto de esta problemática se evidencia en el desempeño operacional acumulado durante el año 2025, periodo en el cual TMLUC registró un indicador de instock promedio del 95,84%, adicionalmente mantuvo un promedio diario de 1.617 combinaciones producto-sala quebrados por día. Esta falta de disponibilidad oportuna de cara al consumidor se tradujo en una venta perdida estimada de $4.883.357.593 debido a quiebres de stock en los puntos de venta. Bajo la arquitectura actual, la ausencia de un procesamiento automático impide contrarrestar diaramente el forecast , las ventas y el Stock MultiEstado, imposibilitando realizar planes de acción tras el cálculo de KPIs de gestión granular y aumentando el riesgo de penalizaciones financieras por incumplimiento de niveles de servicio.

Actualmente, en términos operativos el tratamiento manual de la información de abastecimiento de las salas le consume al Analista CPFR en promedio 90 minutos diarios, lo que equivale 8 horas semanales, de forma agregada representa aproximadamente a una semana hábil al mes dedicada al tratamiento manual de información. Esta deficiencia representa una barrera crítica para la adopción de herramientas de Inteligencia de Negocios (BI).

La falta de automatización significa que el recurso humano de ingeniería se consume en la preparación de datos en lugar de análisis estratégico de los indicadores.

Tal como sostiene (Davenport, 2006), la carencia de flujos de datos integrados y estandarizados constituye la principal barrera estructural que impide a las organizaciones progresar hacia estados superiores de madurez analítica (p.101-103). Al carecer de esta base tecnológica centralizada, las empresas permanecen confinadas en niveles operativos de escaso valor , lo que les

impide explotar estratégicamente su información mediante soluciones de BI.

En consecuencia se identifica la necesidad de diseñar un modelo de integración y gobernanza de datos B2B. Este proyecto no busca alterar la lógica de negocio ni crear modelos de predicción, en cambio, si pretende mejorar el motor y la visualización de datos. El objetivo es garantizar la integridad, disponibilidad y oportunidad de la información a nivel Producto-Sala, permitiendo que TMLUC transite hacia una excelencia operativa basada en datos reales y procesos automatizados en el contexto de su relación comercial con Walmart para mejorar la toma de decisiones.

5. Objetivos

5.1 Objetivo General

Diseñar un modelo de integración y gobernanza de datos B2B que automatice la captura, transformación y consolidación de la información en la última línea de la cadena de suministro (Ventas, Stock y Forecast) proveniente de Walmart, con el fin de mejorar la integridad, disponibilidad y oportunidad del dato granular (Producto-Sala) para la gestión de Tresmontes Lucchetti.

5.2 Objetivos Específicos

Diagnosticar el flujo actual de información para identificar cuellos de botella, latencia temporal en el proceso manual de extracción de datos desde Retail Link y cuantificar el impacto en la carga operativa del equipo analítico.

Establecer estándares de calidad, reglas de negocio y arquitectura lógica para asegurar que los datos mantengan su integridad durante todo el flujo de información.

Diseñar el modelo lógico para la extracción y procesamiento automático de datos (ETL), utilizando herramientas de bajo código o programación para eliminar la intervención manual.

Diseñar e Implementar tableros de control que consoliden la información integrada, permitiendo la visualización en tiempo real de los indicadores operativos a nivel Producto-Sala para la toma de decisiones.

6. Marco Teórico

6.1 Fundamento de la Cadena de Suministro Colaborativa y Logística B2B

Christopher (2016) sostiene que las cadenas de suministro modernas han trascendido los modelos fragmentados y puramente transaccionales para consolidarse como ecosistemas dinámicos e interdependientes. Particularmente en el sector de bienes de consumo masivo, la relación entre proveedores de alimentos y retailers demanda una sincronización operativa de alto nivel, esta integración es la que permite maximizar la disponibilidad de productos en góndola al mínimo costo posible. (p. 537)

6.1.1 El Paradigma de la Integración en las Cadenas de Suministro

Históricamente, los eslabones de una cadena de suministro operaban bajo un enfoque de silos, donde cada organización se preocupaba de maximizar sus propios procesos de almacenamiento y transporte de forma aislada.

Como señalan Mentzer et al. (2001), el paradigma tradicional de gestión basado en silos operativos ha sido superado por un modelo de integración sistémica. Bajo este nuevo esquema, la competitividad de una firma trasciende la mera eficiencia interna y se fundamenta en su capacidad para colaborar con socios estratégicos y adaptarse con flexibilidad a la volatilidad del entorno (pp. 6-10).

La integración en el ámbito de la logística entre proveedor-retailer, Business To Business (B2B) se fundamenta principalmente en el flujo bidireccional y continuo de la información, desde el inicio hasta el final y viceversa de la cadena de suministro.

En la relación proveedor-retailer la integración de la información permite transitar desde una lógica reactiva o una gestión proactiva orientada a la demanda real de consumidor final.

6.1.2 Asimetría de la Información y Efecto Látigo (Bullwhip Effect)

Un desafío crítico en las relaciones entre proveedor y retailer es la distorsión de la información a través de eslabones. El Efecto Látigo, conceptualizado por Lee et al. (1997) describe como pequeñas variaciones en la demanda del consumidor final se amplifican exponencialmente a medida que el requerimiento asciende hacia el proveedor. Este fenómeno es provocado principalmente por la falta de transparencia en los datos de ventas reales (Sell – Out) y el procesamiento de señales de demanda aisladas (pp. 546-547).

Algunas consecuencias logísticas y financieras de no mitigar este fenómeno son las siguientes:

- Sobredimensionamiento de la producción: Ante la falta de datos en los puntos de venta y la falta de análisis de estos, el proveedor sobredimensiona los lotes de fabricación para cubrir picks de demanda artificiales.

- Quiebres de Stock: Las fluctuaciones y asimetrías en la información provocan periodos de desabastecimiento.

- Exceso de Stock de Seguridad: La falta de visibilidad de la información provoca exceso de stock de seguridad, lo cual incrementa los costos de almacenamiento y mantenimiento de inventario. Esto se traduce de forma directa en ineficiencia en capital de trabajo y mermas de por vencimiento de productos.

6.2 Inteligencia de Negocios (BI) en la Gestión de la Cadena de Suministro Proveedor-Retailer

6.2.1 Evolución Histórica y Fundamentos de BI para SCM

El concepto de Inteligencia de Negocios (Business Intelligence o BI) no constituye una corriente tecnológica emergente, sino que es el resultado de una evolución sistemática de los

Sistemas de Soporte a la Decisión (DSS).

Formalmente Negash (2004) define la Inteligencia de Negocios como un sistema que integra la captura, el resguardo y la administración del conocimiento junto con herramientas de análisis avanzado. Esta arquitectura tecnológica tiene como fin último procesar flujos de información compleja para generar conocimiento estratégico, facilitando la toma de decisiones informadas y el desarrollo de ventajas competitivas sostenibles (pp. 177-178).

Por otro lado, al definir la Inteligencia de Negocios (Business Intelligence, BI), Cano (2018) la caracteriza como un proceso iterativo diseñado para explorar y analizar información estructurada sobre un área específica, con el fin de descubrir tendencias o patrones y derivar de ellos ideas y conclusiones estratégicas (p. 345).

Bajo este enfoque, el BI provee a las organizaciones capacidades fundamentales: el conocimiento de los factores operativos internos, el análisis del comportamiento de los clientes y socios comerciales.

Cuando los fundamentos analíticos del BI se trasladan al contexto de la gestión de la cadena de suministro (Supply Chain Management) y a los canales de distribución masiva, el foco sistémico vira hacia lo que la literatura técnica denomina Inteligencia de Negocios Operacional (Operational BI).

De acuerdo con Cano (2018), la adopción de soluciones de BI en la cadena de suministro es un imperativo estratégico en mercados de consumo masivo, caracterizado por márgenes reducidos y penalizaciones críticas por quiebres de stock. El autor sostiene que los flujos de información convencionales padecen una fragmentación estructural, mientras los minoristas poseen la visibilidad directa del punto de venta, los proveedores operan bajo una ceguera operativa, gestionando su capacidad a partir de órdenes de compra que reflejan inventarios

rezagados y no la demanda real del consumidor (pp. 26-30). Por lo tanto, según el autor el BI permite a las organizaciones responder interrogantes fundamentales, tales como: ¿Qué ocurrió?, ¿Por qué ocurrió? y ¿Qué acciones deben tomarse para influir en el futuro?

La integración digital entre proveedores y grandes retailers se sustenta en que, cuando el acceso al dato presenta latencia por procesos manuales, se alimenta el fenómeno conocido como Efecto Látigo.

En este sentido el BI mitiga esta desconexión mediante la construcción de un lenguaje de datos unificado entre ambas corporaciones.

En el ámbito académico hispanohablante Curto (2010) profundiza en los desafíos de implementación de Business Intelligence. El autor enfatiza que el reto fundamental para las organizaciones no radica en la selección de un software de visualización de información, si no en la capacidad de ingeniería interna para orquestar la compleja infraestructura de datos e integración que subyace al sistema (pp. 28-30).

En la gestión de la cadena de suministro, el BI es la herramienta que permite alcanzar una posición llamada, la única verdad corporativa. Este concepto implica la unificación de criterios de cálculo y definiciones de datos entre departamentos, garantizando que todos los actores tomen decisiones basadas en indicadores consistentes y extraídos de un repositorio centralizado. Como sostiene (Davenport, 2006), las organizaciones que logran estandarizar sus flujos de datos fundamentales transitan desde un nivel operativo de escaso valor hacia una madurez analítica superior, permitiendo competir a través del análisis de datos (pp. 103).

Dentro del diseño de una arquitectura de BI, el cálculo del Costo Total de Propiedad (TCO) es imperativo para validar la rentabilidad del proyecto. Tal como describe Cano (2018), para evaluar correctamente la rentabilidad de un proyecto de BI es imperativo calcular el TCO,

el cual no se limita a la adquisición de tecnología, sino que integra los costos de operación y el tiempo del personal interno (p. 48). El autor advierte que las organizaciones con baja madurez analítica operan bajo una ineficiencia estructural, donde los profesionales desperdician un 80% de su jornada en tareas manuales de normalización de datos, sacrificando el tiempo destinado al análisis estratégico (p. 20). Esta desproporción justifica la necesidad de automatización para comprimir el lead time de la información y transformar la visibilidad de datos en un sustituto estratégico del inventario físico.

6.3 Indicadores Abastecimiento y de Disponibilidad en Retail

La medición de desempeño mediante indicadores clave (KPIs) es esencial para garantizar la salud operativa de cualquier proceso. La mejora de la productividad requiere un diagnóstico objetivo de la relación entre los recursos utilizados y salidas generadas.

6.3.1 Días de Inventario (DOH)

Este indicador mide la eficiencia de la gestión del stock, cuantificando el tiempo que el capital de la empresa permanece inmovilizado en forma de mercancía. Un control deficiente del DOH impacta negativamente en el flujo de caja y eleva el riesgo de obsolescencia y merma.

$$DOH_{ps} = \frac{OH_{ps}}{FD_{ps}}$$

El OH_{ps} y el FD_{ps} representan el inventario disponible a la mano (On Hand) y el forecast diario para una combinación producto p y sala s, respectivamente.

6.3.2 Instock Operacional

Representa la métrica de éxito definitiva en el entorno del comercio minorista. El Instock mide la disponibilidad efectiva del producto en la góndola en el momento de la compra. Un quiebre de stock no solo implica una pérdida financiera inmediata, si no que erosiona la lealtad del cliente hacia la marca y el establecimiento.

$$I_{ps} = \begin{cases} 1, & si\ OH_{ps} \geq FD_{ps} \\ 0, & si\ OH_{ps} < FD_{ps} \end{cases}$$

Donde I_{ps} representa el estado de Instock para un producto p en una sala s, siendo 1 cuando hay disponibilidad y 0 cuando hay quiebre de stock para la combinación ps.

6.3.3 Fill Rate

El Fill Rate o Nivel de Servicio del Pedido, mide la eficacia con la que un proveedor satisface una orden de compra específica emitida por el retailer, a diferencia del Instock, que mide la disponibilidad continua del producto en góndola frente a la demanda del consumidor final. Un Fill Rate deficiente refleja problemas en la planificación de la producción o en la logística de despacho del proveedor, y suele derivar en penalizaciones contractuales por parte del retailer.

$$FR = \frac{Unidades\ Entregadas}{Uindades\ Solicitadas}\%$$

Donde FR representa el % de las unidades que solicitó el retailer y el proveedor fue capaz de entregarlas en tiempo y forma.

6.4 Tratamiento de Datos, Procesos ETL.

6.4.1 Tipología de Datos en Entornos Corporativos

Para la estructuración de un modelo analítico robusto, Negash (2004) sostiene que es imperativo diferenciar la información según su tipología técnico estructural, clasificándola en tres categorías fundamentales (pp. 180-181).

- Datos estructurados: Corresponde a aquella información que se alinea estrictamente a un modelo de datos predefinido y a esquemas relacionales rígidos. Su principal ventaja radica en que permite realizar búsquedas indexadas veloces, optimizar el almacenamiento y automatizar la generación de reportes operativos.

- Datos no estructurados: Representan información que carece de un modelo conceptual o sintaxis nativa predefinida, siendo inherentemente complejos de procesar. Requieren de tecnologías de almacenamiento no relacionales y modelos analíticos avanzados para su explotación. Algunos ejemplos de datos no estructurados son: registros de procesos de negocio, chats, correos, noticias, fotos, videos, entre otros. Estos datos resultan críticos para comprender variables cualitativas como el comportamiento, opiniones, fricciones de los clientes y colaboradores.

- Datos semi-estructurados: Definen a aquellos activos de información que, si bien no se ajustan a un modelo relacional estrícto, contienen marcadores, etiquetas o metadatos organizativos que imponen una jerarquía interna, por ejemplo archivos XML, JSON o bitácoras de transferencia electrónica.

Citando estimaciones de Merrill Lynch Negash (2004), uno de los retos principales para los sistemas de BI es que aproximadamente el 85% de la información corporativa reside en formatos semiestructurados o no estructurados. Esta predominancia de datos invisibles para las arquitecturas tradicionales subraya la necesidad de implementar herramientas analíticas capaces de capturar y procesar flujos de información complejos para la toma de decisiones (p. 181).

6.4.2 Procesos de Extracción, Transformación y Carga (ETL)

Los portales B2B de las cadenas de supermercados, como Retail Link de Walmart, actúan como repositorios de datos masivos y crudos que imponen barreras de acceso debido a interfaces de usuarios complejas y formato de descarga rígidos. El BI aplicado exige que la organización proveedora diseñe un motor de datos intermediario que automatice la ingesta y el consumo de estas interfaces.

La validez analítica de cualquier cuadro de mando (dashboard) de BI está estrictamente

condicionada por la calidad del flujo de datos que alimenta sus métricas. En la ingeniería de datos, este principio metodológico está gobernado por el axioma fundamental Garbage In, Garbage Out (Si entra basura, sale basura). Para asegurar que las visualizaciones analíticas reflejen con absoluta fidelidad la realidad operativa de toda cadena de suministro entre proveedor y retailer, la información extraída desde el portal B2B debe someterse a un riguroso proceso de Extracción, Transformación y Carga (ETL).

La fase de transformación constituye la fase más crítica en la gobernanza de datos orientada al soporte de decisiones logísticas. Mientras que la captura y la carga se limitan a la captura física y al almacenamiento final de los registros. La transformación es el espacio lógico donde se inyecta la ingeniería de procesos. Según la Asociación Internacional de Gestión de Datos DAMA-International (2020), es en este punto donde se aplican las reglas de datos y de negocio, las cuales se estructuran en dos dimensiones fundamentales.

- Reglas de integridad técnica: Orientadas a la limpieza, control de duplicidad, completitud y estandarización estructural de dato bruto.
- Reglas de lógica operativa: Encargadas de procesar, normalizar y calcular los indicadores clave de rendimiento (KPI) bajo los criterios específicos de la organización.

A nivel de integridad técnica, la primera regla de gobernanza que debe imponer un modelo de datos diseñado para una organización en donde los datos provienen de un portal externo B2B es el control de redundancia. Debido a que la carga diaria desde un portal externo presenta un riesgo inminente de solapamiento de archivos, el motor de base de datos utilizado debe aplicar restricciones de clave primaria. Esta restricción garantiza de manera matemática la unicidad de registros, evitando distorsiones por duplicidad en los datos agregados.

Una vez consolidada la fiabilidad y unicidad de la data cruda, la infraestructura de BI

ejecuta la fase de transformación avanzada, aplicando algoritmos lógicos matemáticos que traducen los estados de inventario planos provistos en alertas de gestión proactiva. En la relación proveedor-retailer, se establecen reglas operacionales clave que procesan la relación entre la información disponible.

6.5 Gobernanza de la Información: Arquitectura y Modelamiento de Almacenes de Datos.

La calidad del dato constituye uno de los pilares centrales de la gobernanza de datos, definida por DAMA-International (2020), como el conjunto de dimensiones que determinan si un dato es apto para su uso en la toma de decisiones. Entre las dimensiones fundamentales de este marco se encuentran, la completitud, que mide si están presentes todos los valores de datos requeridos; la precisión (o exactitud), que evalúa el grado en que el dato representa correctamente el objeto o evento del mundo real que describe; la consistencia, referida a la uniformidad del dato a través de distintos sistemas y momentos de captura; y la unicidad, que mide el grado en que un registro ocurre una sola vez dentro de un conjunto de datos, evitando duplicidades.

El almacenamiento sistemático y estructurado de los datos constituye la piedra angular para mitigar el efecto látigo en la cadena de suministro proveedor-retailer. Para transformar los registros transaccionales en conocimiento accionable, la literatura de ingeniera de datos define una arquitectura de tres niveles complementarios: Data Lake, Data Warehouse y Data Mart.

6.5.1 Data Lake: Ingestión y Flexibilidad de Datos Crudos

De acuerdo con Gorelik (2019), el Data Lake (DL) se conceptualiza como un repositorio centralizado diseñado para albergar datos estructurados, semiestructurados y no estructurados a gran escala. La ventaja estratégica de esta arquitectura reside en su capacidad para conservar información en su formato nativo, eliminando la rigidez de una estructura previa y permitiendo que los datos permanezcan en su estado natural hasta que sean requeridos para el análisis (pp.1-

2).

En la relación B2B proveedor-retailer, el DL actúa como zona de aterrizaje donde se depositan diariamente los archivos planos extraídos desde el portal externo. La implementación del DL reduce los costos de almacenamiento inicial al evitar transformaciones costosas en tiempo de ejecución e ingesta. Permite retener históricos masivos de transacciones a nivel desagregado, sin restricciones de espacio, sirviendo como insumo base para futuros modelos predictivos de demanda.

El DL no exige un formato rígido de entrada para la información, esto permite que se elimine por completo la latencia de ingesta. Si el portal externo desde donde se extrae la información cruda cambia repentinamente la estructura de su reporte (por ejemplo, agrega una nueva columna), la ingesta automática no se cae ni se corrompe. Esto reduce el costo operativo de mantenimiento de TI y asegura que nunca se pierda el histórico de la información por incompatibilidad de formatos.

En su análisis de arquitecturas de Big Data, Gorelik (2019) señala que la capa de ingesta de datos debe estar completamente desacoplada de los procesos de transformación complejos, lo que garantiza una recolección ligera y de bajo impacto en los sistemas de origen.

6.5.2 Data Warehouse: Consolidación e Integración Corporativa

Luego del DL la información se debe cargar en un espacio centralizado y preparado para su almacenamiento. De acuerdo con la definición clásica de Inmon (2005), el Data Warehouse (DW) representa una infraestructura crítica en la gestión de datos corporativos. Se conceptualiza como un repositorio de datos orientado a temas, integrado, variante en el tiempo y no volátil, diseñado precisamente para robustecer la toma de decisiones empresariales (pp. 29-31). En cuanto a su arquitectura interna, el modelo de Inmon se fundamenta tradicionalmente en la

Tercera Forma Normal (3NF) para garantizar la integridad, aunque paradigmas contemporáneos proponen el uso de Data Vault para incrementar la flexibilidad ante cambios en el negocio.

Mientras que el DL retiene los datos de forma aislada, el DW los unifica con el objetivo de integrar corporativamente sin duplicación. En este punto los datos transaccionales del retailer se pueden cruzar con el ERP interno del proveedor (catálogo productos, información de salas, clientes, costos logísticos, etc.). Al almacenar de forma sistemática esta información en un Sistema de Gestión de Bases de Datos Relacionales (RDBMS), la organización proveedora elimina la asimetría que perjudica la planificación operativa.

El DW rompe los archivos planos en múltiples tablas altamente interconectadas mediante llaves primarias y foráneas para evitar redundancias. Este modelo garantiza la consistencia y la integridad referencial absoluta del negocio.

6.5.3 Data Mart y Modelamiento Dimensional

Debido a la magnitud y complejidad técnica de un almacén de datos corporativo centralizado, el acceso directo por parte de los analistas logísticos suele comprometer la agilidad y eficiencia de las consultas. Para mitigar esta barrera, Kimball y Ross (2002) proponen la arquitectura de Data Mart (DM), conceptualizados como subconjuntos del área de presentación orientados a procesos de negocio únicos o canales específicos, esta estructura permite transformar datos atómicos en información comprensible y de alto rendimiento para la toma de decisiones (pp. 10-11).

Para un canal B2B entre proveedor-retailer se diseña un DM específico parametrizado bajo un modelo dimensional. Este modelo organiza la información en tablas de hechos (stock, venta, forecast, órdenes de compra, costos logísticos, etc.) interconectadas con tablas de dimensiones (calendario, productos, sala, etc.).

Esta infraestructura base habilita la visualización interactiva y dinámica mediante herramientas de visualización modernas de BI. La combinación de un motor de base de datos optimizado para consultas analíticas con una capa de visualización ágil democratiza el uso del dato dentro de la organización para el consumo analítico rápido y la toma de decisiones. Los analistas dejan de actuar como operarios de extracción manual de datos y se transforman en ingenieros de diagnóstico capaces de navegar desde vistas agregadas de familias de productos hasta el detalle específico de un SKU en una sala determinada.

Al proveer consultas con tiempo de respuesta ínfimos, el DM permite automatizar alertas de quiebre diarias a nivel de sala.

El DM aplica una desnormalización de los datos deliberadamente para organizarlos en modelos dimensionales optimizados para herramientas de visualización rápida.

- Modelo Estrella (Star Schema): Consiste en una tabla de hechos central que registra eventos cuantitativos, unida directamente a tablas dimensionales a través de llaves primarias y foráneas, reduciendo las uniones (joins) en las consultas y acelerando el cálculo de KPIs a milisegundos.
- Modelo Copo de Nieve (Snowflake Schema): Extensión del modelo estrella, donde las tablas de dimensiones se normalizan en subjerarquías hasta alcanzar la 3NF, reduciendo la redundancia de almacenamiento a costa de una mayor complejidad computacional.
- Constelación de Hechos (Fact Constellation Schema): Arquitectura avanzada que contiene múltiples tablas de hechos que comparten tablas de dimensiones, requerida cuando conviven procesos de diferente granularidad, tales como ventas diarias, presupuestos mensuales, forecast semanal, etc.

6.5.4 Arquitectura de Medallón: Integración de las Etapas de Refinamiento del Dato

Las tres estructuras descritas en las secciones anterior, el Data Lake, el Data Warehouse y el Data Mart no operan como alternativas excluyentes, sino como etapas sucesivas dentro de un mismo flujo de refinamiento progresivo del dato. Una formalización específica de este flujo es la Arquitectura de Medallón, definida por Databricks como un patrón de diseño de datos utilizado para organizar lógicamente la información dentro de un lakehouse, cuyo objetivo es mejorar de manera incremental y progresiva la estructura y calidad del dato a medida que este avanza a través de las capas Bronce, Plata y Oro (Databricks, 2024).

Bajo este patrón, la capa Bronce corresponde a la zona de aterrizaje del Data Lake descrita en la Sección 6.5.1, que conserva el dato en su formato nativo sin ningún tipo de transformación, la capa Plata corresponde a la etapa de consolidación e integridad referencial propia del Data Warehouse descrito en la Sección 5.5.2, donde el dato es limpiado, tipado y relacionado bajo un modelo normalizado. Finalmente, la capa Oro corresponde al Data Mart descrito en la Sección 6.5.3, donde el dato es modelado dimensionalmente y queda listo para su consumo por parte del usuario de negocio (Databricks, 2024).

Esta arquitectura de tres capas es, a su vez, conceptualmente equivalente a la distinción de dos grandes áreas descritas por Kimball y Ross (2002), el Área de Staging, definido como todo el espacio y los procesos ubicados entre los sistemas operacionales de origen y el área de consumo del usuario final, sin acceso directo por parte de este último (p. 1112), y el Área de Presentación, donde reside el modelo dimensional listo para el análisis. Bajo esta equivalencia, las capas Bronce y Plata de la Arquitectura de Medallón corresponden, en conjunto, al Área de Staging, mientras que la capa Oro corresponde al Área de Presentación.

6.5.5 Motores de Bases de Datos Analíticas Embebidas

Como alternativa a los sistemas gestores de bases de datos tradicionales, que operan como servidores independientes y requieren una infraestructura dedicada para su despliegue, Raasveldt y Mühleisen (2019) han emergido una categoría de motores de datos embebidos (embedded), diseñados para ejecutarse directamente dentro del proceso de la aplicación que los invoca, sin necesidad de un servidor ni de un protocolo de comunicación cliente-servidor (p. 1982).

Según sostiene (Raasveldt & Mühleisen, 2019), si bien este paradigma ya era ampliamente utilizado para cargas de trabajo transaccionales (OLTP) a través de motores como SQLite, no existía en la literatura tecnológica un motor equivalente de uso extendido orientado específicamente a cargas de trabajo analíticas (OLAP), lo que motivó el desarrollo de DuckDB, un sistema gestor de bases de datos relacional embebido de propósito específico (pp. 1981-1982).

Entre los requisitos de diseño que los autores Raasveldt y Mühleisen (2019) identifican para este tipo de motores se encuentra la portabilidad práctica, es decir, el sistema debe poder ejecutarse en cualquier entorno que corra la aplicación anfitriona, minimizando dependencias externas y evitando la necesidad de administrar un proceso de servidor independiente (p. 1982). Esta característica resulta particularmente relevante para escenarios de análisis interactivo de datos, como los realizados mediante lenguajes como Python, en los que las operaciones de manipulación relacional se ejecutan directamente sobre las estructuras de datos del entorno de trabajo.

En cuanto a su arquitectura de almacenamiento , DuckDB utiliza un formato de bloque de datos (DataBlocks) optimizado para lectura, en el cual las tablas lógicas se particionan horizontalmente en fragmentos de columnas que se comprimen mediante métodos de compresión ligera, incorporando además índices de valores mínimos y máximos por columna que permiten

descartar rápidamente los bloques no relevantes para su consulta. Este enfoque de almacenamiento columnar comprimido es el que permite a DuckDB ejecutar consultas analíticas de agregación sobre grandes volúmenes de datos con un rendimiento sustancialmente superior al de los motores tradicionales orientados a filas.

6.6 Metodología Hefesto para el Diseño de Almacenes de Datos

El diseño e implementación de un sistema tecnológico para la relación logística B2B entre un proveedor de consumo masivo y un retailer de alta exigencia no puede responder a un enfoque empírico o de ensayo y error. Requiere el despliegue de metodologías de ingeniería estructuradas que aseguren la robustez del proceso, la confiabilidad de los sistemas de información y la adaptabilidad a los cambios.

A continuación, se analizan y estructuran metodologías idóneas para desarrollar un proyecto como este:

6.6.1 Lean Six Sigma (Metodología DMAIC)

Como sostiene George (2002), Lean Six Sigma es una metodología estratégica diseñada para maximizar el valor del negocio mediante la eliminación sistemática de desperdicios (Muda) y la reducción en la variabilidad en los procesos (pp. 12, 54).

El despliegue operativo de la metodología Lean Six Sigma se instrumenta a través del ciclo DMAIC, un proceso estructurado de resolución de problemas, el cual garantiza que la estrategia organizacional se traduzca en proyectos de ejecución con un impacto medible.

En un contexto de integración y gobernanza de datos B2B, este ciclo se desglosa bajo el siguiente rigor metodológico.

- Definir: En esta fase inicial se clarifican las metas y el valor estratégico del proyecto. El objetivo principal es alinearse con la Voz del Cliente (VOC). Según George (2002),

en esta etapa se deben identificar los requisitos Críticos para la Calidad (CTQ) que pueden ser monitoreados mediante métricas contractuales (p. 3). El uso del diagrama SIPOC (Suppliers, Inputs, Process, Outputs, Customers) es fundamental en esta etapa para delimitar el alcance de flujo de información logística e identificar a todos los actores clave en el proceso estudiado.

- Medir: Con respecto a la etapa de Medición, George (2002) sostiene que esta fase es fundamental para establecer la línea base del desempeño organizacional mediante la recolección de datos objetivos que permitan cuantificar la capacidad actual del proceso (p. 6). Dicha información se extrae de sistemas de información críticos (ERP y portales B2B), lo que permite validar la magnitud de las ineficiencias al identificar la brecha existente entre el desempeño real observado y el nivel de servicio deseado por los socios comerciales.

- Analizar: El autor George (2002) sostiene que el objetivo central es examinar los datos recolectados para localizar las denominadas Trampas de Tiempo (Time Traps), definidas como aquellas actividades específicas que inyectan el 80% o más de la demora total en un flujo de valor (p. 127). Para su identificación y priorización, el autor sugiere el uso de herramientas de control de calidad como el Diagrama de Pareto o el Diagrama de Ishikawa, instrumentos que permiten aislar las causas raíz de los sobrecostos y las deficiencias en el nivel de servicio (p. 6).

- Mejorar: Respecto a esta fase, George (2002) sostiene que esta etapa se dedica exclusivamente al diseño y despliegue de soluciones que mitiguen las causas raíz identificadas en el diagnósticos previos (p. 6). La optimización de los flujos de datos busca eliminar sistemáticamente el desperdicio de información y las actividades que no agregan valor, lo cual permite reducir el Lead Time de respuesta organizacional (p. 12).

- Controlar: Finalmente, en lo que respecta a la etapa de controlar, George (2002) establece que el objetivo primordial es implementar mecanismos de monitoreo sistemático que

garanticen la permanencia de las mejoras en el tiempo (p .312). Para asegurar esta sostenibilidad, el autor George (2002) propone la integración de técnicas de Mistake Proofing (Poka-Yoke) y el uso de tableros de control visual, permitiendo que el flujo de información mantenga su integridad y que los estándares de colaboración alcancen niveles de eficiencia constantes, mitigando la variabilidad natural de proceso (p. 331).

El Diagrama de Ishikawa, también denominado Diagrama Causa-Efecto o Diagrama de Espina de Pescado, constituye una herramienta visual de análisis cualitativo que permite estructurar de forma sistémica las causas potenciales de una desviación. Tal como sostiene George (2002), el efecto o problema central se sitúa en la cabeza del diagrama, mientras que las causas se distribuyen en ramas o categorías jerárquicas que convergen hacia él, facilitando una priorización objetiva mediante instrumentos complementarios como el Diagrama de Pareto (p. 168).

Como técnica de profundización, George (2002) plantea el uso de los 5 Porqué (5 Whys), método que consiste en cuestionar iterativamente el origen de una falla hasta alcanzar su causa raíz estructural, evitando que el análisis se limite a la manifestación superficial del síntoma (p. 111).

Dentro de la ingeniería de procesos, estas causas se agrupan tradicionalmente en categorías operativas: Método (procedimientos), Mano de Obra (capacitación), Máquina (herramientas tecnológicas o analógicas), Medición (indicadores) y Medio Ambiente (entorno organizacional).

La integración de ambas herramientas permite pasar de un diagnóstico descriptivo a una resolución de problemas con base científica.

- Método: Relacionado con la forma en que se ejecuta el proceso, incluyendo la existencia o ausencia de procedimientos estandarizados.
- Mano de Obra: Asociado a la disponibilidad, capacitación y dependencia de factor

humano en la ejecución de tareas.

- Máquina: Vinculado a las limitaciones técnicas del Software o sistema utilizado para ejecutar el proceso.
- Medición: Relacionado con la ausencia de controles o indicadores que permitan detectar desviaciones de forma oportuna.
- Medio Ambiente: Asociado a factores externos u organizacionales que condicionan la ejecución del proceso.

6.6.2 Modelo en Cascada (Waterfall)

El Modelo en Cascada se fundamenta en la propuesta de Royce (1970), quien describe un enfoque de desarrollo lineal y secuencial para sistemas de gran complejidad. Bajo este paradigma, el proceso se divide en fases interdependientes donde la progresión hacia una nueva etapa supone la finalización de la anterior (p. 330).

Este modelo resulta idóneo para proyectos de infraestructura logística donde los requerimientos son estables y conocidos.

Su despliegue se estructura de la siguiente forma:

- Levantamiento de Requisitos: Fase en la que se documentan exhaustivamente las necesidades del sistema. En el marco de una integración B2B, esto implica definir los datos de inventario y ventas (Sell-Out) que deben ser extraídos para alimentar los modelos de reabastecimiento.
- Diseño del Sistema: Etapa centrada en la arquitectura técnica y el modelado de datos. Royce (1970) sostiene que un diseño preliminar robusto es esencial para prevenir fallos críticos en el flujo de datos (pp. 331-332).

- Implementación (Codificación/Configuración): Los ingenieros y desarrolladores construyen la solución técnica basándose exclusivamente en los planos del diseño. Si se desarrolla un software de automatización de reportes o un Bot Robotic Process Automate (RPA) para descargar la información desde el portal externo B2B de manera automática, el código se escribe en esta fase.

- Verificación: Representa la fase de mayor riesgo en términos de cronograma y presupuesto, ya que es el primer evento en donde se validan las interfaces y la capacidad real de almacenamiento del sistema.

- Mantenimiento: Una vez que el sistema entra en producción, esta fase contempla el mantenimiento y la actualización constante. Según Royce (1970), la documentación generada en etapas previas es vital en este punto para permitir adaptaciones ante cambios en las políticas de la organización (p. 332).

6.6.3 Modelo en V (Verificación y Validación)

El Modelo en V representa una metodología en la ingeniería de sistemas que evoluciona de un modelo secuencial de cascada para introducir una estructura tridimensional del ciclo del proyecto. Introducido por Forsberg y Mooz (1991), este modelo organiza las fases de desarrollo de manera simétrica, estableciendo una relación bidireccional entre la rama descendente de definición técnica y la rama ascendente de integración y verificación. (pp. 1-3).

Un principio fundamental de este paradigma es que las actividades de verificación y validación se planifican de manera concomitante con la especificación de requerimientos en cada nivel. Según Forsberg y Mooz (1991), este enfoque proactivo minimiza la propagación de errores de diseño y optimiza la rentabilidad del proyecto al evitar correcciones reactivas en etapas avanzadas de la implementación (p .4).

La rama de la izquierda, denominada Descomposición y Definición, tiene como función determinar las necesidades estratégicas para que se transformen progresivamente en especificaciones técnicas:

- Análisis de Requisitos del Negocio: Se identifican y documentan las necesidades de los usuarios finales y los objetivos estratégicos de la organización. La principal función es definir qué debe resolver el sistema, estableciendo los criterios de éxito y restricciones del entorno corporativo.

- Especificación Funcional del Sistema: Las necesidades del negocio se traducen en funciones específicas que el nuevo sistema o proceso debe realizar. Se deben determinan las principales interfaces del usuario, flujos de datos principales y principales reglas de negocios.

- Diseño de la Arquitectura: Se define la estructura modular de la solución. Se establecen las plataformas, mecanismos de comunicación entre ellas y protocolos de transferencia.

- Diseño de Detalle: Es la fase de mayor granularidad, donde se especifican los algoritmos, las estructuras de datos exactas, la lógica interna de cada módulo, permitiendo que la fase de construcción o codificación se ejecute sin ambigüedades.

La rama derecha, Fase de Ejecución y Construcción implica la codificación, parametrización o construcción física. Aquí, los planos y diseños conceptuales de la rama izquierda se transforman en elementos tangibles. Es el puente de transición entre el diseño abstracto y la realidad operativa.

La Fase que contempla Integración, Verificación y Validación es la rama de la derecha que representa la agregación progresiva de los componentes construidos para verificar su correcto ensamblaje y validar si efectivamente resuelven las necesidades del negocio identificadas en la rama izquierda.

- Pruebas Unitarias: Corresponde a la validación de la fase Diseño de Detalle, se examinan los módulos o componentes individuales de forma aislada para garantizar que funcionen exactamente como se especificó.

- Pruebas de Integración: Se vincula con la fase de Diseño de la Arquitectura, su objetivo es comprobar que los diferentes módulos coordinados interactúen correctamente entre sí, sin pérdidas ni fallas en la integración de información.

- Pruebas del Sistema: Corresponde a la verificación de la fase Especificación Funcional, se evalúa la solución completa e integrada, verificando que funcione correctamente como un todo, bajo los parámetros y requisitos de tiempo, rendimiento y organizacionales.

- Pruebas de Aceptación del Usuario: Se vincula con la fase de Análisis de Requisitos del Negocio, son ejecutadas por los usuarios finales y las partes interesadas, ya que son ellos los que determinan que la solución satisface las necesidades del negocio.

6.6.4 Metodologías Ágiles (Framework Scrum)

En el ámbito del desarrollo de sistemas complejos, Schwaber y Beedle (2001) definen Scrum como un marco de trabajo adaptativo, iterativo e incremental. A diferencia de los modelos lineales tradicionales, Scrum se orienta a la entrega continua de valor funcional mediante ciclos temporales denominados Sprints, cuya duración suele oscilar entre 1 y 4 semanas (pp. 15-17).

Este enfoque es particularmente efectivo en entornos de BI y gobernanza de datos, donde la validación constante de prototipos por parte del usuario final es crítica para el éxito del proyecto.

Su arquitectura operativa se fundamenta en un proceso cíclico compuesto por los siguientes pilares:

- Construcción del Backlog: Representa la fuente única de requisitos, es una lista priorizada y dinámica de todas las funcionalidades, mejoras y correcciones necesarias para el

producto.

- Sprint Planning: Ceremonia de planificación inicial donde el equipo selecciona los elementos del Backlog de mayor prioridad para transformarlos en un incremento de producto terminado durante el ciclo actual.

- Ejecución del Sprint y Daily Scrum: Fase de desarrollo donde el equipo se sincroniza diariamente mediante reuniones breves para inspeccionar el progreso hacia el objetivo del sprint e identificar impedimentos existentes que requieran solución inmediata.

- Sprint Review: Evento de inspección al cierre del ciclo donde se realiza una demostración técnica del incremento funcional. El objetivo es obtener retroalimentación directa de los interesados sobre el producto operando en tiempo real.

- Sprint Retrospective: Instancia de introspección del equipo destinada a la mejora continua del proceso, analizando las dinámicas de trabajo, las herramientas y la comunicación para optimizar el rendimiento en el ciclo subsiguiente.

6.7 Metodologías y Herramientas para la Evaluación y Modelamiento de Procesos

Desde una perspectiva académica, Heizer y Render (2009) sostienen que la optimización y el rediseño de los procesos de negocio demandan una reevaluación fundamental de sus objetivos y suposiciones subyacentes (p. 266). En la ingeniería de procesos este diagnóstico no se limita a la observación casual, sino que exige la implementación de marcos metodológicos estandarizados, tales como, diagramas de flujo, gráficas de función tiempo y mapas de valor, diseñados para visibilizar de forma objetiva las ineficiencias y los cuellos de botella que restringen la productividad (p. 277). Estos instrumentos analíticos permiten descomponer flujos completos de materiales e información, facilitando la identificación de actividades que no agregan valor y garantizando que el sistema final se alinee con los requisitos estratégicos del entorno corporativo

(p. 267).

A continuación, se describen y analizan algunas herramientas para el levantamiento, modelamiento y diagnóstico de procesos operativos.

6.7.1 Modelamiento de Procesos de Negocio (BPMN)

En el ámbito de la gestión de procesos, el Business Process Model and Notation (BPMN) se ha consolidado como el estándar internacional para el modelado conceptual y técnico de flujos de trabajo. Tal como sostiene Hitpass (2012), su propósito fundamental trasciende la mera representación gráfica, buscando ofrecer una notación intuitiva pero dotada del rigor técnico necesario para ser comprendida por analistas de procesos, desabolladores de sistemas y directivos estratégicos (pp. 38, 559).

Desde la perspectiva del análisis operativo, Heizer y Render (2009), argumentan que el uso de estos esquemas constituye una herramienta de diagnóstico esencial que facilita la identificación y supresión sistemática de desperdicios, tales como actividades redundantes, duplicidades y demoras que inyectan ineficiencia al sistema (p. 308). De este modo, la organización puede analizar procesos de transformación con el fin último de maximizar el valor percibido por el cliente.

A nivel académico y estructural, la arquitectura de BPMN organiza el flujo a través de 4 cuatro categorías jerárquicas de objetos.

- Objetos de Flujo: Constituyen el núcleo dinámico del modelo. Incluyen los eventos, que representan instancias de inicio, hitos intermedios o estados finales; las actividades, definidas como tareas específicas que consumen recursos corporativos, y las compuertas (gateways), las cuales gobiernan la lógica de bifurcación y convergencia del flujo según reglas de negocio predefinidas.

- Objetos de Conexión: Representan la secuencia cronológica de las tareas y el flujo de mensajes entre los diversos actores del sistema.

- Canales de Organización (Pools y Lanes): Estructuras visuales compuestas por Pools, que representan entidades independientes o socios comerciales B2B y Lanes que subdividen las responsabilidades por roles o departamentos internos.

- Objetos de Información: Permiten documentar la entrada y salida de datos durante el proceso, tales como archivos de maestra de materiales o registros transaccionales.

La utilidad de BPMN radica en su capacidad para facilitar el rediseño de procesos, lo que implica volver a conceptualizar la forma fundamental del flujo del negocio para lograr mejoras drásticas en el desempeño. Como señalan Heizer y Render (2009), un modelamiento efectivo permite a los administradores cuestionar el propósito de cada actividad y sus suposiciones implícitas, garantizando que el sistema final se alinee con los requisitos estratégicos de la compañía (p. 314). En consecuencia, esta herramienta permite localizar Trampas de Tiempo y fricciones en la transferencia de información, sentando las bases para la posterior mejora del proceso.

6.7.2 Mapeo de la Cadena de Valor (Value Stream Mapping - VSM)

El Mapeo de la Cadena de Valor (VSM) es una herramienta estratégica dentro de la metodología Lean, diseñada para visualizar y analizar de forma integral el flujo de materiales e información necesarios para entregar un producto o servicio al cliente final.

A diferencia de los diagramas de flujo convencionales que se limitan a describir la secuencia de actividades, el VSM realiza una observación ampliada que permite distinguir entre los puntos donde se genera valor y aquellos que solo inyectan latencia o desperdicio en el sistema.

Desde la perspectiva de la ingeniería de procesos, el análisis mediante VSM exige clasificar cada tarea bajo el criterio del valor percibido por el cliente. Según George (2002) las actividades

se categorizan en tres dimensiones fundamentales (pp. 229-231).

- Valor Agregado por el Cliente (CVA): Tareas que transforman la forma, característica o función del producto/servicio por las cuales el cliente está dispuesto a pagar. En el contexto de esta memoria, se traduce en el cálculo y análisis estratégico de las variaciones de stock para la toma de decisiones.

- Valor Agregado por el Negocio (BVA): Actividades que, si bien no agregan valor directo al cliente, son requeridas por normativas, reglamentaciones o controles de riesgo financiero de la organización. Son necesarias bajo el diseño actual, pero deben ser minimizadas.

- Sin Valor Agregado: Representa el desperdicio puro. Incluye actividades como el conteo manual, manipulación excesiva de datos, inspecciones redundantes, esperas (latencia) y retrabajos por corrupción de archivos.

Este diagnóstico busca identificar los recursos consumidos en actividades que no agregan valor, provocados usualmente por la variabilidad y la excesiva duración de ciclos en los procesos.

La utilidad técnica del VSM radica en su capacidad para cuantificar la eficiencia operativa mediante la relación entre el tiempo productivo y el tiempo total de permanencia de la información en el sistema.

El indicador principal utilizado para este fin es la Eficiencia del Ciclo del Proceso (PCE), el cual se calcula mediante la siguiente expresión matemática:

$$PCE = \frac{Tiempo\ con\ Valor\ Agregado}{Tiempo\ Total\ del\ Ciclo}\ \%$$

Otro indicador es la Fuga de Valor de Ingeniería que representa el costo económico directo que desperdicia en tareas manuales de carpintería de datos.

$$FVI = Horas\ Semanales\ de\ Carga\ Manual\ x\ Costo\ HH\ del\ Analista$$

En la relación entre proveedor y retailer, el Mapeo de la Cadena de Valor (VSM) permite identificar las denominadas Trampas de Tiempo (Time Traps) definidas según George (2002), como aquellas estaciones de trabajo o procesos que inyectan el 80% o más del retraso total en el tiempo de entrega de la información estratégica (p. 127). Como sostiene Christopher (2016), los lapsos de tiempo en los flujos de información se traducen directamente en inventario inmovilizado o quiebres en la góndola (p. 195).

Bajo estas premisas, Christopher (2016) justifica que, la visibilidad de los datos actúa como un sustituto estratégico de las existencias físicas, lo cual justifica la automatización de los procesos de captura para comprimir el lead time de entrega de información y mejorar la capacidad de respuesta de la cadena de suministro (p. 178).

6.7.3 Capacidad Tecnológica del Repositorio de Datos

La literatura de gestión de datos identificada la capacidad técnica del repositorio de información como un factor crítico para la integridad analítica. Según Cano (2018), cuando una organización emplea software de propósito general, específicamente planillas de cálculo como sustituto de un repositorio formal, la capacidad de retención de activos queda supeditada a las restricciones físicas de la herramienta y no a las necesidades estratégicas del negocio (p. 92). Esta limitación compromete la ejecución de análisis longitudinales o de tendencias, los cuales requieren la preservación de series históricas que trasciendan el corto plazo.

Bajo esta premisa, la evaluación de la capacidad se determina mediante la relación de Días Históricos Teóricos (DHT):

$$DHT = \frac{L_{max}}{RD}$$

Donde DHT corresponde a los Días Históricos Teóricos que la herramientas puede almacenar, L_{max} representa el límite físico de registros que esta puede procesar sin degradar su

rendimiento y RD corresponde a la tasa de generación de registros diarios del proceso evaluado.

Tal como advierte Davenport (2006), el uso de hojas de cálculo para gestionar volúmenes de datos masivos cercanos a su límite técnico eleva la probabilidad de error en el archivo a un rango entre el 20% y el 40% (p. 101).

6.7.4 Modelo de Madurez de Inteligencia de Negocios (BIMM) - Modelo de Wayne Eckerson

El progreso y la evolución analítica de una organización se miden a través del Business Intelligence Maturity Model (BIMM), un marco de referencia metodológico diseñado para evaluar la capacidad de los sistemas de información al transformar datos crudos en conocimiento accionable y planes estratégicos sostenibles. De acuerdo con Lituus Analytics (2023), estos modelos operan como un dispositivo de navegación estratégica que permite identificar la situación actual de la organización y trazar el camino hacia la consolidación de una cultura orientada a decisiones basadas en datos.

Teóricamente el diseño de estos modelos se inspira en enfoques de evaluación y mejora de procesos consolidados, tales como la evaluación de madurez de capacidades (CMA), el modelo de madurez de capacidades de software (CMM) o en marco de gobernanza y control tecnológico como COBIT.

En la literatura de BI, un referente fundamental para estructurar y el diagnosticar esta evolución es el modelo propuesto por el experto Wayne Eckerson. Este marco conceptual describe con rigurosidad técnica la progresión de la madurez analítica mediante cinco etapas secuenciales e identifica dos puntos de fricción críticos, denominados Gulf y Chasm, que representan las barreras operacionales, tecnológicas y culturales que las organizaciones deben superar para transitar hacia estados superiores de integración semántica. Las etapas se definen en función de características como la percepción de los ejecutivos, el alcance analítico, la estructuración

analítica, el stewardship, la plataforma tecnológica y la gestión del cambio. Las organizaciones avanzan a ritmos distintos y pueden exhibir características de más de una etapa a la vez, por lo que el tránsito entre etapas no es completamente nítido.

Cabe precisar que la formulación original del modelo, desarrollada por Eckerson en The Data Warehousing Institute (TDWI), distingue seis etapas: Prenatal, Infant, Child, Teenager, Adult y Sage. En esta memoria se adopta la versión de cinco etapas, en la que Prenatal e Infant se presentan como una única etapa. Esta agrupación no altera el diagnóstico respecto del primer punto de fricción, ya que en ambas formulaciones el Gulf se ubica inmediatamente después de la etapa Infant.

En la medida que se progresa por esta camino evolutivo, se incrementa la credibilidad, consistencia y unificación de los datos, logrando mitigar la latencia y la fragmentación operativa

- Etapa 1: Prenatal/Infant (Prenatal / Infancia): Se caracteriza por una severa lentitud en la generación de reportes y la carencia absoluta de un entorno interactivo o ágil para la visualización de datos, lo que fomenta el desarrollo de soluciones informales basadas en procesamiento manual, en donde los analistas extraen consultas directas de los sistemas de origen para pegarlas en planillas locales de cálculo, dando origen a la proliferación de silos de información aislados. En esta etapa el analista opera bajo el rol de un Human Data Warehouse (Data Warehouse Humano), consumiendo la mayor parte de su tiempo de ingeniería en la recolección, limpieza y preparación manual de datos, dejando un espacio marginal para el análisis estratégico. Desde la perspectiva del riesgo de integridad, el uso de hojas de cálculo individuales como repositorio presenta una vulnerabilidad, ya que tal como advierte Davenport (2006) , el uso de hojas de cálculo para gestionar volúmenes de datos masivos cercanos a su límite técnico eleva la probabilidad de error en el archivo a un rango entre el 20% y el 40% (p. 101).

- Punto de Fricción Gulf (El Golfo): Representa la primera barrera estructural de la organización. Superar el Gulf para transitar a la siguiente etapa exige conseguir un patrocinador directivo, resolver la escasez de licencias analíticas, capacitar formalmente al personal y solucionar de raíz las asimetrías y problemas de calidad de datos en los sistemas operacionales de origen.

- Etapa 2: Child (Infancia): Los líderes departamentales identifican la urgencia de empoderar a los colaboradores con datos confiables para la toma de decisiones. Se impulsa el diseño de repositorio de datos departamentales independientes (Data Mart). No obstante, persistente la dificultad de unificar reglas de negocio comunes a nivel de toda la organización y se mantiene la ineficiencia por la extracción repetitiva de los mismo datos desde los sistemas de origen.

- Etapa 3: Teenager (Adolescencia): El éxito de las primeras iniciativas analíticas impulsa una demanda masiva de accesos dentro de la compañía, buscando transitar desde un grupo acotado de usuarios avanzados hacia perfiles de visualización casuales. Se introducen herramientas interactivas, pero la falta de un gobierno corporativo claro genera desorden en las definiciones de los indicadores.

- Punto de Fricción Chasm (El Abismo): Es el obstáculo crítico donde se estanca la mayoría de las organizaciones. La barrera deja de ser estrictamente tecnológica y pasa a ser un desafío de cultura organizacional y gobernanza corporativa, puesto que los usuarios se resisten a abandonar sus silos mentales y duplican la generación de informes redundantes y desactualizados.

- Etapa 4: Adult (Adultez): Se alcanzan definiciones y reglas de negocio estándar de toda la compañía, consolidando los repositorios de información departamentales en un único almacén integrado gobernado por el negocio. Se transita formalmente desde los informes analíticos dinámicos hacia sistemas robustos de gestión del rendimiento

- Etapa 5: Sage (Sabiduría): Las herramientas analíticas se embeben por completo en los procesos de negocio en tiempo real, convirtiéndose en un recurso estratégico que permite monetizar o disponibilizar los datos hacia clientes y proveedores para coordinar la cadena de valor conjunta de forma colaborativa e industrializada.

Si bien el modelo de Wayne Eckerson constituye uno de los marcos de referencia más extendidos para evaluar la progresión analítica en las organizaciones, la literatura académica y la consultoría especializada han desarrollado modelos adicionales para diagnosticar la madurez del Business Intelligence (BI) desde diversas perspectivas dimensionales.

6.7.5 El Modelo de Madurez de Gartner (Gartner Maturity Model for BI and Performance Management)

Uno de los marcos más difundidos es el Gartner's Maturity Model for Business Intelligence and Performance Management. Al analizar las metodologías de evaluación de BI, Olszak (2013) expone que el modelo de Gartner establece la hoja de ruta para determinar el grado de aprovechamiento de los activos de información, evaluando simultáneamente tres dimensiones clave de la organización, personas, procesos, métricas y tecnología (pp. 953-954).

- Nivel 1 - Inconsciente (*Unaware*): Definido técnicamente por Olsza (2013) como un estado de anarquía de información, donde los datos son incompletos, inconsistentes y de baja calidad (p. 953). La organización carece de métricas estandarizadas y el uso de herramientas de reportabilidad es sumamente restringido.

- Nivel 2 - Táctico (Tactical): Según Olszak (2013) corresponde a la etapa en que se inician las primeras inversiones en infraestructura de BI (p. 953). Las métricas se definen de manera aislada a nivel departamental, los datos y aplicaciones permanecen confinados en silos y los usuarios carecen de las competencias necesarias para explotar el sistema de forma autónoma.

- Nivel 3 - Enfocado (Focused): Es el punto donde Olszak (2013) define que se deben alcanzar los primeros éxitos operacionales y beneficios de negocio, aunque acotados a áreas específicas de la empresa (p. 953). En este nivel, la gerencia comienza a demandar el diseño de tableros de control ejecutivos (dashboards).

- Nivel 4 – Estratégico (Strategic): Definido por (Olszak, 2013), como el punto en donde la compañía cuenta con una estrategia formal y clara para el desarrollo de BI, respaldada por la alta dirección (pp. 953-954). El uso de la analítica se extiende hacia clientes y proveedores, dando soporte directo a la toma de decisiones tácticas y estratégicas mediante modelos predictivos y minería de datos.

- Nivel 5 – Penetrante (Pervasive): Olszak (2013) dice que el BI asume un rol transversal e inmerso en todas las áreas y en la cultura corporativa (p. 954). La infraestructura analítica proporciona flexibilidad para adaptarse a cambios rápidos del entorno, democratizando el acceso a la información tanto para colaboradores internos como para socios comerciales de la cadena de valor.

6.7.6 El Modelo de Desarrollo de BI (Business Intelligence Development Model - BIDM)

En el ámbito estrictamente académico, Sacu y Spruit (2010) desarrollan el Business Intelligence Development Model (BIDM), un marco de referencia diseñado para estructurar las etapas evolutivas del BI e identificar los requisitos técnicos y operacionales necesarios para transitar hacia un estado analítico maduro (pp. 1-3). Al revisar los antecedentes de su propuesta, los propios Sacu y Spruit (2010) reconocen que su modelo toma como punto de partida la estructura de seis etapas desarrollada por The Data Warehousing Institute (TDWI), mismo linaje conceptual que da origen al modelo de Wayne Eckerson (p. 3).

No obstante, Sacu y Spruit (2010) optan por abandonar la metáfora biológica del TDWI

para definir seis etapas metodológicas explícitas (pp. 3-4)

- Reporte Predefinido (*Predefined Reporting*): Sacu y Spruit (2010) mencionan que en este punto la organización debe estar caracterizada por evaluaciones rígidas de datos históricos presentadas en informes estáticos y periódicos (p. 4). La capacidad de análisis es mínima y la falta de un entorno interactivo obliga a los usuarios a copiar y pegar segmentos de información entre múltiples archivos locales para construir sus propios reportes.

- Data Marts / Almacén de Datos Departamental: Involucra el desarrollo de repositorios analíticos enfocados en áreas funcionales específicas, introduciendo consultas ad-hoc y estableciendo una semántica común para el departamento.

- Almacen de Datos Corporativo (Enterprise-wide Data Warehouse): Sacu y Spruit (2010) argumentan que en este punto ya se debe representar la consolidación de una infraestructura centralizada a escala de toda la organización, garantizando un alto grado de integración física y consistencia semántica (p. 4).

- Analítica Predictiva (Predictive Analytics): Sacu y Spruit (2010) mencionan que la organización ya debe incorporar técnicas de minería de datos, modelamiento predictivo y algoritmos de inteligencia artificial para explorar grandes volúmenes de información y descubrir patrones no evidentes (p. 4).

- BI Operacional (Operational BI): Según Sacu y Spruit (2010) para que la organización se encuentre en este punto ya se debe superar la tradicional carga nocturna por lotes (batch) para integrar la analítica en tiempo casi real dentro de los procesos de negocio en ejecución, apoyándose en el procesamiento de flujos de eventos (pp. 9-10).

- Gestión del Desempeño del Negocio (Business Performance Management): Sacu y

Spruit (2010) mencionan que en este punto ya se constituye el nivel máximo de integración, donde BI se combina con un marco de gestión en ciclo cerrado que alinea los objeticos estratégicos y operacionales con las actividades diarias de la compañía (pp. 10-11).

Para evaluar estas seis etapas, Sacu y Spruit (2010) diseñan una matriz de diagnóstico compuesta por veinte características clasificadas en seis categorías funcionales, tales como, perspectivas temporales, perspectivas de datos, perspectivas de decisión, perspectivas de salida/visualización, enfoques del proceso BI y otras características organizacionales (pp. 11-13).

7. Antecedentes

La compañía Tresmontes Lucchetti (TMLUC) es una de las entidades industriales de mayor trayectoria y prestigio en el sector de bienes de consumo masivo en Chile, respaldada por una herencia manufacturera que supera los cien años de evolución constante. Desde el año 2013, la compañía alcanzó un hito estratégico fundamental al integrarse al Grupo Nutresa (One Nutresa), el conglomerado de alimentos más importante de Colombia y uno de los actores más relevantes en el mercado Latinoamericano.

TMLUC se define por su capacidad para gestionar un portafolio de alta complejidad, compitiendo simultáneamente en diversas categorías estratégicas que incluyen:

- Pastas: Fideos y pastas alimenticias de consumo masivo.
- Snacks: Productos salados y colaciones de consumo individual.
- Sémola: Sémola de trigo utilizada como insumo para preparaciones culinarias.
- Saborizantes: Jugos en polvo y bebidas saborizadas instantáneas.
- Néctar: Bebidas de fruta envasadas de consumo directo.
- Deshidratados: Sopas, cremas y condimentos en formato deshidratado.
- Cremas: Cremas instantáneas y bases para la preparación de comida.
- Bif: Bebidas instantáneas calientes.
- Bic: Bebidas instantáneas frías.

Además de incursiones en el mercado de aceites y aceitunas:

- Aceite: Aceites comestibles envasadas para uso doméstico e industrial.
- Aceto: Vinagres ya aderezos derivados de la aceituna

TMLUC maneja alrededor de 300 productos (SKUs) distintos que van al consumo masivo

a través de distintos canales, tales como grandes supermercados, mayoristas y tradicional (tiendas de conveniencia).

Dada la gran amplitud del portafolio de productos se requiere que varias áreas de la organización deben gestionar sus operaciones y decisiones simultáneamente. Por lo que se requiere un flujo de información unificado que permita balancear las decisiones operativas de TMLUC con las dinámicas de comercialización de cada una de estas categorías a nivel nacional.

Debido a esta diversificación de productos, se puede situar a la compañía en lo que Heizer y Render (2009) definen como una organización de manufactura de alta complejidad, donde la administración de operaciones debe coordinar no solo la producción técnica, sino también la interacción constante entre departamentos para transformar insumos en bienestar social (p. 2).

7.1 Estructura Organizacional y Gobernanza Corporativa.

Para dar soporte a la operación de aproximadamente 300 productos (SKUs) que fluyen a través de múltiples canales, TMLUC se organiza bajo una estructura funcional y matricial diseñada para maximizar la eficiencia operativa y la captura de valor en el mercado. La gobernanza de la compañía se articula a través de una Gerencia General que supervisa diversas direcciones estratégicas y Gerencias de Áreas, las cuales actúan como los pilares del flujo de información corporativo.

- Gerencia de Operaciones y Logística: Responsable de la planificación de la producción, el abastecimiento de insumos y la distribución física del producto hacia los distintos canales de venta.

- Gerencia Comercial (Ventas): A cargo de la relación comercial con los canales de distribución, incluyendo la negociación y administración de cuentas clave como Walmart Chile.

- Gerencia de Marketing: Responsable del posicionamiento de marca, desarrollo de

nuevos productos y la definición de estrategias de comercialización por categoría.

- Gerencia de Administración y Finanzas: Encargada del control presupuestario, la gestión financiera y el soporte administrativo transversal a la operación.

Esta estructura departamental, si bien permite una especialización técnica necesaria, genera desafíos de coordinación. Según los principios de Heizer y Render (2009), los diagramas organizacionales de empresas de manufactura complejas demuestran que las actividades logísticas están intrínsecamente vinculadas con las áreas finanzas y marketing (p. 5).

En TMLUC, cada categoría de producto exige un flujo de información unificado para balancear las decisiones de producción con las dinámicas de comercialización a nivel nacional.

7.2 Modelo de Negocio y Diversificación de Canales

El modelo de negocio de TMLUC se sustenta bajo una capilaridad de su distribución, la cual se despliega a través de tres grandes canales de venta que imponen exigencias logísticas diferenciadas:

- Canal Tradicional: Compuesto por almacenes de barrio y tiendas de conveniencia de atención directa, caracterizado por un alto número de puntos de entrega pero con volúmenes de compra menores.

- Canal Mayorista: Enfocado en grandes distribuidores que actúan como intermediarios para el abastecimiento de zonas rurales o comercios de menor escala.

- Canal Moderno: Representado por las grandes cadenas de supermercados y el retail organizado.

Es precisamente en este último punto donde la organización experimenta su mayor fricción operativa. Mientras que en los otros canales TMLUC ejerce un control más directo sobre la toma de pedidos, en el Canal Moderno, la interdependencia digital con los socios comerciales es

absoluta.

Esta transición de una visión general de la compañía hacia el análisis específico de la relación con el socio comercial más grande, Walmart, es necesaria para comprender porque la ineficiencia en la captura de datos manuales pone en riesgo la estabilidad del contrato comercial más importante de la compañía.

7.3 Canal Moderno y la Relevancia Estratégica de Walmart

El concepto de Canal Moderno en Chile hace referencia a las grandes cadenas de supermercados que han centralizado el consumo masivo a través de economías de escala y una logística de alta eficiencia. Dentro de este ecosistema, Walmart representa el socio comercial más crítico para TMLUC, operando bajo formatos líderes como Hiper Lider, Express Lider, Ekono, SuperBodega Acuenta y Central Mayorista. Walmart no solo es el cliente número uno en facturación, sino que es el termómetro de la salud logística de la compañía.

Dentro del canal moderno, cada cuenta (Walmart, SMU, Cencosud y Tottus) posee su propio analista CPFR, el cual es cargo responsable de la gestión y análisis de toda la cadena de suministro entre cada cuenta (cliente) y TMLUC. Para cumplir sus funciones, el mismo es el responsable de la extracción, consolidación y distribución de la información de su respectiva cuenta a través del portal B2B correspondiente.

Actualmente los productos de TMLUC tienen presencia en 350 salas operativas de Walmart a nivel nacional, con un promedio de 100 SKUs por sala. Esta relación genera una explosión de datos transaccionales, debido a que diariamente se deben monitorear aproximadamente 35.000 combinación producto-sala solo para Walmart.

Desde una perspectiva académica, esta relación B2B está sujeta a lo que Heizer y Render, (2009) definen como una dependencia crítica del desempeño (p. 434). Cualquier desalineación en

el abastecimiento no solo impacta en la venta perdida, sino que gatilla penalizaciones financieras severas por incumplimiento de niveles de servicio (DOH, Fill Rate, Instock).

En este contexto la información logística no es un recurso de apoyo, sino un activo estratégico. La capacidad de Walmart para influir en la última línea del balance de TMLUC obliga a que la gestión de datos deba pasar de un modelo de reporte de lo que pasó a uno de gestión de lo que está pasando en la góndola.

7.4 El Ecosistema Tecnológico Actual

La gestión operativa y comercial de la cuenta Walmart dentro de TMLUC se sustenta sobre una arquitectura de sistemas de información heterogéneos. En la práctica de la ingeniería de procesos, el principal desafío no se centra en la capacidad individual de sus componentes, si no en la conexión y sinergia entre ellos.

Actualmente, el ecosistema tecnológico que soporta la relación B2B entre ambas compañías está compuesto por 3 plataformas principales, cada uno con sus propios objetivos y funciones.

7.4.1 Sistema de Planificación de Recursos Empresariales (ERP)

A nivel interno, TMLUC centraliza su operación financiera y productiva en SAP. Este sistema es el repositorio maestro donde reside la información de inventario propio, costos logísticos, maestros de materiales y las ventas que se realizan a Walmart (Sell-In). Sin embargo, SAP tiene una limitación crítica en este modelo, carece de visibilidad de lo que sucede puertas afuera de la organización, es decir una vez que el producto es entregado al centro de distribución del retailer.

7.4.2 Plataforma Externa de Inteligencia Operacional (Teamcore)

Corresponde a un servicio SaaS (Software as a Service) contratado por TMLUC, que

automatiza la captura de las transacciones de punto de venta (Sell-Out) y el Stock MultiEstado provisto por Walmart. Si bien Teamcore optimiza el acceso a la demanda real en góndola a nivel diario, presenta una limitación técnica estructural, ya que no posee integración ni capacidad de extracción sobre el módulo de planificación y pronósticos algorítmicos (Forecast) de Walmart.

A diferencia del portal Retail Link, donde el Stock MultiEstado solo está disponible durante la jornada en curso, la API de Teamcore permite acceder al stock de los últimos 15 días, lo que posibilita recuperar información no capturada oportunamente.

7.4.3 Portal Retail Link (WMC): Decision Support System

Es la plataforma oficial de retailer donde reside la totalidad de los datos transaccionales, incluyendo los pronósticos de demanda semanales generados por sus algoritmos internos. Debido a que Teamcore no captura esta dimensión de planificación, el Analista CPFR se ve forzado a acceder manualmente al portal Retail Link para extraer el reporte de *Forecast*, generando una fisura en la automatización del flujo de datos.

Esta plataforma entrega variables fundamentales, tales como:

- Ventas POS (Sell-Out): Las unidades de producto de TMLUC que efectivamente fueron comercializadas en las cajas registradoras de cada sala de venta a nivel nacional.
- On Hand: Stock físico en unidades disponible en cada sala Walmart, ya sea en góndola o bodega.
- In Whse: Stock en unidades apartado en los centros de distribución que tienen destino alguna sala Walmart.
- In Transit: Stock en unidades que salió de algún centro distribución y tiene destino alguna sala específica.
- On Order: Stock en unidades solicitado por las salas.

Retail Link es la Voz del Cliente (VOC) en tiempo real.

7.4.4 Herramienta de consolidación y preparación de reportes

Ante la ausencia de una integración nativa entre SAP y Retail Link, la organización utiliza Microsoft Excel como un eslabón de conexión táctica. En esta herramientas, el Analista CPFR descarga, limpia y unifica manualmente los datos maestros internos con la data transaccional externa.

Esta dependencia de archivos locales introduce una vulnerabilidad sistémica, que carece de la robustez para procesar eficientemente millones de registros y degrada la integridad del dato por el riesgo inherente de error humano. Aquí se aplica el axioma de Curto (2010), que dice que la falta de una infraestructura sólida condena a la organización a niveles operativos de bajo valor, donde el tiempo se consume en la preparación del dato y no en su análisis. (p. 30).

7.5 Proceso General de Gestión y Flujo de Información para Indicadores de Abastecimiento

El flujo de información que soporta la relación TMLUC-Walmart opera bajo un ciclo reactivo diario que depende críticamente de la intervención humana. Este proceso se activa cada mañana cuando el Analista CPFR extrae la data generada por Walmart durante la jornada anterior.

El proceso se puede descomponer en tres fases que presentan oportunidades de mejora:

- Extracción Manual de Datos: El analista debe navegar por múltiples módulos de Retail Link para descargar reportes planos (Excel, CSV, TXT). Este paso consume gran parte del tiempo operativo debido a la latencia en el tiempo de respuesta del portal externo Retail Link de WMC.

- Consolidación y Limpieza Local: Una vez obtenidos los datos, se ejecutan macros o fórmulas en Excel para cruzar la información con los maestros. Esta etapa busca normalizar los datos, pero a menudo se ve interrumpida por incompatibilidad de formato, errores humanos o

archivos corruptos, lo que genera retrabajos constantes.

- Distribución de Reportes: Finalmente, se generan visualizaciones estáticas que se envían por correo. Al ser un proceso manual, para cuando el reporte llega al tomador de decisiones, la información tiene al menos 16 horas de desfase, impidiendo acciones correctivas oportunas sobre los quiebre de stock detectados.

La problemática en TMLUC se cuantifica a través del uso ineficiente de talento de ingeniería. Actualmente, el proceso de captura de datos del abastecimiento de las salas Walmart desde Retail Link y tratamiento manual consume 8 horas semanales al Analista CPFR, lo que representa una barrera crítica para alcanzar una madurez analítica superior. Como señala Christopher (2016) cuando el acceso al dato depende de procesos analógicos, se genera una brecha temporal que impide a la organización actuar con la velocidad que la logística moderna exige (p. 244). Esta carga operativa se define como una Fuga de Valor de Ingeniería (FVI), donde el costo económico directo de las horas destinadas a procesos manuales representa un desperdicio de capital financiero y un costo de oportunidad crítico, al impedir que el profesional se enfoque en la gestión proactivo del abastecimiento.

7.6 Indicadores Claves de Desempeño (KPIs) en la relación B2B TMLUC-Walmart

El éxito operacional y la sostenibilidad del contrato comercial entre TMLUC y Walmart están sujetos al cumplimiento de distintos indicadores claves de desempeño (KPIs). Estos indicadores son impuestos contractualmente por el retailer Walmart para garantizar que sus proveedores mantengan un flujo de suministro óptimo que minimice la pérdida de ventas por góndolas vacías. Según Cano (2018), los KPIs son fundamentales para que los miembros de la organización se centren en conseguir objetivos estratégicos (pp. 65-67).

Para el desarrollo de este proyecto, los indicadores críticos que gobiernan la relación B2B

y que configuran el núcleo de análisis se definen conceptualmente a continuación.

7.6.1 Fill Rate (Nivel de Servicio del Pedido)

El Fill Rate es el indicador que mide la eficacia del cumplimiento de las órdenes de compra emitidas por el retailer (Walmart). En la relación TMLUC-Walmart, esta métrica cuantifica la capacidad del proveedor para entregar la cantidad exacta de producto solicitada en el tiempo y lugar acordada.

Desde una perspectiva operativa, un Fill Rate deficiente indica problemas en la planificación de la producción o en la logística de despacho, lo que gatilla automáticamente penalizaciones financieras por parte de Walmart. Este indicador es vital para evaluar la confiabilidad y la capacidad de respuesta de la cadena de suministro puertas adentro del proveedor (TMLUC).

7.6.2 Instock Operacional (Disponibilidad en Góndola)

A diferencia del Fill Rate, el Instock es un indicador de visibilidad externa que mide la disponibilidad efectiva del producto en la sala de venta del retailer (Walmart) para el consumidor final. En el contexto de este proyecto, se define como un indicador binario que evalúa si el inventario físico en sala (On Hand) es suficiente para cubrir la demanda esperada o pronosticada (Forecast) para una jornada determinada.

Matemáticamente la fórmula se expresa de la siguiente forma:

$$I_{ps} = \begin{cases} si\ OH_{ps} \geq\ FD_{ps} \\ si\ OH_{ps} <\ FD_{ps} \end{cases}$$

Donde un valor de 1 representa la disponibilidad y un 0 representa un quiebre de stock para una combinación específica de Producto (p) y Sala (s). Para TMLUC, este indicador es uno de los más críticos, ya que durante el año 2025 se registró un promedio de 95,84%, lo que implica en promedio una cantidad 1.617 productos quebrados por sala diariamente, derivando una venta

perdida estimada de $ 4.883.357.593.

7.6.3 Días de Inventario (DOH – Days On Hand)

Los Días de Inventario (DOH) representan el indicador de eficiencia de las existencias, cuantificando el tiempo estimado en que el inventario actual puede cubrir la demanda proyectada sin recibir nuevo reabastecimiento. Su cálculo vincula el stock disponible con el forecast diario.

$$DOH_{ps} = \frac{OH_{ps}}{FD_{ps}}$$

Para la gestión logística de TMLUC, un DOH excesivamente alto indica una acumulación ineficiente de stock en las salas de Walmart, lo que eleva el riesgo de mermas por vencimiento de productos y capital de trabajo inmovilizado. Por el contrario un DOH muy bajo es una alerta temprana de un posible quiebre de stock inminente.

8. Metodología

El presente capítulo detallada el marco procedimental y los criterios técnicos seguidos para el análisis y diseño del modelo de integración de datos. Se presenta la hoja de ruta que permite capturar, transformar y persistir la información transaccional del abastecimiento de los productos de TMLUC en las salas Walmart.

8.1 Enfoque del Diseño y Alcance del Modelo

El estudio adopta un enfoque mixto (cualitativo y cuantitativo) con un diseño no experimental y transversal, orientado a un alcance descriptivo-propositivo.

El marco cualitativo se sostiene sobre la base metodológica descrita por Hitpass (2012) para la gestión por procesos de negocio, utilizando el modelado en la notación estándar BPMN como el instrumento idóneo para diagnosticar, mapear y comprender la línea base del flujo de información actual (p. 13). Este análisis cualitativo permite identificar empíricamente las compuertas lógicas, loops de reproceso y asimetrías de información que fracturan la relación operativa entre ambas compañías. Paralelamente, el componente cuantitativo incorpora las directrices formuladas por George (2002) para calcular la eficiencia del ciclo del proceso mediante el Mapeo de la Cadena de Valor, contrastándola directamente con la volumetría real de transacciones provistas por el retailer y los límites físicos de procesamiento del software de oficina analizados por Davenport (2006)

El diseño del estudio adquiere una condición no experimental, dado que las variables operativas de la cadena de suministro se observan y recopilan en su contexto natural, (el portal Retail Link de Walmart) sin manipulación externa. Asimismo, se cataloga como un estudio transversal debido a que el levantamiento de datos y el diagnóstico de madurez corresponden a un periodo de tiempo acotado y específico de la relación comercial. De este modo, la investigación

alcanza un alcance descriptivo al caracterizar de forma exacta la problemática del flujo B2B, y un alcance propositivo al diseñar una arquitectura de datos que resuelve de raíz las ineficiencias del ecosistema de información actual.

8.1.1 Componente de Ingeniería de Datos

Al definir la Inteligencia de Negocios (Business Intelligence, BI), Cano (2018) la caracteriza como un proceso iterativo diseñado para explorar y analizar información estructurada sobre un área específica, con el fin de descubrir tendencias o patrones y derivar de ellos ideas y conclusiones estratégicas (p. 345). En este contexto, la construcción de la solución de ingeniería se asienta sobre un enfoque basado en datos, el cual, según argumenta Curto (2010), exige la automatización e industrialización del ciclo de vida del dato mediante procesos de extracción, transformación y carga (ETL) para sustituir de raíz el procesamiento manual e inestable de planillas locales (p. 80).

Para materializar esta solución, se adopta el patrón lógico de refinamiento progresivo o Arquitectura de Medallón, el cual es detallado por Databricks (2024) como un enfoque estructurado en capas de calidad consecutivas que asumen la limpieza y el resguardo de la integridad de los datos de forma secuencial. Siguiendo el planteamiento de Kimball y Ross (2002), la integración de estas capas analíticas permite unificar las reglas operacionales dispersas en una arquitectura única de almacenamiento, separando conceptualmente el área de preparación (staging) del área de presentación lista para el consumo analítico (p. 1122). El propósito final de esta ingeniería de datos es alcanzar lo que Cano (2018) denomina la única versión de la verdad, un principio fundamental que asegura la integridad semántica para que todos los actores de la cadena de suministro tomen decisiones de reabastecimiento sobre la base de un repositorio unificado y consistente (p. 13).

8.1.2 Alcance de Diagnóstico y Visibilidad

El alcance del proyecto se delimita de manera precisa en dos dimensiones complementarias:

- Alcance de Diagnóstico: Se enfoca en la medición cualitativa y cuantitativa del flujo administrativo diario ejecutado por el Analista CPFR. Para esto, se siguen los principios de eliminación de desperdicios de George (2002) con el propósito de cuantificar las Mudas de sobre procesamiento y tiempos de espera (p. 138), vinculando este diagnóstico con los riesgos de alteración de datos advertidos por Davenport (2006) ante el uso masivo de planillas de cálculo corporativas (p. 103).

- Alcance de Visibilidad: La plataforma de BI se restringe exclusivamente al control del abastecimiento en el canal moderno de Walmart Chile. En consecuencia, el modelo unificado recopila, integra y visualiza los datos diarios a nivel Producto-Sala-Día correspondiente a las dimensiones de Ventas y Stock MultiEstado, y a nivel Producto-Sala-Semana correspondiente a la dimensión de Forecast, con el fin de calcular de forma automatizada los indicadores de abastecimiento (Instock y DOH). Se excluye explícitamente del alcance de automatización del indicador de Fill Rate (Sección 6.3.3), dado que su cálculo depende de la comparación entre unidades solicitadas y unidades efectivamente despachadas en el proceso de Sell-In, información que reside en los sistemas internos de gestión de pedidos y logística de TMLUC (SAP) y no en los reportes de Sell-Out, Stock MultiEstado y Forecast provistos por el portal Retail Link de Walmart.

8.1.3 Criterio de no Intervención en la Fuente

Un requisito indispensable para la viabilidad de este proyecto es el Criterio de No Intervención en la Fuente. La arquitectura se diseña de manera desacoplada para observar las variables operativas en su entorno original (el portal B2B Retail Link) sin alterar en ninguna

circunstancia los sistemas transaccionales, código fuente o bases de datos del retailer.

En su análisis de arquitecturas de Big Data, Gorelik (2019) señala que la capa de ingesta de datos debe estar completamente desacoplada de los procesos de transformación complejos, lo que garantiza una recolección ligera y de bajo impacto en los sistemas de origen. Bajo esta directriz metodológica, la automatización del flujo opera como un proceso de lectura desacoplada. Los scripts de recolección automatizada mediante RPA se limitan a capturar y transferir los archivos planos generados (.TXT) depositados de forma nativa por el portal Retail Link de Walmart en su biblioteca de documentos, asegurando un nulo impacto en el rendimiento de los servidores del cliente y respetando los protocolos de gobernanza tecnológica de ambas compañías.

8.1.4 Enfoque del Ciclo de Vida del Proyecto

La ejecución técnica de esta investigación exige adoptar un marco de gestión del ciclo de vida que garantice la trazabilidad bidireccional entre las decisiones de diseño y sus correspondientes fases de verificación. De las metodologías de desarrollo de sistemas analizadas en el Marco Teórico, se selecciona el Modelo en V de Forsberg y Mooz (1991) por sobre el Modelo en Cascada de Royce (1970) y el marco de trabajo (framework) Scrum de Schwaber y Beedle (2001), debido al control de riesgo e incertidumbre de requerimientos.

Frente al Modelo en Cascada, la diferencia sustantiva radica en la gestión y mitigación del riesgo técnico. Bajo el enfoque estrictamente secuencial de Royce (1970), las actividades de verificación se concentran en una única fase posterior a la implementación completa del sistema (p. 329). En el flujo de datos analizado, esto implicaría descubrir eventuales fallas de integridad referencial o inconsistencias en la coherencia dimensional solo después de haber construido la totalidad de las tres capas de la Arquitectura de Medallón, elevando exponencialmente los costos de retrabajo. Por el contrario, el Modelo en V exige que la estrategia de verificación y validación

se planifique de manera concomitante con cada nivel de especificación técnica. Este enfoque simétrico permite validar e integrar de forma independiente y progresiva las capas Bronce, Plata y Oro a medida que se ejecutan los procesos de extracción, transformación y carga (ETL).

Frente al marco de trabajo Scrum, la diferencia radica en la estabilidad de los requerimientos del proyecto. Conforme sostienen Schwaber y Beedle (2001), los enfoques ágiles y adaptativos son óptimos en contextos donde los requerimientos son altamente cambiantes y evolucionan mediante ciclos iterativos de retroalimentación con el usuario final (pp. 15-17). Sin embargo, en la relación comercial estudiada, los requerimientos técnicos centrales, como el grano de datos, las fórmulas de los indicadores clave de desempeño (Instock y DOH) y la estructura rígida de las fuentes provistas por el portal Retail Link de Walmart o TeamCore, se encuentran determinados externamente por el retailer. Al existir un alcance cerrado y no sujeto a negociación iterativa, disminuye la incertidumbre de requerimientos que justificaría una gestión ágil, favoreciendo una disciplina de diseño predictivo, verificación planificada y control de configuración como la que provee el Modelo en V.

En consecuencia, se adopta el Modelo en V como el marco rector para la gestión del ciclo de vida técnico de la solución de ingeniería, cuya estructura y compuertas de control específicas se detallan posteriormente.

8.2 Fase de Diagnóstico y Levantamiento de la Situación Actual

El diagnóstico preliminar de un proceso de ingeniería requiere establecer un marco descriptivo robusto que documente de forma fidedigna la línea base operacional de la compañía, sirviendo como el marco de comparación científica necesario para validar el impacto de la propuesta. Con este propósito, la evaluación del estado actual del flujo de información del Analista CPFR se diseña mediante un enfoque tridimensional que integra el modelamiento cualitativo de

procesos, la cronometración física de tiempos de ciclo y la evaluación del límite tecnológico de las herramientas heredadas. De esta manera, se sistematizan las ineficiencias dispersas en la operación y se traducen en indicadores objetivos que justifican la transición hacia un entorno analítico automatizado.

8.2.1 Levantamiento y Modelamiento del Proceso Mediante BPMN

Para diagnosticar cualitativamente la secuencia de actividades de reportabilidad, se implementará la notación de modelamiento estándar BPMN. Bajo esta perspectiva, Hitpass (2012) expone que el modelado formal de procesos de negocio es una disciplina indispensable para representar visualmente la interacción entre los recursos humanos y los sistemas de información, permitiendo localizar con exactitud los loops de retrabajo, las compuertas de decisión redundantes y los puntos críticos de falla operacional (p. 13).

- Procedimiento: El diagnóstico se llevará a cabo mediante el mapeo exhaustivo de las actividades revisadas por el Analista CPFR desde el inicio de la jornada hasta la distribución del reporte. Se utilizarán las estructuras de Pools y Lanes para delimitar las fricciones en la transferencia de mensajes entre el portal Retail Link, TeamcORE y la estación de trabajo local.

- Justificación y Objetivo: Este modelamiento permitirá localizar con precisión la Mudas de Retrabajo.

8.2.2 Mapeo de la Cadena de Valor (Value Stream Mapping - VSM)

Mientras el modelamiento BPMN describe la lógica procedimental de las tareas, se hace necesario aplicar el Mapeo de la Cadena de Valor (VSM) para cuantificar la eficiencia temporal del flujo de información. Siguiendo los lineamientos operacionales de George (2002), la cuantificación de las latencias se instrumentará mediante el análisis de la Eficiencia del Ciclo del Proceso (PCE), una métrica de ingeniería de operaciones que contrasta el tiempo que añade valor

estratégico real frente al Lead Time total consumido por el ciclo del negocio (p. 137)

- Procedimiento: Se registra cronométricamente el tiempo destinado a la descarga, limpieza, cruce y distribución de reportes. Para cuantificar la ineficiencia, se aplican las siguientes métricas basadas en la metodología Lean Six Sigma.

- Se calcula la Razón de Tiempo de Valor Agregado (PCE) frente al tiempo total del ciclo. El diagnóstico busca evidenciar que proporción del tiempo del ciclo de flujo información de la Situación Actual se utiliza para generar valor a la compañía. Se utiliza la fórmula:

$$PCE = \frac{Tiempo\ con\ Valor\ Agregado}{Tiempo\ Total\ del\ Ciclo}$$

- Se calcula la Fuga de Valor (FVI), representa el costo de oportunidad de usar ingenieros en carpintería de datos, busca evidenciar una fuga de valor, donde el recurso humano de ingeniería se consume en tareas que el sistema debería realizar automáticamente.

$$FVI = Horas\ Semanales\ de\ Carga\ Manual\ x\ Costo\ HH\ del\ Analista$$

- Justificación y Objetivo: Según Heizer y Render (2009), esta herramientas es vital para identificar el desperdicio oculto (p. 277).

8.2.3 Capacidad Tecnológica del Repositorio de Datos

Complementando el levantamiento del proceso (BPMN) y la cuantificación del Lead Time del proceso mediante el VSM, resulta necesario evaluar si la herramienta utilizada como repositorio de información posee la capacidad técnica suficiente para sostener el volumen de datos generado por la cuenta Walmart. Esta evaluación es indispensable porque, a diferencia de las ineficiencias de proceso identificadas mediante BPMN, una limitación de capacidad tecnológica constituye una restricción estructural que ningún ajuste procedimental puede superar.

- Procedimiento: Se calculará la capacidad del repositorio mediante el indicador de Días

Históricos Teóricos (DHT), para contrastar el límite físico de almacenamiento (Lmax) de la herramienta actualmente utilizada frente a la tasa de generación diaria de registros (RD) de ventas, stock y forecast de la cuenta Walmart.

$$DHT = \frac{L_{max}}{RD}$$

- Justificación y Objetivo: Según Davenport (2006), la saturación de una hoja de cálculo al aproximarse a su límite técnico de filas eleva la probabilidad de error del archivo a un rango entre el 20% y el 40%, inhabilitando su uso confiable para análisis de mediano y largo plazo (p. 101). El objetivo de este análisis es demostrar cuantitativamente que la herramienta actual constituye un techo técnico y no únicamente un problema de proceso que impide sostener el historial de datos necesarios para un análisis de tendencias, justificando así la necesidad de migrar hacia un repositorio cuya capacidad de almacenamiento no esté acotada por las limitaciones nativas de una planilla de cálculo.

8.2.4 Análisis Causa Raíz de Problemas

Una vez identificados los problemas estructurales derivados de la aplicación conjunta de BPMN, VSM y DHT, se procede a determinar la causa raíz de cada uno de ellos aplicando de forma secuencial e integrada las dos herramientas presentadas en el Marco Teórico: primero, la técnica de los 5 Porqué (5 Whys), utilizada como mecanismo primario para trazar la cadena causal de cada problema hasta su origen estructural, y luego, las categorías del Diagrama de Ishikawa (Método, Mano de Obra, Máquina, Medición y Medio Ambiente), empleadas como esquema de clasificación de la causa raíz ya identificada, y no como herramienta de tormenta de ideas visual previa. George (2002) argumenta que el cuestionamiento iterativo del origen de cada problema permite alcanzar su causa estructural, evitando que el diagnóstico se limite a la manifestación superficial del síntoma (p. 111), la clasificación posterior dentro de las categorías del Diagrama de

Ishikawa permite, a su vez, discriminar entre ineficiencias de desempeño individual y restricciones estructurales del ecosistema tecnológico actual, sin requerir la construcción de un diagrama de espina de pescado independiente para cada uno de los problemas estructurales que se identifiquen en el diagnóstico.

8.2.5 Diagnóstico de Madurez de Inteligencia de Negocios

La selección de un marco de diagnóstico de Madurez de Inteligencia de Negocios constituye una decisión metodológica relevante, dado que, coexisten distintos modelos con enfoques y niveles de granularidad diferenciados. Frente al Gartner Maturity Model for Business Intelligence and Performance Management Rayner y Schlegel, (2008, citado en Olszak, 2013, pp. 952-953) y al Business Intelligence Development Model (BIDM) de Sacu y Spruit (2010), ambos exigen la evaluación simultánea de múltiples dimensiones organizacionales (personas, procesos, cultura corporativa y tecnología e nivel de toda la compañía), lo cual excede el Alcance de Diagnóstico definido en la Sección 8.1.2, restringido exclusivamente al flujo de información de la cuenta Walmart

Por el contrario, se selecciona el modelo de Wayne Eckerson como marco de diagnóstico de Madurez de Inteligencia de Negocios debido a su facilidad de implementación práctica dentro del alcance acotado a esta investigación. Al estructurarse como una progresión evolutiva de cinco etapas centrada en la infraestructura tecnológica de datos y en el rol de analista responsable de proceso, el modelo de Eckerson permite clasificar directamente la evidencia empírica recopilada mediante BPMN, VSM y DHT, sin requerir un levantamiento adicional de variables organizacionales transversales a la compañía. Esta característica, resulta coherente con el alcance de diagnóstico de relación comercial TMLUC-Walmart, evitando ampliar innecesariamente el levantamiento hacia dimensiones organizacionales fuera del objeto de estudio.

- Procedimiento: Se evalúa el flujo operativo actual en base a las evidencias BPMN, VSM y DHT recopiladas para clasificar la gestión de la información dentro de las etapas del Modelo de Eckerson. Asimismo, se aplicará la métrica de productividad del tiempo para documentar la proporción del tiempo total del proceso destinada a tareas manuales de normalización de datos, frente al tiempo real disponible dentro de ese mismo proceso para el análisis estratégico de abastecimiento.

Justificación Técnica: Este diagnóstico es indispensable para justificar la transición desde sistemas de soporte de decisiones artesanales hacia una arquitectura de almacenamiento de datos industrializada.

8.3 Procedimiento para el Diseño del Modelo B2B (Basado en el Modelo en V)

Conforme a la rama descendente de especificación lógica del Modelo en V de Forsberg y Mooz (1991), el diseño de la solución se estructurará de forma procedimental en cuatro niveles consecutivos de ingeniería.

8.3.1 Procedimiento de Especificación de Requerimientos e Ingesta Híbrida

Se establece la unidad de análisis bajo el concepto de grano atómico de Kimball y Ross, 2002), definiendo el nivel Producto-Sala-Día para los flujos de Ventas POS y Stock MultiEstado, y Producto-Sala-Semana para el Forecast de demanda (pp. 640-641). Para la captura de datos, se define un mecanismo de ingesta híbrido: un script en Python consumirá de forma programada la API de la plataforma externa Teamcore para extraer Ventas y Stock MultiEstado diarios, mientras que un robot de Power Automate Desktop (RPA) navegará por el portal Retail Link de Walmart para extraer el archivo de Forecast semanal, respetando el Criterio de No Intervención en la Fuente.

8.3.2 Procedimiento de Diseño Dimensional (Constelación de Hechos)

Se formula la arquitectura dimensional en estrella mediante el patrón de Tablas de Hechos

Separadas e interconectadas por dimensiones conformadas (Producto, Sala y Tiempo). Para resolver la diferencia de frecuencia temporal del Forecast sin corromper la integridad del grano, se diseñará una Constelación de Hechos, acoplando los hechos diarios a nivel Fecha y los hechos semanales a la jerarquía de la Dimensión Semana.

8.3.3 Procedimiento para el Diseño de la Arquitectura de Datos (Capa Medallón)

El flujo de datos se canaliza mediante la Arquitectura de Medallón de Databricks (2024).

- Capa Bronce (Data Lake / Landing Zone): Zona de aterrizaje inmutable que almacenará los datos crudos extraídos por la API de Teamcore y el RPA.
- Capa Plata (Data Warehouse): Repositorio relacional en Tercera Forma Normal (3NF). Se aplican reglas de calidad bajo el estándar DAMA-International (2020), incluyendo tipado explícito, derivación de excepciones e inserción incremental mediante la operación UPSERT sobre la clave compuesta (SKU-Sala-Fecha), guardando físicamente la información histórica.
- Capa Oro (Data Marts y Visualización): Capa semántica dimensional que preprocesará los KPIs (Instock, DOH) para la explotación ejecutiva.

8.3.4 Procedimiento de Especificación del Detalle Tecnológico

Conforme al cuarto nivel de la rama descendente del Modelo en V, definido por Forsberg y Mooz (1991) como la fase de mayor granularidad donde se seleccionan las herramientas y tecnologías específicas que materializan cada módulo funcional del diseño lógico previamente establecido, se define un procedimiento de selección tecnológica basado en contrastar un escenario ideal frente a un escenario factible. Para cada capa de la Arquitectura de Medallón (Bronce, Plata y Oro), se evaluará la alternativa tecnológica que ofrezca las garantías óptimas de integridad y escalabilidad, contra la alternativa que resulte efectivamente viable dentro de las restricciones de gobierno de tecnología de TMLUC, particularmente la limitación de aprovisionar infraestructura

en la nube fuera del ecosistema Microsoft ya contratado por la organización (Sección 7.4). La tecnología seleccionada en cada capa deberá, además, preservar el Criterio de No Intervención en la Fuente (Sección 8.1.3) y no representar un requerimiento adicional de licenciamiento o aprobación de TI corporativa fuera del alcance de este proyecto.

8.4 Protocolo de Verificación y Pruebas (Rama Derecha del Modelo en V)

Con el propósito de asegurar la calidad del sistema, el procedimiento de control recorrerá de manera ascendente la rama derecha del Modelo en V. Al respecto, Forsberg y Mooz (1991) fundamentan que el seguramiento de la calidad debe estructurarse mediante compuertas secuenciales de evaluación que validen simétricamente cada nivel de especificación técnica definido en la rama descendente (p. 4).

8.4.1 Pruebas Unitarias (Compuerta 1)

Siguiendo las directrices de Forsberg y Mooz (1991) sobre la verificación aislada de los módulos en el menor nivel de abstracción (p. 9), esta etapa contempla la revisión autónoma de cada componente de software antes de su integración general, bajo cuatro focos procedimentales.

Verificación del Script de Ingesta vía API (TeamCore). Se valida la autenticación contra la API de Teamcore, la correcta recepción e ingesta de los registros de Ventas POS y Stock MultiEstado, y el manejo de excepciones antes fallas de conexión o respuestas incompletas de servicio.

Verificación del Script RPA. Se valida que el robot de extracción realice exitosamente el inicio de sesión, la navegación web, la identificación de directorios de Retail Link y la transferencia de archivos hacia la capa Bronce, midiendo la tasa de fallas de conexión y verificando el manejo de excepciones ante la variabilidad en el tiempo de respuesta del portal externo.

Verificación de Parámetros en DuckBD. Se valida que las funciones SQL y los procesos

de extracción e inserción operan correctamente sobre los esquemas locales asignados.

Verificación de Visualizaciones en Power BI. Se comprueba que cada objeto visual, filtro de segmentación y expresión DAX rinda correctamente y no genere errores semánticos por variables huérfanas antes de la sincronización con SharePoint.

8.4.2 Pruebas de Integración (Compuerta 2)

Conforme al objetivo descrito por Forsberg y Mooz, (1991) de comprobar que los módulos tecnológicos se comuniquen sin pérdidas ni corrupción de información a lo largo de las tres capas de la Arquitectura de Medallón (p. 9).

Pruebas de Ingesta (Bronce a Plata). Se audita la conversión de tipos de datos y la correcta derivación de valores nulos hacia las tablas de excepciones. Se testa la solidez del control de unicidad mediante descargas intencionadas de registros duplicados, verificando que la operación UPSERT mantenga la consistencia de la clave primaria compuesta SKU-Sala-Fecha.

Pruebas de Estructura (Plata a Oro). Se audita la migración desde el esquema 3NF hacia el esquema estrella, verificando que la carga incremental sincronice correctamente las tablas de hechos diarias (Venta, Stock) y semanal (Forecast).

Pruebas de Enlace en la Nube (Oro a SharePoint). Se verifica que los Data Marts se compilen en formato Parquet y se sincronicen automáticamente, garantizando que Power BI Online consuma información actualizada.

8.4.3 Pruebas del Sistema: Nivel de Especificación Funcional (Compuerta 3)

Según Forsberg y Mooz, (1991), esta compuerta evalúa el sistema integrado frente a los requerimientos funcionales del diseño dimensional (p. 8). Se simulan escenarios críticos de consulta en Power BI para validar que las tablas de hechos se unan consistentemente con las dimensiones Producto, Sala y Fecha, y que la Constelación de Hechos agregue los valores diarios

y semanales sin duplicidades ni inflación artificial, respondiendo a los consultas de negocio dentro de tiempos de respuesta adecuados para su uso interactivo.

8.4.4 Pruebas de Aceptación del Usuario (Compuerta 4)

Conforme a Forsberg y Mooz, (1991), esta compuerta valida que el sistema responda los requerimientos de negocio identificados originalmente (p. 8). El criterio de aprobación consistente en certificar que Instock y DOH se calculen automáticamente sobre series de largo plazo, sin intervención artesanal. La aceptación exige un proceso de reconciliación en el cual el Analista CPFR contrasta la totalidad de las combinaciones Producto-Sala frente a los reportes crudos de Retail Link, la coincidencia exacta entre ambas fuentes representa el hito técnico final, garantizando que TMLUC opere bajo la única versión de la verdad definida por Cano (2018) (p. 13).

8.5 Procedimiento de Evaluación Económica

Como se estableció en el Marco Teórico, para evaluar correctamente la rentabilidad de un proyecto de Inteligencia de Negocios es imperativo calcular el Costo Total de Propiedad (TCO), el cual integra no solo la adquisición de tecnología, sino también los costos de operación y el tiempo del personal interno (Cano, 2018, p. 48). En coherencia con ello, la evaluación económica de la solución se estructura en cuatro pasos, siguiendo los criterios de evaluación de proyectos propuestos por Sapag Chain et al. (2014).

- Paso 1 – Costo Total de Propiedad (TCO): Se cuantifican los costos de inversión (horas de diseño, desarrollo y pruebas de la solución), los costos de licenciamiento e infraestructura incrementales y los costos de operación y mantenimiento, valorizados con el costo empresa de la hora hombre. El horizonte de evaluación es de 36 meses, correspondiente a la vida útil esperado de la arquitectura antes de su migración al escenario ideal.

- Paso 2 – Beneficio económico: Se valoriza la capacidad liberada del Analista CPFR a partir de la Fuga de Valor de Ingeniería (FVI) diagnosticada en la Sección 9.2.1, dado que la solución elimina las actividades sin valor agregado del proceso manual.
- Paso 3 – Indicadores de Rentabilidad: Se calculan el Período de Recuperación de la Inversión (PRI) y el Retorno sobre la Inversión (ROI) a 36 meses. Dado que el horizonte de recuperación es inferior a un año, se utilizan flujos sin descontar.

$$PRI = \frac{Inversión\ Incial}{Beneficio\ Mensual - Costo\ Operación\ mensual}$$

$$ROI(36) = \frac{Beneficio\ Acumulado\ (36) - TCO\ (36)}{TCO\ (36)} x100\%$$

- Paso 4 – Análisis de Sensibilidad: Se estima la relación entre el indicador de Instock y la venta perdida diagnosticada en el Capítulo 4, bajo el supuesto de proporcionalidad, para determinar el umbral de mejora de disponibilidad que cubriría el TCO. Este paso no constituye una medición del impacto, sino una estimación del potencial económico de la mejora de visibilidad.
- Justificación: Esta evaluación permite contrastar el costo de la solución con el costo de la ineficiencia diagnosticada, completando el análisis económico que el Marco Teórico establece como condición para validar la rentabilidad de un proyecto de Inteligencia de Negocios.

9 Análisis de la Situación Actual y Resultado de la Propuesta de Solución

9.1 Análisis de la Situación Actual

El diagnóstico de la situación actual constituye la base empírica necesaria para justificar el rediseño del flujo de información entre TMLUC y Walmart. A continuación, se detallan las cinco fases que componen el flujo de información de indicadores de abastecimiento de los productos de TMLUC en las salas Walmart.

Para visibilizar las fricciones operativas, se aplicó la notación BPMN, mapeando las actividades que realiza el Analista CPFR (Figura 1). El flujo de información que sustenta la relación comercial entre TMLUC y Walmart opera bajo un ciclo reactivo diario. Este proceso depende en gran parte de la intervención manual del Analista CPFR.

9.1.1 Fase A: Extracción Manual y Desincronizada de Datos (Latencia Inicial)

La jornada comienza a las 8:30AM con el ingreso manual mediante navegador web al portal Retail Link de Walmart. Al ser una plataforma externa sin conexiones estandarizadas, el analista CPFR debe navegar por los módulos de Decision Support System (DSS) para realizar la extracción de datos.

- Tareas de Descarga: Se extraen tres reportes críticos en formato Excel, Ventas POS, Stock MultiEstado (que incluye los estados On Hand, In Transit, In Whse y On Order) y el Forecast Semanal.

- Procedimiento de Ingesta: Debido a la naturaleza del portal, en el flujo BPMN de la situación actual (Figura 1) se utiliza una compuerta paralela, ya que no existe un orden sistémico, el Analista CPFR descarga los archivos de forma paralela o secuencial según la respuesta del servidor externo, depositándolos temporalmente en su estación de trabajo local.

- Ineficiencias Detectadas en la Fase A: Esta etapa inyecta una latencia inicial crítica.

Al depender exclusivamente de la intervención humana para la captura, en este punto se empieza a generar una brecha temporal entre la generación del dato en el punto de venta y su disponibilidad para el análisis. Como sostiene Christopher (2016), cuando el acceso al dato depende de procesos analógicos, se genera una desconexión que impide a la organización actuar con la velocidad que la logística moderna exige (p. 178).

9.1.2 Fase B: Consolidación y Cruce en Excel (Fase Crítica)

Esta etapa representa el núcleo del problema técnico, identificado como el fenómeno de Excel Caos. Es el eslabón donde el Analista CPFR intenta transformar datos masivos en indicadores operativos sin contar con una arquitectura de soporte robusta.

- Manipulación Manual de Datos: El analista debe abrir el Excel de Procesamiento, limpia los registros del día anterior y procede al copiado y pegado manual de la información de los archivos Excel que se extraen en la Fase A. Esta tarea es altamente riesgosa, dado que se procesan aproximadamente 35.000 combinaciones producto-sala diariamente, es decir, más de 1 millón de registros mensuales tanto en la tabla de ventas, como en la de Stock MultiEstado.

- Cruce con Dimensiones Maestras: Se ejecutan fórmulas locales en Excel (BUSCARV, SUMAR.SICONJUNTO, etc.) para unificar la data de Walmart con las dimensiones de fecha, producto y sala, con el fin de poder realizar posteriormente análisis por familia de productos, ubicación geográfica y temporal.

- Cálculo de KPIs: En el Excel de Procesamiento se aplican fórmulas para calcular los KPIs de abastecimiento, Instock y DOH.

- Ineficiencia Identificada en la Fase B: Se identifica una Muda de Retrabajo provocada por 2 posibles causas: La saturación de la herramienta, el motor de Excel frecuentemente se bloquea o corrompe los archivos ante el volumen de datos, por lo que se debe

volver a ejecutar las fórmulas para obtener los KPIs. La otra causa posible es que la información esté mal descargada desde Retail Link, por lo tanto, el flujo se debe reiniciar desde el inicio de la Jornada.

Según George (2002), estas trampas son procesos que inyectan el 80% o más de retraso total del proceso (p. 127).

9.1.3 Fase C: Bifurcación de Visualización

Una vez consolidada la información de los KPIs en el Excel de Procesamiento el proceso experimenta una fragmentación lógica basada en los requerimientos del destinatario, esto se representa en el flujo BPMN (Figura 1) con una compuerta inclusiva, lo que significa que ambos caminos pueden ocurrir simultáneamente o por separado.

- Excel Visualización Semanal (Corporativo): Una vez por semana, el Analista CPFR debe copiar manualmente los resultados consolidados del Excel de Procesamiento en el Excel de Visualización Semanal exigido por la compañía.

- Excel Visualización Interna (Ad-hoc): La generación de este reporte se activa mediante la solicitud de algún miembro de la cuenta Walmart, posteriormente el Analista CPFR debe copiar manualmente los resultados del Excel de Procesamiento en el Excel de Visualización Interna.

- Ineficiencias Detectadas en la Fase C: Esta duplicidad de tareas genera una fuga de valor del recurso humano y viola el principio el principio de la única verdad. Como señala Cano (2018), la existencia de múltiples versiones de la información en una organización provoca que las reuniones de análisis se desvíen hacia la discusión sobre la veracidad de las fuentes en lugar de enfocarse en la toma de decisiones (p. 7).

9.1.4 Fase D: Distribución y Consultas Ad-hoc (Bucle de Pedidos)

La etapa final consiste en la comunicación de los resultados tanto del Excel de Visualización Semanal y del Excel de Visualización Interna a los interesados correspondientes mediante el envío de correos electrónicos con archivos adjuntos pesados y la carga en carpetas compartidas.

- Bucle de Consultas: Debido a que la información reside en archivos Excel locales y no en una arquitectura de Data Mart accesible, cualquier análisis por jerarquías (familia de productos, región o sala específica) u otro análisis que requiera una acción manual debe ser procesado por el analista ante cada solicitud. En este punto, el Analista CPFR entra en un ciclo repetitivo de pedidos de información donde debe filtrar y reformatear las bases para entregar vistas personalizadas (Loop 3).

- Ineficiencias detectadas en la Fase D: Este escenario provoca que los responsables de generar información estén permanentemente desbordados por peticiones de información urgente, continua y no coordinada. Al no existir una única verdad centralizada, se produce una fuga de valor crítica, en donde el recurso humano de ingeniería se consume en la carpintería de datos en lugar de realizar análisis de causa raíz sobre la cadena logística.

9.1.5 Fase E: Verificación y Ciclo de Retrabajo (Muda de Error)

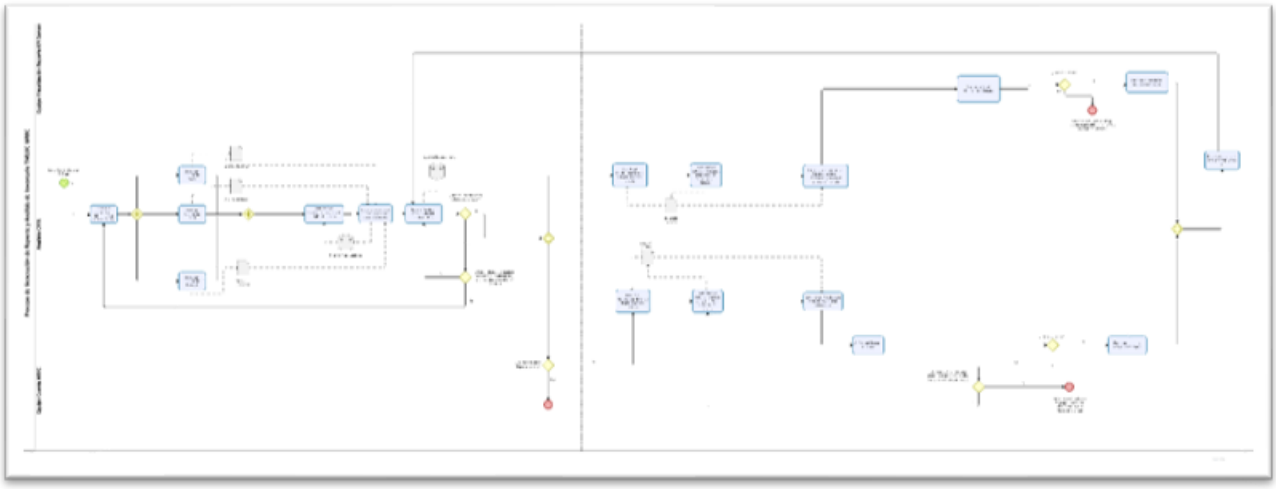
Una vez distribuido el Excel Visualización Semanal y el Excel Visualización Interna, el flujo queda sujeto a una fase de verificación externa por parte de los usuarios finales. En esta etapa, es frecuente la detección de inconsistencias en la asignación de salas de cada miembro del equipo de fuerza de venta o error en cruce de maestro de producto.

- Reinicio del Ciclo: Si se identifica un error, el Analista CPFR debe invalidar el reporte enviado y reiniciar el flujo desde la Fase B (Consolidación y Cruce en Excel).

- Ineficiencias Detectadas en la Fase E: Esta fase contiene una Muda de Retrabajo en su estado más puro. La dependencia de hojas de cálculo es una barrera de calidad, ya que, según Davenport (2006) entre el 20% y el 40% de las hojas de cálculo contienen errores (p. 101). Esta vulnerabilidad sistémica inyecta una variabilidad inaceptable en el proceso, transformando la fase en una trampa de tiempo. En esta última instancia, según Lee et al. (1997) esta falta de fiabilidad provoca que la organización pierda la capacidad de reaccionar oportunamente ante la demanda real, alimentando el Efecto Látigo por la distorsión de las señales de inventario (pp. 546-547).

Figura 1

Diagrama BPMN de la situación actual de flujo de información de TMLUC-Walmart



Nota. Elaboración Propia

9.1.6 Desglose Actividades en Diagrama BPMN de la Situación Actual de Flujo de Información TMLUC-WMC

Con el objetivo de cuantificar objetivamente el Lead Time de flujo de información y de distinguir las actividades que poseen Valor Agregado (VA) de aquellas que constituyen desperdicio (Muda) o carpintería de datos, se aplica el procedimiento de Mapeo de la Cadena de Valor (VSM) descrito en la Metodología. Cada tarea ejecutada por el Analista CPFR, así como

por los demás actores que interactúan con él dentro del proceso, fue registrada documentando el responsable, el sistema o plataforma utilizada, el tiempo promedio de ejecución y la frecuencia con la que se repite dentro de un ciclo semanal.

Dado que no todas las actividades se ejecutan diariamente, si no que algunas responden a una frecuencia semanal, como la consolidación del Excel de Visualización Semanal, o a la concurrencia de un evento específico, como la gestión de una caída de Excel, se utiliza como criterio de estandarización el promedio declarado por el Analista CPFR para cada tarea. Con el fin de homologar estas distintas temporalidades bajo una misma unidad de análisis, el tiempo de cada actividad se transformó a su equivalente en ciclo semanal, multiplicando el Tiempo Promedio de Ejecución por su Frecuencia Semanal.

$$Tiempo\ Semanal = Tiempo\ Promedio\ * Frecuencia\ Semanal$$

La sumatoria de este producto para el total de actividades registradas constituye el Lead Time Total del flujo de información, resultado que se detalla en la Tabla. Adicionalmente, cada actividad se clasifica según la taxonomía Lean de George (2002) en función de si agrega valor al cliente interno o externo del proceso, distinción que permite calcular la Razón de Eficiencia del Ciclo (PCE).

Fase	Tarea	Responsable	Sistema / Plataforma	Tiempo Promedio	Frecuencia Semanal	Clasificación de Actividad Lean
Fase A	Ingresar al portal Retail Link, modulo	Analista CPFR	Retail Link	5 min	5 días	Trabajo operativo (Sin VA)

	DSS.					
Fase A	Extracción y Descarga de Archivos	Analista CPFR	Retail Link	10 min	5 días	Carpintería de Datos (Sin VA)
Fase B	Apertura de Excel de Procesamiento y limpiar registros previos	Analista CPFR	Excel	5 min	5 días	Carpintería de Datos (Sin VA)
Fase B	Copiar y pegar datos de archivos en sus hojas correspondientes	Analista CPFR	Excel	10 min	5 días	Carpintería de Datos (Sin VA)
Fase B	Ejecutar y verificar fórmulas de KPIs	Analista CPFR	Excel	10 min	5 días	Carpintería de Datos (Con VA)
Fase B	Gestión de caídas de Excel y Re-ejecución (Loop 1)	Analista CPFR	Excel	40 min	1 día	Carpintería de Datos (Sin VA) - Muda

Fase C	Abrir Excel Visualización Semanal y limpiar registros previos	Analista CPFR	Excel	5 min	1 día	Carpintería de Datos (Sin VA)
Fase C	Copiar base de datos consolidado y pegar en Excel de Visualización semanal	Analista CPFR	Excel	10 min	1 día	Carpintería de Datos (Sin VA)
Fase C	Abrir Excel Visualización Interna y limpiar registros previos	Analista CPFR	Excel	5 min	5 días	Carpintería de Datos (Sin VA
Fase C	Copiar base de datos consolidado y pegar en Excel de Visualización Interna	Analista CPFR	Excel	10 min	5 días	Carpintería de Datos (Sin VA)
Fase D	Enviar Reporte KPI Semanal	Analista CPFR	Correo	3 min	1 día	Comunicación Manual (Sin

						VA)
Fase D	Enviar Reporte KPI Interno	Analista CPFR	Correo	3 min	5 días	Comunicación Manual (Sin VA)
Fase D	Análisis Excel Visualización Interna	Equipo Cuenta Walmart	Excel	20 min	3 días	Análisis (Con VA)
Fase E	Identifica algún error al visualizar Reporte KPI	Visualizadores		--	1 día	Comunicación (Sin VA)
Fase E	Detecta donde se produjo el error y corrige (Loop 2)	Analista CPFR	Excel	45 min	1 día	Carpintería de Datos (Sin VA)
Fase E	El equipo Cuenta Walmart solicita otro análisis y el Analista lo prepara (Loop 3)	Analista CPFR	Excel	15 min	2 días	Carpintería de Datos (Sin VA)

Al aplicar la sumatoria del Tiempo Semanal sobre el total de las dieciséis actividades

registradas, se obtiene que el Lead Time Total de flujo de información asciende a aproximadamente 483 minutos semanales (equivalente a 8 horas por semana), correspondiente al tiempo acumulado del proceso completo, desde el ingreso al portal Retail Link hasta la resolución de las consultas del Equipo de la Cuenta Walmart, requiere completar un ciclo de descarga manual, consolidación y cruce, generación, distribución y verificación de información.

De la clasificación Lean aplicada a cada actividad, se observa que únicamente dos de las dieciséis tareas registradas están catalogadas como actividades con Valor Agregado, la ejecución y verificación de fórmulas KPI en la Fase B y Análisis de Excel Visualización Interna en la Fase D, mientras que el resto corresponde a Carpintería de Datos, Comunicación Manual o, en el caso de la gestión de caídas de Excel (Fase B), a una Muda explícita. Esta proporción evidencia de forma preliminar que la fracción de tiempo destinada a la generación de valor es marginal frente al tiempo total del ciclo, hallazgo que se cuantificará formalmente mediante el indicador de Eficiencia del Ciclo del Proceso (PCE).

Asimismo, se identifica que las tres actividades asociadas a ciclos de reproceso (Loop 1, Loop 2, Loop 3) concentran en conjunto 115 minutos semanales, equivalentes a cerca de un 24%, lo cual confirma que una porción relevante de tiempo del Analista CPFR no se destina a la ejecución del proceso estándar, sino a la corrección fallas estructurales y reproceso de la información debido a solicitudes generadas.

9.2 Cuantificación de Ineficiencias y Análisis de Criticidad

Tras el modelamiento del flujo BPMN, es imperativo transformar las observaciones cualitativas en métricas para dimensionar el impacto de la situación actual en los activos de la compañía. Para ello, se han seleccionado tres indicadores clave que miden el costo de oportunidad del talento humano, la esbeltez del proceso y la viabilidad técnica del soporte tecnológico utilizado.

9.2.1 Indicador 1: Fuga de Valor de Ingeniería

Este indicador cuantifica el costo económico directo que asume TMLUC al destinar recursos humanos calificados, tales como ingenieros civiles o analistas de proceso a tareas que se definen como carpintería de datos o normalización manual de la información, las cuales no transforman el producto ni agregan valor estratégico al negocio.

Según Cano (2018), en arquitecturas de inteligencia de negocio incipientes, se suele destinar un 80% del tiempo a preparar la información y solo un 20% a analizarla (pp. 20, 86). El cálculo de esta fuga se expresa mediante la siguiente relación:

$$FVI = Horas\ Semanales\ de\ Carga\ Manual\ x\ Costo\ HH\ del\ Analista$$

Considerando los datos levantados por el Analista CPFR:

- Horas semanales destinadas a tareas manuales: 8 horas semanales x 0,77 (1-PCE).
- Costo promedio Horas Hombre (HH) del Analista: $11.000 por hora.
- Resultado Semanal Fuga de Valor (FVI): 68.000 semanales.

Este resultado equivale a un gasto de $272.000 mensuales destinado exclusivamente a financiar una ineficiencia estructural de la arquitectura de información. Al desviar las capacidades del Analista CPFR hacia tareas de soporte que deberían estar completamente automatizadas por un motor lógico de ingeniería de datos, se asume un costo de oportunidad crítico, el tiempo que el profesional debería emplear en la mitigación de quiebres de stock en sala, control preventivo de alertas y optimización logística conjunta, se consume en la manipulación manual de planillas de cálculo.

9.2.2 Indicador 2: Eficiencia del Ciclo del Proceso (PCE)

La esbeltez y la velocidad de respuesta de un flujo de información se evalúan mediante la Eficiencia del Ciclo del Proceso (PCE), un indicador clave de la metodología Lean Six Sigma. De

acuerdo con George (2002), el PCE mide la relación existente entre el tiempo que realmente agrega valor al cliente (actividades de transformación o análisis estratégico) y el Lead Time total que tarda el proceso desde su inicio hasta su entrega final (p. 137).

$$PCE = \frac{Tiempo\ con\ Valor\ Agregado\ (VA)}{Lead\ Time\ Total\ del\ Ciclo\ (LTT)}$$

Donde:

- Tiempo VA (Análisis Estratégico): 110 minutos a la semana.
- Lead Time Total: 483 minutos por semana.
- PCE Resultante: 23%.

De acuerdo con los estándares de Heizer y Render, (2009), en un proceso optimizado las actividades de valor agregado deberían representar la mayor fracción del tiempo total (p. 277). Según lo establecido por George (2002) un PCE igual o inferior al 20% confirma que el proceso está dominado por las denominadas trampas de tiempo, las cuales inyectan el 80% o más del retraso en la entrega de información estratégica hacia los tomadores de decisiones (p. 127), provocando que la información llegue a los tomadores de decisión cuando esa información ya ha sido degradada y ya ha perdido su valor estratégico.

Este resultado justifica técnicamente la necesidad de implementar una arquitectura de datos automatizada. El objetivo de ingeniería del proyecto es comprimir el Lead Time Total mediante la eliminación de la carpintería de datos manual, permitiendo elevar el PCE hacia niveles del 50% o más, garantizando que el Analista CPFR destine la mayor parte de su tiempo disponible a la mitigación proactiva de quiebres de stock para mejorar el abastecimiento en sala.

9.2.3 Indicador 3: Capacidad y Límite Tecnológico

Este indicador evalúa la viabilidad técnica de Microsoft Excel como repositorio o base de

datos analítica frente a la volumetría real de transacciones de la cuenta Walmart. Su justificación se basa en contrastar la tasa de generación diaria de datos del negocio frente al límite físico absoluto de almacenamiento de la herramienta de oficina, determinando si el soporte tecnológico actual constituye una restricción estructural insalvable para el análisis a largo plazo.

Para modelar esta limitación, se aplica la relación de Días Históricos Teóricos (DHT):

$$DHT = \frac{Límite\ de\ Filas\ de\ Excel}{Registros\ Diarios\ (SKU - Sala)}$$

Bajo las condiciones actuales:

- Registros generados diariamente: 35.000 registros de ventas y stock. A esta cifra se añaden 7.000 registros de forecast generados con una cadencia semanal (grano Producto-Sala-Semana), los cuales, al residir en una hoja de cálculo independiente dentro del mismo Excel de Procesamiento (Sección 9.1.2), no inciden en la tasa diaria de generación de registros utilizada para el cálculo del DHT aquí presentado.
- Límite de filas de la herramienta: 1.048.576.
- Capacidad Teórica de almacenamiento histórico: 29 días.

El resultado matemático del indicador de Días Históricos Teóricos (DHT) arroja una capacidad teórica máxima de almacenamiento histórico de 29 días para la planilla de cálculo actual. Sin embargo, la evaluación de esta métrica bajo los estándares de la ingeniería de sistemas y la administración de operaciones revela deficiencias estructurales y críticas que degradan significativamente el flujo de información de la cuenta Walmart. Aunque de manera teórica el cálculo sugiere la viabilidad de albergar un mes de transacciones, pero en la práctica el procesamiento continuo de una volumetría masiva de transacciones diarias satura rápidamente el rendimiento del software de oficina.

Esta saturación tecnológica genera una severa degradación operacional, volviendo el archivo de cálculo local inmanejable debido al procesamiento de cientos de miles de filas interconectadas con fórmulas de búsqueda y consolidación local, tales como BUSCARV y SUMAR.SI.CONJUNTO. La sobrecarga de cálculo no solo provoca congelamientos constantes en la estación de trabajo, sino que induce fallas de procesamiento recurrentes que interrumpen la continuidad del flujo analítico. Este fenómeno obliga al Analista CPFR a reiniciar ciclos completos de preparación de datos, lo que se traduce directamente en la ejecución sistemática de la Muda de Retrabajo identificada como el Loop 1 en el modelamiento del proceso actual bajo la notación BPMN.

A nivel de confiabilidad de la información, el empleo de planillas de hojas de cálculo de propósito general como un repositorio improvisado de bases de datos eleva exponencialmente la probabilidad de errores humanos e instrumentales en la manipulación de registros. Como demuestra la investigación de Davenport (2006), la carencia de flujos de datos integrados y estandarizados constituye la principal barrera estructural que impide a las organizaciones progresar hacia estados superiores de madurez analítica, transformando la proliferación de planillas individuales en un terreno fértil para las inconsistencia de datos (p. 103).

Esta ineficiencia en el tratamiento del dato se traduce directamente en una significativa Fuga de Valor de Ingeniería (FVI), dado que la falta de un repositorio automatizado condena al Analista CPFR a realizar manualmente actividades de limpieza, unificación y cruce de datos maestros. Bajo este enfoque artesanal, se cumple de manera sistemática el axioma de Curto (2010), el cual advierte que la carencia de una infraestructura robusta de inteligencia de negocios obliga a los profesionales de análisis a desperdiciar aproximadamente el 80% de su jornada laboral únicamente en tareas de preparación y normalización de información (p. 30). En consecuencia, se

dispone únicamente de un 20% del tiempo disponible para la ejecución del diagnóstico estratégico de la cadena logística, sacrificando el análisis de causa raíz sobre los quiebres de stock y limitando el potencial analítico del recurso de ingeniería.

En conclusión, las planillas locales de cálculo representan un límite tecnológico y estructural infranqueable que bloquea el desarrollo analítico y la esbeltez operativa de la relación comercial. Al ser una restricción estructural, ninguna mejora procedimental o capacitación del Analista CPFR puede solucionar la limitación física de la herramienta de oficina para almacenar y procesar la volumetría atómica de Walmart. Por lo tanto, se justifica metodológicamente el diseño de una arquitectura de datos que separe de raíz la capa de ingesta mediante un Data Lake (capa Bronce) para garantizar la captura continua, y consolide el histórico unificado en un Data Warehouse (capa Plata) de capacidad y rendimiento virtualmente ilimitados.

9.3 Análisis del Potencial Desaprovechado y Ceguera Operativa

El diagnóstico de la situación actual revela una paradoja crítica, la organización dispone de una masa crítica de datos proveniente de Walmart, pero carece de la capacidad de transformarlos en conocimiento accionable.

A continuación, se detallan las barreras estructurales localizadas en el flujo BPMN que bloquean la creación de valor y justifican la necesidad de una arquitectura industrializada.

9.3.1 Problema 1: El Sistema de Información Artesanal

Esta Problema se genera en la Fase A (Descarga Manual de Archivos) y Fase B (Consolidación y Cruce). Se identifica por la saturación del único carril ejecutor (Analista CPFR), quien debe realizar manualmente la carpintería de datos.

Se observa en la Tarea de Ingresar al portal Retail Link y descargar los tres archivos manualmente en formato Excel, también se observa en todas las tareas de la Fase B, en donde

existe una dependencia manual del Analista CPFR, ya que debe copiar y pegar la información de los archivos, y estar pendiente ante cualquier falla en el Excel de Procesamiento. En este punto, se observa un bucle de Muda de Retrabajo. Al no existir un motor lógico, el Analista CPFR destina un 77% del tiempo del proceso a tareas de normalización y solo un 22% a tareas de diagnóstico estratégico.

Los indicadores FVI ($272.000 mensuales) y PCE (23%), representan una señal inequívoca que la situación actual se considera un sistema artesanal que depende de que una persona prepare la información para que la organización sepa que está sucediendo. Dado que el Analista CPFR se encuentra desbordado por peticiones continuas y no coordinadas, como responsable de los indicadores carece de la posibilidad de realizar análisis causa raíz sobre los indicadores de abastecimiento, por lo tanto la gestión se transforma meramente en acciones reactivas y no proactivas.

9.3.2 Problema 2: Ceguera Operativa Longitudinal

Se localiza en los objetos de datos del BPMN (Excel de Procesamiento y los Excel de Visualización). Este problema es un subproducto del límite técnico de las herramientas utilizadas para el almacenamiento.

Cada vez que algún miembro de la compañía requiere una vista histórica o un análisis jerárquico específico se deben reprocesar los archivos locales. El indicador DHT demuestra que el límite técnico y práctico de Excel de 1.048.576 filas frente a una carga de 35.000 registros diarios y varias fórmulas locales, permite retener entre 1 y 3 días de historia, volviéndose inmanejable.

Esta restricción impide descubrir patrones estacionales o tendencias de largo plazo. La organización permanece confinada en una gestión reactiva de 1 a 3 días, alimentando el Efecto Látigo por la distorsión de señales de demanda histórica necesaria para estabilizar los pedidos.

9.3.3 Problema 3: Fragmentación de la Única Verdad

Esta barrera se origina en la compuerta inclusiva de la Fase C. La bifurcación manual de datos hacia el Excel Semanal Corporativo y el Excel de Visualización Interna permite la coexistencia de versiones discrepantes de los KPIs (Instock y DOH).

Como sostiene Davenport (2006), la falta de una base de datos estandarizada e integrada es el problema principal que impide a las empresas competir a través del análisis de datos (p. 103). En TMLUC, esto deriva en reuniones donde el tiempo se consume discutiendo la veracidad de las fuentes en lugar de ejecutar planes de acción.

9.3.4 Problema 4: Ciclo de Retrabajo Crítico

Se manifiesta en la Fase E (Verificación y Retrabajo) del flujo BPMN, si el cliente detecta una inconsistencia en la Fase E, se invalida gran parte del tiempo ya invertida en ese ciclo del proceso, obligando a reiniciarlo desde la Fase B.

Esta vulnerabilidad, cuantificada en la Sección 9.2, eleva la fiabilidad del diagnóstico a un nivel insuficiente para la alta exigencia del canal moderno.

9.3.5 Problema 5: Pérdida de Información por Fin de Semana, Vacaciones, Motivos Exógenos

Esta barrera se origina en la Fase A del flujo actual (Figura 1) y evidencia una fractura crítica en la disponibilidad de los activos de información de la compañía. Se manifiesta a través de dos dimensiones de ineficiencia técnica.

- Volatilidad del Dato de Stock Multiestado: A diferencia de los registros históricos de ventas, la información contenida en el archivo de Stock Multiestado no es persistente en el portal Retail Link, es decir representa un flujo dinámico de inventario que solo está disponible para su captura durante la jornada en curso. De acuerdo con Curto (2010), para que un dato sea útil en la toma de decisiones, debe cumplir con la expectativa de disponibilidad, lo que implica

que el dato sea capturado y almacenado en un repositorio persistente en el momento adecuado (pp. 102-103). Bajo el modelo actual, si la descarga manual en la Fase A no se ejecuta en el momento exacto, el dato se pierde de forma irreversible, imposibilitando la reconstrucción de la historia de Stock MultiEstado y limitando la capacidad de realizar análisis causa raíz sobre disponibilidad de producto en góndola.

9.4 Análisis Causa Raíz y Oportunidad de Mejora

El Análisis Causa Raíz corresponde a la fase de Análisis de la Situación Actual, aplicando de forma sistemática la técnica de los 5 porqué (5 Whys) junto con la clasificación del Diagrama de Ishikawa sobre cada de uno de los problemas estructurales identificados en la Sección 9.3.

9.4.1 Problema 1: El Sistema de Información Artesanal

- Análisis Causa Raíz: Se examina el problema estructural mediante la técnica de los 5 Porqué, complementada con la clasificación del Diagrama de Ishikawa. ¿Por qué el Analista CPFR destina el 77% del tiempo de este proceso a tareas de consolidación y normalización? Porque debe descargar, copiar y consolidar manualmente tres archivos proveniente de Retail Link en cada ciclo. ¿Por qué debe hacerlo manualmente? Porque no existe un mecanismo automatizado de extracción, transformación (ETL) entre Retail Link y el Excel de Procesamiento. ¿Por qué no existe este mecanismo? Porque el ecosistema Tecnológico actual de TMLUC no cuenta con una arquitectura de datos formal, delegando la integración de sistemas al factor humano.

Esta cadena permite descartar que la causa raíz corresponda a la categoría Mano de Obra, el mismo cuello de botella se replicaría sin importar que profesional ejecute la tarea, por lo que no es un problema de desempeño individual.

La causa se ubica en la categoría Máquina (Herramienta Tecnológica), la ausencia estructural de un motor lógico automatizado. Como sostiene Curto (2010), la falta de una

infraestructura sólida condena a la organización a niveles operativos de bajo valor, donde el tiempo se consume en la preparación del dato y no en su análisis (p. 30). Bajo esta lectura, el 77% de tiempo en normalización no es la causa, es su síntoma más visible; la causa raíz es la inexistencia de un pipeline ETL que separe la ingesta de la interpretación del dato.

9.4.2 Problema 2: Ceguera Operativa Longitudinal

- Análisis Causa Raíz: Se examina el problema estructural mediante la técnica de los 5 Porqué, complementada con la clasificación del Diagrama de Ishikawa. ¿Por qué la organización solo puede analizar entre 1 y 3 días de historia? Porque el Excel de Procesamiento se satura al acercarse a su límite tecnológico de 1.048.576 filas frente a los 35.000 registros diarios generados por la cuenta Walmart. ¿Por qué se utiliza una planilla de cálculo para almacenar este histórico? Porque la organización no cuenta con un repositorio de datos formal (Data Warehouse) que centralice la información. ¿Por qué no existe ese repositorio? Porque la infraestructura tecnológica actual de TMLUC fue concebida para generar reportes puntuales y no sostener series de tiempo de largo plazo.

Esta cadena permite descartar que la causa raíz corresponde a la categoría Mano de Obra, ningún nivel de habilidad o dedicación del Analista CPFR le permitiría retener más historia de lo que la herramienta es técnicamente capaz de almacenar, por lo que el límite no depende del ejecutor, sino del instrumento.

La causa se ubica en la categoría Máquina (Herramienta Tecnológica), la incapacidad estructural de la planilla de cálculo para actuar como repositorio persistente. Un Data Warehouse, por definición de Inmon (2005) debe ser no volátil y variante en el tiempo para sostener el análisis histórico (pp. 29-31), Excel, al no cumplir ninguna de estas dos propiedades, es estructuralmente incapaz de sostener un análisis de tendencias. Bajo esta lectura, la ventana de 1 a 3 días de historia

disponible no es la causa del problema, es su síntoma; la causa raíz es la ausencia de un repositorio con capacidad de almacenamiento no acotada por las limitaciones nativas de una hoja de cálculo.

9.4.3 Problema 3: Fragmentación de la Única Verdad

- Análisis Causa Raíz: Se examina el problema estructural mediante la técnica de los 5 Porqué, complementada con la clasificación del Diagrama de Ishikawa. ¿Por qué existen dos versiones distintas de los mismos KPI (Instock y DOH)? Porque el proceso bifurca manualmente el dato consolidado hacia dos archivos independientes, el Excel de Visualización Semanal Corporativo y el Excel de Visualización Interna, en la compuerta de la Fase C. ¿Por qué se bifurca manualmente en lugar de derivar automáticamente ambos reportes de una misma fuente? Porque no existe una única fuente de datos (Data Mart) de la cual ambos archivos puedan consultar de forma sincronizada. ¿Por qué no existe una única fuente de datos? Porque el modelo de datos actual no está diseñado bajo un esquema de tabla de hechos conectadas por dimensiones maestras compartidas, sino que cada archivo constituye una copia local e independiente.

Esta cadena permite descartar que la causa raíz corresponda a la categoría Medición, los KPI de Instock y DOH están correctamente definidos a nivel conceptual, por lo que el problema no es la calidad de la métrica en sí, sino la forma en que se distribuye.

La causa se ubica en la categoría Método, la arquitectura del proceso obliga a duplicar el dato en dos flujos paralelos sin mecanismo de sincronización. Como sostiene Davenport (2006), la falta de una base de datos estandarizada e integrada es la barrera principal que impide a las organizaciones competir a través del análisis de datos (p. 103). Bajo esta lectura, la discrepancia entre cifras del Excel de Visualización Semanal y del Excel de Visualización Interna no es la causa del problema, es su síntoma; la causa raíz es la ausencia de una única fuente de verdad de la cual ambos reportes deriven automáticamente.

9.4.4 Problema 4: Ciclo de Retrabajo Crítico

- Análisis Causa Raíz: Se examina el problema estructural mediante la técnica de los 5 Porqué, complementada con la clasificación del Diagrama de Ishikawa. ¿Por qué el error en los datos se detecta recién en la Fase E, cuando el cliente lo reporta? Porque el proceso actual no reporta ningún control de calidad en las etapas previas de extracción y consolidación (Fase A y B). ¿Por qué no existen esos controles? Porque el flujo no incorpora validaciones automáticas de integridad referencial ni de consistencia entre los archivos descargados y los maestros de producto. ¿Por qué no las incorpora? Porque el proceso fue construido íntegramente sobre herramientas de oficina de propósito general, que no tiene capacidad nativa de aplicar reglas de negocio ni de bloquear el avance de un dato inconsistente.

Esta cadena permite descartar que la causa raíz corresponda a la categoría Mano de Obra, el error no se origina en un descuido puntual del Analista CPFR, sino que el propio diseño del proceso no ofrece ningún punto de control automático antes de la entrega.

La causa se ubica en la categoría Método, la ausencia de reglas de integridad técnica en el punto de captura. La cifra de error estructural ya cuantificada en la Sección 9.2.3 no es un defecto de la persona que manipula la hoja de cálculo, sino la consecuencia esperable de un proceso sin controles de calidad en ninguna de sus etapas. Bajo esta lectura, la inconsistencia detectada por el cliente en la Fase E no es la causa del problema, es su síntoma; la causa raíz es la ausencia de validaciones automáticas de integridad en las etapas tempranas del flujo.

9.4.5 Problema 5: Pérdida de Información por Fin de Semana, Vacaciones, Motivos Exógenos.

- Análisis Causa Raíz: Se examina el problema estructural mediante la técnica de los 5 Porqué, complementada con la clasificación del Diagrama de Ishikawa. ¿Por qué se pierde de forma irreversible el dato de Stock MultiEstado si no se descarga a tiempo? Porque este dato no

es persistente dentro del portal Retail Link, sino que representa un flujo dinámico de inventario disponible solo durante la jornada en curso. ¿Por qué la organización depende de que alguien lo descargue en el momento exacto? Porque no existe un mecanismo de ingesta automatizada que capture el dato de forma continua, independiente de la intervención humana. ¿Por qué no existe esa ingesta automatizada? Porque el 100% de las actividades de captura del flujo actual dependen de la ejecución manual de una sola persona, sin mecanismo de respaldo ni redundancia.

Esta cadena permite descartar que la causa raíz corresponda a la categoría Medio Ambiente, la ocurrencia de fines de semana, vacaciones o ausencias no constituye una anomalía externa, sino un factor de calendario normal y predecible que cualquier proceso robusto debería poder absorber.

La causa se ubica en la categoría Máquina, específicamente en la dependencia estructural de un único ejecutor sin automatización de respaldo. Según Curto (2010), para que un dato sea útil en la toma de decisiones debe cumplir con la expectativa de disponibilidad, lo que implica que sea capturado y almacenado en un repositorio persistente en el momento adecuado (pp. 102-103). Bajo esta lectura, la pérdida del dato durante una ausencia del Analista CPFR no es la causa del problema, es su síntoma; la causa raíz es la inexistencia de un proceso de captura automatizado que no dependa de la disponibilidad de una sola persona.

9.4.6 Síntesis de Causa Raíz

El análisis de causa raíz aplicado a los 5 problemas estructurales revela que, si bien cada uno se manifiesta en una fase distinta del flujo BPMN, todos convergen en un número reducido de causas estructurales compartidas. Los problemas 1, 2 y 5 comparten una causa común de naturaleza tecnológica, la ausencia de automatización en las etapas de captura, almacenamiento e ingesta de datos. Los problemas 3 y 4, en tanto, comparten una causa común de naturaleza estructural

asociada a la ausencia de una única fuente de datos gobernada. En ningún caso la causa reside en el desempeño del Analista CPFR, sino en la inexistencia de una arquitectura de datos formal que sustente el proceso.

Esta convergencia de causas explica por qué, al evaluar a la organización bajo el Modelo de Madurez de Inteligencia de Negocios de Wayne Eckerson, TMLUC se clasifica en un estado de madurez temprano, las cinco barreras no son fallas aisladas y corregibles de forma independiente, sino síntomas independientes consistentes de una misma etapa de desarrollo analítico, lo que justifica el diseño de una solución arquitectónica única, en lugar 5 soluciones paralelas.

9.5 Diagnóstico de Madurez de Business Intelligence según el Modelo de Wayne Eckerson

A partir de la evidencia empírica levantada mediante el modelamiento del proceso en notación BPMN y el Mapeo de la Cadena de Valor (VSM), la gestión de información en la relación comercial B2B entre TMLUC y Walmart se analiza bajo el marco teórico del Modelo de Madurez de Inteligencia de Negocios de Wayne Eckerson. Este marco de referencia permite evaluar la capacidad real de los sistemas de información de una compañía para unificar reglas de negocio, mitigar la latencia y transformar datos atómicos en activos estratégicos.

9.5.1 Clasificación en la Etapa 1: Prenatal / Infant (La "Infancia" Analítica)

La caracterización del flujo de datos actual determina que la organización se sitúa estructuralmente en la Etapa 1 (Prenatal / Infant) del modelo de Eckerson. Esta fase se define académicamente por la ausencia de entornos interactivos de explotación de información y por la dependencia severa de desarrollos manuales.

En el caso de estudio, la clasificación se fundamenta en la coexistencia de dos fenómenos característicos de este estado de madurez.

- Silos en hojas de cálculo: Toda la información de Venta, Stock MultiEstado y Forecast es descargada de forma desintegrada y almacenada en planillas locales de Microsoft Excel. Al no contar con una base de datos centralizada, cada archivo opera como un silo de información aislado, lo que fragmenta la consistencia semántica y expone el flujo a duplicidades e inconsistencias en la reportabilidad.

- La Configuración de un Human Data Warehouse (Data Warehouse Humano): El análisis cuantitativo del tiempo del Analista CPFR revela que este recurso asume el rol de un procesador manual de datos. El profesional destina aproximadamente el 77% del tiempo del proceso semanal (8 horas semanales de proceso) exclusivamente a tareas operativas de extracción, copiado, pegado, limpieza y normalización física de archivos planos. Esta sobrecarga operativa de carpintería de datos reduce a solo un 23% el tiempo residual disponible para la ejecución de diagnósticos y gestión del abastecimiento a las salas.

9.5.2 Justificación de la Exclusión de Estados Superiores

Para robustecer la clasificación se descarta técnicamente la pertenencia de la organización a otras etapas del modelo:

- Etapa 2 (Child): El modelo de Eckerson establece que la transición a la Etapa 2 (Child) requiere que los líderes departamentales impulsen repositorios de datos independientes (Data Mart). Aunque en la relación comercial existe un rol formal dedicado a la preparación de reportes (Analista CPFR), el flujo carece de una base de datos o repositorio analítico intermedio. La data transaccional sigue residiendo en planillas Excel vulnerables, cuya saturación y degradación de rendimiento ante una volumetría de 35.000 registros diarios limitan la retención histórica a solo 1 a 3 días de historia disponible (véase problema 2, sección 9.4.2). Por lo tanto, no se cumple con la definición de un repositorio departamental unificado.

- Etapa 4 (Adult): Este nivel exige la existencia de una única versión de la verdad corporativa, sustentada en un Data Warehouse integrado y gobernado transversalmente por el negocio con el soporte de un Centro de Excelencia de Inteligencia de Negocios (BICC). La situación observada en la compuerta inclusiva de la Fase C del proceso actual, donde coexisten visualizaciones discrepantes y reportes duplicados (Excel Semanal Corporativo y Excel de Visualización Interna), evidencia una asimetría de información que imposibilita la clasificación en este estado de madurez

9.5.3 El Desafío del Punto de Fricción Gulf (El Golfo)

El diagnostico analítico concluye que la relación B2B de TMLUC con Walmart se encuentra estancada inmediatamente antes del primer gran límite del modelo, el Punto de Fricción Gulf (El Golfo). Cruzar este golfo exige conseguir un patrocinador directivo, resolver la escasez de licencias analíticas, capacitar formalmente al personal y enfrentar directamente los problemas de calidad de los datos de origen.

La dependencia de planillas de cálculo representa una restricción estructural crítica frente a este desafío. Como advierte Davenport (2006), el uso de hojas de cálculo para procesar grandes volúmenes de datos corporativos eleva la probabilidad de error en el archivo a un rango de entre el 20% y el 40% (p. 101). Esta vulnerabilidad sistémica inyecta desconfianza en los reportes, genera ciclos de retrabajo constantes (Loop 1, Loop 2 en el BPMN) y anula la capacidad predictiva de la compañía, alimentando el Efecto Látigo ante variaciones de demanda que no son detectadas a tiempo por la latencia del flujo analógico.

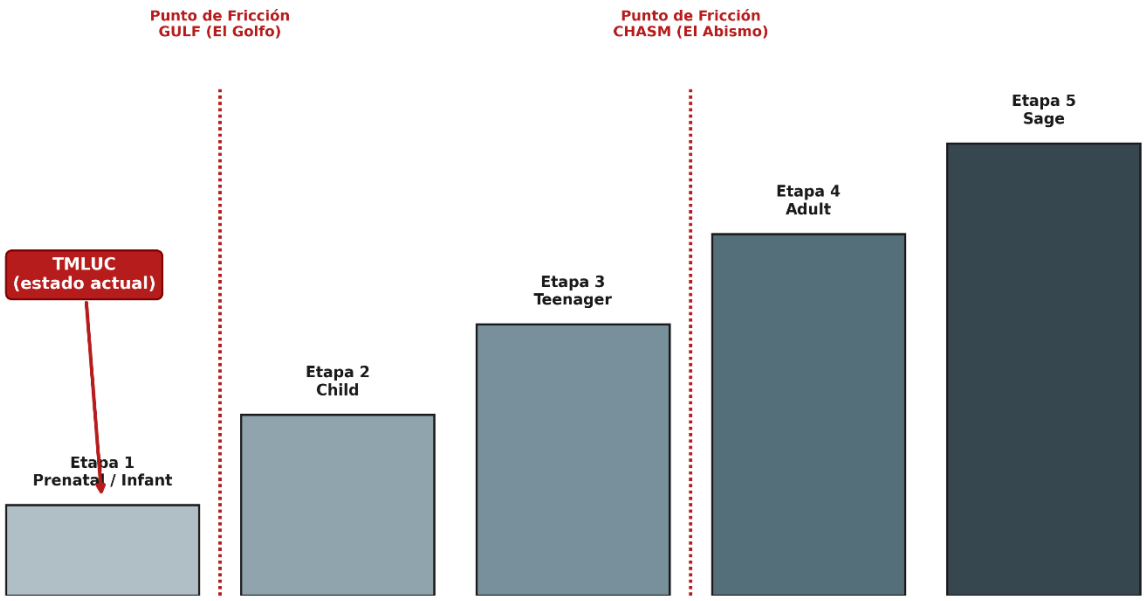
Este diagnóstico justifica metodológicamente la necesidad de implementar el modelo de integración y gobernanza propuesto. La automatización del motor ETL y la estructuración de la Arquitectura de Medallón (Bronce, Plata y Oro) operarán como el vehículo de ingeniería para

cruzar de forma permanente el Punto de Fricción Gulf. Esto permitirá transitar desde una gestión artesanal del dato hacia un estado industrializado que garantice la integridad y disponibilidad de la información para la toma de decisiones estratégicas.

Para sintetizar visualmente el diagnóstico de madurez presentado en esta sección, la Figura 2 representa las cinco etapas del modelo de Eckerson junto con los puntos de fricción críticos (Gulf y Chasm) descritos anteriormente, señalando la posición actual de TMLUC en la Etapa 1 (Prenatal/Infant), inmediatamente antes del primer punto de fricción estructural.

Figura 2

Modelo de Madurez de Inteligencia de Negocios de Wayne Eckerson y posición actual de TMLUC



Nota. Elaboración propia a partir del modelo de madurez de Eckerson, citado en Lituus Analytics (2023).

9.6 Resultado del Desarrollo de la Solución

Habiendo diagnosticado la situación actual y sus causas raíz (Objetivo Específico 1), se presenta a continuación el resultado del diseño e implementación de la arquitectura de datos

propuesta, correspondiente a los Objetivos Específicos 2, 3 y 4, estructurado según los cuatro niveles de especificación técnica del Modelo en V definidos en la Metodología (Sección 8.3).

9.6.1 Nivel 1 del Modelo en V: Análisis de Requisitos del Negocio y Catálogo de Fuentes

De acuerdo con el ciclo de vida simétrico de Forsberg y Mooz (1991), el primer paso en el diseño de un sistema consiste en la captura de las necesidades del negocio para traducirlas en requerimientos técnicos explícitos y medibles (p. 8). En este proyecto, la necesidad de negocio identificada consiste en que TMLUC requiere visualizar los indicadores Instock y DOH desagregados a nivel Producto y Sala, con la posibilidad de filtrar y agregar dicha información según las jerarquías de Fecha, Región/SubCadena y Categoría/Línea Promocional, permitiendo que el diagnóstico operativo diario como el análisis estratégico de tendencias descritos en la Sección 9.3.2.

Esta necesidad, traducida a requerimientos técnicos, delimita de manera estricta el grano de dato, la cobertura y el origen de la información que alimentará la solución, según se detalla a continuación.

Criterio de Granularidad Atómica. Bajo esta directriz de ingeniería de datos, se definieron granularidades diferenciadas según la naturaleza del dato entregado por Walmart. Para los flujos de Ventas POS y Stock MultiEstado se estableció el grano atómico Producto-Sala-Día, siguiendo el criterio de Kimball y Ross (2002) postulan que el diseño de cualquier arquitectura dimensional debe priorizar obligatoriamente el nivel más bajo y detallado de los datos atómicos capturados por el proceso de negocio, dado que esta condición hiperdimensional ofrece la máxima flexibilidad analítica para responder a consultas actuales y futuras de la organización (pp. 640-641). Esta decisión responde directamente a la Ceguera Operativa Longitudinal diagnosticada en la Sección 9.3.2. Al registrar venta y stock por cada SKU y sala de forma diaria, el modelo permite

reconstruir la disponibilidad en góndola día a día y trazar la trayectoria histórica de demanda, sin el sesgo que introduciría una agregación previa a nivel de sucursal o cadena, y habilitando precisamente la desagregación por Producto y Sala que la necesidad de negocio exige.

Universo de Datos y Escalabilidad. El volumen de información que el sistema procesa diariamente está determinado por la amplitud operacional de la cuenta Walmart, la combinación de 100 SKUs promedio de TMLUC por sala genera aproximadamente 35.000 registros diarios de venta e inventario, más 7.000 registros semanales de forecast, la misma volumetría que, en la Sección 9.2.3, se demostró que satura el límite de Excel en menos de un mes. Por eso la arquitectura se diseñó bajo un criterio de escalabilidad modular, ya que como sostiene Gorelik (2019), los sistemas modernos de almacenamiento deben estructurarse de manera modular para absorber incrementos exponenciales en el volumen de datos sin degradar los tiempos de respuesta de los reportes estratégicos, garantizando que el almacenamiento soporte series de tiempo de largo plazo (p. 45). La capacidad de soportar este flujo continuo es lo que permite superar el techo técnico del software de oficina actual, el cual colapsa físicamente al verse expuesto a la acumulación de transacciones de 3 días de operación, garantizando además que la organización pueda sostener el filtrado por jerarquía de fecha (día, semana, trimestre), que la necesidad de negocio exige, sin limitaciones de historia que hoy impone el software de oficina.

Criterio de Carga Total. Metodológicamente, se descartó el uso de muestro probabilístico en la ingesta del modelo analítico, optando por un enfoque de carga total, el universo capturado equivale al 100% de las transacciones generadas por la cuenta Walmart, sin selección estadística de subconjuntos. Esta decisión resuelve directamente el riesgo señalado en el diagnóstico, omitir aleatoriamente el registro de stock de una sola sala podría ocultar un quiebre crítico, distorsionando el cálculo de Instock y DOH, los dos indicadores dentro del alcance definido en la Sección 8.1.2

y alimentando el Efecto Látigo por falta visibilidad real. La carga total es, por tanto, la única vía que garantiza que ambos indicadores se calculen sobre la realidad operativa íntegra del punto de venta, y no sobre una muestra que podría enmascarar quiebres puntuales, asegurando que la visualización por Producto y Sala que exige la necesidad de negocio esté disponible sobre el 100% de las combinaciones, y no únicamente sobre una muestra que podría dejar fuera justamente los quiebres que se busca detectar.

Catálogo de Fuente de Datos. Siguiendo el criterio de catalogación exhaustiva de fuentes definido en la Metodología (Sección 8.3.1), se mapeó cada activo de información B2B según su origen físico, su tipología de dato y su rol dentro del pipeline, antes de definir cualquier lógica de transformación.

Bajo este rigor, la solución define un mecanismo de ingesta híbrido o dual que combina la extracción programada vía scripts/APIs con la automatización por robótica de software (RPA) para responder a la naturaleza de cada fuente.

El primer activo de información corresponde al reporte diario de ventas, el cual registra las transacciones de punto de venta (POS) agregadas bajo la granularidad Producto-Sala-Día. Esta fuente provee campos numéricos esenciales tales como las unidades físicas comercializadas y el valor monetario de las ventas a precio costo, actuando como el indicador base de la demanda real del consumidor.

El segundo activo de información lo constituye el reporte diario de inventario MultiEstado, el cual describe la posición física del stock bajo el mismo grano del punto de venta. Este archivo captura las variables de inventario disponible en góndola (On Hand), stock en tránsito desde los centros de distribución (In Transit), stock en apartado en la CD para una sala específica (In Whse) y pedidos confirmados en firme por el retailer (On Order).

La extracción de las Ventas POS y Stock MultiEstado se realiza mediante la ingesta automática mediante scripts de Python que consumen de forma periódica la API que proporciona la plataforma Teamcore, extrayendo los registros diarios a nivel Producto-Sala-Día.

Finalmente, el tercer activo de información corresponde al reporte semanal de forecast, cuyo grano se sitúa en Producto-Sala-Semana de acuerdo con los ciclos de planificación y reabastecimiento programados por el algoritmo de Walmart. Esta distinción en la frecuencia temporal del forecast exige un tratamiento de ingeniería de datos diferenciado para asegurar su coexistencia analítica con las fuentes diarias de ventas e inventarios sin corromper la veracidad de los reportes integrados. La ingesta de esta información se realiza mediante un robot de RPA en Power Automate Desktop. El robot navega de forma desatendida por el portal Retail Link, descarga el archivo plano y lo deposita en la zona de aterrizaje.

Fuente Técnica	Atributos y Hechos	Tipo de Dato	Justificación en la Arquitectura
Ventas POS	Unidades Sell-Out diarias por SKU y Sala.	Estructurado	Mide el ritmo de salida de inventario del cliente final y alimenta el cálculo de demanda real.
Stock MultiEstado	On Hand, In Transit, In Warehouse.	Estructurado	Define la posición logística y permite detectar asimetrías en el flujo.
Forecast	Demanda proyectada (algoritmo	Estructurado	Actúa como el denominador lógico para las alertas de quiebre de stock.

	Walmart).		

9.6.2 Nivel 2 del Modelo en V: Especificación Funcional y Diseño Dimensional

Según lo planteado por Forsberg y Mooz (1991) una vez que se han delimitado los requerimientos del negocio (Nivel 1), el Modelo en V exige descender al nivel de especificación funcional, el cual define que funciones analíticas y reglas de relación estructurarán el sistema (p. 8). Conforme a las directrices de modelamiento dimensional de Kimball y Ross (2002), la base de datos no debe diseñarse bajo un esquema transaccional normalizado de cara al usuario final, sino organizarse en un modelo en estrella que optimice el rendimiento de las consultas analíticas mediante la separación estricta de tablas de hechos y dimensiones conformadas (p. 55).

Definición del Catálogo de Dimensiones Maestras. Para garantizar que el modelo integrado fuera consistente y libre de asimetrías de información, se diseñó una arquitectura dimensional basada en entidades compartidas. Siguiendo el concepto de Kimball y Ross (2002) que introducen el concepto de dimensiones conformadas, catalogándolas como tablas de referencia estandarizadas y reutilizables que contienen llaves naturales idénticas y atributos como una semántica unificada para toda la organización, lo que habilita la ejecución de consultas cruzadas consistentes sobre múltiples procesos de negocio (p. 77).

Metodológicamente, se definió el diseño lógico de cuatro dimensiones compartidas: la Dimensión Producto (que unificará los códigos del proveedor con los del retailer), la Dimensión Sala (que estructurará la jerarquía geográfica de las sucursales), la Dimensión Fecha (que establecerá un calendario diario alineado al calendario operativo de Walmart) y la Dimensión Semana.

- Dimensión Producto: Se diseña a partir de la integración de los maestros de productos

de TMLUC con los registros de producto en Retail Link. Se incluyen los atributos de SKU (llave natural de producto), Descripción Producto, Unidad de Negocio, Categoría, Línea, Sub-línea, Línea promocional, Unidades por Caja, Cajas x Pallet.

- Dimensión Sala: Dado que Walmart entrega múltiples atributos de tienda, esta tabla contendrá el Código Sala, Nombre Sala, Región, Comuna, Cadena, Sub-Cadena.
- Dimensión Fecha: Tabla generada internamente a nivel grano diario, que descompone la fecha en atributos de valor estratégico, Fecha, Trimestre, Día de la Semana, Semana del Año, Semana según Calendario Walmart.
- Dimensión Semana: Tabla de nivel agregado, conformado como subconjunto estricto de la Dimensión Fecha, que contiene un registro por cada Semana según Calendario de Walmart, con los atributos Semana Walmart (llave), Semana del Año, Mes, Trimestre. Esta dimensión conformada con la Dimensión Fecha, en donde sus atributos comunes, Semana del Año, Mes, Trimestre, están etiquetados, definidos y valorados de forma idéntica en ambas tablas.

Definición de las Tablas de Hechos Separadas. Según Kimball y Ross (2002), los flujos operacionales con distinta granularidad o frecuencia temporal no deben forzarse a coexistir en una única tabla de hechos: cada proceso debe modelarse en su propia tabla, relacionada con las demás exclusivamente a través de dimensiones conformadas, para resguardar la consistencia lógica de las métricas (p. 12). En consistencia con este principio, el modelo se separó en tres tablas de hechos desacopladas: una para Ventas POS (grano Producto-Sala-Día), una para Stock MultiEstado (grano Producto-Sala-Día) y una para Forecast (grano Producto-Sala-Semana).

- Tabla Hechos de Ventas: Fecha, SKU, Código Sala, Ventas (unidades), Venta ($), Costo de Venta ($).
- Tabla Hechos de Stock: Fecha, SKU, Código Sala, Curr/Valid, On Hand, In Transit, In

Whse, On Order. Siendo Curr/Valid 1 si la combinación producto sala se encuentra catalogada.

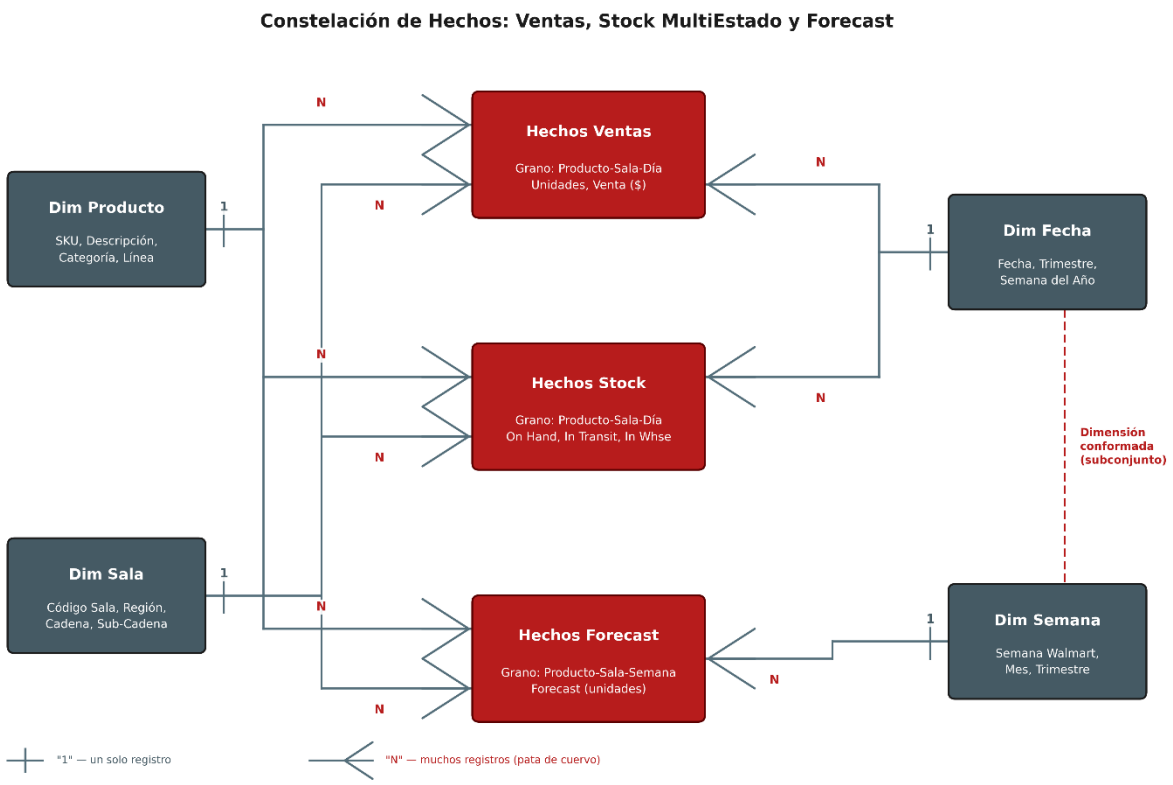
- Tabla Hechos de Forecast: Semana Walmart (llave foránea), SKU, Código Sala, Forecast (unidades).

Para resolver la diferencia de granularidad temporal del forecast sin corromper la integridad del grano, se diseña una Constelación de Hechos (Fact Constellation) donde Ventas y Stock se acoplarán al nivel diario de la Dimensión Tiempo, mientras que Forecast se vinculará a la jerarquía de la Dimensión Semana, permitiendo agregaciones matemáticas automáticas sin manipulación manual del dato por parte del usuario.

La Figura 3 sintetiza la arquitectura dimensional descrita en esta sección, mostrando las tres tablas de hechos desacopladas (Ventas, Stock MultiEstado y Forecast), y su relación con las cuatro dimensiones a través de llaves conformadas, incluyendo la relación de subconjunto entre las dimensiones Fecha y Semana que permite la coexistencia de ambas granularidades temporales sin corromper la integridad del grano.

Figura 3

Constelación de Hechos del modelo de integración B2B TMLUC-Walmart



Nota. Elaboración propia.

9.6.3 Nivel 3 del Modelo en V: Arquitectura del Almacén de Datos (Capa Lógica)

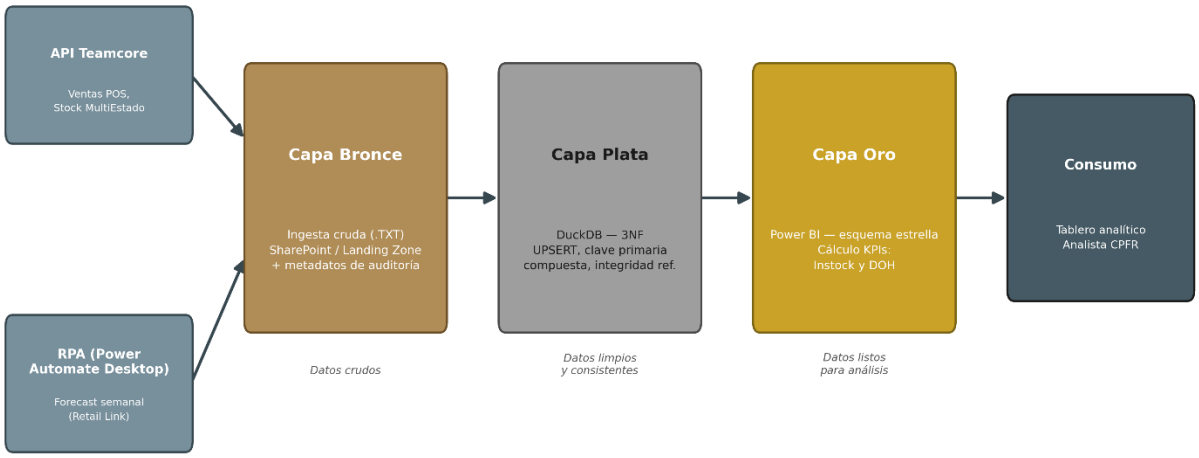
Siguiendo las dos categorías de reglas definidas en la Metodología (Sección 8.3.3), integridad técnica y lógica operativa, cada etapa del Medallón aplicó un subconjunto específico de ellas, como se detalla a continuación.

Conforme a la equivalencia entre capas y áreas de staging/presentación establecida en la Sección 6.5.4, las capas Bronce y Plata se implementaron como área de preparación y la capa Oro como área de presentación final.

Antes de detallar cada nivel lógico, la Figura 4 presenta una visión general del flujo de datos propuesto, desde las dos fuentes de ingesta (la API de Teamcore y el robot RPA sobre Retail Link), hasta el tablero de consumo final, pasando por las tres capas de refinamiento progresivo (Bronce, Plata y Oro) que se detallan a continuación.

Figura 4

Arquitectura de Medallón de la solución implementada



Nota. Elaboración Propia

Capa Bronce (Data Lake), Ingestión y Preservación del Dato Nativo. En esta primera etapa, la prioridad técnica radica en la captura masiva y el resguardo de los activos de información provenientes de Retail Link en su formato nativo (.TXT). Como sostiene Gorelik (2019), desacoplar la ingesta de las transformaciones complejas asegura la máxima flexibilidad operativa ante eventuales cambios en las estructuras de los reportes emitidos por el portal B2B (p. 84).

Bajo esta directriz la capa Bronce opera como una zona de aterrizaje (Landing Zone) desacoplada localizada en una biblioteca exclusiva para esta etapa. En esta capa coexistirán los dos flujos de ingesta en su estado nativo:

- El script de Python extrae las ventas y stocks mediante la API de Teamcore y los almacena en bruto dentro de SharePoint.
- En paralelo, el robot de RPA (Power Automate Desktop) extrae el reporte plano de Forecast desde Retail Link y lo deposita inmutable en la misma ubicación.

Procedimentalmente, ambas fuentes se enriquecerán en la Capa Bronce añadiendo de forma automática tres metadatos de auditoría técnica, la marca de tiempo de carga, el nombre de la fuente de origen y el identificador del lote, garantizando la trazabilidad completa del dato nativo.

Capa Plata (Data Warehouse), Integración, Limpieza y Control de Unicidad. En la Capa de Plata se ejecuta la consolidación de la información bajo una estructura relacional en Tercera Forma Normal (3NF). En esta etapa se resuelven las inconsistencias de formato y se garantiza la integridad referencial entre la data transaccional de Walmart y los maestros internos de TMLUC.

- Estandarización y Limpieza Técnica: Se realiza una conversión explícita del tipo de datos y se tratan los valores nulos o vacíos mediante una imputación condicional para asignar un valor específico o eliminación del registro.
- Control de la Unicidad (Clave Primaria Compuesta): Para evitar la duplicidad por solapamiento de cargas diarias, el motor ETL aplicará una restricción de Clave Primaria Compuesta sobre cada una de las tablas de Hechos. Para las tablas de Ventas y Stock la Clave Primaria Compuesta es SKU-Sala-Día y para la tabla de Forecast la Clave Primaria Compuesta es SKU-Sala-Semana.
- Reglas de Integridad Técnica: Si se detecta un registro entrante cuyo valor de clave compuesta coincide con un registro previo en la base de datos, el motor ETL aplicará una estrategia de actualización (UPSERT) basada en la marca de tiempo más reciente, impidiendo la duplicidad de registros y asegurando la consistencia cuantitativa.
- Integración y Enlace Dimensional (Integridad Referencial): Se realiza el cruce entre los códigos de producto y salas Walmart con los catálogos maestros de TMLUC. Aquellos registros cuya llave natural no exista en las tablas maestras se redirigen automáticamente a una

tabla de excepciones para su auditoria, garantizando que el Data Warehouse mantenga una consistencia relacional absoluta.

Capa Oro (Data Marts), Inteligencia de Negocios y Métricas Operacionales. La capa Oro representa el área de presentación optimizada bajo un modelado dimensional (Esquema Estrella). En este nivel, utilizando las herramientas de Power Query y DAX de Power BI se aplican las reglas de negocio avanzadas y los algoritmos matemáticos para el cálculo automatizada de los Indicadores Claves de Desempeño (KPIs) sobre el grano atómico SKU-Sala-Día.

- Cálculo Automatizado de Días de Inventario (DOH): Para medir el nivel de cobertura del inventario físico respecto al pronóstico de demanda y detectar sobre stocks inmovilizados o riesgos de desabastecimiento, se establece la siguiente relación matemática.

$$DOH_{ps} = \frac{OH_{ps}}{FD_{ps}}$$

- Cálculo de Instock Operacional Binario: Para automatizar la detección de quiebres en la góndola de forma proactiva, se establece un algoritmo lógico binario que determina el estado de disponibilidad del producto en el punto de venta.

$$I_{ps} = \begin{cases} 1, & si\ OH_{ps} \geq FD_{ps} \\ 0, & si\ OH_{ps} < FD_{ps} \end{cases}$$

Si el Stock On Hand es mayor o igual al Forecast semanal dividido en 7, el indicador será 1 (Disponible), de lo contrario será 0 (Quiebre).

- Reglas de Agregación y Optimización Dimensional: En la Capa Oro, las tablas de hechos desnormalizadas calculan resúmenes ejecutivos a nivel de familia de productos, categoría, región y sub-cadena. Esta estructura garantiza que las consultas analíticas ejecutadas desde los tableros de control entreguen tiempos de respuesta en milisegundos, democratizando el consumo ágil de información para la toma de decisiones estratégicas dentro de la organización.

9.6.4 Nivel 4 del Modelo en V: Especificación del Detalle Tecnológico

El último nivel de menor abstracción en el descenso por el Modelo en V, según Forsberg y Mooz (1991) corresponde al diseño de detalle, donde se definen los parámetros físicos por las herramientas y se seleccionan las tecnologías específicas que materializan cada módulo funcional del sistema (p. 9). La selección de herramientas no se realiza por preferencia comercial, sino evaluando que sus características técnicas respondan con precisión a las restricciones del entorno empresarial estudiado.

Selección de la Tecnología de Ingesta y Almacenamiento – Capa Bronce. La Capa Bronce, requiere un mecanismo que reemplace la extracción manual de archivos desde el portal Retail Link, preservando el principio de no intervención en la fuente. El diseño de este mecanismo se estructura en dos etapas diferenciadas.

- Generación Programada del Reporte: El módulo DSS de Retail Link permite configurar de forma nativa la generación periódica de los reportes de Ventas, Stock y Forecast, los cuales quedan disponibles en una biblioteca de documentos dentro del propio portal, sin requerir que un usuario los solicite manualmente en cada ocasión. Esta funcionalidad es provista directamente por la plataforma del retailer y replica la misma cadencia de actualización que hoy sostiene el proceso manual.

- Segunda Etapa: Recolección Automatizada: Dado que los reportes ya se encuentran generados y disponibles al momento de la ejecución programada, la automatización no necesita simular el proceso de solicitud del reporte, autenticación, navegación por menús, configuración de parámetros, sino únicamente verificar la biblioteca de documentos y descargar los archivos nuevos. Esta naturaleza acotada de la interacción corresponde al concepto de Automatización Robótica de Procesos (RPA), como un conjunto de herramientas que operan sobre la interfaz de

usuario de otros sistemas computacionales de la misma manera en que lo haría una persona, replicando una tarea estructurada y repetitiva sin alterar el sistema subyacente.

Para implementar la segunda etapa se selecciona Power Automate Desktop como herramienta de automatización. Esta selección responde al mismo criterio de restricción tecnológica aplicado en el resto de la arquitectura, ya que al formar parte del conjunto de licencias de Microsoft ya contratadas por TMLUC, su adopción no representa un requerimiento de infraestructura, costo ni aprobación de TI adicional para la organización, a diferencia de RPA de terceros que requerirán una licencia independiente.

Por su parte, la ingesta de Ventas POS y Stock MultiEstado se automatiza mediante scripts de Python que leen los datos provistos por Teamcore mediante la API disponible, logrando un equilibrio eficiente que optimiza los tiempos de ejecución del sistema.

Los archivos recolectados se depositan en su formato nativo, sin transformación alguna, en el entorno corporativo de SharePoint, preservando el principio de conservación del dato crudo característico de la Capa Bronce.

Cabe precisar que este diseño no vulnera el criterio de no intervención en la fuente, la automatización se limita a una acción de lectura sobre una biblioteca de archivos ya generados por la propia plataforma, sin modificar, alterar ni interactuar con los sistemas transaccionales de Walmart más allá de lo que la funcionalidad de programación de reportes, provista nativamente por Retail Link ya contempla para sus proveedores.

Selección de la Tecnología de Persistencia – Capa Plata. El diseño lógico de la Capa Plata (descrito en la Sección 9.6.3), establece que esta capa debe sostener una estructura relacional normalizada, con control de unicidad mediante clave primaria compuesta, integridad referencial contra los maestros de la organización y operaciones de actualización incremental (UPSERT).

Cumplir estos requisitos exige, en principio, un motor de base de datos relacional. Sin embargo, la selección de la tecnología que materializa este diseño no puede realizarse de forma aislada de las condiciones reales de TMLUC, por lo que se contrasta el escenario ideal frente al escenario factible dentro de la organización.

El escenario real descartado consiste en la alternativa natural para implementar la Capa Plata, lo cual es un motor de base de datos relacional gestionado en la nube, como Azure SQL Database, que ofrece de forma nativa las garantías de integridad, concurrencia y administración requeridas. No obstante, las políticas de gobierno de tecnología de TMLUC restringen el aprovisionamiento de infraestructura en la nube fuera del ecosistema Microsoft ya contratado por la organización (Sección 7.4), lo que excluye la implementación de este servicio sin un proceso de aprobación de TI corporativo fuera del alcance del presente proyecto.

El escenario factible seleccionado, consiste en seleccionar DuckDB como sistema gestor de bases de datos relacional embebido, ya que según Raasveldt y Mühleisen (2019) está diseñado específicamente para ejecutar consultas analíticas complejas sin requerir un proceso de servidor independiente ni protocolo de comunicación cliente-servidor (pp. 1981- 1982). A diferencia de un motor de base de datos tradicional, DuckDB se ejecuta como una librería dentro del proceso de script de Python que orquesta el pipeline, lo que permite aplicar restricciones de clave primaria, validaciones de integridad referencial y lógica de actualización incremental directamente sobre archivos en formato Parquet, sin necesidad de administrar infraestructura de servidor dedicada.

La siguiente tabla resume la comparación entre ambos escenarios:

Criterio	**Escenario Ideal (Azure SQL Database)**	**Escenario Factible (DuckDB + Parquet)**
Cumplimiento de política	No cumple, requiere	Cumple, se ejecuta como librería

de TI	aprobación de infraestructura cloud externa.	dentro del entorno ya autorizado
Costo de licenciamiento	Costo recurrente de servicio gestionado.	Sin costo, software de código abierto.
Administración requerida	Requiere administración de servidor y credenciales.	No requiere servidor ni proceso administrado.
Garantías de integridad (3NF, UPSERT, clave primaria)	Nativas del motor relacional gestionado.	Equivalentes, ejecutadas mediante consultas SQL embebidas sobre archivos Parquet.
Persistencia física de los datos	Interna, gestionada por el servicio en la nube.	Formato Parquet particionado, almacenado en SharePoint.

Se concluye que DuckDB permite satisfacer los requisitos de integridad dentro de las condiciones reales del entorno de TMLUC, constituyendo una solución técnicamente equivalente al escenario ideal en cuanto a garantías de modelamiento relacional, aunque con las limitaciones de concurrencia de escritura propias de un motor embebido.

Selección de la Tecnología de Modelamiento Dimensional – Capa Oro. El diseño lógico de la Capa Oro, descrito en la Sección 9.6.3, establece un esquema dimensional de tipo estrella, en el cual las tablas de hechos de Ventas, Stock y Forecast se relacionan con las dimensiones maestras de SKU y Sala para el cálculo de los indicadores definidos en la Sección 6.3.

Para materializar este modelo se selecciona Power BI, que permite construir el esquema estrella mediante relaciones definidas en Power Query y calcular los indicadores mediante

expresiones DAX sobre un modelo de datos comprimido en memoria. La selección de esta herramienta, por sobre otras plataformas de Business Intelligence disponibles en el mercado, responde al mismo criterio de restricción tecnológica aplicado en las capas anteriores, ya que Power BI forma parte del conjunto de licencias Microsoft ya contratadas por TMLUC, por lo que su adopción no representa un requerimiento de infraestructura ni de licenciamiento adicional para la organización.

Es necesario precisar el mecanismo mediante el cual los archivos Parquet procesados localmente por DuckDB llegan a estar disponibles para su consumo por parte de Power BI en la nube. Dado que DuckDB opera como un motor embebido dentro del proceso local de Python y no dispone de un componente de red propio, la sincronización se resuelve mediante el cliente de sincronización de SharePoint instalado en el equipo o servidor donde se ejecuta el script de automatización. Este cliente mantiene una réplica local de la biblioteca de documentos de SharePoint, el script de transformación escribe los archivos Parquet resultantes directamente dentro de esa carpeta local sincronizada, y el cliente se encarga de propagar automáticamente los cambios hacia el entorno en la nube, sin que el pipeline de Python requiera implementar lógica de red o autenticación adicional para este propósito. De esta forma, el cliente de sincronización actúa como el puente de comunicación automática entre el motor de procesamiento local (DuckDB) y el servicio de Business Intelligence en la nube (Power BI).

En conjunto, los niveles 1 a 4 descritos dan cumplimiento al Objetivo Específico 2 (estándares y arquitectura lógica, Niveles 2 y 3) y al Objetivo Específico 3 (modelo ETL bajo código, Nivel 1 y 4), mientras que el objetivo específico 4 se aborda en el resultado del tablero de control presentado a continuación.

9.6.5 Resultado del Tablero de Control Implementado

Como resultado de la integración de las cuatro capas descritas anteriormente, se implementó el tablero de control en Power BI que constituye el punto de consumo final de la arquitectura propuesta, dando cumplimiento al Objetivo Específico 4. La Figura 5 presenta la visibilidad principal del tablero, sobre la cual se identifican cinco elementos que evidencian la resolución de las problemáticas diagnosticadas en la Sección 9.3.

Figura 5

Tablero de control implementado en Power BI



Nota. Captura de pantalla del sistema implementado. Los valores numéricos corresponden a una base ficticia por motivos de confidencialidad comercial.

El elemento ① corresponde a las tarjetas de indicadores Instock y DOH, calculadas automáticamente sobre la fecha seleccionada, eliminando el cálculo manual descrito en la Sección 9.1.2. El elemento ② muestra el panel de segmentadores (Fecha, Región, Categoría y Línea Promocional), que permite al Analista CPFR navegar la información de forma autónoma sin

requerir la intervención de un especialista técnico. El elemento ③ corresponde a la tabla de Instock por producto con formato condicional tipo semáforo, que identifica automáticamente los quiebres de stock mediante codificación visual, automatizando el proceso de detección que hoy se ejecuta manualmente. El elemento ④ presenta la serie histórica de Instock y On Hand a lo largo del tiempo, resolviendo directamente la Ceguera Operativa Longitudinal diagnosticada en la Sección 9.3.2. Finalmente, el elemento ⑤ corresponde a la tabla de Instock por Sala, que permite identificar de forma agregada el desempeño de cada punto de venta de la cuenta Walmart.

9.6.6 Resultado del Protocolo de Verificación y Pruebas

Con el propósito de validar la solución implementada, se ejecutaron las cuatro compuertas de verificación definidas en la Sección 8.4, sobre las cuales se reportan a continuación los hallazgos obtenidos.

Compuerta 1: Pruebas Unitarias. En correspondencia con el Diseño de Detalle (Sección 9.6.4), esta compuerta auditó el funcionamiento aislado de cada módulo del sistema. Por un lado, el script de ingesta desarrollado en Python para consumir la API de la plataforma Teamcore operó correctamente desde las primeras ejecuciones, sin registrar incidentes de autenticación ni pérdida de paquetes en la recopilación de los registros de Ventas POS y Stock MultiEstado. Por otro lado, el robot de extracción desarrollado en Power Automate Desktop (RPA) presentó fallas en sus primeras ocho iteraciones de prueba debido a la variabilidad en los tiempos de respuesta del portal externo Retail Link de Walmart, los cuales superaban los umbrales de espera fijos calibrados inicialmente para la detección de elementos de la interfaz. Dado que esta latencia exógena no puede eliminarse en la fuente, la corrección se orientó a reestructurar los mecanismos de espera del robot hacia una lógica condicional basada en la confirmación de carga explícita de los elementos web. Tras este ajuste, el robot alcanzó una ejecución desatendida consistente y libre de

errores en las corridas posteriores. Finalmente, las pruebas unitarias sobre las funciones SQL en el motor DuckDB y sobre los objetos visuales y medidas en lenguaje DAX en Power BI se completaron de manera satisfactoria.

Compuerta 2: Pruebas de Integración. En correspondencia con el Diseño de la Arquitectura (Sección 9.6.3), esta etapa evaluó la integración a lo largo de las tres capas de la Arquitectura de Medallón (Bronce, Plata y Oro) y SharePoint. En este nivel se validó con éxito el mecanismo de control de unicidad mediante la operación UPSERT ejecutada sobre la clave primaria compuesta (SKU-Sala-Fecha), comprobando que el motor absorba cargas duplicadas sin corromper la base histórica. Asimismo, se auditó la rutina de integridad referencial para el tratamiento de combinaciones Producto-Sala no catalogadas previamente en las dimensiones maestras (originadas por la habilitación de nuevas salas o catalogación de productos por parte de Walmart), verificando que dichos registros se redirijan automáticamente a la tabla lógica de excepciones para su posterior auditoría, en lugar de insertarse o descartarse de forma silenciosa. No se registraron fallas en las pruebas de integración a lo largo del pipeline.

Compuerta 3: Pruebas del Sistema (Nivel de Especificación Funcional). En correspondencia con la Especificación Funcional (Sección 9.6.2), se validó la respuesta del modelo en Power BI ante escenarios de consulta complejos. El pipeline completo de extracción, transformación y carga (ETL) a través de las capas Bronce, Plata y Oro demostró un tiempo total de actualización de 5 minutos, mientras que las consultas analíticas sobre el modelo dimensional ya cargado respondieron en escala de milisegundos, ratificando el diseño de rendimiento diferenciado (la latencia del refresco de datos reside en el proceso de integración y no en la capa de consumo analítico). Adicionalmente, se confirmó que al cruzar las tablas de hechos diarias de Ventas y Stock con la tabla de hechos semanal de Forecast mediante las dimensiones conformadas

de Fecha y Semana, no se produjo inflación ni duplicación artificial de métricas, validando que la separación física de las tablas de hechos conectadas por dimensiones conformadas evita los errores de duplicidad.

Compuerta 4. Pruebas de Aceptación del Usuario (UAT). En correspondencia con el Análisis de Requisitos del Negocio (Sección 9.6.1)**,** el Analista CPFR llevó a cabo un proceso de reconciliación de datos sobre la población completa (no muestral) durante cinco días consecutivos, contrastando la totalidad de las combinaciones Producto-Sala generadas por la Arquitectura de Medallón contra las descargas crudas nativas del portal Retail Link. Los resultados mostraron una coincidencia matemática exacta del 100% sin registrar diferencias en ninguno de los días evaluados.

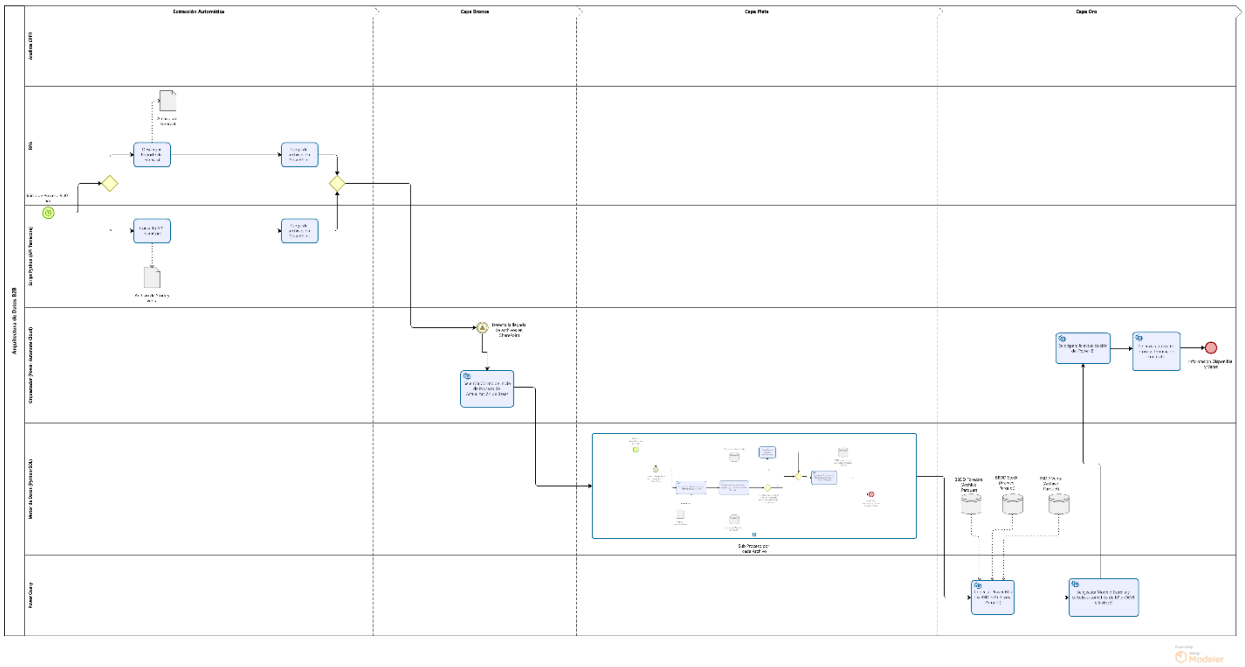
La compuerta 4 constituye el hito técnico decisivo de esta verificación, al certificar directamente el objetivo central del proyecto, la coincidencia exacta y sostenida en el tiempo entre el modelo dimensional y la fuente nativa de Walmart demuestra que TMLUC dispone de una plataforma que opera bajo la única versión de la verdad definida por Cano (2018), eliminando la necesidad de que Analista CPFR reconcilie manualmente la información como ocurría bajo el proceso descrito en la Sección 6.1.

9.6.7 Síntesis del Proceso B2B Rediseñado

A modo de síntesis del resultado global de esta investigación, la Figura 6 presenta el proceso B2B rediseñado de extremo a extremo, permitiendo contrastar directamente su estructura con el proceso manual diagnosticado en la Figura 1 (Sección 9.1.6).

Figura 6

Proceso B2B rediseñado – Arquitectura de Datos B2B TMLUC-Walmart



Nota. Elaboración propia.

9.7 Evaluación Económica de la Solución

Con el propósito de contrastar el costo de la solución implementada con el costo de la ineficiencia diagnosticada en la Sección 9.2, se aplica el procedimiento de evaluación económica definido en la Sección 8.5. Los valores corresponden a pesos chilenos 2026, en un horizonte de 12 meses sin descontar.

9.7.1 Costo Total de Propiedad (TCO)

El TCO de la solución considera la inversión de desarrollo y los costos de operación de primer año, conforme al criterio de Cano (2018), que exige incorporar el tiempo del personal interno y no solo la adquisición de tecnología. Este horizonte corresponde a la vida útil esperada de la arquitectura antes de que el crecimiento del volumen de datos haga necesario su migración. Siguiendo los criterios de evaluación privada de proyectos, los recursos se valorizan a su costo de

oportunidad, es decir, al precio que la organización debería pagar por ellos en el mercado (Sapag Chain et al., 2014). Por ello, el desarrollo no se valoriza con la remuneración de memorista efectivamente pagada, sino con el costo de un profesional equivalente, lo que permite que la evaluación sea válida también ante una eventual réplica de la solución en otras cuentas del canal moderno.

Componente	**Cálculo**	**Monto (CLP)**
Inversión: diseño, desarrollo y pruebas	4 meses × $1.719.151	$6.876.604
Licenciamiento	Power Automate, Power BI y SharePoint incluidos en el ecosistema Microsoft ya contratado; DuckDB y Python de código abierto	$0
Infraestructura	Equipo ya existente en la organización	$0
Operación y mantenimiento	4 h/mes × $11.000 x 36 meses	$1.584.000
TCO a 36 meses		**$8.460.604**

El costo mensual de desarrollo se fijó en $1.719.151, correspondiente al ingreso bruto promedio al primer año de titulación que el portal Mi Futuro del Ministerio de Educación reporta para la carrera de Ingeniería Civil Industrial (Meganoticias, 2024). Se utiliza este valor por provenir de una fuente oficial y por representar el costo de un profesional con la formación requerida para desarrollar la solución.

El mantenimiento se estimó en 4 horas mensuales, distribuidas en tres tareas. La primera

es la revisión semanal de las cargas y de la tabla de excepciones (1 hora). La segunda es el ajuste del robot RPA ante cambios o latencia del portal Retail Link (2 horas), componente que requirió ocho iteraciones de ajuste antes de estabilizarse (Sección 9.6.6). La tercera es la actualización de los maestros de salas y productos (1 hora).

El costo incremental de licencias e infraestructura es nulo como consecuencia directa de la restricción tecnológica adoptada en la Sección 9.6.4: haber optado por el escenario ideal (Azure SQL Database) habría agregado un costo recurrente de servicio gestionado.

9.7.2 Beneficio Económico

El beneficio corresponde a la capacidad del Analista CPFR liberada al eliminar las actividades sin valor agregada del proceso manual. Se valoriza con la Fuga de Valor de Ingeniería diagnosticada en la Sección 9.2.1, equivalente a 373 minutos semanales sin valor agregado y a $272.000 mensuales.

Cabe precisar que este beneficio no corresponde a un ahorro de caja, ya que el Analista CPFR permanece en la organización. Se trata de capacidad liberada que se reasigna desde la preparación manual de datos hacia el análisis estratégico del abastecimiento, valorizada a su costo de oportunidad.

9.7.3 Indicadores de Rentabilidad

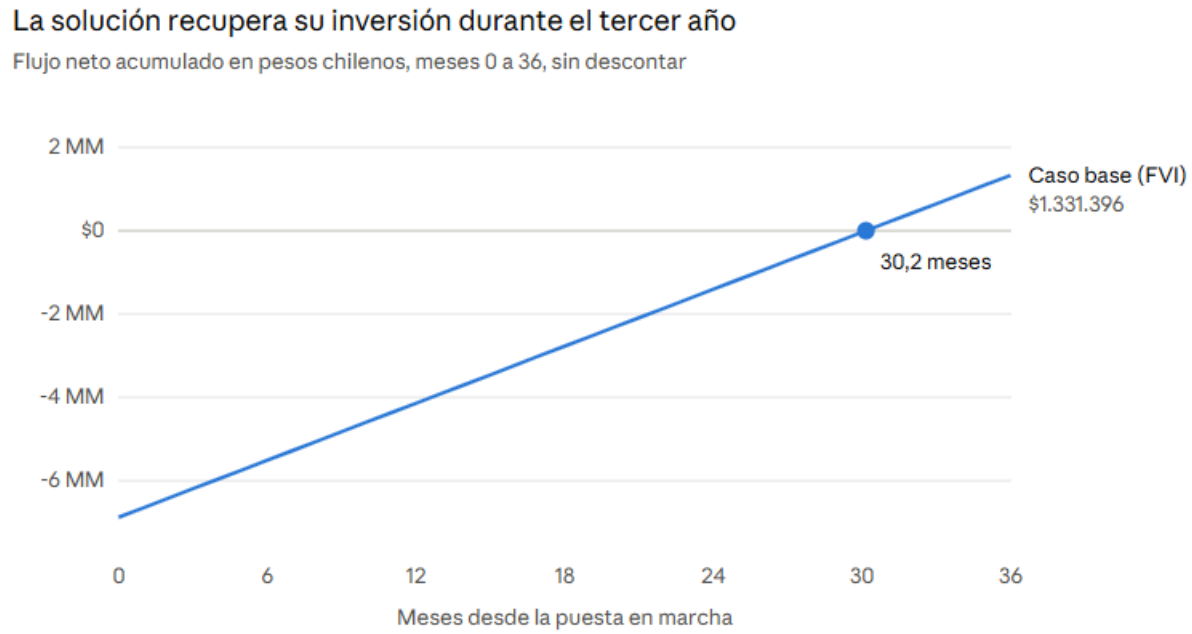
Indicador	**Valor**
Beneficio mensual	$272.000
Costo de operación mensual	$44.000
Beneficio neto mensual	$228.000
Período de recuperación (PRI)	31 meses
Beneficio acumulado a 36 meses	$9.792.000

TCO a 36 meses	$8.460.604
Beneficio neto a 36 meses	$1.331.396
ROI a 36 meses	16%

La inversión se recupera en 31 meses, es decir, durante el tercer año de operación, y el beneficio neto acumulado alcanza $1.331.396, con un ROI de 16%. Este resultado se obtiene considerando únicamente la capacidad liberada del Analista CPFR y valorizando el desarrollo con el ingreso promedio de mercado reportado por una fuente oficial, por lo que constituye une escenario exigente para la solución.

Figura 7

Flujo neto acumulado de la solución en los primeros 12 meses



Nota. Elaboración propia. Flujos sin descontar, mes de 4 semanas.

9.7.4 Análisis de Sensibilidad

El beneficio por capacidad liberada representa solo una fracción de potencial económico de la solución. Durante 2025, TMLUC registró un Instock promedio de 95,84% y una venta perdida estimada de $4.883.357.593 (Sección 7.6.2). Bajo el supuesto de que la venta perdida es proporcional a la tasa de quiebre, cada centésima de punto porcentual de Instock equivale a aproximadamente $11,7 millones de venta anual. Para cubrir el costo anual promedio de la solución ($2.820.201, un tercio del TCO a 36 meses), bastaría con elevar el Instock promedio 0,0024 puntos porcentuales, lo que equivale a recuperar el 0,06% de la venta perdida anual.

Bajo el mismo supuesto de proporcionalidad, la tabla siguiente muestra la venta anual que se recuperaría ante distintas mejores del Instock promedio de 2025 (95,84%)

Mejora de Instock	**Venta recuperada estimada (anual)**
+0,0024 puntos porcentuales (umbral de equilibrio)	$2,8 millones
+0,01 puntos porcentuales	$11,7 millones
+0,1 punto porcentual	$117,4 millones
+0,5 puntos porcentuales	$586,9 millones
+1,0 punto porcentual	$1.173,9 millones

Este cálculo no constituye una medición del impacto de la solución, sino un umbral de equilibrio. Supone que la venta perdida es proporcional a la tasa de quiebre, está expresado en términos de venta ay no de margen, y no considera que el Instock depende de múltiples factores ajenos a la visibilidad del dato, como la producción, el despacho o la gestión del retailer. Adicionalmente, la detección oportuna de quiebres contribuye a reducir el riesgo de penalizaciones contractuales por incumplimiento de niveles de servicio, cuyo monto no fue cuantificado en este

estudio.

9.7.5 Síntesis de la Evaluación Económica

La evaluación demuestra que la solución es económicamente viable aun en un escenario exigente: valorizando su desarrollo con el ingreso promedio de mercado reportado por Mi Futuro y considerando únicamente la capacidad liberada del Analista CPFR, se recupera en 30,2 meses y genera un beneficio neto de $1.331.396 en 36 meses. Además, el umbral operativo necesario para cubrir el costo de la solución, una mejora del Instock de 0,0024 puntos porcentuales, es bajo frente a la magnitud del problema diagnosticado.

10 Conclusiones y Recomendaciones

10.1 Conclusiones

El patrón transversal identificado en el Análisis Causa Raíz permite sostener una primera idea central, la ineficiencia diagnosticada, que consumía el 77% del tiempo que el Analista CPFR le dedicaba a este proceso (Sección 9.2.1), nunca fue, en rigor, un problema de desempeño individual. Tres de los cinco problemas estructurales convergieron en una causa tecnológica compartida y los dos restantes en una causa de arquitectura de datos (Sección 9.4.6), y en ninguna de las cinco cadenas de 5 Porqué se sostuvo la categoría Mano de Obra como casusa raíz. Esta convergencia permite formular, con relativa confianza, un principio que excede el caso puntual de TMLUC: allí donde la clasificación de causa raíz apunta sistemáticamente a categorías de Máquina o Método por sobre Mano de Obra, la respuesta organizacional correcta es la inversión en arquitectura, no la intervención sobre el recurso humano.

Esta lectura se profundiza al contrastar dos hallazgos aparentemente distantes entre sí. Por un lado, la posición diagnosticada de TMLUC en la Etapa 1 del modelo de Eckerson, inmediatamente antes del Punto de Fricción Gulf (Sección 9.5.3), podría interpretarse en solitario como una debilidad organizacional. Por otro lado, el resultado de Nivel 4 del Modelo en V demostró una coincidencia exacta entre el modelo dimensional implementado y la fuente nativa de Walmart durante los cinco días consecutivos evaluados, sin una sola diferencia registrada (Sección 9.6.6, Compuerta 4). Leídos en conjunto, ambos hallazgos revelan que el verdadero aporte de esta memoria no es la resolución de un síntoma operativo puntual, sino la construcción del instrumento de ingeniería que permite a la organización cruzar deliberadamente un punto de fricción estructural que, de otro modo, habría permanecido como una limitación no abordada. La ausencia de diferencias en esa reconciliación de cinco días, a su vez, permite elevar el concepto de

única versión de la verdad formulado por Cano (2018), citado en el Marco Teórico como principio normativo a una condición empíricamente verificada, y no meramente aspiracional, dentro del sistema construido.

No todos los componentes de la solución alcanzaron el mismo grado de estabilidad inmediata, de lo anterior emerge una idea que matiza el resultado global. Mientras el componente de ingesta vía API operó sin incidentes desde su primera ejecución, el robot RPA requirió ocho iteraciones de ajuste antes de estabilizarse (Sección 8.6.5). Esta diferencia sugiere que la fragilidad relativa de un componente de automatización no depende únicamente de la calidad de su implementación, sino del grado de control que la organización ejerce sobre la interfaz de la que depende, dado que el portal Retail Link permanece fuera del control de TMLUC, esta dependencia estructural no queda resuelta por la implementación actual, sino que se traslada hacia una necesidad de gestión activa en el tiempo.

Finalmente, la capacidad evidenciada en el tablero de control implementado (Sección 9.6.5, Figura 5) para sostener series históricas, donde el Excel de Procesamiento no podía mantener más de 1 a 3 días de historia (Sección 9.2.3), permite concluir que el aporte de este proyecto no se agota en su dimensión técnica. Al liberar al Analista CPFR de algunas tareas manuales de consolidación que consumían 8 horas semanales, se habilita un desplazamiento de su función hacia el análisis estratégico del abastecimiento, que constituye la capacidad que la organización requería desarrollar desde un inicio, y que la arquitectura anterior impedía estructuralmente.

La evaluación económica (Sección 9.7) permite sostener que la solución no solo es técnicamente valida, sino también económicamente conveniente, aun valorizando su desarrollo con el ingreso promedio de mercado y sin considerar más beneficio que la capacidad liberada del Analista CPFR, la inversión se recupera en 31 meses y genera un ROI de 16% a 36 meses. Más

relevante aún, bastaría con una mejora del Instock de 0,0024 puntos porcentuales para cubrir su costo, lo que muestra que el valor económico de la solución reside principalmente en la disponibilidad en góndola. Esta asimetría refuerza la conclusión de que la inversión en arquitectura de datos es, ante todo, una inversión en nivel de servicio.

Por último, es importante destacar que la solución fue diseñada e implementada con capas desacopladas y dimensiones compartidas, lo que permite su escalabilidad tecnológica por etapas sin rediseñar su lógica, condición y característica que la distingue de una solución puntual.

10.2 Recomendaciones

Esta investigación definió explícitamente dos escenarios de almacenamiento en la Sección 9.6.4, uno ideal sobre Azure SQL, y uno factible, sobre DuckDB, y el Protocolo de Verificación confirmó que el escenario factible satisface adecuadamente los volúmenes actuales de la cuenta Walmart, respondiendo en milisegundos a las consultas de modelo dimensional (Sección 9.6.6, Compuerta 3). Dado que la elección del escenario factible respondió a una restricción de recursos y no a una limitación técnica del escenario ideal, se recomienda establecer indicadores de umbral, por ejemplo, volumen de registros o tiempo de actualización del pipeline, hoy en torno a los 5 minutos para el ciclo completo, que permitan anticipar el momento en que la migración hacia el escenario ideal resulte necesaria, evitando que la organización enfrente nuevamente la limitación tecnológica no planificada que motivó este proyecto.

Por otra parte, la arquitectura implementada deja disponible, sin requerir nuevas fuentes de información, una base para profundizar el análisis del Sell-Out. La tabla de hechos de Ventas, construida sobre grano atómico Producto-Sala-Día con carga total de las transacciones de la cuenta Walmart (Secciones 9.6.1 y 9.6.2) y sostenida como serie histórica, a diferencia de los 1 a 3 días que permitía el Excel de Procesamiento (Sección 9.2.3), permite analizar el comportamiento de

los productos y el efecto de las promociones sobre las ventas, análisis que el proceso manual impedía estructuralmente (Sección 9.3.2). Se recomienda, por tanto, explotar la información de venta ya integrada para analizar en profundidad el Sell-Out, incorporando el calendario de promociones de TMLUC como una nueva dimensión conformada que permita contrastar las ventas de períodos con y sin promoción. A diferencia de Fill Rate, esta línea de trabajo no depende de información adicional de Walmartm por lo que puede iniciarse con los datos ya disponibles, y su alcance dependerá de la extensión de la historia acumulada desde la puesta en marcha. El beneficio esperado es mejorar la anticipación de la demanda en períodos promocionales, aspecto ligado a la disponibilidad en góndola, y materializar el desplazamiento del Analista CPFR hacia el análisis estratégico del abastecimiento (Sección 10.1), en línea con la necesidad de visibilidad del Sell-Out que plantean Lee et al. (1997).

En la misma línea de proyección futura, la Sección 8.1.2 delimitó explícitamente el alcance de este estudio a los indicadores Instock y DOH, excluyendo el Fill Rate. El resultado del Nivel 1 del Modelo en V (Sección 9.6.1) demuestra, no obstante, que la arquitectura de ingesta y gobernanza construida no está limitada estructuralmente a estos dos indicadores, si no que responde a un diseño de capas generalizable a cualquier métrica que comparta las mismas dimensiones maestras. Se recomienda, por tanto, evaluar la incorporación del Fill Rate en una segunda fase, considerando que el costo marginal de sumar un indicador sobre una arquitectura ya validada es sustancialmente menor al costo de construcción de dicha arquitectura, ya asumido íntegramente por este proyecto.

Por último, el resultado cuantificado de la Eficiencia del Ciclo del Proceso (Sección 9.2.2) demostró que ese mismo 77% del tiempo del proceso actual es consumido en tareas de normalización manual, equivalente $272.000 mensuales en horas hombre, corresponde

específicamente a la cuenta Walmart. La arquitectura propuesta, sin embargo, no contiene ninguna regla de negocio exclusiva de esa cuenta, la separación en capas de refinamiento progresivo y el mecanismo híbrido de ingesta vía API y RPA (Sección 9.6.3) se sostienen sobre estructuras de dimensiones conformadas genéricas. Se recomienda evaluar la replicabilidad de este modelo hacia otras cuentas del Canal Moderno relevantes para TMLUC, bajo la hipótesis de que una fracción comparable de tiempo improductivo cuantificado en este estudio podría estar replicándose, sin haber sido aún diagnosticada, en dichas cuentas

11 Discusión

El presente estudio partió de un marco teórico robusto, construido sobre pilares de gestión de procesos (Lean Six Sigma), modelamiento de datos (Kimball y Ross) e ingeniería de sistemas (Modelo en V), entre otros. Habiendo obtenido los resultados descritos en el Capítulo 9, corresponde ahora confrontar dicho marco con la evidencia empírica recogida, para explicitar en qué medida la teoría se confirmó, se matizó, o expuso los límites de su propio alcance.

La confrontación más reveladora surge la contrastar los resultados con el propio Modelo de Madurez de Wayne Eckerson (Sección 6.7.4), que esta memoria utilizó para diagnosticar a TMLUC en la Etapa 1, inmediatamente antes del Punto de Fricción Gulf. La teoría es explícita respecto a que se requiere para franquear esa barrera: conseguir un patrocinador directivo, resolver la escasez de licencias analíticas, capacitar formalmente al personal, y solucionar de raíz las asimetrías y problemas de calidad de datos en los sistemas de origen, Los resultados de esta investigación muestran que la solución implementada resuelve de manera demostrable dos de esas cuatro condiciones: la calidad y asimetría del dato queda resuelta mediante reglas de integridad referencial y control de unicidad validadas en la Compuerta 2 (Sección 9.6.6), y la escasez de licencias queda resulta por diseño, al restringir deliberadamente la solución tecnológica al ecosistema Microsoft ya contratado por TMLUC (Sección 9.6.4). Sin embargo, las otras dos condiciones que Eckerson exige, conseguir un patrocinador directivo y capacitar formalmente al personal, son de naturaleza organizacional y cultural, no tecnológica, y por tanto exceden el alcance de un proyecto de ingeniería de datos. Esta confrontación permite matizar la teoría, más

que simplemente confirmarla o refutarla, el aporte real de esta memoria consiste en resolver el prerrequisito técnico que hace posible cruzar el Gulf, dejando en evidencia que dicho prerrequisito, aunque necesario, no es suficiente por sí solo según la propia literatura de madurez de Inteligencia de Negocios. La organización todavía debe gestionar la dimensión humana de esa transición para completar el cruce.

Una segunda confrontación, de naturaleza distinta, surge al comparar la infraestructura prescrita clásicamente por Kimball y Ross (2002) con la efectivamente implementada. El modelamiento dimensional que este estudio aplicó, grano atómico, dimensiones conformadas, tablas de hechos desacopladas (Sección 9.6.2), fue formulado por sus autores en un contexto tecnológico que presuponía motores de bases de datos gestionadas en servidor. Esta memoria demuestra, sin embargo, que dichos principios de diseño dimensional son independientes de la infraestructura física que los sostiene, implementados sobre DuckDB, un motor embebido que ni siquiera existía como categoría tecnológica consolidada al momento en que Kimbal y Ross publicaron su obra, el modelo mantuvo garantías de integridad, unicidad y consistencia validadas en la Compuerta 2 y 3 (Sección 9.6.6). Este hallazgo sugiere que la vigencia de la teoría de Kimball y Ross (2002) no reside en la infraestructura que ambos autores asumían disponible en su época, sino en los principios lógicos de organización del dato, que esta investigación confirma como aplicables incluso bajo restricciones tecnológicas que sus propios autores no contemplaron.

Un tercer contraste directa se da en la distribución de tiempo del Analista CPFR. Cano (2018) y Curto (2010) plantean que, en organizaciones de baja madurez analítica, cerca del 80% del tiempo se destina a preparar los datos y solo un 20% a analizarlos. El diagnóstico de esta memoria encontró un 77% y un 23%, lo que es consistente con esa regularidad, con la salvedad de que proviene de un único proceso y de tiempos declarados por el Analista CPFR. En la misma

línea, George (2002) asocia un PCE igual o inferior al 20% con procesos dominados por trampas de tiempo; el 23% obtenido se ubica levemente sobre ese umbral, por lo que la evidencia es coherente con la teoría, pero no la confirma de forma estricta.

Respecto de la capacidad tecnológica, Cano (2018) advierte que, en planillas de cálculo, la retención de datos queda supeditada a los límites físicos de la herramienta. El DHT calculado (29 días) sugería, en teoría, capacidad para un mes de historia; en la práctica, la saturación redujo la ventana útil a 1-3 días. El hallazgo confirma la teoría en lo cualitativo y la matiza en lo cuantitativo. El límite físico de filas sobreestima la capacidad operativa, y el DHT debe leerse como cota superior. Por su parte, el rango de error de 20%-40% de Davenport (2006) se utilizó como referencia de la literatura, este estudio no midió la tasa de error de las planillas de TMLUC, por lo que puede confirmar ni refutar ese rango.

En cuanto a la Metodología, Forsberg y Mooz (1991) sostienen que planificar la verificación junto con cada nivel de especificación evita la propagación de errores. La falla de RPA se detectó y corrigió en las pruebas unitarias (Compuerta 1), antes de integrarse con las capas Plata y Oro, lo que es consistente con ese planteamiento. Sin embargo, no se contrastó con un escenario alternativo (Cascada o Scrum), por lo que la ventaja relativa del Modelo en V fue argumentada (Sección 8.1.4), pero no evaluada empíricamente.

Respecto de la evidencia previa revisada en el Marco Teórico, que describe la arquitectura de BI y de almacenes de datos de manera general (Cano, 2018; Curto, 2010; Inmon, 2005; Kimball y Ross, 2002), el aporte de esta memoria no es teórico, sino de aplicación y verificación en un contexto concreto: (i) una arquitectura de medallón verificada del lado del proveedor de una relación B2B retail, bajo las restricciones de no intervención en la fuente y de licencias ya contratadas; (ii) el contraste explícito entre un escenario ideal y uno factible; y (iii) un protocolo

de cuatro compuertas con reconciliación de la población completa contra la fuente nativa de Walmart.

Corresponde, finalmente, señalar con honestidad académica el límite del alcance teórico verificado. El Efecto Látigo descrito por Lee et al. (1997), motivación teórica del problema de investigación, atribuye la amplificación de la variabilidad de la demanda a la distorsión de las señales de inventario. Esta investigación abordó la causa que la teoría identifica como origen del efecto, la falta de visibilidad y fiabilidad del dato compartido entre TMLUC y Walmart, pero no midió una reducción de la variabilidad de la demanda ni de las amplificaciones de pedido, indicadores que habrían requerido una historia más extensa que la disponible. La contribución debe entenderse, por tanto, como la resolución del prerrequisito de visibilidad y no como una medición empírica de su mitigación.

En el plano económico, Cano (2018) sostiene que el TCO de un proyect de Inteligencia de Negocios debe intrgrar el tempo del personal interno y no solo la tecnología. La evaluación realizada confirma la pertinencia de este criterio, en una solución sin costo incremental de licencias, el TCO queda determinado caso por completo por horas hombre, de desarrollo y de mantenimiento . A su vez, el umbral de equilibrio de Instock es coherente con el planteamiento de Christopher (2016) de que la visibilidad actúa como sustituto del inventario físico (p. 178), el beneficio potencial de la visibilidad sobre la disponibilidad en góndola supera ampliamente al beneficio por productividad. Sin embargo, esta investigación no midió ducho efecto, por lo que la relación se plantea como hipótesis respaldada por la teoría y no como resultado verificado.

12 Limitaciones

Las limitaciones que se describen a continuación no invalidan los resultados alcanzados, sino que delimitan con precisión el alcance dentro del cual dichos resultados deben interpretarse, en coherencia con la reflexión crítica que la propia investigación exige.

En primer lugar, el estudio delimitó su alcance de indicadores a Instock y DOH, excluyendo deliberadamente el Fill Rate desde la Sección 8.1.2. Esta exclusión no respondió únicamente a una decisión de acotación de alcance, sino a una limitación real de acceso a la información. El cálculo del Fill Rate exige contrastar las unidades solicitadas mediantes Órdenes de Compra (OC) emitidas por Walmart con las unidades efectivamente despachadas, y construir una matriz de abastecimiento Centro de Distribución-Sala-Producto que permite identificar desde que CD se abastece cada combinación Sala-Producto, información a la cual TMLUC no dispone actualmente de acceso suficiente dentro de Retail Link. Las OC constituyen, además, un catálogo de fuente de datos no contemplado dentro del universo de información abordado por esta arquitectura. Si bien la Sección 10.2 plantea incorporar este indicador como recomendación de trabajo futuro, dicho incorporación no depende únicamente de una decisión de diseño, sino de una gestión para ampliar y conocer como se distribuyen los productos desde los Centros de Distribución hacia las Salas.

En segundo lugar, esta investigación no pudo validar la solución contra una serie histórica extensa de Stock MultiEstado previo a su implementación. A diferencia de las Ventas POS, suyo registro es persistente en el portal Retail Link, el dato de Stock MultiEstado no es persistente

durante mucho tiempo. Bajo el proceso manual previo a esta investigación, el Analista CPFR preservaba este dato día a día en archivos Excel independientes, sin consolidarlo en un repositorio único, lo que impide en la práctica reconstruir una serie histórica confiable anterior a la implementación de la arquitectura propuesta. En consecuencia, la validación de este estudio se circunscribió a los datos capturados de manera prospectiva a partir de la puesta en marcha del proyecto.

En tercer lugar, la validación de la solución mediante el Protocolo de Verificación se sostiene sobre un período de observación acotado y único evaluador. La compuerta 4 se ejecutó durante cinco días consecutivos, por parte del Analista CPFR. Una venta de observación más extensa habría permitido descartar con mayor certeza la ocurrencia de inconsistencias de baja frecuencia, tales como anomalías asociadas a cierres de mes o periodos de alta estacionalidad comercial, y la participación de más usuarios habría permitido contrastar la aceptación de tablero desde distintos perfiles.

En cuarto lugar, esta investigación no midió el estado posterior a la implementación ni evaluó su retorno económico. No se registraron el PCE, las horas semanales de trabajo ni la Fuga de Valor de Ingeniería posteriores a la solución, por lo que liberación de las horas semanales de consolidación manual del Analista CPFR es una consecuencia derivada de eliminar las actividades manuales diagnosticadas y no una reducción observada empíricamente.

En quinto, la selección de DuckDB como motor de persistencia para la Capa Plata respondió a una restricción real de política de TI de la organización, y no a una comparación exhaustiva de todas las alternativas tecnológicas disponibles en el mercado (Sección 9.6.4). El estudio no evaluó empíricamente el comportamiento del sistema ante volúmenes significativamente mayores a los actuales, ni las limitaciones de concurrencia de escritura propias

de un motor embebido bajo un escenario de múltiples usuarios simultáneos.

En sexto lugar, y en línea con lo discutido en el Capítulo 11, esta investigación acotó su alcance a la dimensión técnica del problema de madurez analítica diagnosticado, dejando fuera de su ámbito de acción las condiciones organizacionales que la propia literatura de Eckerson identifica como necesarias para completar la transición hacia etapas superiores de madurez, tales como la obtención de un patrocinador directivo o la capacitación formal del personal

13 Referencias

Cano, J. L. (2018). *Business Intelligence Competir Con Informacion.* Fundación Cultural Banesto.

Christopher, M. (2016). *Logistics & Supply Chain Management* (5ta ed.). Pearson Education Limited.

Curto, J. (2010). *Introducción al Business Intelligence y Big Data: Generando valor a partir de los datos.* UOC.

DAMA-International. (2020). *Guía del conocimiento para la gestión de datos* (Vol. II). Technics Publications.

Databricks. (2024). Obtenido de https://docs.databricks.com/aws/en/lakehouse/medallion

Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review, 84(1)*, 98-107.

Forsberg, K., & Mooz, H. (1991). The relationship of System Engineering to the Project Cycle. *National Council On Systems Engineering (NCOSE) and American Society for Engineering Management (ASEM) Joint Conference*.

George, M. L. (2002). *Lean Six Sigma: Combining Six Sigma Quality with Lean Production Speed.* McGraw-Hill.

Gorelik, A. (2019). *The Enterpise Big Data Lake: Delivering the promise of Big Data and Data.* O'Reilly Media.

Heizer, J., & Render, B. (2009). *Principios de administración de operaciones* (7ma ed.). Pearson Education.

Hitpass, B. (2012). *Business Process Management (BPM) - Fundamentos y Conceptos de Implementación* (4ta ed.). BHH Ltda.

Inmon, W. H. (2005). *Building the Data Warehouse* (4ta ed.). Wiley Publishing, Inc.

Jhon, M., DeWitt, W., James, K., Soonhong, M., Smith, C., Nix, N., & Zach, Z. (2001). Defining supply chain management. *Journal of Business Logistics, 22(2)*.

Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling* (2da ed.). John Wiley & Sons, Inc.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science, 43(4)*, 546-558. doi:https://doi.org/10.1287/mnsc.43.4.546

Lituus Analytics. (2023). *Modelos de Madurez BI*. Obtenido de https://lituus.cl/Analytics/business-intelligence/index.html

Meganoticias. (9 de Octubre de 2024). *¿Te gusta la Ingeniería Civil Industrial? Conoce los sueldos promedio que ganan sus egresados*.

Negash, S. (2004). Business Intelligence. *Communications of the Association for Information Systems, 13*. doi:https://doi.org/10.17705/1CAIS.01315

Olszak, C. M. (2013). Assessment of Business Intelligence Maturity in the Selected Organizations. *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems.* IEEE.

Raasveldt, M., & Mühleisen, H. (2019). DuckDB: An Embedded Analytical Database. Association for Computing Machinery. doi:https://doi.org/10.1145/3299869.3314047

Royce, W. W. (1970). Managing the development of large software systems. *Proceedings, IEEE WESCON*.

Sacu, C., & Spruit, M. (2010). BIDM: The Business Intelligence Development Model. Institute of Information and Computing Sciences, Utrecht University.

Sapag Chain, N., Sapag Chain, R., & Sapag Puelma, J. M. (2014). Preparación y evaluación de proyectos. McGraw-Hill Educatio.

Schwaber, K., & Beedle, M. (2001). Agile Software Development with Scrum. *Prentice Hall*. doi:https://dl.acm.org/doi/abs/10.5555/559553

van der Aalst, W. M., Bichler, M., & Heinzl, A. (2018). Robotic Process Automation. *Business & Information Systems Engineering, 60(4)*. doi:https://doi.org/10.1007/s12599-018-0542-x