



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería de Sistemas e Informática

Escuela Profesional de Ingeniería de Software

**Optimización de un proceso ETL para el sistema de
alertas tempranas en la gestión de riesgos de clientes en
una entidad financiera peruana**

TRABAJO DE SUFICIENCIA PROFESIONAL

Para optar el Título Profesional de Ingeniero de Software

AUTOR

Héctor Daniel HUAPAYA VÁSQUEZ

ASESOR

Luis Alberto ALARCÓN LOAYZA

Lima, Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Huapaya, H. (2025). *Optimización de un proceso ETL para el sistema de alertas tempranas en la gestión de riesgos de clientes en una entidad financiera peruana*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Escuela Profesional de Ingeniería de Software]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Héctor Daniel Huapaya Vásquez
Tipo de documento de identidad	DNI
Número de documento de identidad	72854259
URL de ORCID	https://orcid.org/0000-0003-3616-9046
Datos de asesor	
Nombres y apellidos	Luis Alberto Alarcón Loayza
Tipo de documento de identidad	DNI
Número de documento de identidad	00456684
URL de ORCID	https://orcid.org/0000-0001-8340-7161
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	William Martin Enríquez Maguiña
Tipo de documento	DNI
Número de documento de identidad	06179457
Miembro del jurado 1	
Nombres y apellidos	Norberto Antonio Osorio Beltrán
Tipo de documento	DNI
Número de documento de identidad	08799230
Datos de investigación	
Línea de investigación	No aplica
Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	País: Perú Departamento: Lima

	Provincia: Lima Distrito: Cercado de Lima Jr. Carlos Amezaga No. 375 Universidad Nacional Mayor de San Marcos Latitud: -12.0564232 Longitud: -77.0843327
Año o rango de años en que se realizó la investigación	2024
URL de disciplinas OCDE	2.02.04 -- Ingeniería de sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
Escuela Profesional de Ingeniería de Software

Acta Virtual de Sustentación
del Trabajo de Suficiencia Profesional

Siendo las 19:00 horas del día 15 de enero del año 2025, se reunieron virtualmente los docentes designados como Miembros del Jurado del Trabajo de Suficiencia Profesional, presidido por el Mg. Enríquez Maguiña William Martin (Presidente), Ing. Osorio Beltrán Norberto Antonio (Miembro) y el Lic. Alarcón Loayza Luis Alberto (Miembro Asesor), usando la plataforma Meet (<https://meet.google.com/jtg-ystx-bsz>), para la sustentación virtual del Trabajo de Suficiencia Profesional intitulado: **“OPTIMIZACIÓN DE UN PROCESO ETL PARA EL SISTEMA DE ALERTAS TEMPRANAS EN LA GESTIÓN DE RIESGOS DE CLIENTES EN UNA ENTIDAD FINANCIERA PERUANA”**, por el Bachiller **Huapaya Vásquez Héctor Daniel**; para obtener el Título Profesional de Ingeniero de Software.

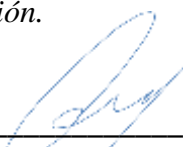
Acto seguido de la exposición del Trabajo de Suficiencia Profesional, el Presidente invitó al Bachiller a dar las respuestas a las preguntas establecidas por los miembros del Jurado.

El Bachiller en el curso de sus intervenciones demostró pleno dominio del tema, al responder con acierto y fluidez a las observaciones y preguntas formuladas por los señores miembros del Jurado.

Finalmente habiendo efectuado la calificación correspondiente por los miembros del Jurado, el Bachiller obtuvo la nota de **16 (DIECISÉIS)**.

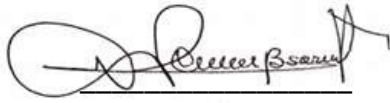
A continuación el Presidente de Jurados el Mg. Enríquez Maguiña William Martin, declara al Bachiller **Ingeniero de Software**.

Siendo las 19:50 horas, se levantó la sesión.



Presidente

Mg. Enríquez Maguiña William Martin



Miembro

Ing. Osorio Beltrán Norberto Antonio



Miembro Asesor

Lic. Alarcón Loayza Luis Alberto



CERTIFICADO DE SIMILITUD

Yo, ALARCÓN LOAYZA LUIS ALBERTO en mi condición de asesor acreditado con la Resolución Decanal N° 001059-2024-D-FISI/UNMSM y Oficio N° 001-PTTSP-FISI-2024 del Trabajo de Suficiencia Profesional cuyo título es Optimización de un Proceso ETL para el Sistema de Alertas Tempranas en la Gestión de Riesgos de Clientes en una Entidad Financiera Peruana, presentado por el bachiller HUAPAYA VÁSQUEZ HÉCTOR DANIEL para optar el título de Ingeniero de Software CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 8% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del título.

Lima, 8 de enero de 2025

Firma del Asesor

Nombres y Apellidos del asesor: ALARCÓN LOAYZA LUIS ALBERTO

DNI 00456684

FICHA CATALOGRÁFICA

Optimización de un Proceso ETL para el Sistema de Alertas Tempranas en la Gestión de Riesgos de Clientes en una Entidad Financiera Peruana

AUTOR: HUAPAYA VÁSQUEZ, HÉCTOR DANIEL

ASESOR: ALARCON LOAYZA, LUIS ALBERTO

LIMA – PERÚ, 2024

TÍTULO PROFESIONAL: Título Profesional de Ingeniero de Software

ÁREA / PROGRAMA / LÍNEA DE INVESTIGACIÓN:

Ingenierías / Tecnología de Información y Comunicación / Ingeniería de Software

PREGRADO: Universidad Nacional Mayor de San Marcos - Facultad de Ingeniería de Sistemas e Informática - Escuela Profesional de Ingeniería de Software

Formato 21 x 29.7 cm

Páginas: viii,55

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA DE SOFTWARE

**Optimización de un Proceso ETL para el Sistema de Alertas
Tempranas en la Gestión de Riesgos de Clientes en una Entidad
Financiera Peruana**

Autor: Bach. Huapaya Vásquez, Héctor Daniel
Asesor: Alarcon Loayza, Luis Alberto
Título: Trabajo de Suficiencia Profesional para optar el Título Profesional de
Ingeniería de Software
Fecha: Diciembre del 2024

RESUMEN

La creciente digitalización y el aumento de la inclusión financiera en Perú han impulsado a las entidades bancarias a mejorar sus sistemas y mejorar la obtención de información de sus clientes. Esto ha conllevado que el proceso de ingesta de datos al modelo de alertas tempranas en la entidad financiera sea lento e ineficiente, lo que provoca no cumplir con las metas de entregar el modelo indicado. En este contexto, se ha llevado a cabo un proyecto para optimizar un proceso ETL (Extracción, Transformación y Carga), con el objetivo de reducir los tiempos de procesamiento y proporcionar información oportuna para un sistema de alertas tempranas de clientes que puedan presentar problemas financieros. El proceso ETL implementado recoge datos de múltiples orígenes, incluyendo bases de datos transaccionales, archivos de texto y sistemas legados. La fase de extracción garantiza la recopilación completa de datos, mientras que la transformación aplica reglas de negocio y técnicas de normalización para asegurar la calidad y coherencia de la información. La fase de carga transfiere los datos transformados a un almacén centralizado, optimizando su accesibilidad y análisis. El proyecto se desarrolló utilizando un método ágil propio de la entidad con base en Scrum, que facilitó una gestión dinámica y colaborativa del trabajo. Se ajustaron los procesos, incorporando procesos almacenados y automatizaciones de cargas de trabajo. Este trabajo no solo automatiza procesos manuales y mejora la eficiencia operativa, sino que también establece una base sólida para tener un modelo de ingesta de datos. Estas mejoras permitieron disminuir significativamente los tiempos de procesamiento mensual, incrementando la eficiencia y efectividad del sistema de alertas tempranas.

Palabras claves: ETL, evaluación riesgos, aprovisionamiento datos, entidad bancaria, inteligencia de negocios.

MAJOR NATIONAL UNIVERSITY OF SAN MARCOS
FACULTY OF SYSTEMS AND INFORMATICS ENGINEERING
PROFESSIONAL SCHOOL OF SOFTWARE ENGINEERING

**Optimization of an ETL Process for the Early Warning System
in Customer Risk Management in a Peruvian Financial Entity**

Author: Bach. Huapaya Vásquez, Héctor Daniel
Advisor: Alarcon Loayza, Luis Alberto
Title: Professional Sufficiency Work to opt for the Professional Title of Software Engineer
Date: December 2024

ABSTRACT

Growing digitalization and increasing financial inclusion in Peru have prompted banks to improve their systems and customer information intake. This has generated a high demand for efficient processes that ensure the accuracy and quality of the data. In this context, a project has been carried out to optimize an ETL (Extraction, Transformation and Loading) process, with the objective of reducing processing times and providing timely information for an early warning system for customers that may present financial problems. The ETL process implemented in this project collects data from multiple sources, including transactional databases, text files and legacy systems. The extraction phase ensures complete data collection, while transformation applies business rules and normalization techniques to ensure information quality and consistency. The loading phase transfers the transformed data to a centralized warehouse, optimizing its accessibility and analysis. The project was developed using the agile Scrum methodology, which facilitated dynamic and collaborative work management. Processes were reviewed and adjusted, incorporating stored processes and workload automation. This work not only automates manual processes and improves operational efficiency, but also establishes a solid foundation for the development of advanced analytical models and business intelligence tools. These improvements significantly reduced monthly processing times, increasing the efficiency and effectiveness of the early warning system.

Key words: ETL, risk assessment, data sourcing, banking, business intelligence, business intelligence.

TABLA DE CONTENIDOS

RESUMEN.....	III
INTRODUCCIÓN	1
CAPÍTULO I - TRAYECTORIA PROFESIONAL.....	3
CAPÍTULO II – CONTEXTO EN EL QUE SE DESARROLLA LA EXPERIENCIA	7
2.1 EMPRESA - ACTIVIDAD QUE REALIZA	7
2.2 VISIÓN	7
2.3 MISIÓN	7
2.4 ORGANIZACIÓN DE LA EMPRESA	9
2.5 ÁREA, CARGO Y FUNCIONES DESEMPEÑADAS.....	10
2.6 EXPERIENCIA PROFESIONAL REALIZADA EN LA ORGANIZACIÓN.....	11
CAPÍTULO III – ACTIVIDADES DESARROLLADAS.....	12
3.1 SITUACIÓN PROBLEMÁTICA	12
3.1.1 DEFINICIÓN DEL PROBLEMA	14
3.2 SOLUCIÓN	14
3.2.1 OBJETIVOS	17
3.2.2 ALCANCE.....	17
3.2.3 ETAPAS Y METODOLOGÍA	18
3.2.4 FUNDAMENTOS UTILIZADOS.....	21
3.2.4.1 METODOLOGÍAS ÁGILES	21
3.2.4.2 KANBAN.....	22
3.2.4.3 MEDOLOGÍA EN CASCADA	24
3.2.4.4 PROGRAMACIÓN EXTREMA	25
3.2.4.5 ETL	25
3.2.4.6 JIRA	26
3.2.4.7 PROCEDIMIENTOS ALMACENADOS.....	28
3.2.5 IMPLEMENTACIÓN DE LAS ÁREAS, PROCESOS, SISTEMAS Y BUENAS PRÁCTICAS.....	30
3.2.5.1 USO DE LA METODOLOGÍA.....	30

3.3 EVALUACIÓN	46
3.3.1 EVALUACIÓN COSTO – BENEFICIO	46
3.3.1.1 ANÁLISIS DE COSTOS.....	46
3.3.1.2 ANÁLISIS DE BENEFICIOS.....	47
3.3.1.3 EVALUACIÓN DE RESULTADO.....	48
3.3.2 INTERPRETACIÓN DEL VAR Y TIR	49
CAPÍTULO IV – REFLEXIÓN CRÍTICA DE LA EXPERIENCIA	50
CAPÍTULO V – CONCLUSIONES Y RECOMENDACIONES	51
5.1 CONCLUSIONES.....	51
5.2 RECOMENDACIONES.....	51
BIBLIOGRAFÍA	53

ÍNDICE DE FIGURAS

FIGURA 1: MISIÓN DE LA ENTIDAD FINANCIERA	8
FIGURA 2: ESQUEMA ORGANIZATIVO DE LA EMPRESA	9
FIGURA 3: INCREMENTO DE BANCARIZACIÓN EN PERÚ.....	14
FIGURA 4: FLUJO DE LA SOLUCIÓN: PROCESO DIARIO DE VALIDACION DE ARCHIVOS	16
FIGURA 5: FLUJO DE LA SOLUCIÓN: PROCESO DE LLENADO Y CRUCES DE DATOS.....	16
FIGURA 6: PRIMERA ETAPA PROYECTO: IDENTIFICACIÓN DE PROBLEMAS.....	18
FIGURA 7: SEGUNDA ETAPA DEL PROYECTO: IMPLEMENTACIÓN DEL PROYECTO	19
FIGURA 8: MARCO DE TRABAJO SCRUM DE LA ENTIDAD: DE TRABAJO EN EQUIPO CON LAS REUNIONES DIARIAS.....	20
FIGURA 9: COMUNICACIÓN EN EL TRABAJO EN EQUIPO.....	22
FIGURA 12: VISUALIZACIÓN DE TRABAJO EN KANBAN.....	23
FIGURA 13: ESTRUCTURA DE LA METODOLOGÍA EN CASCADA.....	24
FIGURA 10: ARQUITECTURA DE UN PROCESO ETL	26
FIGURA 11: FLUJO DE TAREAS EN JIRA.....	27
FIGURA 14: FLUJO DE UN PROCEDIMIENTO ALMACENADO	28
FIGURA 15: ETAPAS DE LA GESTIÓN DEL PROYECTO.....	30
FIGURA 16: SECUENCIA DE TRABAJO EN JIRA	31
FIGURA 17: CRONOGRAMA DE PLANIFICACIÓN DEL PROYECTO	32
FIGURA 18: DIAGRAMA DE ARQUITECTURA.....	39
FIGURA 19: JOBS DE VALIDACIONES POR ARCHIVOS EN LA HERRAMIENTA CONTROL M.....	40
FIGURA 20: CARGAR DATA EN TABLAS INDEXADAS.....	41
FIGURA 21: CARGA DE DATOS	41
FIGURA 22: CARGA Y TRANSFORMACIÓN Y CRUCE DE DATA	42
FIGURA 23: FORMATO DE ARCHIVOS	43
FIGURA 24: CASOS DE PRUEBA	44
FIGURA 25: SECUENCIA DE TRABAJO EN JIRA	45

ÍNDICE DE TABLAS

TABLA 1: EXPERIENCIAS LABORALES	3
TABLA 2: EDUCACIÓN ACADÉMICA	5
TABLA 3: CURSOS DE CAPACITACIONES	6
TABLA 4: CERTIFICACIONES	6
TABLA 5: IDIOMAS.....	6
TABLA 6: COMPARATIVA DEL INCREMENTO EN LA INCLUSIÓN FINANCIERA Y LA ADOPCIÓN DE SERVICIOS BANCARIOS DIGITALES EN PERÚ (2020-2023).....	12
TABLA 7: INCREMENTO DE CONTRATOS POR AÑO	13
TABLA 8: REDUCCIÓN TIEMPOS DE ENTREGA.....	15
TABLA 9: FASES DEL PROYECTO	30
TABLA 10: ROLES DEDICADOS AL PROYECTO	33
TABLA 11: HERRAMIENTAS Y ENTORNOS DE TRABAJO PARA EL PROYECTO.....	35
TABLA 12: ARCHIVOS NECESARIOS DE ENTRADA PARA EL PROCESO ETL	35
TABLA 13: VALIDACIONES PRINCIPALES AL PROCESO	37
TABLA 14: ERRORES CONSIDERADOS PARA REINTENTOS.....	42
TABLA 15: COSTO TOTAL DE IMPLEMENTACIÓN DEL PROYECTO.....	46
TABLA 16: PROYECCIÓN DE COSTES EN VALIDACIÓN EN EL PROCESO MENSUAL DE APROVISIONAMIENTO	47

INTRODUCCIÓN

La gestión y aprovechamiento adecuado de la información son fundamentales para el éxito en el sector financiero. Las entidades financieras enfrentan constantemente desafíos relacionados con la precisión, calidad y disponibilidad de los datos, elementos críticos para una toma de decisiones informada y efectiva. En este contexto, el uso de un buen proceso de Extracción, Transformación y Carga (ETL, por sus siglas en inglés) se presenta como una solución eficaz para tener consolidados los datos y favorecer a tener la información disponible, mejorando la capacidad del equipo de riesgos para evaluar y gestionar a clientes.

El presente Trabajo de Suficiencia Profesional se centra en la optimización y mejora de un proceso ETL diseñado para aprovisionar de información a un sistema de alertas tempranas en evaluaciones de riesgos a clientes en una entidad financiera. La relevancia de esta iniciativa radica en su contribución a la mejora de la gestión de los tiempos de entrega y de información íntegra, permitiendo a la entidad anticipar posibles amenazas de clientes en riesgo y tomar decisiones estratégicas basadas en datos precisos y oportunos. Este proyecto no solo busca optimizar los procesos internos y la calidad de los datos, sino también establecer una base sólida para el desarrollo de futuros modelos analíticos y herramientas de inteligencia de negocios.

La adopción de un enfoque ETL bien estructurado y robusto garantiza que la información recopilada de diversas fuentes sea transformada y cargada de manera eficiente en un almacén de datos centralizado. Esto facilita su análisis y utilización por parte del sistema destino, alineándose con las mejores prácticas de la industria y contribuyendo a la estabilidad y crecimiento sostenido de la entidad financiera.

El proceso ETL no solo representa una mejora operativa significativa, sino también un paso crucial en la capacidad de la entidad para gestionar riesgos de manera proactiva y eficiente, abarcando los siguientes capítulos:

El Capítulo 1, presenta la trayectoria del autor, los conocimientos adquiridos y la experiencia laboral.

El Capítulo 2 presenta los antecedentes del estudio y los diversos aspectos de la industria bancaria a los que se refiere este estudio.

En el Capítulo 3 se explica el ETL, los problemas que llevaron a su implementación, sus objetivos y alcances, las bases teóricas utilizadas, su desarrollo e implementación gradual, que es la base de este estudio.

El Capítulo 4 presenta una reflexión crítica sobre la carrera y sus aportes, herramientas y conocimientos adquiridos, así como un análisis de las estadísticas anuales, confirmando que los procesos mensuales constituyen la exitosa implementación del nuevo proceso.

Finalmente, en el Capítulo 5 encontramos las conclusiones y recomendaciones, las cuales se presentan al final del trabajo entregado.

CAPÍTULO I - TRAYECTORIA PROFESIONAL

El autor se desempeña como gestor técnico de proyectos de software, especialmente en el área financiera, utilizando el lenguaje de programación java, así como otras tecnologías complementarias para tal función. Tiene amplia experiencia en análisis, desarrollo y pruebas de nuevos componentes así como mantenimiento a sistemas productivos.

Actualmente trabaja para una institución financiera, liderando desarrollos y mantenimiento de los sistemas de uso interno de la empresa. El autor ha trabajado anteriormente para otras empresas tecnológicas en este campo financiero. Los detalles se muestran en la Tabla 1:

TABLA 1:

EXPERIENCIAS LABORALES

EMPRESA/PERIODO	DESCRIPCIÓN DE ACTIVIDADES REALIZADAS	LOGROS
Entidad Financiera (Agosto 2022 - Actualidad)	Cargo: SOFTWARE ASSOCIATE SOLUTIONS DEVELOPMENT I Actividades: Dirigir y diseñar features orientados a mejoras de aplicaciones existentes, como solución de vulnerabilidades y bugs encontrados en estos y así como migrar aplicativos de servidores o de versiones.	Logros: Implementación de componentes asociados al área de riesgos área de finanzas a nivel de gestionar desarrollos y mejoras en aplicativos propios de la entidad.
Indra (Enero 2022 - Julio 2022) Empresa consultora tecnológica que ofrece software financiero a diferentes entidades.	Cargo: Backend Senior Actividades: Análisis y desarrollo de requerimientos por parte del cliente RIMAC en el producto Vida, participar de propuestas de solución y el desarrollo Backend y Frontend.	Logros: Implementación de componentes asociados al producto Journey Vida, inclusión de nuevas reglas comerciales, nuevas validaciones, correcciones de seguridad a la aplicación.

Culqi (Septiembre 2021 – Diciembre 2021) Empresa de tecnología en el sector financiero que ofrece servicios de transacciones financieras	Cargo: Development Lead Actividades: Liderar el equipo de gestión de comercios, diseñar arquitecturas para los requerimientos, establecer la visión técnica del equipo para el logro de objetivos.	Logros: Liderar equipo para implementación de un nuevo producto pos para comercios asociados
Culqi (Abril 2020 – Agosto 2021) Empresa de tecnología en el sector financiero que ofrece servicios de transacciones financieras	Cargo: Full Stack Semi Senior Actividades: Realizar desarrollos que permitan mas estabilidad a la plataforma, trabajo bajo una arquitectura orientada a microservicios, mediante metodologías ágiles.	Logros: Análisis y Desarrollo frontend/backend de onboarding Afiliate, FFVV, admin y dashboard. Técnico: Creación de módulos en vuejs usando la librería culqi ui, consultando microservicios a través de axios. Gestión de incidencias producto mpos.
Llamabus (Mayo 2020 – Junio 2020) Empresa de tecnología en el sector financiero que ofrece servicios de transacciones financieras	Cargo: Full Stack Actividades: Desarrollo bajo la modalidad freelance. Desarrollo componentes frontend de plataforma de compra de boletos.	Logros: Fomentar el desarrollo bajo micro frontends
Nanky Studio (Enero 2020 – Marzo 2020) Empresa de tecnología en el sector financiero que ofrece servicios de transacciones financieras	Cargo: Full Stack Actividades: Desarrollo bajo la modalidad freelance. Frontend de la plataforma de e-commerce.	Logros: Desarrollo de módulos de compras
UNMSM - Quipucamayoc (Junio 2019 – Marzo 2020) Oficina parte de la red telemática de la Universidad Nacional Mayor de San Marcos encargada de los sistemas contables y administrativos.	Cargo: Full Stack Developer Actividades: Desarrollo de nuevas herramientas que brinden a la universidad la automatización de sus procesos, bajo una metodología híbrida. Mantenimiento a Sistema de Contabilidad Quipucamayoc.	Logros: Análisis y desarrollo del sistema de Asesoría Legal. Análisis y desarrollo del sistema para Bienes Culturales.

Interbank (Enero 2019 – Marzo 2019) Entidad Financiera	Cargo: Programador Java Actividades: Desarrollo de nueva herramienta que permita la mejor gestión del producto inmobiliario.	Logros: Desarrollo de una web donde los ejecutivos y la gerencia inmobiliaria podía ver el flujo de los proyectos inmobiliarios, como también tableros de control que permitan un mejor seguimiento a la actividad y un tablero comerciales del mercado inmobiliario.
UNAC - OTIC (Abril 2018 – Diciembre 2018) Oficina parte de la red telemática de la Universidad Nacional del Callao encargada de los sistemas administrativos.	Cargo: Analista Programador Junior Actividades: Desarrollo de nuevas herramientas para los estudiantes, utilizando nuevas tecnologías con metodología ágil.	Logros: Desarrollo del sistema de subida de archivo. Desarrollo y mantenimiento del Sistema de Bolsa de Trabajo. Desarrollo del sistema de seguimiento al egresado.

Nota: Preparado por el autor

La educación académica del autor se detalla en la Tabla 2:

TABLA 2:

EDUCACIÓN ACADÉMICA

ESTUDIO REALIZADO	PERIODO	INSTITUCIÓN
Grado Académico de Bachiller en Ingeniería de Software	2015 - 2020	Escuela Académico Profesional de Ingeniería de Software Facultad de Ingeniería de Sistemas e Informática Universidad Nacional Mayor de San Marcos
Secundaria Completa	2008 - 2012	Institucion Educativa Emblematica José María Eguren

Nota: Preparado por el autor

Los cursos realizados por el autor se detalla en la Tabla 3:

TABLA 3:

CURSOS DE CAPACITACIONES

CAPACITACIÓN	NIVEL	INSTITUCIÓN
• Curso de Informática	Avanzado	Instituto de Informática de la UNMSM (CINFO)
• PHP y MySQL • Programacion Orientada a Objetos • Android • Administración en Linux	Avanzado	Sistemas UNI

Nota: Preparado por el autor

Las certificaciones realizados por el autor se detalla en la Tabla 4:

TABLA 4:

CERTIFICACIONES

CERTIFICACIÓN	PERIODO	INSTITUCIÓN
SCRUM fundamentals	2018	SCRUM STUDY
AWS Cloud Practitioner	2023	AWS

Nota: Preparado por el autor

Los estudios de idioma realizados por el autor se detalla en la Tabla 5:

TABLA 5:

IDIOMAS

IDIOMA	PERIODO	NIVEL	INSTITUCIÓN
Inglés	2019 - 2020	Intermedio	WallStreetEnglish

Nota: Preparado por el autor

CAPÍTULO II – CONTEXTO EN EL QUE SE DESARROLLA LA EXPERIENCIA

2.1 EMPRESA - ACTIVIDAD QUE REALIZA

La institución financiera goza de una alta reputación en el Perú y el extranjero, pertenece a un holding con accionistas del Perú y extranjero.

La entidad se encuentra autorizada para operar por la Superintendencia de Banca, Seguros y AFP (SBS) según la Ley N° 26702 que establece un marco regulatorio para las entidades de este sector que son Finanzas y Seguros.

A través de su extensa red de oficinas en provincias y sedes en todo el país realiza sus principales actividades financieras, con ubicación de su central en el distrito de San Isidro de Lima, capital constitucional del Perú.

2.2 VISIÓN

La entidad financiera define su visión como la búsqueda continua del bienestar futuro de sus clientes, por lo que ha desarrollado una visión basada en tres pilares:

- Integridad y siempre anteponiendo la ética en sus actividades y en todas las relaciones con diversos grupos de interés.
- Sea cauteloso y tenga siempre presente el principio de precaución a la hora de tomar decisiones por la entidad financiera.
- La transparencia es un pilar importante que esta entidad siempre busca para tener información transparente y auténtica al alcance dentro de la legalidad.

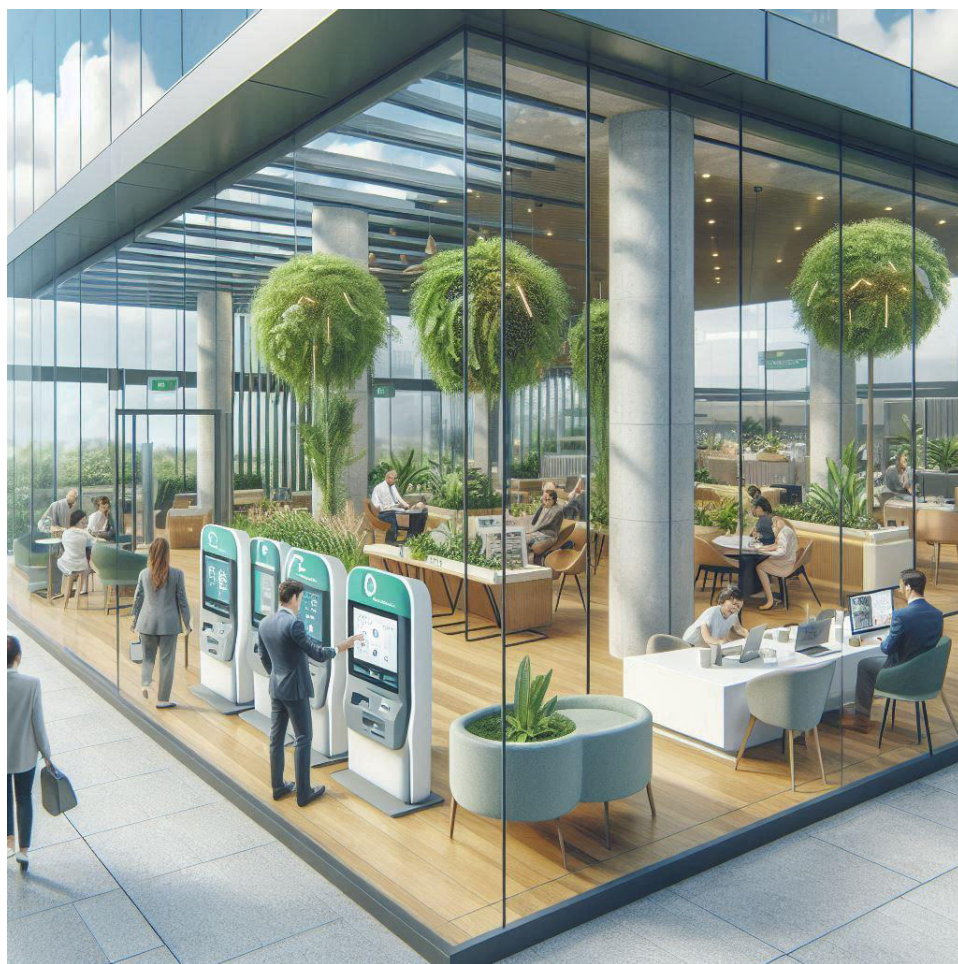
2.3 MISIÓN

La responsabilidad principal de las instituciones financieras es cumplir con satisfacer adecuadamente las necesidades de sus clientes y con ello poder agregar valor al capital de

sus accionistas, sin descuidar el desarrollo de la entidad en las diferentes comunidades de los países en los que opera, en la Figura 1 podemos observar la misión que se tiene en la entidad financiera.

FIGURA 1:

MISIÓN DE LA ENTIDAD FINANCIERA



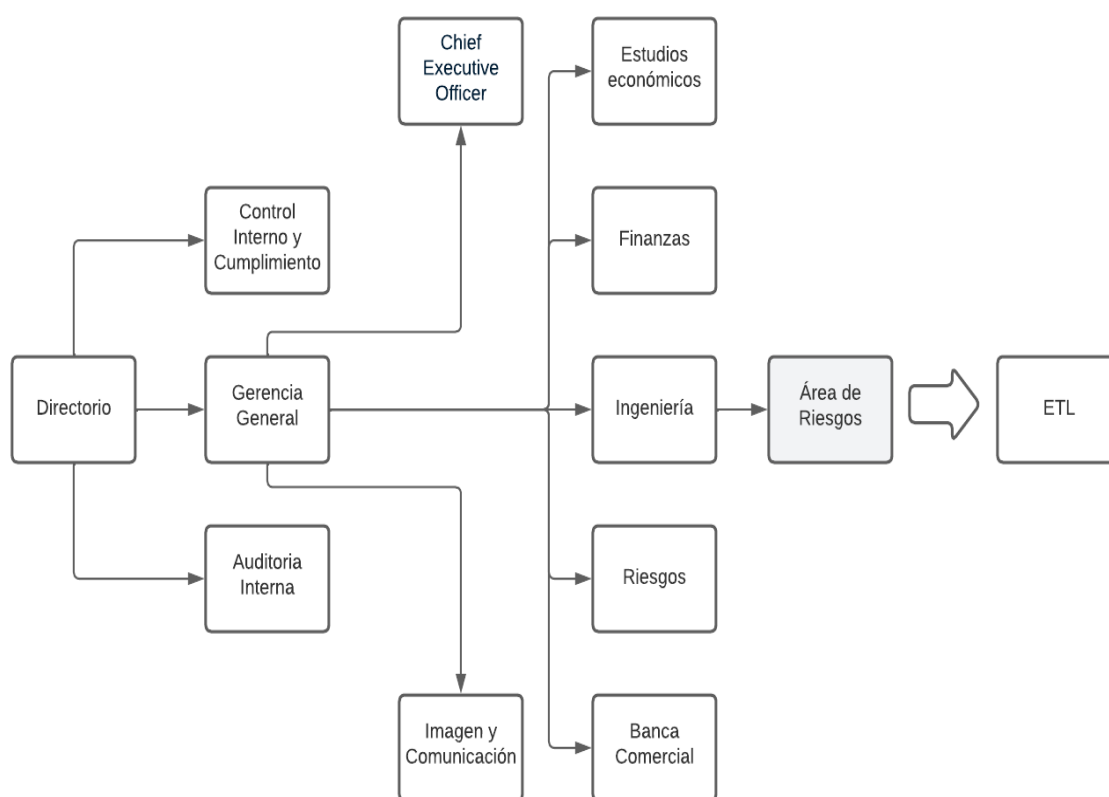
Nota: Elaborado por IA Copilot

2.4 ORGANIZACIÓN DE LA EMPRESA

La estructura organizacional de esta entidad financiera se muestra en la Figura 2 a continuación, se enfoca en la ubicación del área de riesgos donde se implementó el ETL del presente trabajo.

FIGURA 2:

ESQUEMA ORGANIZATIVO DE LA EMPRESA



Nota: Preparado por el autor

2.5 ÁREA, CARGO Y FUNCIONES DESEMPEÑADAS

El autor de este trabajo de suficiencia profesional desempeñó las responsabilidades de líder técnico del proyecto, en el equipo de riesgos, durante el proceso del mismo, con las siguientes principales responsabilidades:

- Analizar, desarrollar y modificar componentes de los aplicativos propios de la entidad por el área de riesgos.
- Realizar evaluación técnica y funcional de requerimientos propios del área, identificando el tiempo y el personal necesario para llevar a cabo el proyecto.
- Atención de migraciones por obsolescencia tanto de servidores como de aplicativos.
- Realizar pruebas técnicas y funcionales de los aplicativos y servicios del área.
- Gestionar la instalación de componentes desarrollos o modificados de acuerdo a los procedimientos de la entidad.
- Resolver incidencias reportadas por los usuarios de los aplicativos del área.
- Atender solicitudes y consultas de las áreas de negocio.
- Brindar el soporte de atención por bajas en los servicios o problemas asociados que interrumpan la operativa normal de los usuarios.

2.6 EXPERIENCIA PROFESIONAL REALIZADA EN LA ORGANIZACIÓN

El autor desde que comenzó a trabajar en la entidad financiera, ha estado involucrado en múltiples proyectos de desarrollo y mantenimiento para diversas aplicaciones en el dominio de riesgos, entre los que destacan:

- Creación del módulo de mantenimiento de carterizados para la creación de programas financieros en la entidad financiera.
- Migración del aplicativo de finanzas por oficina, por motivos de obsolescencia de tener un servidor con una versión antigua con muchas vulnerabilidades.
- Creación del módulo de fichas de contratos históricas, donde los gerentes y sub gerentes de oficina pueden consultar fichas de contratos generadas por oficina anteriores.
- Migración de los servicios transaccionales para estados financieros por obsolescencia.
- Optimizaciones a la descarga de reportes de credito que dificultaba la operativa de los usuarios de riesgos.
- Optimizaciones al aprovisionamiento mensual para la evaluación de riesgo de clientes

CAPÍTULO III – ACTIVIDADES DESARROLLADAS

3.1 SITUACIÓN PROBLEMÁTICA

En el contexto financiero actual de Perú, la inclusión financiera y la adopción de servicios digitales han registrado un notable crecimiento en los últimos tres años. Durante este periodo, el porcentaje de la población con cuentas en el sistema financiero se incrementó de un 53.7% en 2020 a un 61.1% en 2023 (INEI, 2023). Asimismo, el uso de plataformas en línea por parte de los clientes de entidades financieras ha pasado del 28% al 60% (SBS, 2020), así como lo muestra en la Tabla 6, reflejando una tendencia hacia la digitalización de los servicios bancarios.

TABLA 6:

COMPARATIVA DEL INCREMENTO EN LA INCLUSIÓN FINANCIERA Y LA ADOPCIÓN DE SERVICIOS BANCARIOS DIGITALES EN PERÚ (2020-2023)

Año	Porcentaje de la población con cuenta bancaria en el sistema financiero	Porcentaje de clientes online en entidades financieras	Número de usuarios de la banca digital
2020	53.7%	28%	Aumento del 15%
2021	56.3%	50%	Aumento del 20%
2022	57.4%	53%	Aumento del 25%
2023	61.1%	60%	Aumento del 30%

Nota. Fuente:

INEI: Informe sobre la Inclusión Financiera en el Perú, primer trimestre de 2023. Lima, 2023

SBS: Reportes anuales sobre el incremento de la digitalización de servicios bancarios entre 2020 y 2023. Lima 2020

El aumento en el número de clientes por la entidad financiera fue de 17.5% con más de 700 000 nuevos clientes y transacciones digitales presenta desafíos significativos para las entidades financieras. La necesidad de manejar grandes volúmenes de datos de manera eficiente y segura se convierte en una prioridad para garantizar la precisión y la calidad de la información utilizada en la toma de decisiones.

El Sistema de Alertas Tempranas es una herramienta de seguimiento preventivo de la cartera empresarial en la entidad financiera, cuyo objetivo es anticipar el deterioro de la calidad crediticia de los clientes en el corto plazo. Este sistema se basa en un conjunto de variables construidas a partir de indicadores o archivos de entrada de riesgo, las cuales son suministradas por el ETL, que evalúan diferentes aspectos del desempeño financiero del cliente, otorgándole una puntuación y nivel de gravedad asociado.

Es en este escenario el aprovisionamiento de información mensual para el modelo de alertas temprana de riesgos cada mes era mayor, el proceso parte de acuerdo a la cantidad de contratos de clientes, el incremento se puede visualizar en la Tabla 7, por lo que el proceso demoraba más y más, es por esto que la implementación de procesos ETL se vuelve crucial.

TABLA 7:

INCREMENTO DE CONTRATOS POR AÑO

AÑO	NÚMERO DE CONTRATOS	PORCENTAJE DE INCREMENTO
2022	1,577,915	-
2023	1,601,295	1.48%
2024	1,655,960	3.41%

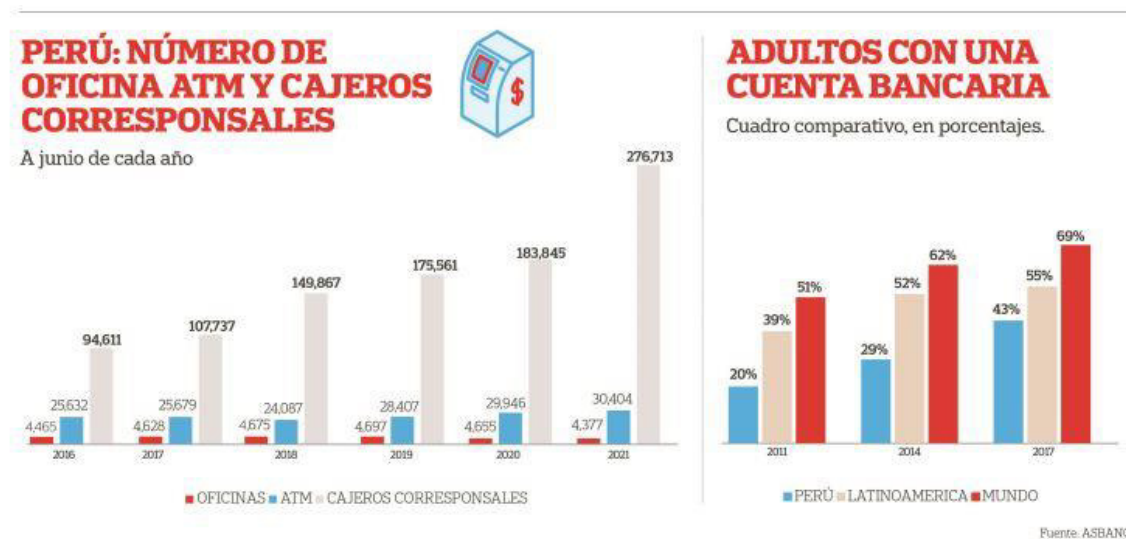
Nota: Se muestra el incremento de contratos sobre fines de año que es la base donde parte del aprovisionamiento al modelo de alertas tempranas

Es por esto que un proceso ETL bien estructurado va a permitir la integración y la calidad de datos de los múltiples fuentes, asegurando su transformación y carga en una base de datos local del aplicativo final. Esto no solo optimiza la accesibilidad y el análisis de los datos, sino que también mejora la capacidad de las entidades para anticipar y gestionar riesgos financieros.

La creciente digitalización y el incremento de clientes en el sistema financiero peruano, como lo detalla la Figura 3, subrayan la necesidad de adoptar soluciones tecnológicas avanzadas como los procesos ETL, que permiten enfrentar de manera efectiva los desafíos asociados con la gestión de datos y la toma de decisiones estratégicas.

FIGURA 3:

INCREMENTO DE BANCARIZACIÓN EN PERÚ



Nota: Fuente: Diario El Peruano. Inclusión financiera: por un país moderno, formal y conectado. 2022

3.1.1 DEFINICIÓN DEL PROBLEMA

La entidad financiera enfrenta un impacto significativo en los tiempos de procesamiento de datos debido a un proceso ETL que no es óptimo ni escalable. Este problema se manifiesta en procesos cada vez más lentos, lo que conlleva a demoras en la toma de decisiones, mayor consumo de recursos computacionales e impacto negativo en la calidad del servicio.

3.2 SOLUCIÓN

Como solución se propuso la implementación de un ETL estructurado y escalable para soportar las grandes cantidades de datos y poder entregar el resultado en menos tiempo, se tuvo en cuenta el proceso anterior para verificar que validaciones de calidad de la data que se realizaban, se mantuvieron y se agregaron más validaciones para evitar que haya errores

en el proceso mensual, lo que permitió reducir los reintentos así acortando el tiempo de entrega del resultado al usuario que se aprecia en la Tabla 8.

TABLA 8:

REDUCCIÓN TIEMPOS DE ENTREGA

TIEMPOS	PROCESO ACTUAL	PROCESO SOLUCIÓN
Tiempo Validación	1 día	20 minutos
Tiempo Procesamiento	3 días	1 día
Tiempo Total	4 días	1 día

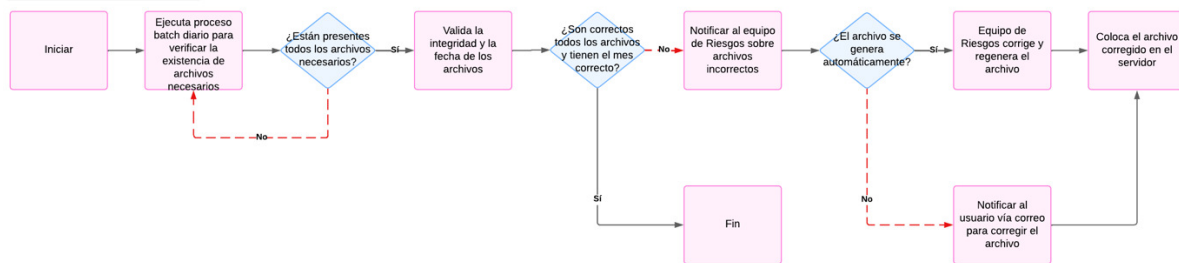
Nota: Preparado por el autor

En el flujo de la solución se propone diariamente que se ejecuta un proceso batch que valida la existencia de los archivos de entrada necesarios para el aprovisionamiento. Este batch al encontrar el archivo, lanza una notificación al equipo de Riesgos para dar conformidad la existencia del archivo necesario, en caso aún no se completen todos los archivos, se ejecutará el día siguiente validando de nuevo la existencia de los archivos, así sucesivamente hasta encontrar todos los archivos. Cuando se valide la existencia de todos los archivos, se correrá un proceso batch que valida la integridad y la fecha correspondiente al mes del archivo, en caso algún archivo esté corrupto, vacío o tenga otra fecha incorrecta al mes de ejecución, se enviará una notificación al equipo de Riesgos indicando los archivos que tienen problemas para su revisión. En caso el archivo sea uno generado automáticamente, el equipo de Riesgos es el responsable de su corrección y regeneración del mismo. En caso el archivo sea uno generado por el usuario, se deberá comunicar al usuario vía correo para su revisión y corrección. Una vez se corrijan los archivos incorrectos tienen que ser llevados a el servidor de pase. Al correr el batch diario de validación y se confirme que todos los archivos necesarios están correctos, como se grafica en la Figura 4. Comenza el siguiente pase que es un proceso batch de llenado de cada archivo, que está asociado una tabla externa de oracle cada uno, hacia tablas indexadas, como vemos en la Figura 5. Este proceso previamente limpia todas las tablas indexadas para así no tener acumulamiento de

información innecesaria, estas tablas indexadas solo son usadas para este fin de aprovisionar. Luego se realizarán los cruces de tablas aplicando reglas de negocio que conlleven hacia tablas secundarias y tablas finales, al concluir este proceso, se valida que no haya errores en el proceso y se envía una notificación vía correo donde indica el término del proceso exitosamente, en caso haya algún error en el proceso de los flujos de negocio, sean por temas de infraestructura o por algún error en la data, se enviará una notificación vía correo con el detalle del error para poder reprocesar la parte con error. Se podrán realizar reprocesos a partes específicas, para así no tener la necesidad de reprocesar todo. En el capítulo 3.2.5.1.2 se detallará el digrama de arquitectura para tener mayor entendimiento de la solución.

FIGURA 4:

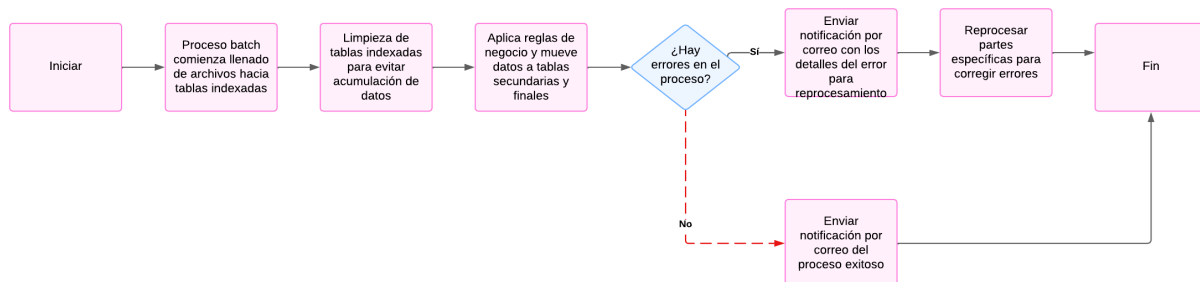
FLUJO DE LA SOLUCIÓN: PROCESO DIARIO DE VALIDACION DE ARCHIVOS



Nota: Preparado por el autor

FIGURA 5:

FLUJO DE LA SOLUCIÓN: PROCESO DE LLENADO Y CRUCES DE DATOS



Nota: Preparado por el autor

3.2.1 OBJETIVOS

3.2.1.1 Objetivo General

Mejorar la eficiencia y calidad del proceso ETL para el aprovisionamiento de información hacia la generación del modelo de alertas tempranas en evaluaciones de riesgos en clientes, utilizando las nuevas herramientas de administración de las cargas de trabajo para poder reducir los tiempos totales de validación y procesamiento, que esto conlleva a reducción de costos operacionales.

3.2.1.2 Objetivos Específicos

Se indican los objetivos específicos de acuerdo al objetivo general:

1. Diagnosticar la Causa de la Lentitud en el Proceso ETL.
2. Establecer Validaciones de Integridad y Calidad de los Datos en los Archivos
3. Realizar Validaciones de Fechas en los Archivos de Datos.
4. Diseñar y Optimizar la Carga de Trabajo del Proceso ETL.
5. Implementar Funcionalidad para Reprocesar Datos Específicos.
6. Reducir la Necesidad de Reprocesos Totales.
7. Desarrollar un Sistema de Monitoreo y Notificaciones para el Proceso ETL.

3.2.2 ALCANCE

En el alcance de este proyecto solo se contempla las ejecuciones mensuales del ETL para el aprovisionamiento de información más no el proceso diario, el proceso está orientado para el rubro de clientes mayoristas, los clientes minoristas no están en el ambito por decisión de negocio.

En caso de errores en el proceso por archivos propios de otras áreas, las correcciones de los archivos input o fuente deben ser realizadas manualmente por dichas áreas y subidas a los respectivos repositorios para que pueda ejecutarse el reintento del proceso ETL, el proceso no corregirá errores de data. Los errores solo pueden ser corregidos por criterio experto, no pueden ser corregidos por automatizaciones por ser data sensible.

Las alertas solo serán dirigidas a las personas previamente establecidas por el área de ingeniería a cargo, estas solo incluirán información del proceso, más no datos de los archivos que puede contener información sensible de clientes.

3.2.3 ETAPAS Y METODOLOGÍA

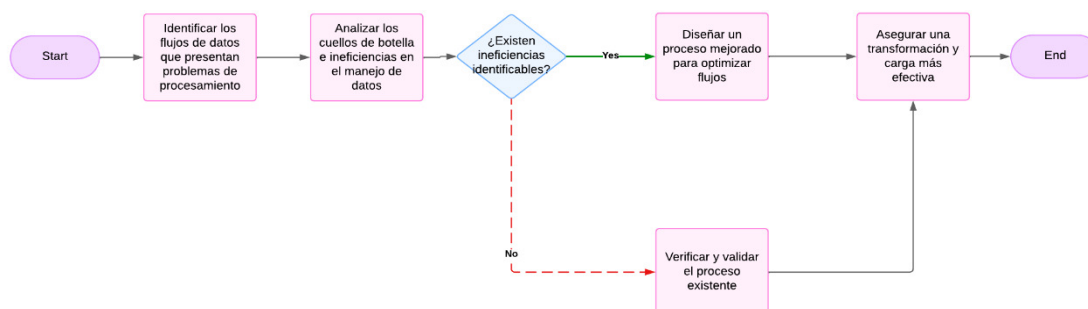
3.2.3.1 ETAPAS

El proyecto se ejecutó en dos etapas donde se brinda entregables a los usuarios, en línea con las fases del proyecto establecidas por la entidad:

La primera etapa se dedicó a la identificación de los flujos de datos que presentaban problemas de procesamiento. Se analizaron los cuellos de botella y las ineficiencias en el manejo de datos, con el fin de diseñar un proceso mejorado que optimizara estos flujos, asegurando una transformación y carga más efectiva, esta etapa se puede apreciar en la Figura 6.

FIGURA 6:

PRIMERA ETAPA PROYECTO: IDENTIFICACIÓN DE PROBLEMAS

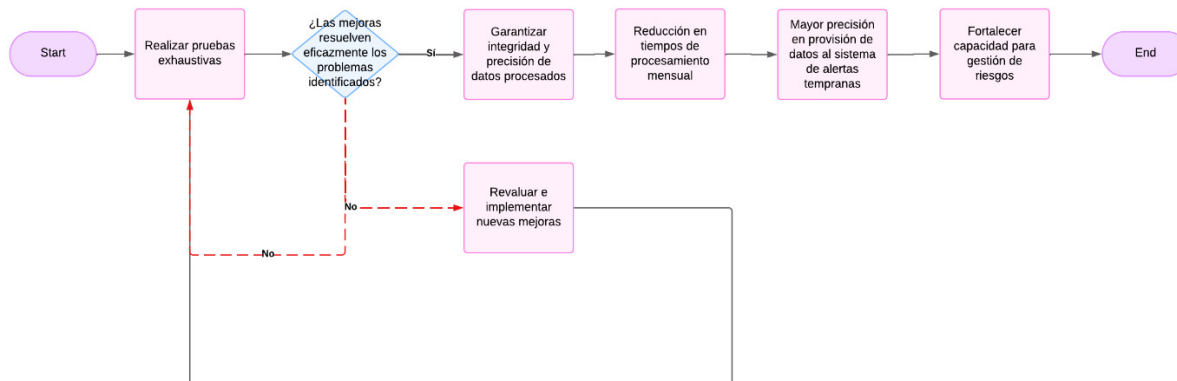


Nota: Preparado por el autor

En la segunda etapa, se implementó el proceso ETL optimizado y se llevaron a cabo pruebas exhaustivas. Estas pruebas fueron cruciales para verificar que las mejoras implementadas resolvieran eficazmente los problemas identificados en la etapa anterior, garantizando la integridad y precisión de los datos procesados. Como resultado, se logró una notable reducción en los tiempos de procesamiento mensual y una mayor precisión en la provisión de datos al sistema de alertas tempranas, fortaleciendo la capacidad de la entidad financiera para gestionar los riesgos de clientes de manera más eficiente, esta etapa se puede apreciar en la Figura 7.

FIGURA 7:

SEGUNDA ETAPA DEL PROYECTO: IMPLEMENTACIÓN DEL PROYECTO



Nota: Preparado por el autor

3.2.3.2 METODOLOGÍA

Para implementar el ETL se utilizó el método propio de la entidad financiera utilizando algunas prácticas de SCRUM, administrado por la herramienta JIRA. SCRUM es un marco de trabajo ágil que promueve la gestión eficiente de proyectos a través de ciclos de trabajo iterativos e incrementales llamados sprints, permitiendo ajustes rápidos en respuesta a la retroalimentación constante (Schwaber & Sutherland, 2020).

El uso de SCRUM ofrece varios beneficios, incluyendo una mejor comunicación y colaboración del equipo, una mayor adaptabilidad a los cambios y una entrega continua de valor al cliente. Estas características son particularmente útiles en proyectos tecnológicos

complejos como la optimización de un proceso ETL, donde la precisión y la capacidad de respuesta son esenciales (Rubin, 2012). La transparencia y las reuniones diarias (daily stand-ups) aseguraron que todos los miembros del equipo estuvieran alineados con los objetivos y pudieran resolver rápidamente cualquier impedimento, como se ve en la Figura 8.

En este proyecto, la implementación de SCRUM permitió una gestión flexible y eficiente, asegurando que las mejoras en el proceso ETL estuvieran alineadas con los objetivos estratégicos de la entidad financiera. Gracias a este marco de trabajo, se lograron reducir significativamente los tiempos de entrega y con ello se mejoró la precisión en el aprovisionamiento de datos para el sistema de alertas tempranas en un tiempo no prolongado de desarrollo, fortaleciendo así la capacidad de la entidad para anticipar y gestionar riesgos financieros (Cohn, 2010). La iteración constante y la revisión de cada sprint garantizaron que los requisitos del negocio se cumplieran a cabalidad, proporcionando un producto final que no solo automatizó procesos manuales sino que también sentó una base sólida para futuros desarrollos y mejoras.

FIGURA 8:

MARCO DE TRABAJO SCRUM DE LA ENTIDAD: DE TRABAJO EN EQUIPO CON LAS REUNIONES DIARIAS



Nota: Preparado por el autor

3.2.4 FUNDAMENTOS UTILIZADOS

3.2.4.1 METODOLOGÍAS ÁGILES

Las metodologías ágiles se han consolidado como un enfoque esencial en la gestión de proyectos, especialmente en el desarrollo de software. Esta transformación no es casualidad, sino una respuesta a la necesidad de adaptarse a un entorno empresarial dinámico y competitivo.

En primer lugar, la flexibilidad que proporcionan estas metodologías es altamente necesaria. A diferencia de los métodos tradicionales que siguen un plan rígido de principio a fin, las metodologías ágiles permiten ajustes continuos basados en la retroalimentación y cambios en el proyecto. Esto facilita que los equipos se adapten rápidamente a nueva información o requerimientos, incrementando así las posibilidades de éxito del proyecto (Bravo-Huivin et al., 2020).

Además, la colaboración es un pilar fundamental de las metodologías ágiles. Promueven la comunicación constante entre los miembros del equipo y los interesados, asegurando que todos estén alineados con los objetivos y trabajando de manera cohesionada, como lo vemos en la Figura 9. Esta interacción frecuente no solo mejora la transparencia del proyecto, sino que también fomenta un ambiente de trabajo más colaborativo y motivado (Bravo-Huivin et al., 2020).

Otro beneficio importante de las metodologías ágiles es la entrega incremental de valor para el negocio. En lugar de esperar hasta el final del proyecto para ver resultados, los equipos ágiles entregan funcionalidades completas y utilizables en ciclos cortos. Esto no solo permite a los interesados ver el progreso tangible del proyecto, sino que también facilita la identificación y resolución temprana de problemas (Bravo-Huivin et al., 2020).

Finalmente, la mejora continua es un principio esencial de las metodologías ágiles. Mediante revisiones y retrospectivas periódicas, los equipos pueden identificar áreas de mejora y ajustar sus procesos para ser más eficientes. Esta cultura de aprendizaje y adaptación constante asegura que los equipos no solo cumplan con sus objetivos, sino que también evolucionen continuamente (Bravo-Huivin et al., 2020).

FIGURA 9:

COMUNICACIÓN EN EL TRABAJO EN EQUIPO



Nota: Fuente: LinkedIn. Innovación: Metodologías ágiles en compras. 2018

3.2.4.2 KANBAN

Kanban es una metodología ágil de gestión de proyectos que se originó en el sistema de producción de Toyota, diseñado por Taiichi Ohno. Aunque inicialmente fue concebida como una técnica para la gestión de inventarios en el ámbito manufacturero, su aplicación se ha expandido a diversas industrias, incluyendo el desarrollo de software y la gestión de servicios, debido a su eficacia en la optimización de flujos de trabajo (Anderson, 2021).

Uno de los pilares principales de Kanban es la visualización del trabajo. Este principio se implementa mediante el uso de un tablero Kanban. Esta visualización en tablero permite a los equipos tener una clara percepción del estado actual de las tareas y facilita la comunicación y la colaboración (Medeiros dos Santos, Beltrão, de Souza, & Travassos, 2018).

La limitación del trabajo en proceso o por su siglas en inglés WIP (Work In Progress) es otro principio esencial de Kanban. Al restringir la cantidad de tareas que pueden estar en progreso simultáneamente, se evita la sobrecarga de trabajo y se promueve la finalización de

tareas antes de comenzar nuevas. Este enfoque ayuda a identificar cuellos de botella y mejora la eficiencia general del flujo de trabajo (Rosin, 2021).

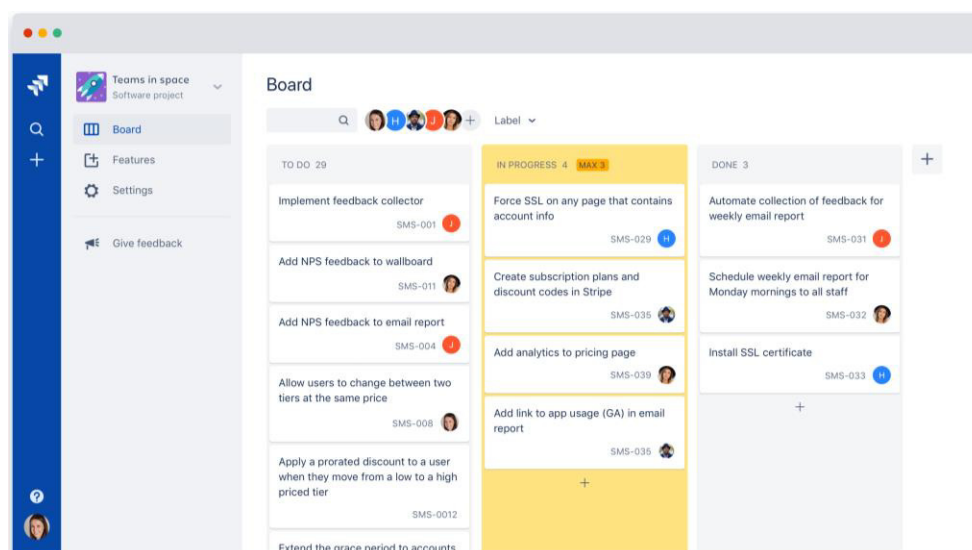
Kanban también se enfoca en la gestión del flujo, lo que implica monitorear y optimizar continuamente el movimiento de tareas a través del sistema. Al reducir el tiempo de ciclo y aumentar la previsibilidad en las entregas, los equipos pueden mejorar su rendimiento y cumplir con los plazos establecidos (Anderson, 2021).

La retroalimentación regular es fundamental en Kanban. A través de reuniones diarias y revisiones periódicas, los equipos pueden discutir problemas, compartir aprendizajes y planificar mejoras, fomentando un ciclo de mejora continua (Medeiros dos Santos et al., 2018).

En resumen, Kanban es una metodología flexible y adaptable que puede mejorar significativamente la gestión de proyectos y el desarrollo ágil. Al visualizar el trabajo, como lo vemos en la Figura 10, limitar el trabajo en proceso, gestionar el flujo, establecer políticas claras y fomentar la retroalimentación y la mejora continua, los equipos pueden lograr una mayor eficiencia, productividad y satisfacción en sus proyectos (Anderson, 2021; Medeiros dos Santos et al., 2018; Rosin, 2021).

FIGURA 10:

VISUALIZACIÓN DE TRABAJO EN KANBAN



Nota: Fuente: Atlassian. ¿Qué es un tablero de kanban?

3.2.4.3 MEDOLOGÍA EN CASCADA

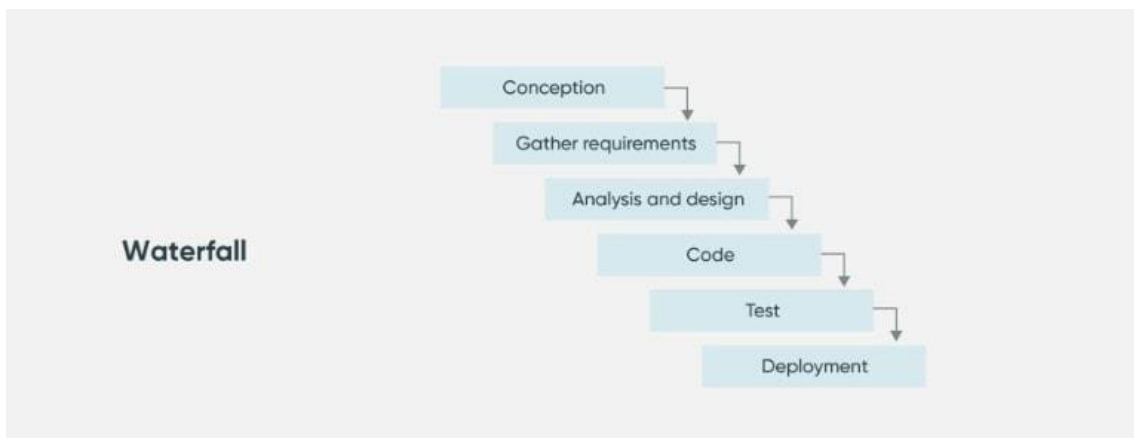
La metodología en cascada, también conocida en inglés como "Waterfall", es un enfoque tradicional en la gestión de proyectos que se caracteriza por su estructura lineal y secuencial. Este modelo, popularizado en los años 70, sigue una serie de fases bien definidas: recopilación de requisitos, diseño, implementación, verificación y mantenimiento (Balaji & Murugaiyan, 2021). Cada fase debe completarse antes de que la siguiente pueda comenzar, lo que proporciona un control riguroso del progreso y una clara definición de tareas y responsabilidades.

Una de las principales ventajas de la metodología en cascada es su claridad y estructura, como vemos en la Figura 11, lo que facilita la gestión del proyecto y la documentación detallada en cada etapa (Mishra & Dubey, 2020). Sin embargo, esta rigidez puede ser una desventaja en entornos de cambio rápido, ya que dificulta la adaptación a nuevos requisitos o cambios durante el desarrollo del proyecto (Royce, 2022).

En resumen, la metodología en cascada es adecuada para proyectos con requisitos bien definidos y estables, donde la previsibilidad y el control son críticos. Aunque su enfoque lineal puede ser una limitación en proyectos más dinámicos, su estructura clara y secuencial ofrece ventajas en términos de gestión y documentación (Rastogi et al., 2021).

FIGURA 11:

ESTRUCTURA DE LA METODOLOGÍA EN CASCADA



Nota: Fuente: ServiceNow. ¿Cuál es la diferencia entre las metodologías Agile y en cascada?

3.2.4.4 PROGRAMACIÓN EXTREMA

La programación extrema (XP) y la metodología en cascada son enfoques distintos en la gestión de proyectos de software. XP, desarrollada por Kent Beck, es una metodología ágil que enfatiza la colaboración, la comunicación constante y la retroalimentación rápida. Sus principios incluyen la simplicidad, el coraje para realizar cambios necesarios y el respeto mutuo entre los miembros del equipo. Las prácticas clave de XP son el desarrollo iterativo, la programación en pareja, las pruebas automatizadas, la refactorización continua y la integración continua. Estas prácticas permiten una rápida adaptación a los cambios y aseguran una alta calidad del software (Beck, 2004; Fowler, 2019).

Por el contraparte, la metodología en cascada es un enfoque lineal y secuencial popularizado en los años 70. Este se divide en fases bien estructuradas: requisitos, diseño, implementación, verificación y mantenimiento. Cada fase debe completarse antes de pasar a la siguiente, lo que proporciona claridad y estructura. Sin embargo, su rigidez puede dificultar la adaptación a cambios durante el desarrollo (Balaji & Murugaiyan, 2021; Mishra & Dubey, 2020).

Ambas metodologías tienen sus beneficios y desventajas. La cascada es adecuada para proyectos con requisitos bien definidos y estables, mientras que XP es ideal para entornos dinámicos donde la adaptabilidad y la calidad del software son cruciales. Para elegir entre ambas se debe ver las necesidades específicas del proyecto y del equipo.

3.2.4.5 ETL

El proceso de Extracción, Transformación y Carga (ETL) es crucial en la gestión contemporánea de datos. Este proceso permite integrar información de múltiples fuentes en un solo repositorio centralizado, lo cual es vital para la inteligencia empresarial y aplicaciones avanzadas como el análisis de datos y el aprendizaje automático (IBM, 2023). Al consolidar datos de diversas plataformas, el ETL garantiza que la información esté disponible de manera coherente y accesible, facilitando la toma de decisiones estratégicas, el flujo se visualiza en la Figura 12.

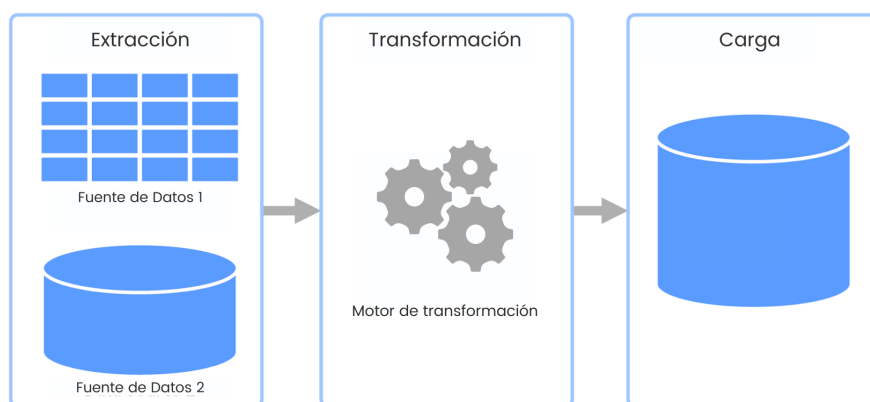
El ETL también juega un papel fundamental en la limpieza y normalización de datos, eliminando duplicados, corrigiendo errores y unificando formatos. Esto resulta en una mejora

sustancial de la calidad de los datos, proporcionando una visión precisa y completa de la realidad empresarial (Astera, 2024). La correcta organización de los datos es esencial para generar informes exactos y análisis confiables, lo que a su vez mejora la eficiencia operativa y la capacidad de respuesta de la organización.

Además, el ETL permite procesar grandes volúmenes de datos de manera eficiente. La automatización de las etapas de extracción, transformación y carga ahorra tiempo y recursos, reduciendo la carga de trabajo manual y minimizando el riesgo de errores humanos. Esto libera al equipo para centrarse en tareas de mayor valor, como la interpretación de datos y la toma de decisiones basadas en información precisa (Bravo-Huivin, Cieza-Mostacero, Flores-Rodriguez, & Uceda-Davila, 2020).

FIGURA 12:

ARQUITECTURA DE UN PROCESO ETL



Nota: Fuente: Algoritmia. La importancia de un proceso ETL en el Business Intelligence.

2022

3.2.4.6 JIRA

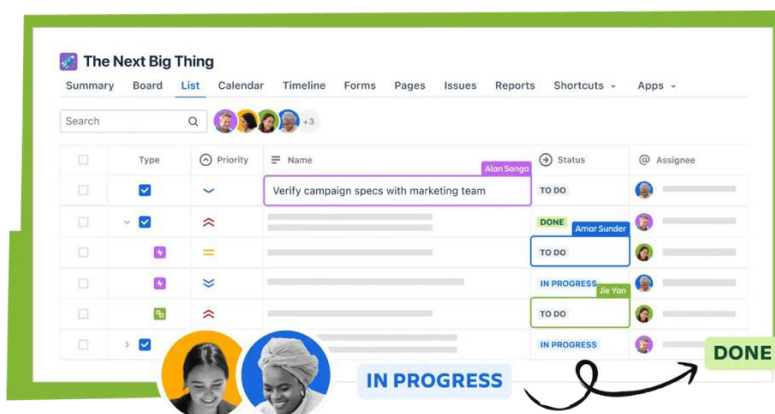
Jira es una herramienta desarrollada por Atlassian que ha transformado la gestión de proyectos en la industria del software. Su propósito inicial era el seguimiento de errores, pero con el tiempo ha evolucionado para convertirse en una plataforma integral que facilita la planificación y administración de proyectos ágiles. Gracias a su adaptabilidad, Jira permite a los equipos de trabajo implementar metodologías ágiles como Scrum y Kanban, optimizando la organización y priorización de tareas (Smith & Johnson, 2021).

La importancia de Jira radica en varios aspectos. En primer lugar, promueve la colaboración efectiva entre los miembros del equipo, ofreciendo una visión clara y compartida del progreso de las tareas y del estado de los proyectos. Además, su capacidad de personalización permite ajustar flujos de trabajo, campos y pantallas según las necesidades específicas de cada proyecto, lo que lo hace extremadamente versátil (Williams, 2022). Jira también se integra con una amplia gama de herramientas y servicios, proporcionando una gestión de proyectos más coherente y unificada.

Su uso en el presente proyecto es crucial para manejar las tareas y cada estado de las mismas, donde el movimiento de estados indica el avance del proyecto y los tiempos por cada estado son indicadores que muestran o alertan posibles retrasos o alertas para los entregables. También que su uso es necesario para poder seguir el flujo de despliegue, porque está asociado al repositorio de los proyectos y estos van a generar versionamiento del aplicativo que permite la generación de artefactos para poder realizar pases a entornos previos y entornos productivos.

FIGURA 13:

FLUJO DE TAREAS EN JIRA



Nota: Fuente: Atlassian. Jira Software – Features

Por último, pero no menos importante, Jira ofrece herramientas avanzadas de reporte y análisis, que permiten a los equipos monitorear su desempeño y identificar áreas de mejora continua, tenemos una visualización en la Figura 13. En resumen, Jira es una herramienta

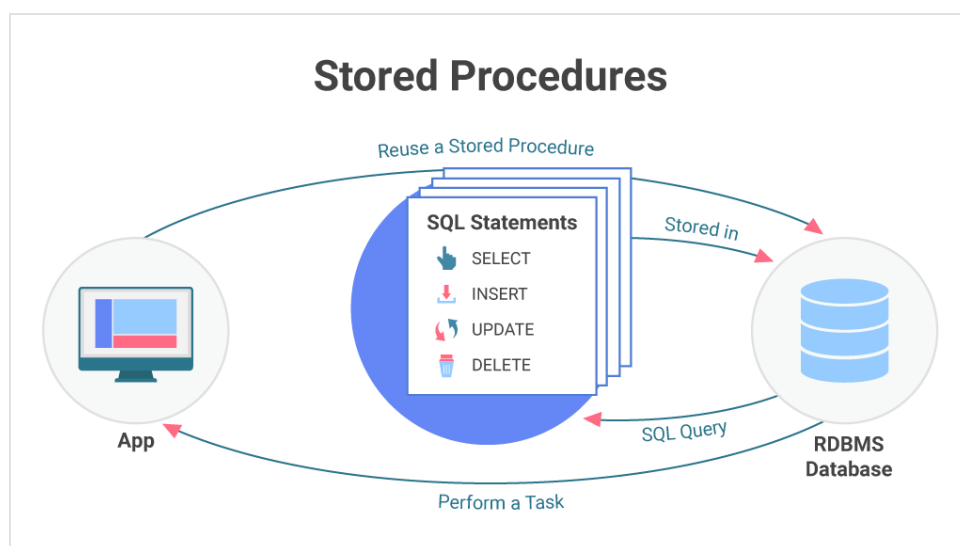
esencial para cualquier equipo de desarrollo de software que busque mejorar su eficiencia y alcanzar sus objetivos de manera efectiva (Atlassian, n.d.).

3.2.4.7 PROCEDIMIENTOS ALMACENADOS

Los procedimientos almacenados son conjuntos de instrucciones SQL que se guardan en una base de datos y que pueden ser ejecutados por el sistema de gestión de bases de datos (SGBD). Estos procedimientos permiten realizar tareas complejas y recurrentes de manera eficiente, ya que se ejecutan directamente en el servidor de la base de datos, lo que reduce la cantidad de datos transferidos entre el servidor y el cliente, mejorando así el rendimiento y reduciendo la latencia en aplicaciones con grandes volúmenes de datos (Ramakrishnan & Gehrke, 2022). Además, al almacenar la lógica de negocio en procedimientos almacenados, se facilita la reutilización del mismo código en diferentes partes de una aplicación o incluso en diferentes aplicaciones, lo vemos en la Figura 14, lo que simplifica el mantenimiento y la actualización del código (Elmasri & Navathe, 2021).

FIGURA 14:

FLUJO DE UN PROCEDIMIENTO ALMACENADO



Nota: Fuente: Revealbi. Uso de procedimientos almacenados en Análisis integrado

En términos de seguridad, los procedimientos almacenados ofrecen una capa adicional de control sobre el acceso a los datos. Al encapsular las operaciones de la base de

datos en procedimientos, se puede restringir el acceso directo a las tablas y permitir que las operaciones se realicen únicamente a través de estos procedimientos, lo que ayuda a prevenir accesos no autorizados y a mantener la integridad de los datos (Ramakrishnan & Gehrke, 2022). Además, los procedimientos almacenados permiten implementar políticas de seguridad y auditoría más robustas, ya que todas las operaciones sobre los datos pueden ser monitoreadas y registradas de manera centralizada (Elmasri & Navathe, 2021).

En términos de mantenimiento, los procedimientos almacenados permiten una mayor modularidad y organización del código SQL. Al tener la lógica de negocio encapsulada en procedimientos almacenados, es más fácil realizar cambios y actualizaciones sin tener que modificar múltiples aplicaciones. Esto no solo simplifica el proceso de mantenimiento, sino que también reduce el riesgo de introducir errores al actualizar el código (Ramakrishnan & Gehrke, 2022).

Otro aspecto importante es la portabilidad. Los procedimientos almacenados pueden ser movidos y reutilizados entre diferentes bases de datos y entornos de desarrollo, lo que es especialmente útil en entornos de desarrollo y producción donde se necesita mantener la coherencia del código y la lógica de negocio a través de múltiples sistemas y versiones de bases de datos (Ramakrishnan & Gehrke, 2022).

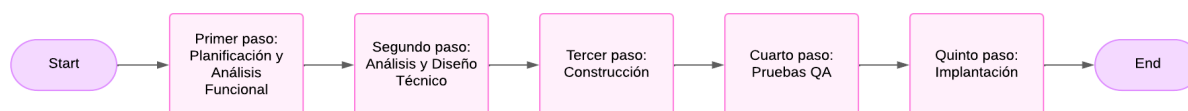
Finalmente, los procedimientos almacenados permiten la automatización de tareas rutinarias y repetitivas, como la generación de informes, la sincronización de datos y la validación de reglas de negocio. Esto no solo ahorra tiempo y esfuerzo, sino que también asegura que las tareas se realicen de manera consistente y precisa (Ramakrishnan & Gehrke, 2022).

3.2.5 IMPLEMENTACIÓN DE LAS ÁREAS, PROCESOS, SISTEMAS Y BUENAS PRÁCTICAS

Para este desarrollo del proyecto se siguió las etapas de gestión de proyectos establecidas por la entidad de acuerdo a sus estándares y se visualiza en la Figura 15.

FIGURA 15:

ETAPAS DE LA GESTIÓN DEL PROYECTO



Nota: Preparado por el autor

En cada entregable a desarrollar se emplearon las fases indicadas en la Tabla 9.

TABLA 9:

FASES DEL PROYECTO

FASE	DESCRIPCIÓN
Planificación y Análisis Funcional	Se definió los componentes previamente usados para validaciones, a crear modificar y se diseña el documento funcional del entregable.
Diseño Técnico	Se diseñó la propuesta técnica del entregable.
Construcción	Se desarrolló la solución técnica.
Pruebas QA	Se realizó las pruebas al proceso acordes a los lineamientos ya establecidos por la entidad.
Despliegue productivo o Implantación	Se gestiona la instalación en el ambiente productivo de la entidad.

Nota: Preparado por el autor

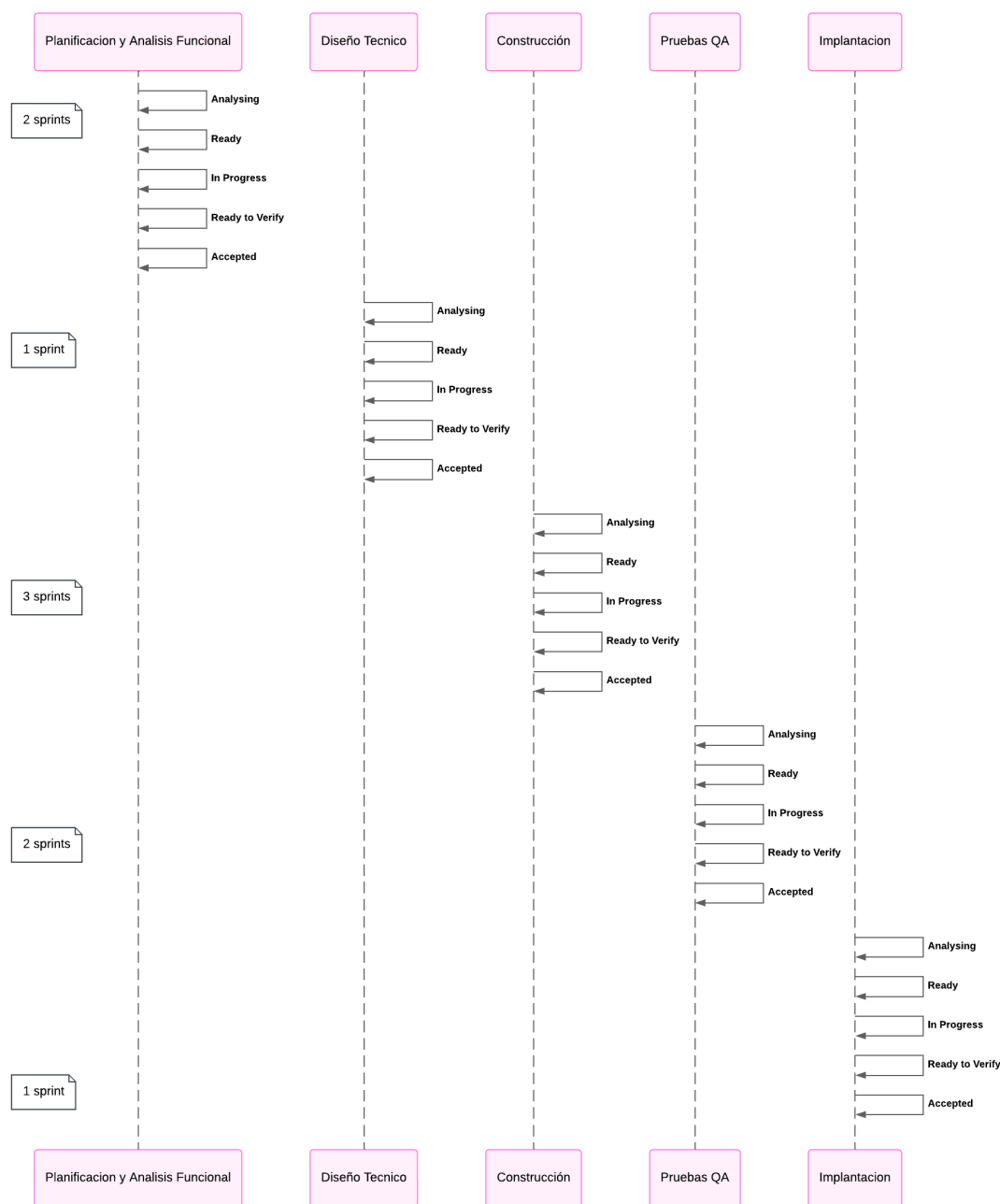
3.2.5.1 USO DE LA METODOLOGÍA

En el proyecto se usó una metodología ágil híbrida propia de la entidad, usando principalmente SCRUM que es un metodo que otorga desarrollos basados en sprints, frecuentemente usada en los proyectos de TI actualmente.

En este proyecto también se usó la herramienta JIRA, la cual administró las tareas creadas en cada sprint del proyecto, el cual conllevó ocho sprints, donde cada sprint dura dos semanas. La secuencia de trabajo se indica en la Figura 16.

FIGURA 16:

SECUENCIA DE TRABAJO EN JIRA



Nota: Preparado por el autor de acuerdo al uso de la herramienta de JIRA por la entidad

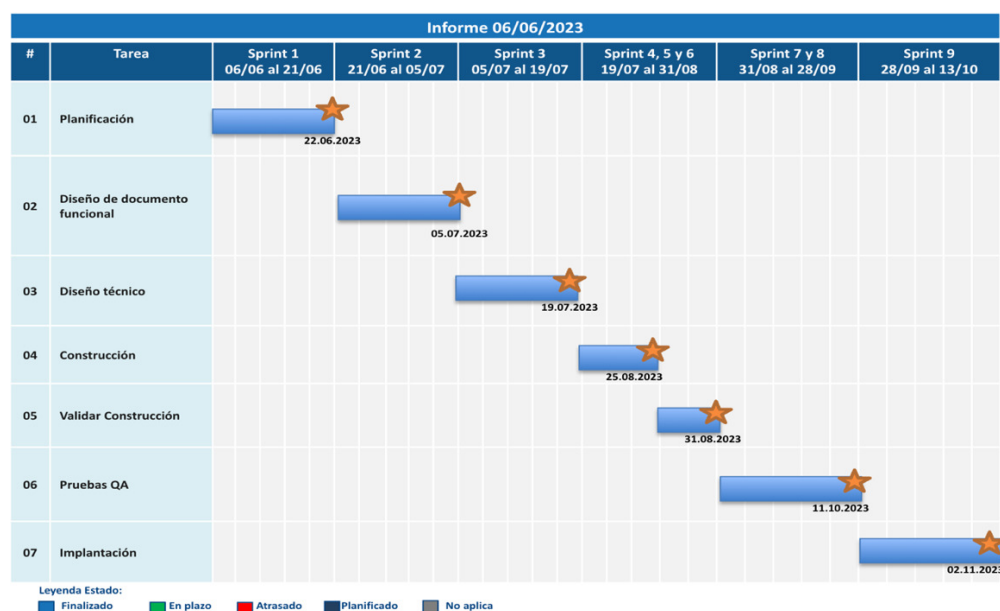
3.2.5.1.1 PLANIFICACIÓN Y ANÁLISIS FUNCIONAL

SPRINT 0

En el Sprint 0, el primer paso fue realizar un diagnóstico exhaustivo del estado actual del proceso ETL para identificar los cuellos de botella y áreas de mejora. Encontrando que la causa de la lentitud es debido a procesos manuales de validación y reintentos completos por cada error. A continuación, se definieron los requerimientos técnicos necesarios para optimizar el proceso, incluyendo especificaciones para la implementación de nuevas herramientas de administración de cargas de trabajo. Seguidamente, se desarrollará un plan detallado del rediseño del proceso ETL, que incluirá el cronograma, como lo vemos en la Figura 17, recursos necesarios y responsables de cada tarea. Además, se configurarán y validarán las nuevas herramientas de trabajo, asegurando que todos los miembros del equipo estén capacitados para utilizarlas eficientemente. Para medir la eficacia del nuevo proceso ETL, se establecerán métricas de desempeño claras, alineadas con los objetivos de reducción de tiempos y costos.

FIGURA 17:

CRONOGRAMA DE PLANIFICACIÓN DEL PROYECTO



Nota: Elaborado por el autor. Se muestra la separación de semanas y las tareas programadas por sprint, la estrella y fecha indican la finalización

Se definieron los siguientes roles y responsabilidades, visualizados en la Tabla 10:

TABLA 10:

ROLES DEDICADOS AL PROYECTO

RECURSO	DEDICACIÓN	DESCRIPCIÓN
Product Owner	Tiempo Completo	Define la visión del proyecto
Analista Programador Java	Tiempo Completo	Desarrolla las funcionalidades
Analista Programador Java	Tiempo Parcial	Desarrolla las funcionalidades
Analista QA	Tiempo Completo	Asegura la calidad de las entregas

Nota: Elaborado por el autor

Cuando se trata de roles, diferimos entre:

- El Product Owner es el representante del cliente final encargado de gestionar el producto o proyecto. Actúa como un intermediario clave entre los clientes y el equipo de desarrollo. Sus responsabilidades incluyen definir y priorizar las características del producto, establecer las fechas de entrega y evaluar los resultados. El Product Owner tiene la autoridad para aceptar o rechazar los entregables que se le presentan, asegurándose de que el producto final cumpla con las expectativas y requisitos del cliente.
- El Scrum Master es el facilitador del equipo de desarrollo, asegurándose de que todos los miembros del equipo comprendan y sigan la metodología Scrum. El Scrum Master se encarga de eliminar obstáculos que puedan impedir el progreso del equipo, promover un entorno de trabajo productivo y garantizar que el equipo alcance sus objetivos. Además, el Scrum Master educa al equipo sobre las prácticas ágiles y actúa como un enlace entre el equipo y las partes interesadas externas.
- Equipo de Desarrollo que está compuesto por los profesionales responsables de llevar a cabo el trabajo necesario para completar el proyecto. El equipo de

desarrollo es multifuncional y autoorganizado, lo que significa que los miembros del equipo colaboran estrechamente, comparten responsabilidades y se apoyan mutuamente para lograr los objetivos del proyecto. Las decisiones técnicas y de implementación son tomadas colectivamente por el equipo de desarrollo.

Estos roles no implican jerarquía ni autoridad, sino que definen las responsabilidades dentro del proceso. En relación a los artefactos, podemos distinguir:

- El Product Backlog es una lista creada en el momento en que se hace el pedido, con la participación del propietario del producto. Este documento registra todas las funcionalidades del proyecto y las organiza en orden de prioridad.
- El Sprint Backlog es la lista que incluye características seleccionadas del conjunto completo de elementos registrados en el Product Backlog, que se desarrollarán durante la Fase 1 o Sprint. La duración del sprint está determinada. El equipo es responsable de seleccionar estas características y siempre sigue las instrucciones del Product Owner al trabajar con el Product Backlog.

Respecto a las reuniones distinguimos entre:

- La Planificación del Sprint o Sprint Planning conlleva un tablero de Sprint Backlog claro y bien definido, con el cual el equipo se compromete. En esta reunión participa el Product Owner.
- Las Reunión Diaria o Daily Meeting se realizan encuentros diarios para discutir el progreso del proyecto. Se abordan desafíos, dificultades y estrategias para avanzar en el proyecto.
- La Revisión del Sprint o Sprint Review se lleva a cabo una revisión informal de los resultados del sprint con el Product Owner u otras personas que puedan aportar valor. El objetivo es mejorar el contenido del estudio.

- La Retrospectiva del Sprint o Sprint Retrospective consiste en que al final de cada sprint, el equipo se reúne para analizar su trabajo y buscar formas de mejorarlo. El objetivo principal es la mejora continua del desempeño del equipo.

Una vez completados el sprint, la revisión y la retrospectiva, se comienza a preparar el siguiente sprint. Este ciclo de trabajo se repite tantas veces como sea necesario hasta la fecha de entrega o hasta agotar el presupuesto.

Las herramientas utilizadas para el trabajo del proyecto y los entornos de trabajo, lo vemos en la Tabla 11.

TABLA 11:

HERRAMIENTAS Y ENTORNOS DE TRABAJO PARA EL PROYECTO

HERRAMIENTA	USO
Jira	Gestión de tareas y seguimiento
Git	Control de versiones
Google Chat	Comunicación del equipo

Nota: Preparado por el autor

SPRINT 1

En este sprint lo que se llevó a cabo fue que el Product Owner en compañía del equipo técnico obtuvieron toda la información del proceso actual que se tenía, las validaciones que se realizaban y la lista de archivos necesarios que fue crucial para poder identificar cuales eran usados y cuales no, los archivos recopilados necesarios se detallan en la Tabla 12.

TABLA 12:

ARCHIVOS NECESARIOS DE ENTRADA PARA EL PROCESO ETL

ARCHIVO	ORIGEN	PERIODICIDAD ACTUALIZACIÓN
H_XX_XXDQCRXX.BIN	Tarjetas de Crédito	Mensual
H_XX_XXDQDEXX.BIN	Tarjetas de Débito	Mensual

H_XX_ARCHCXX.BIN	Contratos Tarjetas	Mensual
H_XX_ARCHTXX.BIN	Tarjetas	Mensual
H_XX_CCC0IMXX.BIN	Cartera de deuda fuera de fecha de vencimiento	Mensual
H_XX_FRGXX.BIN	Pasivos	Mensual
H_XX_SBDQRCXX.BIN	Registro Crediticio del Cliente en el sistema financiero	Mensual
H_XX_ENTMOXX.BIN	Recuperaciones	Mensual
H_XX_XGCLIEXX.BIN	Rating clientes	Mensual
H_XX_XGVALOXX.BIN	Estado Financieros	Mensual
H_XX_XGRATIXX.BIN	Ratios	Mensual
H_XX_XGEJERXX.BIN	Ejercicios rating	Mensual
H_XX_RVGTDX.BIN	Evaluaciones	Mensual
H_XX_MPTAR0XX.BIN	Tarjetas	Mensual
H_XX_RCDTRXX.BIN	Deuda de colocación mayor que cero	Mensual
H_XX_CKMCLXX.BIN	Maestro de Clasificación del Cliente	Mensual
H_XX_BURONAXX.BIN	Naturales	Diario
H_XX_BUROJUXX.BIN	Jurídicos	Diario
H_XX_SUNATXX.txt	Sunat	Mensual
H_XX_CRDT0XX.BIN	Maestro de Contratos	Mensual
H_XX_CRDT0XX.BIN	Vinculación de Contratos	Mensual
H_XX_CRDT0XX.BIN	Hipotecario	Mensual
H_XX_RCDIXX.BIN	Impagados	Mensual
H_XX_SLECRXX.BIN	Mis Productos	Diario
H_XX_ITDTJXX.BIN	Motor Estados Financieros	Mensual

H_XX_ITDTBXX.BIN	Motor Estados Financieros	Mensual
H_XX_LA3CCNXX.BIN	Préstamos	Mensual
H_XX_LA3CLFXX.BIN	Préstamos	Mensual
H_XX_LA3CLFXX.BIN	Préstamos	Mensual
H_XX_LA3CLFXX.BIN	Préstamos	Mensual
H_XX_ABDTMXX.txt	SalDOS/Cuotas	Mensual

Nota: Preparado por el autor

SPRINT 2

El objetivo de este sprint fue poder identificar cuales son las validaciones necesarias que se tenían actualmente y esta información recopilada en la Tabla 13.

TABLA 13:

VALIDACIONES PRINCIPALES AL PROCESO

VALIDACIÓN	DESCRIPCIÓN	ETAPA
Validación de existencia de data	Confirmar que los archivos requeridos existen en el servidor de carga.	Extracción
Validación de fechas	Verificar que los archivos contengan como fecha de calculo la fecha correspondiente al proceso por ejecutar	Extracción
Validación de formato	Asegurar que la información de los archivos cumplen con el formato, asociado a la tabla externa.	Extracción
Validación de integridad referencial	Verificar que las relaciones entre la data de los archivos están relacionadas o preservadas después de la transformación.	Transformación
Validación de duplicados	Asegurar que no hay registros duplicados en el destino que es la base de datos oracle	Carga

Nota: Elaborado por el autor

Con esto obtenido, más las necesidades propias del usuario se preparó y elaboró un documento funcional que brinda el detalle de qué se va a desarrollar y cómo se plantea las mejoras al proceso. Este documento es compartido con el usuario de riesgos y con el equipo técnico para poder ver viabilidad y conformidad. Tanto el usuario de riesgos como el equipo técnico tienen que dar sus comentarios y al final su conformidad, la cual es necesaria para poder continuar a la siguiente etapa del sprint.

3.2.5.1.2 DISEÑO TÉCNICO

SPRINT 3

En el tercer sprint tiene como objetivo poder diseñar la propuesta técnica de acuerdo al documento funcional, para este fin se desarrollaron reuniones con el equipo técnico y los responsables de arquitectura.

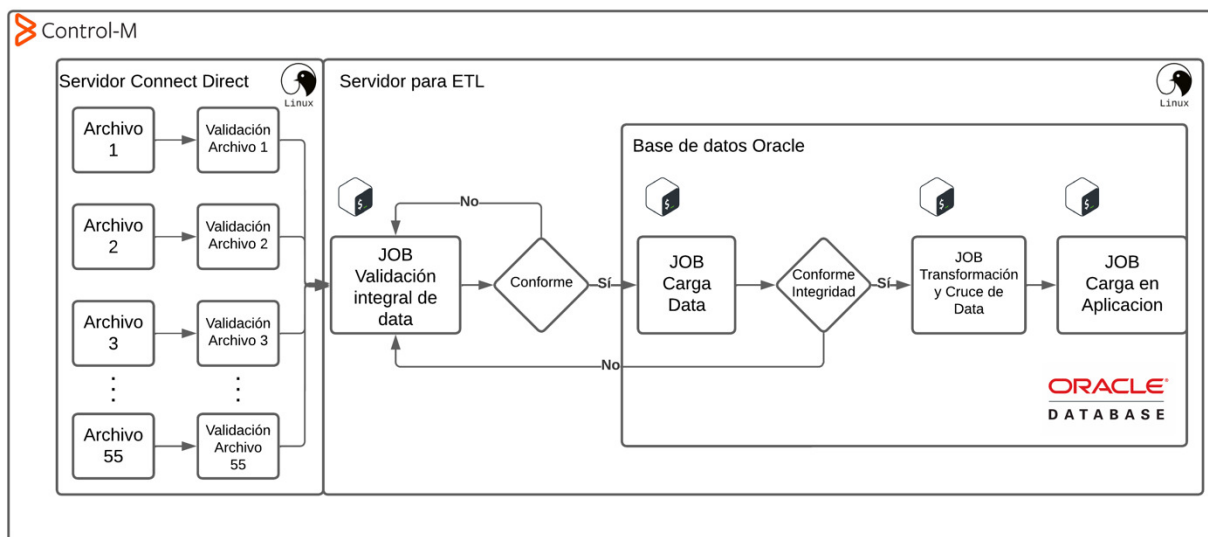
Esta propuesta contempla las herramientas que van a ser utilizadas de acuerdo a los estándares en tecnología de la entidad financiera, así como jobs de ejecución de acuerdo a las correctas horas por cargas de trabajo de la red de oficinas y red comercial. En la la Figura 18 se detalla el diagrama de arquitectura, se van a diseñar los jobs de validación y decargas de trabajo utilizando la herramienta Control M, en esta herramienta se puede seleccionar el servidor donde se realizarán los jobs. El servidor connect direct, que es en base a linux, es un servidor común entre varios aplicativos donde se realizan validaciones de existencia y copias de archivos que son cargas básicas, se va a realizar la existencia de los archivos en el servidor origen, la lista de archivos se encuentran en la Tabla 10, al encontrar existencia del archivo se enviará una notificación vía correo electrónico para notificar del archivo correctamente subido. El siguiente paso es la validación del archivo que indicará la integridad y fecha correcta del archivo con las funcionalidades propias de la herramienta Control M. Una vez se valide todos los archivos necesarios, se enviará la notificación vía correo electrónico de validación conforme y se lanzará el job diario con bash en linux que verificará que todos los archivos hayan sido validados, comenzará el job de validación integral de la data de los archivos.

Si la validación no es conforme, se volverán a subir los archivos con problemas resolviendo su casuística de error, regresando a la validación integral de la data y continuando

el proceso. El job de Carga Data pasará todas las tablas externas asociadas a los archivos input a tablas indexadas para su procesamiento más rápido. En caso falle por formatos incorrectos en la tabla se volverá a cargar los archivos correctamente o se revisará el motivo del cambio en el formato del mismo y ejecutará al siguiente día el job Validación integral de data.

Al confirmar la integridad de la Carga Data, seguirá el siguiente job de Transformación y Cruce de Data donde cruzará toda la información de las tablas indexadas de acuerdo a las reglas negocio que ya se encontraban previamente establecidas. Luego con la finalización de este proceso, se procede el job de Carga en Aplicación que cargará toda la información transformada y cruzada en el aplicativo final, donde se visualizará el modelo de alertas tempranas.

FIGURA 18:
DIAGRAMA DE ARQUITECTURA



Nota: Elaborado por el autor de acuerdo a estándares de la entidad

3.2.5.1.3 CONSTRUCCIÓN

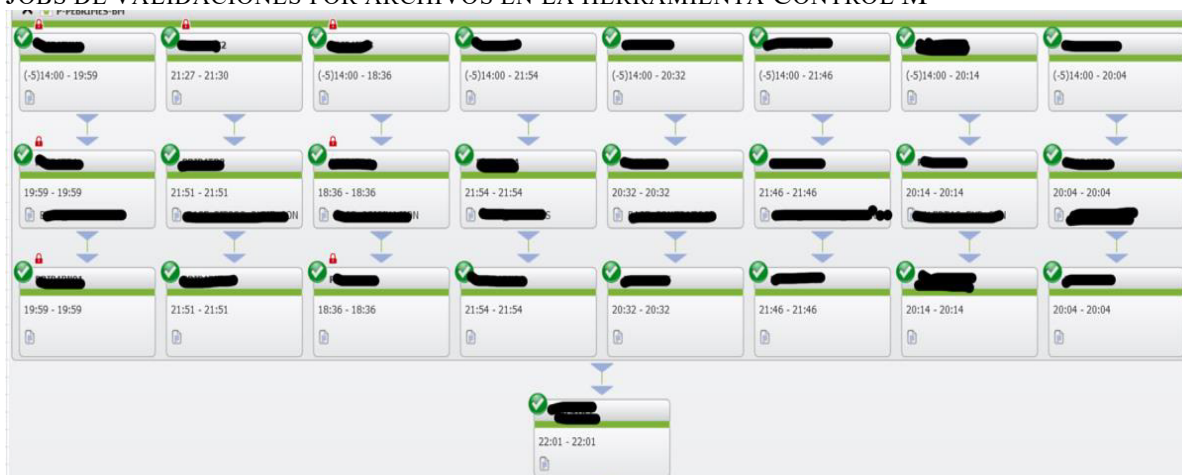
La etapa de construcción se dividió en 3 sprints, donde se llevaron a cabo los desarrollos de las partes de validaciones, ejecuciones y reintentos.

SPRINT 4

En el cuarto sprint del proyecto el objetivo fue construir las validaciones de los archivos que consiste en jobs asociados a cada archivo input donde al encontrar un archivo va al siguiente paso de validar la integridad del archivo y en caso sea exitoso quedaría en espera a que todas las validaciones sean exitosas para continuar al siguiente paso, en la Figura 19 vemos una muestra con un bloque de archivos donde se realizaron las validaciones, esta es la visualización gráfica de Control M, donde cada recuadro son los jobs, se muestra la hora de ejecución y el flujo que siguen, como se puede observar cada recuadro de validación está asociado al archivo input.

FIGURA 19:

JOBS DE VALIDACIONES POR ARCHIVOS EN LA HERRAMIENTA CONTROL M



Nota: Elaborado por el autor de acuerdo a estándares de la entidad

SPRINT 5

En el quinto sprint se tuvo como objetivo desarrollar los jobs de ejecución de procesos propios de negocio, primero se definieron las cargas, que consistían en consultar las tablas externas asociadas a los archivos input y cargarlos en tablas indexadas, esto a través de stored procedure o procedimientos almacenados que fueron creados con código PLSQL y ejecutados en la base de datos oracle intermediaria, así para cada archivo input necesario para el proceso dentro de un procedimiento almacenado orquestador, como podemos ver en la Figura 20.

FIGURA 20:
CARGAR DATA EN TABLAS INDEXADAS

[illegible]

Nota: Elaborado por el autor de acuerdo a estándares de la entidad

En la Figura 21 vemos como se contruyeron estas cargas desde tablas externas a tablas indexadas, considerando la fecha del proceso y cada tamaño de columna.

FIGURA 21:
CARGA DE DATOS

```

INSERT /*+ APPEND NOLOGGING */
INTO S [REDACTED]_BIN
(
  SELECT
    V [REDACTED] AS FECHA_PROCESO,
    trim(SUBSTR(DATOS, 0, 10)) AS NO [REDACTED],
    trim(SUBSTR(DATOS, 11, 5)) AS NA [REDACTED],
    trim(SUBSTR(DATOS, 16, 5)) AS NM [REDACTED],
    trim(SUBSTR(DATOS, 21, 1)) AS CT [REDACTED],
    trim(SUBSTR(DATOS, 22, 5)) AS NN [REDACTED],
    trim(SUBSTR(DATOS, 27, 5)) AS NU [REDACTED],
    trim(SUBSTR(DATOS, 32, 3)) AS CD [REDACTED],
    trim(SUBSTR(DATOS, 35, 1)) AS CC [REDACTED],
    trim(SUBSTR(DATOS, 36, 5)) AS NE [REDACTED],
    trim(SUBSTR(DATOS, 41, 5)) AS NI [REDACTED],
    trim(SUBSTR(DATOS, 46, 1)) AS CS [REDACTED],
    trim(SUBSTR(DATOS, 47, 60)) AS C [REDACTED],
    trim(SUBSTR(DATOS, 107, 5)) AS C [REDACTED],
    trim(SUBSTR(DATOS, 112, 10)) AS [REDACTED],
    trim(SUBSTR(DATOS, 122, 17)) AS [REDACTED],
    trim(SUBSTR(DATOS, 139, 2)) AS C [REDACTED],
    trim(SUBSTR(DATOS, 141, 10)) AS [REDACTED],
    trim(SUBSTR(DATOS, 151, 8)) AS C [REDACTED],
    trim(SUBSTR(DATOS, 159, 5)) AS N [REDACTED],
    trim(SUBSTR(DATOS, 164, 5)) AS N [REDACTED],
    FROM EXT I [REDACTED]
  WHERE l=1
);

```

Nota: Elaborado por el autor de acuerdo a estándares de la entidad

Luego se definieron las transformaciones y cruce de data que al igual que el job anterior son procedimientos almacenados que se ejecutan en la base de datos donde cargan tablas a través de lógica de negocio, en la Figura 22 podemos ver una ejemplo de una carga a una tabla incluyendo filtros de clientes.

FIGURA 22:

CARGA Y TRANSFORMACIÓN Y CRUCE DE DATA

```

INSERT /*+ APPEND NOLOGGING */
INTO C [REDACTED]
SELECT
  F [REDACTED] PROCESO AS FECHA_P [REDACTED]
  S [REDACTED] (NCCLIEN,2,8) AS C [REDACTED] TRAL,
  N [REDACTED] ER AS AN [REDACTED] EERCICIO,
  N [REDACTED] ER AS ME [REDACTED] ERICIO,
  C [REDACTED] A AS DIV [REDACTED],
  C [REDACTED] EI AS TI [REDACTED],
  N [REDACTED] NCE AS N [REDACTED],
  S [REDACTED] (CDESSE [REDACTED]) AS T [REDACTED]
CASE
  WHEN N [REDACTED] = '4001' THEN 3
  WHEN N [REDACTED] IN ('3000','3001','3003','3004') THEN 2
  WHEN N [REDACTED] IN ('8000','8001','9998','9999') THEN 1
  ELSE 0
END AS FLAG
FROM ST [REDACTED]
WHERE F [REDACTED] = listaPeriodo(i).PERIODO
AND N [REDACTED] NOT IN ('3020','4020')
AND T [REDACTED] (PER) >= (SUBSTR(V [REDACTED] DO,1,4))-3
AND NMESEJER>2;
SP_WRITE_LOG('Registros a Insertar : '|| SQL%ROWCOUNT);

```

Nota: Elaborado por el autor de acuerdo a estándares de la entidad

SPRINT 6

En el sexto sprint, el objetivo fue definir la lógica de reintentos, que consistía en acordar que jobs son los que van a realizar los reintentos con el objetivo de evitar reprocesos totales, esto lo podemos ver en el diagrama de arquitectura de la Figura 18. Otra actividad fue también definir qué tipo de errores están asociados a los reintentos, los errores acordados se pueden ver en la Tabla 14.

TABLA 14:

ERRORES CONSIDERADOS PARA REINTENTOS

ERROR	IMPACTO	CONLLEVA REINTENTO
Formato Incorrecto	La tabla externa genera error	Sí

Fecha errónea	El proceso no lo va a incluir	Sí
Registros duplicados	El proceso daría error	Sí

Nota: Elaborado por el autor

3.2.5.1.4 PRUEBAS QA

Para las pruebas se contemplaron dos sprints, donde realizaron pruebas funcionales al flujo de validación, así como al flujo de carga. Estas pruebas fueron realizadas en entorno de calidad con data de pruebas.

SPRINT 7

En el séptimo sprint del proyecto que fue el primero de las pruebas, el objetivo fue preparar el entorno de calidad y esto llevó a que se recopile los formatos de los archivos productivos, como podemos ver en la Figura 23, luego fueron cargados al servidor y se consultaron las tablas externas para validar su integridad.

FIGURA 23:

FORMATO DE ARCHIVOS

CAMPO	TIPO	PUEDE SER NULO
	VARCHAR2(8 BYTE)	Yes
	VARCHAR2(10 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(1 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(3 BYTE)	Yes
	VARCHAR2(1 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(1 BYTE)	Yes
	VARCHAR2(60 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(10 BYTE)	Yes
	VARCHAR2(20 BYTE)	Yes
	VARCHAR2(2 BYTE)	Yes
	VARCHAR2(10 BYTE)	Yes
	VARCHAR2(8 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes
	VARCHAR2(5 BYTE)	Yes

Nota: Preparado por el autor en base a los datos de un archivo asociado a una tabla externa

SPRINT 8

En el octavo sprint se tuvo como objetivo hacer pruebas a las cargas o ejecuciones de los subprocesos asociados a la transformación y cruce de data, estos procesos debían dar como salida un reporte correspondiente al modelo de alertas tempranas, en estos casos todos los resultados fueron correctos ya que al tener los inputs correctamente validados, el proceso de transformación y carga final procedían sin problemas. Se detalla el inventario de casos de prueba en la Figura 24.

FIGURA 24:

CASOS DE PRUEBA

Cod Caso Prueba	Caso de Prueba	Prioridad
CP001	Validar mostrar error al tener formato incorrecto	Alta
CP002	Validar continuar al tener formato correcto	Alta
CP003	Validar horario de ejecución programada	Alta
CP004	Validar continuidad del proceso	Alta
CP005	Validar conformidad de integridad de archivos	Alta
CP006	Validar inicio y termino de ejecución	Alta
CP007	Validar seguimiento en reglas de negocio	Alta
CP008	Validar salida final del proceso	Alta

Nota: Preparado por el autor en conjunto con el área de calidad y el usuario de riesgos

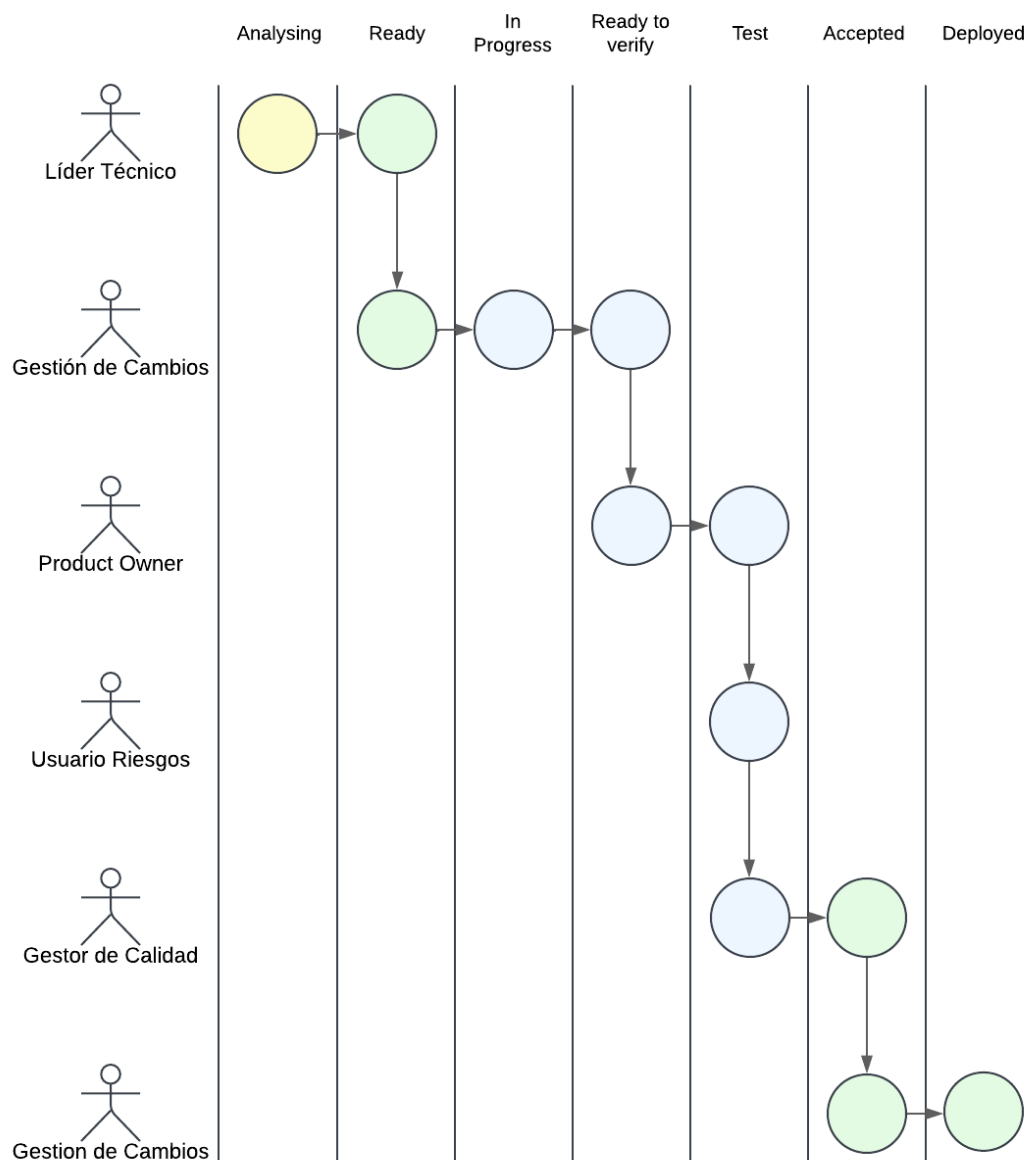
3.2.5.1.5 DESPLIEGUE PRODUCTIVO O IMPLANTACIÓN

SPRINT 9

En el noveno sprint que tiene como objetivo la implantación se debió tener conformidades por diferentes equipos y disciplinas, se siguió la secuencia de trabajo ya establecido por la entidad gestionado desde la herramienta JIRA que está descrito en la Figura 25.

FIGURA 25:

SECUENCIA DE TRABAJO EN JIRA



Nota: Preparado por el autor de acuerdo al uso de la herramienta de JIRA por la entidad

3.3 EVALUACIÓN

La puesta en marcha del proyecto de optimización del proceso ETL para el modelo de alertas tempranas, detallada en este trabajo de suficiencia profesional, proyecta una reducción de costos inmediatamente después de su implementación en el proceso mensual de aprovisionamiento.

3.3.1 EVALUACIÓN COSTO – BENEFICIO

3.3.1.1 ANÁLISIS DE COSTOS

El proyecto descrito se llevó a cabo durante un periodo de 5 meses, culminando en el segundo trimestre de 2023. Este desarrollo requirió una inversión para la entidad, que incluyó tanto el pago al personal del equipo de proyecto como el ahorro proyectado debido a la reducción en los tiempos de atención al eliminar la necesidad de realizar reintentos y validaciones manuales. Esta disminución en el tiempo de atención permitió un manejo más eficiente de los recursos, optimizando así los costos operativos. Además, al implementar el proyecto ETL, se proyecta un ahorro en los costes asociados con las labores manuales repetitivas, permitiendo al personal enfocarse en tareas de mayor valor agregado. El monto total necesario para la implementación del proyecto ETL ascendió a s/. 109,000, según se detalla en la Tabla 15, lo cual incluye todas las inversiones iniciales y los costos operativos necesarios para asegurar el éxito del proyecto.

TABLA 15:

COSTO TOTAL DE IMPLEMENTACIÓN DEL PROYECTO

RECURSO	PORCENTAJE PARTICIPACIÓN	SUELDO MENSUAL	SUELDO TOTAL POR 5 MESES
Analista	150%	s/. 7,200	s/. 54,000
Programador Java			
Product Owner	100%	s/. 6,000	s/. 30,000

Analista QA	100%	s/. 5,000	s/. 25,000
Costo Total			s/. 109,000

Nota: Elaborado por el autor. Tratándose el porcentaje de participación del número de colaboradores asociados al recurso requerido

En la Tabla 16 se especifica el costos asignados para la implementación del proyecto durante el primer año, así como las proyecciones de reducción de costos al disminuir los gastos asociados con los procesos manuales de validación que se efectúan en el aprovisionamiento mensual para el modelo de alertas tempranas en los tres primeros años tras la implementación del ETL.

TABLA 16:

PROYECCIÓN DE COSTES EN VALIDACIÓN EN EL PROCESO MENSUAL DE APROVISIONAMIENTO

AÑO	2023	2024	2025	2026
Total de costos incurridos en la validación del aprovisionamiento mensual para el modelo de alertas tempranas	s/.120,000	s/. 0	s/. 0	s/.0

Nota. Preparado por el autor. Proyección orientada en el sueldo mensual de un colaborador de la entidad que dedica su tiempo en validaciones manuales, como se aprecia en 2024 en adelante ya no hay costo porque se realizan automaticamente.

3.3.1.2 ANÁLISIS DE BENEFICIOS

Para poder demostrar los beneficios del proyecto se tomarán en cuenta los beneficios económicos y los beneficios estratégicos.

Los beneficios económicos están basados en el ahorro en costos operacionales por la automatización de procesos anteriormente señalados. Con la implementación de este proyecto, se espera una reducción de s/. 10,000 anuales en costos operativos, que están basados en el sueldo de las personas que están a cargo de los procesos mensuales de validación, en ejecuciones como en reprocesos. En el aspecto operacional, se prevé una disminución en el tiempo de 30% en los procesamiento de las ejecuciones de reglas de negocio, lo cual mejorará significativamente la eficacia operativa. Además, la eliminación de procesos manuales reducirá la probabilidad de errores humanos y mejorará la precisión de los datos.

Por los beneficios estratégicos, la automatización del procesamiento de datos incrementará la precisión y confiabilidad de la evaluación de riesgos, mejorando la toma de decisiones. Esta mejora en la precisión contribuirá a una mejor respuesta ante riesgos, proporcionando una ventaja competitiva en el mercado. La entidad podrá responder más rápidamente a las necesidades del mercado y de los clientes, fortaleciendo su posición competitiva.

3.3.1.3 EVALUACIÓN DE RESULTADO

La relación costo-beneficio se calculó en base a los valores estimados para el primer año después de la instalación del proyecto, por ello para efecto de este análisis se tomó en cuenta la inversión en la puesta productiva del proyecto así como la reducción de costes a futuro.

Con estas consideraciones se calculó la relación costo-beneficio, el proyecto tuvo un costo de s/. 109,000 tal cual lo muestra la Tabla 15 y como reducción anual de s/. 120,000 por lo que ya no se dará presupuesto a validaciones manuales y pasarían a ser automáticas, entonces se puede determinar que durante el primer año el costo-beneficio es de 0.11, en base a ello se puede estimar que por cada S/ 1 invertido, alcanzando un beneficio de S/ 1.11 generando que el proyecto a desarrollar sea beneficioso económicamente a partir del segundo año en adelante.

3.3.2 INTERPRETACIÓN DEL VAR Y TIR

Para evaluar la viabilidad económica del proyecto, se realizó un análisis del Valor Actual Neto (VAN) y la Tasa Interna de Retorno (TIR) utilizando una tasa de descuento del 10%. La inversión inicial requerida es de S/ 109,000 y se proyecta una reducción anual de costos de S/ 120,000 a partir del segundo año.

El VAN se calculó para un período de cinco años. En el primer año no se generan beneficios, mientras que en los años siguientes se proyectan flujos de caja anuales de S/ 120,000. La fórmula utilizada para calcular el VAN fue:
$$VAN = \sum (\text{Flujo de Caja}_t / (1 + r)^t) - \text{Inversión Inicial}$$
 Aplicando estos datos, se obtuvo un VAN de S/ 162,385, lo cual indica que el proyecto es rentable.

Para determinar la TIR, se identificó la tasa de descuento que hace que el VAN sea igual a cero. Utilizando las mismas proyecciones de flujo de caja y herramientas financieras como Excel, se encontró que la TIR es aproximadamente del 44%. Esta tasa de retorno es significativamente superior a la tasa de descuento del 10%, lo que indica que el proyecto no solo recuperará la inversión inicial, sino que también generará beneficios considerables.

De acuerdo a lo anterior mencionado, el análisis del VAN y TIR demuestra que el proyecto es económicamente viable. Un VAN positivo de S/ 162,385 y una TIR del 44% resaltan la rentabilidad del proyecto desde una perspectiva financiera, justificando así su implementación.

CAPÍTULO IV – REFLEXIÓN CRÍTICA DE LA EXPERIENCIA

El autor de este trabajo de Suficiencia Profesional desempeñó el rol de líder del proyecto en todas sus etapas, incluyendo el análisis, desarrollo, pruebas de calidad e implementación del sistema en el entorno productivo de la plataforma tecnológica de la entidad financiera. Esta experiencia brindó una valiosa oportunidad para fortalecer las habilidades de liderazgo y adquirir un entendimiento más profundo del ciclo de vida de proyectos establecido por la entidad financiera. De los logros obtenidos se destacan los siguientes puntos:

- El desarrollo del proyecto ETL representó la primera ocasión en la que el autor asumió el liderazgo en todas las fases del proyecto. La metodología Kanban y propia de la entidad fue crucial para gestionar las tareas del proyecto de manera eficiente y comunicar claramente los avances alcanzados a través de la herramienta JIRA.
- Las reuniones al inicio y al final de cada etapa de desarrollo permitieron obtener retroalimentación del equipo de desarrollo con los usuarios. Al compartir las experiencias adquiridas, se pudieron establecer las tareas adecuadas para alcanzar los objetivos y entregables requeridos en cada sprint.
- Durante la instalación del proyecto en el entorno de calidad, surgieron problemas asociados a no tener la misma cantidad de data o información respecto a producción, por lo que se tuvo que simular mucha data y en casos no era data tal cual producción por lo que se tuvo que cocinar data y tomó más tiempo en pruebas, este punto se pudo haber previsto.
- Una lección aprendida es que se tiene que planificar adecuadamente y adherirse a los cronogramas establecidos. Subestimar el tiempo necesario para tareas puede llevar a retrasos y al incumplimiento de plazos.

CAPÍTULO V – CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- Con la implementación de un correcto ETL se puede llegar a la eficacia y eficiencia en un proceso de aprovisionamiento de información. Si se determinan correctamente qué archivos y flujos son los necesarios, se puede llegar a reducir los tiempos significativamente en un proceso.
- Se logró establecer los tiempos correctos de ejecución para así no impactar en otros procesos por consumo de computador, esto conlleva tanto de reducir tiempos de ejecución como la alerta temprana de errores, ya no hay encolamiento.
- Se elaboró una conveniente definición y planificación de las tareas a realizar en cada etapa del desarrollo y con ello obtener un importante espacio de tiempo para enfrentar complicaciones técnicas y funcionales que fueron apareciendo en las etapas de desarrollo y pruebas de calidad, pero sin impactar con la entrega final del proyecto.

5.2 RECOMENDACIONES

- Para el presente proyecto se deben realizar mejoras en los procesos de desarrollo de software, que permita un proceso más ágil y se tenga menos obstáculos para poder llegar al proceso de implantación.
- Comunicar a los demás equipos de esta solución y que pueda ser base para futuros planteamientos que impliquen aprovisionamiento de información, los flujos de carga de trabajo pueden ser usados para otras cargas de trabajo.
- Utilizar nuevas herramientas para la conversión y transformación de data que se destaquen por su eficiencia, con esto proponer a la entidad financiera el uso de las nuevas tecnologías para que puedan ser incluidas en sus estándares.

- Proveer formación continua a todos los miembros del equipo sobre las nuevas metodologías y herramientas implementadas. Esto asegurará que todos estén alineados y capacitados para utilizar las tecnologías de manera eficiente.
- Desarrollar una documentación estandarizada y accesible de todos los procesos y herramientas utilizadas. Esto facilitará la transferencia de conocimientos y la incorporación de nuevos miembros al equipo.
- Establecer un mecanismo de retroalimentación regular donde los miembros del equipo puedan compartir sus experiencias y sugerencias sobre las nuevas implementaciones. Esto permitirá realizar ajustes oportunos y mejorar continuamente.

BIBLIOGRAFÍA

Congreso de la República del Perú. (1996). *Ley N° 26702, Ley General del Sistema Financiero y del Sistema de Seguros y Orgánica de la Superintendencia de Banca y Seguros*. Diario Oficial El Peruano.

Diario Oficial El Peruano. (2024). *Inclusión financiera: por un país moderno, formal y conectado*. Recuperado de <https://elperuano.pe/noticia/141949-inclusion->

INEI (Instituto Nacional de Estadística e Informática). (2023). *Informe sobre la Inclusión Financiera en el Perú, primer trimestre de 2023*. Recuperado de <https://infomercado.pe/inclusion-financiera-en-el-peru-crecio-casi-4-en-el-primer-trimestre-del-2023-010823-gb/>

SBS (Superintendencia de Banca, Seguros y AFP). (2020-2023). *Reportes anuales sobre el incremento de la digitalización de servicios bancarios en Perú*. Recuperado de <https://www.bcrp.gob.pe/docs/Proyeccion-Institucional/Encuentro-de-Economistas/2020/ee-2020-choy.pdf>

Algoritmia. (2024). *La importancia de un proceso ETL en el Business Intelligence*. Recuperado de <https://algoritmia8.com/2022/11/23/la-importancia-de-un-proceso-etl-en-el-business-intelligence/>

Cohn, M. (2010). *Succeeding with Agile: Software Development Using Scrum*. Addison-Wesley.

Rubin, K. S. (2012). *Essential Scrum: A Practical Guide to the Most Popular Agile Process*. Addison-Wesley.

Schwaber, K., & Sutherland, J. (2020). *The Scrum Guide*. Scrum.org.

Bravo-Huivin, E. K., Cieza-Mostacero, S. E., Flores-Rodriguez, L. A., & Uceda-Davila, L. (2020). *Revisión sistemática de la literatura sobre metodologías ágiles en el desarrollo de software aplicadas a la gestión de proyectos*. Universidad Privada del Norte.

Bravo-Huivin, E. K., Cieza-Mostacero, S. E., Flores-Rodriguez, L. A., & Uceda-Davila, L. (2020). Revisión sistemática de la literatura sobre metodologías ágiles en el desarrollo de software aplicadas a la gestión de proyectos. Universidad Privada del Norte.

IBM. (2023). ¿Qué es ETL (extracción, transformación, carga)? Recuperado de <https://www.ibm.com/es-es/topics/etl>

Astera. (2024). ¿Qué es ETL? La guía definitiva para extraer, transformar y cargar. Recuperado de <https://www.astera.com/es/type/blog/etl/>

Atlassian. (n.d.). Jira Software. Recuperado de <https://www.atlassian.com/software/jira>

Smith, J., & Johnson, L. (2021). Agile Project Management with Jira. Tech Publishing.

Williams, R. (2022). Modern Software Development: Tools and Techniques. Software Press.

Anderson, D. J. (2021). *Kanban: Successful evolutionary change for your technology business*. Blue Hole Press.

Medeiros dos Santos, P. S., Beltrão, A. C., de Souza, B. P., & Travassos, G. H. (2018). On the benefits and challenges of using kanban in software engineering: A structured synthesis study. *Journal of Software Engineering Research and Development*, 6(13).

Rosin, D. (2021). *Kanban: Knowledge and References*. Taylor & Francis.

Balaji, S., & Murugaiyan, M. S. (2021). Waterfall vs. V-Model vs. Agile: A comparative study on SDLC. *International Journal of Information Technology and Computer Science*, 12(1), 1-8.

Mishra, A., & Dubey, S. K. (2020). A comparative study of different software development life cycle models in different scenarios. *International Journal on Recent and Innovation Trends in Computing and Communication*, 8(3), 59-63.

Rastogi, A., Singh, P., & Sharma, R. (2021). Software development methodologies: A comparative analysis. *International Journal of Software Engineering & Applications*, 12(3), 23-33.

Royce, W. (2022). Managing the development of large software systems. *Proceedings of IEEE WESCON*, 29(12), 1-9.

ServiceNow. (n.d.). *What is Agile vs. Waterfall?* Recuperado de <https://www.servicenow.com/es/products/strategic-portfolio-management/what-is-agile-vs-waterfall.html>