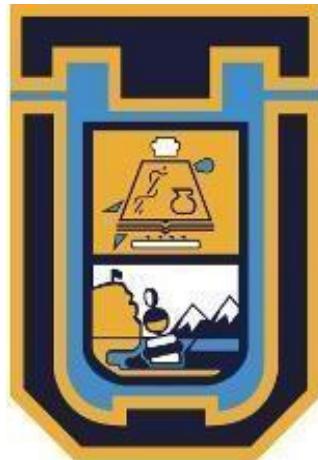


UNIVERSIDAD DE TARAPACÁ
FACULTAD DE INGENIERÍA
INGENIERÍA CIVIL EN COMPUTACIÓN E INFORMÁTICA ARICA
– CHILE



Documento de requisitos de
**Sistema de Análisis de RRHH
para Casino Luckia**

Equipo de Desarrollo:

- Tomás Ignacio Silva Muñoz
- Juan Carlos Yampara Rojas

Empresa o Unidad: Casino Luckia

Curso: Proyecto IV ICCI

Profesor: Diego Aracena Pizarro

1. Índice

2. Resumen	3
3. Definición del Problema y Justificación	4
3.1. Contexto	4
3.2. Problema	4
3.3. Solución escogida	5
3.4. Otras soluciones	5
4. Alcance, Objetivos y Planificación	6
4.1. Alcance del proyecto	6
4.2. Objetivo General	7
4.3. Objetivos Específicos	7
4.4. Carta Gantt	8
5. Requerimientos del proyecto	10
5.1. Funcionales	10
5.2. No funcionales	11
5.3. De documentación	12
6. Metodología	13
6.1. Macro-Proceso Analítico: CRISP-DM	13
6.2. Gestión Ágil del Desarrollo: Kanban	14
6.3. Integración de la Metodología Híbrida	14
6.4. Definición de tareas	15
7. Modelado del sistema	20
7.1. Modelo de contexto	20
7.2. Modelo de casos de uso	21
8. Modelado del Procesos de Negocio	24
9. Diseño de Interfaz de Usuario y Prototipo	25
9.1. Mockups	25
10. Conclusiones	27
11. Bibliografía	28

falta indice de figuras y tablas

2. Resumen

La gestión de Recursos Humanos en Casino Luckia Arica se enfrenta actualmente a desafíos significativos debido a la fragmentación de sus datos, los cuales residen en sistemas dispares y archivos manuales como SAP, GeoVictoria y planillas Excel. Esta desconexión impide un análisis estratégico y fomenta una toma de decisiones reactiva ante problemas críticos como la rotación de personal, el ausentismo y la optimización de la dotación.

Este proyecto propone el diseño y la futura implementación de una Plataforma Analítica Integral para abordar dicho problema. El objetivo principal es planificar la centralización de esta información a través de procesos automatizados de Extracción, Transformación y Carga (ETL) en un Data Warehouse (DW) unificado.

Para guiar la ejecución, se utilizará una metodología híbrida: CRISP-DM estructurará el proceso de minería de datos, desde la comprensión del negocio hasta el eventual despliegue de modelos predictivos; mientras que Kanban se empleará para la gestión ágil del desarrollo de software (ETLs, BBDD, Dashboards).

El resultado esperado de este plan es una solución funcional que automatizará el cálculo de KPIs (Tasa de Rotación, Ausentismo) y los presentará en dashboards interactivos, con el fin de dotar al área de RRHH de una herramienta que facilite la migración de una gestión operativa reactiva a una gestión estratégica y basada en datos.

debe estar justificado ambos lados

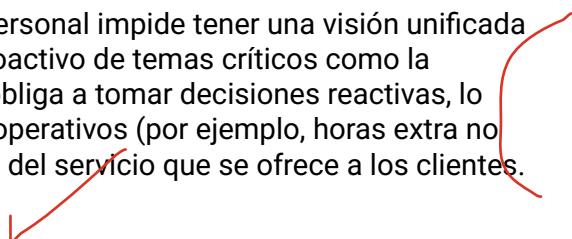
3. Definición del Problema y Justificación

3.1. Contexto

Actualmente, el equipo de Recursos Humanos de Casino Luckia Arica gestiona información esencial del personal (como ingresos, asistencia, turnos y licencias) a través de múltiples archivos y sistemas separados. Esta información se consolida de forma manual y fragmentada, lo que limita la capacidad del área para anticipar y tomar decisiones estratégicas sobre la rotación de colaboradores, sobre el ausentismo por área o turno, y sobre la dotación óptima requerida para la operación del casino.

3.2. Problema

La gestión fragmentada y manual de los datos de personal impide tener una visión unificada y rápida de la plantilla, lo cual dificulta el análisis proactivo de temas críticos como la rotación y el ausentismo. Esta falta de integración obliga a tomar decisiones reactivas, lo que se traduce en un mayor riesgo de sobrecostos operativos (por ejemplo, horas extra no planificadas) y una potencial afectación a la calidad del servicio que se ofrece a los clientes.



3.3. Solución escogida

Se propone implementar una Plataforma Analítica Integral que unifique automáticamente todas las fuentes de datos de Recursos Humanos en un almacén de datos centralizado y permitir desde ahí el procesamiento analítico de los datos. Esta plataforma permitirá calcular métricas clave de rotación, ausentismo y dotación; además de permitir desarrollar modelos predictivos para anticipar el comportamiento del personal, y visualizar toda la información de manera sencilla a través de paneles de control interactivos. Esto transformará la gestión de personas en un proceso estratégico y proactivo. Este proyecto también contempla la identificación y recolección de nuevos datos de relevancia analítica, garantizando un ciclo de mejora continua donde los modelos se realimentarán con más información para aumentar progresivamente su precisión e impacto en las decisiones estratégicas de la empresa.



3.4. Otras soluciones

Antes de decidir la arquitectura final, se consideró una alternativa más simple: usar una herramienta como Power BI para conectarla directamente a las fuentes de datos (reportes de SAP, planillas Excel, etc.), sin construir un Data Warehouse (DW) intermedio.



Esta opción se analizó por ser, en teoría, mucho más rápida de implementar. Sin embargo, se descartó porque traía problemas importantes a futuro:

1. **Mantenimiento Complicado:** Toda la lógica para limpiar y unificar los datos tendría que vivir dentro de Power BI. Si una fuente de datos cambia, arreglar los reportes sería muy difícil.
2. **Dificultad para Modelos Predictivos:** Entrenar un modelo (como el de predicción de rotación) requiere datos limpios, históricos y consolidados. Hacer esto "al vuelo" desde las fuentes originales es ineficiente y muy complejo.
3. **Poca Flexibilidad:** Si en el futuro se quisieran añadir nuevas métricas o cruzar datos de formas no previstas, sería necesario rehacer gran parte del trabajo.



Se concluyó que, aunque la construcción de un Data Warehouse toma más tiempo al inicio, es la base necesaria para que el sistema sea mantenible, escalable y, sobre todo, para que los modelos predictivos puedan funcionar.

4. Alcance, Objetivos y Planificación

A continuación, se define el alcance y los objetivos del proyecto, integrando los requerimientos funcionales y técnicos.

4.1. Alcance del proyecto

El proyecto contempla el diseño, desarrollo e implementación de una plataforma analítica de datos para el área de RRHH de Casino Luckia Arica.

El alcance del sistema incluye:

1. **Ingesta de Datos:** Desarrollar un módulo que permita la carga de archivos (ej. planillas Excel) exportados desde los sistemas actuales (SAP y Geovictoria).
2. **Procesamiento y Almacenamiento:** Implementar procesos ETL (Extracción, Transformación y Carga) para limpiar, validar, transformar y unificar los datos en un Data Warehouse (DW) centralizado que almacenará el historial de la información.
3. **Analítica Descriptiva:** Calcular y exponer métricas clave de forma automatizada, incluyendo Tasa de Rotación, Tasa de Ausentismo y Dotación de personal por área/turno.
4. **Analítica Predictiva:** Entrenar un modelo predictivo inicial, basado *exclusivamente* en los datos históricos disponibles durante el proyecto, para estimar la probabilidad de rotación (baja) de colaboradores por área.
5. **Visualización y Reportes:** Disponer de una interfaz gráfica (dashboard) que permita a los usuarios finales consultar las métricas y los resultados del modelo de forma interactiva, así como generar reportes automatizados.
6. **Seguridad:** Aplicar controles de ciberseguridad para asegurar la confidencialidad, integridad y disponibilidad de los datos (ej. control de acceso, logs).

Fuera del Alcance (Exclusiones):

1. El proyecto **no modificará los sistemas fuente** (SAP o Geovictoria).
2. El proyecto **no desarrollará nuevos módulos de captura de datos** (ej. formularios de ingreso); solo consumirá la información existente.
3. La implementación se realizará utilizando la infraestructura y tecnologías dispuestas o aprobadas por la organización.

4.2. Objetivo General

Implementar una plataforma analítica de datos para el área de RRHH de Casino Luckia Arica, que automatice la ingesta y procesamiento de información, permitiendo el análisis de métricas clave y la generación de modelos predictivos para facilitar la toma de decisiones proactiva sobre la gestión del personal.

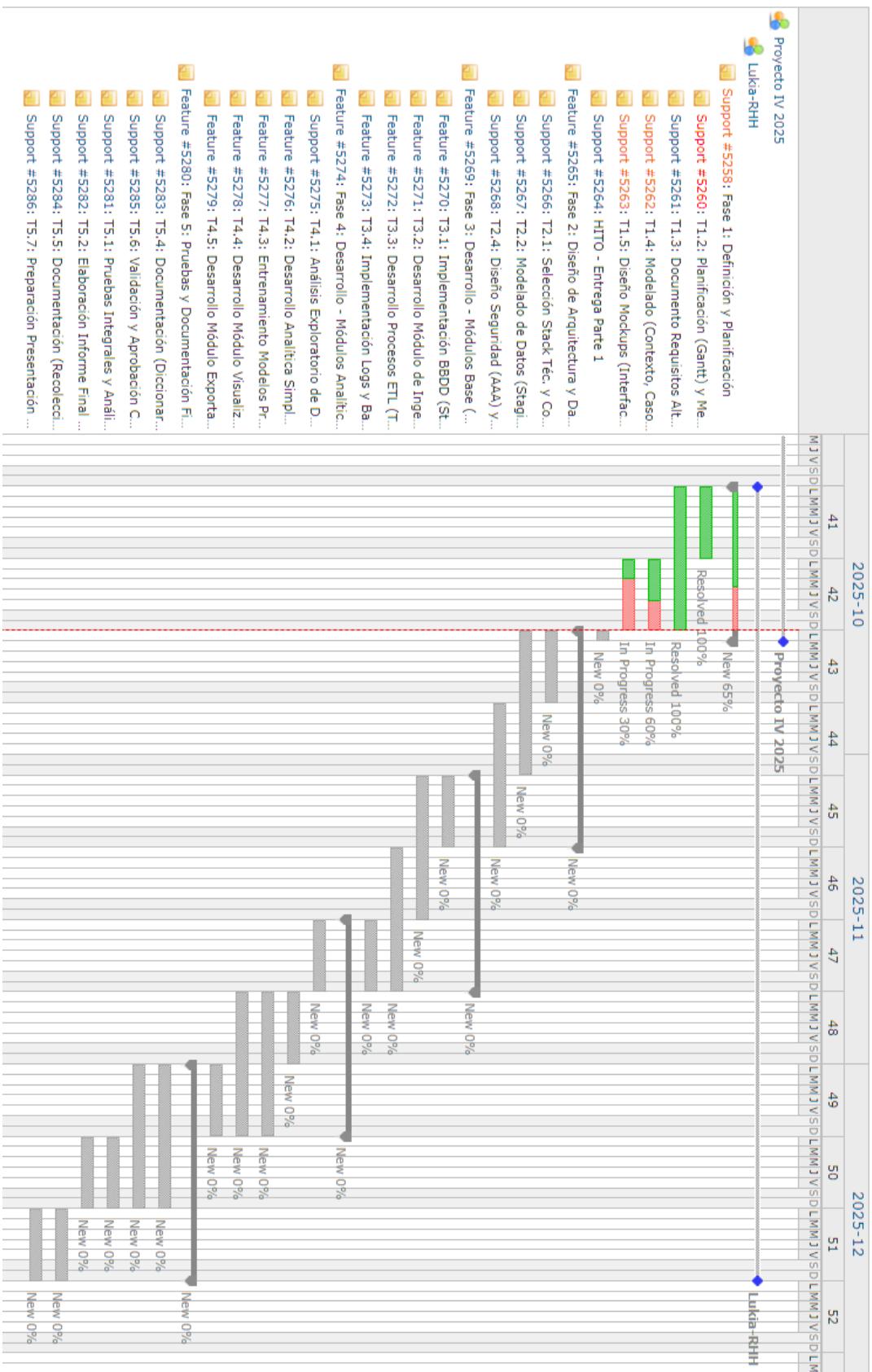
4.3. Objetivos Específicos

1. **Diseñar y construir** un proceso ETL (Extracción, Transformación y Carga) capaz de ingestar, validar y unificar datos de múltiples fuentes (SAP y Geovictoria).
2. **Implementar** un Data Warehouse (DW) centralizado para consolidar y almacenar de forma histórica y optimizada la información del personal.
3. **Analizar** los datos históricos para desarrollar y entrenar un modelo predictivo inicial, enfocado en la probabilidad de rotación de personal por área.
4. **Desarrollar** un módulo de visualización con reportes automatizados y dashboards interactivos que muestren las métricas clave (rotación, ausentismo, dotación).
5. **Garantizar** la seguridad y privacidad de la información mediante la aplicación de principios de ciberseguridad (ej. control de acceso, logs, comunicaciones seguras) en la arquitectura de la solución.
6. **Documentar** la arquitectura del sistema, el diccionario de datos y los flujos ETL para asegurar su mantenibilidad y facilitar la futura captación de nuevos datos.

rehacer los objetivos.. el general debe estar uno a uno con el título del proyecto
Los específicos, estan muy de hacer y parecen actividades..
-Estudiar los Datos para construir el ETL y Data Warehouse
-Analizar los datos para..
-Desarrollar un módulo.. y garantizar la security.
-Validar probar los resultados ..documentar

4.4. Carta Gantt

Título	Descripción	Semana (Relativa)	Fecha Inicio	Fecha Término
Fase 1: Definición y Planificación	Análisis de requisitos, alcance, modelado (Contexto, Casos de Uso, BPM) y mockups para la primera entrega.	1 - 2	06-oct-25	17-oct-25
HITO 1: Entrega 1	Entrega formal de la documentación de la Fase 1.	3	20-oct-25	20-oct-25
Fase 2: Diseño de Arquitectura y Datos	Selección de stack, configuración de Git, modelado de BBDD (Staging/DW) y diseño de procesos ETL y seguridad (AAA).	3 - 6	21-oct-25	14-nov-25
Fase 3: Desarrollo - Módulos Base (ETL/DW)	Implementación de BBDD, desarrollo de ingesta (Excel), procesos ETL (transformación/carga) y sistemas de logs/backups.	5 - 8	03-nov-25	21-nov-25
HITO 2: Entrega 2	<i>Presentación de avances de implementación y diseño detallado.</i>	8 (Aprox)	(Sin Fecha)	(Sin Fecha)
Fase 4: Desarrollo - Módulos Analíticos	Análisis exploratorio (EDA), desarrollo de KPIs (Rotación, etc.), entrenamiento de modelos predictivos y dashboards.	8 - 11	24-nov-25	16-dic-25
Fase 5: Pruebas y Documentación Final	Pruebas integrales, redacción de informes, manuales (Usuario, ETL, RPO/RTO) y validación con el cliente.	10 - 12	10-dic-25	23-dic-25
HITO 3: Entrega 3	<i>Entrega del sistema implementado y el informe final/pruebas.</i>	11 (Aprox)	(Sin Fecha)	(Sin Fecha)
HITO 4: Entrega Final	Presentación final del producto y entrega de manuales aprobados.	12	24-dic-25	24-dic-25



5. Requerimientos del proyecto

5.1. Funcionales

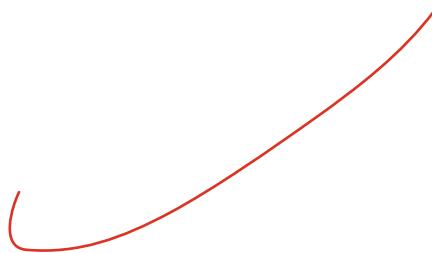
ID	Título	Descripción
RF-01	Plataforma de integración de datos	El sistema debe contar con una plataforma para subir datos de distintas fuentes no relacionales (principalmente Planillas Excel), transformándolos en un formato común y trabajable.
RF-02	Base de datos de staging	El sistema debe contar con una base de datos separada para almacenar temporalmente los datos extraídos antes de su carga final en el Almacén de Datos.
RF-03	Almacén de Datos (DW)	<p>El sistema debe disponer de un Data Warehouse estructurado que permita almacenar los datos consolidados de forma histórica y optimizada para consulta.</p> <p>El sistema debe permitir la integración de datos desde múltiples fuentes mediante procesos ETL, incluyendo la validación, limpieza y transformación de la información antes de su carga en el Almacén de Datos.</p>
RF-04	Analítica simple de los datos	El sistema debe mostrar métricas simples calculadas automáticamente con los datos disponibles. Debe mostrar métricas como: Tasa de rotación, Tasa de Ausentismo, Dotación de personal, entre otros.
RF-05	Modelos predictivos	El sistema debe implementar modelos predictivos entrenados con los datos recopilados. Debe predecir: Probabilidad de rotación,
RF-06	Visualización de métricas y KPIs	El sistema debe mostrar gráficamente métricas relevantes obtenidas de los modelos y otros cruces de información complejos.
RF-07	Registro de logs	El sistema debe registrar eventos relevantes, como accesos, ejecuciones de procesos ETL, fallos y cambios de configuración.
RF-08	Backups automáticos	El sistema debe realizar copias de seguridad periódicas de las bases de datos y configuraciones del sistema.
RF-09	Control de acceso AAA	El sistema debe implementar autenticación, autorización y auditoría de usuarios (AAA) con distintos roles de acceso.
RF-10	Exportación de datos	El sistema debe permitir exportar los datos transformados o consolidados en formatos como CSV, Excel o JSON.

5.2. No funcionales

ID	Título	Descripción
RNF-01	Comunicaciones seguras	Toda comunicación entre módulos o usuarios debe realizarse mediante protocolos cifrados (TLS/SSL).
RNF-02	Disponibilidad	El sistema debe mantenerse disponible y operativo de forma continua, considerando réplicas y ventanas de mantenimiento programadas.
RNF-03	Escalabilidad	El sistema debe poder ampliar su capacidad de procesamiento y almacenamiento sin rediseñar su arquitectura base.
RNF-04	Mantenibilidad	El sistema debe diseñarse de forma modular para facilitar actualizaciones, pruebas y resolución de errores.
RNF-05	Compatibilidad	El sistema debe ser compatible con las tecnologías y herramientas de la organización (Microsoft, PostgreSQL, Power BI, etc.).
RNF-06	Confiabilidad	El sistema debe garantizar la integridad de los datos y la continuidad operativa ante fallos.
RNF-07	Eficiencia	El sistema debe procesar y cargar los datos en tiempos razonables según el volumen esperado.
RNF-08	Legalidad y privacidad	El sistema debe cumplir con la normativa vigente sobre protección de datos personales y confidencialidad laboral.
RNF-09	Auditabilidad	Todas las acciones del sistema deben poder rastrearse para fines de control y cumplimiento normativo.
RNF-10	Usabilidad dual	El sistema debe presentar una interfaz sencilla para usuarios no técnicos y funcionalidades avanzadas para expertos.

5.3. De documentación

ID	Título	Descripción
RD-01	Diccionario de datos	Se debe elaborar un diccionario que describa el significado, formato, dominio y relaciones de cada campo del sistema.
RD-02	Documentación de procesos ETL	Se debe documentar cada flujo de extracción, transformación y carga, incluyendo fuentes, reglas y errores posibles.
RD-03	Informe de Análisis Exploratorio de Datos (EDA)	Debe elaborarse un informe técnico que resuma la completitud, consistencia, representatividad y calidad de los datos disponibles; con el objetivo de analizar la viabilidad de los modelos predictivos y observar la necesidad de la captación de nuevos datos.
RD-05	Plan de recolección de nuevos datos	Debe elaborarse un documento con estrategias para mejorar la cobertura y calidad de datos.
RD-06	Manual de usuario administrativo	Debe documentarse el uso de la plataforma, incluyendo carga, validación, reportes, backups y administración de usuarios.
RD-07	Plan de recuperación ante fallos (RPO/RTO)	Se debe definir un plan para restaurar datos y servicios dentro de los tiempos objetivos de recuperación (RPO y RTO).
RD-08	Plan de adopción	Se debe elaborar un plan que facilite la incorporación del sistema en la organización, incluyendo capacitación y soporte inicial.



6. Metodología

Para abordar la naturaleza dual de este proyecto, que combina el **desarrollo de un producto de software** (la plataforma) con un proceso de **descubrimiento de conocimiento** (la analítica de datos), se implementará una metodología híbrida.

Esta metodología combina el macro-proceso **CRISP-DM** para guiar el proyecto de datos, con la agilidad y simplicidad de **Kanban** para la gestión diaria del desarrollo y las tareas.

6.1. Macro-Proceso Analítico: CRISP-DM

Para el componente de analítica y modelos predictivos, se utilizará el **Proceso Estándar Inter-industria para Minería de Datos (CRISP-DM)**. Este modelo es el estándar de la industria porque es iterativo y se centra en cómo los datos pueden responder a los problemas del negocio.

Las fases de CRISP-DM se aplicarán de la siguiente manera:

1. **Comprensión del Negocio (Business Understanding)**: Completada. Corresponde a la "Definición del Problema" y "Objetivos", donde se identificó la necesidad de analizar la rotación (RF-05) y el ausentismo (RF-04).
2. **Comprensión de los Datos (Data Understanding)**: Se realizará un Análisis Exploratorio de Datos (EDA) (RD-03) para identificar la calidad, completitud y potencial de los datos de SAP y Geovictoria.
3. **Preparación de los Datos (Data Preparation)**: Es la fase más intensiva en desarrollo. Implica construir los flujos ETL (RF-03) para limpiar, transformar y consolidar los datos en el Data Warehouse (RF-02, RF-03).
4. **Modelado (Modeling)**: Se seleccionarán y se usarán diferentes técnicas y algoritmos de minería de datos para intentar predecir la rotación de personal (RF-05), utilizando los datos preparados.
5. **Evaluación (Evaluation)**: Se medirán los modelos para asegurar que su precisión y rendimiento sean aceptables y que respondan correctamente a la pregunta de negocio.
6. **Despliegue (Deployment)**: El modelo predictivo final se integrará en la plataforma para ser consumido a través de los dashboards y reportes (RF-06).

6.2. Gestión Ágil del Desarrollo: Kanban

Para la gestión diaria del proyecto y el desarrollo de software (ETLs, BBDD, App Web, Dashboards), se utilizará **Kanban**. Esta metodología ágil se prefiere sobre Scrum por su simplicidad y flexibilidad, eliminando ceremonias complejas y centrándose en el flujo de trabajo visual.

Nuestra implementación de Kanban se basará en tres pilares:

1. **Visualizar el Flujo (Tablero Kanban):** Se utilizará un tablero digital simple para visualizar todas las tareas. Las columnas básicas serán:
 - **Backlog (Por Hacer):** Lista priorizada de todas las tareas.
 - **En Proceso (Doing):** Tareas que se están trabajando activamente. Se limitará el "Trabajo en Progreso" (WIP) para mantener el foco y simplificar el proceso.
 - **En Pruebas / Validación:** Tareas terminadas que están siendo probadas por el equipo o validadas por el cliente.
 - **Hecho (Done):** Tareas completadas y aprobadas.
2. **Cadencias (Reuniones Ágiles y Ligeras):** Para mantener la comunicación y la cadencia de desarrollo, adoptaremos solo las ceremonias más valiosas y simples:
 - **Reunión Diaria (Daily Stand-up):** Una reunión de 15 minutos cada mañana frente al tablero Kanban para sincronizar al equipo. Cada miembro responde: ¿Qué hice ayer? ¿Qué haré hoy? ¿Qué impedimentos tengo?
 - **Revisión de Entrega (Review/Demo):** Al finalizar un hito o una funcionalidad importante (ej. "Módulo de Reportes terminado"), se realizará una demo al cliente (Profesor y/o Luckia) para obtener feedback temprano, en lugar de esperar al final de un Sprint.
3. **Gestión del Flujo (Mejora Continua):** En lugar de planificar en Sprints, el equipo se guiará tomando la siguiente tarea de mayor prioridad del Backlog, tan pronto como termine una tarea actual.

6.3. Integración de la Metodología Híbrida

Ambas metodologías funcionan en conjunto:

- **CRISP-DM** define las **grandes fases analíticas** del proyecto, las mencionadas anteriormente.
- **Kanban** se usa para **desglosar y ejecutar** esas fases en tareas atómicas y manejables.

6.4. Definición de tareas

Para gestionar el trabajo siguiendo la metodología Kanban, el proyecto se desglosa en un **Backlog** (lista de trabajo pendiente). Este backlog se estructura jerárquicamente en **Épicas** y **Tareas**.

1. **Épicas (Epics):** Son las grandes unidades de trabajo o módulos del proyecto (ej. "Módulo de Ingesta y ETL"). Una Épica es demasiado grande para trabajarla directamente, pero sirve para agrupar tareas relacionadas.
2. **Tareas (Tasks):** Son las unidades de trabajo individuales y accionables que se moverán por el tablero Kanban (ej. "Crear tabla de Staging"). Cada tarea nace de uno o más requisitos (RF, RNF, RD).

Sistema de Etiquetas (Tags)

Cada tarea en el backlog se clasifica con un sistema de etiquetas simple para entenderla a simple vista dentro del mismo:

- **Fase (#fase[1-5]):** Indica a qué fase de la Carta Gantt pertenece (ej. #fase3).
- **Tipo (#[tipo]):** Define la naturaleza del trabajo (ej. #doc para documentación, #feature para funcionalidad, #config para infraestructura, #análisis para investigación).
- **Componente (#[componente]):** El módulo técnico que afecta (ej. #etl, #dw, #ui, #seguridad).
- **Referencia (#[ref]):** El requisito principal que cumple (ej. #rf01, #rnf02).

A continuación, se presenta el backlog inicial del proyecto.

Épica 1: Definición y Planificación (Entrega Parte 1)

Agrupa todas las tareas documentales, de análisis y de diseño de alto nivel requeridas para cumplir con la primera entrega del 20 de octubre.

Nombre Tarea	Descripción	Tags	Prioridad
Redactar Doc. Problema/Solución	Definir y justificar el problema y la solución escogida.	#fase1 #doc #proyecto	Crítica
Redactar Doc. Alcance y Objetivos	Detallar el alcance, objetivos (Gral/Específicos) del proyecto.	#fase1 #doc #proyecto	Crítica
Crear Carta Gantt	Planificar las fases, tareas e hitos del proyecto (06-Oct al 24-Dic).	#fase1 #doc #proyecto	Crítica
Redactar Doc. Metodología	Explicar el uso de CRISP-DM y Kanban (Sección 5).	#fase1 #doc #proyecto	Crítica

Crear Doc. Requisitos Alto Nivel	Enumerar las características principales y requisitos funcionales.	#fase1 #doc #proyecto #rf-all	Crítica
Desarrollar Modelo de Contexto	Crear diagrama de contexto (Sistema vs. Entorno externo).	#fase1 #doc #diseño	Crítica
Desarrollar Modelo Casos de Uso	Crear diagrama y descripción de Casos de Uso (Admin, Analista).	#fase1 #doc #diseño	Crítica
Desarrollar Diagrama BPMN	Mapear el proceso de negocio propuesto (flujo de datos y análisis).	#fase1 #doc #diseño	Crítica
Crear Mockups de Interfaces	Diseñar un prototipo de baja fidelidad de la UI (Carga y Dashboard).	#fase1 #doc #ui #rnf10	Crítica

Épica 2: Diseño de Arquitectura y Datos

Define los "planos" técnicos detallados del sistema, las bases de datos y la seguridad antes de construir.

Nombre Tarea	Descripción	Tags	Prioridad
Documentar Stack Tecnológico	Decidir y documentar SO, BBDD, Lenguajes, Frameworks.	#fase2 #doc #arquitectura #rnf05	Crítica
Diseñar Topología de Red	Crear diagrama de componentes (Servidor, BBDD, App Web).	#fase2 #doc #arquitectura #rnf01	Media
Modelar BBDD Staging	Diseñar el esquema de tablas para la BBDD de Staging.	#fase2 #doc #bd #rf02 #rd01	Crítica
Modelar BBDD Data Warehouse	Diseñar el modelo dimensional (tablas de Hechos y Dimensiones).	#fase2 #doc #dw #rf03 #rd01	Crítica
Mapear Flujo ETL	Documentar el mapeo de campos (Origen -> Staging -> DW).	#fase2 #doc #etl #rd02	Crítica
Definir Reglas de Limpieza	Especificar las reglas de negocio para transformar y validar datos.	#fase2 #doc #etl #rd02 #rnf06	Crítica
Definir Roles y Permisos	Detallar qué puede hacer cada rol (Admin, Analista, Visitante).	#fase2 #doc #seguridad #rf09	Alta
Definir Estrategia de Backup	Especificar frecuencia, retención y lugar de almacenamiento.	#fase2 #doc #backup #rf08	Baja

Épica 3: Configuración de Infraestructura y BBDD

Tareas para "levantar" el entorno de desarrollo, instalar servicios y desplegar las bases de datos vacías.

Nombre Tarea	Descripción	Tags	Prioridad
Instalar y Configurar Servidor	Instalar Aplicaciones y Servicios para soportar el sistema (BBDD, Git, Python, Docker).	#fase3 #config #servidor	Crítica
Configurar Comunicaciones Seguras	Implementar certificados TLS/SSL autofirmados o Let's Encrypt.	#fase3 #config #seguridad #rfn01	Baja
Desplegar BBDD Staging	Crear y ejecutar script SQL para crear la base de datos de Staging.	#fase3 #feature #bd #rf02	Alta
Desplegar BBDD Data Warehouse	Crear y ejecutar script SQL para crear el Data Warehouse (esquema).	#fase3 #feature #dw #rf03	Alta

Épica 4: Desarrollo Módulo de Ingesta y ETL

Construcción del "corazón" del sistema: la plataforma para cargar archivos y los scripts que mueven y transforman los datos.

Nombre Tarea	Descripción	Tags	Prioridad
Desarrollar UI Carga de Archivos	Crear la interfaz web para que el usuario suba planillas Excel.	#fase3 #feature #ui #rf01 #rfn10	Alta
Desarrollar API Carga (Excel -> Staging)	Crear el backend que recibe el archivo y lo inserta en BBDD Staging.	#fase3 #feature #api #rf01	Crítica
Desarrollar Script ETL (Staging -> DW)	Crear el proceso (Transform/Load) que aplica reglas y mueve datos al DW.	#fase3 #feature #etl #rf03	Crítica
Orquestar Flujo ETL	Crear el script principal que llama a la Carga y al T-L en orden.	#fase3 #feature #etl #rf03 #rfn07	Crítica

Épica 5: Desarrollo Módulo de Seguridad y Operaciones

Implementación de los sistemas de acceso, auditoría (logs) y los respaldos automáticos.

Nombre Tarea	Descripción	Tags	Prioridad
Desarrollar Módulo de Autenticación	Crear o integrar (SSO) la lógica de Login/Logout.	#fase3 #feature #seguridad #rf09	Media
Desarrollar Módulo de Autorización	Implementar la lógica de Roles y Permisos en la UI/API.	#fase3 #feature #seguridad #rf09	Media
Implementar Sistema de Logs	Configurar logging para accesos, errores y eventos ETL.	#fase3 #feature #auditoria #rf07 #rnf09	Alta
Automatizar Backups BBDD	Crear y programar (ej. cronjob) el script de backup automático.	#fase3 #config #backup #rf08 #rnf06	Baja

Épica 6: Análisis Exploratorio y Modelado Predictivo

Tareas de Ciencia de Datos (CRISP-DM) para entender los datos y construir el modelo de predicción de rotación.

Nombre Tarea	Descripción	Tags	Prioridad
Ejecutar EDA (Análisis Exploratorio)	Analizar calidad, distribución y completitud de los datos históricos.	#fase4 #análisis #datos #rd03	Alta
Preparar Dataset para Modelo	Crear el dataset de entrenamiento (Feature Engineering, limpieza).	#fase4 #análisis #modelo #rf05	Media
Entrenar y Evaluar Modelo	Probar algoritmos (ej. Regresión Logística, Random Forest) y evaluarlos.	#fase4 #análisis #modelo #rf05	Media
Desplegar Modelo (Job/API)	Implementar el modelo entrenado para generar predicciones.	#fase4 #feature #modelo #rf05	Media

Épica 7: Desarrollo Módulo Analítico y Visualización

Descripción: Construcción de la interfaz de usuario final (Dashboards) para mostrar KPIs y predicciones.

Nombre Tarea	Descripción	Tags	Prioridad
Implementar Cálculos de KPIs	Crear Vistas o Procedimientos Almacenados para KPIs.	#fase4 #feature #dw #rf04	Alta
Desarrollar Dashboard Sección de KPIs	Crear la UI para visualizar la Analítica Simple (Gráficos).	#fase4 #feature #ui #rf04 #rf06	Alta
Desarrollar Dashboard Sección Predictivo	Crear la UI para visualizar los resultados del modelo de rotación.	#fase4 #feature #ui #rf05 #rf06	Media
Implementar Módulo de Exportación	Desarrollar la función para exportar datos (CSV/Excel).	#fase4 #feature #api #ui #rf10	Baja

Épica 8: Documentación Final y Cierre

Descripción: Tareas de empaquetado del proyecto, creación de manuales y planes para la entrega final.

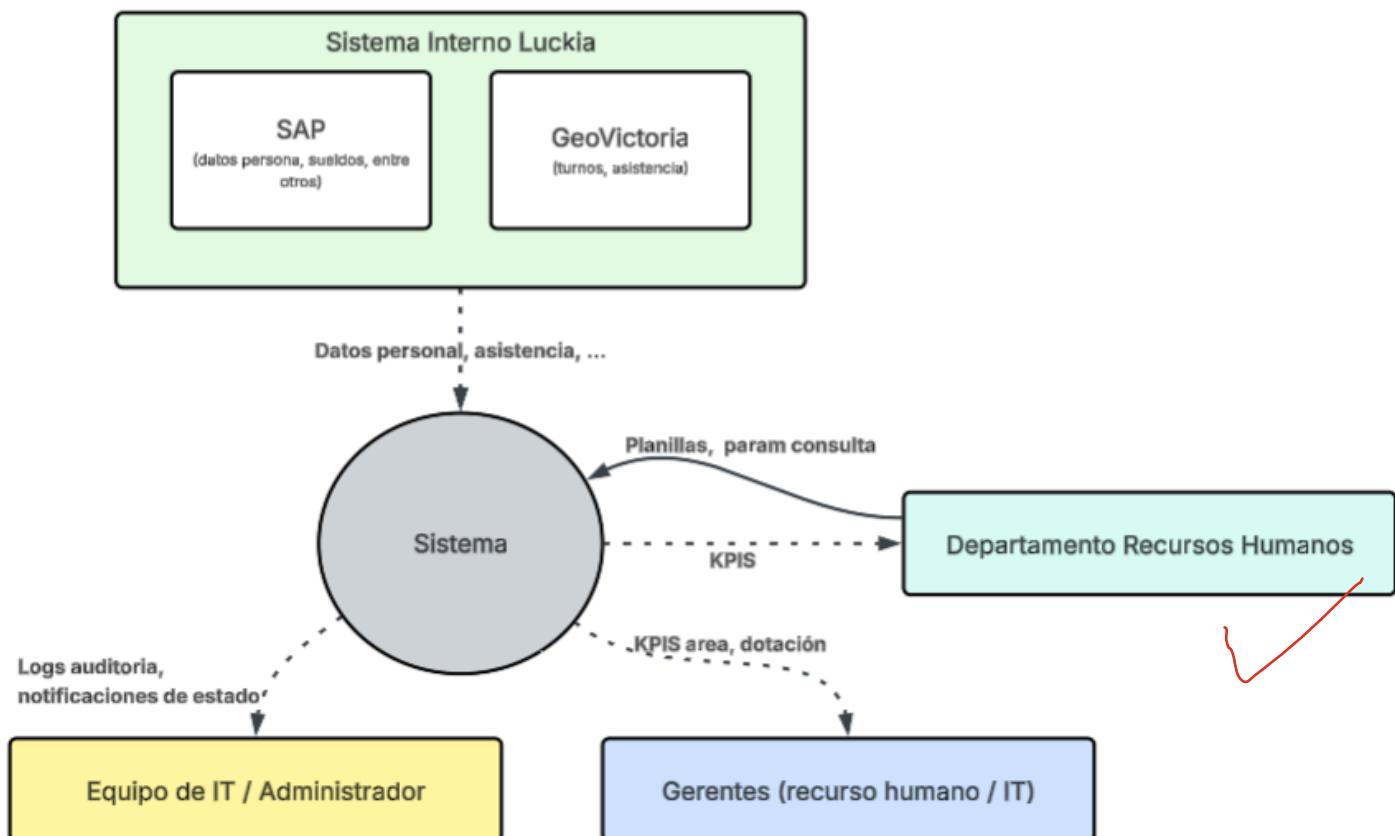
Nombre Tarea	Descripción	Tags	Prioridad
Redactar Diccionario de Datos	Documentar el esquema final del DW (tablas, campos, tipos).	#fase5 #doc #bd #rd01	Alta
Documentar Flujos ETL	Crear los diagramas y descripción final de los flujos ETL.	#fase5 #doc #etl #rd02	Media
Redactar Manual de Usuario/Admin	Crear guía de uso de la plataforma (Carga, Reportes, Usuarios).	#fase5 #doc #usuario #rd06	Media
Redactar Plan de Recuperación (RPO/RTO)	Documentar el plan de recuperación ante desastres.	#fase5 #doc #riesgos #rd07	Baja
Redactar Plan de Recolección de Datos	Proponer estrategias para capturar nuevos datos a futuro.	#fase5 #doc #análisis #rd05	Baja
Redactar Plan de Adopción	Documentar el plan de capacitación y transferencia.	#fase5 #doc #usuario #rd08	Baja
Pruebas Finales y Validación Cliente	Ejecutar pruebas integrales y obtener la aprobación final.	#fase5 #qa #cliente	Media

7. Modelado del sistema

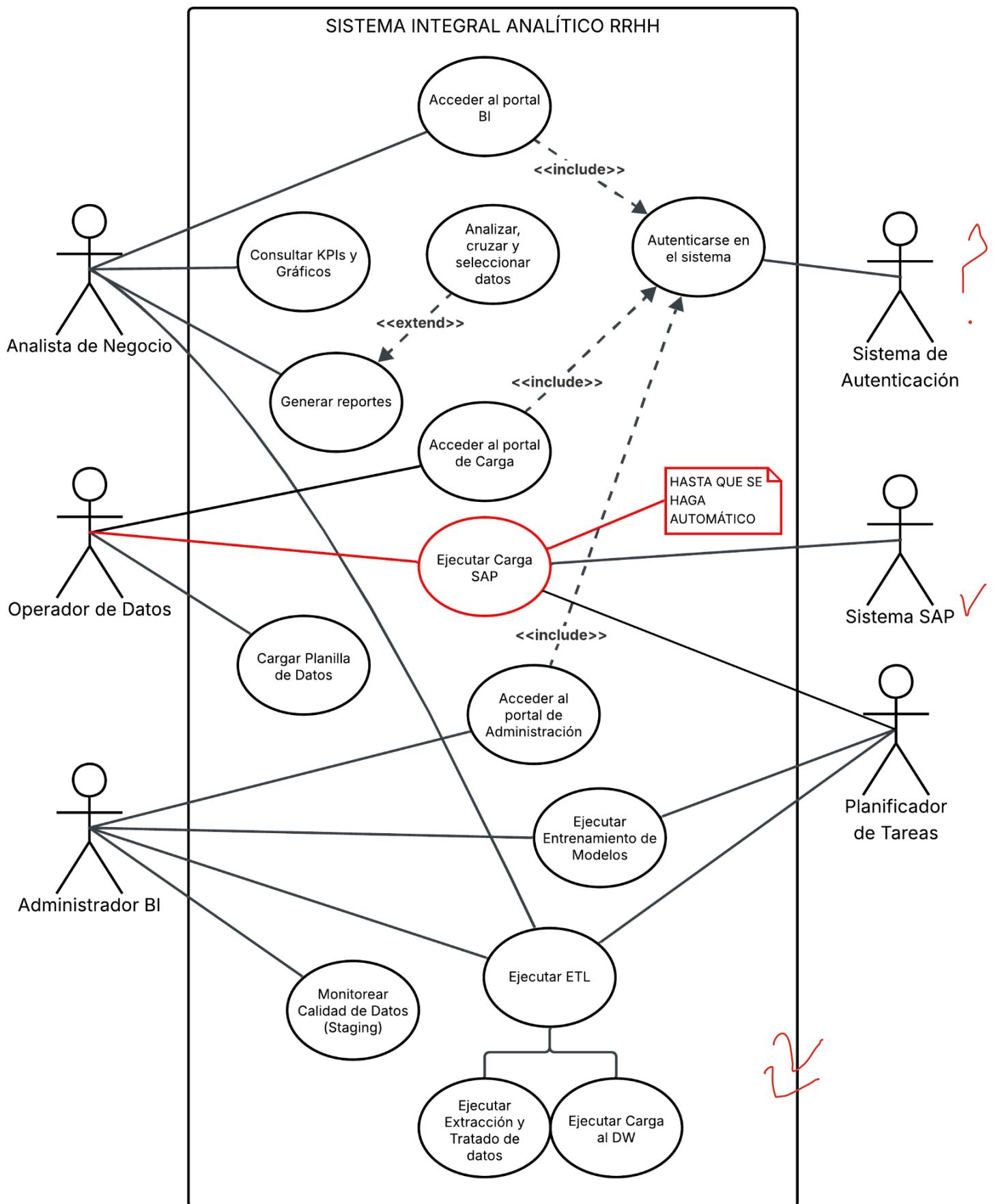
7.1. Modelo de contexto

El diagrama de contexto ilustra la plataforma analítica, denominada "Sistema", como la entidad central del flujo de información. Este sistema se alimenta desde dos canales principales: por un lado, ingiere los datos provenientes de los sistemas transaccionales de la empresa (SAP y GeoVictoria), los cuales proveen la información maestra del personal, como sus detalles contractuales, sueldos, turnos y registros de asistencia. Por otro lado, también recibe entradas de datos manuales, como planillas Excel, directamente desde el Departamento de Recursos Humanos.

Una vez que el sistema procesa esta información, la distribuye a sus usuarios a través de dos canales principales. Primero, provee la inteligencia de negocio (incluyendo indicadores de gestión (KPIs), métricas especializadas como la dotación por área y los resultados de los modelos de predicción) tanto al Departamento de Recursos Humanos como a los Gerentes. Segundo, reporta toda la información técnica, como los registros de auditoría y las notificaciones de estado del sistema, al Equipo de IT o Administrador para fines de monitoreo y control.



7.2. Modelo de casos de uso



1. Actor: Analista de Negocio

Es el consumidor principal de la información.

- **Acceder al portal BI:** El analista ingresa a la plataforma de visualización.
 - **(Include) Autenticarse en el sistema:** Para acceder, el sistema siempre le pedirá que inicie sesión.
- **Consultar KPIs y Gráficos:** Visualiza los reportes y dashboards predefinidos.
- **Generar reportes:** Utiliza la plataforma para crear reportes personalizados.
 - **(Extend) Analizar, cruzar y seleccionar datos:** Opcionalmente, durante la generación de reportes, puede realizar análisis más profundos y cruces de datos específicos.
- **Ejecutar ETL:** Tiene la potestad de iniciar un proceso de actualización de datos fuera del horario programado.

2. Actor: Operador de Datos

Es el encargado de la ingestión de los datos más manuales.

- **Acceder al portal de Carga:** Ingresa a la sección específica para subir archivos.
- **Cargar Planilla de Datos:** Sube los archivos (ej. Excel) que contienen datos que no se pueden obtener automáticamente.

3. Actor: Administrador BI

Es el rol técnico que mantiene y monitorea los datos del sistema analítico.

- **Acceder al portal de Administración:** Ingresa a la consola de administración del sistema.
- **Monitorear Calidad de Datos (Staging):** Revisa los datos en la base de datos temporal (staging) para asegurar que se cargaron correctamente antes de moverlos al DW.
- **Ejecutar ETL:** Tiene la potestad de iniciar un proceso de actualización de datos fuera del horario programado, ya sea para probar el sistema o depurar el proceso.
- **Ejecutar Entrenamiento de Modelos:** Inicia manualmente el proceso de re-entrenamiento del modelo predictivo.

4. Actores: Sistemas Externos

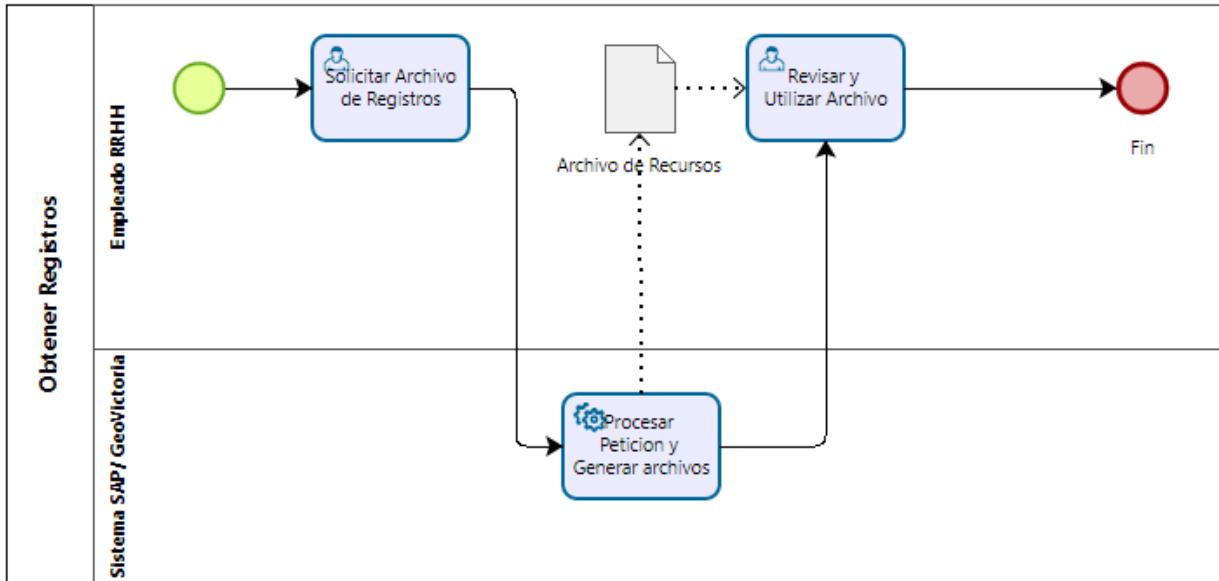
Son sistemas automáticos que interactúan con la plataforma.

- **Sistema SAP (Actor):**
 - **Ejecutar Carga SAP:** El sistema SAP exporta sus datos.
- **Planificador de Tareas (Actor):**
 - **Ejecutar Carga SAP:** Invoca la exportación de datos desde el sistema SAP.
 - **Ejecutar ETL:** Inicia automáticamente el proceso de Extracción, Transformación y Carga (ETL) que mueve los datos de Staging al Data Warehouse.
 - **Ejecutar Entrenamiento de Modelos:** Inicia automáticamente el proceso de re-entrenamiento del modelo predictivo.
- **Sistema de Autenticación (Actor):**
 - **Autenticarse en el sistema:** Valida las credenciales de los usuarios.

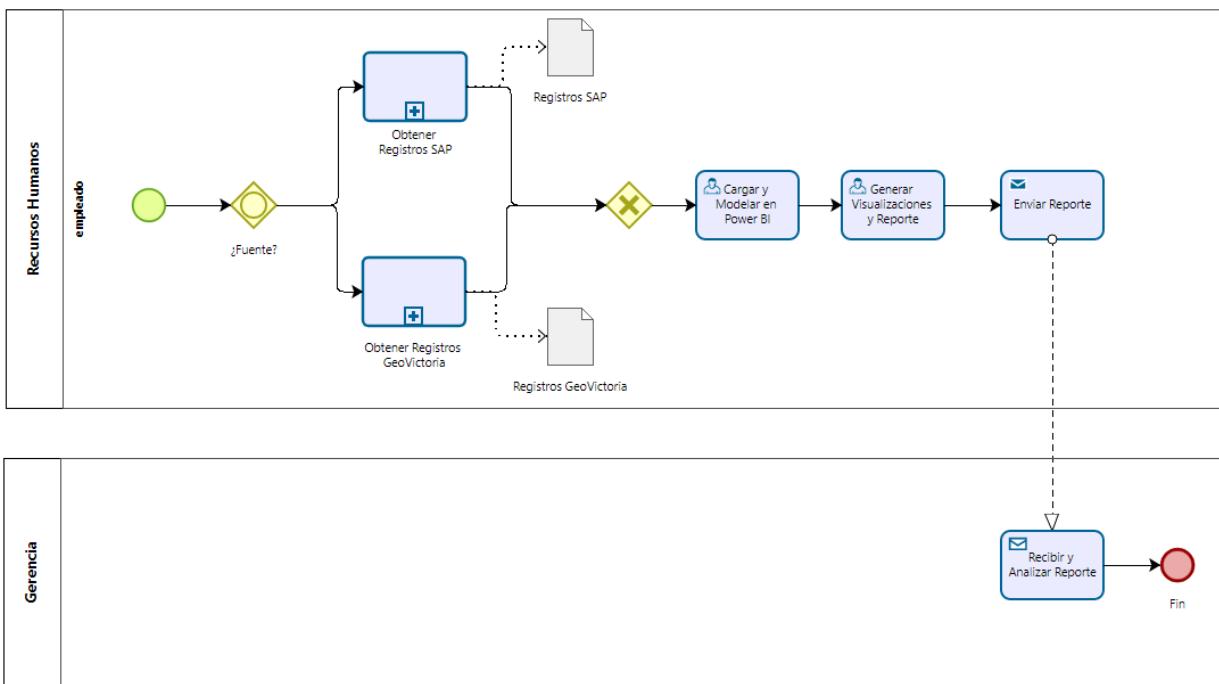
8. Modelado del Procesos de Negocio

El modelado de procesos de negocio es bastante sencillo comparado al sistema propuesto. Previo al sistema, los datos se extraen manualmente y las métricas se calculan a pulso.

Extracción manual de los datos de los principales sistemas presentes en la empresa (SAP-RRHH y GeoVictoria-Asistencia)



Unificación, modelamiento y cálculo de métricas desde los datos, para presentarse en gerencia y toma de decisiones.



9. Diseño de Interfaz de Usuario y Prototipo

9.1. Mockups

Panel de Modelos Predictivos

The screenshot shows a web-based application for HR analysis. The left sidebar has a dark blue header 'Análisis RRHH' with three menu items: 'Dashboard General' (selected), 'Análisis de Rotación' (highlighted in blue), and 'Administración'. The main content area is titled 'Análisis Predictivo de Rotación' with a 'Bienvenido, Juan Carlos' message and a 'Salir' button. It features a 'Filtros de Predicción' section with dropdowns for 'Área / Departamento' (set to 'Todas') and 'Cargo' (set to 'Todos'). Below this is a table titled 'Empleados con Alta Probabilidad de Rotación (Próximos 3 meses)'. The table lists three employees with their ID, name, role, department, and rotation probability:

ID Empleado	Nombre	Cargo	Área	Probabilidad de Rotación
1054	Ana Martínez	Croupier	Mesas de Juego	85% (Alta)
2311	Carlos Solo	Croupier	Mesas de Juego	78% (Alta)
1890	Sofía Vega	Cajera	Cajas	72% (Media Alta)

Panel de KPIs

The screenshot shows a 'Dashboard General' section within the HR analysis application. The left sidebar is identical to the previous panel. The dashboard displays four key performance indicators: 'Tasa de Rotación (Anual)' at 15.2% (▲ 2.1% vs año anterior), 'Tasa de Ausentismo (Mensual)' at 4.8% (▼ 0.5% vs mes anterior), 'Dotación Actual (Headcount)' at 485 (▲ 12 nuevas contrataciones), and 'Antigüedad Promedio' at 3.2 Años (Constante). Below these metrics is a section titled 'Dotación de Personal por Área' which contains a large black rectangular placeholder with white text: 'The image you are requesting does not exist or is no longer available.' and 'imgur.com'.

Panel de Carga y Administración

Bienvenido, Juan Carlos [Salir](#)

Análisis RRHH

- Dashboard General
- Análisis de Rotación
- Administración**

Administración del Sistema

Carga de Datos (ETL)

Sube los archivos Excel con los registros de asistencia, altas/bajas, etc.

Seleccionar Planilla Excel

Sin archivos seleccionados

Exportación de Datos

Exporta los datos consolidados del Data Warehouse a diferentes formatos.

Seleccionar Formato

Gestión de Acceso (AAA)

Administra los roles y permisos de los usuarios del sistema.

Usuario	Rol	Último Acceso	Acciones
jyampara	Administrador	2025-10-18 15:30	Editar Revocar
tsilva	Analista	2025-10-18 14:00	Editar Revocar
gerencia	Visualizador	2025-10-17 11:00	Editar Revocar



10. Conclusiones

Este documento ha detallado el plan y la arquitectura para abordar un problema tangible en el área de RRHH de Casino Luckia Arica: la fragmentación de sus datos. La situación actual se caracteriza por la dispersión de la información del personal, lo que deriva en un trabajo manual intensivo y limita la toma de decisiones a un plano reactivo.

La solución propuesta es diseñar un Data Warehouse (un almacén de datos) para centralizar esta información. Se planea usar procesos ETL (Extracción, Transformación y Carga) para mover y limpiar los datos. El objetivo es que el área de RRHH pueda consultar métricas clave, como la rotación o el ausentismo, sin tener que armar reportes a mano cada vez.

Para gestionar el desarrollo, se eligió Kanban por su flexibilidad y menor carga administrativa, lo que parece adecuado para el tiempo ajustado del proyecto. El proceso de análisis de datos seguirá las fases de CRISP-DM, que es un estándar probado para proyectos de este tipo.

Si la ejecución de este plan es exitosa, se espera que la plataforma resultante reduzca significativamente la carga operativa. Más allá de la mera visualización de métricas en un dashboard, la verdadera prueba de éxito del proyecto residirá en la capacidad del modelo predictivo para anticipar la rotación de personal con un grado de precisión que resulte útil para la gestión.

El plan es ambicioso. Su viabilidad dependerá críticamente de dos factores principales: primero, la calidad, completitud e integridad de los datos históricos que se logren extraer de los sistemas fuente (SAP y Geovictoria) y, segundo, la disciplina en la ejecución de las tareas definidas en el backlog. Se espera que el diseño aquí presentado siente una base técnica sólida para alcanzar dichos objetivos.

Buen Informe, debe ser más coherente con la justificación del texto
Debe evitar al máximo la intervención humana, porque los errores son directamente proporcional al numero de intervenciones.
La definición del problema es acotada, y la solución seleccionada se logra entender el problema y la buena solución que se estima.
Faltan numerar las figuras, tablas y referenciarlas desde el texto.
El BPN no esta tan claro, como también el contexto.. hay que darle una refinada.
excelente carta gant,

11. Bibliografía

- [1] D. J. Anderson, *Kanban: Successful Evolutionary Change for Your Technology Business*. Sequim, WA: Blue Hole Press, 2010.
- [2] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in *Proc. 4th Int. Conf. Practical Application of Knowledge Discovery and Data Mining*, Manchester, UK, 2000, pp. 29–39.
- [3] S. K. Wagh, A. A. Andhale, K. S. Wagh, J. R. Pansare, S. P. Ambadekar, and S. H. Gawande, "Customer churn prediction in telecom sector using machine learning techniques," *Results Control Optim.*, vol. 14, p. 100342, Mar. 2024, doi: 10.1016/j.rico.2023.100342.
- [4] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3rd ed. Hoboken, NJ: John Wiley & Sons, 2013.
- [5] E. Turban, R. Sharda, and D. Delen, *Business Intelligence: A Managerial Approach*, 2nd ed. Upper Saddle River, NJ: Pearson, 2014.
- [6] F. Diez, M. Bussin, and V. Lee, *Fundamentals of HR Analytics: A Manual on Becoming HR Analytical*. Bingley, UK: Emerald Publishing Limited, 2019.