

| **1. Informe final Proyecto APT** |
| --- |
| El objetivo de este informe es que describas los aspectos más relevantes de tu Proyecto APT. Es importante que fundamentes las decisiones que tuviste que tomar a lo largo del proceso.  A continuación, encontrarás distintos campos que deberás completar con la información solicitada, los que dan cuenta del resumen de tu proyecto APT y sus principales resultados. |

| Nombre del proyecto | Sistema y Gestión Middleware ETL |
| --- | --- |
| Área (s) de desempeño(s) | Automatización, Ingeniería de datos y Desarrollo de software |
| Competencias | Ofrecer propuestas de solución según los requerimientos de la organización.  Integración y correcta administración de los distintos ambientes informáticos, logrando una correcta operatividad por parte de la empresa.  Desarrollo de la solución utilizando software de alto calibre, escalable y mantenible. asegurando el logro de los objetivos.  Implementar soluciones sistémicas integrales para automatizar y optimizar procesos de negocio de acuerdo a las necesidades de la organización.  Resolver las vulnerabilidades sistémicas para asegurar que el software construido cumple las normas de seguridad exigidas por la industria.  Programar consultas o rutinas para manipular información de una base de datos de acuerdo a los requerimientos de la organización. |

| **Contenidos del informe final** | |
| --- | --- |
| 1. Relevancia del proyecto APT | El proyecto busca solucionar la problemática de la fragmentación de datos en Recursos Humanos (RR.HH.) dentro de la empresa Cencosud, una empresa con operaciones a nivel internacional. La necesidad de acceder a información unificada de todos sus empleados en diferentes países se ve obstaculizada por la dispersión de los datos de RR.HH. en sistemas y bases de datos con estructuras y enfoques distintos, dificultando la creación de una visión global y coherente de la fuerza laboral.  Esta fragmentación dificulta la toma de decisiones estratégicas basadas en datos a nivel global, la gestión del talento en diferentes regiones y la creación de programas de desarrollo efectivos que se adapten a las necesidades específicas de cada país. Esta problemática es crucial para el campo laboral de la gestión de datos y la analítica, ya que la necesidad de integrar y analizar datos de diferentes fuentes es una realidad en la mayoría de las empresas multinacionales.  El proyecto propone la implementación de un "middleware" que actúe como puente entre las diferentes fuentes de datos a nivel internacional. Este middleware centralizará la información en un modelo unificado de entidades relacionadas entre sí, permitiendo:   * **Automatizar el proceso ETL:** Se implementará un proceso automatizado de extracción, transformación y carga de datos (ETL) para integrar la información de las diferentes fuentes, creando una base de datos centralizada y consistente a nivel global. * **Integración de Clusters de Machine Learning:** El análisis de datos se verá enriquecido mediante la incorporación de algoritmos de aprendizaje automático, lo que permitirá obtener insights más profundos y generar predicciones a nivel global y regional. * **Accesibilidad:** Los datos procesados podrán ser consumidos por diferentes plataformas, como Power BI, la plataforma web interna de la empresa, el equipo de científicos de datos, el departamento de inteligencia de negocio y RR.HH., tanto a nivel central como en las diferentes regiones.   Con esto aportamos una solución real al contexto laboral de Cencosud a nivel internacional, proporcionando una herramienta para integrar y analizar datos de RR.HH. de manera eficiente y accesible. Este sistema de datos centralizado y unificado, basado en un modelo de entidades interrelacionadas, facilitará la gestión de los diferentes sectores que requieren acceder a información de RR.HH. a nivel global, permitiendo optimizar la gestión del talento, mejorar la toma de decisiones estratégicas y fortalecer el desarrollo de la empresa a nivel internacional. |
| 2. Objetivos | Como objetivos principales, se busca proporcionar un software autogestionable, escalable y mantenible en el tiempo. Para esto, desglosamos los objetivos específicos para este proyecto.   * Proporcionar un sistema agnóstico a diferentes fuentes de datos. * Generar Gráficos proporcionales con gran valor estadístico. * Procesar tareas automáticas frecuentemente. * Sistema autogestionable. * Sistema con salida a escalabilidad. * Datos con gran valor estadístico y con salida a crecimiento constante. |
| 3. Metodología | Actualmente contamos con metodología ágil, para este proyecto, específicamente SCRUM. Donde contaremos con una serie de factores positivos para la realización y la plenitud constante de este proyecto.  Por un lado contaremos con los requerimientos principales, y levantamiento de la propuesta a realizar, esto nos ayudará a tener mejor conocimiento de las funcionalidades que debe cumplir el proyecto, evitando fallas de conceptos por parte del equipo SCRUM, esta fase ya fue realizada por parte de nuestro miembro con la función de Product Owner, donde realizó todo el proceso de gestión y comprensión inicial con nuestros stakeholders.  Luego, una vez definido nuestros requerimientos y los estándares a cumplir, comenzaremos con los primeros SPRINTS, donde se recalcó el utilizar las historia de usuario como tareas a resolver. Estos sprints van por fase y se dió como máximo, 3 semanas para completar todas las tareas o historias de usuario dentro de un SPRINT. Para llevar la fase de los sprints por un buen camino, realizaremos reuniones diarias, o más conocidas como Daily Meeting, donde utilizaremos estas reuniones para dar a conocer nuestros avances o retrasos al estar resolviendo una tarea definida.  Estos básicos conceptos de metodologías nos bastará para llevar por un buen camino el proyecto que se está realizando, además nos ayuda a mantener los requerimientos a flote ante cualquier cambio repentino, donde aprovecharemos los Sprints para dar a conocer nuevos requerimientos que surgieron, o conceptos de estos. |
| 4. Desarrollo | Actualmente contamos con una serie de objetivos o requerimientos a utilizar, divididos por historias de usuarios. Donde utilizaremos los sprints como etapas en el desarrollo del proyecto. Al día de hoy, contamos con un mvp preliminar donde este cuenta con las funcionalidades principales del proyecto, el cual abordamos todo el proceso ETL, estadística descriptiva, y gestión de datos.  Las dificultades al abordar este proyecto, es el tema de los datos generados artificialmente con lógica muy específica, ya que este proyecto cuenta con una solución enfocada al sector de recursos humanos de la empresa Cencosud, pero al ser datos sensibles y al estar protegido por regimens, se dificulta la utilización de estos para abordar el proyecto APT. A pesar de este problema muy importante, supimos resolverlo, y actualmente contamos con una fase muy avanzada del proyecto, el cual cumple con las expectativas impuestas inicialmente.  En un inicio, presentamos la solución principal, el cual aborda la problemática presentada al inicio de este documento. Pero como se recalcó, este desarrollo tuvo que realizarse con un enfoque de datos diferente, pero seguir con la idea de un desarrollo autogestionado, y sirviendo como intermediario entre varios sistemas de datos. |
| 5. Evidencias | **Principales pipelines de datos:**    **Bases de datos pobladas exitosamente con procesos ETL:**          **Dashboard con datos de la base de datos:** |
| 6. Intereses y proyecciones profesionales | El actual proyecto nos ayudó a profundizar en los conceptos de ciencias de datos, manejo de bases de datos, desarrollo web, estadísticas descriptivas, entre otros. Al desarrollar proyectos con mayor diversidad y abordando varios sectores y saber correlacionar los, nos proporciona mayor visión hacia donde nos queremos dedicar en el futuro, ya que estos trabajos son tan complejos que nos permiten tener que profundizar a grandes rasgos, todo lo que aprendimos durante la carrera.  Al principio de la carrera, o a los primeros años, teníamos una visión diferentes sobre el mundo de la informática en general, pero al pasar de los semestres, encontramos varios sectores muy interesantes, como por ejemplo, el mundo de los datos, que es en donde nos basamos para construir nuestro proyecto, y en donde aguarda el corazón de nuestro sistema. Actualmente nuestras preferencias van más por este camino, de integración continua con múltiples plataformas de tecnología, interactuando con datos, desarrollo web, estadística, big data, etc.  En el mundo laboral, planeamos empezar por estos sectores que manejamos, profundizando y buscando aprender cada vez más de este fascinante sector de los datos. |