



## Guía2. Desarrollo Proyecto APT Asignatura Capstone

### 1. Resumen avance Proyecto APT

Resumen de avance proyecto APT	Hasta ahora, hemos desarrollado una arquitectura robusta y escalable para el Middleware de Cencosud, centralizando los datos de RR.HH. a nivel internacional. Hemos implementado mecanismos avanzados de validación y limpieza de datos para asegurar su calidad e integridad. El proceso diario de ETL ha sido configurado para extraer datos de diversas fuentes, transformarlos y cargarlos en un Datawarehouse que facilita la inteligencia de negocios. Se han realizado varias pruebas de transformación y calidad de datos, ajustando el proceso para mejorar la precisión y consistencia.
Objetivos	No se han realizado ajustes a los objetivos hasta ahora.
Metodología	No se han realizado ajustes a la metodología hasta ahora.

<p>Evidencias de avance</p>	<p>Implementación del proceso ETL utilizando Apache Airflow.</p> <p>Configuración y carga de un Datawarehouse en Azure.</p> <p>Utilización de Great Expectations para la validación de datos.</p> <p>Documentación técnica actualizada que detalla la arquitectura del sistema y los procesos ETL.</p> <p>Estas evidencias permiten demostrar que se han implementado las fases críticas del proyecto, asegurando la calidad y accesibilidad de los datos de RR.HH.</p> <p>En este apartado adjuntar la(s) evidencia(s) seleccionada(s) para ser evaluada por el docente.</p>
-----------------------------	---

## 2. Monitoreo del Plan de Trabajo

Plan de Trabajo							
Competencia o unidades de competencias	Actividades	Recursos	Duración de la actividad	Responsable <sup>1</sup>	Observaciones	Estado de avance	Ajustes
Integración de datos	Implementar ETL con Apache Airflow	Airflow, Python, Azure	4 semanas	Jhopcel Rosales	Configuración inicial compleja	En curso	Ninguno
Validación de datos	Uso de Great Expectations	Great Expectations	3 semanas	Benjamin Pavez	Requiere ajuste de reglas de validación	En curso	Ajuste de reglas de validación
Desarrollo de Datawarehouse	Configurar y optimizar Datawarehouse en Azure	Azure	5 semanas	Jhopcel Rosales	Requiere mantenimiento continuo	En curso	Ninguno
Documentación técnica	Actualización de documentación	Documentos técnicos	2 semanas	Benjamin Pavez	Necesita revisión periódica	En curso	Ninguno

<sup>1</sup> En caso de que el Proyecto APT sea grupal, en esta columna deben indicar el nombre de los responsables de cada tarea o actividad. Esto posteriormente permitirá diferenciar la evaluación por cada integrante.

Pruebas y calidad de datos	Realizar pruebas de transformación y calidad de datos	Datos de prueba	3 semanas	Sofía Martínez	Resultados deben ser precisos	En curso	Mejora de procedimientos de prueba
----------------------------	---	-----------------	-----------	----------------	-------------------------------	----------	------------------------------------

### 3. Ajustes a partir del monitoreo

Factores que han facilitado y/o dificultado el desarrollo de mi plan de trabajo:

**Facilitadores:** La disponibilidad de herramientas avanzadas como Apache Airflow y Azure, y la colaboración efectiva del equipo han facilitado el desarrollo del proyecto.

**Dificultades:** La configuración inicial de los sistemas y la necesidad de ajustar las reglas de validación de datos han sido los principales obstáculos. Para solucionarlos, se han implementado sesiones de revisión y ajustes continuos.

Actividades ajustadas o eliminadas:

**Ajustadas:** Las reglas de validación de datos con Great Expectations se ajustaron para mejorar la precisión.

**Eliminadas:** Ninguna actividad ha sido eliminada.

Actividades que no has iniciado o están retrasadas:

*Las actividades de mantenimiento continuo del Datawarehouse están en curso pero podrían requerir más tiempo del estimado. La estrategia es asignar recursos adicionales para cumplir con los plazos establecidos.*