



Business Analytics

Ejercicio teórico técnico


Data

Engineer

Semi-Senior

 **Sitios Web**

<https://www.taligent.es>

 **Emails**

info@taligent.es

 **Teléfono**

+34 682 25 55 15

 **Ubicaciones**

C. de Caravaca, 11, El Llano del
Real, 46021 Valencia, España

Criterios de Evaluación	3
Duración	3
Consigna	4
1. Airflow	4
2. OpenMetadata	4
3. PostgreSQL	4
Integración y Demostración	5

Criterios de Evaluación

- **Instalación y Configuración:** Se evaluará la capacidad para instalar y configurar cada una de las herramientas (utilizar Docker).
- **Funcionalidad de los DAGs:** Los DAGs deben ejecutarse sin errores y realizar correctamente las tareas de ETL.
- **Integración:** La integración entre Airflow, OpenMetadata y PostgreSQL debe ser fluida y sin problemas.
- **Documentación:** La documentación debe ser clara, detallada y fácil de seguir.
- **Demostración:** La demostración debe mostrar claramente el funcionamiento completo de las herramientas y su integración.

Duración

- **Tiempo Estimado:** ≈30 horas.

Consigna

Para este ejercicio técnico, los candidatos deberán demostrar su capacidad para implementar y utilizar tres herramientas clave en un entorno on-premise: Airflow, OpenMetadata y PostgreSQL, utilizando Docker. Aquí está la descripción detallada del ejercicio:

1. Airflow

- **Objetivo:** Implementar Apache Airflow en un entorno local y crear DAGs (Directed Acyclic Graphs).
- **Requerimientos:**
 - Instalar y configurar Airflow en un entorno local.
 - Crear al menos dos DAGs que realicen tareas de ETL (Extract, Transform, Load).
 - Asegurarse de que los DAGs estén programados y se ejecuten sin errores.
 - Documentar los pasos de instalación, configuración y creación de DAGs.

2. OpenMetadata

- **Objetivo:** Implementar OpenMetadata y generar un entendimiento claro de su funcionalidad.
- **Requerimientos:**
 - Instalar y configurar OpenMetadata en un entorno local.
 - Integrar OpenMetadata con Airflow y PostgreSQL.
 - Documentar el proceso de instalación y configuración.
 - Explicar las funcionalidades clave de OpenMetadata y cómo puede ayudar en el gobierno de datos.

3. PostgreSQL

- **Objetivo:** Crear y gestionar bases de datos utilizando PostgreSQL.
- **Requerimientos:**
 - Instalar y configurar PostgreSQL en un entorno local.
 - Crear al menos tres tablas en PostgreSQL.
 - Implementar un proceso mínimo de ETL utilizando DAGs en Airflow, donde los datos se extraigan, transformen y carguen en las tablas de PostgreSQL.

- Documentar el proceso de creación de tablas y el flujo ETL.

Integración y Demostración

- **Objetivo:** Demostrar la integración completa y el funcionamiento de las tres herramientas.
- **Requerimientos:**
 - Asegurarse de que los DAGs de Airflow se ejecuten correctamente y se reflejen en OpenMetadata.
 - Verificar que la metadata de PostgreSQL sea visible en OpenMetadata.
 - Documentar todo el flujo de trabajo y proporcionar una demostración que muestre la ejecución completa desde la creación de tablas en PostgreSQL, pasando por la ejecución de DAGs en Airflow, hasta la visualización de la metadata en OpenMetadata.