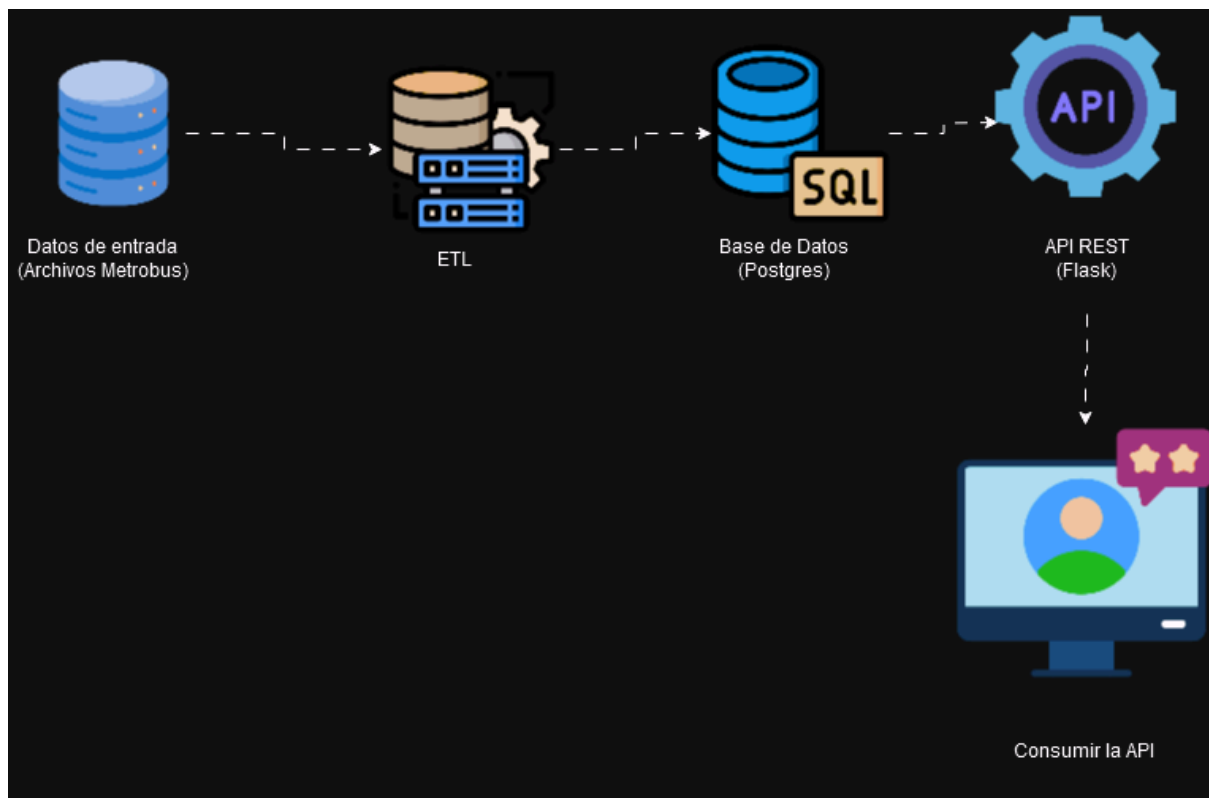


## Diagrama Data Pipeline



### 1. Datos de Entrada:

- Los datos de entrada son los archivos estáticos que contienen información sobre las rutas, paradas, horarios, etc., del metrobús.
- Estos datos se obtienen de fuentes abiertas de la Ciudad de México.

### 2. Procesamiento (ETL):

- **Extracción:** Los datos se extraen de los archivos
- **Transformación:** Los datos se transforman para asignar la alcaldía correspondiente a cada parada y se preparan para su almacenamiento.
- **Carga:** Los datos transformados se cargan en la base de datos.

### 3. Base de Datos (PostgreSQL):

- Los datos procesados se almacenan en una base de datos PostgreSQL.
- Las tablas principales incluyen:
  - **rutas:** Información sobre las rutas del metrobús.
  - **stops:** Información sobre las paradas.
  - **stop\_times:** Horarios de las paradas.
  - **alcaldias:** Asignación de alcaldías a las paradas.

### 4. API REST (Flask):

- La API expone los datos almacenados en la base de datos a través de endpoints REST.
- Los endpoints principales son:
  - **/unidades:** Obtener la lista de unidades disponibles.
  - **/unidades/<id>/ubicaciones:** Obtener las ubicaciones de una unidad específica.
  - **/alcaldias:** Obtener la lista de alcaldías disponibles.
  - **/alcaldias/<nombre>/paradas:** Obtener las paradas dentro de una alcaldía específica.

