



BIG DATA
ACADEMY

LABORATORIO 53

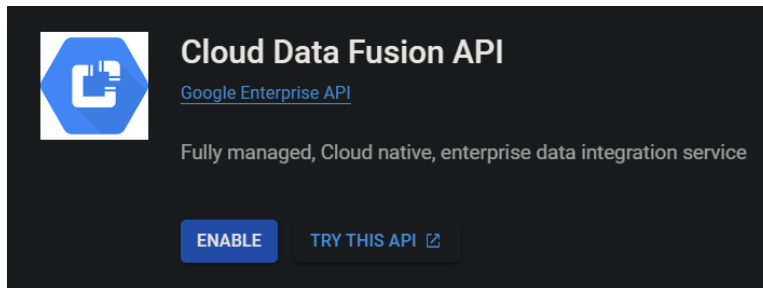
DESPLIEGUE DATA FUSION

FORMADOR: ALONSO MELGAREJO
alonsoraulmgs@gmail.com

LABORATORIO 53 – DESPLIEGUE DATA FUSION

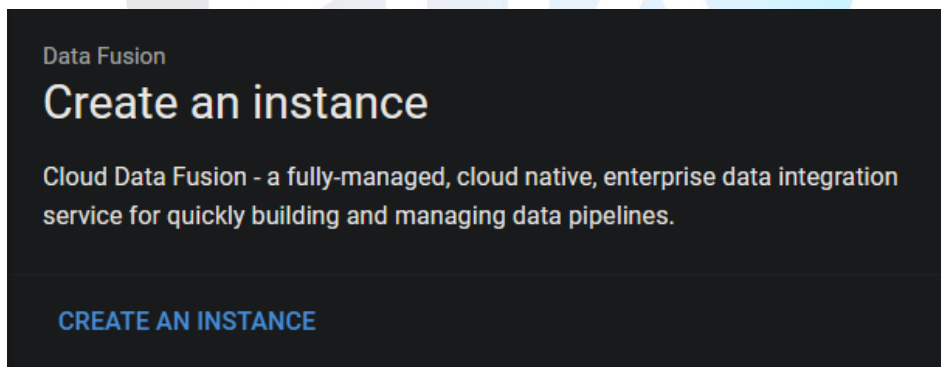
1. Data Fusion nos permite construir pasos de procesamiento (TO_SILVER, TO_GOLD).
2. Desde el buscador de servicios, buscamos:

3. Habilitamos el servicio dando clic en “Enable”



TIEMPO: 1 MINUTO (IMPORTANTE: SI LA HABILITACIÓN TOMA MÁS DE 1 MINUTO, REFRESCAR LA PÁGINA Y PROBAR NUEVAMENTE)

4. Seleccionamos “Create an instance”




5. Configuramos:

Instance name	serveretl
Región	us-west1
Versión	6.x
Edition	Developer

6. Para poder usar “Data Fusion” debemos de indicar quién es el dueño es el (ownerdatalake), además el dueño debe tener los permisos de uso de Data Fusion. Seleccionamos al dueño “ownerdatalake” y damos clic sobre “Grant Permission” para otorgarle los permisos

Authorization

 Cloud Data Fusion does not have permission to use the selected Dataproc Service Account. These permissions are required to launch Dataproc clusters in customer projects for executing pipelines. It is necessary to grant permission before the Cloud Data Fusion instance can be created.

[GRANT PERMISSION](#) [MORE INFO](#)

Dataproc Service Account

ownerdatalake@datalake-381415.iam.gserviceaccount.com

A service account is used to authorize the instance to run a Dataproc job.

Advanced Options

The instance creation will take approximately 20 minutes.

[CREATE](#) [CANCEL](#)

Damos clic en "Create"

TIEMPO: 20 MINUTOS

