

**Parcial 1**

**Big data e ingeniería de datos**

**Juan Carlos Tarazona Contreras**

**Diana Carolina Quintero Bedoya**

**Bogotá D.C.**

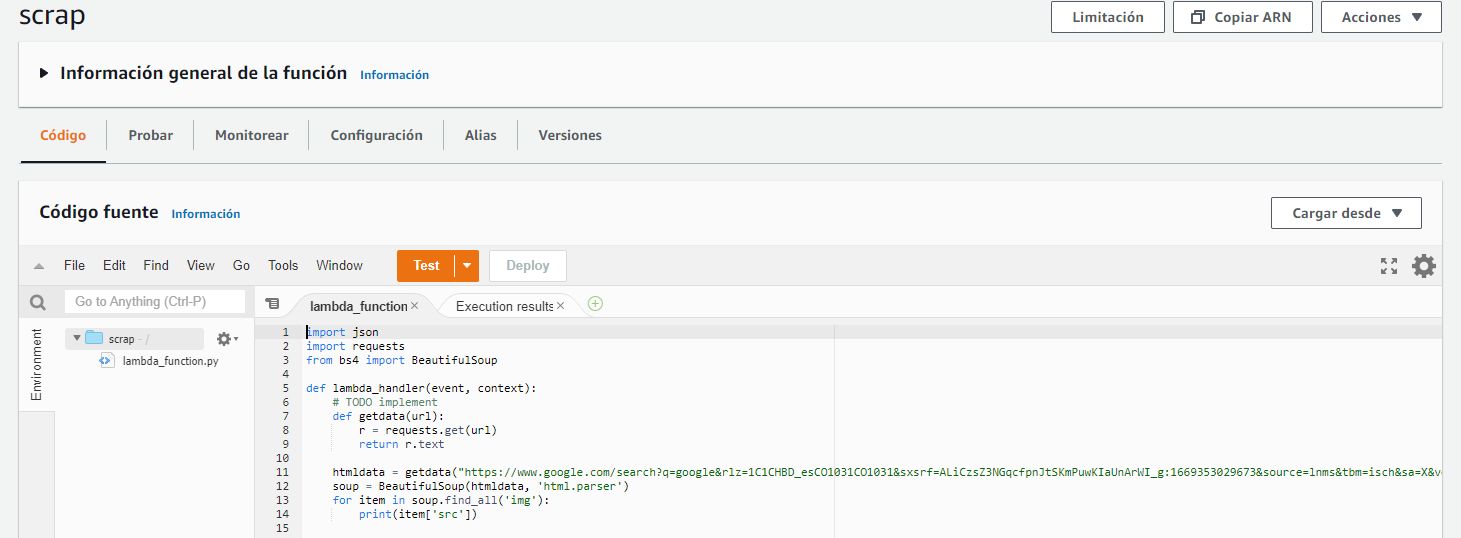
**1) Crear un lambda que realice scraping de una página y descargue las imágenes que se encuentren en dicha página(jpg o png) a un bucket en s3. La url de la página se debe leer de un archivo en S3. (10 puntos).**

Se realiza la creación del lambda teniendo el llamado “scraping” como la prueba de esta primera función y siguiente el lambda usado llamado “scrap” el cual podemos notar como una principal diferencia que este fue realizado en python 3.9.

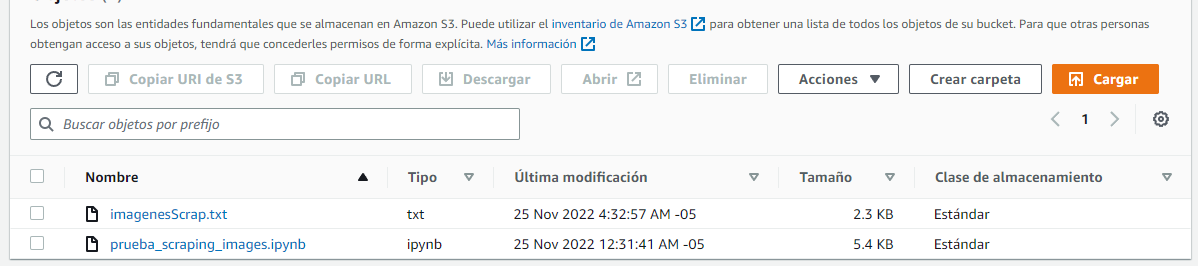
Para su completo funcionamiento se le agrega el labrol.



Se presenta el código implementado.

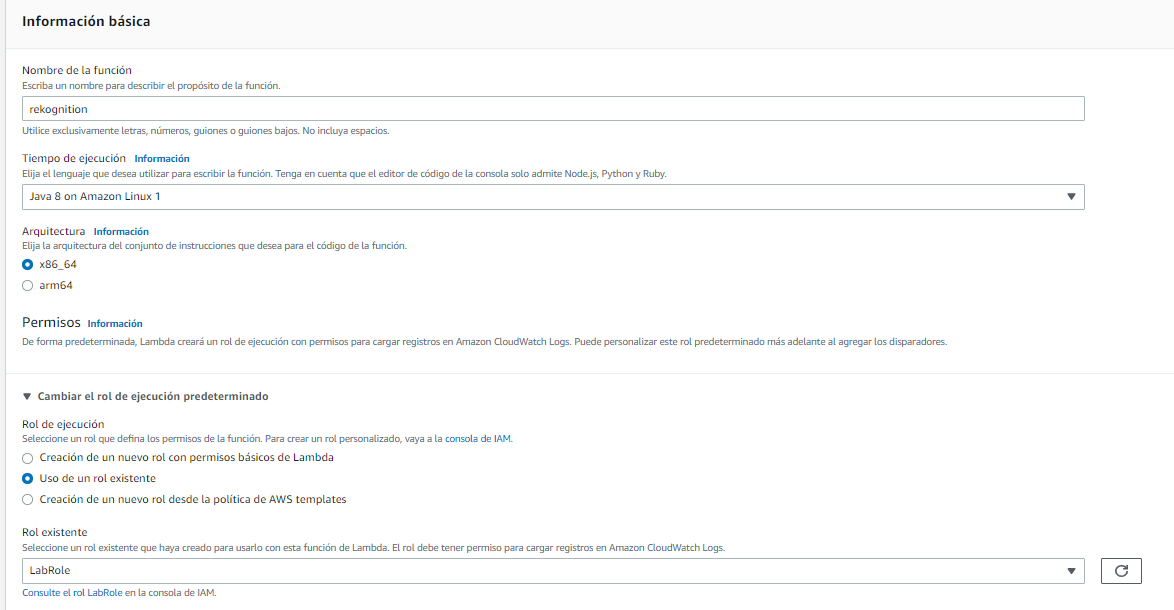


Enviando en un txt las imágenes como archivos url.

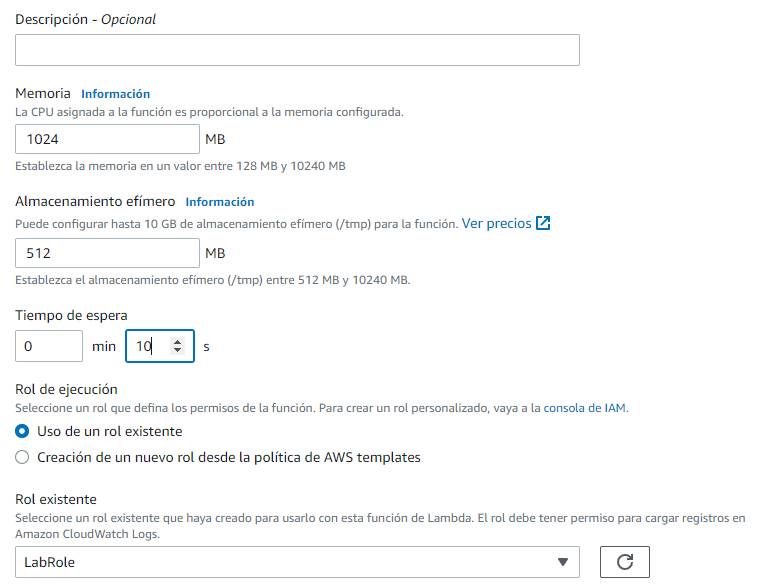


**2) Crear un segundo lambda (disparado cuando llega el archivo), que tome la imagen y utilice AWS Rekognition para determinar qué objetos hay en la imagen(5 puntos).**

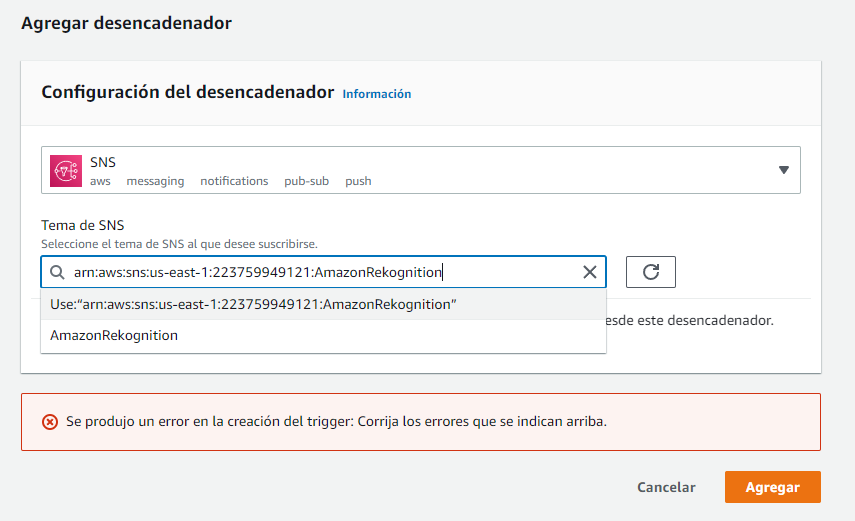
Se hace la configuración necesaria para la nueva función.



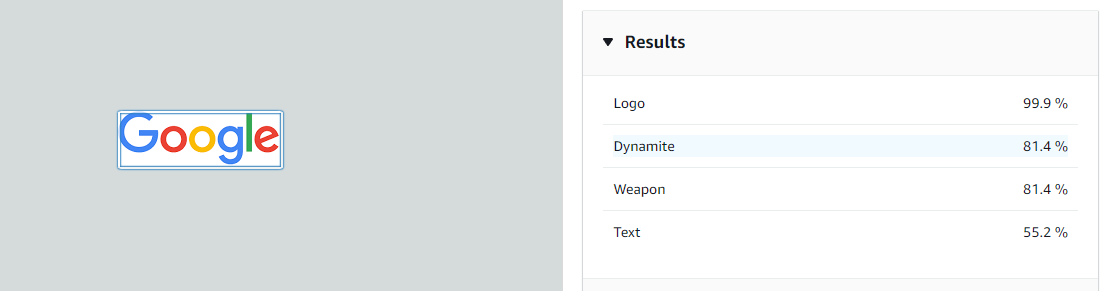
Se cambia la configuración básica.

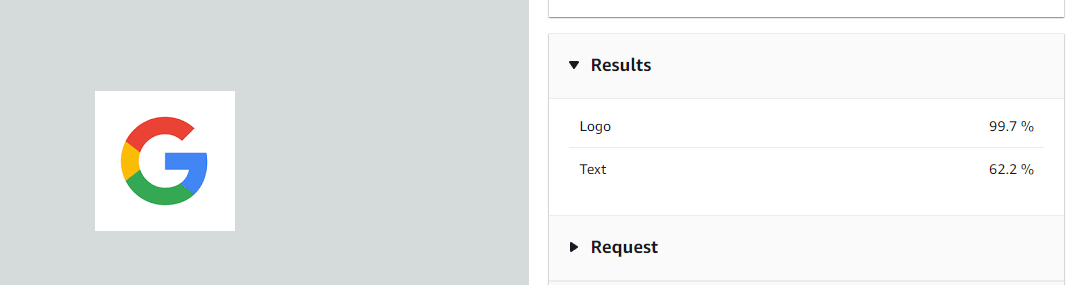


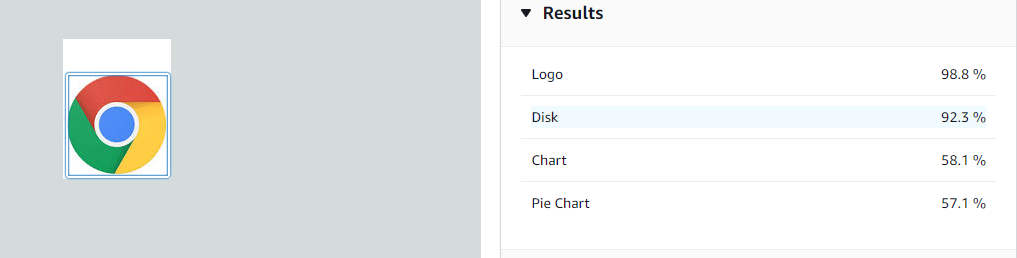
Agregamos el sns en el desencadenador para que este realice la ejecución por medio de la función lambda.

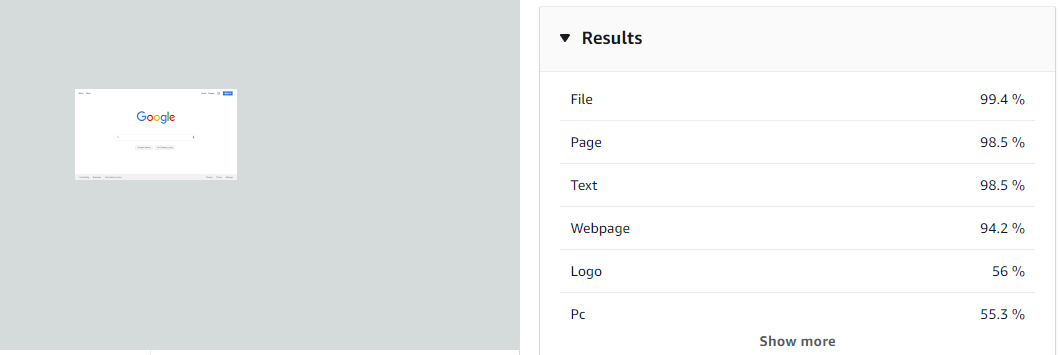


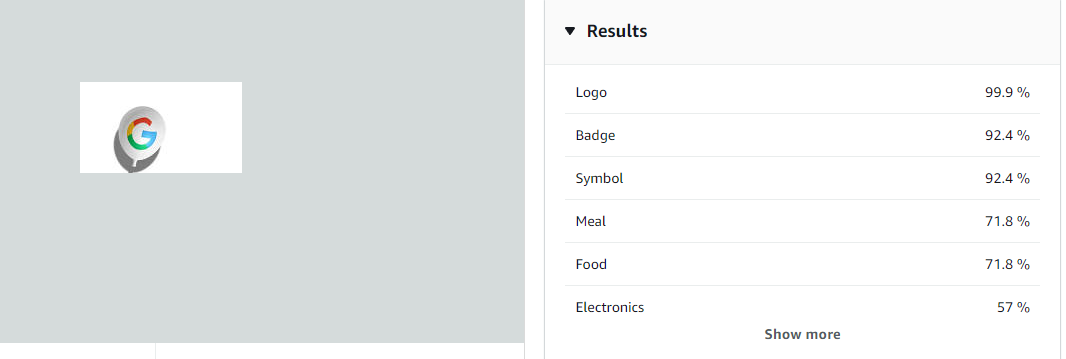
Tenemos como reconocimiento de algunas imágenes.











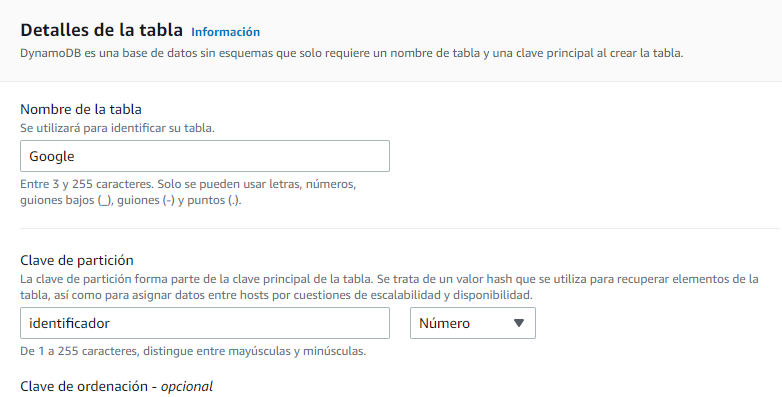
**3) Insertar en una BD dynamo la siguiente información: (10 puntos)**

**índice: Label principal de la imagen (por ej. Lion)**

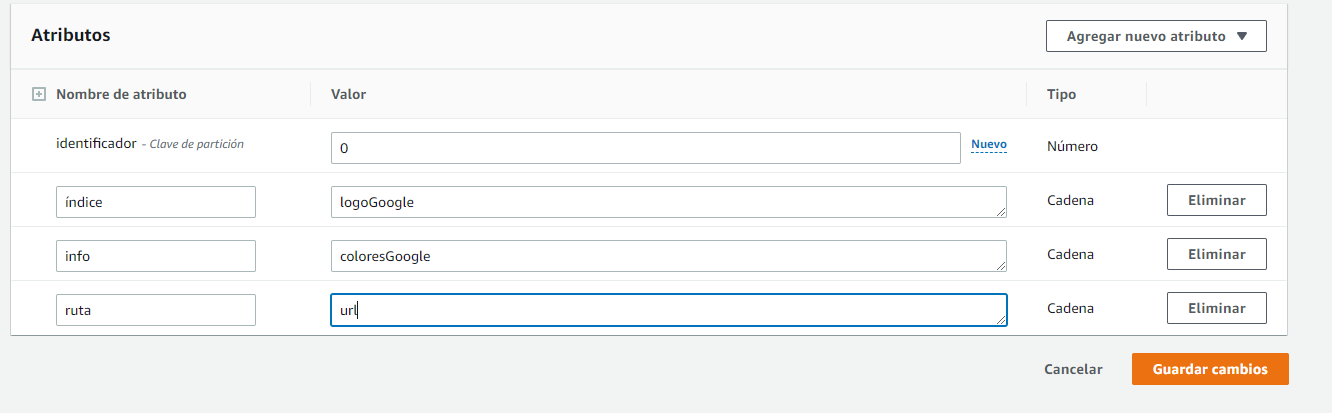
**info: Labels secundarios (Animal, mammal, ...)**

**ruta: ubicación de la imagen**

Se crea la tabla de nuestra base de datos Dynamo



Inserción de los datos de la tabla.



Finalmente nos encontramos con las importaciones de las imágenes a el DynamoDB.

