Diseño del DAaaS

### **Definición la estrategia del DAaaS**

Crear un servicio web para la automatización de la búsqueda de literatura académica. Con la web Leidrac MSL, los investigadores podrán reducir los tiempos en la creación del mapeo. Para lograrlo, se deberá acceder al sitio web y copiar la URL de la página o páginas que contienen la información que se desea para construir el MSL (https://www.scopus.com y/o https://www.elsevier.com y/o https://scholar.google.com/). Una vez hecho esto, se activará la creación de un usuario para la facturación, creación de máquinas virtuales, buckets y acciones de crawler para generar un archivo CSV por cada URL.

Las máquinas virtuales y el almacenamiento se eliminarán una vez finalizado el proceso. Cada archivo contendrá los detalles del título del artículo, autores, revista/conferencia, año de publicación, resumen, palabras clave, área temática, metodología, resultados principales, conclusiones, calidad del estudio y referencias.

El investigador podrá descargar un archivo de Power BI para visualizar los datos obtenidos y los resultados de las URL en archivos csv. Tiene un costo asociado por cada descarga de CSV. El límite máximo de registros por URL es de 100. El servicio web actuará con una API HTTP RESTful.

**

*Imagen aprox del PBI*

### Arquitectura DAaaS

*Fase de Salida*

Cloud function + WEB + URL de Bucket de los crawler + URL de descarga de Archivo de PBI + eliminación de los Buckets

*Fase de Proceso*

Cloud function + ejecución del crawler en MV + bucket de los crawler + Crawler repositorio

Cloud function + creación de MV + creación de Bucket de los crawler + mongoDB+ extracción de URL de bucket URL

*Entrada de URLs a través de web + creación de usuario + activación de perfil + sistema de facturación + creación de bucket de URL*

*Fase de Inicio*

Cloud function +Crear *Formulario de usuario y datos de método de pago*

*App Emgine para la creacion de la web + Python + SDK + Flask*

### DAaaS Operating Model Design and Rollout

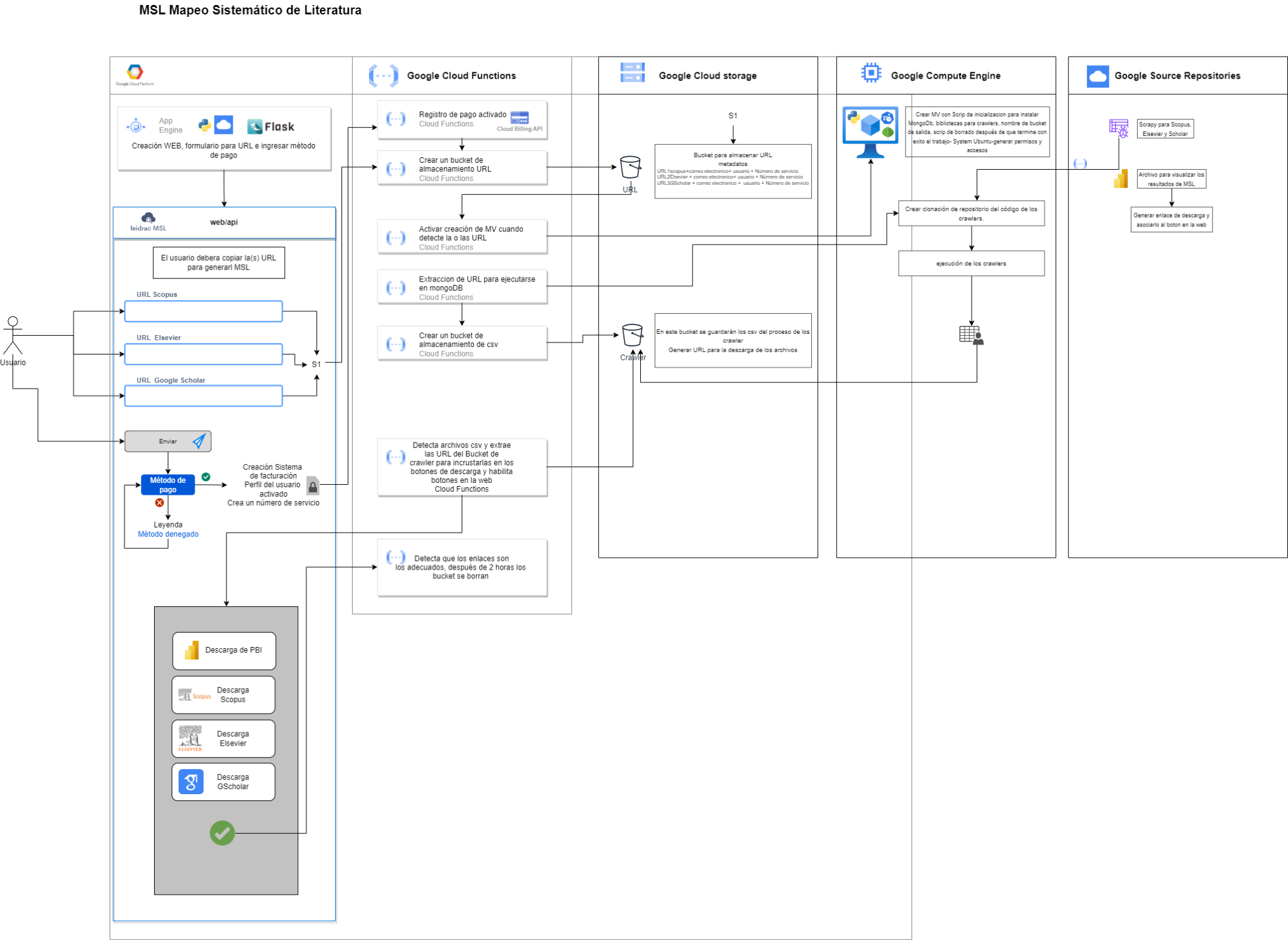
1. *Se construye una interfaz web con dos formularios: uno para ingresar las tres* URL y otro para la creación del usuario con método de pago encriptado. Al diseño se agregan cuatro botones de descarga: uno para cada URL y otro para descargar el archivo de Power BI. Estos se activan y desactivan en función del proceso.
2. Se realiza manualmente la importación del crawler de cada URL y del archivo de Power BI al Source Repositories.
3. Cuando el usuario copia una o varias URL y el método de pago está activo, se genera a través de una función el Bucket “URL” para guardar cada una de ellas con un nombre único y con metadatos asociados.
4. Mediante una serie de funciones, se activa la creación de una máquina virtual (MV) en la que se instala MongoDB, las bibliotecas necesarias para ejecutar los crawlers, zona, tipo de SO, memoria… También se activan los puertos necesarios y se genera el bucket del crawler.
5. A través de una serie de funciones, se genera una clonación del crawler en la MV, se extraen las URL del Bucket, se ejecuta el crawler dando como resultado un archivo CSV por cada URL. Estos archivos se guardan en el bucket del crawler.
6. Otra función detecta los archivos CSV en el Bucket crawler, extrae las URLs del bucket para incrustarlas en los botones de la web. De esta forma, el usuario puede activar la descarga haciendo clic en los botones, los cuales contienen enlaces de descarga asociados.
7. Posteriormente, otra función verifica que los enlaces y los archivos estén disponibles en los botones. Transcurridas dos horas desde la activación del proceso, se eliminan los buckets y la MV.

*Personalizar los modelos operativos DAaaS para cumplir con los procesos, la estructura organizacional, las reglas y el gobierno de los clientes individuales. Realizar seguimiento de consumo y mecanismos de informe.*

### Desarrollo de la plataforma DAaaS. (ligera descripción del desarrollo)

*Construcción iterativa de todas las capacidades de la plataforma, incluido el diseño, desarrollo e integración,* ***pruebas****, carga de datos, metadatos y población de catálogos, y despliegue.*

Link a Diagrama:

https://drive.google.com/file/d/1T\_aUx00vMBRN96v-G3THXBLqWt5T0tRt/view?usp=sharing