Diseño del DAaaS

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

Servicio de envíos periódicos de informes por email, con insights valiosos, para un club de karate, que permita la toma de decisiones sobre sus deportistas.

LOS OBJETIVOS DE LA PLATAFORMA:

\* Facilitar la generación de informes a través de Power BI.

\* Establecer un entorno para la gestión y para el análisis de los datos de forma centralizada

\* Mejorar la toma de decisiones de carácter deportivo y empresarial.

Para ello hay que determinar que personas dentro del club pueden acceder a cada tipo de información. Asegurar un acceso fácil y seguro a los informes relevantes para cada usuario tipo: Presidente del club, profesores y deportistas.

La solución debe ser escalable para adaptarse al crecimiento de datos y usuarios.

**INPUT DE LOS DATOS:**

\* Base de datos SQL del club con los datos de los alumnos.

\* Datos de ranking nacional obtenidos por un crawler en la página Karate scoring

\* Resultados provinciales y regionales de la página del Federación Castellano Leonesa de Karate.

**CATALOGO DE SERVICIOS:**

\* Definir procesos de carga desde las diferentes fuentes de los datos disponibles necesarios.

\* Implementar flujos de trabajo para la limpieza y transformación de los datos, que garanticen la calidad y coherencia de los mismos para su posterior análisis.

\* Datapedias: Desarrollar los recursos documentales necesarios para la comprensión de los datos.

\* Identificar y ofrecer, si las necesidades del club lo requieren, otros servicios adicionales como análisis predictivos.

### **Arquitectura DAaaS**

\* Una Google Function para ejecutar API Gmail cuando se actualice el report, en google Storge, que envíe un email a los diferentes clientes objetivo.

\* Uso de Power BI y otras herramientas, si fuera necesario, para el análisis de los datos y la generación de visualizaciones, reportes, etc.…

\* Uso de un Dataproc de Google Cloud donde realizar los trabajos de ETL, a través de jobs programados, para cargar, limpiar, transformar y generar un CSV para su posterior uso en las visualizaciones de PowerBI.

\* Google Function que se ejecute cuando se actualicen los datos de entrada en Google Storage para levantar el Dataproc.

\* Uso de la herramienta Google Storage para el almacenamiento de los Input, Output y funciones auxiliares que pueden usarse en el proceso.

\* Google Function para actualizar el CSV de la base de dato del club de forma periódica y programada

\* Google Function que levante una MV para actualizar el CSV generado por un crawler que extraiga datos de la página <https://karatescoring.com/> útiles para su posterior análisis.

### **DAaaS Operating Model Design and Rollout**

\* Una persona cargará en Google Storage los archivos con resultados de competiciones en formato .xls que se envían a través de las federaciones autonómicas y de la federación nacional.

\* Una vez al mes se ejecutará el Google Function que levantará una MV en Compute Engine para ejecutar un crawler que actualizará el archivo CSV de los rankings.

\* Una vez al mes se ejecutará el Google Function para conectarse a la base de datos SQL del club y actualizar los datos relevantes del CSV.

\* Una vez actualizados estos archivos un Google Function programado levantará un Dataproc con los jobs necesarios para procesar todos esos datos en un CSV que se guardará en Google Storage.

\* Un analista de datos, a través de PowerBI, que se actualizará de forma programada, procederá a diseñar o modificar las visualizaciones y reportes que ofrezcan los insights adecuados a cada usuario objetivo.

\* El o los reports los cargará el analista en Google Storage.

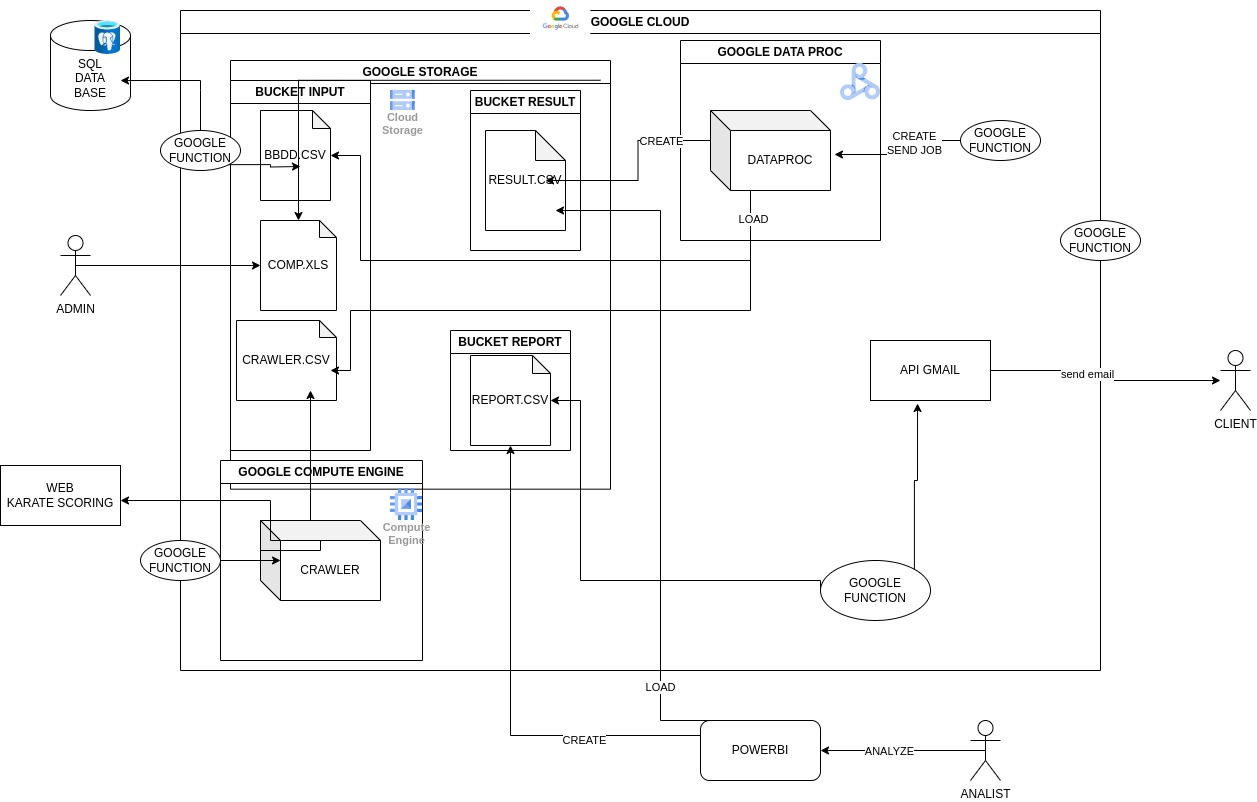
\* Una vez cargado en Google Storage se ejecutará una Google Function para conectar el API Gmail y enviar los emails necesarios a los clientes finales.

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

Crawler para descargar los datos de ranking nacional senior maculino y femenino de kata y kumite.

<https://colab.research.google.com/drive/12_KeV4YOnhDpJBXpnsVL-aA-GqVkdepJ?usp=sharing>

**DIAGRAMA**

****