Brainstorm

* Newsletter que te diga las frass del día de los jugadores/entrenadores de tu equipo favorito.
* Te suscribes en una web con tu mail y el equipo al que quieres seguir.
* Recibes semanalmente un email con las frases el día de tus jugadores.

Diseño del DAaaS

### Definición la estrategia del DAaaS

*Definir el catálogo de servicios que proporcionará la plataforma DAaaS, que incluye incorporación de datos, limpieza de datos, transformación de datos, datapedias, bibliotecas de herramientas analíticas y otros.*

De la web sacamos los mailers al que enviar semanalmente un mail. De la API de football-data org sacamos los nombres de los futbolistas del equipo favorito seleccionado por cada Mailer y vamos a Marca.com donde sacamos los titulares de noticia del día para cada equipo favorito. Luego matcheamos los nombres de jugadores con los titulares y si aparece texto entre comilla añadimos ese texto junto con el nombre del jugador al texto que vamos a enviar a cada usuario inscrito.

### Arquitectura DAaaS

*Definir la selección de componentes, la definición de procesos de ingeniería y el diseño de interfaces de usuario. Diseño y ejecución de Proofs-of-Concept (PoC) para demostrar la viabilidad del enfoque DAaaS.*

Alojada en Google Cloud Platform:

-Creamos bucket para crear un CNAME y llamar a nuestro dominio previamente creado través de gsutil.

-Con Gsutil, subimos al bucket el html para la web con sólo un form que pida dos input: mail y equipo favorito de una lista de equipos de la liga española.

-Creamos otro bucket para el procesamiento de archivos

-Cada vez que se rellene correctamente el form con un mail válido se añadirá una fila a un archivo csv que se autorellenará con los nuevos usuarios y sus equipos.

-Función 1: Creo una cloud function que le des un nombre de equipo de la liga española y te devuelva la lista de todos los jugadores de ese equipo a través de la api de football-data.org y me cree un csv con todos los jugadores de cada equipo.

- Función 2: Creo otra cloud function que haga Scrapping a Marca.com para sacar todos los titulares de todos los equipos de la liga española que sean nuevos de semana comparando con los existentes de la semana anterior.

- Función 3: Creo una cloud function que me coja de las noticias por cada equipo, las que incluyan “ y cojo el archivo o output de la función 2 para junto con el output de la función 1 sacar por jugador las frases que hayan salido suyas esa semana: por ejemplo:



Como vemos, Marca para las frases tiene siempre el mismo formato y se trata de hacer coincidir un nombre de algún jugador o entrenador de ese equipo y además encontrarte texto entre dos corchetes. Ese texto sería una frase nueva de ese jugador/entrenador esa semana y se guardaría en un nuevo dataset a través de Hive que genere las frases nuevas por equipo por jugador.

-Función 4 :Con Map reduce (tarea Dataproc) acabo con un dataset final que tenga todos los usuarios inscritos, junto con sus jugadores de sus equipos favoritos y sus nuevas frases de esa semana.

-Creo una instancia de máquina virtual para usar Sendgrid y así configurarme un mail que cree un texto por cada cursor del dataset final y mande un mail a cada usuario con las nuevas frases de los jugadores del equipo favorito.

### DAaaS Operating Model Design and Rollout

*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.*

-El bucket con la web está siempre levantado.

-Cada lunes se arranca el proceso por la mañana. El bucket de los archivos y las funciones cloud.

-Vía Cloud scheduler llamo a la función 1, 2 y 3 para tener listos todos los jugadores y todas las nuevas frases de la semana.

-Se dispara luego la cloud function 4 que inicializa Hive y crea el dataset final entre los dos buckets que utilizará Sendgrid.

-A las 9:00 AM con cloud scheduler se llama a la instancia de Sendgrid para que mande el email a todos los usuarios con sus frases

### Desarrollo de la plataforma DAaaS.

*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.*

Utilizaríamos Scrappy para scrapear, Python en las funciones de cloud .

Output de las funciones:

Función 1.head(3):

Equipo Jugadores (separados por coma)

1 SevillaFC Papu Gómez, Rakitic…..

2 Real Madrid Odegaard,Vinicius JR…..

3 FC Barcelona Messi, Piqué…..

Función 2.head(3):

Equipo Noticias (separados por punto y coma)

1 SevillaFC Lopetegui: “El Sevilla es el mejor equipo”; El Sevilla gana…..

2 Real Madrid Zidane en duda; Vinicius JR: “De momento me quedo…..

3 FC Barcelona Messi es sorprendido en negociaciones…..

Función 3.head(2):

Equipo Jugador Frase

1 SevillaFC Lopetegui El Sevilla es el mejor equipo

2 Real Madrid Vinicius JR De momento me quedo en el Real Madrid

Función 4.head(2):

Equipo Jugador Frase Usuario

1 SevillaFC Lopetegui El Sevilla es… luissevillista@gmail.com

2 Real Madrid Vinicius JR De momento… luismadridista@gmail.com

Mail a enviar:

Hola {{usuario}},

Aquí tienes las nuevas frases de esta semana de los jugadores de tu equipo favorito, el {{Equipo}}:

{{Jugador}} : “{{Frase}}”

{{Jugador}} : “{{Frase}}”

{{Jugador}} : “{{Frase}}”

Hasta la semana que viene!

Link a Diagrama: