

Base de Datos Multidimensional



Proyecto 1er Bimestre

España Bryan, Lizeth Toasa
ESCUELA POLITÉCNICA NACIONAL
Quito-Ecuador

bryan.espana@epn.edu.ec, grimaneza.toasa@epnl.edu.ec

RESUMEN: Realizar una recolección de datos de animales de varias partes para ingresar en couchDB para la visualización.

1 INTRODUCCIÓN

Realizar una arquitectura para el ingreso de varios tipos de datos relaciones y no relacionales para visualizar en un archivo HTML.

2 Fundamento teórico

En la actualidad el uso de las redes sociales como: Twitter, LinkedIn, Instagram, Facebook, ha permitido que los individuos puedan compartir interés o actividades. El presente proyecto utilizó la red social Twitter, para poder recoger tweets con el #Mysteryland, contando al momento con 224379 de tweets.

Twitter puede ser utilizado para recopilar datos y analizar los mismos, debido a que su metadata tiene información como: keywords, hora, fecha, geotagging, hashtags, seguidores, etc.

Existen actualmente miles de herramientas para Twitter, algunas aportan un soplo de aire fresco en el desierto de Twitter, ya que te permiten gestionar tu comunidad con tu imaginación como único límite Twitter nos ofrece su API (Interfaz de Programación de Aplicaciones, Application Programming Interface). Entonces, el API de Twitter conecta a nuestras aplicaciones o sitios web con el tráfico de Twitter a través de varios servicios. El festival de Mysteryland, uno de los festivales de música electrónica más conocidos mundialmente, ha anunciado que retransmitirá en live stream de realidad virtual, por lo que se convierte en el primer festival del mundo en realizar este tipo de retransmisión.

3 Objetivo General

- Diseñar un Data warehouse en el cual confluyan varias fuentes de información estructurada, no estructurada y semi estructurada. Esta parte del

proyecto pertenece a las etapa de “preparar y explorar datos”.

• **Objetivos específicos**

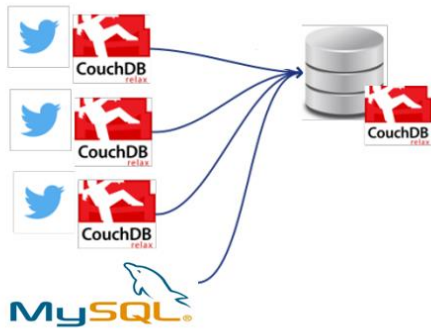
- Recopilar tweets acerca del animales con los subtemas de perros, gatos y cerdos.
- Importar datos de mysql a couchdb.
- Exportar datos de un archivo cvs a couchdb a través de python.
- Preparar y exportar datos a couchdb.
- Visualizar los datos a través de una interfaz en HTML.

4 Recursos y herramientas utilizadas

- XAMPP
- Python 3.5
- Notepad++
- CouchDb
- MySql

5 Arquitectura de la solución

El tema principal es Animales de los cuales existe 3 bases de datos que se accederá remotamente a couchdb principal. Las 3 bases son los subtemas que son perros, gatos, cerdos. Mientras que la base MySQL se exporta a un archivo csv y se la ingresa al couchdb principal.



6 Extracción de datos

La extracción de datos está destinada a mostrar cualquier correlación dentro del volumen de datos del sistema de información con el fin de detectar la tendencia de los datos.

Primero creamos una [API](#) de twitter que nos asignará un conjunto de credenciales necesarios para la extracción de la información.

Application Management

Mysteryland

Details Settings **Keys and Access Tokens** Permissions

Application Settings

Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

Consumer Key (API Key)	w721sk5kmzyH7ZZf8UfPJUDI
Consumer Secret (API Secret)	1UL5sNwITAXoSyKZh9izpAL19ALJUaPRWbKtETrBh8Q2jYWoEd
Access Level	Read and write (modify app permissions)
Owner	KerlyTonSita
Owner ID	872556070471913473

Crear los archivos tweets.py donde se definen las palabras o la geo localización de donde se tomará los tweets.

Nombre	Fecha de modifica...	Tipo	Tamaño
cosech_cerdo	27/11/2018 3:09	Python File	2 KB
cosech_gatos	27/11/2018 2:38	Python File	2 KB
cosech_perros	27/11/2018 2:39	Python File	2 KB

Correr los archivos y verificar que en couchDB.

```

C:\Windows\System32\cmd.exe
Microsoft Windows [versión 10.0.17134.487]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.
C:\Users\Bryan\Downloads\Proyecto_Final\Tweets\Males\python cosech_cerdo.py

C:\Windows\System32\cmd.exe
Microsoft Windows [versión 10.0.17134.487]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.
C:\Users\Bryan\Downloads\Proyecto_Final\Tweets\Males\python cosech_gatos.py

C:\Windows\System32\cmd.exe
Microsoft Windows [versión 10.0.17134.487]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.
C:\Users\Bryan\Downloads\Proyecto_Final\Tweets\Males\python cosech_perros.py
  
```

7 Exportar e importat datos de MySQL.

Seleccionar la base de datos que se requiera trasnformar la informacion a csv.

phpMyAdmin

La selección actual no contiene una columna única. La edición de la grilla y los enlaces de copiado, eliminación y edición no es posible.

Mostrando filas 0 - 24 (total de 50, La consulta tardó 0,000 segundos)

SELECT * FROM "catsvogs"

COL 1	COL 2	COL 3	COL 4	COL 5	COL 6	COL 7	COL 8	COL 9	COL 10	COL 11	COL 12
Alabama	1.020	59.5	1.000	44.1	907	1.7	1.410	27.1	601	2.5	1.262
Arizona	2.515	59.5	1.497	40.1	1.008	1.8	1.790	29.6	743	1.9	1.430
Arkansas	11.68	62.4	716	47.9	550	2.0	1.097	30.6	351	2.3	810
California	12.974	50.5	6.865	32.8	4.200	1.6	6.107	26.3	1.607	1.9	7.118

Exportar el archivo csv.

phpMyAdmin

Exportando filas de la tabla "catsvogs"

Exportar plantillas:

Nueva plantilla: Crear

Plantillas existentes: Actualizar Borrar

Método de exportación:

☒ Rápido: mostrar sólo el mínimo de opciones de configuración

☐ Personalizado: mostrar todas las opciones de configuración posibles

Formato:

CSV

Continuar

Ejecutar el script Python para ingresar datos a couchdb

Importar

Nombre	Fecha de modifica...	Tipo	Tamaño
importarcsv	27/11/2018 1:11	Carpeta de archivos	
importar_csv_mysql	27/11/2018 3:48	Python File	1 KB

```

C:\Windows\System32\cmd.exe
Microsoft Windows [versión 10.0.17134.487]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.
C:\Users\Bryan\Downloads\Proyecto_Final\Importar\python importar_csv_mysql.py
  
```

9 Resultados

- Index.html



• Perros



• Gatos



• Cerdos



• Estadística



10 Conclusiones

- Para el ingreso de datos a la base couchDB se debe transformar los archivos.
- Se puede extraer información de cualquier base de datos para transformarla e ingresarla a una nueva base de datos.
- Para la visualización se puede realizar de plantillas creadas en HTML.

11 Recomendaciones

Utilizar Python 3.5 para el ingreso de librería couchDB y Tweepy.

12 REFERENCIA

[1]"Introducción — Conociendo GitHub 0.1 documentation", Conociendogithub.readthedocs.io, 2017. [Online]. Available: <http://conociendogithub.readthedocs.io/en/latest/data/introduccion/>.

[2]"¿Qué es CouchDB? - Definición en WhatIs.com", SearchDataCenter en Español, 2017. [Online]. Available: <http://searchdatacenter.techtarget.com/es/definicion/CouchDB>.

[3]d. Que es Twitter and d. Que es Twitter, "¿Qué es Twitter y cómo funciona?", IIEMD, 2017. [Online]. Available: <https://iiemd.com/twitter/que-es-twitterfunciona>.

