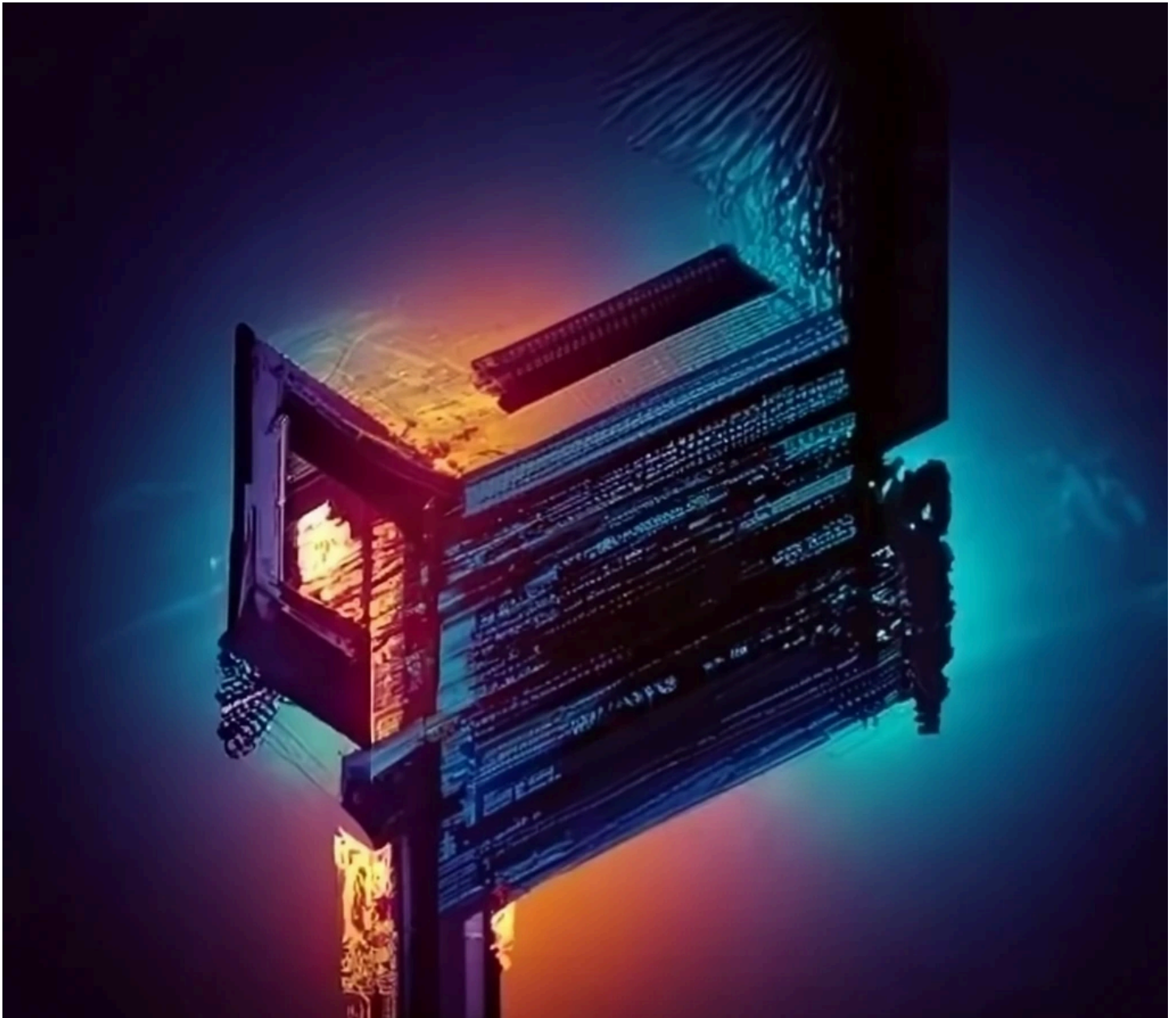




**Universidad de Guadalajara**  
**Centro Universitario de Ciencias Exactas e Ingenierías**



Sección: D06      Profesor: MICHEL EMANUEL LOPEZ FRANCO

Tema: Airflow is a platform to programmatically author, schedule and monitor workflows

Ciclo: 2024A

## Introducción

Apache Airflow es una plataforma de orquestación de flujo de trabajo open-source desarrollada por Airbnb y actualmente mantenida por la Apache Software Foundation. Permite a los usuarios definir, programar y monitorear flujos de trabajo (también conocidos como "pipelines") de manera programática.

En términos simples, Airflow ayuda a automatizar tareas complejas y repetitivas, como el procesamiento de datos, la ingesta de datos, la generación de informes, entre otros. Utiliza un concepto llamado "DAGs" (Grafos Acíclicos Dirigidos) para definir la lógica de los flujos de trabajo, lo que permite a los usuarios representar visualmente la secuencia de tareas y las dependencias entre ellas.

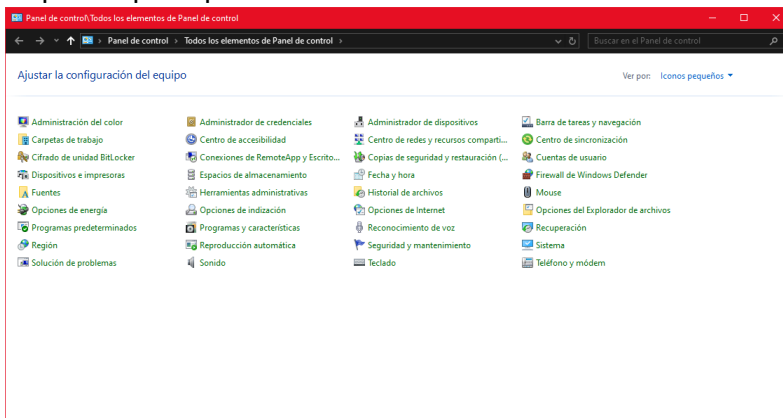
En resumen, Airflow es una herramienta poderosa para la gestión y automatización de flujos de trabajo de datos y procesamiento de tareas en entornos empresariales.



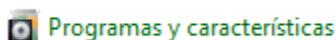
## Contenido

Desde la perspectiva de la instalación como un usuario se deben encontrar maneras poco o nualmente conflictivas con este elemento ya que apache airflow es un poco complejo y se resiste a ser instalado como si de un programa cualquiera se tratase, por lo cual se encuentra una manera diferente de instalar sin tener que depender de tantos programas aparte como lo es la mención de DOCKER, que a resumidas cuentas es un ambiente que crea pequeños campos con inmersión LINUX para que este complemento funcione de una manera adecuada, como se mostrará a continuación la experiencia que he tenido con la instalación de este elemento aparte:

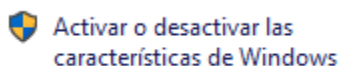
en primer paso para hacer la instalación en windows 10 entramos al panel de control:



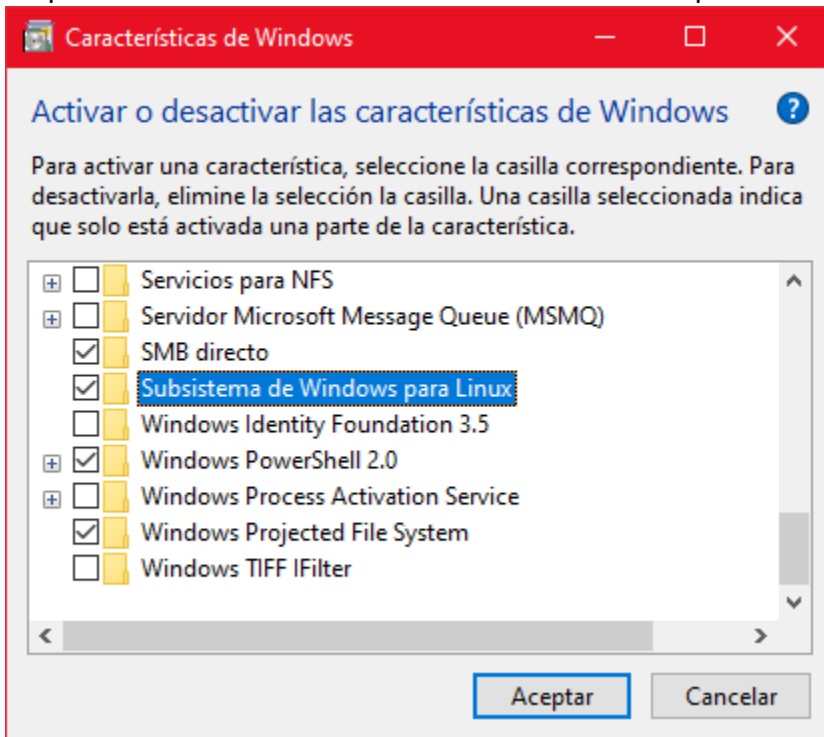
por siguiente elegimos la opción de programas y características



una vez dentro usamos la opción de activar o desactivar programas y características



despues en el menu de características usamos una opción donde aplica un subsistema de windows para linux



pero para no perder costumbres tambien he usado el comando de pip install en el entorno deseado que donde es aplicado pues es usando VS code, para asi tenerlo de cierta manera ya instalado.

tambien despues de aplicar los cambios deseados para el subsistema, instalamos desde la microsoft store ubuntu, mas concretamente la herramienta.



para mi sorpresa no me permitia avanzar ya que no concordaba el nombre de usuario que habia puesto asi como el siguiente ejemplo de lanzamiento de error:

Please enter a username matching the regular expression configured via the NAME\_REGEX[\_SYSTEM] configuration variable. Use the '--force-badname' option to relax this check or reconfigure NAME\_REGEX.

como consiguiente al no darme forma de usar ubuntu para dar inicio del ejemplo use visual studio pero: lo malo es que de igual manera no me dejo hacer uso con las siguientes imagenes de codigo.

```
AirFlow-Apache.py > ...
1  from datetime import datetime, timedelta
2  from airflow import DAG      No se ha podido resolver la importación "airflow".
3  from airflow.operators.python_operator import PythonOperator      No se ha podido resolver la importación "air
4
5  def print_hello():
6      return 'Hello Airflow!'
7
8  def sleep():
9      return 'Sleeping for 5 seconds...'
10
11  default_args = {
12      'owner': 'airflow',
13      'depends_on_past': False,
14      'start_date': datetime(2024, 4, 13),
15      'retries': 1,
16      'retry_delay': timedelta(minutes=5),
17  }
```

```
PS C:\Users\USER\Downloads\Micro Analizador Lexico> & C:/Python312/python.exe "c:/Users/USER/Downloads
Traceback (most recent call last):
  File "c:\Users\USER\Downloads\Micro Analizador Lexico\AirFlow-Apache.py", line 2, in <module>
    from airflow import DAG
ModuleNotFoundError: No module named 'airflow'
PS C:\Users\USER\Downloads\Micro Analizador Lexico> airflow dags unpause simple_workflow
>>
airflow : El término 'airflow' no se reconoce como nombre de un cmdlet, función, archivo de script o p
nombre o, si incluyó una ruta de acceso, compruebe que dicha ruta es correcta e inténtelo de nuevo.
En línea: 1 Carácter: 1
+ airflow dags unpause simple_workflow
+ ~~~~~
+ CategoryInfo          : ObjectNotFound: (airflow:String) [], CommandNotFoundException
+ FullyQualifiedErrorId : CommandNotFoundException

PS C:\Users\USER\Downloads\Micro Analizador Lexico> 
```

## Referencias

- Cardona, J. [@JorgeCardona1]. (2021, octubre 15). *Instalación, Configuración Apache Airflow - Windows sin Docker - Ubuntu, Creación y Ejecución DAGs*. Youtube.

<https://www.youtube.com/watch?v=YPIX-qT6d8M>

•