#### Propuesta Caso de Uso:

# Air Quality Monitoring using Web Scraping & AWS Services



Sistemas y Servicios en la Nube Máster Universitario en Ingeniería Informática

> Manuel Brazález Cañadas manuel.brazalez@alu.uclm.es Abel Ortiz Villaescusa abel.ortiz1@alu.uclm.es Álvaro Mayorga Zarco alvaro.mayorga@alu.uclm.es

> 23 de diciembre de 2024

### Índice general

1. Propuesta del Caso de Uso

3

## Índice de figuras

1.	Diagrama	de la	arquitectura	propuesta														4
----	----------	-------	--------------	-----------	--	--	--	--	--	--	--	--	--	--	--	--	--	---

#### 1. Propuesta del Caso de Uso

Nuestro objetivo en este proyecto será crear una arquitectura en AWS capaz de recolectar información relacionada con la calidad del aire a través del uso de web scraping. Para ello, utilizaremos la página web troposfera.es como fuente de datos. En base a los datos recopilados, podremos monitorizar la calidad del aire en la ciudad de Albacete con el objetivo de estudiar la calidad del aire en tiempo real y poder tomar decisiones en base a esta información.

En la figura 1, se muestra la arquitectura propuesta junto con los puntos claves de la misma que se detallan a continuación:

- Paso 1. Definiremos cada cuánto queremos que se ejecute la lambda que se encargará del webscraping usando un schedule event.
- Paso 2. Utilizaremos el servicio de AWS EventBridge con el schedule previamente configurado, para programar la ejecución de la función lambda que contiene el código para realizar el web scraping que se encargará de recopilar los datos de la página web mencionada anteriormente.
- Paso 3. Los datos recopilados se almacenarán en una base de datos NoSQL de AWS DynamoDB.
- Paso 4. Con la ayuda del servicio de AWS QuickSight podremos ofrecer a un usuario los datos almacenados en DynamoDB de una forma visual y sencilla.
- Paso 5. Al añadir un nuevo dato a la base de datos, se disparará una lambda encargada de enviar una notificación a un usuario a través del servicio de mensajería SNS.
- Paso 6. La función lambda que se encarga de enviar la notificación al tópico de SNS, para que los usuarios suscritos reciban un email indicando que se ha almacenado en la base de datos nueva información relacionada con la calidad del aire.

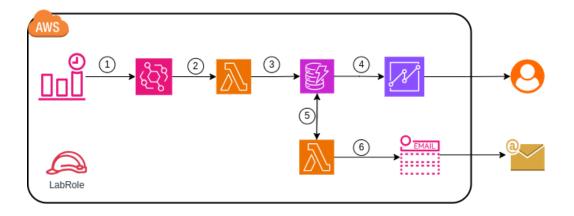


Figura 1: Diagrama de la arquitectura propuesta