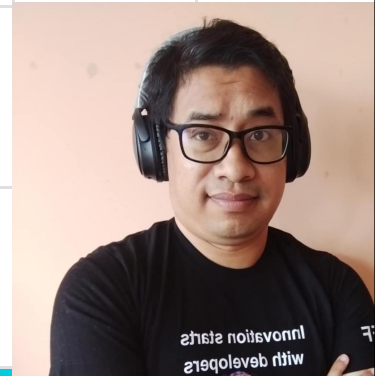




AWS User Group Arequipa



# Implementando patrones de calidad de datos en AWS

Guía técnica y teórica

21/01/2025

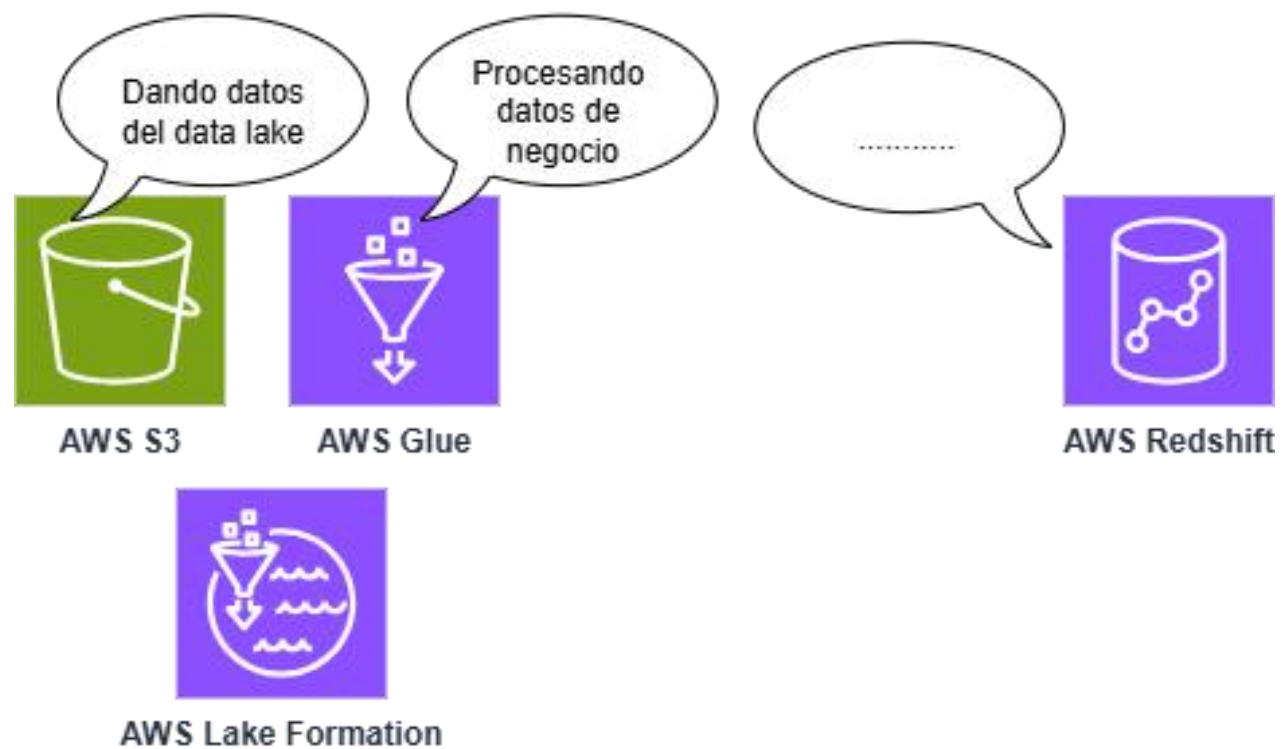
# Agenda

- Entendiendo patrones de calidad de datos.
- Ecosistema de AWS para data pipeline: conectando mundos.
- Consejos prácticos y ejemplos

Entendiendo patrones de calidad de datos.



# La realidad en la arquitectura de datos



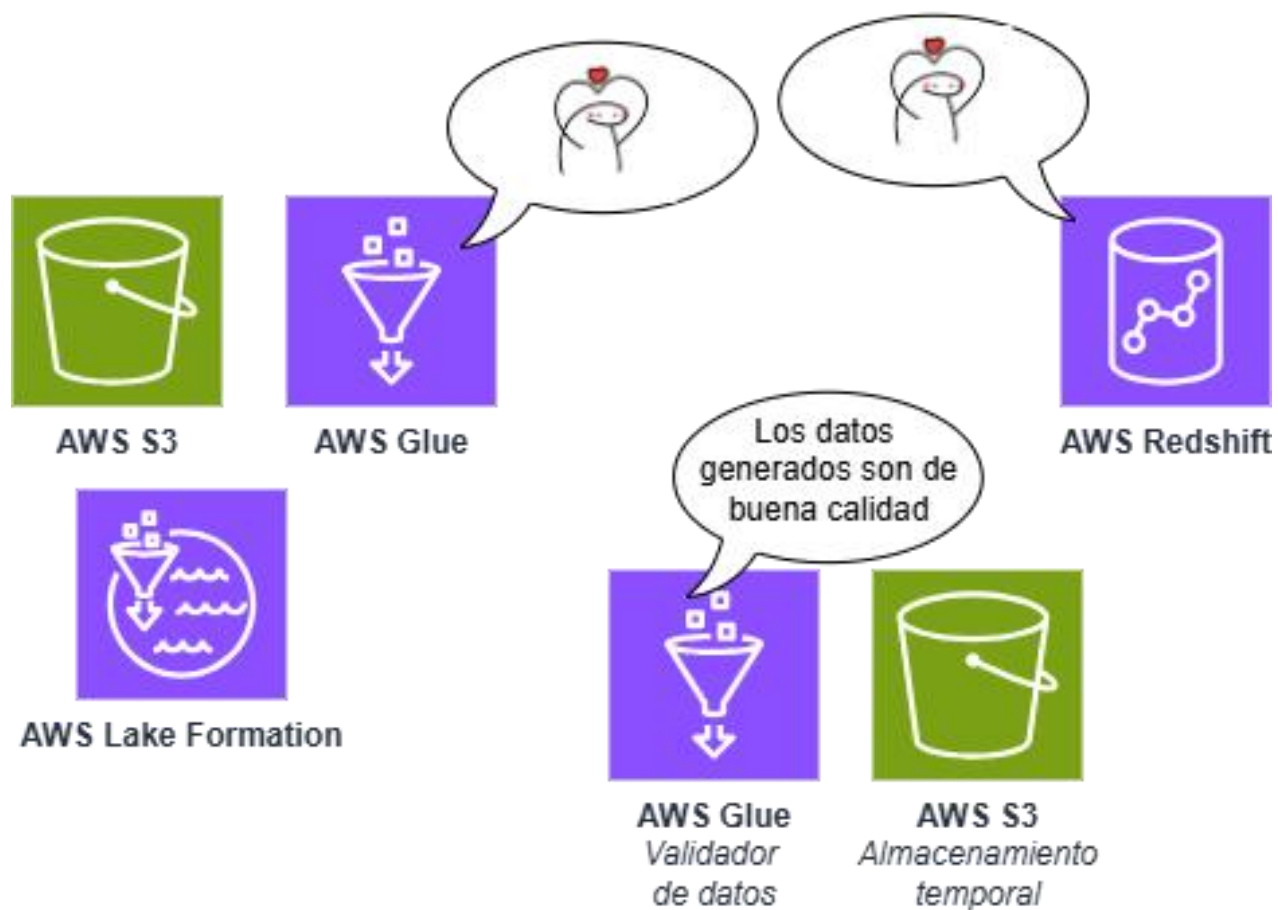
# La realidad en la arquitectura de datos



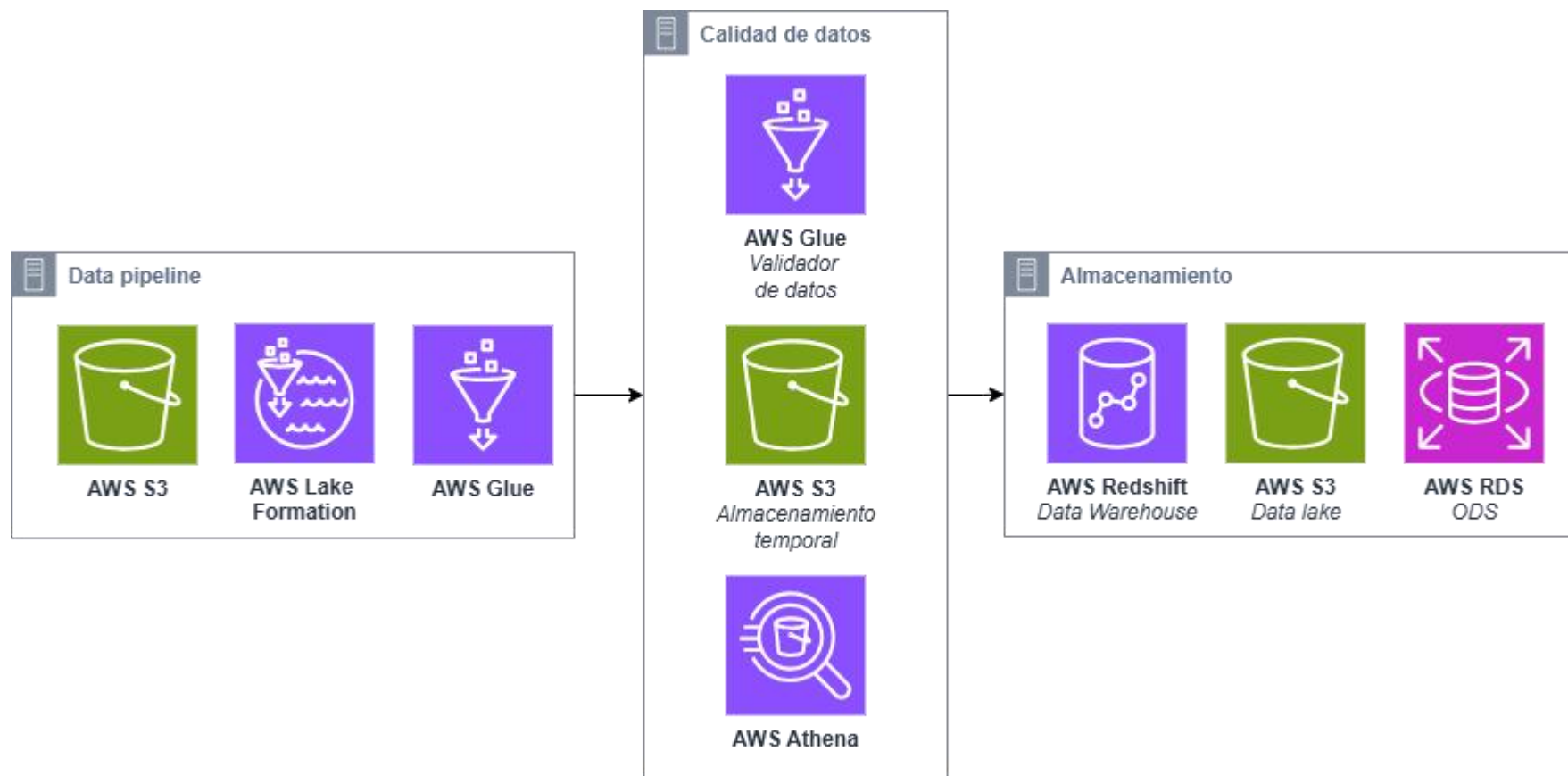
# La realidad en la arquitectura de datos



# La calidad de datos



# La calidad de datos: el patrón en el data pipeline

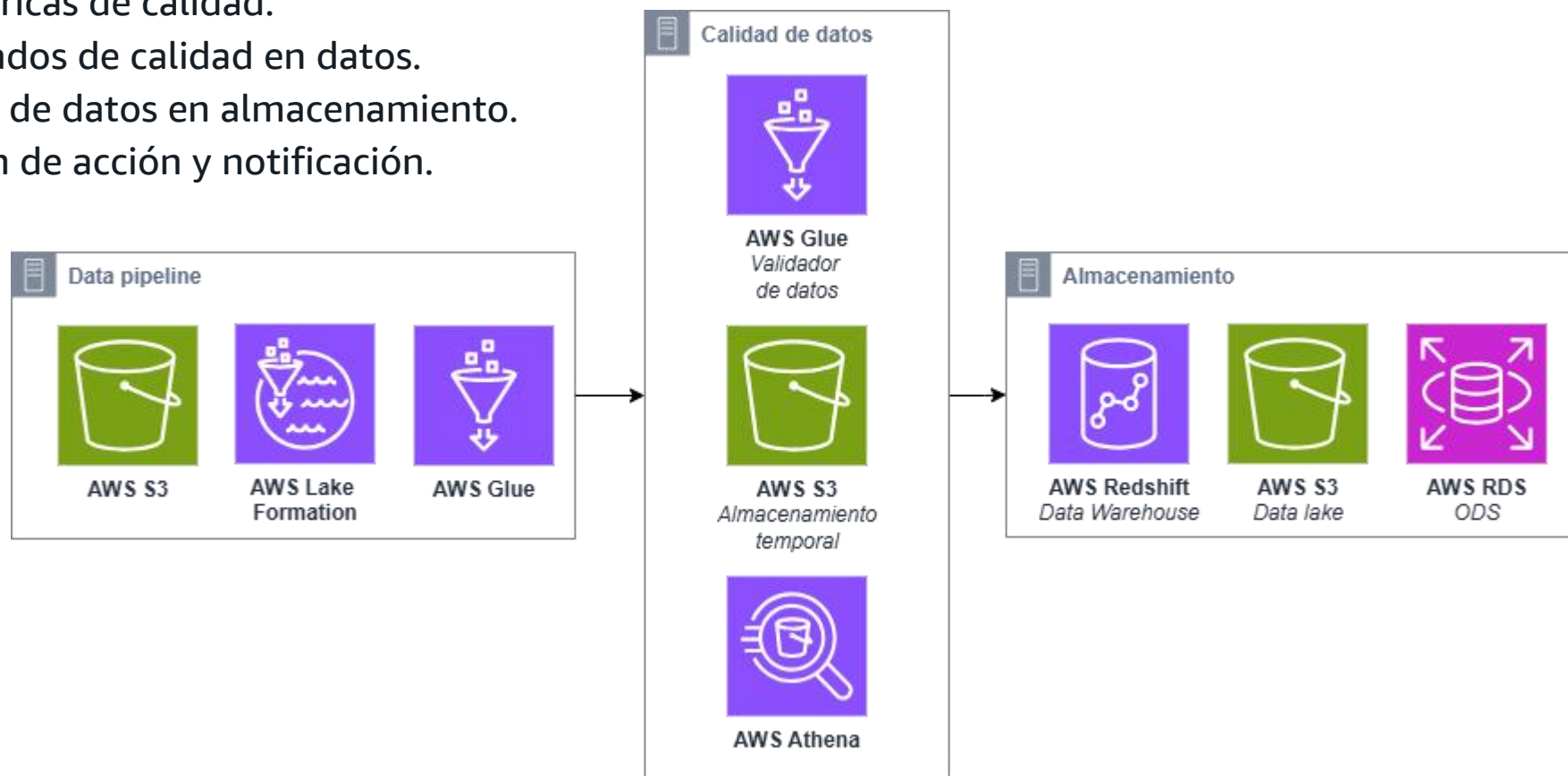




# La calidad de datos: el patrón en el data pipeline

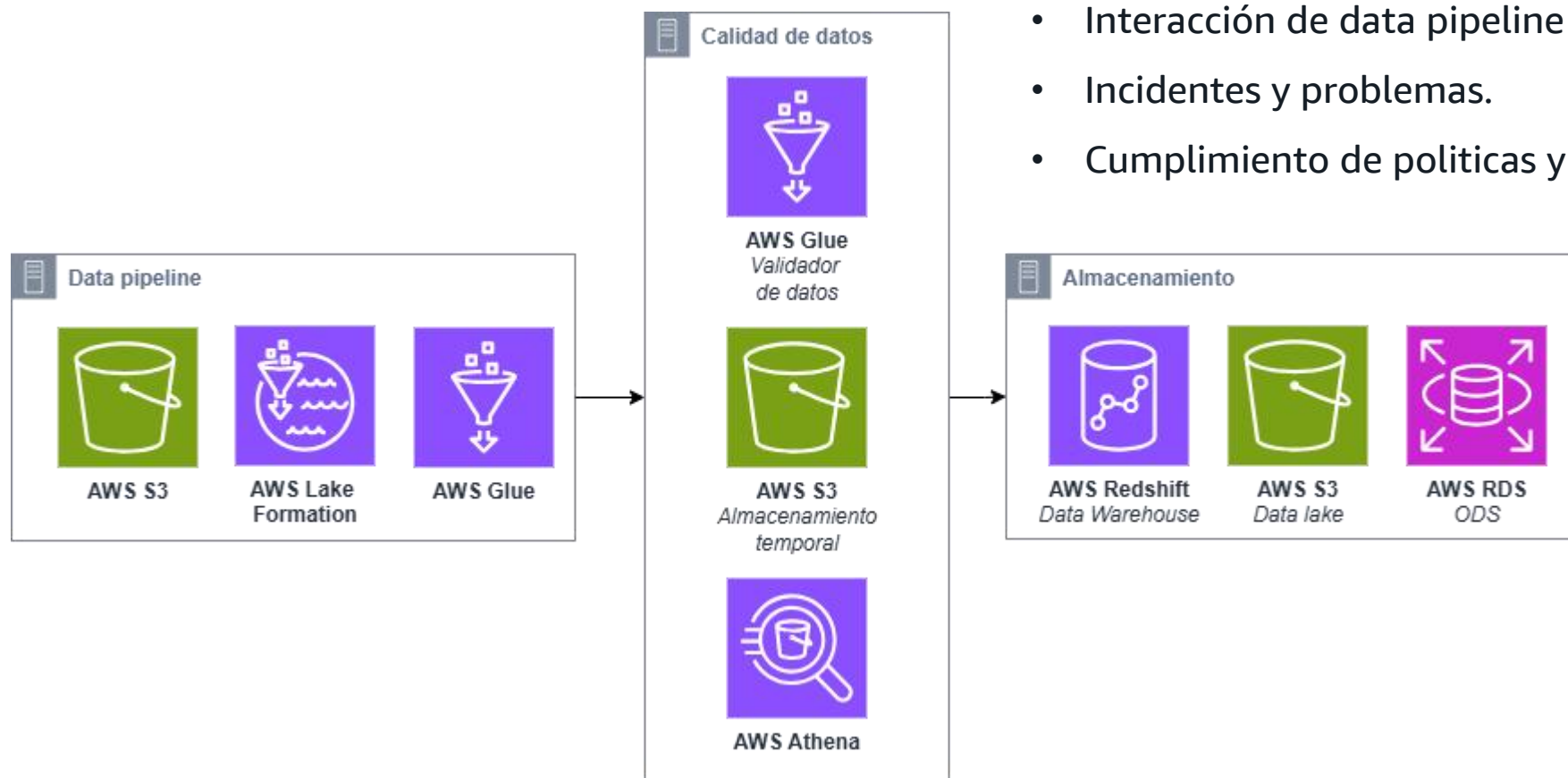
Una combinación de:

- Metricas de calidad.
- Estados de calidad en datos.
- Uso de datos en almacenamiento.
- Plan de acción y notificación.



Entendiendo patrones de calidad de datos.

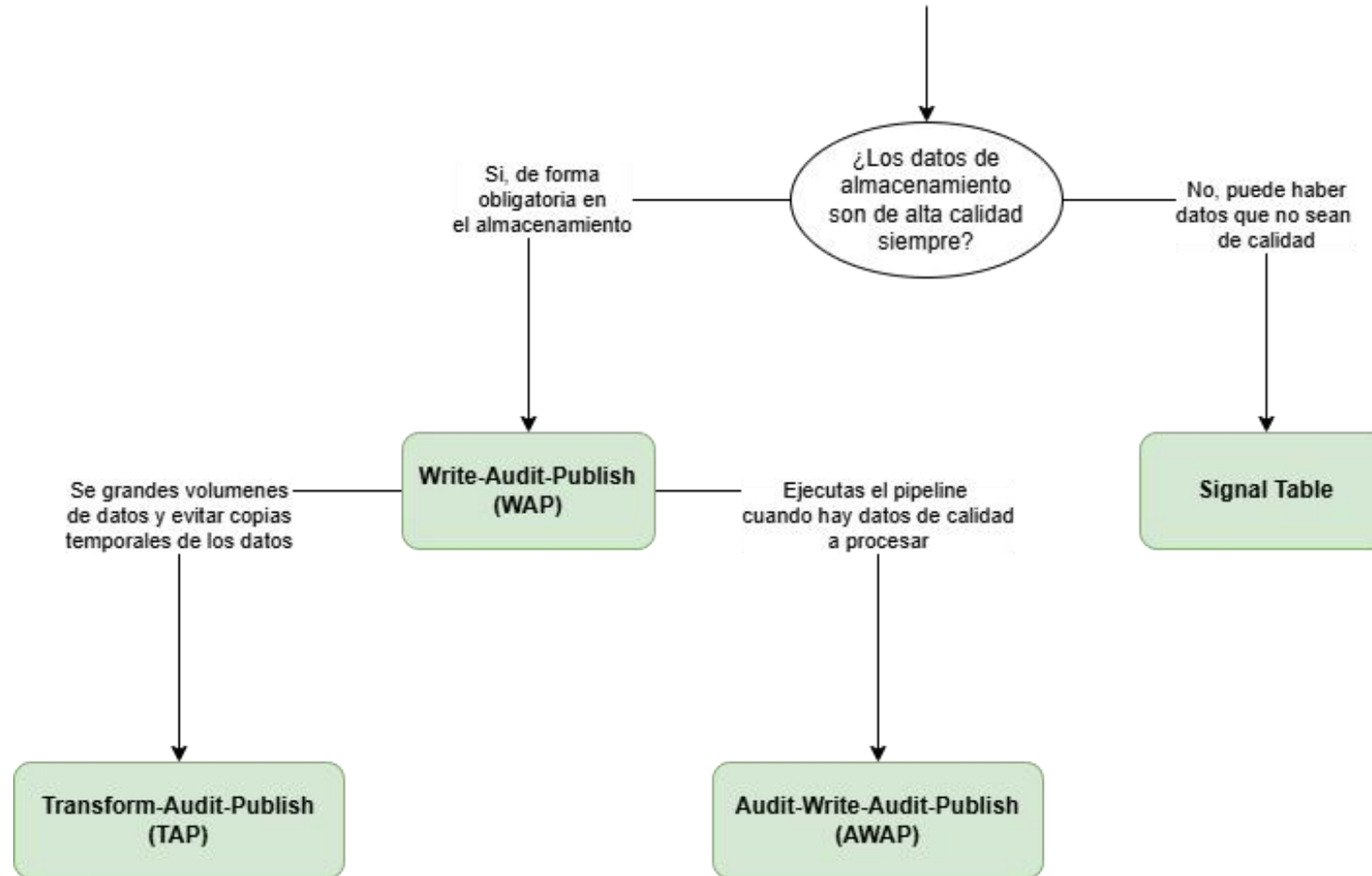
# La calidad de datos: el patrón en el data pipeline



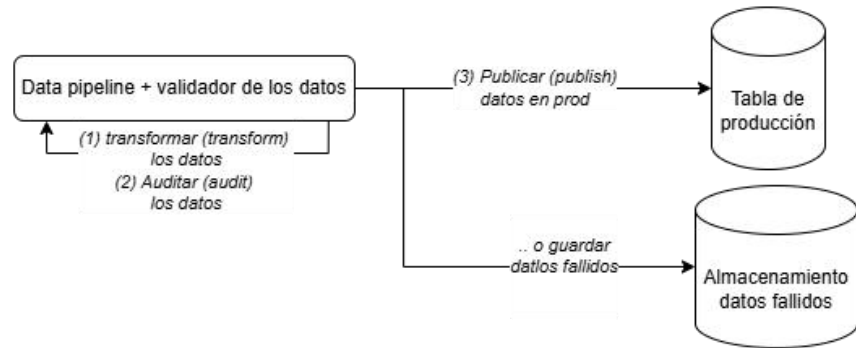
Esto afecta en varios temas:

- Interacción de data pipeline y almacenamiento.
- Incidentes y problemas.
- Cumplimiento de políticas y acuerdos con cliente.

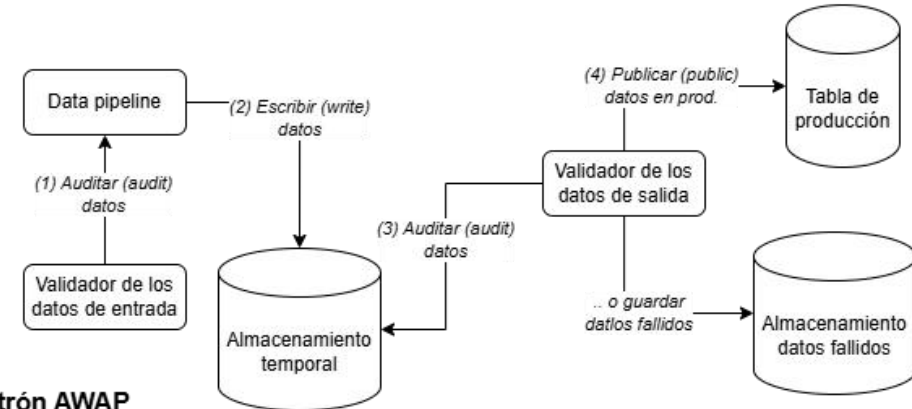
# Patrones de calidad de datos: mapa de elección



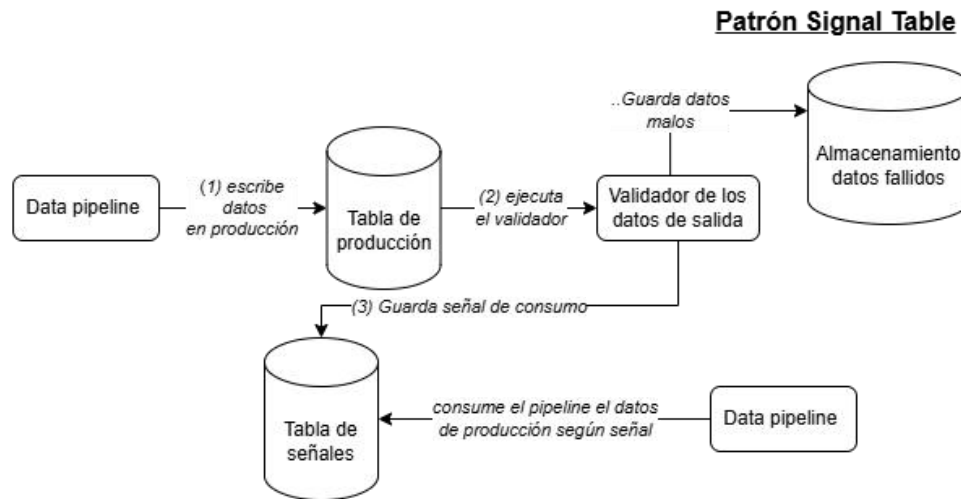
# Patrones de calidad de datos: visión técnica



**Patrón TAP**

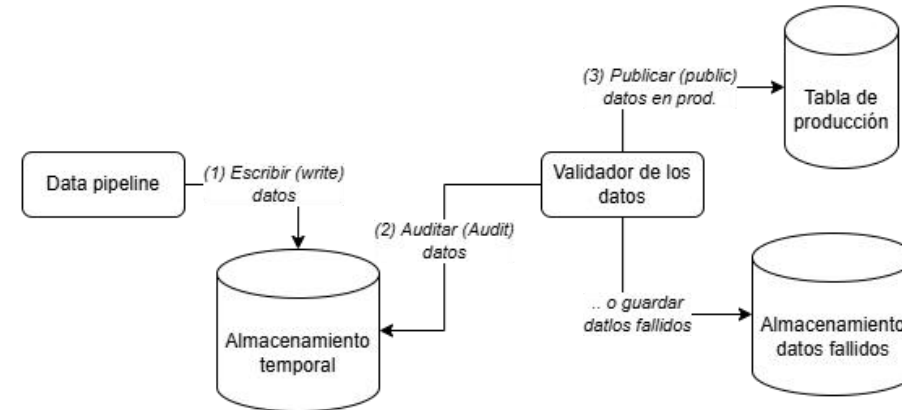


**Patrón AWAP**



**Patrón Signal Table**

**Patrón WAP**



Ecosistema de AWS para data  
pipeline: conectando mundos.



# Servicios AWS comunes en varios roles



**AWS EMR**



**AWS Redshift**



**AWS S3**



**AWS Glue**



**AWS Athena**



**AWS EventBridge**



**AWS lambda**



**AWS SQS**



**AWS Aurora**



**AWS Kinesis  
Data Streams**



**AWS MSK**

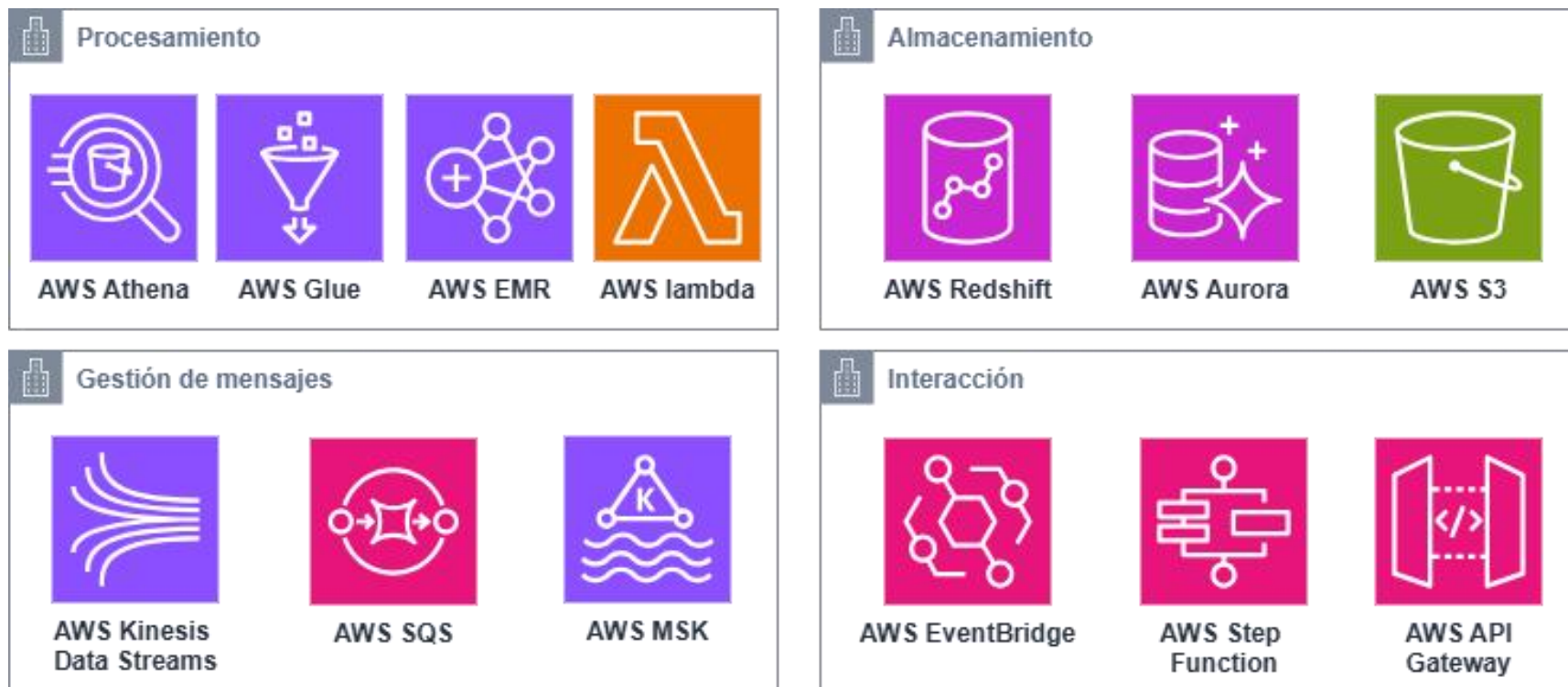


**AWS Step  
Function**

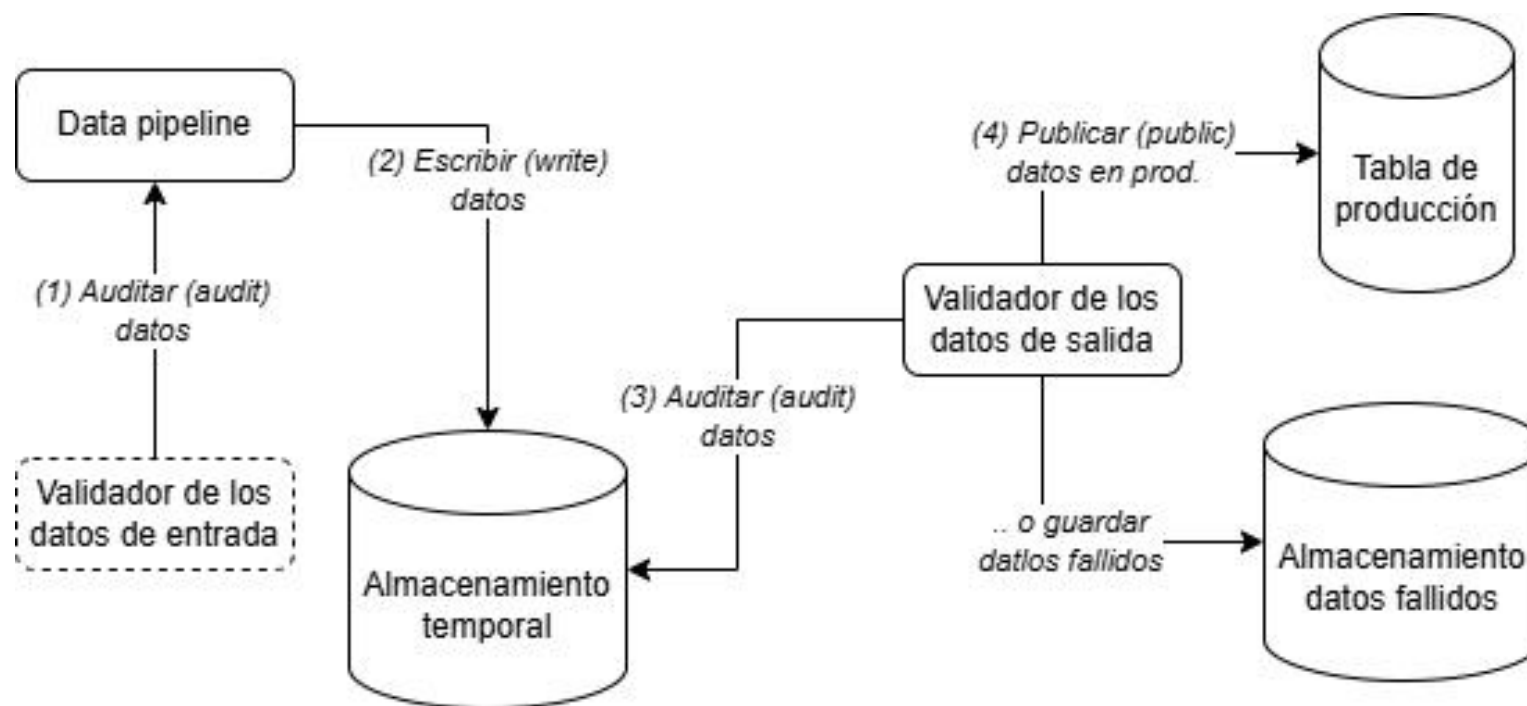


**AWS API  
Gateway**

# Una ayuda memotecnica como data engineer

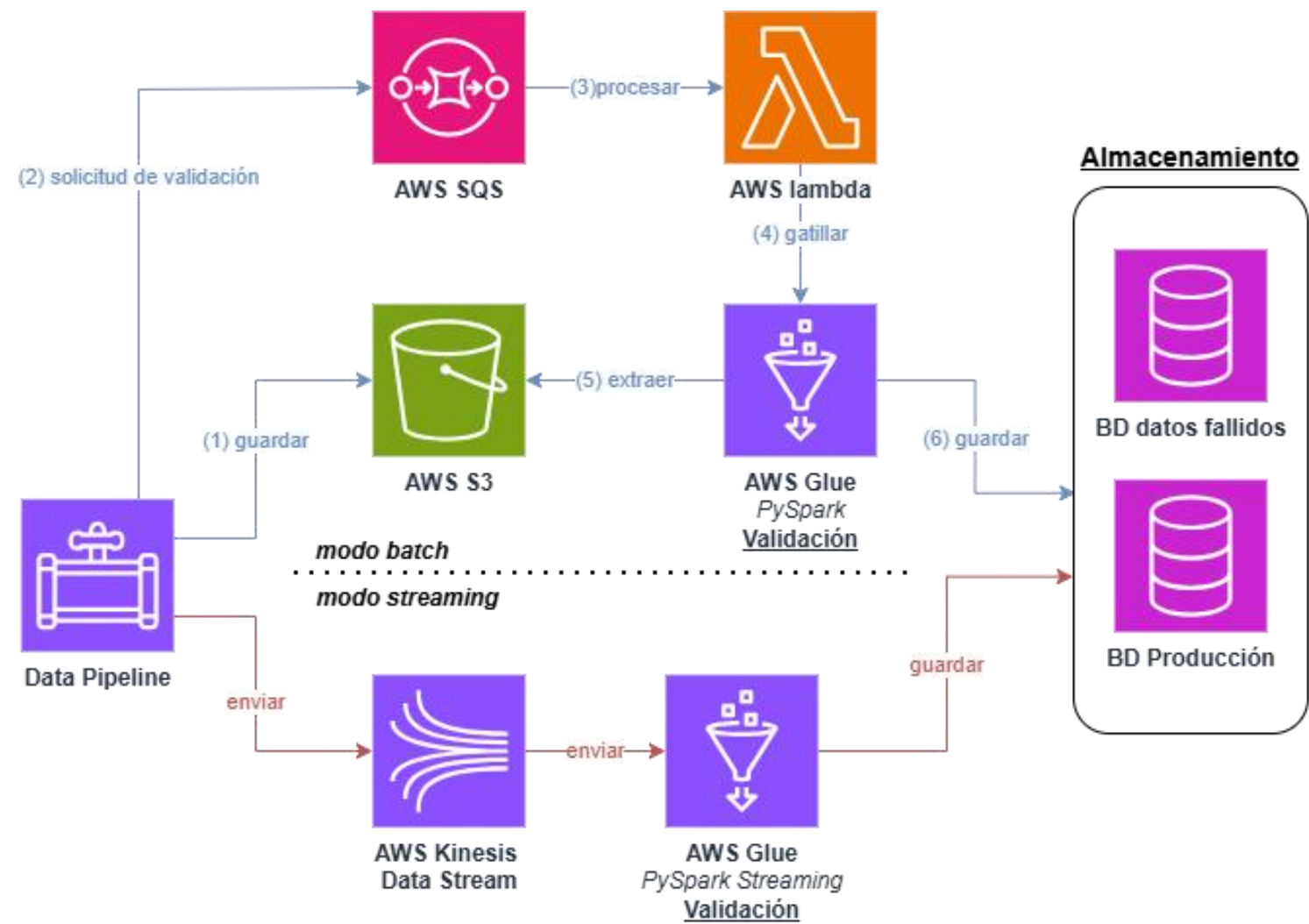


# Patrón tipo WAP para data pipeline





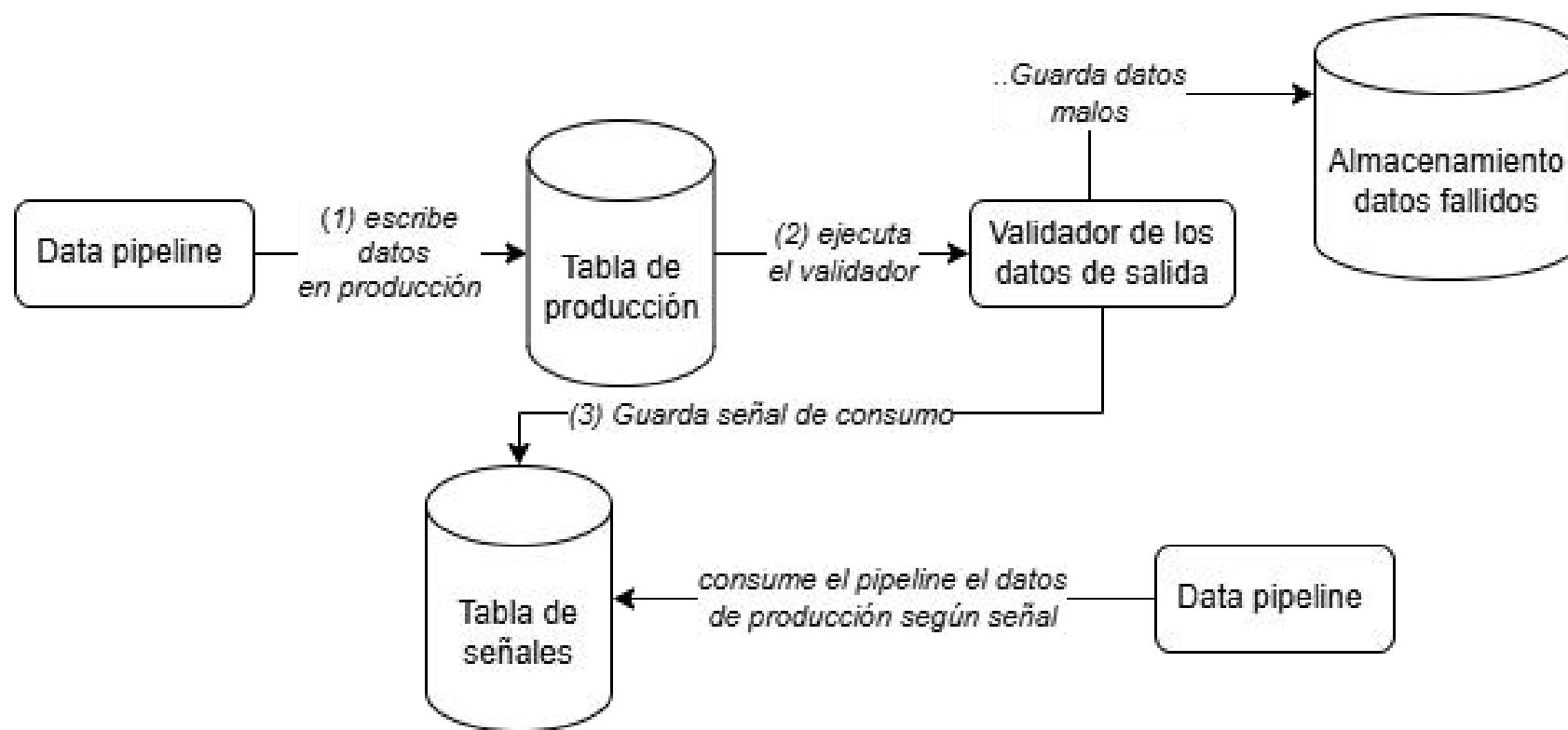
# Patrón tipo WAP para data pipeline



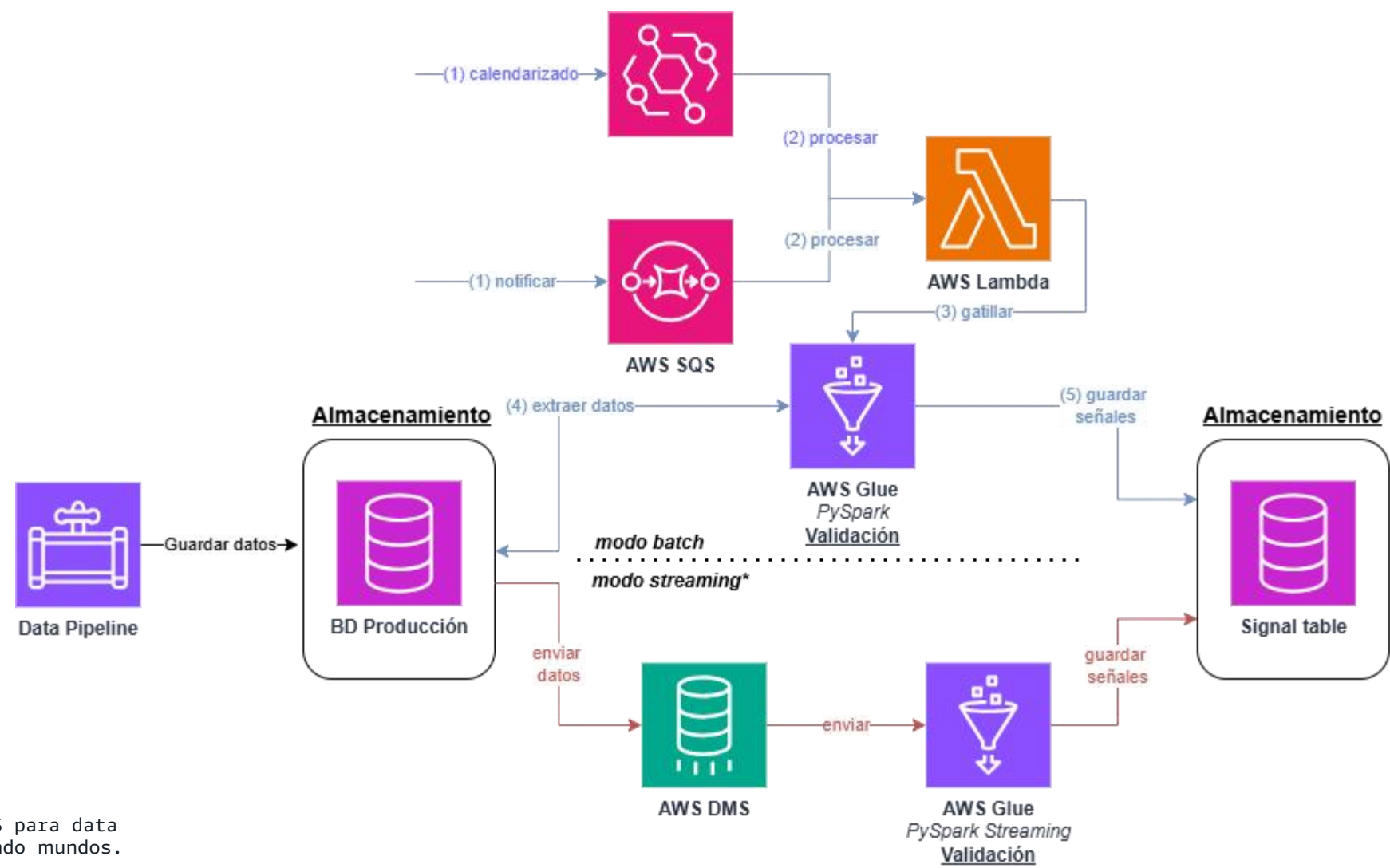
# Ejemplo del patrón WAP en streaming



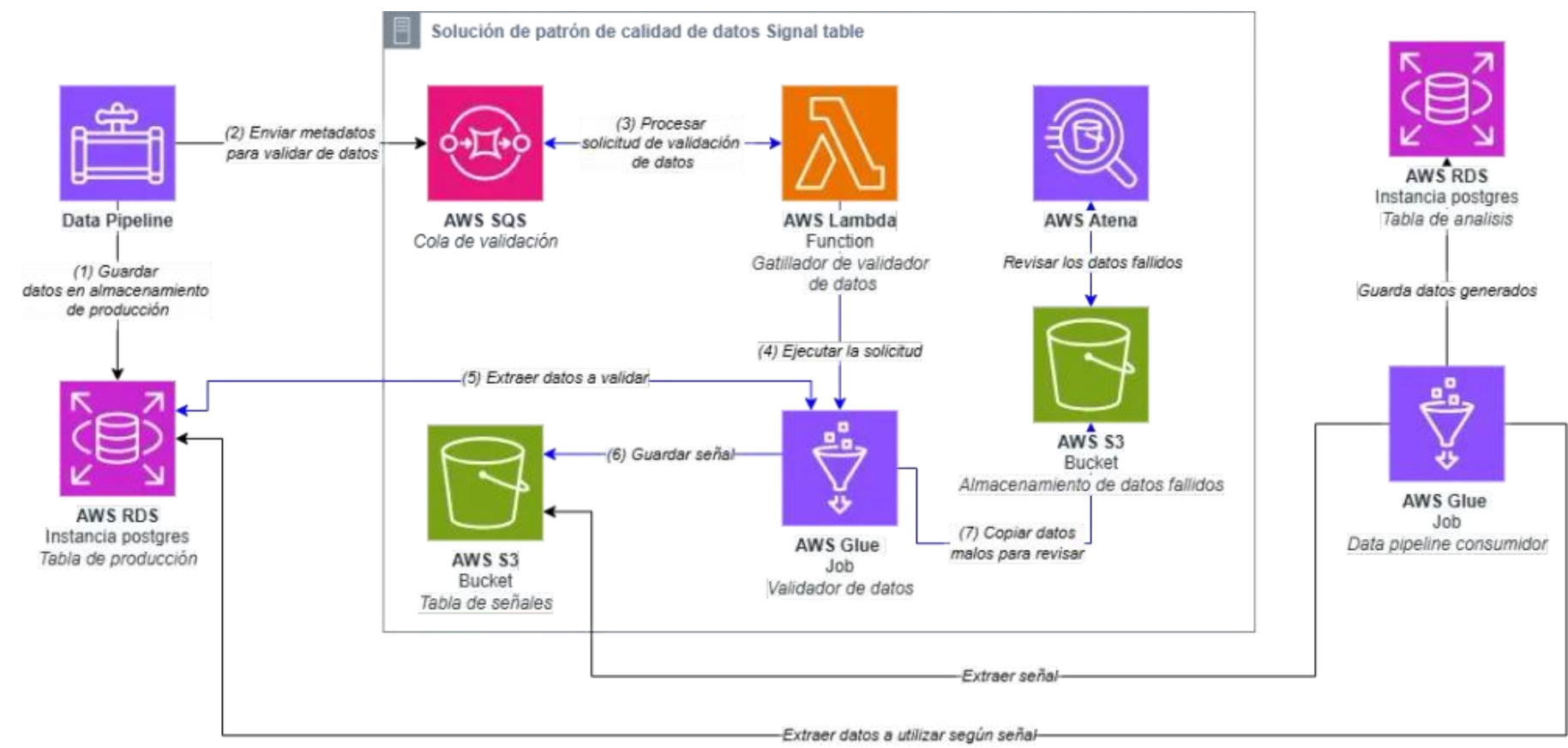
# Patrón Signal table para data pipeline



# Patrón Signal table para data pipeline



# Ejemplo del patrón Signal table



Consejos practicos y ejemplos.



# Algunos consejos practicos

## Ve lo que representa para el negocio

Es importante que el negocio entienda lo que implica el patrón de calidad de datos para sus procesos y relaciones con los equipos.



## Metadatos simplifica mucho

En los procesos de recuperación de datos como análisis de datos, simplifica el proceso de identificar cuales datos manipular.



## Diseña la gestion de errores

Entiendo como se hace la revisión de los datos fallidos, que metadatos requieres para ello, y cuanto metadatos necesitas para entender la situación.



# ¡Ejemplos al detalle para que armes tu PoC!

**Patrón de calidad de datos AWAP:**  
**solución dirigida por eventos para la**  
**gestión uso de un data pipeline y**  
**resultados generados**



**Patrón de calidad de datos Signal**  
**Table: una propuesta de solución**  
**arquitectural dirigida por eventos**



**Patrón de calidad de datos WAP**  
**clasico para data pipeline: una**  
**propuesta de solución para casos en**  
**streaming**





# Síguenme en mis redes



AWS Builder  
Center



LinkedIn

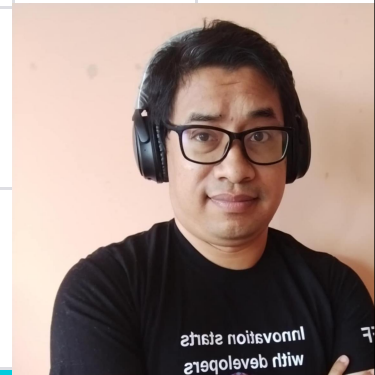


Youtube

15/01/2025

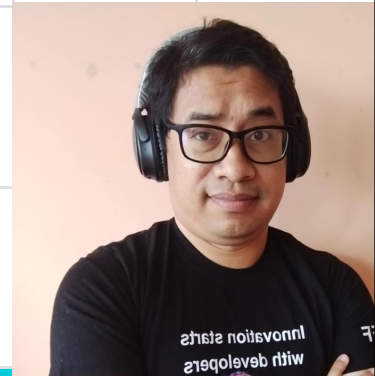


AWS User Group Arequipa





AWS User Group Arequipa



# Implementando patrones de calidad de datos en AWS

Pregunta y respuestas

15/01/2025

¡Gracias  
totales!