

AWS Glue Crawler / Data Catalog

- Glue crawler escanea datos en S3, crea esquemas
- Puede ejecutarse periódicamente
- El catálogo de datos de Glue (Data Catalog):
 - Solo almacena la definición de la tabla
 - Los datos originales permanecen en S3
- Una vez catalogados, puedes tratar tus datos no estructurados como si fueran estructurados:
 - Amazon Redshift Spectrum
 - Amazon Athena
 - Amazon EMR
 - Amazon QuickSight

