5.9.2. Actividad 2

Dado el siguiente caso de estudio:

Un hospital está buscando implementar un sistema de gestión de datos más eficiente para abordar los desafíos de la diversidad de datos médicos, como registros de pacientes, resultados de pruebas, imágenes médicas, informes de laboratorio, etc. El objetivo es mejorar la accesibilidad, la seguridad y la capacidad de análisis de estos datos para mejorar la atención al paciente, la investigación médica y la eficiencia operativa.

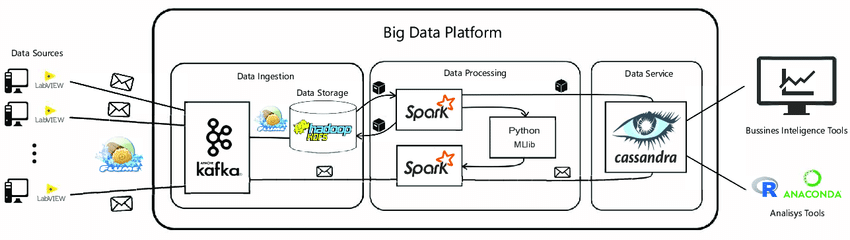
1. Analízalo e identifica los desafíos que podrían enfrentar y discute cómo un Data Lake podría abordar esos desafíos.

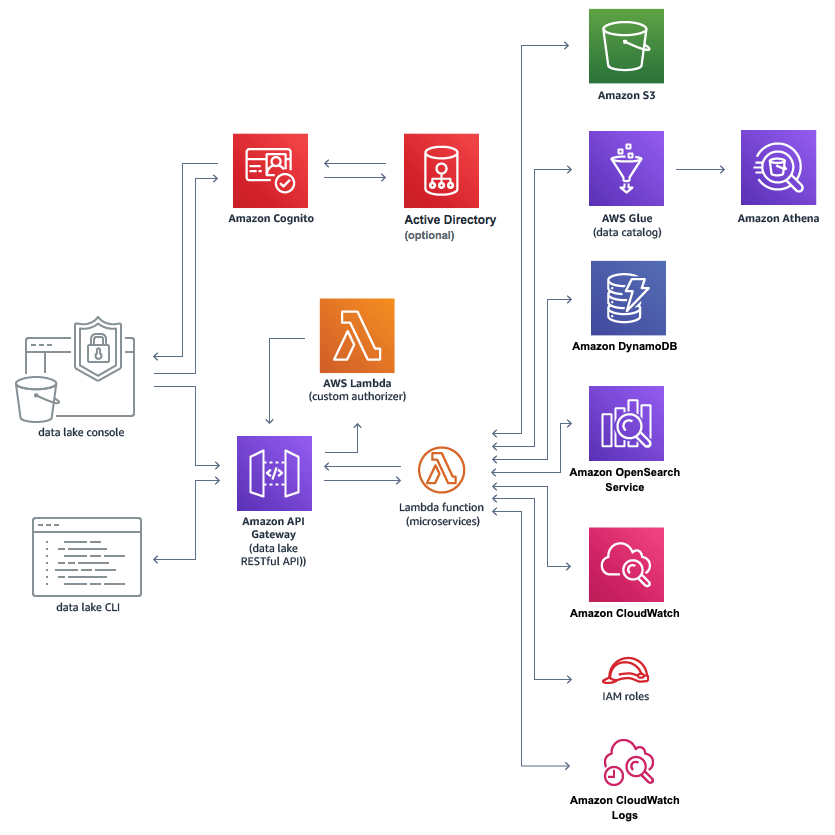
* El principal reto al que nos presenta el problema es la capacidad de analizar y ser capaces de extraer información útil de datos con las siguientes características:
  + Fuentes distintas de datos (batch, tiempo real)
  + Variedad de datos (estrucurados, no estrucutrados, semi-estructurados)
  + Datos no procesados
  + Sensibles
  + Sistemas de gestión y aplicaciones: Obtener datos sin modificar los datos de las fuentes de las que se obtiene

1. Recopila información sobre la arquitectura típica de un Data Lake.

Una arquitectura de típica de un Data Lake es:

* Capa de ingesta: se obtienen datos de diversas fuentes
* Capa de almacenamiento: se almacenan los datos sin procesar
* Capa de procesamiento de datos: Transformar los datos en información que pueda ser útil.
* Capa de organización: para organizar y catalogar la información con el fin de que sea fácil localizarla
* Análisis de datos y consumidores: donde los programas de BI o análisis de datos buscan la información
* Seguridad:
* Data lifecycle: controlar los datos según políticas de empresas, leyes, etc





1. Diseña la arquitectura de un Data Lake para esa empresa, identificando componentes clave, herramientas utilizadas y cómo abordarían los desafíos específicos del escenario.
2. Presenta tu diseño conceptual al resto de la clase.