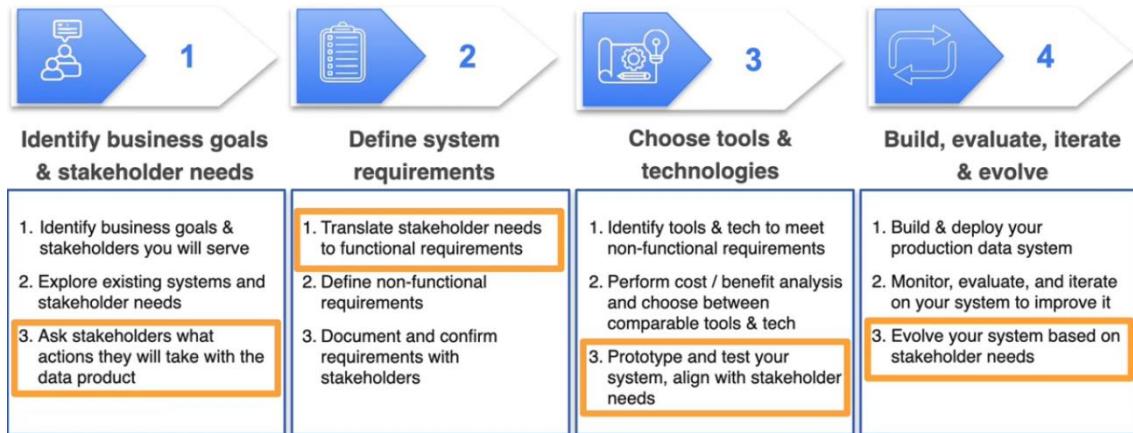


## Pensando con un ingeniero de datos

Relevamientos de Requisitos, definición de los requerimientos del sistema, selección de la tecnología, y construcción, evaluación, iteración y resolución y entrega de la solución.

## Thinking Like a Data Engineer



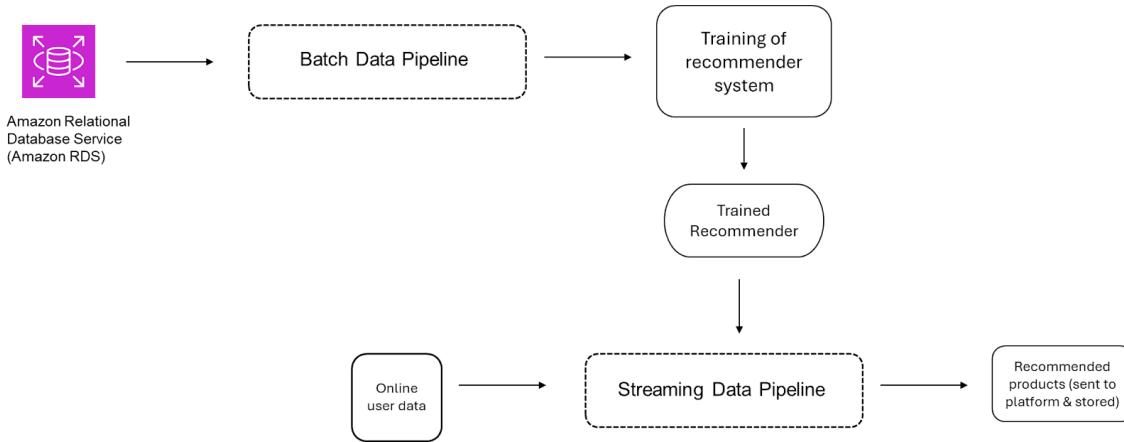
## Conversación para llevar

**Estado: Traducido automáticamente del Inglés**

**Traducido automáticamente del Inglés**

Estos son los puntos clave de esta conversación. Su tarea consiste en crear dos canalizaciones de datos:

- Una canalización de datos por lotes que sirve datos de entrenamiento para que los científicos de datos entrenen el sistema de recomendación;
- Un flujo de datos que proporcione las mejores recomendaciones de productos a los usuarios en función de su actividad en línea. El flujo de datos utiliza el sistema de recomendación entrenado para proporcionar las recomendaciones.



El sistema de recomendación recomendará productos a los clientes basándose en su información de usuario y en los productos que el cliente ha estado viendo o que están en su cesta.

#### Notas:

- El laboratorio se centra en proporcionar un ejemplo de canalización de datos de extremo a extremo. Sus tareas principales consistirán en interactuar con los componentes de los procesos por lotes y de flujo.
- El laboratorio no se centra en el ajuste fino del rendimiento del modelo de recomendación. Además, los datos que se utilizan son los mismos que se usaron en la semana 2, y puede que no contengan buenas características predictivas de usuarios y productos. El objetivo del laboratorio no es evaluar el rendimiento del problema de la recomendación, sino incorporar sus cálculos en el flujo de datos.
- El modelo de recomendación calculará dos vectores: el vector de incrustación del producto y el vector de incrustación del usuario. A continuación, generará recomendaciones de productos basadas en estos vectores. La incrustación del producto es un vector que contiene información sobre las características del producto, y la incrustación del usuario es un vector que contiene información sobre las preferencias del usuario para cada una de las características del producto representadas en el vector de incrustación del producto. Para saber más sobre las incrustaciones vectoriales y cómo funciona el sistema de recomendación, no dudes en consultar el siguiente punto de lectura opcional.