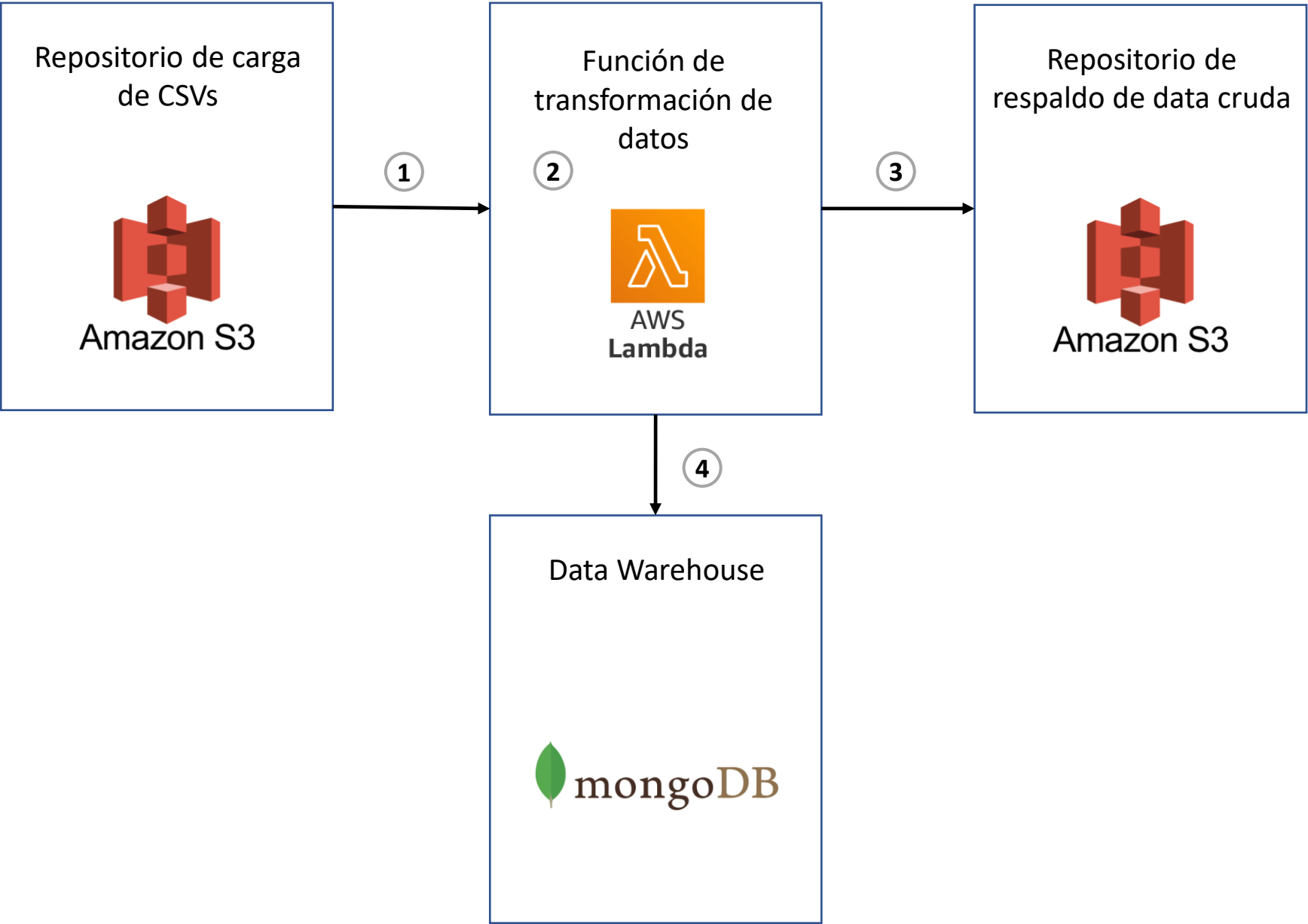
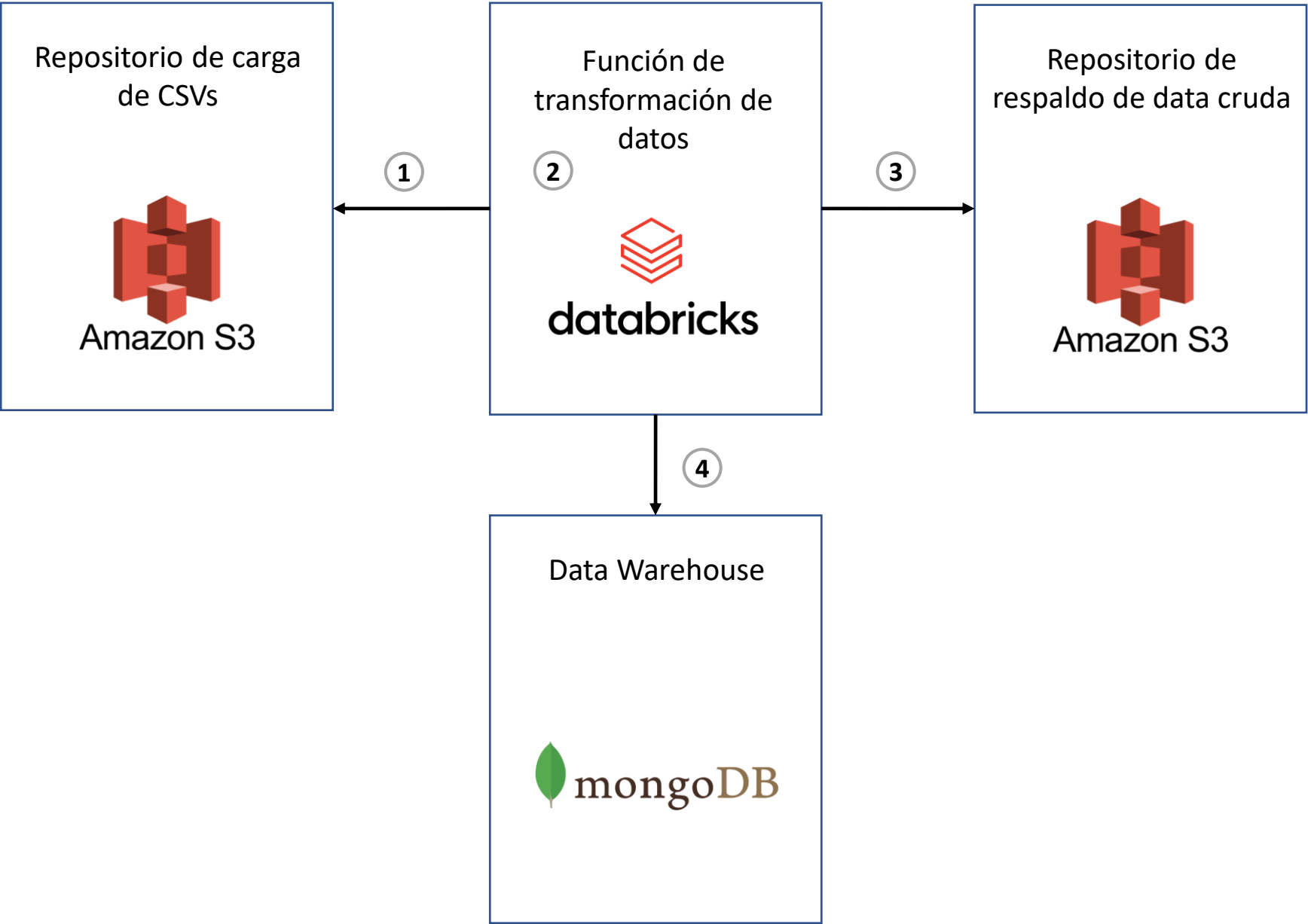


Arquitectura - Opción 1



1. Se sube un CSV al bucket de S3.
2. Se “dispara” una función lambda, la cuál realiza limpieza sobre los registros del CSV, además de nutrir la data con información adicional de postcodes.io.
3. Se respalda la data cruda (CSVs) en un bucket distinto de S3.
4. Se carga la información corregida y aumentada al Warehouse en MongoDB.

Arquitectura - Opción 2



1. Databricks consulta el bucket de S3 al que se suben los CSVs crudos.
2. Se activa una función lambda, la cuál realiza limpieza sobre los registros del CSV, además de nutrir la data con información adicional de postcodes.io.
3. Se respalda la data cruda (CSVs) en un bucket distinto de S3.
4. Se carga la información corregida y aumentada al Warehouse en MongoDB.