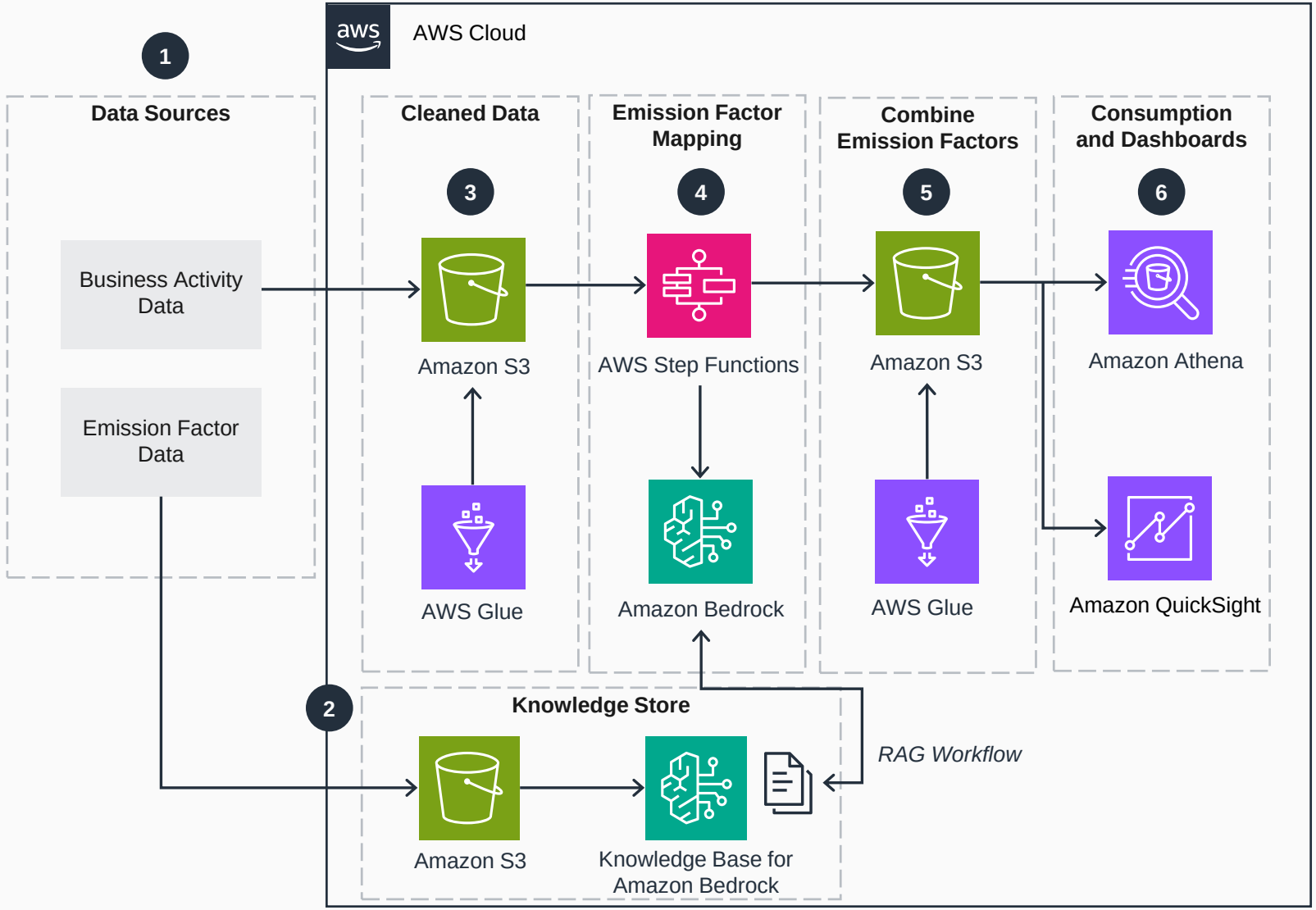
Guidance for Environmental Impact Factor Mapping on AWS

Option 1: Amazon S3 and Amazon Bedrock Knowledge Store

This architecture diagram shows how to automate selection of EIFs using foundation models to scale carbon footprint assessments.



1. Collect business activity data and emission factors from appropriate sources.
2. Index relevant activity category datasets using knowledge bases for **Amazon Bedrock** to build vector knowledge stores for searching during mapping.
3. Ingest the activity data to **Amazon Simple Storage Service (Amazon S3)**. **AWS Glue** cleans and formats incoming data for mapping activities to impact factors.
4. **AWS Step Functions** iterates through the business activities. It calls large language models (LLMs) hosted by **Amazon Bedrock** to match the given activity to the most relevant emission factor category using RAG. Matches are stored in **Amazon S3**.
5. **AWS Glue** combines the matched activity and category with its corresponding kg CO2e from relevant emission factor datasets. This combined dataset is registered in **AWS Glue Data Catalog**.
6. Report on mapped emission factor data with **Amazon QuickSight** or **Amazon Athena**.

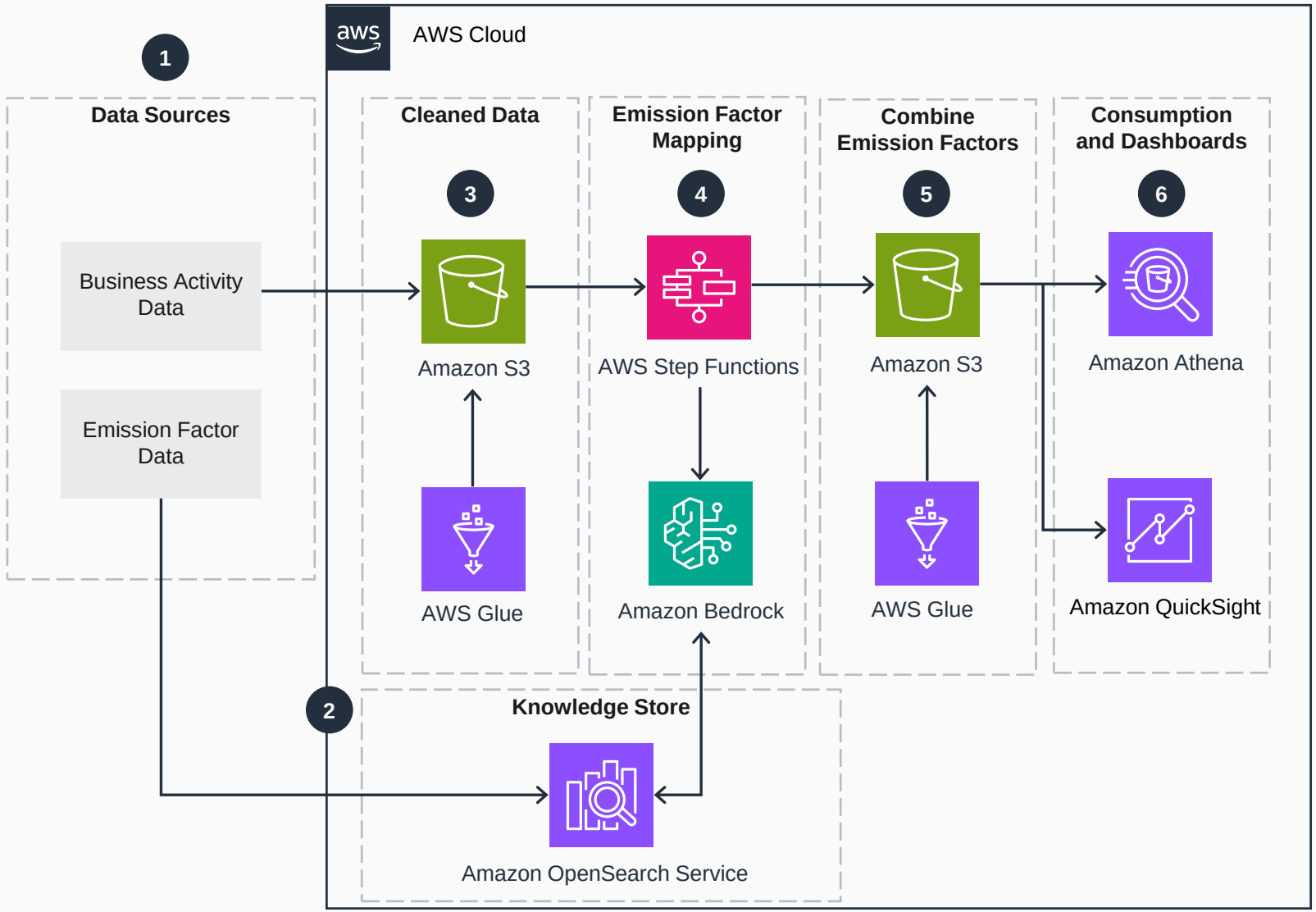
Reviewed for technical accuracy August 19, 2024

AWS Reference Architecture

Guidance for Environmental Impact Factor Mapping on AWS

Option 2: Amazon OpenSearch Service Neural Search Knowledge Store

This architecture diagram shows how to automate selection of EIFs using foundation models to scale carbon footprint assessments.



1. Collect business activity data and emission factors from appropriate sources.
2. Use **Amazon OpenSearch Service** Neural Search feature to convert emission factor query text into vector embeddings, and return close matches.
3. Ingest the activity data to **Amazon S3**. **AWS Glue** cleans and formats incoming data for mapping activities to impact factors.
4. **Step Functions** iterates through the business activities. It calls large language models (LLMs) hosted by **Amazon Bedrock** to match the given activity to the most relevant emission factor category using RAG. Matches are stored in **Amazon S3**.
5. **AWS Glue** combines the matched activity and category with its corresponding kg CO2e from relevant emission factor datasets. This combined dataset is registered in **Data Catalog**.
6. Report on mapped emission factor data with **QuickSight** or **Athena**.

AWS Reference Architecture