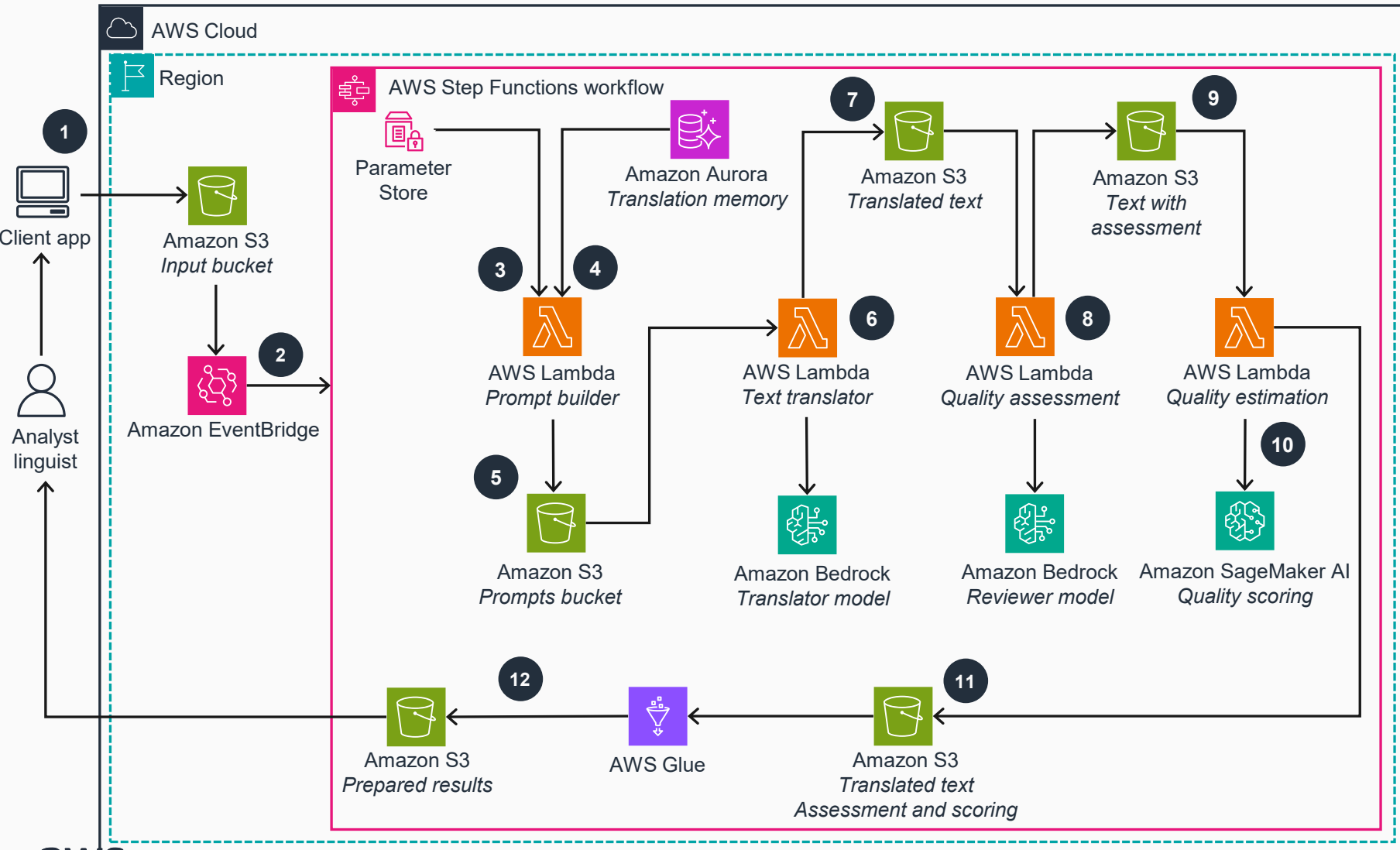


Guidance for Machine Translation Pipelines Using Generative AI on AWS

This architecture diagram demonstrates a serverless pattern enabling highly personalized batch machine translation using foundation models from Amazon Bedrock. AWS Step Functions orchestrates the various steps, including LLM-driven quality assessment and quality score estimation through a COMET ML model on Amazon SageMaker AI.



- 1 The user uploads source text to the **Amazon Simple Storage Service (Amazon S3)** input bucket to initiate the translation process.
- 2 **Amazon EventBridge** invokes **AWS Step Functions** to start the translation workflow.
- 3 An **AWS Lambda** function begins the **Step Functions** execution, receiving configuration parameters from **Parameter Store**, a capability of **AWS Systems Manager**, for secure management.
- 4 **Lambda** fetches the translation memory from **Amazon Relational Database Service (Amazon RDS) Aurora PostgreSQL** and generates translation prompts.
- 5 **Lambda** stores translation prompts in the **S3** input bucket.
- 6 **Lambda** invokes foundation models hosted on **Amazon Bedrock** to perform machine translation.
- 7 The translated outputs from **Amazon Bedrock** are stored in the **S3** model output bucket.
- 8 **Lambda** invokes **Amazon Bedrock** for LLM-driven qualitative assessment.
- 9 Quality assessment results from **Amazon Bedrock** are stored in the **S3** consolidated results bucket.
- 10 **Lambda** invokes the **Amazon SageMaker AI** endpoint for COMET ML score estimation.
- 11 **SageMaker AI** evaluation results are consolidated with the output and stored in the **S3** consolidated results bucket.
- 12 **AWS Glue** prepares the consolidated results for end-user consumption and analysis.