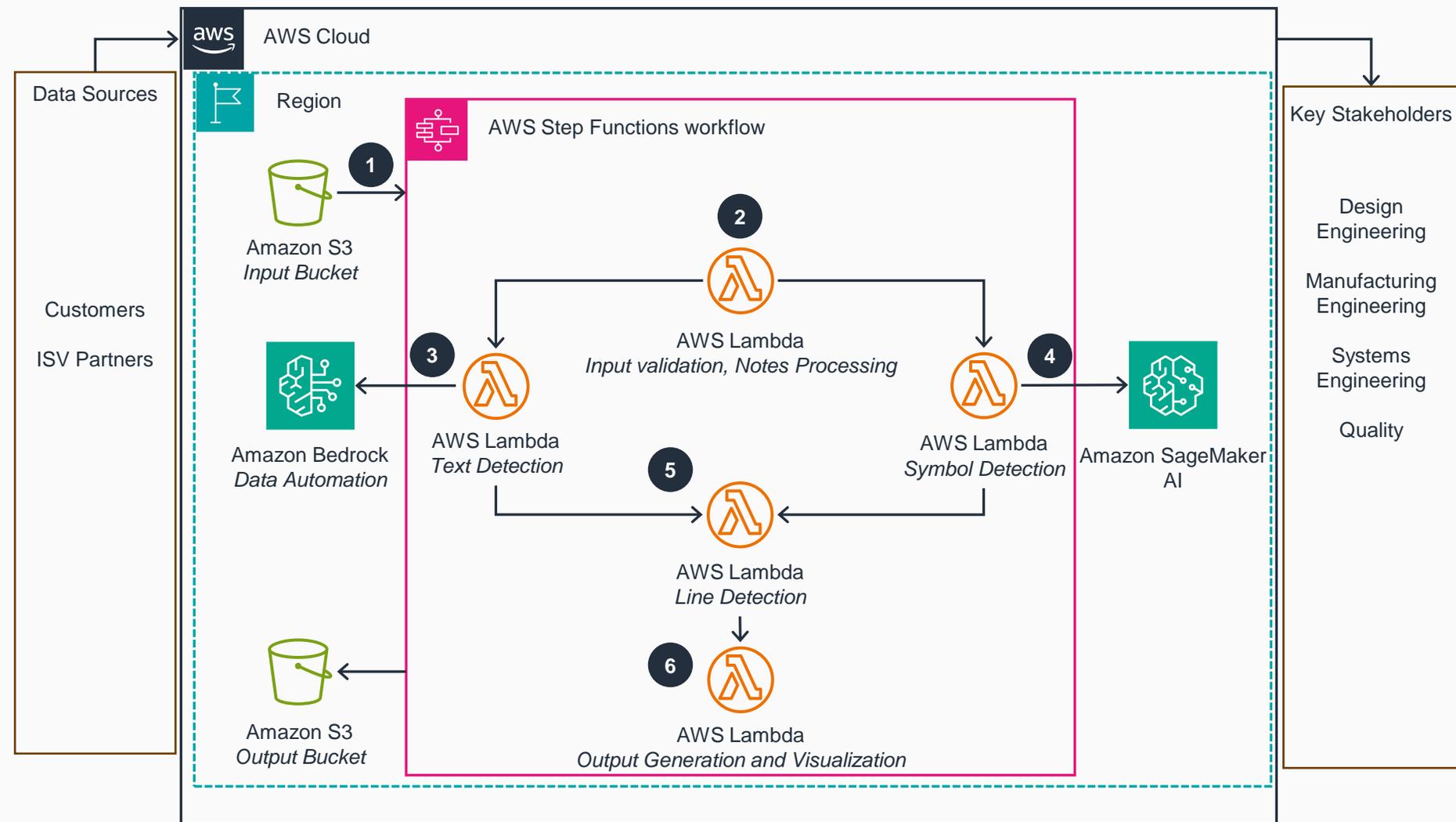# Guidance for Piping and Instrumentation Diagram Digitization on AWS

This architecture diagram shows how to convert paper and digital P&ID drawings into structured engineering data on AWS, helping customers automate technical drawing digitization and standardization. This slide shows steps 1-5.

1. User uploads a Piping and Instrumentation Diagram (P&ID) diagram to the **Amazon Simple Storage Service (Amazon S3)** input bucket.

2. The user invokes the **AWS Step Functions** workflow. The **AWS Lambda** Input Validation and Notes Processing function validates the input file, processes any notes and annotations, and prepares the document for parallel processing. The workflow splits into parallel processes for text and symbol detection.

3. The **AWS Lambda** Text Detection function identifies and extracts text elements from the P&ID Diagram leveraging **Amazon Bedrock** for Data Automation. **Amazon Bedrock** offers superior accuracy in technical drawing text extraction compared to other services, with specialized capabilities for engineering notation, multiple orientations, and handling large number of elements typical in P&ID diagrams in a fully managed service.

4. The **AWS Lambda** Symbol Detection function integrates with the fully managed service **Amazon SageMaker AI** to deploy two deep learning models: a Faster R-CNN (Region-based Convolutional Neural Network) model that detects and localizes P&ID symbols within the diagram, and a Siamese Neural Network that compares detected symbols against a reference database for accurate classification.

5. This custom model architecture enables precise identification of industry-specific P&ID symbols and connections.

   The **AWS Lambda** Line Detection function detects and processes the pipeline connection and flow lines.

AWS Reference Architecture

Guidance for Piping and Instrumentation Diagram Digitization on AWS

This architecture diagram shows how to convert paper and digital P&ID drawings into structured engineering data on AWS, helping customers automate technical drawing digitization and standardization. This slide shows step 6.

AWS Reference Architecture

---

6 The **AWS Lambda** Output Generation and Visualization function processes and combines the extracted data to generate multiple output formats: machine-readable JSON or Data Exchange in the Process Industry (DEXPI) files containing symbol coordinates and relationships.

Customers can access these outputs through the **Amazon S3** bucket, where they'll find both the structured data files and visual representations that highlight detected symbols, connections, and text annotations from the original static P&ID input.