Cost Comparison Analysis

This document provides a cost comparison analysis of two approaches for processing documents: 1) Using AWS Comprehend for PII reduction and metadata extraction, with AWS Bedrock LLMs for specific documents, and 2) Using LLMs for every document. The analysis assumes a total of 10,000 documents with varying sizes and types.

# Assumptions

Document Distribution for 10,000 Documents:  
 - Large Documents (4,000 docs): Average size of 5 MB  
 - Medium Documents (3,000 docs): Average size of 1 MB  
 - Small Documents (3,000 docs): Average size of 0.2 MB  
  
Document Types and Requirements:  
 - Structured Documents (30%) - 3,000 docs (50% generic, 50% need LLMs)  
 - Semi-structured Documents (40%) - 4,000 docs (30% generic, 70% need LLMs)  
 - Unstructured Documents (30%) - 3,000 docs (10% generic, 90% need LLMs)

# Scenario 1: Using AWS Comprehend + Bedrock LLM for Specific Documents

Step 1: Document Size and S3 Storage Cost  
Total document size: 23.6 GB  
S3 Storage Cost: $0.5428/month  
  
Step 2: AWS Lambda Cost  
Lambda Request Cost: $0.002/month  
Lambda Duration Cost: $0.0011/month  
  
Step 3: AWS Comprehend Costs  
PII Reduction: $15.00/month  
Entity Recognition & Keyphrase Extraction: $10.00/month  
  
Step 4: AWS Bedrock (LLMs) Costs  
Documents requiring LLMs: 7,000  
Total tokens processed: 21,000,000  
Cost of Bedrock (LLMs): $420.00/month  
  
Total Monthly Cost for Scenario 1: $445.55/month

# Scenario 2: Using LLM for Every Document

Step 1: Document Size and S3 Storage Cost  
Total document size: 23.6 GB  
S3 Storage Cost: $0.5428/month  
  
Step 2: AWS Lambda Cost  
Lambda Request Cost: $0.002/month  
Lambda Duration Cost: $0.0011/month  
  
Step 3: AWS Comprehend Costs  
PII Reduction: $15.00/month  
Entity Recognition & Keyphrase Extraction: $10.00/month  
  
Step 4: AWS Bedrock (LLMs) Costs  
Documents requiring LLMs: 10,000  
Total tokens processed: 30,000,000  
Cost of Bedrock (LLMs): $600.00/month  
  
Total Monthly Cost for Scenario 2: $625.55/month

# Comparison Summary for 10,000 Documents

|  |  |  |
| --- | --- | --- |
| Cost Component | Using Comprehend + LLM (Specific Docs) | Using LLM for Every Document |
| S3 Storage | $0.5428 | $0.5428 |
| S3 Event Notification | $0.006 | $0.006 |
| AWS Lambda | $0.0031 | $0.0031 |
| AWS Comprehend (PII + Metadata) | $25.00 | $25.00 |
| AWS Bedrock (LLMs) | $420.00 | $600.00 |
| Total Monthly Cost | $445.55 | $625.55 |

# Conclusion

Using AWS Comprehend + Bedrock LLM for specific documents is more cost-effective, with a total monthly cost of $445.55/month compared to $625.55/month for using LLM for every document. The savings arise from processing only the specific documents with LLMs, significantly reducing token usage and processing time.