Automated Receipt Processing System using AWS Textract

Leveraging Serverless Architecture to Digitize and Organize Receipt Data

Presented by: Pravin Balaji K N
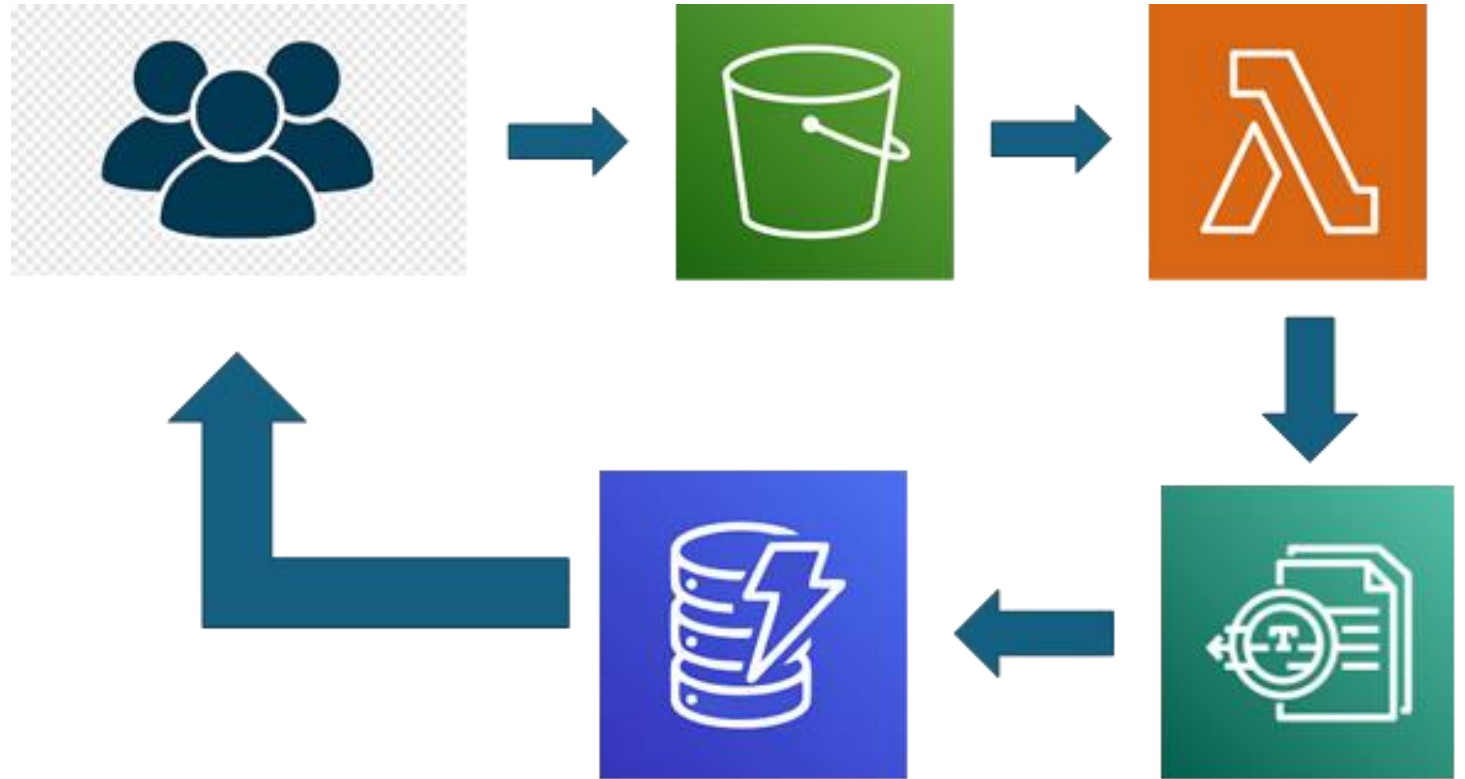
Introduction

The "Automated Receipt Processing System" is a serverless application designed to transform how businesses manage and digitize expense receipts. Built entirely on AWS services, this project demonstrates an efficient, scalable, and intelligent approach to automating what was traditionally a manual, error-prone process.

Using Amazon Textract's advanced AI-powered extraction capabilities and integrating it with event-driven Lambda functions, this system processes and stores structured receipt data in real-time — all without human intervention. It showcases how modern cloud technologies can be applied to streamline operational workflows and deliver business value.

Architecture Overview

- **Amazon S3:** For uploading and storing receipt images.

- **AWS Lambda:** For orchestration and automation.

- **Amazon Textract:** For extracting structured data from receipts.

- **Amazon DynamoDB:** For storing and querying extracted receipt data.

Flow Summary

1. User uploads a receipt image to an S3 bucket.

2. S3 triggers a Lambda function.

3. Lambda invokes Textract to analyze the image.

4. Extracted data is parsed and stored in DynamoDB.

5. Logging and diagnostics handled via CloudWatch.

Problem-Solution Summary

The Problem:

- Manual receipt processing slows operations.

- High potential for human error.

- Unstructured data is hard to analyze.

Our Solution:

- Upload → Analyze → Store → Query – all automated.

- Accurate data extraction using AWS Textract.

- Real-time processing with Lambda triggers.

- Durable storage and easy retrieval from DynamoDB.

Primary Use Cases

- **Expense Management Systems** – Automatically log and analyze business expense receipts.

- **Accounting & Auditing Firms** – Speed up data collection from physical receipts.

- **Retail Chains** – Analyze purchasing trends and customer receipts.

- **Travel & Logistics** – Automate reimbursement workflows.

- **Healthcare Providers** – Process pharmacy or medical receipts for insurance claims.

Benefits & Value Proposition

- **Efficiency:** Saves hours of manual data entry.

- **Accuracy:** Reduced risk of human errors.

- **Cost-Effective:** Pay-as-you-go pricing with AWS serverless.

- **Rapid Deployment:** Easily deployable with minimal setup.

- **Scalable:** Capable of handling thousands of receipts.

Future Enhancements:

- Add a web-based dashboard to view and search receipts.

- Include user authentication and upload limits.

- Add OCR support for multilingual receipts.

Conclusion:

The Automated Receipt Processing System is a robust and elegant solution for modernizing expense management. It demonstrates the power of AWS serverless services and AI to eliminate inefficiencies, drive accuracy, and reduce operational load.