Automated Data Extraction for Healthcare Claims Processing

Overview:

The client wanted us to develop an application that would help them automate the high volume and arduous task of extracting data arriving from multiple sources and assist in claim processing. We built them a solution that extracts data from the ROI request forms, presents it to an operator for validation, and after doing so sends the results back to ROI Online through its API. Mindfire Solutions has implemented custom machine vision, natural language processing (NLP), and Robotic Process Automation (RPA) within the application to facilitate attended automation of the logging-and-assignment workflows. The logging-and-assignment application also enables the operator to pragmatically forecast the productivity gains that automation yields when deployed at scale on the Kubernetes cluster.

Client details:

Name: Confidential | Type: Healthcare | Location: USA

Technology:

Python, Django, GraphQL, Vue.js, PWA, Azure, REST, Azure SQL, Google Vision, Amazon Textract, Solr, GIT, Docker, Kubernetes, Azure Key Vault, Azure BLOB storage, Jenkins

Project Description:

It is a GraphQL based Artificial Intelligence platform that allows admins to add tasks with a bunch of forms from different sources like manual upload, SFTP servers, O365 email, Retarus (Fax) APIs etc. to identify and extract specific information. Forms are internally processed via Google Vision, Amazon Textract and Azure Form Recognizer and predefined extraction rules are using the pull the required information.

Human Operator dashboard is populated with the extracted information for validating the machine vision (MV) results. The analytics page allows administrators to oversee performance analytics by form.

The application has three types of OCRs to extract the contents from the request forms:

1. Azure Form Recognizer
2. Google Vision
3. Amazon Textract

The OCR can be easily switched from one to another from the application configuration. After trying many solutions we decided to use Google Vision due to the data accuracy.

Azure Services Used:

Azure SQL (Database):
- Used SQL Database to store all the application related data in relational format.

Azure Key Vault:
- Used to store all the confidential data related to the application.

Azure BLOB Storage:
- Used to store all the images and pdfs of the processed request forms.

Azure Kubernetes Service (AKS):
- Used to deploy the application in Kubernetes cluster
- AKS is Kubernetes service provided by azure

Azure Container Registry (ACR):
- Used to store Docker images on the Azure cloud
- AKS uses Docker images present on ACR

Azure Form Recognizer:
- Used to identify the kind of template a request form belongs to
- AFR takes a set of form for a particular type of template as training set and it trains a ML model
- It uses the trained ML model to identify the kind of template for a request form

Google Services Used:

Google Vision (OCR):
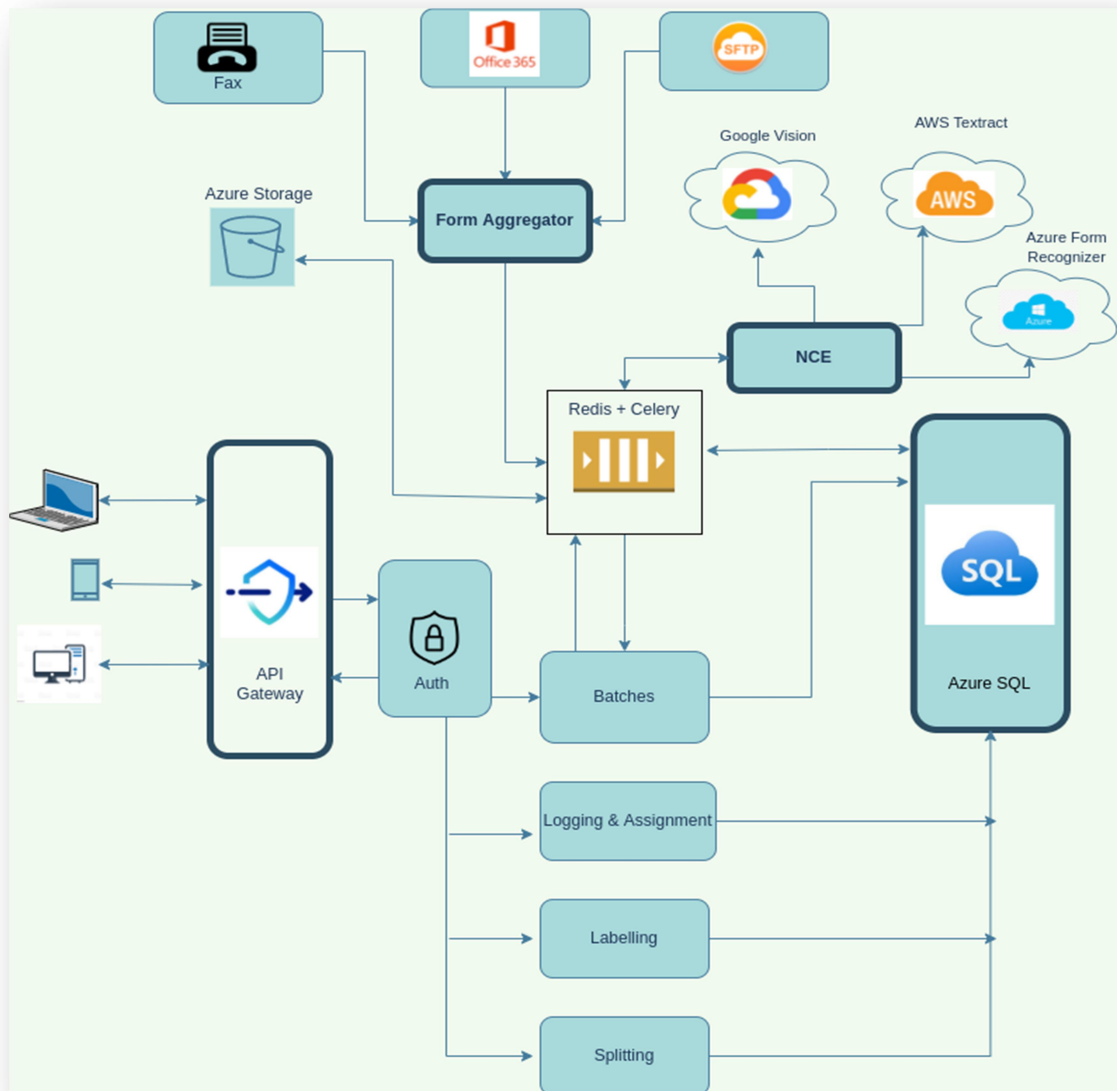- Used to extract the contents of request forms

- The Google Vision response is used to identify the kind of template for a request form
- All the required information are extracted using the predefined rules

AWS Services Used:

Textract (OCR):
- Used to extract the contents of request forms
- The Textract response is used to identify the kind of template for a request form
- All the required information are extracted using the predefined rules

Work-Flow Diagram:

Architecture Diagram:

Kubernetes Cluster