

This document is based on a prototype and is not meant to be a full solution.

# Arizona Department of Homeland Security Prototype Report



For prototyping purposes

## Table of Contents

<b>Overview</b>	<b>3</b>
<b>Problem Description</b>	<b>3</b>
<b>Architecture Diagram</b>	<b>3</b>
<b>AWS Cloud Services</b>	<b>4</b>
Bills of Services	
Total cost	
<b>Process</b>	<b>5</b>
Uploading the documents into the S3 bucket and Kendra Sync	5
Querying using the User Interface	6
<b>Project Outcomes</b>	<b>7</b>
<b>Out of Scope Items</b>	<b>8</b>
<b>Lessons Learned</b>	<b>9</b>
<b>Appendix</b>	<b>9</b>
Visual	9
Credits	10
Developers	10
License	10
References	10

For prototyping purposes

## Overview

The Arizona State University Cloud Innovation Center, in partnership with the Arizona Department of Homeland Security (AZ DOHS), is developing the Cyber Homeland Information Portal (CHIP), a prototype solution that uses generative AI to provide real time support for organizations participating in AZ DOHS's Cyber Readiness Program on cyber security tool implementation. Organizations with little or no IT resources can now leverage tools to secure their information systems with the help of AZ DOHS and the CHIP application.

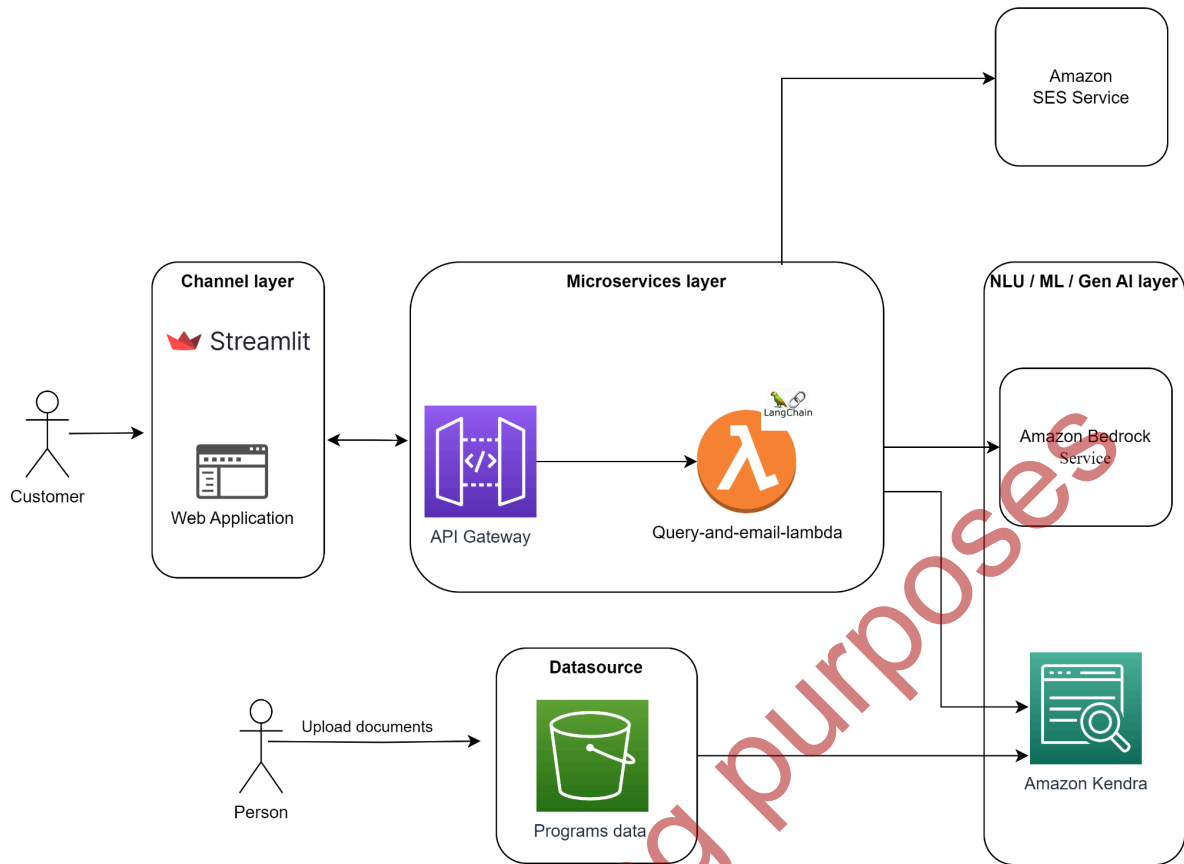
## Problem Description

Small organizations participating in AZDOHS's Cyber Readiness Program encounter difficulties in grasping and applying the various cyber security resources. This results in delayed responses to their inquiries due to the constraints of a limited support staff. In response to this challenge, the innovative CHIP application has been developed, incorporating cutting-edge technologies such as AWS, Machine learning and Generative AI. The primary objective of this application is to furnish participants with real-time and concise answers to their queries.

Accessible through the AZDOHS website, the CHIP application streamlines the user experience by not only providing immediate assistance but also offering document links for additional reference. This ensures that participants can verify the information they receive. Moreover, the application has the capability to escalate to human support when the complexity of a query demands personalized assistance, thereby guaranteeing a smoother overall Cyber Readiness Program experience for participating organizations.

## Architecture Diagram

The project is developed using AWS and the Streamlit framework. The documents are stored in the S3 bucket as a data source. When a user enters a query through the user interface, the request goes to the API gateway, triggering the Lambda function. This function handles the retrieval of relevant text data and documents from Amazon Kendra, using the Langchain framework and bedrock service to work with the LLM model. Afterward, the LLM's response is passed as a response and displays on the user interface to the user. If the LLM can't provide a proper answer based on the obtained data, the query is passed on to a human. This is done by sending an email using Amazon SES service that includes the user's details and the specific question.



## AWS Cloud Services

- AWS S3 as the datastore to store the documents.
- Lambda functions is employed fetching relevant data, providing LLM answers and sending emails.
- SES ts used to send emails to the support team, which includes user details and specific query.
- UI to provide an interface for submitting the Query
- EC2 instance to host the UI website and to manage communicate with the API Gateway and Lambda functions.

## Bills of Services

## Process

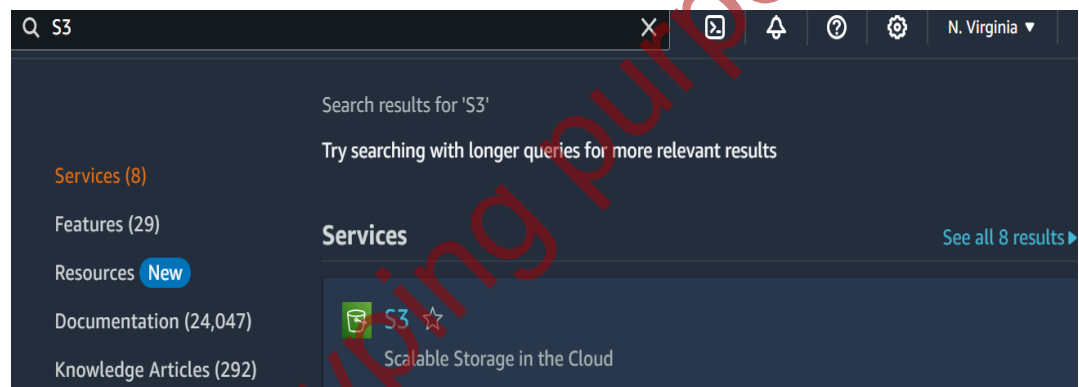
The process involves the two steps: Uploading documents and Querying using User interface.

- **Uploading the documents into the S3 bucket and Kendra Sync:**

It is a manual process. The uploading of the latest version of an existing document or uploading a new document has to be done manually. Once the documents are uploaded, the kendra index has to be synced. The below shows the steps with the screenshots.

1. **Access S3 Service:**

In the AWS Management Console, navigate to the "Services" dropdown. Search for "S3".

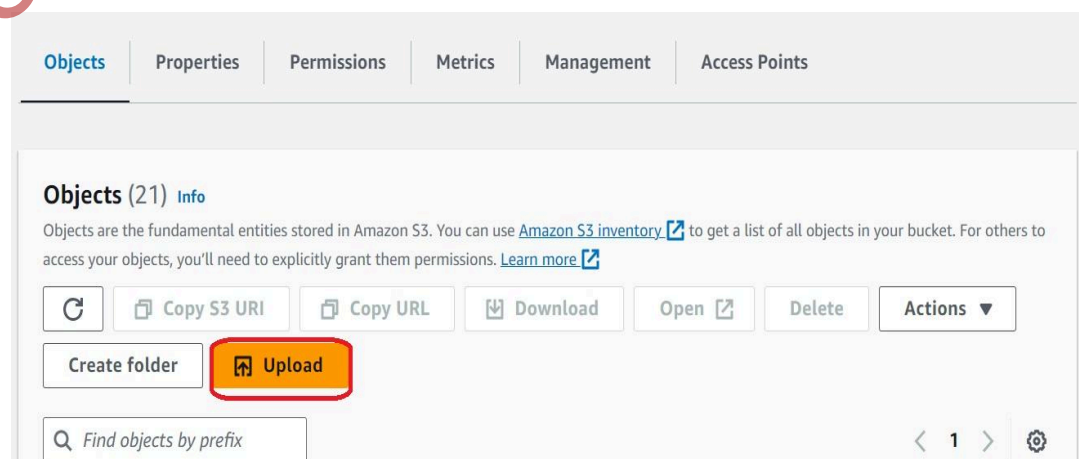


2. **Select Existing Bucket:**

From the list of buckets, click on the name of the existing bucket where you want to upload documents.

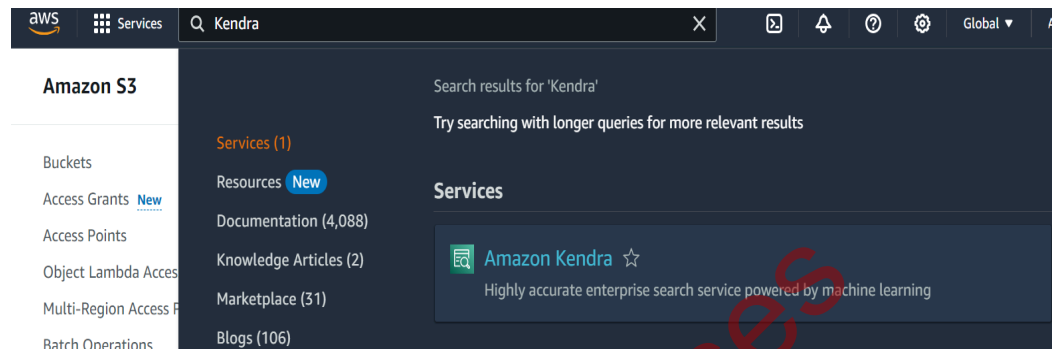
3. **Upload Documents:**

Within the selected bucket, click on the "Upload" button. Choose the files you want to upload and follow the prompts to complete the upload.



#### 4. Access Amazon Kendra:

In the AWS Management Console, navigate to the "Services" dropdown. Search for "Amazon Kendra"



#### 5. Select Index:

Choose the Kendra index that corresponds to the data source you want to synchronize.

#### 6. Navigate to Data Sources and Sync:

In the Kendra index settings, find the section related to "Data Sources". Click on the data source and click in the Sync option.

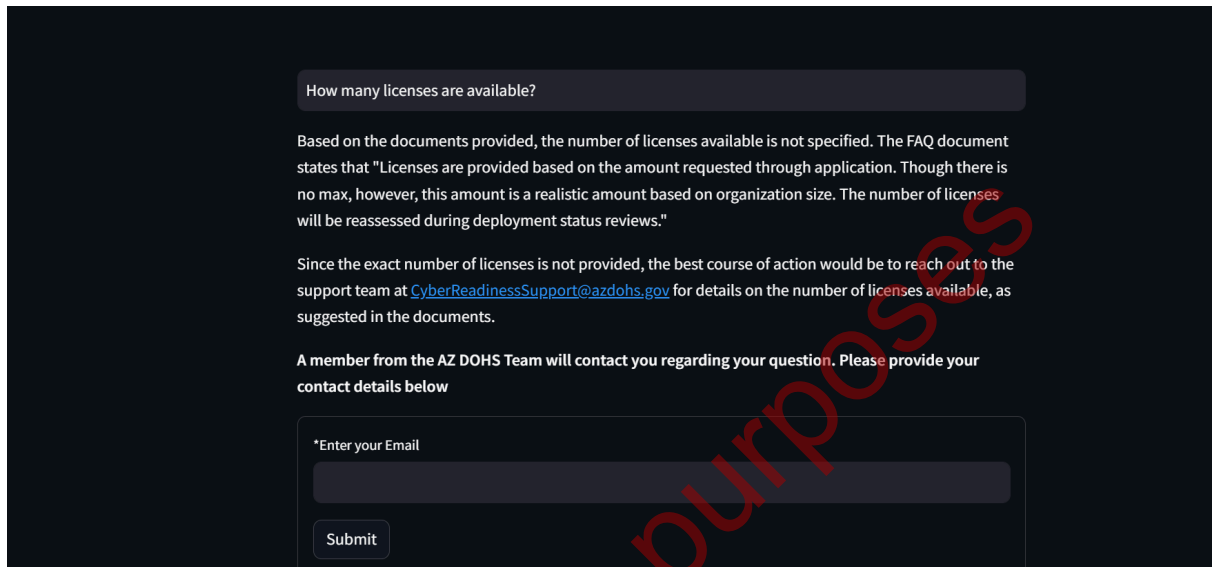
#### • Querying using the User Interface

The user interface features a text box for users to input their queries. Users can enter a query and receive a concise response along with relevant documents. To access the document, users can simply click on the provided link. The example query output and link are illustrated in the screenshot below.



## Providing an Email option if the tool not able to answer the query:

The tool additionally offers the option for users to input their email addresses, facilitating the sending of queries to the support team. The screenshot below displays an example query and the user interface.



The screenshot shows a dark-themed user interface. At the top, a search bar contains the query "How many licenses are available?". Below the search bar, the system's response is displayed in white text. The response explains that the number of licenses is not specified in the provided documents and suggests reaching out to the support team at [CyberReadinessSupport@azdohs.gov](mailto:CyberReadinessSupport@azdohs.gov). Below the response, there is a section titled "A member from the AZ DOHS Team will contact you regarding your question. Please provide your contact details below". This section contains a form with a label "\*Enter your Email" and a text input field. A "Submit" button is located below the input field.

## Project Outcomes

The Arizona State University Cloud Innovation Center (ASU CIC) collaborates with the Arizona Department of Homeland Security to develop an AWS-based prototype tool aimed at achieving the following objectives:

### 1. Answering User Questions with LLM Models and AWS Infrastructure:

Harnessing the capabilities of Language Models (LLM) and the AWS infrastructure, the system extracts relevant information from uploaded PDFs to provide precise responses to user queries.

### 2. Email Support Option:

Users have the opportunity to furnish their email addresses. In instances where the user's question lacks relevant data in the uploaded PDFs, the support team can subsequently reach out to the user via the provided email for further assistance.

### **3. Offloading Customer Associates' Work with for Frequently Asked Questions:**

The system efficiently offloads the workload of customer associates by deploying a tool to address frequently asked questions.

## **Out of Scope items**

### **Conversational Chatbot Functionality:**

The system is not designed for extended back-and-forth interactions typical of conversational chatbots. Its primary purpose is to provide specific responses to user queries rather than engaging in conversational dialogue.

### **Scheduling Demos with Vendor:**

The system lacks functionality for scheduling vendor demonstrations by inputting available dates and times. Users should seek alternative methods for coordinating vendor demos.

### **Feedback Loop for User Response:**

The project omits a feedback loop mechanism for users to express the helpfulness of responses through a thumbs-up or thumbs-down feature. User feedback on the system's performance is not part of the current scope.

### **Analytics Tracking on Usage and Questions:**

The system does not incorporate analytics tracking to monitor user usage patterns or record specific questions over time. Comprehensive analytical insights into user interactions are not within the project's current scope.

### **CHIP Tool Integration into Slack:**

Integration with Slack for the CHIP tool is not within the project's scope. Users should not anticipate tailored features or functionalities specifically designed for seamless integration with the Slack platform.

### **Use of YouTube Videos for CHIP Model Training:**

The project does not utilize YouTube videos as part of the training data for the CHIP model. Video content from YouTube is not factored into the model's learning process.

### **Exporting Embedded Links from PDF Source Documents as Answers:**

The system does not export embedded links from PDF source documents as part of its response. Users should be aware that the system focuses on extracting textual information and does not include embedded link extraction.



## **Returning Images, Graphs, or Visual Results:**

The system does not support the presentation of images, graphs, or visual results as part of its response. Users should recognize that the system prioritizes textual information delivery and does not incorporate visual elements in its output.

## **Lessons Learned**

### **AWS CDK Deployment Proficiency:**

Developed a strong understanding of the AWS Cloud Development Kit (CDK) and used it effectively to manage the deployment of infrastructure in AWS. This involved creating and handling AWS resources through code, providing a more dynamic and code-centric approach to setting up and managing infrastructure.

### **Framework Expertise with Streamlit and Langchain:**

Demonstrated technical proficiency by integrating and working with Streamlit for building the user interface of the application. Additionally, explored Langchain framework, for connecting the LLM models into project. These experiences have enriched my abilities to adapt to emerging frameworks.

### **Explored and implementation of AWS Services:**

Explored the different AWS services like the Amazon Kendra, Amazon SES etc. The Amazon Kendra, an AI-driven search service integrated it into the project's architecture. This included configuring, indexing, and querying data to optimize search functionalities. The Amazon SES is implemented in the project to send out emails in specific scenarios.

### **Explored and implementation of different LLM models:**

Explored and implemented various Language Model (LLM) models by integrating them into the project application using the Amazon Bedrock Service. This experience helps in understanding and effectively applying diverse LLM models.

## **Appendix**

### **Visual**

## Credits

Arizona-Department-of-Homeland-Security ([link](#)) is an open source software. The following people have contributed to this project.

## Developers:

- Vishnusai Kandati
- Loveneet Singh
- *Sr. Program Manager, AWS:* Jubleen Vilku
- *Solution Architect leader, AWS:* Arunachalam Arun
- *General Manager, ASU:* Ryan Hendrix

This project is designed and developed with guidance and support from the ASU Cloud Innovation Center and the Arizona Department of Homeland Security.

## License

This project is distributed under the Apache License 2.0 ([link](#))

## References

1. <https://docs.streamlit.io/>
2. <https://docs.aws.amazon.com/kendra/>
3. <https://aws.amazon.com/ses/>
4. <https://www.langchain.com/>

For prototyping purposes