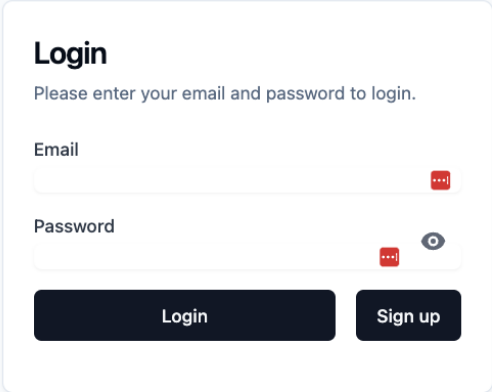


# Build a chatbot using RAG on the ThirdAI platform

This guide will walk you through building a chatbot using Retrieval-Augmented Generation (RAG) on the ThirdAI platform. Follow these steps to set up your chatbot with retrieval capabilities and customizable LLM guardrails.

## 1. Login to ThirdAI Platform [here](#)

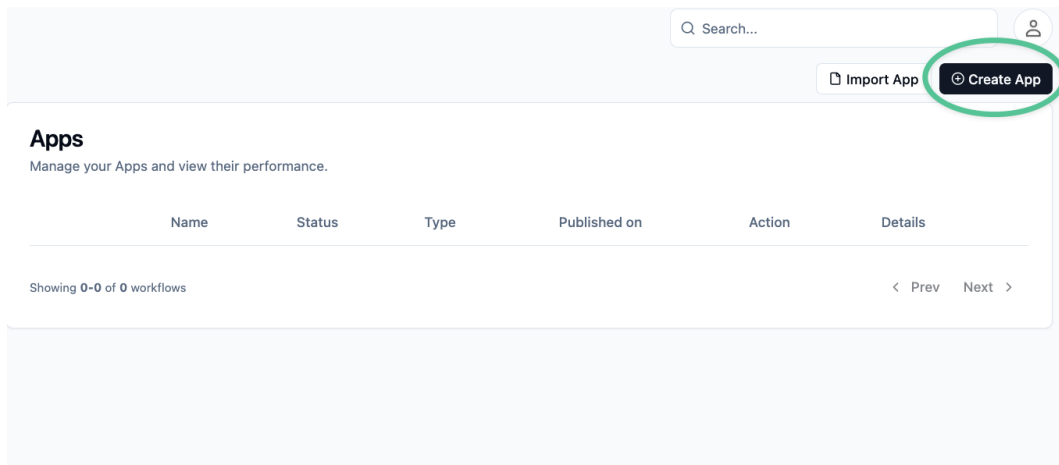
- After signing up, you'll receive an email with a verification link. Click the link to verify your email address.

A screenshot of the ThirdAI Platform login form. The form is titled "Login" and includes the instruction "Please enter your email and password to login." It features two input fields: "Email" and "Password". The "Email" field has a red error icon on the right. The "Password" field has a red error icon and a toggle icon (an eye) on the right. Below the input fields are two buttons: "Login" and "Sign up".

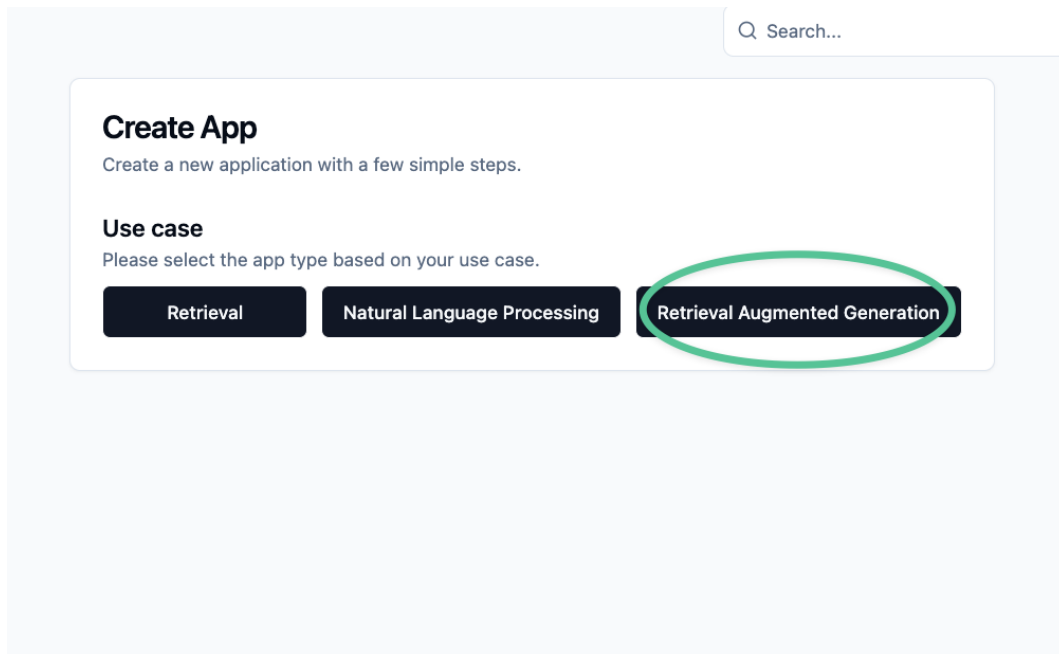
- Once verified, login to the platform using your credentials.

## 2. Create a New App

- After logging in, you'll be taken to the dashboard.
- Click on the "Create New App" button.



- You'll see three options: choose **RAG Workflow** to start building your chatbot.



### 3. Select or Create Retrieval App

- Start by specifying an app name of your choice (please avoid any spaces in the name)

## Create App

Create a new application with a few simple steps.

### Use case

Please select the app type based on your use case.

Retrieval

Natural Language Processing

Retrieval Augmented Generation

App Name

Retrieval App

LLM Guardrail

Chat

### App Name

Next

- Then you can either select an existing retrieval app if you have one or create a new one by specifying the name. In this case, let's go ahead and create a new one.

**Create App**  
Create a new application with a few simple steps.

**Use case**  
Please select the app type based on your use case.

Retrieval Natural Language Processing Retrieval Augmented Generation

App Name Retrieval App LLM Guardrail Chat

**Retrieval App**  
Use an existing retrieval app?

Yes No, create a new one

**App Name**  
Insurance-chat-Retrieval

**Sources**  
Select files to search over.

Add Local File Add S3 File

Create

Previous Next

- Click **Next** to proceed.

## 4. Load Documents

- You will now have the option to load documents from your local system or an S3 bucket. Choose the appropriate source and upload the documents. In this case, let's upload the documents included in the package.

**Sources**  
Select files to search over.

Choose Files No file chosen

Add Local File Add S3 File

Delete

## 5. Add LLM Guardrails for PII Redaction

- During setup, you'll be prompted to add guardrails to your app.

- You can choose which LLM guardrail model to use.
  - Let's create a RAG application without LLM guardrails in the interest of time.
  - Once configured, click **Next**.
- 

## 6. Choose an LLM Option

- Now, you'll select an LLM to power the chatbot.
  - The platform offers three options:
    - **OpenAI API** for a managed solution.
    - **Self-Hosted LLM** if you have your own models.
    - **On-Prem LLM** for businesses with strict data requirements.
- 

## 7. Finish Creating Your App

- After setting everything up, click the **Create** button to finalize your RAG-based chatbot.
- 

## 8. Access the Model Screen

- You will be redirected to the dashboard.
  - Here, you'll see various models listed. Find the **Retrieval Augmented Generation** model you just created.
  - Click **Start** to activate it.
  - You may also notice "retrieval" and "NLP" models that are auto-created—these are for backend purposes and can be ignored for now.
- 

## 9. Open the Generative Search Interface

- Once the RAG model is activated, you'll see an **Endpoint** button.
  - Click on it to open the **Generative Search** interface for your documents.
- 

## 10. Enable Chat Mode - Aha!

- In the Generative Search interface, you'll see a **Chat** icon in the top right corner.
  - Click on the chat icon to enable chatbot mode instead of traditional search.
- 

## 11. View Loaded Documents

- You can also click on the **View Documents** option to review all the documents you've uploaded for the chatbot to retrieve from.

## How can we help?

Ask anything...

View Documents

## 12. Upvote and Associate to fine-tune the application

- Each reference has a buttons to **upvote** and **downvote** it, which updates the model so that it improves.



### References

scan balance.pdf [Read source ↗](#)

For inpatient mental health stays you pay per benefit period: \$0 copayment per day for days 1-90. Your inpatient benefits are based upon the date of admission. If you are admitted to an inpatient mental health facility in 2024 and are not discharged until 2025 the Services that are covered for you What you must pay when you get these services 2024 Evidence of Coverage for SCAN Balance Chapter 4 Medical Benefits Chart (what is covered and what you pay) 80 Inpatient services in a psychiatric hospital (continued) 2024 copayments will apply until you have not received any inpatient care in an acute hospital a skilled nursing facility or an inpatient mental health facility for 60 days in a row.



- On the top right corner, you'll see a **Teach** icon. When you click on it, you can see text boxes that let you associate key phrases with each other.



### Teach

Question goes here

Answer goes here

Add answer

Associate