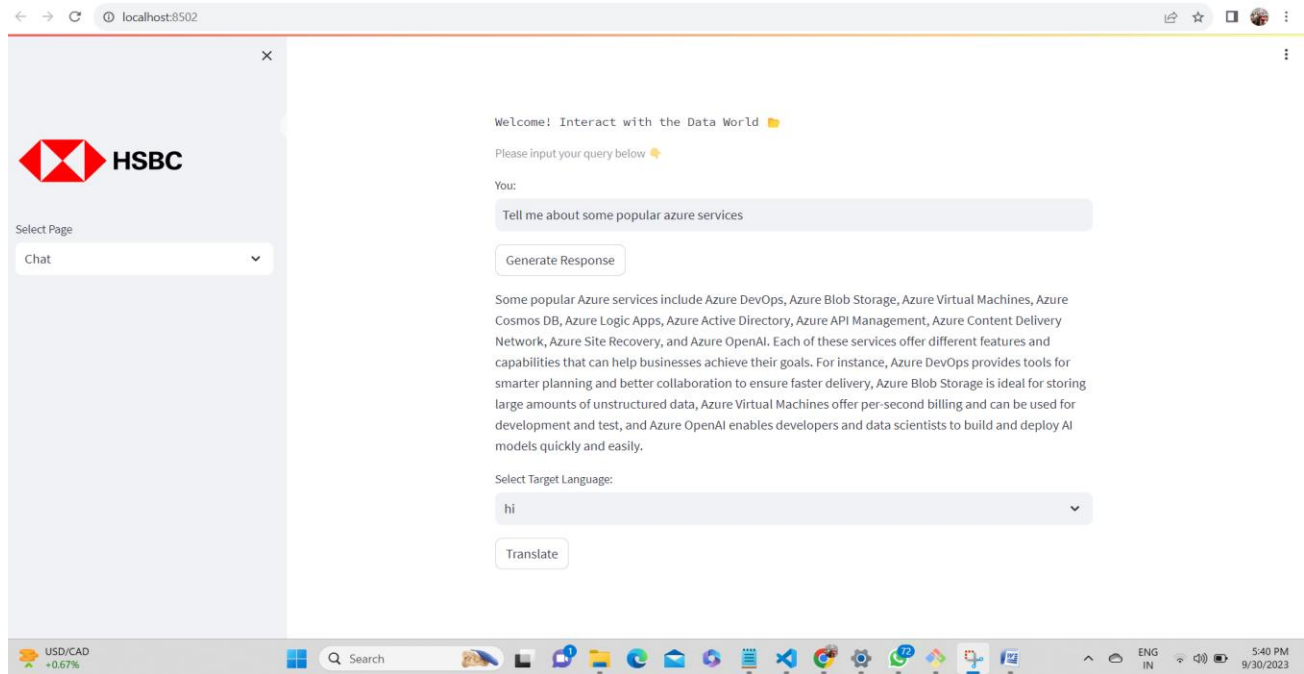


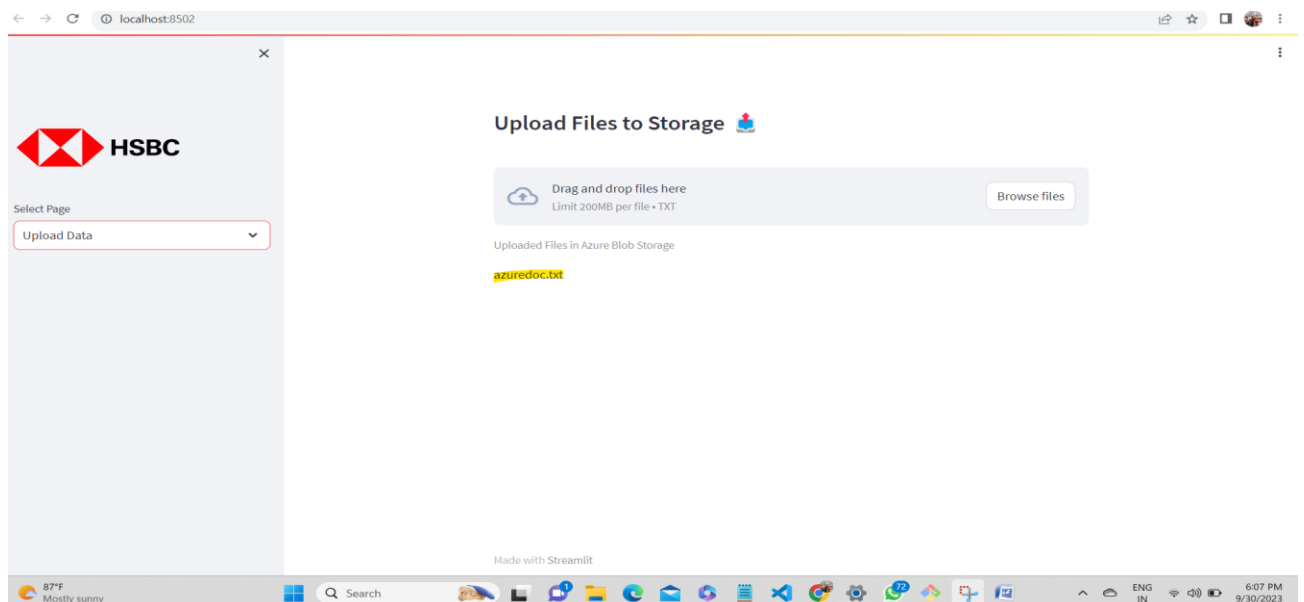
## Main Chat page



Earlier, Output was coming in single line. I have fixed it and now we have improved visibility of the output.

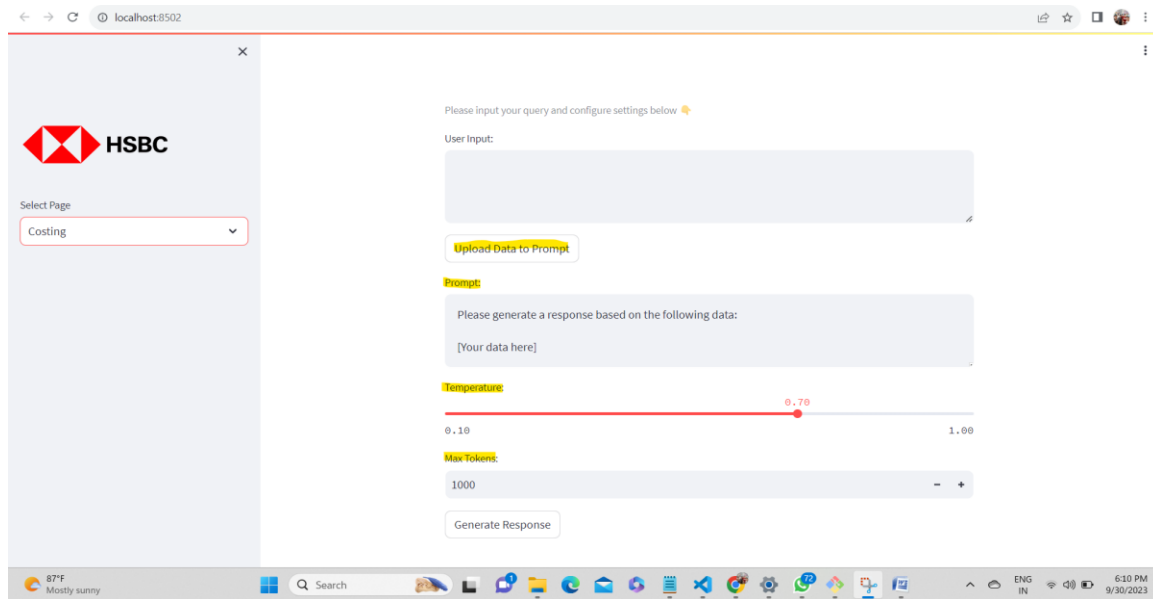
Logo moved to left corner.

## Uploading files page



List view of uploaded files in storage moved from sidebar to bottom.

## Costing Page



Costing page upgraded.

Features added:

1. Temperature
2. Tokens size
3. Prompt- We can edit this prompt as per will and it will respond appropriately
4. Added upload data to prompt button to put the data uploaded in storage as prompt if desired.
5. User can input query in user input
6. Generate button on costing page produces output and the cost/tokens info in a tabular format.
7. All prices are in USD.

## EDITED PROMPT

Please input your query and configure settings below 🗨️

User Input:

Who is Virat Kohli?

Upload Data to Prompt

Prompt:

You are a chatbot. Give adequate reply to user only if the question is about food. If irrelevant question is asked, just say Please ask something about food as this is a food bot.

Temperature:



Max Tokens:

1000

Generate Response

Please ask something about food as this is a food bot.

	Measure	Value
0	Estimated cost	0.001280
1	Total Tokens	64
2	Input Prompt Cost	0.001000
3	Prompt Tokens	50
4	Completion Cost	0.000280
5	Completion Tokens	14

Please note that all the price is in USD\$

## UPLOAD DATA TO PROMPT

On clicking this button, storage data gets added in prompt

Please input your query and configure settings below 🗨️

User Input:

Who is Virat Kohli?

Upload Data to Prompt

Prompt:

Introduction to Microsoft Azure Cloud Services  
It is a Cloud Computing service, provided by Microsoft. It connects data centers to the cloud effortlessly and supports 42 regions like no other cloud provider. Familiar tools such as Visual Studio, ASP.NET, and programming languages like Visual Basic, C++, C#, etc., are used to develop

Other features details:

Temperature (temperature):

Temperature is a parameter that affects the randomness of the generated text.

A higher temperature (e.g., 1.0) makes the output more random, while a lower temperature (e.g., 0.1) makes it more deterministic and focused.

Max Tokens (max\_tokens):

max\_tokens is used to limit the length of the generated text. It specifies the maximum number of tokens (words or subwords) that the model can generate in its response.