

LANGUAGE UNDERSTANDING IN NATURAL LANGUAGE PROCESSING USING MICRO SOFT AZURE

(Bridging the Gap between Humans and Machines)

Submitted by 99220040579
(K . Manoj Kumar Reddy)

A large orange circle is positioned at the top left of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom edge of the slide.

Table of contents

- Introduction
- About Software Tools
- Usage of Tools
- Reported Literature
- Objective Of Project
- Timeline of work proposal
- Used Algorithm
- Work done in step by step Description
- Result and Discussion
- Summary
- References

A large orange circle is positioned in the top-left corner of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom edge of the slide.

Introduction

- NLP stands for Natural Language Processing.
- It's a subfield of artificial intelligence (AI) that focuses on the interaction between computers and human language.
- Language is one of the primary ways humans communicate.
- It's rich in context, ambiguity, and nuance, making it challenging for machines to understand.
- NLP aims to bridge the gap between human language and computers.
- Its primary objectives include understanding, interpreting, and generating human language.
- NLP is pervasive in our daily lives.
- Examples of applications include virtual assistants (e.g., Siri, Alexa), search engines (e.g., Google), and chatbots.

A decorative orange circle is positioned at the top left of the slide, with a thin orange vertical line extending downwards from its center.

About software tool

- In this language understanding we used Microsoft Azure tool.
- Microsoft Azure is a cloud computing platform and service created by Microsoft. It provides a wide range of cloud-based services, including infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) offerings.
- Azure is designed to help organizations build, deploy, and manage applications and services through Microsoft-managed data centers worldwide.

A large orange circle is positioned at the top left of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom edge of the slide.

Usage of Tools

- Set up Azure Services – creating an Azure account
- Choose the Right Azure Service – We choose *Azure QnA Maker*
- Collect and Prepare Data – we choosen real time Sample Data
- Train the Model
- Integration with Chatbot
- Test and Iterate
- Deploy and Monitor

A large orange circle is positioned at the top left of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom of the slide.

Reported Literature

- A literature survey on Language Understanding :Vision, status, and research topics of Natural Language Processing by Xieling Chen a, Haoran Xie b, Xiaohui Tao c
- The Power of Natural Language Processing by Ross Gruetzemacher
- Advances in natural language processing JULIA HIRSCHBERG AND CHRISTOPHER D. MANNING

A decorative orange circle is positioned at the top left of the slide, with a vertical orange line extending downwards from its center.

Objective Of Project

- Enable the chatbot to comprehend the meaning and intent behind user questions, regardless of variations in language, syntax, or wording.
- Retrieve relevant information from a knowledge base or database to provide accurate answers to user queries.
- Ensure that the chatbot's responses are not only accurate but also relevant to the user's question
- Create engaging interactions that encourage users to interact with the chatbot and find value in the information it provides.

A large orange circle is positioned at the top left of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom edge of the slide.

Timeline of work proposal

- Week 1: Planning and Research & Data Collection and Preparation
- Week 2: Model Training
- Week 3: Integration and Development
- Week 4: Testing and Iteration
- Week 5: Deployment and Optimization

A large orange circle is positioned at the top left of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom edge of the slide.

Algorithm Used

- **Azure QnA Maker**

QnA Maker uses a combination of natural language processing techniques, including machine learning algorithms for language understanding and information retrieval. It employs techniques such as word embeddings, semantic similarity, and ranking algorithms to match user questions to the most relevant answers in the knowledge base.

Work done in step by step Description

- 1 Selecting Q&A model

The screenshot displays the Microsoft Azure Language Studio web interface. The browser's address bar shows the URL `language.cognitive.azure.com/home`. The page header includes the text "Cognitive Services | Language Studio" and a user profile for "neeraj@azure-training.com". The main content area features a "Welcome to Language Studio" message and a section titled "Recent custom projects you've worked on", which currently shows no projects. A "Create new" button is visible, and a dropdown menu is open, listing several options: "Conversational language understanding" (Build natural language into apps, bots, and IoT devices), "Orchestration workflow" (Connect and orchestrate CLU, Custom question answering & LUIS projects), "Custom question answering" (Customize the list of questions and answers extracted from your content...), "Custom text classification" (Train a classification model to classify text using your own data), and "Custom named entity recognition" (Train an extraction model to identify your domain categories using your o...). Below the dropdown, there are three featured cards: "Post call transcription and analytics" (Batch transcribe call center recordings and extract valuable information such as Personal Identifiable Information (PII), sentiment, and call summary), "Summarize information" (Summarize the most important or relevant information within documental and conversational text), and "Document translation" (Batch translate documents into one or more languages either from local storage or Azure Blob Storage). The bottom section of the page is titled "Learning resources" and includes links to "Read the documentation", "Explore our code samples", "Watch a video (coming soon)", and "Microsoft Learn (coming soon)". The Windows taskbar at the bottom shows the time as 11:06 on 04-05-2023.

• 2 Naming The Project

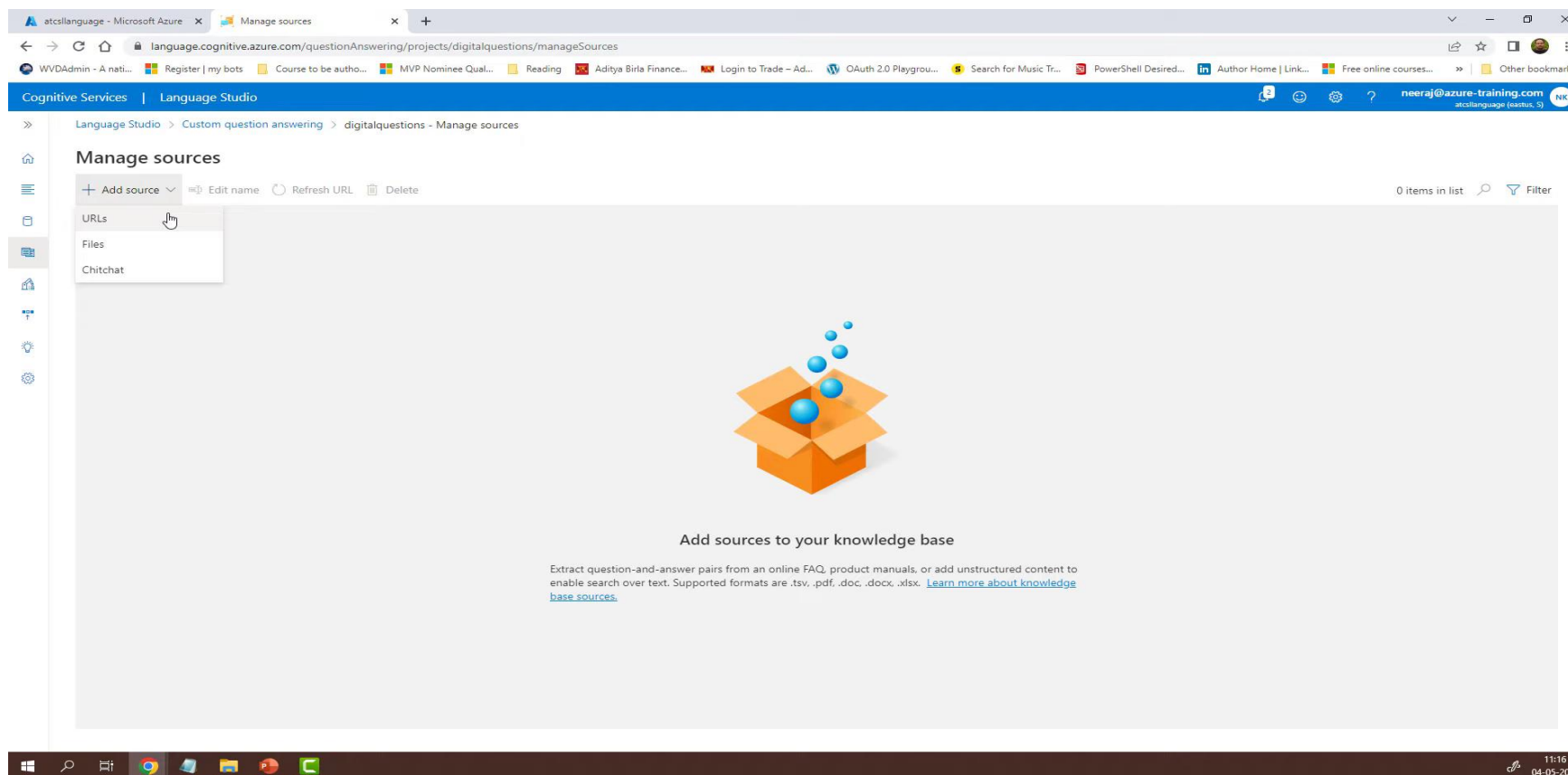
The screenshot shows the 'Create a project' dialog box in the Azure Language Studio interface. The dialog is titled 'Create a project' and has a close button (X) in the top right corner. It contains a progress indicator on the left with three steps: 'Choose language setting' (completed), 'Enter basic information' (current step), and 'Review and finish' (not started). The 'Enter basic information' section includes the following fields and options:

- Enter basic information**
Enter the basic information for your custom question answering knowledge base such as name and description.
- Azure search resource**
Dropdown menu showing 'atcsllanguage-asfcwnvcyhrik4c'. Below it, a link says 'To change your resource go to [Azure Search](#)'.
- Name ***
Text input field containing 'digitalquestions'.
- Description**
Text input field containing 'Type project description'.
- Source language ***
Dropdown menu showing 'English'.
- Default answer when no answer is returned ***
Text input field containing 'No answer found. Contact the administrator'.

At the bottom of the dialog, there are four buttons: 'Back', 'Next' (highlighted in blue), 'Create project' (disabled), and 'Cancel'.

The background shows the Azure Language Studio interface with the breadcrumb 'Language Studio > Custom question answering'. The top navigation bar includes 'Cognitive Services | Language Studio' and a user profile 'neeraj@azure-training.com'. The bottom status bar shows the time '11:08' and date '04-05-2023'.

• 3 Adding Our Dataset



The screenshot shows the 'Manage sources' page in the Azure Cognitive Services Language Studio. The browser address bar indicates the URL: `language.cognitive.azure.com/questionAnswering/projects/digitalquestions/manageSources`. The page header shows 'Cognitive Services | Language Studio' and a user profile 'neeraj@azure-training.com'. The main content area is titled 'Manage sources' and shows '0 items in list'. A dropdown menu is open under the '+ Add source' button, listing 'URLs', 'Files', and 'Chitchat'. The 'Files' option is highlighted. Below the dropdown, there is a large orange box icon with blue bubbles, and the text 'Add sources to your knowledge base'. A paragraph below this text explains: 'Extract question-and-answer pairs from an online FAQ, product manuals, or add unstructured content to enable search over text. Supported formats are .tsv, .pdf, .doc, .docx, .xlsx. [Learn more about knowledge base sources.](#)' The Windows taskbar at the bottom shows the time as 11:16 on 04-05-2023.

• 4 Deploying All our Datasets and Creating Our

The screenshot displays the Microsoft Azure portal interface. The browser's address bar shows the URL: `portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2Fa231df6b-78a5-4329-b255-c01e3762f224...`. The page title is "Microsoft.Template-20240225225854 | Overview". The main content area features a green checkmark icon and the text "Your deployment is complete". Below this, the deployment details are listed: "Deployment name : Microsoft.Template-20240225...", "Subscription : Azure for Students", and "Resource group : atcsindia". The start time is "25/02/2024, 22:59:04" and the correlation ID is "e60f3bb3-f761-4b10-870f-14...". A blue button labeled "Go to resource group" is visible. On the right side, there are three promotional cards: "Cost management" (Get notified to stay within your budget), "Microsoft Defender for Cloud" (Secure your apps and infrastructure), and "Free Microsoft tutorials" (Start learning today). The left sidebar contains a search bar and a list of navigation items: Overview, Inputs, Outputs, and Template. The bottom of the screen shows the Windows taskbar with various application icons and the system clock indicating 11:03 PM on 25-02-2024.

Microsoft Azure

Home >

Microsoft.Template-20240225225854 | Overview

Deployment

Search

Delete Cancel Redeploy Download Refresh

Overview

Inputs

Outputs

Template

✓ Your deployment is complete

Deployment name : Microsoft.Template-20240225... Start time : 25/02/2024, 22:59:04

Subscription : Azure for Students Correlation ID : e60f3bb3-f761-4b10-870f-14...

Resource group : atcsindia

> Deployment details

∨ Next steps

Go to resource group

Give feedback

Tell us about your experience with deployment

Cost management

Get notified to stay within your budget and prevent unexpected charges on your bill.

Set up cost alerts >

Microsoft Defender for Cloud

Secure your apps and infrastructure

Go to Microsoft Defender for Cloud >

Free Microsoft tutorials

Start learning today >

Tomorrow's high
Near record

Search

ENG IN

11:03 PM
25-02-2024

• 5 Configuration

The screenshot displays the Microsoft Azure portal interface. The top navigation bar includes the 'Microsoft Azure' logo, a search bar, and user information for '99220040579@klu.ac.in'. The breadcrumb trail indicates the path: Home > Microsoft.Template-20240225225854 | Overview > atcsindia > atcslangugae-bot.

The main content area is titled 'atcslangugae-bot | Configuration' and features a left-hand sidebar with navigation options: Overview, Activity log, Access control (IAM), Tags, Settings, Bot profile, Configuration (selected), Channels, Pricing, Test in Web Chat, Encryption, Networking, and Properties.

The configuration settings for the bot are as follows:

- Messaging endpoint:** `https://atcslangugae.azurewebsites.net/api/messages`
- Enable Streaming Endpoint:** ☐
- Bot Type:** UserAssignedMSI
- Microsoft App ID (Manage):** a2a79a82-e4d6-45b5-9281-b378d6d79899
- App Tenant ID:** c69dcaa2-96bf-460b-a8fe-880d700a09bb
- App MSI Resource ID:** /subscriptions/a231df6b-78a5-4329-b255-c01e3762f224/resourceGroups/atcsindia/providers/Microsoft.ManagedIde...
- Application Insights Instrumentation key:** atcslangugae

At the bottom of the configuration panel are buttons for 'Apply', 'Discard changes', and a 'youtube.com' link.

The Windows taskbar at the bottom shows the system clock at 11:05 PM on 25-02-2024, along with various application icons and system status indicators like temperature (27°C) and network connectivity.

• 6 Channels

Microsoft Azure

Search resources, services, and docs (G+)

Home > Microsoft.Template-20240225225854 | Overview > atcsindia > atcslangugae-bot

atcslangugae-bot | Channels

Azure Bot

Search

- Overview
- Activity log
- Access control (IAM)
- Tags
- Settings
 - Bot profile
 - Configuration
 - Channels**
 - Pricing
 - Test in Web Chat
 - Encryption
 - Networking
 - Properties

You are using the updated channels page. Let us know what you think by providing feedback [Feedback](#)

This bot is connected with the following channels.

Channel ↓	Health status	Details	Actions
Direct Line	Healthy	REST API for communicating directly with a bot	
Web Chat	Healthy	Embeddable Web Chat control	

Available Channels

Connect the bot with channels. [Learn more](#)

Channel ↓	Details
Alexa	Alexa Ch... Microsoft Store

27°C Haze

Search

11:06 PM 25-02-2024

- 7 We can Add more channels

The screenshot shows the Microsoft Azure portal interface. The top navigation bar includes the 'Microsoft Azure' logo and a search bar. The breadcrumb trail indicates the path: Home > Microsoft.Template-20240225225854 | Overview > atcsindia > atcslangugae-bot. The main heading is 'atcslangugae-bot | Channels'. On the left, a sidebar menu lists various settings and monitoring options, with 'Channels' highlighted. The main content area displays a list of channels with columns for the channel name and a description. The channels listed are: Facebook (Support for text messaging via Facebook), GroupMe (GroupMe Channel), LINE (Support for LINE Channel), Microsoft 365 (Enable message extensions in Outlook, and Microsoft 365 apps), Microsoft Teams (Microsoft Teams Channel), Omnichannel (Omnichannel Channel), Outlook (Outlook Channel), Skype (Skype Channel), Slack (Slack Channel), and Telegram (Telegram Channel).

Channel	Description
Facebook	Support for text messaging via Facebook
GroupMe	GroupMe Channel
LINE	Support for LINE Channel
Microsoft 365	Enable message extensions in Outlook, and Microsoft 365 apps
Microsoft Teams	Microsoft Teams Channel
Omnichannel	Omnichannel Channel
Outlook	Outlook Channel
Skype	Skype Channel
Slack	Slack Channel
Telegram	Telegram Channel

This screenshot shows the 'Channels' page for the 'atcslangugae-bot' in the Microsoft Azure portal. The breadcrumb trail is: Home > Microsoft.Template-20240225225854 | Overview > atcsindia > atcslangugae-bot. The main heading is 'atcslangugae-bot | Channels'. The page includes a search bar and links for 'Get bot embed codes', 'Refresh', and 'Feedback'. A notification banner states: 'You are using the updated channels page. Let us know what you think by providing feedback'. Below this, a message indicates: 'This bot is connected with the following channels.' A table lists the connected channels with columns for Channel, Health status, Details, and Actions. The channels listed are: Direct Line (Healthy, REST API for communicating directly with a bot), Microsoft Teams (Healthy, Microsoft Teams Channel, with a link to 'Open in Teams'), and Web Chat (Healthy, Embeddable Web Chat control). Below the table, there is a section for 'Available Channels' with a link to 'Learn more'.

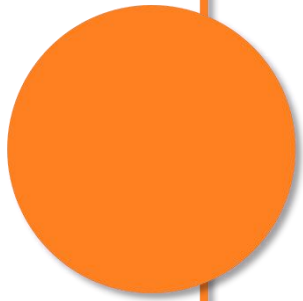
Channel	Health status	Details	Actions
Direct Line	Healthy	REST API for communicating directly with a bot	
Microsoft Teams	Healthy	Microsoft Teams Channel	Open in Teams
Web Chat	Healthy	Embeddable Web Chat control	

- 8 Interface out Bot

The screenshot displays the Microsoft Azure portal interface. The browser tabs at the top include 'Sign in to access KLU net', 'atcslangugae - Microsoft', 'Deploy knowledge base', and 'atcslangugae-bot - Micro'. The address bar shows the URL: portal.azure.com/#/@99220040579kluac.onmicrosoft.com/resource/subscriptions/a231df6b-78a5-4329-b255-c01e3762f224/resourceGroups/atc.... The Microsoft Azure header is visible with a search bar and user information: '99220040579@klu.ac.in' and 'DEFAULT DIRECTORY (992200405...)'. The breadcrumb trail reads: 'Home > Microsoft.Template-20240225225854 | Overview > atcsindia > atcslangugae-bot'. The main section is titled 'atcslangugae-bot | Test in Web Chat' with a star icon and a close button. On the left, a sidebar menu lists various options: 'Overview', 'Activity log', 'Access control (IAM)', 'Tags', 'Settings', 'Bot profile', 'Configuration', 'Channels', 'Pricing', 'Test in Web Chat' (highlighted), 'Encryption', 'Networking', and 'Properties'. The 'Test in Web Chat' area shows a chat interface with a 'Test' button and a 'Start over' link. A message bubble contains the text 'Hello and Welcome'. Below the message, it says 'Just now'. At the bottom of the chat area is a text input field with the placeholder 'Type your message' and a 'WPS Office' watermark. The Windows taskbar at the bottom shows the system clock as 11:13 PM on 25-02-2024, along with various application icons and system status indicators like temperature (27°C Haze) and network connectivity.

Result and Discussion

The screenshot displays the Microsoft Azure portal interface for managing the 'atcslangugae-bot'. The browser window shows the URL 'portal.azure.com/#@99220040579kluac.onmicrosoft.com/resource/subscriptions/a231df6b-78a5-4329-b255-c01e3762f224/resourceGroups/atc...'. The page title is 'atcslangugae-bot | Test in Web Chat'. The left sidebar lists various management options, with 'Test in Web Chat' currently selected. The main area shows a chat interface with a 'Hello and Welcome' message and a 'Type your message' input field. The bottom of the screen shows the Windows taskbar with the date 25-02-2024 and time 11:13 PM.



Sign in to access KLU net | atclangugae - Microsoft | Deploy knowledge base | atclangugae-bot - Micro | New Tab

portal.azure.com/#@99220040579kluac.onmicrosoft.com/resource/subscriptions/a231df6b-78a5-4329-b255-c01e3762f224/resourceGroups/atc...

Microsoft Azure | Search resources, services, and docs (G+/)

99220040579@klu.ac.in | DEFAULT DIRECTORY (992200405...

Home > Microsoft.Template-20240225225854 | Overview > atcsindia > atclangugae-bot

atclangugae-bot | Test in Web Chat

Azure Bot

Search

Overview

Activity log

Access control (IAM)

Tags

Settings

Bot profile

Configuration

Channels

Pricing

Test in Web Chat

Encryption

Networking

Properties

Test

Start over

How many long was Lincoln's formal education?

How many long was Lincoln's formal education?

None of the above.

Just now

How many long was Lincoln's formal education?

Just now

18 months

Just now

Type your message

27°C Haze

Search

ENG IN

11:15 PM 25-02-2024

A large orange circle is positioned in the top-left corner of the slide. A thin orange vertical line extends from the bottom of the circle down to the bottom edge of the slide.

Summary

- The project aims to develop a question-answering chatbot leveraging Azure's AI services for natural language processing (NLP). The chatbot will be capable of understanding user queries and providing accurate responses by extracting relevant information from a knowledge base or database.



References

- A. Kalyanpur, S. Patwardhan, B. K. Boguraev, A. Lally, J. Chu-Carroll, 2012. Fact-based question decomposition in DeepQA. IBM J. RES. & DEV. VOL. 56 NO. 3/4, pp. 1-11.
- A. Lally, J. M. Prager, M. C. McCord, B. K. Boguraev, S. Patwardhan, J. Fan, P. Fodor, J. Chu-Carroll., 2012. Question analysis: How Watson Reads a clue. IBM J. RES. & DEV. VOL. 56 NO. 3/4, pp. 1-14.
- B. L. Lewis, 2012. In the game: The interface between Watson and Jeopardy!. IBM J. RES. & DEV. VOL. 56 NO. 3/4, pp. 1-6.
- Benjamin S. Bloom, 1984. Taxonomy of educational objectives. s.l., s.n.C. Wang, A. Kalyanpur, J. Fan, B. K. Boguraev, D. C. Gondek., 2012. Relation extraction and scoring in DeepQA. IBM J. RES. & DEV. VOL. 56 NO. 3/4, pp. 1-12.