

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/325607065>

# Implementation of a Chat Bot System using AI and NLP

Research · May 2018

DOI: 10.21276/ijircst

CITATION

1

READS

7,498

5 authors, including:



**Shashank Bhalotia**

Cynapto Technologies

3 PUBLICATIONS 12 CITATIONS

[SEE PROFILE](#)



**Shreya Bisen**

Shri Ramdeobaba Kamla Nehru Engineering College

2 PUBLICATIONS 12 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Content Security Management System [View project](#)



College Information Chatbot System [View project](#)

# Implementation of a Chat Bot System using AI and NLP

Tarun Lalwani, Shashank Bhalotia, Ashish Pal, Shreya Bisen, Vasundhara Rathod

**Abstract—** For using software applications, user interfaces that can be used includes command line, graphical user interface (GUI), menu driven, form-based, natural language, etc. The mainstream user interfaces include GUI and web-based, but occasionally the need for an alternative user interface arises. A chat bot based conversational user interface fits into this space. The chat bot is a class of bots that have existed in the chat platforms. The user can interact with them via graphical interfaces or widgets, and the trend is in this direction. They generally provide a stateful service i.e. the application saves data of each session. On a college's website, one often doesn't know where to search for some kind of information. It becomes difficult to extract information for a person who is not a student or employee there. The solution to these comes up with a college inquiry chat bot, a fast, standard and informative widget to enhance college website's user experience and provide effective information to the user. Chat bots are an intelligent system being developed using artificial intelligence (AI) and natural language processing (NLP) algorithms. It has an effective user interface and answers the queries related to examination cell, admission, academics, users' attendance and grade point average, placement cell and other miscellaneous activities.

**Keywords:-** Artificial intelligence, chat bot, knowledge base, lemmatization, natural language processing, semantic sentence similarity, wordnet.

## I. INTRODUCTION

Nowadays, we see the chat bots everywhere Chat bots are the source of answers to the users questions in any particular domain where it is operating. Chat bots are the source of answers to the users questions in any particular

domain where it is operating. The most popular example today is the Amazon's Alexa. Chat bots are at almost every place, one can see it at every second website they visit. A bot is helpful in answering queries related to information which might be unreachable at that website easily. Most of the websites avail users with chat bots to aid them to go through what the websites facilitate. They are turning out to be our virtual assistants in everyday lives.

### A. Basics of chat bot

A chat bot is an artificially intelligent creature which can converse with humans. This could be text-based, or a spoken conversation (in case of voice-based queries). Chat bots are basically used for information acquisition. It can run on the local PCs and mobile phones, though most of the time it is accessed through the internet. It can be compelling, captivating and spell-binding. It is a conversational agent which interacts with users in a certain domain or on a particular topic with input in natural language sentences. Mainly a chat bot works by a user asking some question or initiating a new topic of discussion. Chat bots can be referred as software agents that pretend as human entity. These are the agents with AI embedded and using NLP they can answer to user questions. Predefined knowledge base helps develop a response to the query.

### B. Chat bot for College

The need for college inquiry system arises due to various reasons which include: the slow nature of college website, an outsider would not know where to search for a particular piece of information, difficult for the person outside college's domain to extract information. The smart solution for all the drawbacks leads to the need of the system. The college inquiry system will provide the response by summarizing the query and then output answers, it also provides selective information what the user wants. A college system will dispense all answers relating to domains such as admission, examination cell, notice board, attendance, placement cell and other miscellaneous domains.

The major features of the chat bot are:

- College admission related queries could be answered through it.
- Viewing user profiles and retrieves attendance and grade/ pointers.
- College students can get information about examinations to be held.
- College students can fetch particulars about placement activities.

**Manuscript Received April 25, 2018.**

**Tarun Lalwani, Shashank Bhalotia, Ashish Pal, Shreya Bisen**, Student, Dept. of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur, Maharashtra, India

**Vasundhara Rathod**, Assistant Professor, Dept. of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur Maharashtra, India Phone(91)-(712)-2580011

## Implementation of a Chat Bot System using AI and NLP

College inquiry system will act as a fast, standard and informative widget to enhance college website's user experience and bestow users with righteous information. The bot will analyze users queries and understand users' message and then reply accordingly. It uses AI & NLP. This way users' time and efforts will be saved and s/he will be equipped with effective answers.

The objectives of this application are:

- To analyze users queries and understand users message.
- To provide an answer to the query of the user very effectively.
- To save the time of the user since s/he does not have to personally go to the college for inquiry.
- This system will help the student to be updated about the college activities.
- The system will reply using an effective GUI which implies that as if a real person is talking to the user.

## II. PROPOSED SYSTEM

### A. Context Identification:

Pre-processing is applied to the input text to standardize the input as per the system's requirement. Based on the keywords used in the text, appropriate context is recognized.

### B. Personal Query Response System (Module-1):

Upon receiving personal queries like CGPA, attendance, etc., the authenticity of the user is checked through user-id and password. If the user detail is invalid, an appropriate response is sent.

If the user authenticates successfully, the input text is processed to extract keywords. Based on the keywords, information required by the user is understood and the information is provided from the database.

### C. AIML Response System (Module-2):

If the user is trying to make a normal conversation with the bot, the input is mapped to an appropriate pattern in Artificial Intelligence Modeling Language (AIML) files. If the response is available, it is sent to the user. Other data provided to the chat bot such as username, gender, etc. are also saved. If the pattern is not available in AIML files, a random response is sent suggesting "Invalid Input".

### D. Query Analysis and Response System (Module-3):

When a user wants some information pertaining to college, the response will be provided through this module.

If the input matches a pattern in the AIML files, the appropriate response will be sent to the user. If the AIML files have no entry for that particular query pattern, keywords are fetched from the input.

An algorithm to check sentence similarity (NLP) is applied to the modified input to check its similarity with the questions of a predefined question-set, whose answers are available.

If a sentence is retrieved with confidence > 0.5, we return the answer of that question as the response.

If no questions map to the user input, the input is saved in a log file for improvement of the system by the admin. The administrator can incorporate the answer to that query

in the knowledge base if s/he finds it convenient. Also, a random response is sent to the user suggesting "Answer not available".

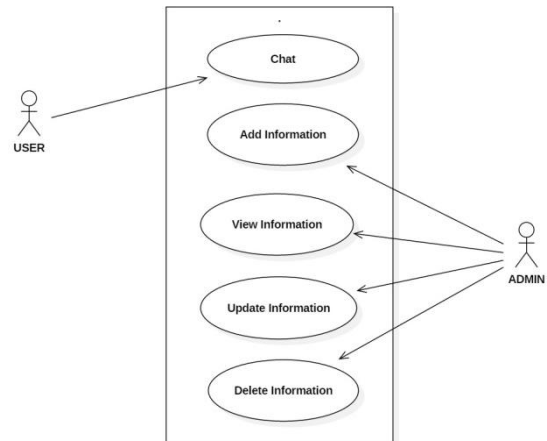
### E. Context Reset:

Once the user is satisfied with the response of bot and does not wish to chat further, he/she has the option to log out of the system or simply exit. Once the user exits the system, all input parameters are automatically reset.

## III. DESIGN

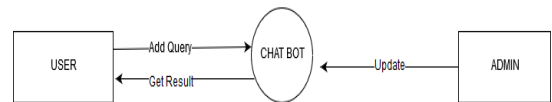
### Use Case Diagram:

(User and admin roles)

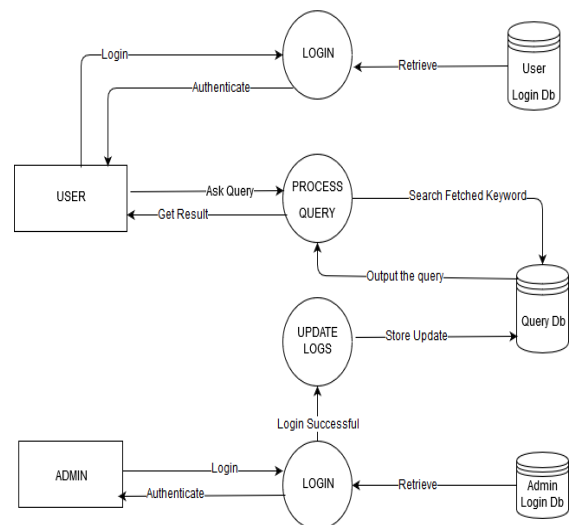


### Data Flow Diagram:

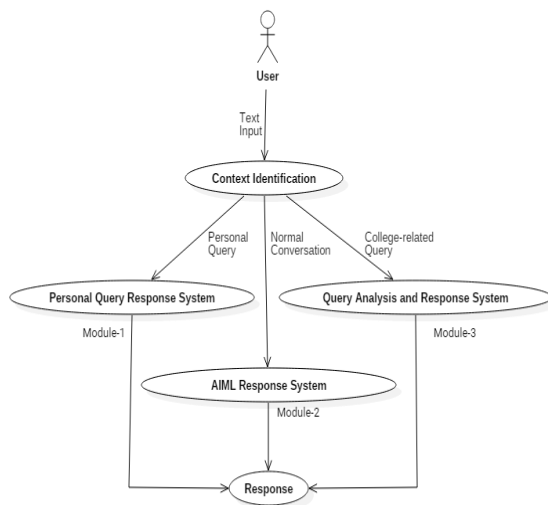
Level 0:



Level 1:

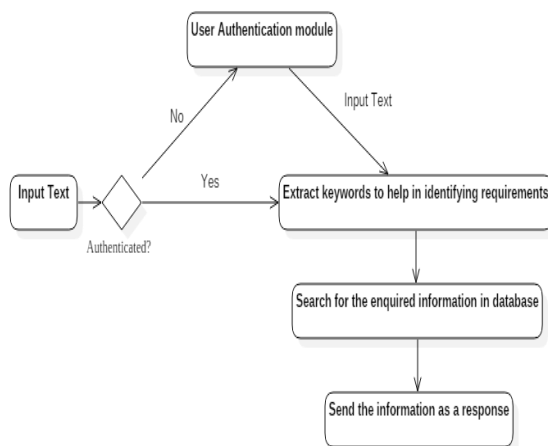


### Use Case Diagram: (Context identification)

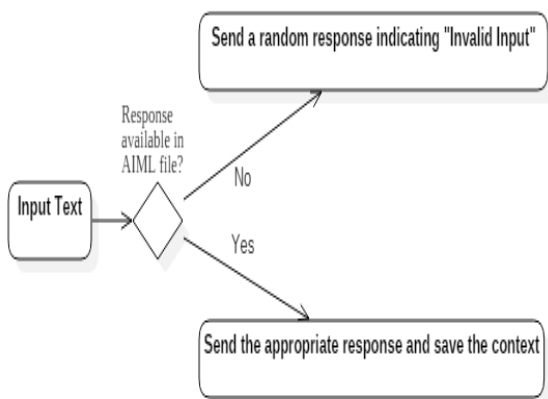


### Activity Diagram:

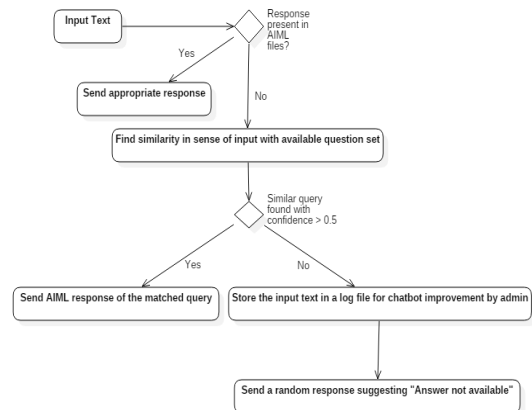
#### Personal Query Response Activity (Module-1):



#### Normal Conversation Response Activity (Module-2):



#### College Related Query Response Activity (Module-3):



## IV. IMPLEMENTATION

### A. AIML:

To create our knowledge base for normal conversation, we have used AIML files to store the question and answers pair. When user converses with our chat bot, the input is matched to patterns listed in AIML files and corresponding answer is returned as response.

The sample AIML file structure is as:

```

<aiml version = "1.0.1" encoding = "UTF-8"?>
<category>
  <pattern> HELLO USERNAME</pattern>
  <template> Hello User!</template>
</category>
</aiml>

```

### B. Lemmatization and POS Tagging Using WordNet:

Information extraction from the input text was done by extracting keywords. For example, "What is the current placement scenario?" contain "current", "placement" and "scenario" as the keywords. Appropriate Lemmas of the keywords were found using Lemmatization and POS tagging, to group together the different inflected form of the words. For example, requiring, require and required should map to require. WordNet from Python's "nltk" package was used for this purpose.

### C. Semantic Sentence Similarity:

There are various combinations in which user can input the same query. For example,

Q1: What is the notice regarding PG courses re-registration?

Q2: Tell me about re-registration in PG courses in our college.

Q1 and Q2 both mean the same thing (same sense). Also, there will be many more combinations for this same query and finding all such combinations will not be feasible. The scalability and performance of the system will also get affected. To overcome this problem, similarity is found out between the user input and the queries present in the available question set (whose answers are available with the system). The query which matches with the input with maximum score

gets selected (if greater than threshold) and the appropriate response is returned.

Similarity score for two sentences is calculated by averaging the similarity of the individual keywords of those sentences. Each keyword of the first sentence is matched with every keyword of the second sentence to find the word with maximum similarity to it. Then the similarity score of individual words are averaged to represent the sentence similarity.

To find the word similarity- Path Similarity and Wu-Palmer (WUP) Similarity is used.

Path similarity computes shortest number of edges from one word sense to another word sense, assuming a hierarchical structure like WordNet. In general, word senses which have a longer path distance are less similar than those with a very short path distance, e.g. man, dog versus man, tree (expectation is that man is more similar to dog than it is to tree). The Wu-Palmer metric weights the edges based on distance in the hierarchy.

### D. Log File:

We have maintained a log file which stores the inputs which the chat bot was not able to answer. Admin can see the log and add the response of relevant sentences to the knowledge base. This would help in improvement of the chat bot knowledge system i.e. the databases.

## V. APPLICATIONS

- It enables the students to be updated with college activities.
- It saves time for the students as well as teaching and non-teaching staffs.
- It is providing us a readily available information source without taking any physical efforts.
- It is easily accessible and saving time and money also.

## VI. CONCLUSIONS

It is often impossible to get all the data on a single interface without the complications of going through multiple forms and windows. The college chat bot aims to remove this difficulty by providing a common and user-friendly interface to solve queries of college students and teachers.

The purpose of a chat bot system is to simulate a human conversation. Its architecture integrates a language model and computational algorithm to emulate information online communication between a human and a computer using natural language.

The college student and employees can freely upload their queries. The chat bot provides fast and efficient search for answers to the queries and gets the relevant links to their question. A background research took place, which included an overview of the conversation procedure and tries to find out the relevant keywords related to that query to provide the proper link. The database storage includes information about questions, answers, keywords, and logs. We have also developed an interface. The interface developed will have two parts, one for users and the other for the administrator.

## VII. FUTURE ENHANCEMENTS

Instead of AIML based bot, other algorithms can be implemented. We can include voice-based queries. The users will have to give voice input and the system will give the text output. Also, after successful execution of chat bot in college domain, we can implement it in other domains like medical, forensic, sports, etc. It will be beneficial in all the fields as without spending much time, we are accessing the relevant information and that too without any sorting.

## VIII. ACKNOWLEDGMENT

We would like to extend our sincere thanks to Prof. Vasundhara Rathod for giving us good guidelines throughout numerous consultations.

## REFERENCES

- [1] Yuhua Li, David McLean, Zuhair A. Bandar, James D. O'Shea, Keeley Crockett, "Sentence Similarity Based on Semantic Nets and Corpus Statistics", IEEE Transactions on Knowledge and Data Engineering, Volume 18 - No. 8, August 2006.
- [2] Emanuela Haller, Traian Rebedea, "Designing a Chat-bot that Simulates an Historical Figure", IEEE Conference Publications, July 2013.
- [3] Pratik Slave, Vishruta Patil, Vyankatesh Gaikwad, Girish Wadhwa, "College Enquiry Chat Bot", International Journal on Recent and Innovation Trends in Computing and Communication, Volume 5, Issue 3, March 2015.
- [4] "AIML Based Voice Enabled Artificial Intelligent Chatterbot", International Journal of u- and e- Service, Science and Technology Volume 8 - No. 2, 2015.
- [5] Amey Tiwari, Rahul Talekar, Prof. S. M. Patil, "College Information Chat Bot System", International Journal of Engineering Research and General Science, Volume 2, Issue 2, April 2017.
- [6] Rachit Kulkarni, Ankit Methwani, Nakul Pawar, Charmi Valecha, Pooja Shetty, "College Chat-bot", International Journal of Advanced Research in Computer Engineering & Technology, Volume 6, Issue 4, April 2017.
- [7] Chaitrali S. Kulkarni, Amruta U. Bhavsar, Savita R. Pingale, Prof. Satish S. Kumbhar, "BANK CHAT BOT - An Intelligent Assistant System Using NLP and Machine Learning", International Research Journal of Engineering and Technology, Volume 4, Issue 5, May 2017.
- [8] Yash Mehta, Shreya Sawkar, "The college chatbot", International Journal of Computer Applications, Volume 173 - No. 7, September 2017.
- [9] Prof. K. Bala, Mukesh Kumar, Sayali Hulawale, Sahil Pandita, "Chat-Bot For College Management System Using A.I", International Research Journal of Engineering and Technology, Volume 4, Issue 11, Nov 2017.



### Author's Profile



Tarun Lalwani is currently in his last semester pursuing the degree of Bachelor of Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur - 440 013 (M.S.) (India). He is a student at Department of Computer Science and Engineering. His academic projects include: AVOID Q : Quick Billing System, an innovative approach to minimize problems faced by people in billing section; Real Time Intrusion Detection System using Machine Learning, etc.



Shashank Bhalotia is currently in his last semester pursuing the degree of Bachelor of Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur - 440 013 (M.S.) (India). He is a student at Department of Computer Science and Engineering. His academic projects include: Content Security Management System (to minimize piracy of paid content); Gestured Computer Controller (controlling Computer actions through hand gestures) and Gestured Calculator (a calculator taking all inputs through fingers and providing output in text and speech). He was also a part of an open source project called "XCOS-ON-WEB" during his internship (Summer 2017) at IIT Bombay. He is a member of the Computer Society of India (CSI). His research paper publication include a research paper titled: "A novel Approach to Compute Steiner point in Graph: Application for Network Design" (IJCSE).



Ashish Pal is currently in his last semester pursuing the degree of Bachelor of Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur - 440 013 (M.S.) (India). He is a student at Department of Computer Science and Engineering. His academic projects include: Media content Security System; Dynamic/Static Hand Gesture Recognition System.



Shreya Bisen is currently in her last semester pursuing the degree of Bachelor of Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur - 440 013 (M.S.) (India). She is a student at Department of Computer Science and Engineering. His academic projects include: Customer Relationship Management(CRM) Software System, as an industry project; Enhance security for Cloud storage using hybrid Cryptography (a way to secure files on clud storage using encryption algorithms); Emergency Response Application, which which help users through an androd app to locate nearby help centers. She is also a member of the Computer Society of India (CSI).



Vasundhara Rtahod is currently the Assistant Professor at Department of Computer Science and Engineering at Shri Ramdeobaba College of Engineering and Management, Nagpur, Ramdeo Tekdi, Gittikhadan, Katol Road, Nagpur - 440 013 (M.S.) (India). She has 3 years of teaching experience as of now and 1 year industry experience. 3 international journal research papers have been published in her account. The subjects of her specialization include: Object Oriented Technologies, Computer Networks, Web Application Development.