

# Database Access for Non-technical Users using NLP

<sup>#1</sup>Kritika Chaurasia, <sup>#2</sup>Mrunali Golait, <sup>#3</sup>Diksha Parashar, <sup>#4</sup>Komal Yewale,  
<sup>#5</sup>Prof.Priyanka Kedar



<sup>#1234</sup>UG Student, Department of Computer Engineering  
<sup>#5</sup>Assistant Professor, Department of Computer Engineering

Dhole Patil College of Engineering, Pune, India

## ABSTRACT

The field of “Natural Language Processing” (NLP) has seen a dramatic shift in both research direction and methodology in the past several years. In the past, most work in computational linguistics tended to focus on purely symbolic methods. Recently, more and more work is shifting towards hybrid methods that combine new empirical corpus-based methods, including the use of probabilistic and information theoretic techniques, with traditional symbolic methods. The main purpose of Natural Language Query Processing is for an input sentence(s) to be interpreted by the computer and appropriate action taken; asking questions to databases in natural language is very convenient and easy method of data access, especially for casual users who do not understand complicated databases such as SQL. This project proposes the architecture for translating input sentence(s) into SQL query using Semantic Grammar with added feature of Voice Recognition. This system has been developed to enable secure access of data to a voice-based user interface (UI) by enabling voice-based authentication and integration with an existing Natural Language Processing (NLP) system. Gaining secure access to existing NLP systems also served as motivation. In this system we have described the implementation of a personal virtual assistant which can take the human voice commands to perform tasks which otherwise would need the dependence on others.

**Keywords:** NLP, Languages and compilers, Optimization, Verification, Voice Recognition, Machine-independent microcode generation.

## ARTICLE INFO

### Article History

Received: 12<sup>th</sup> November 2019

Received in revised form :

12<sup>th</sup> November 2019

Accepted: 16<sup>th</sup> November 2019

**Published online :**

**16<sup>th</sup> November 2019**

## I. INTRODUCTION

Natural Language Processing (NLP) is an area of application and research that explores how computers can be used to understand and manipulate natural language speech or text to do useful things. The foundation of NLP lie in a number of disciplines, namely, computer and information sciences, linguistics, mathematics, electrical and electronic engineering, artificial intelligence robotics, and psychology. NLP researchers aim to gather knowledge on how human beings use and manipulate natural languages to perform desired tasks so that appropriate tools and techniques can be developed. Applications of NLP include a number of fields of study such as

multilingual and cross-language information retrieval (CLIR), machine translation, natural language, text processing and summarization, user interfaces, speech recognition, artificial intelligence and expert systems.

While natural language may be the easiest system for people to learn and use, it has proved to be the hardest for a computer to understand. The goal of NLP is to enable communication between people and computers without resorting to memorization of complex commands and procedures.

In other words, NLP is a technique, which can make the computer understand the languages naturally used by humans. In this project, we are translating

English query into a SQL query using semantic grammar. The system will accept users query in natural language as an input. The program will check whether the query is valid or not.

Then we will generate tokens by performing the division of the question clause. Each token represents a single word in the users query. The tokens from the query clause are compared with clauses already stored in the dictionary. The dictionary needs to be constantly updated. Then the algorithm scans the tokens and tries to find attributes present in the query. Then we find all the tables in the database which contain the attributes by comparing syntax and semantics. Then we build the final SQL query and execute it on the database and return the result dataset to the user.

### **Problem Statement:**

Most of used don't know about the data access, data searching, data query so, user have better knowledge that area. A voice-controlled system for non technical, which transceivers information in the form of audio: a personal virtual assistant which can take the human voice commands to perform tasks which otherwise would need the dependence on others. Most of used don't know about the data access, data searching, data query so, user have better knowledge that area.

## **II. LITERATURE SURVEY**

Natural language processing can be done in two-way communication with device one is written communication as well as verbal communication with device written communication is much easier than the verbal communication. In written communication syntax, semantic, lexical and morphological analysis is done. Whereas in verbal communication includes all the process in written as well as additional process include additional knowledge about phonology as well as enough added information to handle the further ambiguities that arise in speech [1].

This paper places an interest in some emerging capabilities for incremental speech understanding and processing in virtual human dialogue systems. This work is part of an in-progress effort that aims to enable practical spoken dialogue with virtual humans in multiparty arbitration scenarios. These scenarios are designed to allow trainees to practice their intervention

skills by engaging in face-to-face spoken negotiation with one or more virtual humans. An important factor in achieving naturalistic behavior in these arbitration scenarios, which ideally should have the virtual humans representing fluid turn-taking, composite reasoning, and responding to factors like trust and emotions, is for the virtual humans to begin to understand [2].

The current custom in virtual human dialogue systems is to use skilled human recordings or limited-domain speech synthesis. Both approaches lead to good show but at an elevated cost. To determine the best trade-off between performance and cost, we perform an evaluation of a human and synthesize voices with respect to naturalness, conversational aspect, and likability. Varying the type, length, and content of utterances, and take into account the age and native language of ratters as well as their expertise with speech synthesis. The results suggest that a professional human voice can surpass both an amateur human voice and synthesized voices. Also, a high-quality general-purpose voice or a good limited-domain voice can execute better than part-time human recordings. As expected, in most cases, the high-quality general-purpose voice is rated higher than the limited-domain voice. There is also a non-statistically significant trend that has been observed for long or negative utterances to receive lower ratings [3].

The aim of this paper is to explore business applications of chat bots, as well as to propose several extent metrics to evaluate practice, usability and overall quality of an embodied conversational agent. On the basis of these metrics we examine existing Polish-speaking commercial chat bots that, firstly, work in the B2C subdivision. Secondly, reach the widest possible range of users. And lastly, are most probably the most advanced commercial deployments of their creators. The system analyses various aspects of functioning of each personified conversational agent: optical look, form of operation on the website, speech amalgamation unit, built-in knowledge base, presentation of knowledge and supplementary functionalities, conversational abilities and perspective sensitiveness, personality traits, personalization options, emergency responses in unforeseen situations, possibility of rating chatbot and the website by the user [4]

### III. PROPOSED SYSTEM

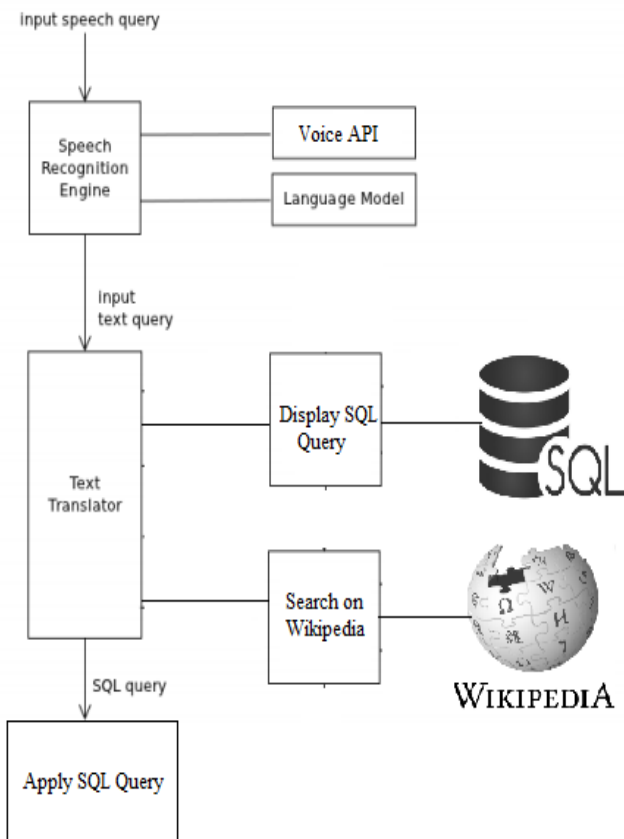


Fig 1. System architecture

#### System Feature 1

The system will help non-technical people in learning sql queries using voice and text.

#### System Feature 2

User can fire their own queries for demo database using voice or text.

#### System Feature 3

User can also search their query in Wikipedia.

### IV. CONCLUSION

Use of Natural Language brings ease for any human being. This system helps user to easily retrieve data from database using simple English language. The user need not learn complex query language like SQL. We can add more synonyms for column names and table names so that system is able to handle more queries. The system also stores the successfully executed queries based on voice generation. This system

provides some recommendations so that it is helpful for user. In future we can add some strong recommendation framework in this system so that user will have to take fewer efforts.

### REFERENCES

- [1] F.Siasar djahantighi, M.Norouzifard, S.H.Davarpanah, M.H.Shenassa, "USING NATURAL LANGUAGE PROCESSING IN ORDER TO CREATE SQL QUERIES", in Proceedings of the International Conference on Computer and Communication Engineering, Kuala Lumpur, May 2008, pp. 600-604.
- [2] NATURAL LANGUAGE PROCESSING USING PYTHON International Journal of Scientific Engineering Research Volume 8, Issue 5, May-2017 19 ISSN 2229-5518[IEEE]
- [3] Natural Language Processing Techniques Applied in Information Retrieval-Analysis and Implementation in Python, TulikaNarang, International Journal of Innovations Advance- ment in Computer Science IJIACS ISSN 2347 – 8616 Volume 5, Issue 4 April 2016
- [4] Anuradha Mohite, Varunakshi Bhojane, "Natural Language Interface to Database Using Modified Co-occurrence Matrix Technique 2015 International Conference on Pervasive Computing (ICPC)
- [5] Levin E., Pieraccini R., Eckert W. "Learning dialogue strategies within the Markov decision process framework", Automatic Speech Recognition and Understanding IEEE Proceedings, pp. 72-79, 1997.[IEEE]
- [6] Liddy, E. D. In Encyclopedia of Library and Information Science, 2nd Ed. Marcel Decker, Inc.
- [7] J. Allan. Perspectives on information retrieval and speech. In Information Retrieval Techniques for Speech Applications: LNCS 2273, pages 1–10, 2002.