International Journal of Oceans and Oceanography
ISSN 0973-2667 Volume 13, Number 1 (2019), pp. 251-257

http://www.ripublication.com

# Natural Language Processing System for Fetching Ocean Related Information Based on Ontology

**K.N.Sivabalan, S.Balakrishnan**

*Professor*
*Sri Krishna college of Engineering and Technology, Coimbatore, Tamilnadu, India.*

**Abstract**

Acquisition of knowledge has made a major contribution in research and development nowadays. The earth contains 29% of land mass and remaining 71% occupied by water. Since gathering information about the hydrosphere has always remained as a herculian task since exploring the water resources always remained beyond the reach of human being. This work focus on tapping valuable information about the ocean using natural language processing by deconstructing the natural language statements based on ontology. The system process the natural text based statements which has been entered in the BOT and classifies, groups it into different segments and then a SQL query is constructed to retrieve the information from the database which contains ocean related information. The system uses natural language processing toolkit (NLTK) of python and MYSQL Server for storing the data.

**Keywords:** Ontology, natural language processing, oceanography

## 1. INTRODUCTION

Getting Profound knowledge about unexplored domains always remains a challenging task. Since majority of the planet is covered by oceans still getting valuable information or knowledge continues to remain as daunting task. This research work focus on developing a natural language processing system which accepts users natural questions with stipulated keywords process it into database query and gives the intended result about the oceans to the user. The questions may focused such as what is the salinity of sea water, which is the deepest point in pacific ocean, which places are good for fishing in Atlantic ocean. All this questions can be entered in any pattern or form in the technical BOT. R.Mitra et.al suggested in his work that Artificial Intelligence algorithms has the capability to expose the micro organisms in within marine organisms at species level. This invention facilitates to incorporate the algorithm in robots and identify the various species in ocean water. Spencer chin of

North Carolina university has developed a efficient Artificial Algorithm to classify the micro organisms in sea water based on clear aquatic images. Hue Li et.al highlighted the various applications of artificial intelligence in ocean development. Artificial intelligence has given a new dimension and boost to overcome the traditional research methodologies and the way how inferences are drawn. The artificial intelligence focus on three major areas in ocean based research. The establishments of specialized idea piece, organize extension and keen route are laid in Maritime Internet of Things. Furthermore, in the field of sweeping ocean investigation, unmanned test (remote ocean robot) depended on man-made reasoning innovation has step by step turned into a noteworthy power in universal challenge; thirdly, as a sea data and the executives far reaching coordination stage, advanced sea is connected to split sea data singular island, confused data and other troublesome issues through collaboration improvement, work joining and partiality administrations, upgrading the open sea cognizance. Tuneer Mukherji spilled beans about protecting the maritime resources with fortifications by Artificial intelligence in naval operations. The use of artificial intelligence in safeguarding the maritime resources takes a pivotal role rather than other domains because of the ocean sheer size, hostility and unpredictable circumstances. This also reveals where ever this system exists it has made manned operations more effective. Dr.Timm Schoening of GEOMAR research institute concentrated on using Artificial intelligence in understanding deep sea images. Using this technology diving robots and deep exploration devices in ocean can inherit the capabilities of capturing high resolution images which gives exposure to under water resources to mankind. This paved to study about the aquatic ecosystem around the manganese modules in pacific ocean. Tim Sandle concentrated on exploring the aquatic resources using high resolution images by interpreting it and develop a new workflow for image analysis. The challenge is vested on researchers how to quantify huge data. The Art of artificial intelligence is to decipher marine images. Micah J Dean et al built up a basic way to deal with separate between the sympatric sub-populaces that depends on inside otolith structures.. In any case, notwithstanding ruling more seasoned age classes and including an enormous portion of the producing stock, these spring cod contribute little to enlistment. This clear sub-populace contrast in the stock–enrollment relationship has significant ramifications for evaluation models and the accomplishment of the board estimates intended to modify the stock. Irina I Rypina examined the spatio-worldly dispersion of oceanographic conditions that are helpful for effective producing by bluefin in this locale. In particular, we considered shift in weather conditions examples and water temperatures dependent on another high-goals sea dissemination model. In the wake of approving model speeds and temperatures utilizing perceptions, three criteria were utilized to assess the achievement of reproduced bluefin producing amid 2013: water temperature at generating areas, mean water temperature along larval directions, and larval living arrangement time inside the Slope Sea. Examinations of satellite-based, decade-long (2008–2017) datasets propose that conditions, explicitly water temperatures and shift in weather conditions designs, in the Slope Sea in 2013 were illustrative of run of the mill years. The proposed system is segmented into technical BOT user interface, Natural Language

parser, ontology database, SQL Query Builder, Ocean Database. The user interface takes the responsibility of getting the user query and dispatching it to the natural language parser and finally it will get the result and display it in the intended widget. Here widget is a textbox which displays the output. The natural language parser bifurcates the sentences into tokens and searches for the keywords in ontology database.

The ontology database is considered as the kernel of the system which is entrusted with the responsibility of classification of words such as subject, verbs and nouns. This ontology database is the repository of keywords. The SQL Query builder plays a pivotal role of converting the natural language statements into database specific language which can be understandable by the system. The Ocean database contains the factual information on the various streams about ocean. For example most salinity area name its latitude and longitude, range of salinity which is conducive for fisheries, part of the ocean for economical importance and so on.

## 2. SYSTEM ARCHITECTURE

The system is segregated into BOT User interface, Natural language parser, Ontology database, SQL query Builder, Ocean Database. The user interface gets the user queries in simple English format and then it dispatches to the natural language parser for further processing. It is not concerned about the nitty gritty and the intricacies of the natural language. Once the user feeds the question the NLTK takes the responsibility of converting the sentence into system specific commands.

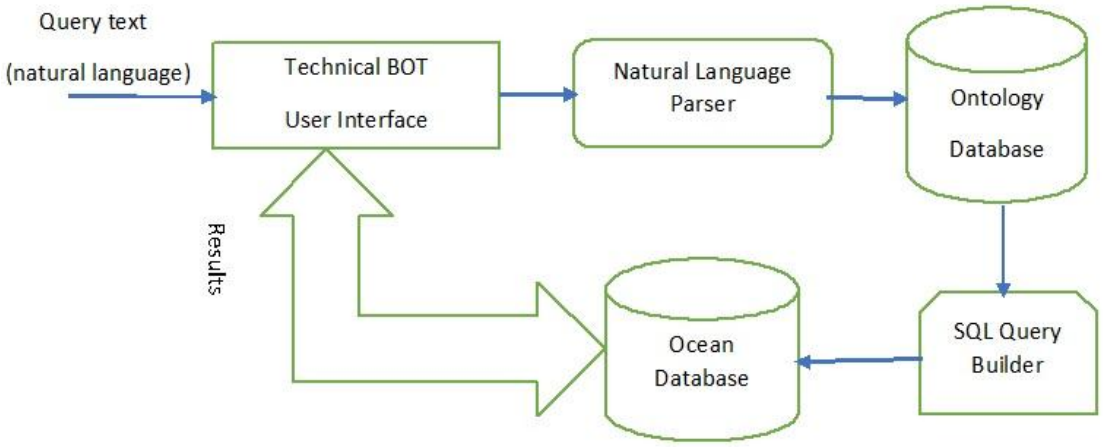


**Fig 1.** System Architecture of Proposed System

Natural language parser splits the sentence and forms them into tokens. This tokenare searched in the ontology database. Once there was a match in ontology database then the corresponding classification is revealed to form the SQL query. The system also maintains two types of databases ontology database and ocean database. The ontology database is used to provide classification for the individual tokens which are formed

from natural language parser. The ocean database contains various information about the ocean and its resources. After the SQL query is formed it is used to get information form the ocean database and in turn the answer is given to the user interface.

### 2.1 Natural Language Parser

The Natural language parser is a python program which uses Natural language processing toolkit (NLTK) to parse the given question in the user interface and form it into individual tokens. The tokens can be classified as subject, verb and object. The subject can be conflated to table in real world database. Objects can be compared to parameters or columns in real world database. Verbs are called as keywords in SQL query.

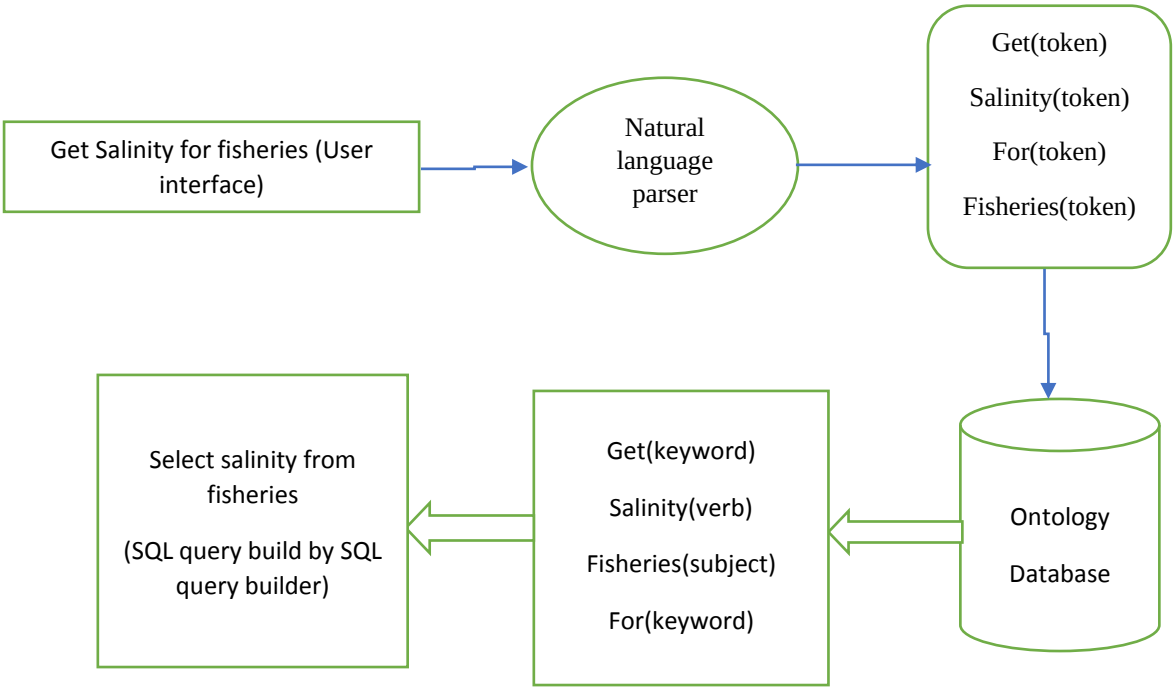


**Fig 2.** Conversion of Natural language to SQL Query

Fig 2. explains how a natural language can be converted into SQL query. Once the classifications of the tokens are done then the query can be build by SQL Query builder.

### 2.2 Ontology Database

This is the kernel of the database which takes the pivotal role in classifying the tokens which forms the sentence in natural language. The tokens are searched in the ontology database where all possible words are maintained as a repository along with the

classification of words such as verbs, object, subject. Every token is searched in the database if the token is found in the ontology database the corresponding classification is fetched and merged along the token. Once the classification is given then the system is ready for framing SQL query.

### 2.3 SQL Query Builder

The classified output from the ontology database is given as input for SQL Query Builder. When a keyword is given as output then it will be replaced by SELECT.eg) GET will be replaced by SELECT. When object is given then it will be replaced as fields in SQL Query. The subject is considered as a Table in the database. Once the query is constructed it will be used to retrieve the data from database.

## 3. CONCLUSION

This proposed project is successful in retrieving information about ocean and its resources. When the verbs, objects and subjects are properly keyed in ontology database this system is most successful in converting the natural language into query. However if non classified words are given in natural language it lacks of categorizing it. So in such kind of scenario more work in deep learning and Artificial intelligence has to done to turn the coin in our favour.

## REFERENCES

[1] R. Mitra, T.M. Marchitto, Q. Ge, B. Zhong, B. Kanakiya, M.S. Cook, J.S. Fehrenbacher, J.D. Ortiz, A. Tripati, E. Lobaton. "Automated species-level identification of planktic foraminifera using convolutional neural networks, with comparison to human performance", *Marine Micropaleontology*, 2019; 147: 16

[2] Spencer Chin, Using Artificial Intelligence to Study the History of Oceans by Spencer chin, North Carolina State University

[3] Hui Li, Qunjie Xu and HonghuaGe ,"The Application of Artificial Intelligence in Ocean Development", Advanced Materials Research (Volumes 864-867).

[4] TuneerMukerjee, "Securing the maritime commons: The role of artificial intelligence in naval operations".

[5] Dr. Timm Schoening," .Understanding deep-sea images with artificial intelligence", Digital Week in Kiel.

[6] Tim Sandle,"Interpreting deep-sea images with artificial intelligence", Science, 2018

[7] Micah J Dean,Scott P Elzey,William S Hoffman,Nicholas C Buchan,"The relative importance of sub-populations to the Gulf of Maine stock of Atlantic cod ",ICES Journal of Marine Science, Marine Science Center, Northeastern

University, Nahant, MA 01908, USA

[8] Irina I RypinaKechen,Christina M Hernandez,Lawrence J Pratt,Joel K Llopiz,"Investigating the suitability of the Slope Sea for Atlantic bluefin tuna spawning using a high-resolution ocean circulation model", I*CES Journal of Marine Science.*

[9] A M Ratheeshkumar, M.Rajkumar, S.Balakrishnan, T R Kalaiarasan, "An Effective Method for Mapping Web User Profile Based on Domain Ontology", International Journal of Engineering and Technology (UAE). Vol. 7, (4.19) (2018), pp. 1-4.

[10] A.Jebaraj Rathnakumar, Venkatachalam K , S.Balakrishnan, "Natural Language User Query To Sparql Conversion For Web Service Discovery From Ontology Based Web Services Registry", Journal of Web Engineering. Vol.17, No.6, May 2018. Pp. 2693-2712.

[11] K. N. Sivabalan, S. Balakrishnan, Securing Sensitive Web Based Student Academic Performance System with Base64 Encoding and Systematic Mirroring", International Journal of Pure and Applied Mathematics, Volume 119, No. 12, 2018, pp. 1117-1126.

[12] S.Balakrishnan, S.Sheeba Rani, K.C.Ramya, "Design and Development of IoT Based Smart Aquaculture System in a Cloud Environment", International Journal of Oceans and Oceanography, ISSN 0973-2667, Volume 13, Number 1 (2019), pp. 121-127.

[13] J.Janet, S.Balakrishnan, S.Sheeba Rani, "IOT Based Fishery Management System", International Journal of Oceans and Oceanography, ISSN 0973-2667, Volume 13, Number 1 (2019), pp. 147-152.

[14] J.Janet, S.Balakrishnan, S.Sheeba Rani, "IoT based lake and reservoir management system", International Journal of Lakes and Rivers (IJLR).

[15] S.Sheeba Rani, S.Balakrishnan, V.Kamatchi Sundari, K.C.Ramya, IoT Based Water Level Monitoring System for Lake in a Cloud Environment, International Journal of Lakes and Rivers (IJLR).

[16] S. Balakrishnan, A. Jebaraj Rathnakumar and K. N. Sivabalan, "Information Security in D-Media (Digital Media)", ARPN Journal of Engineering and Applied Sciences. May 2016, Vol. 11, No. 9, pp. 5707- 5710.

[17] Balakrishnan. S and K L Shunmuganathan. Article: A JADE Implementation of Integrated Agent System for E-Mail Coordination (IASEC). International Journal of Computer Applications 58(5): 5-9, November 2012.

[18] S.Balakrishnan, "An Overview of Agent Based Intelligent Systems and Its Tools", CSI Communications magazine, Volume No. 42, Issue No. 10, January 2019, pp. 15-17.

[19] Balakrishnan S and Steven Uatuaromuinjo Tjiraso, "Integration of Agent Based Computing with Cloud Computing: Towards Cloud Intelligent

Systems", International Research Publication House, Delhi. Engineering and Technology: Recent Innovations & Research, ISBN- 978-93-86138-06-4, pp. 1-17.

[20] S. Balakrishnan, K.N. Sivabalan and J. Janet "MASFE - Mutliagent System for Filtering E-Mails Using JADE", Advanced Engineering Research and Applications (AERA), Research India Publications, ISBN- 978-93-84443-42-9, pp. 148-167, 2017.

[21] P.Arivazhagan, Balakrishnan. S and K L Shunmuganathan. "An Agent Based Centralized Router with Dynamic Connection Management Scheme Using JADE", International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 11, Number 3 (2016) pp 2036-2041.

[22] Balakrishnan. S and K L Shunmuganathan, R. Sreenevasan, "Amelioration of Artificial Intelligence using Game Techniques for an Imperfect Information Board Game Geister" International Journal of Applied Engineering Research (IJAER). ISSN 0973-4562. Vol 9, Number 22 (2014) pp. 11849-11860.

[23] Balakrishnan. S and K L Shunmuganathan, An Agent Based Collaborative Spam Filtering Assistance Using JADE", International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 21 (2015) pp 42476-42479.

[24] A.Jebaraj Rathnakumar, S.Balakrishnan, Design Of Multi-Agent Based Systems For Entrusted Communication Using JADE", Taga Journal of Graphic Technology, Vol. 14, pp. 766-774, 2018.