Database Management of Automated Weather Station and Rain gauge of IMD

Jomin John Felix[1], Nikhil Oke[2], Masoom Raza[3]

[1,2,3]BE Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

Abstract: **The paper exemplify aspects concerning the importance of using cloud computing technology for processing of the information from IMD Station Incharges. The IMD stations are used for monitoring application via variety of sensors. There are some limitations in memory, communication, computation, energy, quantifiability and economical management of the sensors from time to time from IMD stations. The cloud computing is required for powerful, immense superior computing and large storage infrastructure for data processing which can be used for storing IMD maintenance information from completely different stations. The chat-bot will be used for help, reminder, knowledge retrieve etc. Here officers will act with operators, the information should be hold on by utilizing cloud computing technology as a result of this permits computing, storage and software system services (SaaS) with customization prospects and virtualization at low value. The Cloud Computing will offer associate open, flexible, and reconfigurable platform (PaaS) for observance and dominant maintenance applications.**

Keywords: **Cloud computing, IMD, AWS**

1. Introduction

The Asian nation meteoric Department (IMD), conjointly remarked as the Met Department, is place of work of the Ministry of Earth Sciences of the govt. of India. It's the principal agency answerable for meteoric observations, meteorology and geophysical science.IMD is headquartered in metropolis and operates many observation stations across India and Antarctica. Regional offices are at Bombay, Pune, Nagpur, Kolkata.

IMD is additionally one amongst the six Regional specialized earth science centre of the earth's science Organization. It's the responsibility for prognostication, naming and distribution of warnings for tropical cyclones within the Northern Indian Ocean region, as well as the Malacca Straits, the Bay of geographic region, the sea and also the gulf. An automatic weather station (AWS) is an automatic version of the standard meteorological observation post, either to save lots of human labour or to modify measurements from remote areas.

An AWS can usually incorporates a weather-proof enclosure containing the information log-recoder, revitalizable battery, telemetry (optional) and meteoric sensors connected to solar array or turbine mounted upon a mast. The specific configuration may vary according to the requirement of the system. The system could report in close to real time through the Argos System and also the world Telecommunications s system or it can save the information for additional use.

The machine-driven pluviometer (ARG) is developed to collect and record the quantity of precipitation over a group of period of time and automatically passes the data to a central based station on a planned interval basis.The precipitation information are sent through the cellular network as a text message or Short electronic messaging System (SMS). Designed to be rugged and complete, the station may be deployed even within the harsh and remote areas and may operate unceasingly, because it gets power from the sun as well as powered by the interior rechargeable battery.

The readying of this ARG is an element of the event of Hybrid Weather observance System and Production of Weather and Rain machine-controlled Stations Project, that aims to develop and deploy a network of machine-controlled weather stations and ARGs to observe periodic change in weather occurring in specific localities of the country. During maintenance work, an explicit operator will be allotted to a station for maintenance work and all maintenance knowledge is required to be provided by the operator into allotted website.

2. Brief problem to be solved

As of now, time to time support record of IMD station sensor is composed and archived in record books, each upkeep work and change done in AWS and ARG sensors needs to referenced in the record book. Such documentation need to change and to be made Digitalized. At the point when whole upkeep information need to filled by the administrator into a distributed storage and it will be open by approved AWS workstations all over India through a web-based interface.

3. Motivation and scope

Limited information can be accessed to inside constrained era (i.e., seven days) which could be enhanced. Past climate information couldn't be dissected, which could be possible to foresee climate condition in an explicit area. Simplicity of openness of the status of any computerized climate station and rain check stations all over India.

Notice of battery status of Automated climate stations to maintain a strategic distance from information misfortune. The gateway would be accessed to all over India by the AWS areas. The portal will be accessed to by different AWS Stations for refreshing information and the information will be seen and

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

192

analyzed at AWS station, Pune. The data of the AWS station would be useful for the following assessor (Operator) visiting the station.

4. Literature survey

In the survey [1] proposed by Aileni Raluca and Maria Valderrama Carlos, describes the current aspects concerning importance of using cloud computing technology for giant processing from biosensors.

In the survey [2] proposed by Xiao HongJu, Wang Fei, Wang FenMei and Wang XiuZhen, describes the challenge of information management in army data engineering, like huge information volume, heterogeneous information, high rate of knowledge generation and update, time demand of knowledge process, and wide separated information sources.

In the survey [3] proposed by Quan Zou, describes the solutions for enormous remote sensing information analysis and speedy information extraction using high outturn cloud computing interfaces.

In the survey [4] proposed by Chi Yang, Chang Liu, Xuyun Zhang and Surya Nepal, overviews a completely unique information error detection approach that exploits the total computation potential of cloud platform, and also the network feature of WSN.

n the survey [5] proposed by Chao Wu, David Birch, Dilshan Silva, Chun-Hsiang Lee and Orestis Tsinalis, describes a platform known as Concinnity, that permits the cooperative contribution, sharing, and use of massive device knowledge supported cloud-based knowledge repository and easy-to-use work-flow system.

In the survey [6] proposed by Radu Tudoran, Alexandru Costan and Gabriel Antoniu, portrays the Overflow, a uniform information management framework for scientic work process, running crosswise over topographically appropriated locales, meaning to harvest financial benets from this Geo-assorted variety by checking and demonstrating the intercontinental cloud services and architecture.

In the survey [7] proposed by Cheng Hongbing, Rong Chunming, Hwang Kai and Wang Weihong, describes another approach that divides massive information into sequenced elements and stores them among multiple Cloud storage service suppliers so as to guard the mapping of the assorted information components to every provider using a trapdoor function.

In the survey [8] proposed by Luis M. Vaquero, Antonio Celorio, Felix Cuadrado and Ruben Cuevas, describes the large information provisioning service that includes stratified and peer-to-peer information distribution techniques to speed-up information loading into the virtual machines used for processing.

In the survey [9] proposed by Gagangeet Singh, Rajat Chaudhary, Neeraj Kumar and Ashok Kumar Das, proposes, verification, secure storage, and auditing (SecSVA) of huge information in cloud atmosphere.

In the survey [10] proposed by Fabrizio Marozzo, Domenico Talia and Paolo Trunfio, exhibits how Cloud programming innovations can be coordinated to actualize a compelling situation for planning and executing versatile information investigation work processes.

In the survey [11] proposed by Bingquan Liu, Zhen Xu, Chengjie Sun, Baoxun Wang, Xiaolong Wang, Derek F. Wong and Min Zhang, presents a neural system based structure to join client data into the discussion displaying module for tending to the customized reaction positioning issue in the working of programmed chat-bots.

In the survey [12] proposed by Asli Celikyilmaz, Ruhi Sarikaya, Minwoo Jeong and Anoop Deoras, presents the enhancement in execution that to a great extent desire free and demonstrates that such highlights give increases to an extensive variety of undertakings from semantic characterization and space labeling in NLU to named element acknowledgment (NER).

In the survey [13] proposed by Sharefah A. Al-Ghamdi, Joharah Khabti and Hend S. Al-Khalifa, presents the NLP APIs that bolsters Arabic dialect, in which Programmable Web archive was picked as a hotspot for the APIs. It established 23 APIs that help Arabic dialect with various functionalities and dimensions of etymology investigation.

In the survey [14] proposed by Chetan Arora, Mehrdad Sabetzadeh, Lionel Briand and Frank Zimmer, presents a computerized and apparatus bolstered approach for checking conformance to necessities formats. This methodology constructs a mature Natural Language Processing system, known as content chunking.

In the survey [15] proposed by Everton da S. Maldonado, Emad Shihab and Nikolaos Tsantalis, acquaints a methodology with consequently recognize structure and necessity of self-conceded specialized obligation utilizing Natural Language Processing (NLP).

In the survey [16] proposed by Maryam Nafari and Chris Weaver, acquaints an inquiry with inquiry (Q2Q) framework that consequently records client cooperations and presents them semantically utilizing characteristic dialect.

In the survey [17] proposed by Verena Rieser, Oliver Lemon and Simon Keizer, acquaints a methodology with natural dialect generation (NLG) in factual talked discourse frameworks (SDS) utilizing an information driven measurable enhancement structure for steady data introduction, where there is an exchange off to be fathomed between displaying "enough" data to the client while keeping the articulations short and reasonable.

In the survey [18] proposed by Davide Falessi, Member, IEEE Computer Society, Giovanni Cantone, and Gerardo Canfora, presents that NLP methods performed on a given data-set as indicated by both capacity and the chances of recognizing identical necessities accurately.

In the survey [19] proposed by Haiying Shen, Guoxin Liu, Haoyu Wang and Nikhil Vithlani, presents the enhancement of

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

193

execution of QA frameworks by effectively sending inquiries to clients who are fit and willing to answer the inquiries.

In the survey [20] proposed by Zhe Liu and Bernard J. Jansen, presents the subjectivity introduction of inquiries asked on Twitter. They proposed a prescient model dependent on highlights built from lexical, linguistic, and logical viewpoints utilizing machine learning systems.

In the survey [21] proposed by Xiang Cheng, Shuguang Zhu, Sen Su and Gang Chen, presents the detailed inquiry directing as a multi-target positioning issue, and present a multi-target figuring out how to-rank methodology for inquiry steering (MLQR), which can at the same time streamline the noting probability and answer nature of steered clients.

In the survey [22] proposed by Wei-Nan Zhang, Zhao-Yan Ming, Yu Zhang, Ting Liu and Tat-Seng Chua, presents another inquiry recovery demonstrate which can consistently incorporate the key ideas and their interpretations.

5. Existing methodologies

Right now the information of different divisions of AWS is dealt with by individual people independently. Along these lines there is tremendous trouble in keeping up by and large information of AWS stations. The present information records of AWS isn't stored permanently. The information is overwritten following seven days. So they have an absence of information for examination and support.

6. Objectives goals

- To provide single platform for accessing AWS data all over India through a Web portal.
- To provide access to required data of specific center in time specific manner.
- To provide data analysis of data accumulated from different centers i.e. rain gauge data, AWS data and maintenance record of AWS etc.
- A single log for any change in AWS.

7. System feature of IMD maintenance web-page

- The database used is from cloud so it could be accessed from anywhere at anytime.
- The website would provide a single platform for the maintenance operator and the Admin and IMD head-quarters.
- The system would provide an automated environment such giving time to time reminders of AWS stations and its problems and errors to the admin.
- The system would provide a detailed report of the AWS station to the maintenance operator prior to its visit to the station.
- The system would check the images uploaded the by the maintenance operator prior the maintenance work and after the maintenance work if the images are valid or not and does it follow the image constraints.

8. System feature of IMD chat-bot

- There would be a chat-bot on the website to handle user and admin queries related to the system.
- The activity of a talk bot is to have the capacity to decide the best reaction for some random message that it gets. This "best" reaction ought to either (1) answer the senders´ inquiry, (2) give the sender applicable data, (3) ask follow-up inquiries, or (4) proceed with the discussion practically.
- The chat-bot should have the capacity to comprehend the aims of the sender's message, figure out what sort of reaction message is required, pursue right linguistic and lexical guidelines while framing the reaction.
- There is an application layer, that is a database and APIs to call outer administrations. For a situation of the talk bot, User interface is supplanted with chat interface.

A. How Chat-bots works in actual?

The chat-bots work by utilizing 3 grouping techniques:

- *Pattern Matching:*

The chat-bot use design coordinating to arrange the con- tent and create an appropriate reaction for the clients. A standard structure of these examples is A¨ rtificial Intelligence Markup Language¨(AIML). The talk bot can react to anything relating it to the related examples. Be that as it may, it cannot go past the related example.

- *Algorithm (Multinomial Naive Bayes):*

The calculation means to lessen the classifiers and create the more reasonable structure. Computer researchers consider it a "Reductionist" approach so as to give a streamlined arrangement, which decreases the issue. Multinomial Naive Bayes is the exemplary calculation for content characterization and NLP.

1. With new info sentence, each word is meant its event and is represented in shared trait and then each class is allocated a score.
2. The most noteworthy scored class is the well on the way to be related with the information sentence.
3. With the assistance of condition, word matches are evaluated for given some example sentences for each class.

$$P(A\,B) = \frac{P(B|A) * P(A)}{P(B)} \qquad (1)$$

4. Classification score identifies the class with the most elevated term matches.
5. The score connotes which purpose is well on the way to the sentence however does not ensure that it's an ideal match.
6. Most elevated score just gives the relativity base.

- *ANN (Artificial Neural Networks):*

Neural Networks are a method for figuring the yield from the

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

194

information utilizing weighted associations which are determined from rehashed cycles while preparing the information. Each progression through the preparation information changes the weights bringing about the yield with exactness.
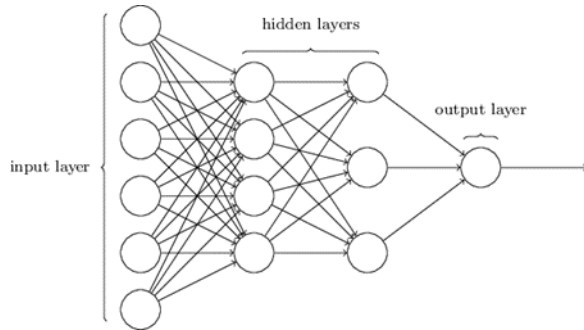

Fig. 1. Structure of ANN

Each sentence is separated into various words and each word is utilized as contribution for the neural systems. The weighted associations are then determined by various cycles through the preparation information a large number of times. Each time enhancing the weights to making it precise.

The prepared information of neural system is a practically identical calculation and less code. The chat-bot can be prepared to retrieve the information with respect to inquiries identified with same informational index. In any case, measure increments by n that can cause an enormous number of mistakes.

NLU (Natural Language Understanding):

Basically it has 3 notions:

- *Entities:* This fundamentally speaks to an idea in your Chat-bot. In IMD chat-bot the substances are climate forecast information, IMD data, AWS and ARG.
- *Intents:* It is basically the activity that talk bot ought to perform when the client say something. The content show trigger single command. In IMD, chat-bot could give required information in timely manner.
- *Context:* At the point when a NLU calculation examines a sentence, it doesn't have the historical backdrop of the client discussion and won't store there client's question. The IMD chat-bot ought to have the capacity to recollect the clients past inquiries as per which it ought to have the capacity to get related information.

NLP (Natural Language Processing):

IMD Chat-bot finds a way to change over the client's content or discourse into organized information that is utilized to choose the related answer.

A portion of the Natural Language Processing steps are:

- *Lexical Analysis:* Lexicon is the words and expressions in dialect. It manages the acknowledgment of structure of the sentence. This

procedure partitions the given printed section into sentences, expressions and words.

- *Syntactic Analysis (Parsing):* It includes examination of words in the sentence for language and organizing words in a way that demonstrates the relationship among the words.
- *Semantic Analysis:* It extract the correct importance or the word reference significance from the content. The content is checked for weightiness. It is finished by mapping syntactic structures and objects in the assignment area.
- *Discourse Integration:* The importance of any sentence relies on the significance of the sentence just before it. Likewise, it additionally achieves the importance of quickly succeeding sentence.
- *Pragmatic Analysis:* Amid this, information exchanged is re-deciphered on what it really implied. It includes inferring those parts of dialect which require true information.
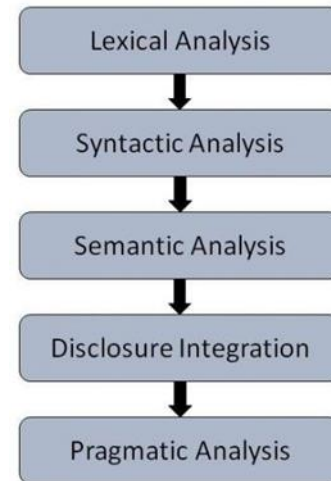

Fig. 2. Steps of NLP

9. System architecture and design
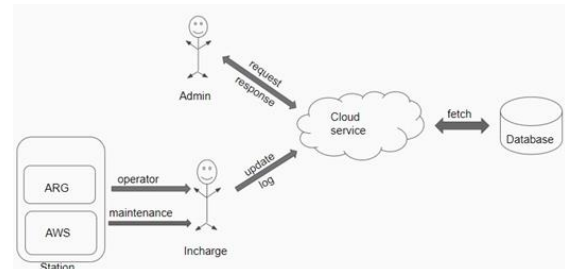
A. *IMD Maintenance Web page*

Fig. 3. Structure of IMD Maintenance Web page

- The station contains AWS (Automated weather stations) and ARG (Automatic Rain Gauge).
- It isn't vital that all stations ought to contain the two parts, it might contain any of the segments.
- The in-control is the sole in charge of upkeep and filling data in the database.

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792
195

- The Administrator will approve the data filled by the station in-control and in the event that any revise is to be done, will organize it, else authorize the information and the upkeep work.

B. IMD chat-bot

- A talk bot is an arrangement of reactions that it provides for a specific message. These are put away in Intents which resemble ideas.
- Expressions are the expressions that you need this purpose to answer to, for instance "what do you like to be called". It utilizes Natural Language Understanding (NLU) to work out what the client is endeavoring to state. On the off chance that they say "What's your name" rather than "What is your name", Lex will in any case coordinate the expressions.
- Accordingly, we have to answer to this message. In here we can type in whatever we need and the bot will react to it. We can enter numerous answers so the client can get fluctuated and progressively characteristic reactions for instance "You can call me Sam-Bot".
- The incite is sent if the Required check box is ticked and the aim does't know the incentive for the variable. Tap- ping the settings machine gear-piece opens up another menu where you can set numerous incite messages and show a few answers.
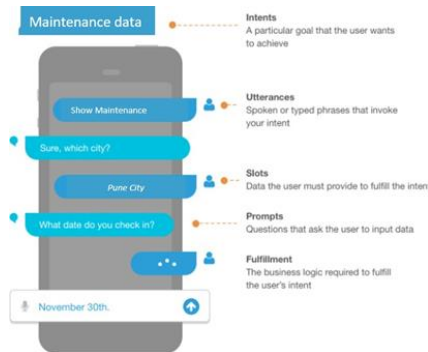

Fig. 4. Components of IMD Chatbot
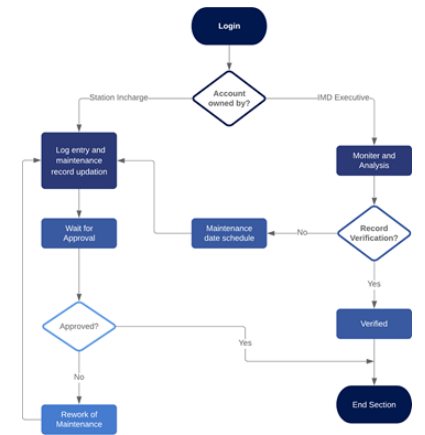
10. Dataflow diagram

Fig. 5. Dataflow of IMD Maintenance Webpage

11. Conclusion

To give a stage to IMD's Automatic climate stations with the assistance of cloud benefit which won't just give support log yet will likewise give security, adaptability and simplicity of getting to required information from any area.

Acknowledgement

References

[1] Aileni Raluca, Maria Valderrama Carlos, Pasca Sever, "Cloud computing for big data from biomedical sensors monitoring, storage and analyze", IEEE Conference Grid, Cloud amp; High Performance Computing in Science (ROLCG), October 2015.
[2] Xiao HongJu, Wang Fei, Wang FenMei, Wang XiuZhen, "Some key problems of data management in army data engineering based on big data", IEEE 2nd International Conference on Big Data Analysis (ICBDA), March 2017.
[3] Quan Zou, "Research on Cloud Computing for Disaster Monitoring Using Massive Remote Sensing Data ", IEEE 2nd International Conference on Cloud Computing, and Big Data Analysis (ICCCBDA), April 2017.
[4] Chi Yang, Chang Liu, Xuyun Zhang, Surya Nepal, Jinjun Chen, "A Time Efficient Approach for Detecting Errors in Big Sensor Data on Cloud ", IEEE Transactions on Parallel and Distributed Systems, (Volume: 26, Issue: 2), Feburary 2015.
[5] Chao Wu, David Birch, Dilshan Silva, Chun-Hsiang Lee, Orestis Tsinalis, "Concinnity: A generic Platform for Big Sensor Data Appli- cation", IEEE, Cloud Computing(Volume: 1, Issue: 2), October 2014.
[6] Radu Tudoran, Alexandru Costan, Gabriel Antoniu, "Overflow: Multi-Site Aware Big Data Management for Scientic Workows on Clouds ", IEEE Transactions on Cloud Computing(Volume: 4, Issue: 1), March 2016.
[7] Cheng Hongbing, Rong Chunming, Hwang Kai, Wang Weihong. "Secure big data storage and sharing scheme for cloud tenants", IEEE, China Communications (Volume: 12, Issue: 6), June 2015.
[8] Luis M. Vaquero, Antonio Celorio, Felix Cuadrado, Ruben Cuevas," Deploying Large-Scale Data Sets on-Demand in the Cloud: Treats and Trickson Data Distribution " , IEEE Transactions on Cloud Computing (Volume: 3, Issue: 2), June 2015.
[9] Gagangeet Singh, Rajat Chaudhary, Neeraj Kumar, Ashok Kumar Das." SecSVA: Secure Storage, Verification, and Auditing of Big Data in the Cloud Environment ", IEEE, Communications Magazine (Volume: 56, Issue: 1), January 2018.
[10] Fabrizio Marozzo, Domenico Talia, Paolo Trunfio, "A Workflow Management System for Scalable Data Mining on Clouds ", IEEE Transactions on Services Computing, (Volume: 11, Issue: 3), July 2016.
[11] Bingquan Liu, Zhen Xu, Chengjie Sun, Baoxun Wang, Xiaolong Wang, "Content-Oriented User Modeling for Personalized Response Ranking in Chat-bots ", IEEE/ACM Transactions on Audio, Speech, and Language Processing, January 2018.
[12] Asli Celikyilmaz, Ruhi Sarikaya, Minwoo Jeong, Anoop Deoras, "An Empirical Investigation of Word Class-Based Features for Natural Language Understanding", IEEE/ACM Transactions on Audio, Speech, and Language Processing, June 2016.
[13] Sharefah A. Al-Ghamdi, Joharah Khabti, Hend S. Al-Khalifa, "Exploring NLP web APIs for building Arabic systems", IEEE, Twelfth International Conference on Digital Information Management (ICDIM), September 2017.
[14] Chetan Arora, Mehrdad Sabetzadeh, Lionel Briand, Frank Zimmer, "Automated Checking of Conformance to Requirements Templates Using Natural Language Processing", IEEE Transactions on Software Engineering (Volume: 41, Issue: 10) May 2015.
[15] Everton da Silva Maldonado, Emad Shihab, "Using Natural Language Processing to Automatically Detect Self-Admitted Technical Debt", IEEE

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

196

Transactions on Software Engineering (Volume: 43, Issue: 11), January 2017.

[16] Maryam Nafari, Chris Weaver, " Query2Question: Translating Visualization Interaction into Natural Language", IEEE Transactions on Visualization and Computer Graphics (Volume: 21, Issue: 6), January 2015.

[17] Verena Rieser, Oliver Lemon, Simon Keizer, "Natural Language Generation as Incremental Planning Under Uncertainty: Adaptive Information Presentation for Statistical Dialogue Systems", IEEE/ACM Transactions on Audio, Speech, and Language Processing, (Volume: 22, Issue: 5), May 2014.

[18] Davide Falessi, Giovanni Cantone, Gerardo Canfora, "Empirical Principles and an Industrial Case Study in Retrieving Equivalent Requirements via Natural Language Processing Techniques", IEEE Transactions on Software Engineering, (Volume: 39, Issue: 1), December 2011.

[19] Guangyou Zhou, Jimmy Xiangji Huang, "Modeling and Learning Distributed Word Representation with Meta data for Question Retrieval",

IEEE Transactions on Knowledge and Data Engineering, (Volume: 29, Issue: 6), February 2017.

[20] Haiying Shen, Guoxin Liu, Haoyu Wang, Nikhil Vithlani, " Social Q&A: An Online Social Network Based Question and Answer Sys- tem", IEEE Transactions on Big Data, (Volume: 3, Issue: 1), November 2016.

[21] Zhe Liu, Bernard J. Jansen, "Understanding and Predicting Question Subjectivity in Social Question and Answering", IEEE Transactions on Computational Social Systems, (Volume: 3 , Issue: 1), June 2016.

[22] Xiang Cheng, Shuguang Zhu, Sen Su, Gang Chen, "A Multi-Objective Optimization Approach for Question Routing in Community Question Answering Services", IEEE Transactions on Knowledge and Data Engineering, (Volume: 29, Issue: 9), April 2017.

[23] Wei-Nan Zhang, Zhao-Yan Ming, Yu Zhang, Ting Liu, Tat-Seng Chua, "Capturing the Semantics of Key Phrases Using Multiple Languages for Question Retrieval", IEEE Transactions on Knowledge and Data Engineering, (Volume: 28, Issue: 4), November 2015.