International Research Journal of Engineering and Management Studies (IRJEMS)
Volume: 03 Issue: 03 | March -2019 e-ISSN: 2395-0126

Automation of database management of AWS and ARG at India Meteorological Department by providing cloud platform

Manisha Desai, Jomin John Felix, Nikhil Oke, Masoom Raza

Department of Computer Engineering, RMD Sinhgad School of Engineering

Abstract— The paper describes the working and implemen-tation of IMD(India Meteorological Department) maintenance website and IMD conversational bot. It explains how various Amazon AWS services are being used to implement the conver-sational bot along with the maintenance website, with the use of AWS services how it is possible to maintain and scale the usage upto various IMD AWS(Automatic Weather Station) across the country(India). It shows the use of AWS Lex, AWS Lambda and Slack messenger for the processing of NLP(Natural language processing) during the implementation of conversational bot and also the use of AWS polly to build the voice bot in-order to integrate along with IMD conversational bot. It also gives brief about the use of AWS RDS, AWS Dynamo, AWS Cognito and AWS CloudWatch for the implementation of IMD Maintenance website and how it could be used to generate the up-to-date AWS(Automatic Weather Station) report from across the country.

Keywords : Cloud computing, IMD, AWS

I. INTRODUCTION

The India Meteorological Department (IMD), additionally alluded to as the Met Department, is an organization of the Ministry of Earth Sciences of the Government of India. It is the foremost organization in charge of meteorological perceptions, climate determining and seismology. IMD has ceaselessly wandered into new zones of use and adminis-tration, and relentlessly based upon its infra-structure in its history of 140 years. It has at the same time sustained the development of meteorology and air science in India. Today, meteorology in India is balanced at the edge of an energizing future.

The need to develop a dynamic and digital database for the maintenance records of the Automated weather stations (AWS) and the Automated rain gauges (ARG) of IMD is very much necessary as the updating of the records in the record books is not possible all the times . So there is a need to develop an online portal which would be accessible to all the AWS stations in India to update the records online. To make the the end user easily access the interface there is also a need of a conversational bot to make understand the user the new interface and solve his queries at that time itself .

So to satisfy all the requirements of the given problem the solution we thought of was to use the amazon web services which provides us web hosting services , cloud storage , and natural language processing services such as AWS lambda and AWS lex.

Google Cloud Platform (GCP) gives a progression of secluded cloud administrations including processing, infor-mation stockpiling, information investigation and machine learning.Google Cloud Platform likewise gives Infrastructure as an administration, Platform as an administration, and Serverless registering situations . Google cloud database is a database that regularly keeps running on a distributed com-puting stage, and access to the database is given as-an admin-istration.Database administrations deal with adaptability and high accessibility of the database. Database administrations make the fundamental programming stack straightforward to the client.

Amazon Lex is a service for building conversational interfaces into any application using voice and text. Amazon Lex provides the advanced deep learning functionality of automatic speech recognition (ASR) for converting speech to text, and natural language understanding (NLU) to recognize the intent of the text, to enable you to build applications with highly engaging user experiences and lifelike conversational interactions.With Lambda, we can run code for basically any kind of utilization or back-end administration .

Amazon Polly is a service that turns text into lifelike speech, allowing you to create applications that talk, and build entirely new categories of speech-enabled products. Amazon Polly is a Text-to-Speech service that uses advanced deep learning technologies to synthesize speech that sounds like a human voice.

II. 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.

In 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 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.

III. PROPOSED WORK

To provide single platform for accessing AWS data all over India through a Web portal. the web portal would be developed by using the latest technology such as html 5 , CSS , javascript.

To provide access to required data of specific center in time specific manner. this work would be done by using google cloud platform . The users would be authenti-cated using the amazon cognito service. The amazon cognito service provides accurate authentication.

To make the conversational bot we are going to use the amazon web services such as amazon lex ,lambda, and amazon polly. Amazon lex and lambda are used for the natural language processing(NLP) purpose while the amazon polly would be used for further increasing the functions of the conversational bot to use the text to speech functionality .

A single log for any change in AWS.

A. 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.

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 conversational-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.

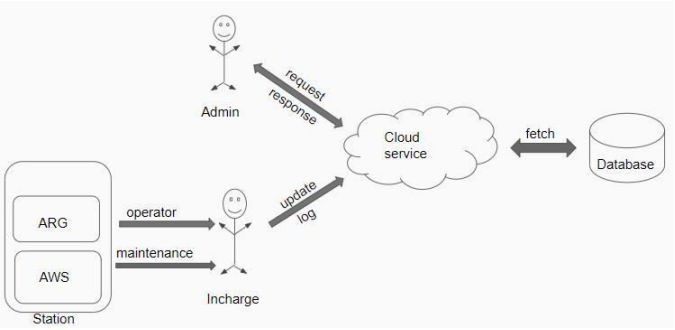

Fig. 1. Structure of IMD Maintenance Web page.

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.

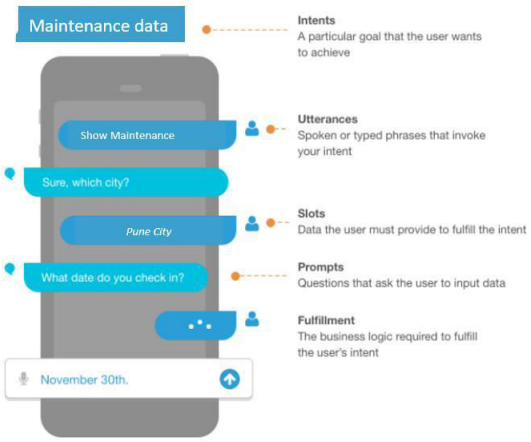

Fig. 2. Components of IMD Conversational-bot.

C. 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.

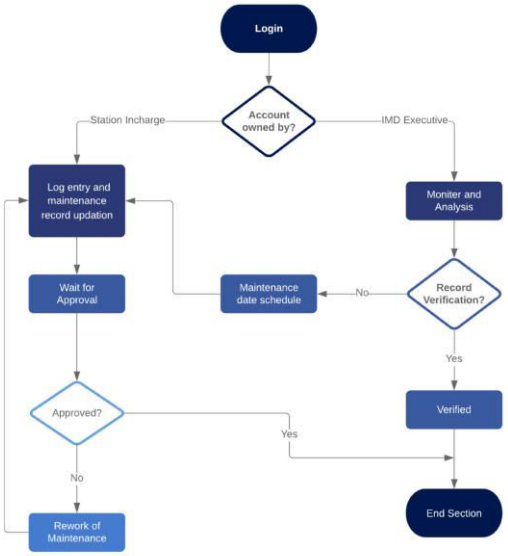

Fig. 3. Dataflow of IMD Maintenance Webpage.

D. System Feature of IMD Conversational-bot

There would be a conversational-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 conversational-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.

IV. SETUP

A. How conversational-bots works in actual ?

The conversational-bot works by utilizing 3 AWS services:

Amazon Lex :

The Amazon Lex is mainly responsible for natural language understanding (NLU) to recognize the intent of the text, to enable us to build applications with highly engaging user experiences and lifelike conversational interactions.

1) Natural Language Understanding :

– Amazon Lex uses some terms in order to implement the bot, Here’s where we give our bot the ability to understand conversations :

– 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, conversational-bot could give required information in timely manner.

– Utterances: It is the sentences provided by the user in-order to invoke an intent, the amazon lex with the help of utterances is able to map most suitable intent and invoke it.

– Slots: It acts like a variable in amazon lex were values must supply to an intent by the user, The values are stored in the slots with the helps of which lex and responsible intent is able to respond with corresponding responses.

– Lex has the functionality of of providing multiple responses to the user, it can also provide a respond card which is used to let user choose from the given options which direct them to another intent.

AWS Lambda :

– The bot is able to understand questions and get inputs, we need to add the logic to perform the skill. We’ll be adding this logic through AWS Lambda.

– AWS Lambda lets us run code without having to create host a server.

– Once lex understands user’s sentence, it produces an input JSON. This input JSON goes to the code performing logic. Lex expects an output JSON in return. Both the input and output JSON have specific formats described in detail.

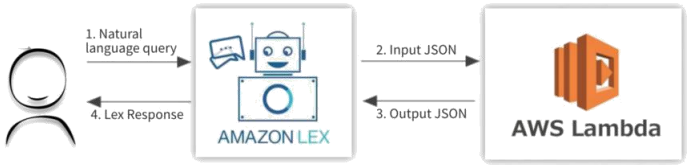

Fig. 4. Workflow of lex with lambda

B. How maintenance website works in actual ?

The working of the IMD maintenance website is based on three main modules . They are Google cloud platform ,AWS cognito and the website . The users are well managed by the AWS cognito service . The users are verified authenticated and then given the access to enter the maintenance report webpage . The users are verified by sending a verification code via the email id .The website acts as the User interface for the AWS incharge and the admin . The website asks the detailed report of the maintenance work form the AWS incharge . It takes all the details of the maintenance report such as physical conditions , action report during the visit , and ask the detailed report of the reading sensors working condition . The website accepts the images of the AWS site before and after the maintenance for verification process of the admin . The website also provides an admin account to verify all the maintenance records done by the aws incharge . If only the admin accepts the maintenance work then only the maintenance report is accepted . The data of the maintenance report is added to the google cloud storage . The google cloud platform provides cloud storage and SQl services .Cloud SQL is a fully-managed database service that makes it easy to set up, maintain, manage, and administer your relational databases on Google Cloud Platform .

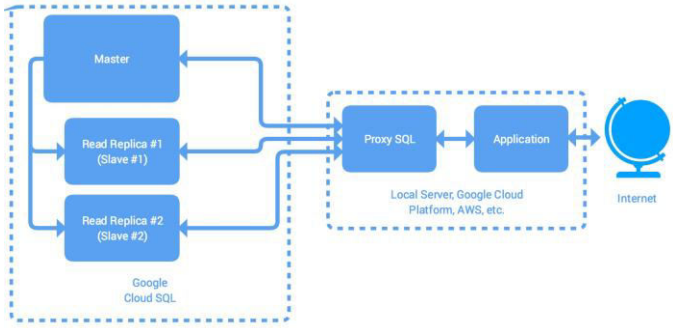

Fig. 5. Working with Google Cloud

V. EXPERIMENTAL SETUP

The experimental setup conducted is to test performance and the responses of conversational-bot with respect to the users queries.

The Amazon Lex provides the flexibility to publish the conversational-bot onto various platforms.

hence we used Slack messenger,the slack has to be integrated with lex to work. The work-space has to be created and the a app has to be generated which will provide secret verification token will be provided which has to be given to lex to be able to interact with slack.

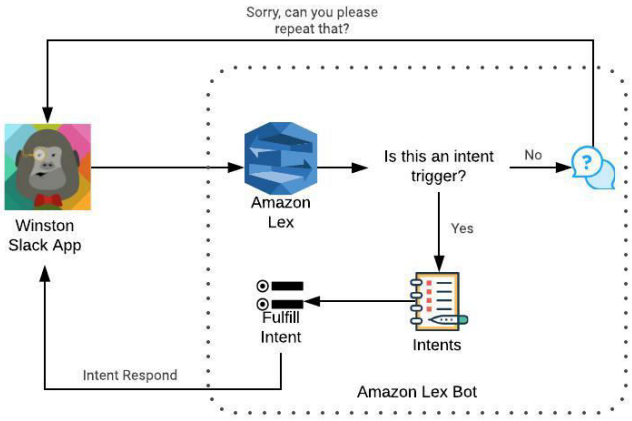

Fig. 6. Workflow of lex with Slack

With help of slack we were able to test the various responses regarding various intents, we were also able to test the respond cards and the hyperlink responses from the lex.

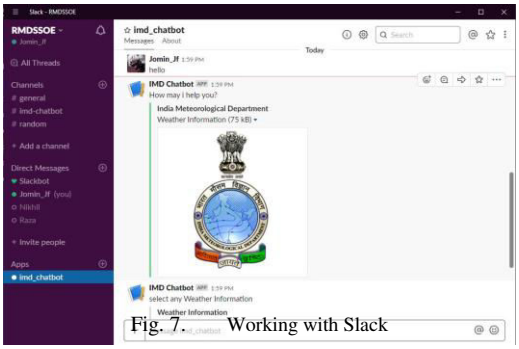

Fig. 7. Working with Slack

VI. DATASET

This project involves the data being divided into 9 classes -:

1) AWS Station :
   This class represents the information about a particular AWS station of India. The information is stored in the form of attributes like Station_ID, Station_name, Address, Power_status, Latitude, Longitude, Datum_pressure, Datum_resistance, Azimuth_antenna_elevation, Sunshine_sensor_elevation, Date_of_commissioning, Commissioning_officer.
2) Sensor :
   This class contains the record of various sensors which

includes Thermometer, Anemometer, Wind-vane, Hy-grometer, Barometer, Ceilometer, Visibility sensor, Rain gauge, Ultrasonic depth sensor, Pyranometer etc. used in AWS and ARG to monitor the weather con-dition of an area. The information of these sensors are stored in the form of attributes like Sensor ID, Sensor speed, Sensor direction, Housing temperature, Sensor sensitivity, Sensor elevation.

3) Data logger :
The data measurement of the above sensors are man-aged in this class. The class will have the attributes to represent these stored data along with the informa-tion of there relative power supply, calculations and communication services.
4) Enclosure :
These are the equipment which is used to provide safety to the sensors in harsh condition.These are made up of fibreglass, Stainless steel etc. And are used as a cover for devices.This class will store the informa-tion of these enclosure in form of there Model type, Date of manufacture, Expected lifetime, Color, Design etc
5) Power supply :
All the sensors and equipment of AWS executes by consuming power from various sources like So-lar energy, Batteries, Regulators, Wind turbine, Lo-cal electricity grid. So, this class will be responsible for maintaining the information of these power sources.
6) Maintenance :
Every equipment degrades with passage of time. So, these equipments needs to be maintained in order to get optimal result every time.The class deals with the management of the information related to the maintenance of all the equipment of AWS across the country. The information is stored in the form of attributes like Station ID, Start date, End date, Cost, Time, Supervisor, Resources, Images.
7) Servicing :
This class stores the record of the activities performed during the maintenance work. The attributes representing the information are Date of servicing,
Service officer, Faults observed, Battery type, Charger details, Date of failure, Details of service, Remarks.
8) Mast :
These are the vertical poles on which the equipments are installed and aids to measure parameters over stratified distances. The class is responsible for the data maintenance of the information related to the mast like Location coordinates, Type, Area covered, Date of installation, Functionalities etc.
9) Report :
This class provides the summary of the maintenance of the AWS and also predicts the next session of mainte-nance of particular equipment or station in accordance of date. It inherits all the properties of the above class and and forms a relevant report for guidance and analysis.

Meta-Data -:
All the above attributes mentioned are in the format of standard SQL data type like varchar, Integer, date etc.

VII. RESULT

Effectively provided automation to traditional data man-aging method performed at IMD station.
Provided cloud platform for data storage which offers benefits like fast accessibility, Metered cost, scalability, security etc.
Provided a web portal through which the manipulation and authorization of maintenance data takes place.
Provided an intelligent conversational chat-bot built upon using amazon Lex and Lamda to enhance user interaction.
Effectively generated report of maintenance work date wise.

VIII. CONCLUSIONS

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.

ACKNOWLEDGMENT

The maintenance data structure is been provided by India Meteorological Department, the guidance of Mrs. Manisha Desai Asst. Professor Department of Computer Engineering RMD SSOE, India.

REFERENCES

[1] Aileni Raluca, Maria Valderrama Carlos, Pasca Sever, "Cloud com-puting 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 conversational-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 Lan-guage Understanding", IEEE/ACM Transactions on Audio, Speech, and Language Processing, June 2016.

[13] Sharefah A. Al-Ghamdi, Joharah Khabti, Hend S. Al-Khalifa, "Explor-ing NLP web APIs for building Arabic systems", IEEE, Twelfth In-ternational 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 Transactions on Software Engineering (Volume: 43, Issue: 11), January 2017.

[16] Maryam Nafari, Chris Weaver, " Query2Question: Translating Visu-alization 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 In-formation 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 Prin-ciples and an Industrial Case Study in Retrieving Equivalent Require-ments via Natural Language Processing Techniques", IEEE Trans-actions on Software Engineering, (Volume: 39, Issue: 1), December 2011.

[19] Guangyou Zhou, Jimmy Xiangji Huang, "Modeling and Learning Dis-tributed 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.