



Student Query Trend Assessment with Semantical Annotation and Artificial Intelligent Multi-Agents

Kaleem Razzaq Malik

COMSATS Institute of Information Technology, & University of Engineering and Technology, Pakistan

Rizwan Riaz Mir

Virtual University of Pakistan, Pakistan

Muhammad Farhan

COMSATS Institute of Information Technology, & University of Engineering and Technology, Pakistan

Tariq Rafiq

COMSATS Institute of Information Technology, Pakistan

Muhammad Aslam

University of Engineering and Technology, Pakistan

• Received 17 November 2016 • Accepted 12 May 2017 • Accepted 19 May 2017

ABSTRACT

Research in era of data representation to contribute and improve key data policy involving the assessment of learning, training and English language competency. Students are required to communicate in English with high level impact using language and influence. The electronic technology works to assess students' questions positively enabling semantics and intelligence in the field concerning education and health. Assessing the importance and complexity of the statement used in a query can save the effort needed to automate the questionnaire system involving better skill testing and formalization. Parts of Speech (POS) for a sentence can be assessed for improving and enhancing the utilization in students' querying skill in writing. Computer aided systems built-up on trained agents to assess data orientation for measuring the strength of the questionnaire as being plotted to test skill of the examinees. These agent needs to be made trained and provided with the data format capable to use for intelligent assessment and strong linkage. This can be done using platform of semantic web data model; well known as Resource Description Framework (RDF). To achieve this purposed study, we represent a methodology to identify each query statement tagged per its parts of speech. Then train agents to assess data impact in calculating complexity of each query. This tagged query is further transformed into RDF to give semantics and hierarchal attachment between parts of the speech.

Keywords: human-computer-interactions, e-learning, web 2.0, knowledge management

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

Correspondence: Kaleem Razzaq Malik, *Department of Computer Science, COMSATS Institute of Information Technology, Sahiwal, Pakistan.* Phone: +92-333-761-6116.

✉ krmalik@gmail.com

State of the literature

- Learning management systems are used to manage the student academic activities and for the interaction with the teachers but lack the intelligent trend assessment of student grades (Kaula, 2016).
- Instructional Design can be personalized and customized according to the needs of the students
- Distance Education and e-Learning are new frontiers of exploration for the patterns and trends of students' evaluation and assessment (Kaula, 2016)

Contribution of this paper to the literature

- Knowledge management and e-assessment used to achieve the automated evaluation of students' query trends
- This paper presents a novel approach to measure and analyze the students' queries trends using computer intelligent software called artificial intelligent agents
- An algorithmic approach for the patterns finding from the query data for student and finding the patterns in them.

INTRODUCTION

Research shows monitoring students activities, capabilities, and skills can help building educational system based on e-learning more prominent and useful (Lin, Yen, Liang, Chiu, & Guo, 2016). Among big challenges the on challenge where student query needs to be understood and analyzed for assessing the problem requires to be addressed (Çalık, 2013). Whereas assessment is now a day considered as a tool in the field of e-learning based education (Qayyum & Smith, 2015). Web data and its representation plays a big role in defining a mechanism to overcome any delays in producing results accurate and effective for electronic assessment (e-assessment).

In the field of Natural Language Processing (NLP) where WordNet is a considerable lexical database of English statements. English statements POS consists information of Noun, Verbs, Adjectives, and Adverbs are assembled into sets of subjective synonyms (Synsets), each communicating an unmistakable concept (Bond & Foster, 2013). Synsets are groups of English words that are formed as collection / Set of synonyms. Synsets are interweaved by method for calculating semantic and lexical relations (Meng, Huang, & Gu, 2013). The succeeding system of genuinely related words and ideas can be explored with the program. In addition, WordNet is utterly and openly accessible for download. WordNet's structure is helpful instrument for computational linguistics and natural language processing (NLP) (Bond & Foster, 2013; G. A. Miller, 1995). The process is showed in **Figure 1**.

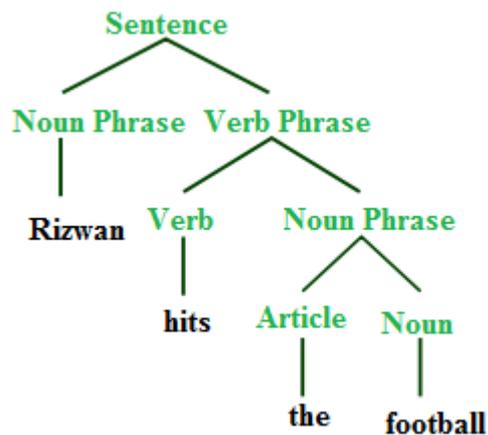


Figure 1. Structure of a Sentence using WordNet

Most of the WordNet's relations join words from the same grammatical feature (POS). In this way, WordNet is truly comprised of four sub-nets, one each for things, verbs, descriptors, and intensifiers, with few cross-POS pointers. Cross-POS relations incorporate the "morphosemantic" joins that hold among semantically comparable words offering a stem to the same importance: watch (verb), perceptive (modifier) perception, observatory (things) (G. Miller & Fellbaum, 1998).

The gigantic increment in the sum and many-sided quality of reachable data in the Internet brought about an over the top interest for instruments and systems that can deal with information semantically (M. Zhang, 2014). With the difficulties in data recovery technologies generally depends on the e-assessment based questioning built over complete examination system, which is demonstrated with sack of-words (Castillo, Herrera, Carrillo, & McCalman, 2016; Qayyum & Smith, 2015). On the other hand, such a model misses the genuine semantic data in content (Meng et al., 2013). To manage this issue, ontologies are proposed for knowledge representation, which are these days the foundation of semantic web applications (Mitri, 2003). Both the data extraction and recovery procedures can profit by such metadata to support assessments which offers semantics to educational systems (Qayyum & Smith, 2015). As a middle and common data representation on web Document Type Description (DTD), a subset of SGML DTD, is used and accepted as standard. An XML structure representation dialect has been used over a significant time span and is well on the way to decoration until Extensible Markup Language (XML) Schema at last arrived (Diao & Franklin, 2016). It has restricted abilities contrasted with other mapping dialects (Ibrahim, Atif, Shuaib, & Sampson, 2015). Its primary building block comprises of an element and an attribute. This presents data's metadata in simplified form by the utilization of various levelled element structures. XML Schema is a continued exertion of World Wide Web Consortium (W3C) to help, overall replacing DTD in the XML world. XML Schema means to be more expressive than DTD and more usable by a more extensive and expressive data

structure utilizations. It has numerous novel systems, for example, inheritance for attributes and elements, client characterized datatypes, and so forth (Fallside & Walmsley, 2004).

Now this XML data becomes simplified data format and target to get transformed into Resource Description Framework (RDF) as a part of Semantic Web (M. Zhang, 2014). RDF data format brings the possibilities of involving intelligence in the system with the help of agents. Where RDF provides data in the form of triples comprised of source, predicate, and resource. Data transformed from queries into RDF can further sent to agents trained to gain assessment out of measuring parts of each sentence (Qayyum & Smith, 2015).

As tag / Synsets conversion is based on procedures by Stanford Parser, so we are bound to accommodate limitations of this package used by WordNet (Bond & Foster, 2013). So, it requires to be sure about syntax and semantics of entered text as WordNet / Stanford Parser would not cater, till this time. If we are talking about benefits of using agents, it is not up to the mark now for the field of education in sense to achieving parallelism and learning collaborative agents. Through conversion of input stream into DB, XML and RDF, we only propose assessing mechanism on well-known datatypes.

This research will formulate a solution to convert WordNet (large lexical database of English language) Synsets into RDF triples. RDF triples are used as raw data to formulate ontologies. So, our future work will focus on a step forward to form ontologies. Another thing in mind is that to form multi-agents that can not only communicate with each other but also will help to form a network based collaborative agents leaning. Handling data types that are not usually used in Relational DB, we can mold our scope to draft a solution in terms of unified procedure to get Big Data representation into standard web resources (Ibrahim et al., 2015).

This study investigates student questionnaire to check their level of composition by looking at their POS effective utilization with the use of scoring function. This study is further divided into several sections starting next from literature survey to cover updates in area of NPL and Semantic Web for assessment (Qayyum & Smith, 2015). Then material and methods section covering implementation, modeling and data transformation along with scoring function to capture each query impact by training agents. Afterwards, in results and discussion section different queries taken from Microsoft free QA corpus are used to see outcome of the system purposed. Next section of verification is about transformed data format verification. In last there are sections of conclusion and references.

LITERATURE REVIEW

This section provides some review with information about the techniques, terms and technologies used throughout this study. We begin with the background of this research. Then, we continue with the important aspects of the Semantic Web, such as ontologies, languages, data models and rules for acquiring data oriented assessments (Ibrahim et al., 2015). In computational lexicology (a branch of computational etymology) one can see a slow move (as far as wording and semantic substance) from machine comprehensible lexicons to

lexical information bases and afterward to lexical ontologies (Ibrahim et al., 2015). A machine decipherable lexicon (Amsler, 1984) speaks to paper word reference information in electronic shape in this manner empowering these information to be handled on a PC. The lexical learning base differs from the machine lucid lexicon in that the word implications are unequivocally shown and connections between the comparing implications are indicated, which makes it conceivable to utilize these information for logical induction (Calzolari, 1991).

Lexical metaphysics contains organized data about the words and incorporates semantic connections between the implications of words (Bach, Chao, Kempson, Fernando, & Asher, 2012). On the other hand, a significant piece of connected ontologies is built for a particular branch of knowledge with the sign of relations between the ideas of corresponding range (Gawrysiak, Ryzko, Więch, & Kozłowski, 2012). As of now, in any case, it is the inadequate size of dictionaries, thesauri, and ontologies that represent an awesome issue for application. Subsequently, a broadly useful lexical metaphysics contains organized data about words and semantic relations. In the meantime, there is no connection to a specific branch of knowledge (Mitri, 2003). WordNet is viewed as a standout amongst the best ventures of this kind (G. A. Miller, 1995) as shown in **Figure 2**.

To support improved lexical features through ontologies as a resource for Natural Language Processing (NLP) applications (Eckle-Kohler, McCrae, & Chiarcos, 2015).

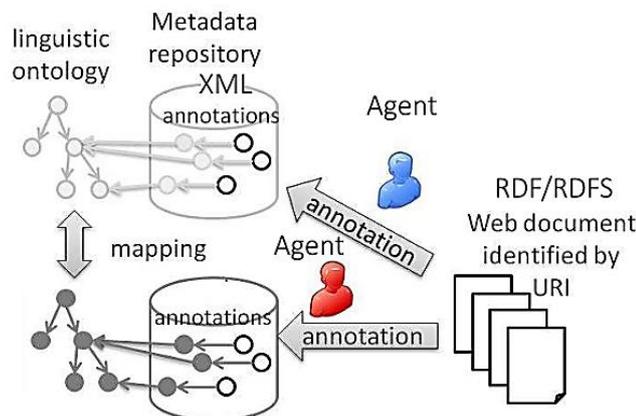


Figure 2. Annotation Methodology

These days NLP based applications face problems when it comes towards domain specific translation. The need to overcome this issue, an agent can be trained on specific domains to support better assessment results for translation. Furthermore, there is a big need of interlinked grammar for resolving sentence understanding when used for reasoning under artificial intelligence. In this, ontologies can play a key role as they have capabilities of data

linkage. This research work will help in addressing an agent based system to capture all possible relations built for an English statement Lexicon. With help of these lexicons, we would be able to convert it to XML format that is portable and supported by many frameworks. RDF triple conversion is followed by XML that would be used as resource for ontology creation (Diao & Franklin, 2016).

There are several NLP techniques in use. Different tools in market have built in natural language processing techniques either for developing XML, RDFs and making semantic clouds (Diao & Franklin, 2016). Artificial Intelligence is also being applied to easily process natural language. Considerations in this research will be to less the computational efforts in terms of natural language processing. Our developed service / program will take natural language as an input and convert it in Lexicons with the help of WordNet Synsets (Bond & Foster, 2013). It will automatically be converted to XML and RDF through use of multi agents. Agents can communicate with each other and act as collaborative agents. XML and RDF are according to standards defined by W3C and end results obtained will be in the form of triples (Subject, Predicate, and Object).

“Asserting an RDF (Resource Description Framework) triple says that some relationship, indicated by the predicate, holds between the resources denoted by the subject and object. This statement corresponding to an RDF triple is known as an RDF statement. The predicate itself is an IRI and denotes a property, that is, a resource that can be thought of as a binary relation. (Relations that involve more than two entities can only be indirectly expressed in RDF).” (Klyne & Carroll, 2006). Semantic Web, every one of the efforts towards expanding the recovery execution while safeguarding the ease of use will in the end get to the heart of the matter of enhancing semantic looking with catchphrase based interfaces (Ibrahim et al., 2015). This is a testing errand as it requires complex inquiries to be replied with just a couple of watchwords. Besides, it ought to permit the induced information to be recovered effortlessly and give a positioning system to reflect semantics and ontology significantly as a tool to perform better e-assessments.

Software agents play a vital role to get trained in NLP and contribute to semantic cloud. An agent or a software agent is a computer program that is used by a user or program (sub-program or independent). To use the services of agent-program an agreement is resulting between programs to act on one's behalf. Such “action on behalf of” infers the power to choose improving the assessment possibilities with which, if any, action is appropriate and applicable. Different types of agents include intelligent agents (in particular revealing some aspect of artificial intelligence, such as learning and reasoning), autonomous agents, distributed agents (being executed on physically distinct computers), multi-agent systems (distributed agents that do not have the capabilities to achieve an objective alone and thus must communicate), and mobile agents (agents that can relocate their execution onto different processors) (Kaula, 2016).

MATERIAL AND METHODS

After the literature review, devised a proposed solution in which multi-agent are developed to get tags from WordNet dictionary and then convert them into lexicons. Lexicons are then converted into XML (W3C standard). XML is then transformed into RDF triples (W3C standard) that is used as a supplement to develop ontologies to form semantic cloud. Proposed solution is shown in **Figure 3**.

Proposed Model

Natural Language processing (NLP) study is being started on motivation of improved semantical annotation mechanism as shown in **Figure 4**. After study of literature review, a flaw is identified that techniques and mechanism lack in translating lexicons to RDF. Previously, tools are available but not suitable in generic sense.

Steps needed for completely transforming sentences into ontology are:

1. Capture whole data on which an Ontology Cloud is needed
2. Separate each sentence followed by period ignoring cases like "e.g.", "etc."
3. Tag each word of the sentence using WordNet dictionary API
4. Save collected information into concern table under the control of concerned agent e.g. Noun, Verb etc. And then assessment is using the scoring function for measuring complexity

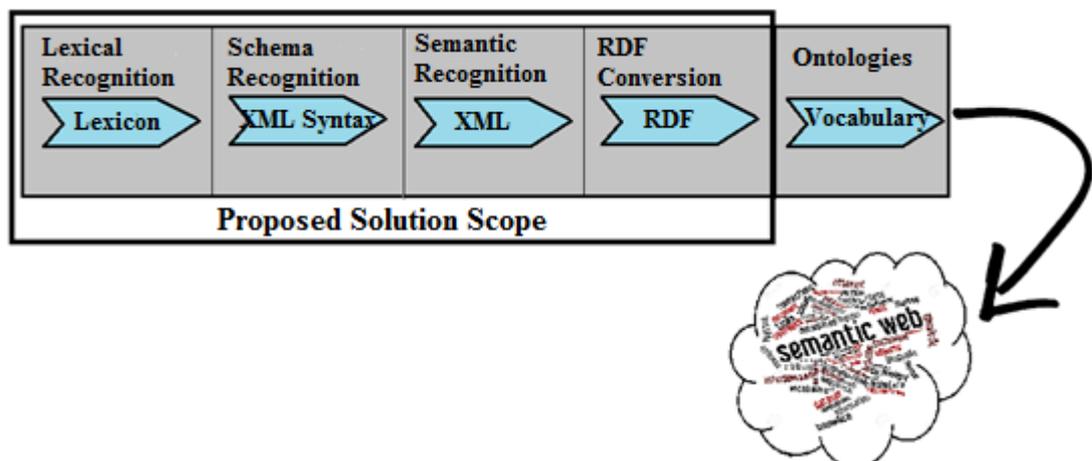


Figure 3. Proposed Solution

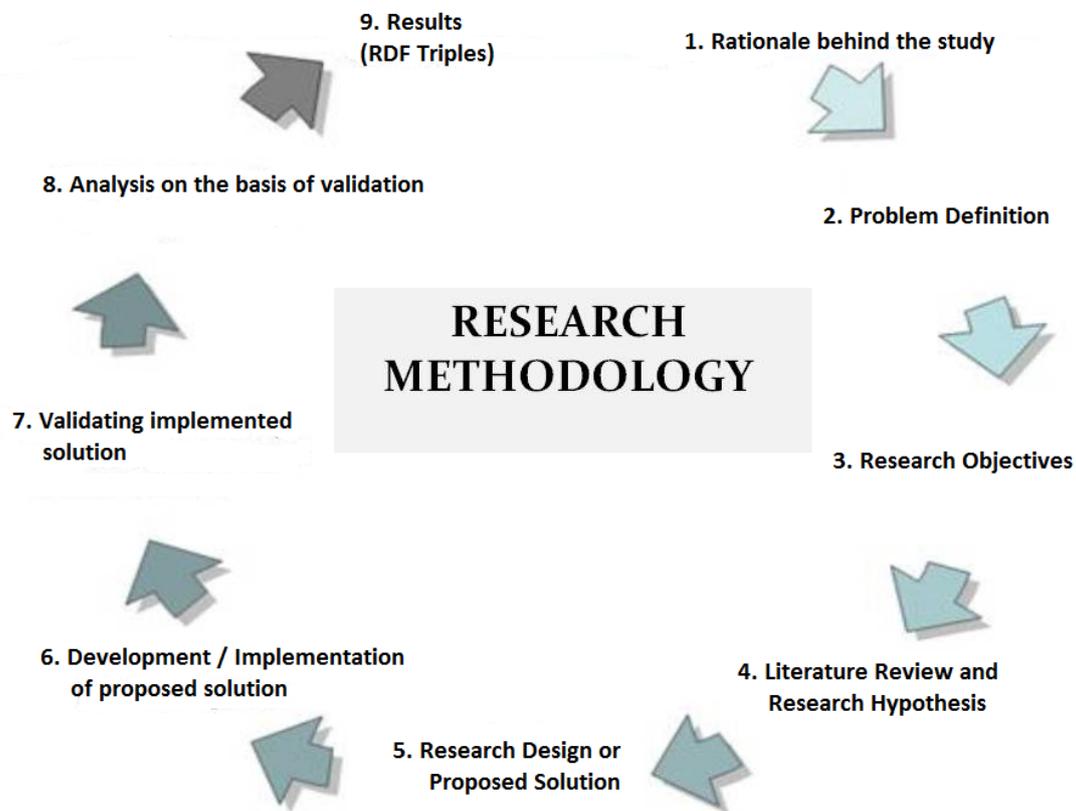


Figure 4. Research Methodology

1. Transform DB (schema and data) into XML
2. Validate XML (schema and data)
3. Transform XML into RDF triple
4. Validate RDF (schema and data)
5. Generate RDF Graph (schema and data)

Validation is performed in two phases. In 1st phase of validation; XML document is verified and validate for XML schema by W3C validate service that is available online. And in 2nd phase, RDF triples are validated online from W3C validation service to check RDF schema. Graphs are available to show validation of results. Process of implementing idea and its validation is shown in **Figure 5**.

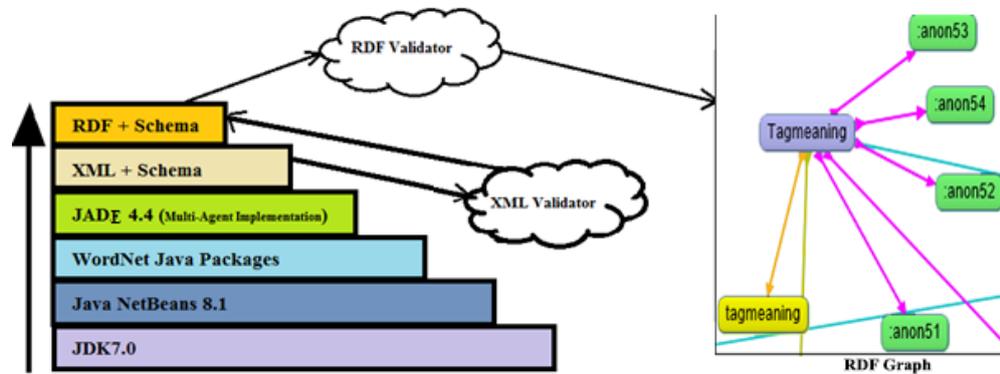


Figure 5. Implementation and Validation Process

Question Scoring Model

In Figure 6, questionnaire complexity scoring model (Shi, He, Wang, Fan, & Guo, 2016) is represented showing a query is sent picked from a corpus of questionnaire as a sentence to the WordNet API for further tagging and POS conversion. Each part of speech combined with their occurrences are computed for assessing for score of complexity by passing through a scoring function s as shown in Eq. (1).

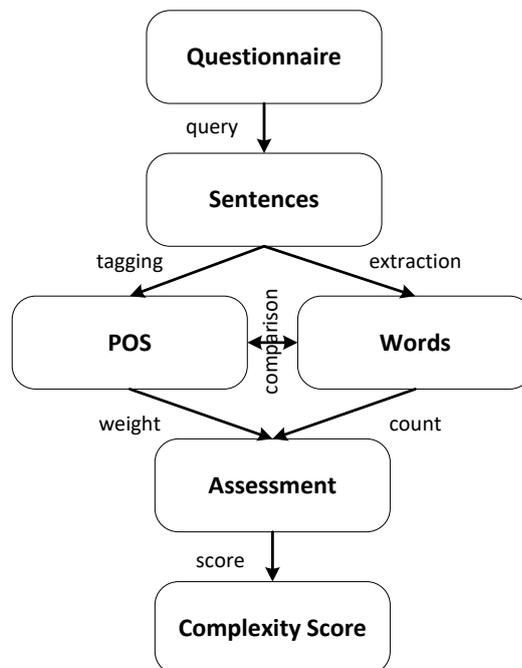


Figure 6. Questionnaire Complexity Scoring Model

$$\begin{aligned}
 & N(Noun) \\
 POS = & V(Verb) \\
 & Ad(Adverb) \\
 & Aj(Adjective) \\
 s_o^i = & \begin{cases} +2 & \text{if } w_i \in N \text{ and } V \\ -1 & \text{if } w_i \in Ad \text{ and } Aj \\ 0 & \text{if } w_i \notin N, V, Ad \text{ and } Aj \end{cases} \quad (1)
 \end{aligned}$$

$\therefore w$ belongs to the set of words in a query statment

IMPLEMENTATION

Proposed process is translated into flow chart, as shown in **Error! Reference source not found.**; that was latterly assisted in systematic implementation of proposed work. This flow chart is created by keeping the implementation details in view. For implementation of research work tools used are, Java NetBeans 8.2 is used having JDK 7. WordNet packages is imported in NetBeans to use services of WordNet to get Synsets. JADE (Java Agent Development Framework) 4.4 is used to simplify the implementation of multi-agents. Java based technologies are selected due to its strong features of object oriented language, portability, multi-platform support, secure and of course freely available to use.

Initially text is passed to Stanford Parser that parse text into tokens and for each token its parts of speech type (noun, verb, adverb or adjective) is also returned in a list. In the meanwhile, agents are created using JADE framework against each type of part of speech. Hence there are four agents created named as Noun-Agent, Verb-Agent, Adverb-Agent, and Adjective-Agent. Returned list of tags and lexicons are searched by agents. The agents get each token of its required type and then entered in MySQL database named as "mydb". Meanwhile, these agents also update a list of their respective POS type along with tags, the lexicons have (Telepovska, 2014). After all the tokens entered in database and updating POS lists, agents are terminated. These agents are working collaboratively with each other for the task described above. Termination of agents are followed by writing text file containing information of lexicons and tags.

Next task after writing entering in Db and writing text file is creation of XML schema per Database "mydb" structure (Telepovska, 2014). Then XML is generated on basis of values in database. XML and XSD are verified and validated against each other per World Wide Web Consortium (W3C). If validation is performed for XML against XSD, an important and last task of study work is performed that is generating RDF Schema and RDF files. RDF Schema is designed per XSD file generated earlier. And RDF file is generated against XML file. RDF files creation ends with validation process from online provided resource. That will lead to triple generation. The whole described process is shown in provided below. Another representation of implementation flow chart for thesis work is represented in Figure 8.

Transformation of English Lexicon Tags into RDB Tuples

First phase of transformation is started with query that is a string. When system receives an input, this is converted into tags / Synsets considered as very first transformation task. Meanwhile, agents are activated (JADE) to get their respective tag / lexicon, then its core responsibility of agents to save these lexicons in their respective table of relational database named as “mydb”.

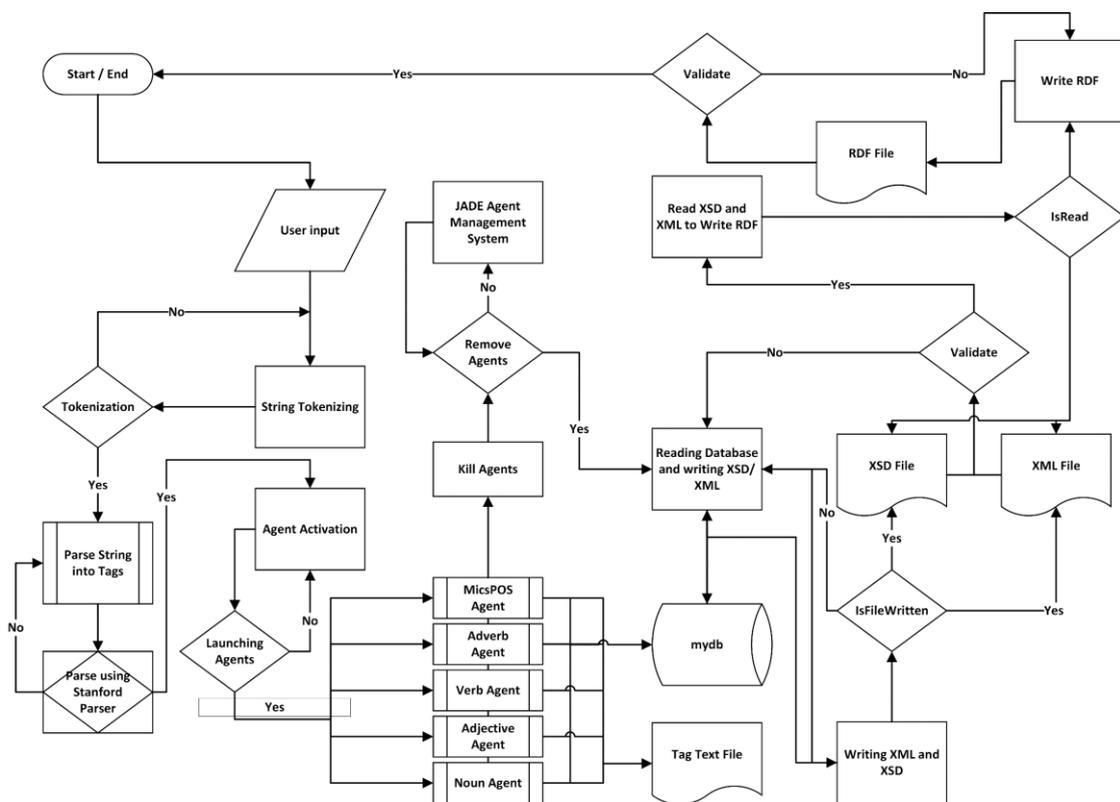


Figure 7. Complete Query POS Tagging and Transformation Process

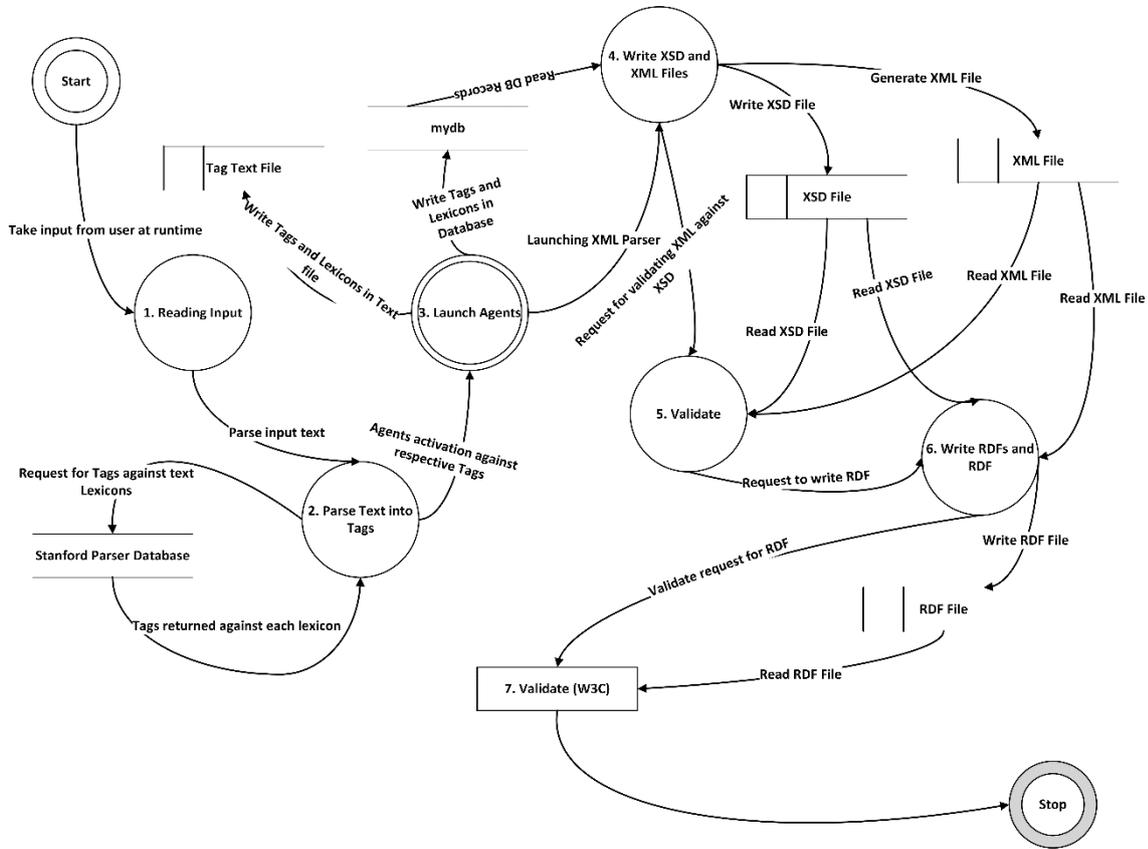


Figure 8. State wise representation of Proposed Mechanism

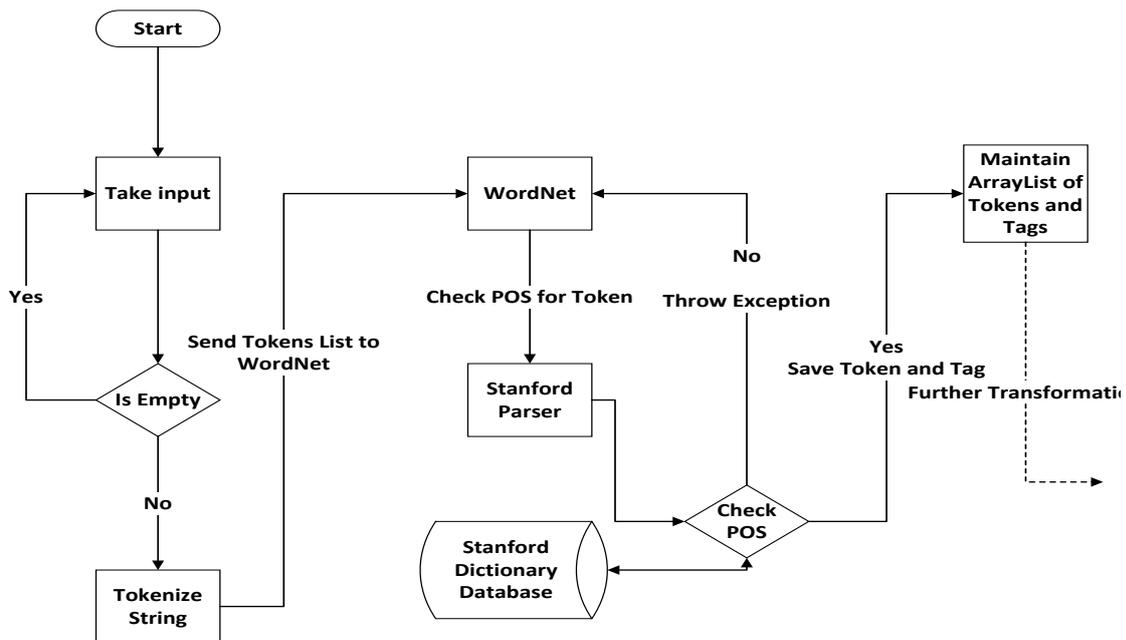


Figure 9. Transformation of Text into Tags / Synsets

Transformation of Simple Text into Lexicons

A user could input a string in the system that is transferred to WordNet. WordNet will further share the text with Stanford parser. Stanford parser would parse the received string into tokens. Then against each token, its parts of speech (POS) type is returned that is rather the token is of noun, verb, adverb or adjective type POS. Stanford parser is instructed through code to merge each token and its POS into an ArrayList and after completion of task, this Array list is forwarded as reference to the sub task to perform further action. A flow of activities performed are shown in **Figure 9**. Instructions to tokenize the entered string and then transform the string into tokens, are written in Java. Then through Java instructions, Tokenized string is inputted to WordNet for further processing that will pass tokenized string to Stanford Parser. A snapshot of implemented code is provided below in **Figure 10**.

Creating and managing JADE Agents

As discussed in Implementation chapter 4, Jade is used to create and manage agents in Java to perform certain tasks on run time. When tags are returned in ArrayList, agents are created by using Jade to check the tags one by one and get information of tag and token to store in relational database for their respective POS (Parts of Speech) information (Telepovska, 2014).

```

28 public static ArrayList parseString(String sent) throws IOException, Exception
29 {
30
31
32     String grammar = "edu/stanford/nlp/models/lexparser/englishPCFG.ser.gz";
33     String[] options = { "-maxLength", "500", "-retainTmpSubcategories" };
34     LexicalizedParser lp = LexicalizedParser.loadModel(grammar, options);
35     TreebankLanguagePack tlp = lp.getOp().langpack();
36     GrammaticalStructureFactory gsf = tlp.grammaticalStructureFactory();
37
38     Iterable<List<? extends HasWord>> sentences;
39
40     Tokenizer<? extends HasWord> token =
41     tlp.getTokenizerFactory().getTokenizer(new StringReader(sent));
42     List<? extends HasWord> sentence2 = token.tokenize();
43
44     List<List<? extends HasWord>> tmp =
45     new ArrayList<List<? extends HasWord>>();
46     tmp.add(sentence2);
47     sentences = tmp;
48
49     Tree parse = lp.parse(sent);
50
51     taggedString = parse.taggedYield();
52
53     return taggedString;
54 }

```

Figure 10. Java Code to get tags for tokens and then store into ArrayList variable taggedString

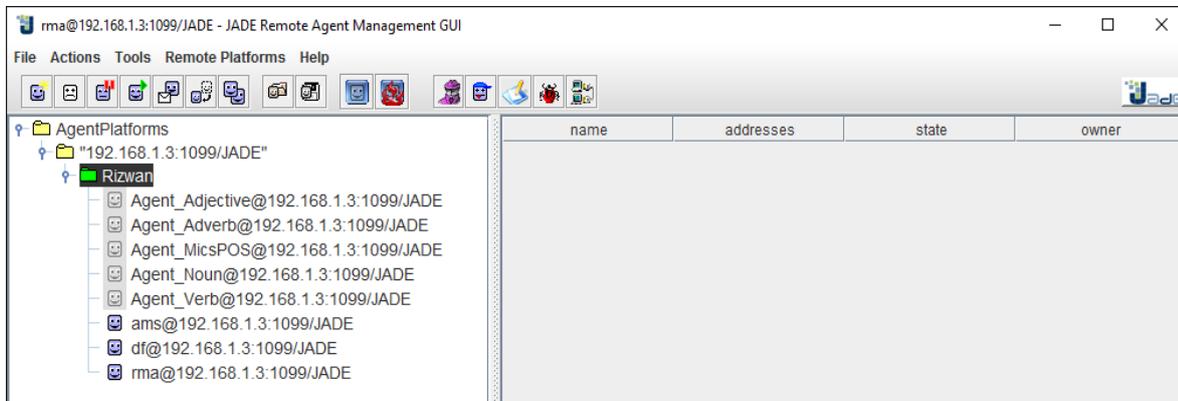


Figure 11. Agents creation and suspended state shown

There are basically five agents are created and their names are based on functionality / characteristics they have. Agents created are, Agent_Noun, Agent_Verb, Agent_Adverb, Agent_Adjective and Agent_MicsPOS. Agent_Noun is responsible to get noun type of tag and its token, initiate connection string with database that is “mydb” created in MySQL. The information is stored in relational DB is the core responsibility of Noun_Agent, if the type of tag received is of noun type. Same set of responsibilities are for verb, adverb, adjective and micsPOS type agents to receive their respective tag and token information and will store in database. After storing these information, agents are transformed their state from active to sleep until unless next token of same type is received to perform same action repeatedly. Agents created initially are in suspended mode and it is also being visible from Jade Remote Agent Management GUI screen shot provided below as **Figure 11**.

As shown in figure, all agents are in suspended mode initially. When a respective tag is received, concerned agent is resumed and forced to perform its action and then again go in suspension state. In **Figure 12**, an active agent of type Agent_Adverb is shown in GUI that is busy in storing information in database.

Storing information in DB and Assessing Complexity Measurements by JADE Agents

Information of tag and token / text / lexicon is collected by agent then connection is established between an agent and MySQL “mydb” database and record is inserted in database. Then agent would close the database connection session and then turned its active state again into suspended state. The agents are in interaction with relational database. Relational Database Entity relation model is presented below to have a bird eye review of mydb database logical design and modeling as presented in Figure 13; there are six entities involved with each other in ERD of mydb.

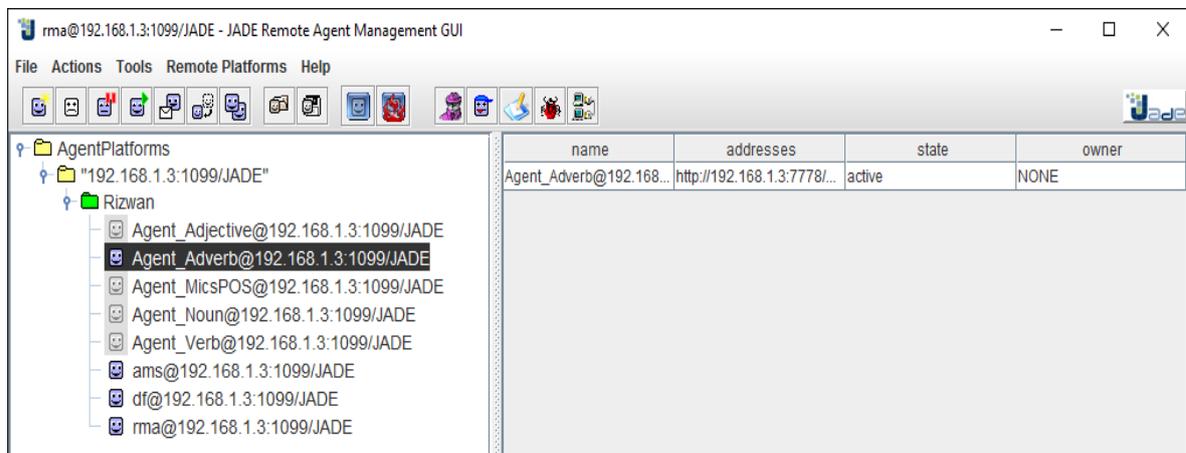


Figure 12. Agent_Adverb in active state

As there are four major types of Parts of Speech (POS), that are Noun, Verb, Adverb and Adjective. But few times a text has POS that does not fall in any of major category, to handle this event MicsPOS entity is framed. And finally, sentence entity is available to handle all types of sentences / text / strings that are passed to WordNet for tokenizing / getting tags. One to many relationships exist between Sentence and all other five entities. Number of attributes are also same for all entities except Sentence entity. Attributes include one attribute that is used to handle ID of the POS, second attribute is handling text / lexicon and finally tag that is returned by WordNet / Stanford Parser. While, sentence entity has only two attributes that are ID and text. When mapping techniques (Ibrahim et al., 2015) are used to transform ER diagram into relational schema, the resultant schema obtained for mydb database is provided in **Figure 13**.

Few additions that are in schema is discussed here. As one attribute is added in five tables handling POS that is Sentence ID. This is due to relational mapping of ER (Ibrahim et al., 2015), as Sentence table has relations of one to many with every other table in database, so a foreign key is added in every table that is also used as member of primary key with ID of respective table to form composite key. A diagram representing whole process of transformation is provided below as an extension is **Figure 14**.

DISCUSSION AND RESULTS

First as set of sixteen questions (as shown in Table 1) picked using Microsoft Students' QA Corpus, available online questionnaire free to be utilized, for passing through the proposed process of query assessment and transformation into RDF Graphs (Qayyum & Smith, 2015). Graphs are generated by two methods to show two different kinds of graph representation.

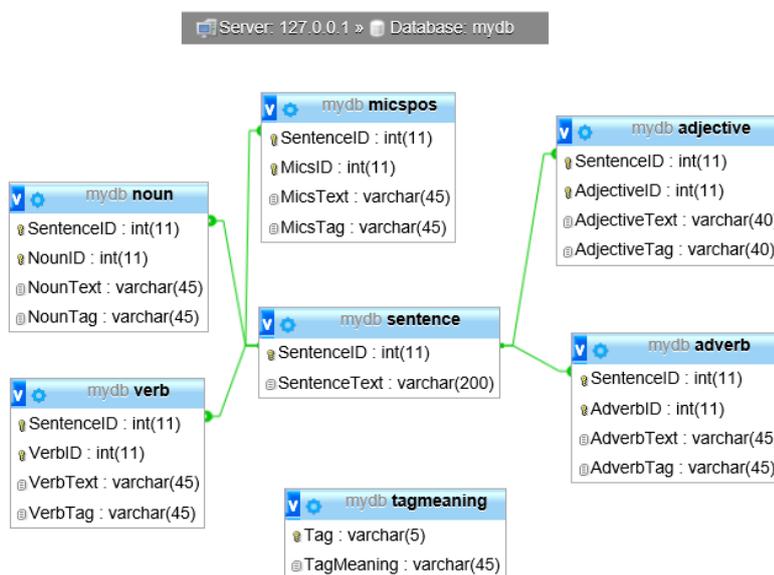


Figure 13. Relational Schema obtained for mydb after transforming its ERD

One is generated by online web validator and second is by a visualization tool to visualize RDF triple graph (Qayyum & Smith, 2015).

First way adopted to generate RDF triple graph is by using widely visualization tool that is RDF Gravity is shown in **Figure 15**. RDF Schema Graph. In RDF Gravity, RDF file is input in the package to generate triple graph and it also allows to search the graph by attribute / directory or values within graph along with save the graph options in many graphics file formats. RDF Gravity will only generate the graph if input file is validated means it follows the W3C standards. If it will not, graph of RDF will not be generated. Sample RDF triple graphs for RDF Schema and RDF that are generated by RDF gravity is attached below to show another representation of graph triples as shown in **Figure 16** and **Figure 17**. Close-up of RDF Store graph (a) Legends of the graph (b) Some of the queries linked with table (c) Remaining queries up close.

Agents after running received data for each part of each query through scoring function s as represented in Eq. (1) gained weightage of verb, noun, adverb, and adjective is represented after calculation passing through developed solution to the problem of assessment in gaining results shown in **Table 2**. Where in **Table 2**. Resultant scoring as TWeight assessment for corresponding query passed through transformation **process** gives impact of POS in understanding and complexity by observing the values. Larger the value of total weight (TWeight) complexity altogether increases and as a counter impact will be decrease in understandability of a question of student.

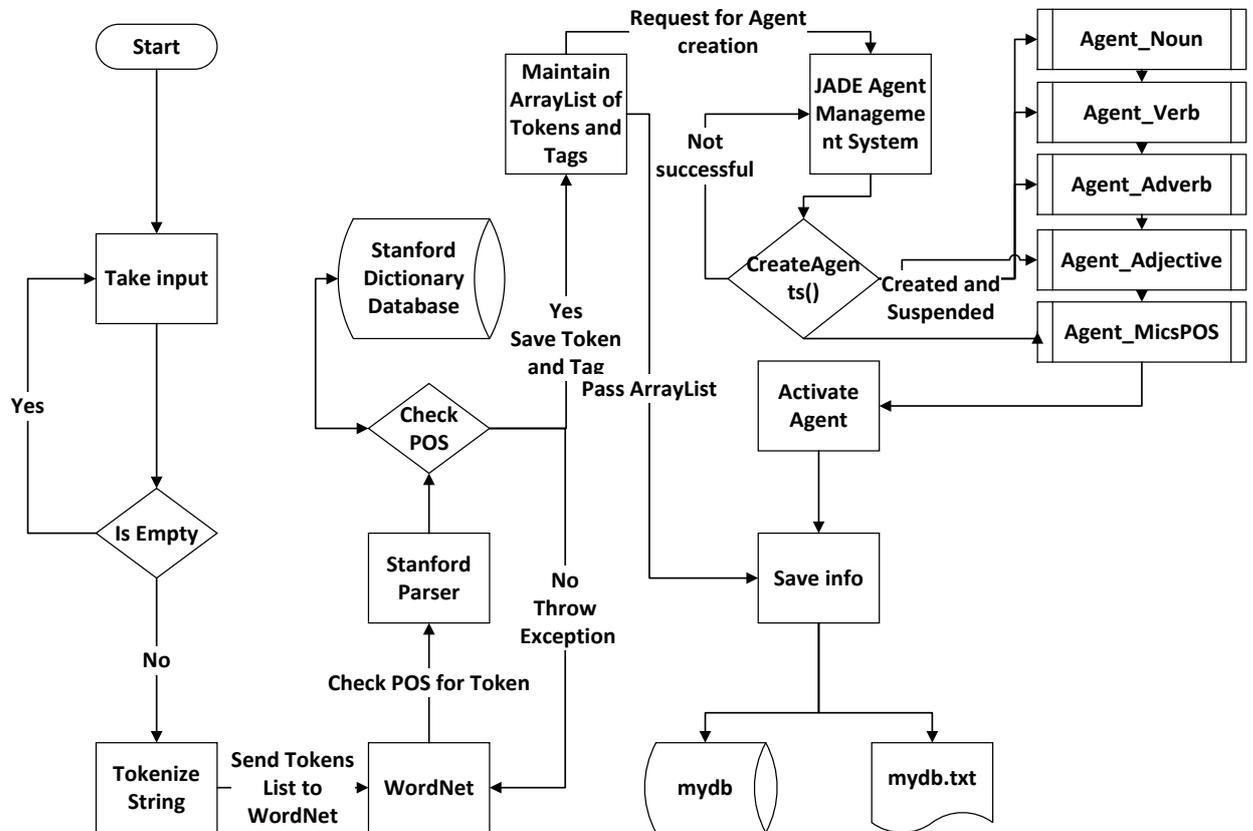


Figure 14. Complete flow diagram of Transformation from text to Relational DB

In Figure 18, Assessment Graph projecting TWeight as impact of each query weight wise focus can easily be seen represented in green color. In Java code, XML and XSD files are provided as input for validation in validator function. Validate XMLSchema function takes these files as input parameters to check files in single capacity and along with each other. Exception is thrown by java if any process encounter an error (s). Function returns a boolean value as zero or one.

Previously shown result and data representation shows the impact of assessing importance and complexity of the statement as used in a query can be used to save the effort needed when automate the questionnaire system involving better skill testing and formalization. Whereas, represents different Question Answering Systems (QAS) not able to assess or translate data into RDF. This shows how important it is to have transformation of available QAS compatible with intelligence for e-assessment to play its part in the field of education.

Validation Process

Method espoused to validate XML files (XSD and XML) is by Java code.

CONCLUSION AND FUTURE WORK

This study will benefit to get standard conversions of simple text to ontologies / RDF triples. Although a lot of work done in this regard, but a single solution is not found which is capable to provide a single go operation.

This study is hoped to become benchmark in NLP / Semantic Web field (M. Zhang, 2014). As our methods of converting relational DB into XML and RDF does not depends on database and hence any database with known data types can be catered by proposed system to get standard conversions devised by W3C. Parts of Speech (POS) for a query improves and enhances the chances of utilizing in students' querying skill of writing. Whereas, built process as trained agents in assessing data for measuring the complexity of the questionnaire as being plotted to test skill of the examinees. These agents are provided the access towards data format capable to use for intelligent assessment and strong linkage for future use (Kaula, 2016).

Table 1. Set of Sample picked questions from Microsoft Students' QA Corpus

SentID	Questions
1	Why does the grass turn brown and die during the winter if it still receives water, sun, and air?
2	What types of insects pose a problem to humans?
3	What happened when the Incas made a sacrifice?
4	What are the dangers of different kinds of drugs?
5	Does the government have anything to do with the problems in Israel?
6	What is chloroplast?
7	Where did rocks come from?
8	What was the Gettysburg address?
9	What is the exchange rate between the dollar and the Euro?
10	What is the name of the biggest castle in the world?
11	How did Michael Jordan earn his title as the MVP of all time?
12	How does the cable people hook cable up to every t.v. in the world?
13	How does second-hand smoke affect non-smokers?
14	Where can I find information about world war veterans?
15	What type of cloud produces rain?
16	How do homing pigeons know where to go and how to come home? How are they trained?

In this study, we devised a mechanism to facilitate in following step by step procedure towards semantic web creation. An adoptive mechanism is implemented to convert Stanford parser returned tags (SynSets in WordNet) into relational database.

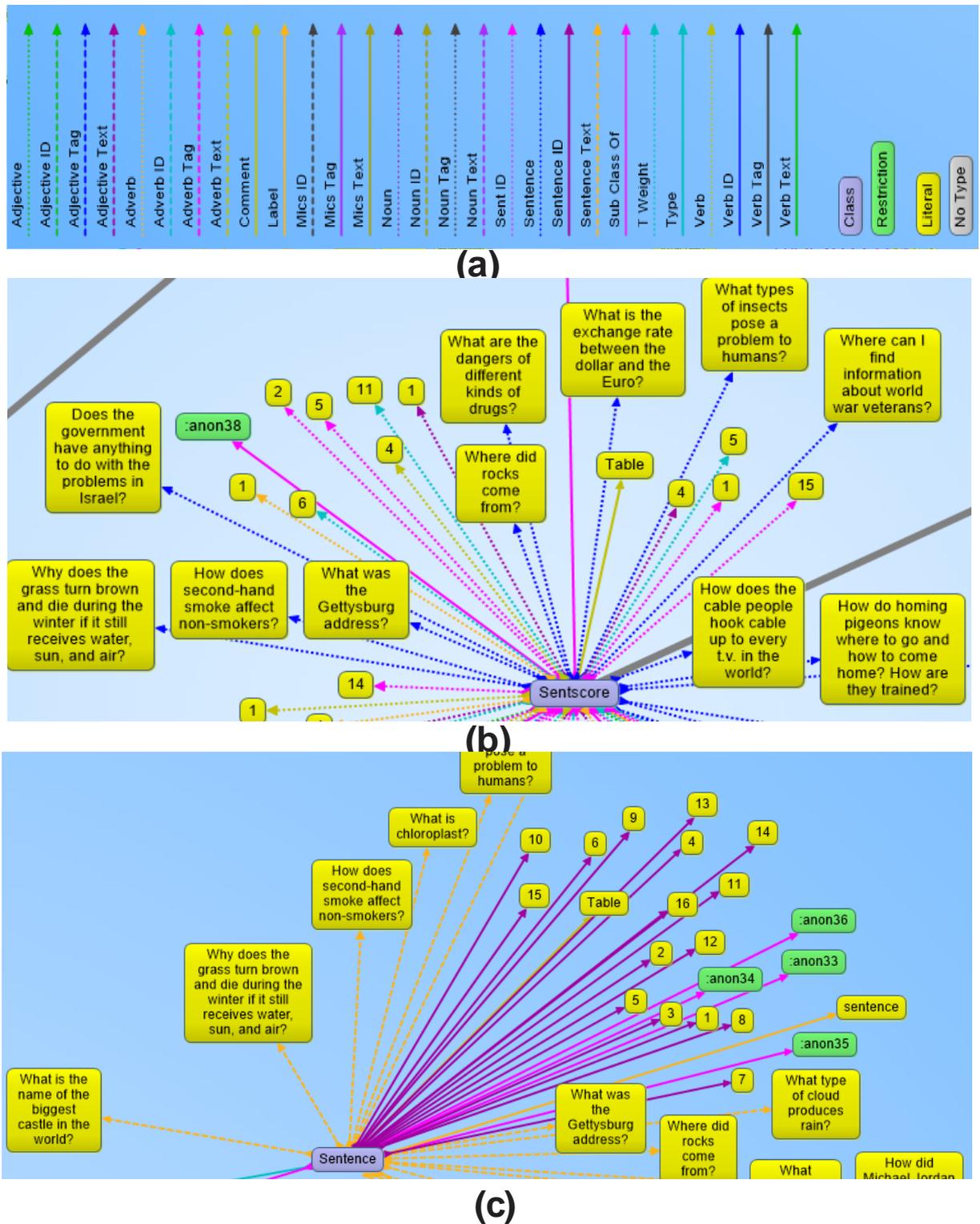


Figure 17. Close-up of RDF Store graph (a) Legends of the graph (b) Some of the queries linked with table (c) Remaining queries up close

Table 2. Resultant scoring as TWeight assessment for corresponding query passed through transformation process

SentID	Adverb	Adjective	Noun	Verb	TWeight
1	2	1	5	4	15
2	0	0	4	1	10
3	1	0	2	2	7
4	0	1	3	1	7
5	0	0	5	2	14
6	0	0	1	1	4
7	1	0	1	2	5
8	0	0	2	1	6
9	0	0	4	1	10
10	0	1	3	1	7
11	1	0	5	2	13
12	1	0	5	2	13
13	1	1	2	2	6
14	1	0	4	1	9
15	0	0	3	1	8
16	4	1	2	6	11

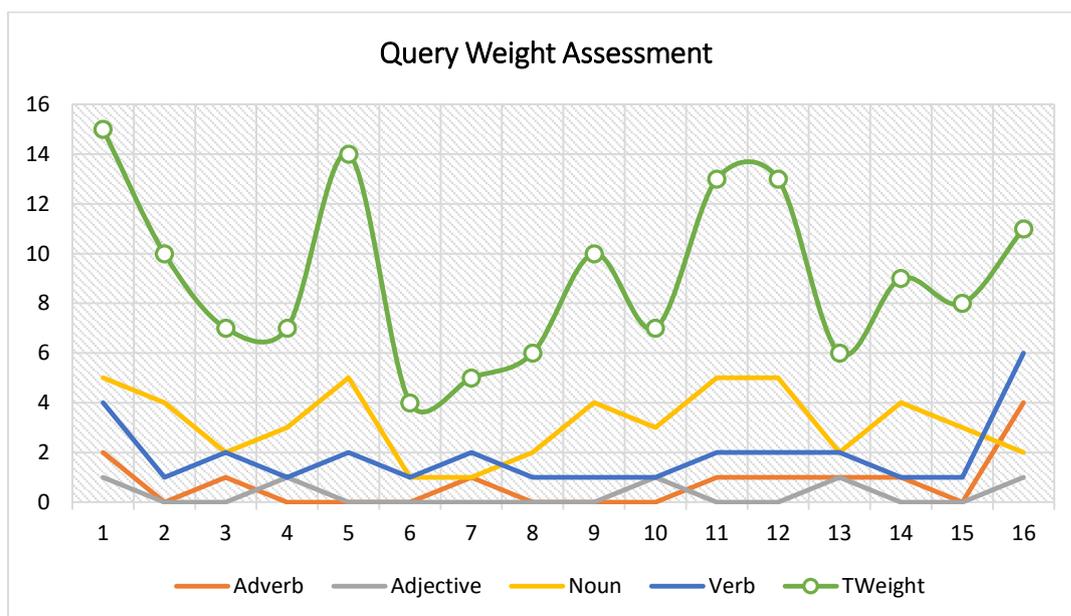


Figure 18. Assessment Graph projecting TWeight as impact of each query weight wise

Table 3. Comparing multiple Question Answering System

Authors, Year Published	Research on Question Answering System	Description
(Li, Zhang, Huang, Bai, & Liu, 2002; K. Zhang & Zhao, 2010)	Chinese Question Answering System	Classification of question which utilizes Word, Part of Speech (POS) tagging, the named entity (NE) and linguistics as an exemplary element for classifying the respective question.
(Ray, Singh, & Joshi, 2010)	Query Classification	Explained about a portion of the current methodologies for query classification and proposed another strategy through the utilization of the WordNet
(Stoyanchev, Song, & Lahti, 2008)	Document Retrieval for a Query Answering System	A procedure of the document retrieval for a query answering system, and adopted the process of utilization of named elements, nouns, verbs, and phrases as accurate match phrases in a query of document retrieval
(Moreda, Llorens, Saquete, & Palomar, 2011)	Utilizing Linguistics	Utilizing linguistics data as a part of Query Answering system, particularly in the step of answers extraction.
(Shen & Lapata, 2007)	Semantic Role	Determined the involvement of semantic roles to factoid type of query answering and demonstrated promising results.

ACKNOWLEDGEMENT

I would like to acknowledge my father, brother, friends and my colleagues for their support, suggestions and reviews, while conducting this study. I am also thankful to them on successful completion of this part of the project.

REFERENCES

- Amsler, R. A. (1984). Machine-readable dictionaries. *Annual Review of Information Science and Technology*, 19, 161-209.
- Bach, E., Chao, W., Kempson, R., Fernando, T., & Asher, N. (2012). The metaphysics of natural language (s). *Philosophy of Linguistics*, 14, 175.
- Bond, F., & Foster, R. (2013). *Linking and Extending an Open Multilingual Wordnet*. Paper presented at the ACL (1).
- Çalık, M. (2013). Effect of technology-embedded scientific inquiry on senior science student teachers' self-efficacy. *Eurasia Journal of Mathematics, Science & Technology Education*, 9(3), 223-232.
- Calzolari, N. (1991). Lexical databases and textual corpora: perspectives of integration for a Lexical Knowledge Base. *Lexical Acquisition: Exploiting on-line Resources to build a Lexicon*, Lawrence Erlbaum Associates, Hillsdale, NJ, 191-208.
- Castillo, J. G. D., Herrera, P. J. C., Carrillo, J. A. O., & McCalman, D. (2016). Raising the technological competence of high school science and mathematics teachers of Mexico through delivery of an online program. *International Journal of Technology, Policy and Management*, 16(2), 163-180.
- Diao, Y., & Franklin, M. J. (2016). High-Performance XML Message Brokering *Data Stream Management* (pp. 451-471): Springer.
- Eckle-Kohler, J., McCrae, J. P., & Chiarcos, C. (2015). LemonUby-A large, interlinked, syntactically-rich lexical resource for ontologies. *Semantic Web Journal-Special issue on Multilingual Linked Open Data*.
- Fallside, D. C., & Walmsley, P. (2004). XML schema part 0: primer second edition. *W3C recommendation*, 16.
- Gawrysiak, P., Ryzko, D., Więch, P., & Kozłowski, M. (2012). Retrieval and Management of Scientific Information from Heterogeneous Sources *Intelligent Tools for Building a Scientific Information Platform* (pp. 37-48): Springer.
- Ibrahim, W., Atif, Y., Shuaib, K., & Sampson, D. G. (2015). A Web-Based Course Assessment Tool with Direct Mapping to Student Outcomes. *Educational Technology & Society*, 18(2), 46-59.
- Kaula, R. (2016). Performance trends for operational intelligence through value chain model. *Journal of Advanced Computer Science & Technology*, 5(2), 34.
- Klyne, G., & Carroll, J. J. (2006). Resource description framework (RDF): Concepts and abstract syntax. *World Wide Web Consortium*.
- Li, S., Zhang, J., Huang, X., Bai, S., & Liu, Q. (2002). Semantic computation in a Chinese question-answering system. *Journal of Computer Science and Technology*, 17(6), 933-939.
- Lin, J.-W., Yen, M.-H., Liang, J.-C., Chiu, M.-H., & Guo, C.-J. (2016). Examining the Factors That Influence Students' Science Learning Processes and Their Learning Outcomes: 30 Years of Conceptual Change Research. *Eurasia Journal of Mathematics, Science & Technology Education*, 12(9), 2617-2646.
- Meng, L., Huang, R., & Gu, J. (2013). A review of semantic similarity measures in wordnet. *International Journal of Hybrid Information Technology*, 6(1), 1-12.

- Miller, G., & Fellbaum, C. (1998). Wordnet: An electronic lexical database: MIT Press Cambridge.
- Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39-41.
- Mitri, M. (2003). Applying tacit knowledge management techniques for performance assessment. *Computers & Education*, 41(2), 173-189.
- Moreda, P., Llorens, H., Saquete, E., & Palomar, M. (2011). Combining semantic information in question answering systems. *Information Processing & Management*, 47(6), 870-885.
- Qayyum, M. A., & Smith, D. (2015). Learning from Student Experiences for Online Assessment Tasks. *Information Research: An International Electronic Journal*, 20(2), n2.
- Ray, S. K., Singh, S., & Joshi, B. P. (2010). A semantic approach for question classification using WordNet and Wikipedia. *Pattern Recognition Letters*, 31(13), 1935-1943.
- Shen, D., & Lapata, M. (2007). *Using Semantic Roles to Improve Question Answering*. Paper presented at the EMNLP-CoNLL.
- Shi, W.-Z., He, X., Wang, Y., Fan, Z.-G., & Guo, L. (2016). PISA and TIMSS Science Score, Which Clock is More Accurate to Indicate National Science and Technology Competitiveness? *Eurasia Journal of Mathematics, Science & Technology Education*, 12(4), 965-974.
- Stoyanchev, S., Song, Y. C., & Lahti, W. (2008). *Exact phrases in information retrieval for question answering*. Paper presented at the Coling 2008: Proceedings of the 2nd workshop on Information Retrieval for Question Answering.
- Telepovska, H. (2014). Support of Database Skills Testing. *International Journal for Innovation Education and Research*, 2(5), 67-75.
- Zhang, K., & Zhao, J. (2010). *A Chinese question-answering system with question classification and answer clustering*. Paper presented at the Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on.
- Zhang, M. (2014). Who are interested in online science simulations? Tracking a trend of digital divide in Internet use. *Computers & Education*, 76, 205-214.

<http://www.ejmste.com>