

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/350526594>

State of Art for Semantic Analysis of Natural Language Processing

Article in *Qubahan Academic Journal* · March 2021

DOI: 10.48161/qaj.v1n2a40

CITATIONS

79

READS

555

5 authors, including:



Dastan Hussen Maulud

Ministry of Higher Education and Scientific Research

6 PUBLICATIONS 833 CITATIONS

[SEE PROFILE](#)



Subhi R. M. Zeebaree

DPU

248 PUBLICATIONS 9,140 CITATIONS

[SEE PROFILE](#)



Karwan Jacksi

University of Zakho

70 PUBLICATIONS 2,583 CITATIONS

[SEE PROFILE](#)



Mohammed A. M. Sadeeq

Duhok Polytechnic University

74 PUBLICATIONS 4,008 CITATIONS

[SEE PROFILE](#)

A State of Art for Semantic Analysis of Natural Language Processing

1st Dastan Hussen Maulud
IT Dept.
Duhok Polytechnic University
Duhok, Iraq
dastan.mawlud@mhe-krq.org

2nd Subhi R. M. Zeebaree
Culture Center
Duhok Polytechnic University
Duhok, Iraq
subhi.rafeeq@dpu.edu.krd

3rd Karwan Jacksi
Department of Computer Science
University of Zakho
Duhok, Iraq
karwan.Jacksi@uoz.edu.krd

4th Mohammed A. M.Sadeeq
Quality Assurance
Duhok Polytechnic University
Duhok, Iraq
mohammed.abdulrazzaq@dpu.edu.krd

<https://doi.org/10.48161/qaj.v1n2a44>

5th Karzan Hussein Sharif
Department of Computer Science
University of Human Development
Sulaimani, Iraq
karzan.hussein@uhd.edu.iq

Abstract—Semantic analysis is an essential feature of the Natural Language Processing (NLP) approach. It indicates, in the appropriate format, the context of a sentence or paragraph. Semantics is about language significance study. The vocabulary used conveys the importance of the subject because of the interrelationship between linguistic classes. In this article, semantic interpretation is carried out in the area of NLP. The findings suggest that the best-achieved accuracy of checked papers and those who relied on the Sentiment Analysis approach and the prediction error is minimal.

Keywords— Semantic Analysis, Natural Language Processing NLP, Ontology, Vocabulary, OWL.

I. INTRODUCTION

People access social media nowadays because of the intense tendency to share their viewpoints, and express their thoughts in a very casual fashion [1, 2]. It is important to consider the sentiment of the passage clearly [3]. Research into word meanings is required. Most Semantic Analysis systems have been developed only in English and European languages [4].

Natural language is the most potent form of communication. In computer science and artificial intelligence, voice is how machines communicate with humans (natural language). There is a wide range of algorithms, including sorting, clustering, text mining, and so on [5]. NLP is this widespread nowadays and involves Chabot and text clustering. Word2vec is a technique for processing, it is templates are indeed used in areas beyond NLP, but it is not something to be embarrassed by [6].

Sentiment analysis is a type of text classification where the subjectivity of the statements is essential. Opinion mining is the extraction of opinions from a large number

of sources [7]. Sentiment analysis and opinion mining are interchangeable throughout this document [8]. It collects opinion words through the NLP. Sentiment analysis is described as detecting the sentiment of people about a specific topic and its features. Opinion mining services are popular as people want to use others' advice to perform sensibly [9]. A big issue in data mining and NLP is subjective research on big data [10].

This paper is organized as follows: Introduction is presented in section I. Background theory in section II. Literature Review and experimental Comparison are presented in section III. Next, Discussion is explained in section IV. Section V describes the Conclusion.

II. BACKGROUND THEORY

A. Semantic Analysis

In NLP, it is studied how to use NLP strategies to users' emotions and decide what users are expressing through them [11]. Culture may affect this area differently, and This could be misinterpreted if it has taken too literally. "This new gadget is bad!" Although it was evident that the title alludes to the user's dislike of the gadget, the title might endorse the gadget to a particular age group of the community [12]. The sentiment analysis will determine the time at which you express your opinion [13]. To gather statements on a time axis can provide a better insight into peoples' feelings, Facebook and Twitter both provide challenges and opportunities for social movements. On the positive side, it allows people to express and express themselves freely [14]. The records can be carefully observed for a specified time to study trends. The data will provide a preponderance of the evidence that support the researcher's hypothesis. With the advent of the Internet, many fields of research have chosen to gather data from the web [15]. Companies like Google,

YouTube, and Amazon know how to customize the content for the customer's best interests [16]. Depending on objective metrics such as social media likes, the number of consumers, and sales. So, there is little data to study this topic. It is challenging because of a: using different languages on one topic or blog, b: using non-standard words that cannot be found in a dictionary, and c: using emoji and symbols. These are questions relevant to both the emotion and sentiment analysis domain [17].

There is a need to provide social scientists and psychiatrists the necessary vocabulary and tools to analyze the web's content and get the necessary data. This work intends to advance this area of study. The paper describes the following:

- Define NLP and explain its scope of application.
- To explain NLP in traditional and statistical ways are great things you have to accomplish.
- The concept of sentiment analysis gained significant attention in recent times.
- Tell the reader briefly about how NLP concepts and ideas can be applied to mental health issues and sentiment analysis.

B. Natural Language Processing

NLP is amongst the most complicated techniques in the world of artificial intelligence, text mining findings are inputs for NLP [17, 18]. NLP's capacity is that humans can speak words. It is the method of converting natural language output (spoken or written) into usable results. NLP is an exciting challenge because it requires computer and human interaction to implement it [15]. NLP is a field of computer studies concerned with studying and understanding the link between computers and the human language [19]. These tools help developers design practical tech applications. Several areas of interest have been established in NLP. Therefore, the core areas' most important activities concentrate on mining named persons, extracting knowledge from texts, translating texts between languages, summarizing written works, inferring answers by inference algorithms, and classifying and clustering papers [20]. It is common practice to discuss theoretical ideas in an academic setting. NLP is a subset of data science that uses dynamic mathematical computations and statistics. In the previous, naive Bayes, k-nearest neighbors, hidden Markov structures, conditional random fields (CRFs), decision trees, random forests, and support vector machines were used by great ML tools [21].

C. Ontology

Ontology [22, 23] is the philosophy of being, which is concerned with 'what is, i.e., the nature of reality and which is concerned with structural aspects of existence as such (Crotty, 1998) or whether it is possible to discover about the cosmos (Snape & Spencer, 2003) (Snape & Spencer, 2003). The SAGE Online Dictionary of Social Science Methods (2006) describes ontology as "a concept concerned with the existence of, and relationships between, different aspects of society such as social actors, cultural norms and social structures." Ontology is the study of reality and our convictions about things [24]. Ontology is the essence of the universe and the existence of truth, but

it also determines what can be said about it. Bryman (2008) defines social ontology as the philosophical significance ascribed to social actors' social world. They assume what people think is deciding whether there is a truth that exists separately from human conceptions and interpretations and whether there is an everyday reality or several context-specific ones [10, 25].

III. LITERATURE REVIEW

In 2020, Xiaolong Wang et al. [6], Uses an NLP and a combination word2vec to create an optimized LDA model which refers to a significance sample theory to extract subject matter words and use a cosine similarity in order to estimate repeat rate. For assessment of the NLP and text2vec combined. They are suitable for text processing and use the Latent Dirichlet assignment model to achieve the best possible results. Value sampling is used to train the model, which increases the accuracy and recall of the model. Cosine similarity is used to boost the outcome of the measurements by speeding up the operation.

In 2017, Kazar Okba et al. [10] authors suggested conversion methods using different sources of knowledge, hoping to improve the quality of translation. The uniqueness of the proposed approaches is that the source language can be used to understand the receiver language's context. Next, they compare the suggested method to accurate data and find that the system yields encouraging results. They employed a progressive and systematic approach towards natural language translation. To deal with the ambiguities, they developed a translation framework based on semantic analysis, which helped them detect and resolve conflation. The proposed method is tested using the OWLXML file that contains grammar, terminological and annotation details. They have created algorithms using proprietary tools to translate an unedited product.

In 2020, Artem A. Maksutov, et al.[18], Authors clarify non-standard details and then take this knowledge to create a more comprehensive account. An article explores how a graph database can help in identifying similarities between texts. It provides storage space for text content. The outline helps to define the relationships between expressions, words, and sentences. The algorithms are precise and rely on the correct input. Clean source text is critical for automatic plagiarism checks. A graph reveals both errors and the loss of information. The dependency parsing algorithm is central. Grammar-based algorithms than other grammatical algorithms better support some grammatical sentences. Improvements may also be extended to other systems. A specific parser may be named based on the complexity of the sentence it receives. An analysis technique is utilized to explain similarities between individuals found in a text. Graphs can view patterns in named entities.

In 2015, Monisha Kanakaraj et al. [26] proposed a system that employs text recognition techniques and other NLP technologies to clean data sets. The ensemble techniques are applied to form the predictive model. After analysis, the derived feature vectors are classified

according to the model, and alerts are shown. The program continues by analyzing data collection, data processing, and planning, classifying, and forecasting patterns. The central theory of the proposed approach is to boost NLP classification accuracy by using additional training data. The computer in question is set to gather data from Twitter and does NLP analysis on those Tweets. The text is tested for emotion using the author's classification system. Different evaluation techniques are analyzed to determine whether the evidence is accurate, harmful, or neutral. The ensemble system does better than traditional methods at classifying a dataset. Among those ensemble approaches, Extremely Randomized Trees turned out to be the most successful.

In 2015, Sally S. Ismail et al. [27], the paper explained the Rich Semantic Graph to Text module's solution method. They illustrate the steps taken and the mechanism used to produce the module. Two issues were established which could benefit from the course. Using mathematical grammatical analysis is a pioneering approach to Arabic Abstractive Text Summarization, and a primary area of study is the evaluation of various systems. The researchers have designed the RSG reduction module and are also designing the prototype for their design.

In 2016, Taimoor Hassan et al. [28], this paper provided a solution to the lack of thorough semantic information for optimal semantic understanding. The approach is to discover the significance of representations of software requirements and then determine the effect of semantic analysis. It was using a semantic structure specifically built to interpret and disambiguate the texts. The system's architecture is based on semantic technology that can be subsumed within the program's documentation and implementation. The presented methodology reveals that by adapting existing required environment ontologies and using them to construct performing experiments and methods that aid in knowledge management to the software system.

In 2016, Muhammad Taimoor Khan et al. [29], Describes the possible applications of sentiment analysis and the challenges it poses. Unless we simplify NLP, our approach will likely remain limited. In this area, the most pressing problem is the lack of fully-flexible and reproducible databases and measurement techniques. They suggested that the vocabulary could be improved to capture more of what people say and nearby. The knowledge mining industry crosses several fields and locations. Sentiment analysis is a widely used linguistic tool used by advertisers and social media firms. The authors utilized machine learning in order to solve the problem. They agreed that NLP has several unanswered questions that are beyond its power. These additional nuances supercede NLU tasks like inter-sentential harmony and word sense disambiguation. Sentiment analysis requires the analysis of only words used to convey a sentiment. Complex network analysis can yield beneficial results with arbitrary text. Knowledge-based systems are successful at procuring better results on unique assignments. Other machine learning solutions are

proposed, but all are not without fault. Knowledge obtained from the device should be reduced in the material dimension.

In 2017, Su-Zhen Wang et al. [30] presented a methodology that allowed a thorough analysis of semantic content extraction technologies. They proposed a new kind of webpage de-duplication algorithm that uses word vector distance and tf-idf. The thesis proposes a model for constructing a semantic cloud. A new web page elimination algorithm is proposed and significantly more effective than standard page recommendations based on TF-IDF and word vector distance. The proposed method has been shown to filter out non-repeating webpages based on webpages' similarity in the corpus from Hbase. The suggested solution provides for memory preservation of speech records.

In 2017, Prashant Gupta et al. [31], the framework for developing an Intelligent Querying Framework, allows the user to develop his own questions to be asked initially. The system has a module that converts English sentences to sentences that a SQL-like query will use—creating queries to answer consumer needs. As a consequence, it decreases the amount of study and makes it easier to do research. The proposed architecture allows non-experts to query a database by providing structure and ease of use. QS is very effective in generating complex queries, but it can also turn more difficult queries into routine ones. It is a simple method for accessing information about the commodity.

In 2017, Kittiphong Sengloiluean et al. [32], Adopted a semantic approach to answering the query by using DBpedia and WordNet. The study aimed to decide the best methods for solving questions. The paper included solutions to how named entities are classified, how named entities are equated, and how named entities can be addressed. The study tested the accuracy with which many replied to the questions. Using the TREC question collection, DBpedia, and the proposed solution, the proposed method obtained an overall F-measure score of 93.43%, an average recall of 94.15%, and an average precision of 92.73%.

In 2018, Amin Sleimi et al. [33], The paper presented an approach of semantic legal metadata so the reader can better understand and interpret the legal clauses. Metadata is essential for the identification of uniform legal requirements. There is a lack of literature on how to determine the consistency of metadata for formal specifications analysis. Furthermore, our ability to automate semantic legal metadata retrieval is not optimally applied; it does not take full advantage of Natural Language Processing. [The methods] they use to test each of the proposed semantic metadata types and address any disputes that occur. First, we do a qualitative review to classify the kinds of metadata best captured by tracking tools. The model recommends a suitable research model for legal requirements and offers accurate extraction rules for applicable metadata. Researchers analyzed extraction methods for different individual cases. Our review shows that the metadata is proper. The analysis is hopeful. When proposing penalties for

annotation inaccuracies, they rate consistency between 87.4 percent and 97.2 percent, and a recall between 85.5 percent and 94.9 percent.

In 2018, Lutfi Kerem Senel et al. [34], the approach relies on a novel word embedding-based method to build a word similarity model among 76 different languages worldwide. Researchers will use the atlas to help them narrow down their target word class of interest. A comparative analysis shows that the spatial similarities between two languages are similarities in word meanings. Research that compares the cross-lingual comparative analysis of 76 different languages was conducted. The results show that English is somewhat similar to other languages in terms of semantic meaning, but English is not closely related to other languages.

In 2018, Fatima Zait et al. [35], The authors promote semantic web and natural language processing methods to detect ambiguity in requirements as early as possible. Therefore, the viewer is likely to have more options and would be able to evaluate ambiguous phrases. They created a solution to explain potentially vague criteria and increase the awareness of how a system should work. Our POS tagger may disambiguate semantic concepts within user requirements and how these concepts are linked to others.

In 2018, Mattia Atzeni et al. [36] developed a method to translate natural language requests and instructions into computer code. To overcome the challenge, Computer Science researchers capitalize on Semantic Web technologies to develop CodeOntology, an open collaboration, a collaborative platform to make open-source code a first-class citizen on the Internet, where it can be interlinked with other resources. This approach uses Code Ontology to retrieve a variety of methods and code samples. These are graded and combined to transform a natural language specification into a Java source code. The results show that the solution should be comparable in efficiency to proprietary ones such as Wolfram Alpha.

In 2019, Jayden Sarker et al. [37] proposed a new algorithm to address questions. The POS marking of user responses and semantic sorting have been used to identify the documents. The proposed approach is a logical structure used to sort sentences based on logical principles. Information is gathered by comparing significant features in an attempt to discover the correct response. Compared to traditional syntactic approaches, the suggested approach improves the study of syntax by using lambda calculus. The proposed framework uses a generic search technique that achieved an average mean accuracy of 83%, surpassing the current approach by about 11%, and achieving a 95% accuracy for Yes/No questions. The method is simple, elegant, and more natural than systems that involve the use of computers. Furthermore, the analysis approach yields a higher degree of precision in answering those kinds of questions.

In 2019, Lakindu Gunasekara et al. [38], The paper introduced approaches using semantic technologies such as NLG and explores the pros and cons of such systems.

The study is based on the use of semantic natural language generation with chatbots, with the lower computational expense, and on the potential to reuse it for similar domains with less coding on natural language generation portion for small-scale level domains. The researchers build a novel architecture to support chatbot systems and indicate the value of understanding natural language. An ontological framework could be built to construct synthetic languages. The platform they provide acts as an API-based framework that offers an interface for text responses. A software API that can easily be integrated into any system can provide the basis for a framework to be used on many other domains.

In 2020, Shane Peeler et al. [39], This paper extended previous work on natural language query interfaces for online data by adding transitive verbs and prepositional sentences and an approach for accommodating queries that require chained complex verbs. This method had received good examples when a query interface was built to access a comprehensive repository of facts. Within a relational database, this property can define n-ary relations in which the property is transitive. Our company is developing a variety of applications that use vast quantities of data.

In 2020, Wendi Li [40] built a technique for clarifying the algorithm of natural language semantics understandings and a review appropriate for preprocessing technology answers to robotic questions, the term sense discrepancy, the semantic examination of reputation, etc., comparative assessments, and performance assessment. They have preprocessed the text using powerful algorithms that illustrate the way the words relate to each other in the text. However, the device is constrained in terms of precision and flexibility.

In 2020, Mengzhe Li et al. [41] provided a system to categorize 300 transgender people's social media posts, and they are a representation. Five separate machine learning models are employed to construct classifiers for sentiment analysis. They cluster the terms in a logistic regression to see if words are related to impact classification.



Fig 1: A word cloud with the most relevant 20 words is used in the article.

Figure 1 shows, top 20 words are visualized to a word cloud figure will see that their significance is distinctly visible. In addition to this, the comment segment discusses the sense of dysphoria. The annotations indicate a high degree of consensus (Cohen's Kappa = 0.8) in all three

Algorithm	Accuray	Averaged AUC
Naive Bayes	0.777	0.720
Random Forest	0.821	0.868
Support Vector Machine	0.832	0.842
Logistic Regression	0.829	0.851
K-Nearest Neighbour	0.764	0.767
Convolutional Neural Network (CNN)	0.861	0.834
Long Short-Term Memory (LSTM)	0.852	0.876

Fig 2: PERFORMANCE OF 5 MACHINE LEARNING CLASSIFIERS AND 2 DEEP NEURAL NETWORKS ON TRANSGENDER SENTIMENT ANALYSIS CLASSIFIER.

groups. Long Short-Term Memory (LSTM) is the most effective classification model, which has a precision of 0.85 and an overall accuracy of 0.876 as shown in figure 2.

In 2020, Mr. Machhirke Vinodkumar Sadhuram et al. [42], a proposal which constructs a theory in which the types of questions asked can be answered in discrete mathematics. NLP is used based on the data from different types of media that come in. NLP stands for natural language processing. It is used in the field of question answering (QA). Language processing is used to classify a sequence of responses to the question asked. The logic system demonstrates how to pose the question in an efficient way. Scientists used SQUAD to perform the

analysis. The test results' accuracy showed that the passage retrieval accuracy using TFIDF was as high as 69.69%, while students averaged 69.93%.

I. DISCUSSION

Almost all of the papers were observational studies that examined findings using a heterogeneous adaptive linguistic study methodology (2015 to 2020). Table 2 provides an overview of each analysis selected (reviewed) in this paper. The explanation includes the year of publication of each paper, technique/ method, the work's goal, methods used in the paper, actual results, and this research's objective. Semantic analysis techniques, Machine Learning algorithm built by pattern mining called "Machines Learning" the Ontology, Tokenization in which spaces separate words and phrases, and NER, which is an automatic extraction and Subject Modeling. Semantic analysis using a sophisticated algorithm for online text marking and language comparison. This system is chiefly used for translation, and other systems are predominantly used for text checking and question answering. As seen in Table 2. A research paper that uses advanced semantic techniques has a good level of accuracy and performance between 90% and 95%. The generalized searching technique has a factual accuracy in yes/no questions but has a marginal accuracy in more than one answer. Much of the articles use Python and Java to develop educational analysis methods. As given in Table 2, there are different techniques and methods used, fifty percentage of papers [26, 30, 34, 39-41] used the Sentiment Analysis technique, the findings demonstrate that sentiment analysis is an exciting way to test word duplication and developing question answering systems.

TABLE 1: AN OVERVIEW OF MOST RECENT NLP TECHNIQUES

Ref.	techniques	The goal of the work	Tool	Significant result	Accuracy
2020, [39]	Semantic Analysis	They are building interactive interfaces to annotate online documents.	Haskell X-SAIGA	The viability of accessing NL data using an event-based semantic-web triplestore was demonstrated, which contained thousands of facts.	92%
2020, [40]	Semantic Analysis	Described the natural language semantic processing of ECA algorithm and research progress on preprocessing technology.	Word2Vec, GloVe	XLNet on the GLUE dataset are significantly better than BERT	—
2020, [6]	Topic Modeling	Check for duplicate words in the language.	Python	improve the accuracy and recall rate	—
2020, [41]	Semantic Analysis	Transgender Community Sentiment Analysis From Social Media Data	Python	Among the 300 reddit comments, 72 are categorized as negative post, while 85 are annotated to be positive. The rest 48% are considered as neutral sentiment. The Cohen's Kappa score is over 0.8 across all classes	95%
2020, [18]	Tokenization	developing text preprocessing system	Python	All algorithms given in the solution produce correct results with a small amount of input data.	—
2020, [42]	Named Entity Recognition	Answering the questions that are asked by the users	Python	Overall Average over the entire dataset was 64% correct answer prediction	69.69%
2019, [37]	generalized searching	answering questions efficiently	bAbI-10k	Single Supporting Facts 90%, two Supporting Facts 90%, and three Supporting Facts 80%, Yes/No questions	95% for Yes/No questions,

				95%, List/Sets 60%.	overall 83%
2019, [38]	Natural language generation	Utilize NLP techniques with semantic technologies.	OWL, API	The recall rate of the system, based on impressions of 69 out of 11, is about 90%. A system that is accurate to within 0.862% is very close to precise.	86%
2018, [33]	automated extraction	Outlined a framework for the various types of legal metadata and then based it on traffic laws.	Python	Depending on whether or not a penalty will be levied for premature or inaccurate annotation, we can obtain precision 97.2% to 89.5%.	Between 85.5% and 94.9%.
2018, [34]	Semantic Analysis	Construct a score of 76 different languages' similarity based on a publicly available corpus.	t-SNE	English is in the top rank with nine other languages while Chinese is at the bottom.	—
2018, [35]	supervised machine learning	Accurately parse natural language, interpret ambiguous input.	POS tagger	Resolving coordination ambiguity, resolving lexical ambiguity, resolving semantic ambiguity are Yes, and low in terms of user interaction.	—
2018, [36]	Ontology	Translate a natural language command into an object-oriented program.	Java	Number of Questions 120, Processed Questions 116, Correct Answers 109, Precision 0.91	94%
2017, [10]	ontology	Translate a text from one language to another.	OWLXML, API, java	Highly helpful! Particularly on the disambiguation and understanding measures.	—
2017, [30]	Semantic Analysis	A cloud service corpus would establish a vocabulary for representing natural language. A new webpages removal algorithm is proposed.	de-duplication algorithm, PageRank, Simhash,	Actually repeated documents 50, Detected repeated documents 41, Detected correctly repeated documents 46, Recall rate 92%	95%
2017, [31]	Tokenization	A question answering system that converts questions in natural language.	Oracle, Microsoft SQL	Presents pragmatic ways to translate database queries into natural language that users can search.	92%
2017, [32]	Named Entity Recognition	Proposed how to analyze and answer questions.	DBpedia, WordNet	The F-measure of 93.43%, an average precision of 92.73%, and over 500 questions successfully administered	94.15%
2016, [28]	ontology	Formulation and analysis of natural language software requirements specifications.	OntoGen	Software Requirements: Sample 15, Correct 10, Incorrect 04, Missing 1, Rec % 0.66 7.	71%
2016, [29]	supervised machine learning	Evaluation techniques that use sentiment analysis and focuses on the need to address NLP open challenges. Without solving NLP challenges, ML techniques will not make significant advances.	java, notepad ++	Require training yes, use training data to classify No, probabilistic approach No, driving factor Centroid vector, similarity metric: Vector distance, strength Classify on vector distance, weakness Sensitive to noise, support for streaming data No.	Centroid 100%
2015, [26]	Semantic Analysis	Increase the accuracy of classifying with NLP by including Word Sense Disambiguation and semantics.	WordNet	Random Forest: F-measure rate of 91%, average precision of 92%, Decision Tree: F-measure rate of 89%, average precision of 89% over 7086 classified tweets.	Random Forest: 90%, D.T:90%

IV. CONCLUSION

In recent years, both academia and business companies have drawn the interest of the IoT. It is now an important component of our lives. It has the potential to link almost everything in our world to everything else. IoT systems are complex in design and have limited capacities for storage and retrieval. The integration of cloud computing with IoT would offer multiple advantages to numerous IoT applications. We have discussed in this article the state-of-the-art cloud infrastructure, including cloud features, architecture, and benefits. The topic also centered on

numerous technologies for IoT that would be expanded across the Cloud. Challenges of cloud IoT deployment and transparent problems are also discussed. In general, this paper's purpose was to include an overview to summarize up-to-date research contributions on cloud computing and the IoT and its applications in our environment and illustrate potential research directions and genuine concerns regarding the integration with the IoT of cloud computing.

REFERENCES

- [1] S. Y. A. Subhi R. M. Zeebaree, Mohammed A. M. Sadeeq, "Social Media Networks Security Threats, Risks and Recommendation: A Case Study in the Kurdistan Region," *International Journal of Innovation, Creativity and Change*, vol. 13, pp. 349-365, 2020/7/4 2020/7/4.
- [2] A. M. Abdulazeez, S. R. Zeebaree, and M. A. Sadeeq, "Design and Implementation of Electronic Student Affairs System," *Academic Journal of Nawroz University*, vol. 7, pp. 66-73, 2018.
- [3] R. R. Zebari, S. R. Zeebaree, and K. Jacksi, "Impact Analysis of HTTP and SYN Flood DDOS Attacks on Apache 2 and IIS 10.0 Web Servers," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 2018, pp. 156-161.
- [4] H. Jelodar, Y. Wang, M. Rabbani, G. Xiao, and R. Zhao, "A collaborative framework based for semantic patients-behavior analysis and highlight topics discovery of alcoholic beverages in online healthcare forums," *Journal of medical systems*, vol. 44, pp. 1-8, 2020.
- [5] K. Jacksi, N. Dimililer, and S. Zeebaree, "State of the art exploration systems for linked data: a review," *Int. J. Adv. Comput. Sci. Appl. IJACSA*, vol. 7, pp. 155-164, 2016.
- [6] X. Wang, X. Dong, and S. Chen, "Text Duplicated-checking Algorithm Implementation Based on Natural Language Semantic Analysis," in *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, 2020, pp. 732-735.
- [7] K. Jacksi, S. R. Zeebaree, and N. Dimililer, "LOD Explorer: Presenting the Web of Data," *Int. J. Adv. Comput. Sci. Appl. IJACSA*, vol. 9, 2018.
- [8] J. T. Rayz and V. Raskin, "Logic of natural language: Through the eyes of ontological semantics," in *2016 IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, 2016, pp. 511-515.
- [9] S. A. A.-Z. A. Selamat, "Electronic Learning Management System Based on Semantic Web Technology: A Review," *Int. J. Adv. Electron. Comput. Sci.*, vol. 4, pp. 1-6, 2017.
- [10] K. Okba, S. Hamza, B. Hind, A. Amira, and B. Samir, "Semantic natural language translation based on ontologies combination," in *2017 8th International Conference on Information Technology (ICIT)*, 2017, pp. 315-321.
- [11] R. Ibrahim, S. Zeebaree, and K. Jacksi, "Survey on Semantic Similarity Based on Document Clustering," *Adv. sci. technol. eng. syst. j.*, vol. 4, pp. 115-122, 2019.
- [12] K. Hood and P. K. Kuiper, "Improving student surveys with natural language processing," in *2018 Second IEEE International Conference on Robotic Computing (IRC)*, 2018, pp. 383-386.
- [13] N. Hadiya and N. Nanavati, "Indic SentiReview: Natural Language Processing based Sentiment Analysis on major Indian Languages," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, 2019, pp. 322-327.
- [14] K. Jacksi, N. Dimililer, and S. R. Zeebaree, "A survey of exploratory search systems based on LOD resources," 2015.
- [15] K. Jacksi, S. Zeebaree, and N. Dimililer, "Design and Implementation of LOD Explorer: A LOD Exploration and Visualization Model," *Journal of Applied Science and Technology Trends*, vol. 1, pp. 31-39, 2020.
- [16] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *ieee Computational intelligenCe magazine*, vol. 13, pp. 55-75, 2018.
- [17] X. Liao and Z. Zhu, "Classification of Natural Language Semantic Relations under Deep Learning," in *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 2020, pp. 1025-1027.
- [18] A. A. Maksutov, V. I. Zamyatovskiy, V. N. Vyunnikov, and A. V. Kutuzov, "Knowledge Base Collecting Using Natural Language Processing Algorithms," in *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIconRus)*, 2020, pp. 405-407.
- [19] A. Rokade, B. Patil, S. Rajani, S. Revandkar, and R. Shedge, "Automated grading system using natural language processing," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018, pp. 1123-1127.
- [20] M. J. Sadeeq and S. R. Zeebaree, "Semantic Search Engine Optimisation (SSEO) for Dynamic Websites: A Review," *International Journal of Science and Business*, vol. 5, pp. 148-158, 2021.
- [21] D. W. Otter, J. R. Medina, and J. K. Kalita, "A survey of the usages of deep learning for natural language processing," *IEEE Transactions on Neural Networks and Learning Systems*, 2020.
- [22] H. Al-Saadi, "Demystifying Ontology and Epistemology in research methods," *Research gate*, vol. 1, pp. 1-10, 2014.
- [23] A. Zeebaree, A. Adel, K. Jacksi, and A. Selamat, "Designing an ontology of E-learning system for duhok polytechnic university using protégé OWL tool," *J Adv Res Dyn Control Syst Vol*, vol. 11, pp. 24-37.
- [24] Z. S. Ageed, R. K. Ibrahim, and M. A. Sadeeq, "Unified Ontology Implementation of Cloud Computing for Distributed Systems," *Current Journal of Applied Science and Technology*, pp. 82-97, 2020.
- [25] A. AL-Zebari, S. Zeebaree, K. Jacksi, and A. Selamat, "ELMS-DPU ontology visualization with Protégé VOWL and Web VOWL," *Journal*

of *Advanced Research in Dynamic and Control Systems*, vol. 11, pp. 478-85, 2019.

- [26] M. Kanakaraj and R. M. R. Guddeti, "Performance analysis of Ensemble methods on Twitter sentiment analysis using NLP techniques," in *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015)*, 2015, pp. 169-170.
- [27] S. S. Ismail, M. Aref, and I. F. Moawad, "A model for generating Arabic text from semantic representation," in *2015 11th International Computer Engineering Conference (ICENCO)*, 2015, pp. 117-122.
- [28] T. Hassan, S. Hassan, M. A. Yar, and W. Younas, "Semantic analysis of natural language software requirement," in *2016 Sixth International Conference on Innovative Computing Technology (INTECH)*, 2016, pp. 459-463.
- [29] M. T. Khan, M. Durrani, A. Ali, I. Inayat, S. Khalid, and K. H. Khan, "Sentiment analysis and the complex natural language," *Complex Adaptive Systems Modeling*, vol. 4, pp. 1-19, 2016.
- [30] S.-Z. Wang, Q.-C. Zhang, and L. Zhang, "Natural language semantic corpus construction based on cloud service platform," in *2017 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2017, pp. 670-674.
- [31] P. Gupta, A. Goswami, S. Koul, and K. Sartape, "IQS-intelligent querying system using natural language processing," in *2017 international conference of electronics, communication and aerospace technology (ICECA)*, 2017, pp. 410-413.
- [32] K. Sengloileuan, N. Arch-int, S. Arch-int, and T. Thongkrau, "A semantic approach for question answering using DBpedia and WordNet," in *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2017, pp. 1-6.
- [33] A. Sleimi, N. Sannier, M. Sabetzadeh, L. Briand, and J. Dann, "Automated extraction of semantic legal metadata using natural language processing," in *2018 IEEE 26th International Requirements Engineering Conference (RE)*, 2018, pp. 124-135.
- [34] L. K. Şenel, İ. Utlu, V. Yücesoy, A. Koç, and T. Çukur, "Generating semantic similarity atlas for natural languages," in *2018 IEEE Spoken Language Technology Workshop (SLT)*, 2018, pp. 795-799.
- [35] F. Zait and N. Zarour, "Addressing lexical and semantic ambiguity in natural language requirements," in *2018 Fifth International Symposium on Innovation in Information and Communication Technology (ISIICT)*, 2018, pp. 1-7.
- [36] M. Atzeni and M. Atzori, "Translating natural language to code: an unsupervised ontology-based approach," in *2018 IEEE first international conference on artificial intelligence and knowledge engineering (AIKE)*, 2018, pp. 1-8.
- [37] J. Sarker, M. Billah, and M. Al Mamun, "Textual Question Answering for Semantic Parsing in Natural Language Processing," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp. 1-5.
- [38] L. Gunasekara and K. Vidanage, "UniOntBot: Semantic Natural Language Generation based API approach for Chatbot Communication," in *2019 National Information Technology Conference (NITC)*, 2019, pp. 1-8.
- [39] S. Peelar and R. Frost, "A Compositional Semantics for a Wide-Coverage Natural-Language Query Interface to a Semantic Web Triplestore," in *2020 IEEE 14th International Conference on Semantic Computing (ICSC)*, 2020, pp. 257-262.
- [40] W. Li, "Analysis of Semantic Comprehension Algorithms of Natural Language Based on Robot's Questions and Answers," in *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 2020, pp. 1021-1024.
- [41] M. Li, Y. Wang, Y. Zhao, and Z. Li, "Transgender Community Sentiment Analysis from Social Media Data: A Natural Language Processing Approach," *arXiv preprint arXiv:2010.13062*, 2020.
- [42] M. V. Sadhram and A. Soni, "Natural Language Processing based New Approach to Design Factoid Question Answering System," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 276-281.