TECHNISCHE UNIVERSITÄT
CHEMNITZ

Department of Computer Science

Data Management System Group

Systematic Literature Review

Design and Implementation of a Web-Based Software for OPC UA Communication Protocol Integrated to a Semantic Question Answering System

Orcun Oruc

Chemnitz, 22 November 2018

These pages intentionally left blank

Abstract

Industrial production is characterized by the increasing interconnection of machines and devices (Industrial Internet of Things, IIoT). The IEC 62541 (OPC UA) standard is one of the most important representatives in the heterogeneous view of available communication protocols, which is also recommended by the VDMA as preferable M2M communication. The OPC UA standard works according to the server-client principle, whereby a server usually represents one or more machines with a static data model. For the reason that this information model is not standardized, the structuring can be done individually for each server implementation. As a result, querying current or historical data requires knowledge of the structure.

The service set of the standard allows to load, query and edit the data model, but relationships between nodes can not be interpreted. This lack of semantic is recognized as one of the current challenges in "OPC Unified Architecture Pioneer of 4th Industrial (R) Evolution".

The goal is this thesis is to create a web service which is providing information with regards to OPC UA Protocol that connected to an internal Semantic Data Platform (Enilink). In order to connect with semantic web, a semantic question answering system is performed on a different layer of the web service and comparing it to other solutions which is currently existed in the market. A closed-domain question answering system provides information to an end-user with natural questions which is converting into SPARQL in order to use linked data source of Fraunhofer IWU.

This thesis investigates the linking of an OPC UA Information Model with Semantic Web technologies in a detailed way. A new query possibility for OPC UA sources is to be created, which allows to query structure information from a OPC UA Server and dynamic values from a remote server named eniLink with natural language commands. An upstream semantic query facility for OPC UA servers and their structured data allows a much faster and more flexible search of the data nodes.

In the practical part of this thesis is related to state of the art frameworks and technologies which is mainly used by the market. In order to provide best integration of question answering and Opc UA Web Service, Frontend and Backend technologies has been assessed with regards to loosely-coupled and asynchronous environment. A semantic question answering system has also been integrated to main web service software that can use linked data sources which is created by OPC UA Server.

Table of Contents

List of Abbreviations **vii**

1 Introduction **1**
1.1 Scope 1
1.2 Problem Definition 1
1.3 Limitations 1
1.4 Overview of Systematic Literature Review Methods 2
1.5 Strategy 3

2 Overview of studies **5**
2.1 OPC Unified Architecture 5
2.2 OPC UA Concept and Existing Applications 5
2.3 OPC UA Address Space Extensible Markup Language to Semantic Data 6
2.4 Open-Domain, Closed-Domain and Semantic Question Answering Systems 6
2.5 Existing Applications of Question Answering 11

3 Conclusions and Research Methods **13**
3.1 Answers of Research Questions 13

Bibliography **15**

Glossary **XIX**

List of Abbreviations

NLP	Natural Language Processing
HMI	Human Machine Interface
REST	Representational State Transfer
JSON	Javascript Object Notation
JWT	Json Web Token
OPC UA	OPC Unified Architecture
RQ	Research Question
SQL	Structured Query Language
SPARQL	SPARQL Protocol And Resource Description Framework Query Language
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network
COM	Component Object Model
NMT	Neural Machine Translation
HMM	Hidden Markov Model
KB	Knowledge Base

1 Introduction

1.1 Scope

Emerging technologies has been shaping day by day in Industry 4.0. Smart Factories which specifically focus on production technologies require a large amount of data processing in order to gain more benefits by implementing an optimization process. From top to down, all units of manufacturing process continuously transfer information to one another. In the lowest level, state-of-the-art technologies provides better sustainability and interoperability in the context of standardization which has enforced by Industry 4.0. OPC Unified Architecture (OPC UA) has released in 2008, is a platform independent service-oriented architecture that integrates all the functionality of the individual backward-compatible specifications into one extensible framework.

1.2 Problem Definition

Lowest communication protocol has been evolved nowadays in accordance with interoperability and sustainability. Automation industry requires a next generation standard such as OPC UA, which uses a scalable platforms, multiple security models, multiple transport layers and a sophisticated information model to allow the smallest dedicated controller [1]. OPC UA protocol mainly consists of two parts which OPC UA Client and Server, thus one can state a client-server communication framework.

1.3 Limitations

Although we commenced our research by setting our aims, due to the time limitation we should do our research with existing technologies. The tools has been chosen based on the requirements defined by Fraunhofer IWU. OPC UA Web Service part is not intended to extend full scaled software solution which can be deployable in a full manner. It might be a Minimum Viable Product where it can deploy in a smart factory environment. In the context of Semantic Question Answering, we will provide both open-domain and closed-domain question answering for the purpose of demo. However, our research will aim a limited implementation of closed-domain question answering.

1.4 Overview of Systematic Literature Review Methods

We used the OPC UA Guideline and Specifications to handle with all information resources relevant to the protocol as well as used research papers to get a deeper insight into Semantic Question Answering Systems.

The following research questions are deriving directly from the thesis objectives. We are interested in answering the following research questions (RQ):

RQ1: What are the existing solutions to problem?

RQ2: What is the strength of the evidence in support of the different solutions?

RQ3: How does the different solutions found by addressing RQ1 compare to each other with respect to approaches or methods?

RQ4: What implications will these findings have when creating algorithm?

RQ5: Is there a clear statement of the aim of the research?

RQ6: Is the study put into context of other studies and research?

RQ7: Are system or algorithmic design decision justified?

RQ8: Is the data set reproducible?

RQ9: Is the study algorithm reproducible?

RQ10: Is the experimental procedure thoroughly explained and reproducible?

RQ11: Is it clearly stated in the study which other algorithms the study’s algorithm(s) have been compared with?

RQ12: Are the performance metrics used in the study explained and justified?

RQ13: Are the test results thoroughly analysed.

RQ14: Does the test evidence support the findings presented?

1.5 Strategy

This section outlines the process for generating search terms, the strategy for searching, the database search, and the documentation for the search. After the strategy outlined, we identified key items for our search from prior experience with the subject area. Our main search terms "opc ua web service", "opc ua code generation", "open domain question answering", "closed domain question answering" and, "semantic question answering" respectively. We gathered the list of potential databases that has been suggested before. We searched following databases:

- ACM Digital Library
- IEEE Xplore
- ScienceDirect
- Springer Link
- Elsevier

Selection of studies for inclusion in the OPC UA Web Service and Question Answering Systems is a three-stage process: Initial selection of studies based upon title; selection of studies based upon reading the abstract; and the further selection of studies based upon reading the paper. Before study selection, duplicate papers identified by different database keyword searches were removed. The study data were stored in Mendeley Software which is used generally by research communities. For each study, we maintained some or all of the following information:

- Title
- Author(s)
- Conference (abbreviation and full time)
- Publication year
- Abstract

2 Overview of studies

2.1 OPC Unified Architecture

OPC Unified Architecture (OPC UA) was developed by taking into consideration the drawbacks of traditional OPC Classic Platform. Microsoft conceptualized OPC Framework with Component Object Model (COM). However, OPC Classic did not purpose for connecting end-user devices to existing protocol. To remedy this drawback, OPC UA Platform was standardized and released in 2008, is a platform independency service-oriented architecture that integrates all the functionality of the individual OPC Classic into one extensible framework [2]. In the master thesis, we will discuss the OPC UA Architecture in conjunction with definition of layers of schema, their usages and impact of our conclusion part.

2.2 OPC UA Concept and Existing Applications

[Salvatore Cavalieri, Marco Guiseppe Salafia, Marco Stefano Scroppo] [3] make an effort to enhance interoperability with web technologies to comply with Web Service using Representational State Transfer mechanism. The system purposed for the end user who does not know technical information about OPC UA Protocol and one can use by means of a web service which provides a token-based authentication[4]. The architecture of system [4] sets apart from other solution platform independence, loosely-coupling and monitoring of message-broker protocol such as MQTT and SignalR. According to the authors of paper, main differences between the proposal and the solutions available in literature is that all these solutions require, on the front-end side, the knowledge of communication services and the data modelling provided by OPC UA Standard [3]. Moreover, by holding a variety of username-password pair, or X509 certificates, this platform assure granting of the services. Services is also splitted up in a distinct manner such as SecureAccess, GetDataSources, ReadInfo, WriteInfo and Monitoring. [Tatu Paronen] [5] states that examines the requirements for the generic client and concerned with the specific technology choices, architecture, interfaces between different components, and other technical decisions related to the implementation [5]. The author organized the structure with three-tier architecture such as Presentation Layer, which consists of Server Connections, Address-Space Browser, Subscriptions and Monitored Item Services. It has composed with Service Layer and OPC UA Servers Layer concerning low-level communication as well.

2.3 OPC UA Address Space Extensible Markup Language to Semantic Data

There is a research gap between OPC UA Web Service and Semantic Question Answering. While searching literature so as to find OPC UA Information Model mapped into Semantic Data Link, it appears that researchers community has not been conducted sufficient surveys, researches or statements. [Eero Laukkanen] [6] has conducted a research about source code generations from OPC UA Servers to create a source code generator in Java by using instances of metamodels and templates with metadata [6]. [Badarinath Katti, Christiane Plociennik and Michael Schweitzer] [7] emphasized the absence of a standardized information model for machines and this could lead us to each vendor has produced on their own standardization. The authors introduced a concept of a semantic markup language called OWL-S to the OPC-UA protocol particularly [7]. In the process of research, a remarkable tool that has published by python-freeopcua is used for creating an extensible markup language of its address space.

2.4 Open-Domain, Closed-Domain and Semantic Question Answering Systems

When the point comes to a Semantic Question Answering, researchers mostly focus on algorithms how to transform from a natural language query to Structured Query Language (SQL) and SPARQL Protocol and Resource Description Framework Query Language (SPARQL). Principally, a question answering system is a system to answer question by human interaction with respect to information retrieval and natural language processing theories. Two types of main question answering systems has been analyzed in this work which are:

Open Domain Question Answering: The question can be asked to general type of data sources such as Dbpedia, Freebase and Wikidata. Not only specific topic can be asked but also a user may ask in any type of question so as to get an answer from a data source.

Closed Domain Question Answering: The question can be asked to restricted type of data source which has been defined in a restricted data source. A user may not ask any type of question so as to get an answer from a data source.

[Victor Zhang, Caiming Xiong, Richard Socher] [8] created a model to leverage the structure of SQL queries to significantly reduce the output space of generated queries by using reinforcement learning method [8]. SQL Queries used for a relational data source but the paper significantly important to understand how to transform a natural language query into a query language. Their algorithm Seq2SQL [8] requires a significant amount of question schemas, which need to be retrieved by the author. Seq2SQL takes an input as a question and columns of a table. It generates the corresponding SQL query during training phase, thus it executes against a database. The output space is representing via a Softmax Layer, which is a way of representation in a probabilistic view of the Machine Learning. However, the author [8] stated that the output space should be limited in order to do so. For the sake of brevity, Seq2SQL produces three basic structure for SQL query such as Aggregation classifier, SELECT column pointer and WHERE clause pointer. Aggregated input encoding such as COUNT, MIN, MAX is applied by Softmax Function to obtain the distribution over the set of possible aggregation operations [8]. The authors has tested their system WikiSQL which is a large crowd-sourced dataset for developing natural language interfaces for relational databases [8]. Towards solving the problem, the SQLNet Algorithm is an approach to employ a sequence-to-sequence style model . In particular, [Xiaojun Xum Chang Lie, Dawn Song] employ a sketch-based approach where the sketch contains a dependency graph so that one prediction performs by taking into consideration only the previous predictions that it depends on [9]. Main contributions of the SQLNet Algorithm express as below:

Avoiding “order-matters” problems in a sequence-to-sequence model and thus avoids necessity to employ a reinforcement learning algorithm [9].

A novel attention structure called column attention, and show that this helps to further boost the performance over a raw sequence-to-set model [9]. To handle with column names, Stanford CoreNLP Tokenizer used for parsing the sentence [10]. Shortly, each token represents a vector and inserts into bi-directional Long Short Term Memory. [F.F. Luz, M. Finger] used an LSTM encoder-decoder model capable of encoding natural language (English) and decoding query language (SPARQL) [11]. The algorithm uses a new type of encoder – decoder model that implements an alignment model in a feedforward neural network, which is concurrently trained with all of the components [11]. In the approach of [F.F. Luz, M.Finfer] is used to learn word vector representation and a LSTM neural network is used to encode natural language sentences and decode SPARQL Query [11]. Main contribution in algorithm part that they used a novel approach in Recurrent Neural Network (RNN) with Neural Attention. Mainly, the researchers attempted to solve the long-range dependency problem and the approach is separated into two steps, which is described as below:

- The first step is to find a good vector representation for the target language lexicon.
- The second step, [F.F. Luz, M. Finger] concerned with implementing and finding the settings so that the architecture can translate from natural language to SPARQL [11].

Most question answering systems translate questions into triples which are matched against the similarity metric. However, in many cases, triples do not represent a faithful representation of the semantic structure of the natural language question, with the result that complicated queries may not be answered [12]. [Christina Unger, Lorenz Bühnmann, Jens Lehnmann, Axel Cyrille Ngonga, Daniel Geber, Philipp Cimano] adopt the parsing and meaning construction mechanism of a question answering system which consists of two algorithm used Lexicalized Tree Adjoining Grammar and Underspecified Discourse Representation Theory respectively [12]. The authors used Stanford Part of Speech (POS) Tagger on-the-fly while parsing to sentences. Nouns entities are noun phrases and are usually modelled as resources, thus a lexical entry is built comprising a syntactic noun phrase representation together with a corresponding semantic representation containing resource slot [12]. Both algorithms serves one single purpose to identify light verbs (to be, to have and imperatives like give me), question words (what, which, how many, when, where) and other determiners (some, all, no , at least, more/less than, the most/least) , together with negation words, coordination and the like [12]. [Christina Unger et. al.] mainly contributed with this paper as below:

- Pure comparisons in natural language specific tools and broken into correct stages while creating SPARQL Query.
- In particular, an active learning method was implemented for end users to give feedbacks on the presented query results [12].
- Clearly defined goals of question answering system and designed for large-scale heterogeneous knowledge base.

[Tommaso Soru, Edgard Marx, Diego Moussalle, Gustavo Publio] have implemented an architecture Neural Machine Translation (NMT) approach that relies on three main components: a generator, a learner and an interpreter [13]. The generator takes inputs to create a query template. A query template is an alignment between a natural language query and its corresponding SPARQL Query[13]. At the final step, the learner obtains a natural language question to encode SPARQL queries. The authors have published a preliminary result in accordance with epochs on training phase. They plan to address current limitations by investigating how to generate domain-independent templates and minimize the burden on the end user [13]. [Shanthi Palaniappan, U.K. Sridevi, Jayasudha Subburaj] focused on a question answering system by improving the semantic similarity with Javascript Object Notation – Linked Data (JSON-LD) [14]. This work aims on different type of questions with different patterns which have to be matched with the ontology tree structure [14]. Chiefly, they specified an architecture in a clear manner to reach their goals. Main drawback of this architecture of ontology mapping part is not outlined. As for main contributions of this paper, we can take into consideration that the authors provide a clear-cut steps that entirely can carry out with a specific semantic data source such as JSON-LD .

Hidden Markov Model (HMM) is observing by [Cristina Giannone, Valentina Bellomaria and Roberto Basili] in conjunction with QALD-3 evaluation [15]. The authors clearly indicate how to parse a RDF Data Source with HMM method in order to create a precise SPARQL Query. The architecture comprises of HMM Initialization, Modeling, Decoding and Query Computation [15]. They have clearly indicated the combinations of sentences with dependency parse tree how affected from changes of order of items. Mainly, HMM Methods used for assigning properties of DBpedia source into their states and compilation of SPARQL queries [15].

[Question Answering in Restricted Domains] overview a main characteristic of question answering in restricted domains is the integration of domain-specific information that is either developed for question answering or that has been developed for other purposes [16]. [Diego Molla, Jose Luis Vicedo] defines main characteristics of question answering system over restricted domains as below [16]:

- It should be circumscribed
- It should be complex
- It should be practical

The authors has compared between open-domain and restricted-domain question answering by figuring out key points. According to their paper, they have defined three clear-cut subjects which are:

- The size of data
- Domain Context
- Resources
- Use of Domain-Specific Resources

[Sebastian Ferre] has published one of the detailed report that expresses natural language processing pitfall, essential points while consolidating SPARQL query and morphological definitions [17]. SQUALL is a solution for querying and updating RDF graphs by exploiting a controlled natural language which restricts grammar structures of a sentence in order to diminish complexities [17]. It has grouped all substantial features of a morphological language and pointed out what type of features in a natural language harnesses with regarding priorities and orders. Main contribution of SQUALL is categorizing ambiguities of natural languages and how turned out an advantage when using a controlled natural language [17]. The authors sketched a translation from their intermediate language to SPARQL to gain more accuracy with their system [17]

2.5 Existing Applications of Question Answering

In this section has been surveyed existing application without distinguishing of open-domain, closed-domain or restricted domain. AquaLog is a portable question-answering system which takes queries express in natural language and an ontology as input, and returns answers drawn from one or more knowledge bases (KB). By using NLP Platform, [Vanessa Lopez Garcia, Enrico Motta, Victoria Uren] aimed to choose an ontology and then ask queries with respect to its universe of discourse [18]. AquaLog uses a Relation Similarity Service which is purposed for dealing with the various sources of ambiguity. The authors gives explicitly a comparison between open domain question answering and AquaLog. Main contribution of AquaLog is capable of learning the user's jargon in order to improve quality of result [19]. The approach of NLP-Reduce is simple and does not use any complex natural language processing technologies [20]. Principally, it utilizes a synonym expansion and stemming techniques. NLP-Reduce contains a user interface, query generator and ontology mapper to send a query against open-domain resources. In the preliminary evaluation part the more NLP query is trained by NLP-Reduce, the easier it is getting precise results associated to a geographical dataset. [Danica Damljanovic, Milan Agatonovic and Hamish Cunningham] developed a system named FREyA, which combines syntactic parsing with the knowledge encoded in ontologies in order to reduce the customization effect [19]. In the research world, main contributions with FREyA can be broke into improving understanding of the question's semantic meaning, providing the concise answer to the user's question and communication the system's interpretation of the query to the user [19]. As the authors stated, the FREyA has been evaluated with Mooney Geoquery dataset in accordance with correctness, ranked suggestions and learning mechanism [19]. Bio2RDF is creating a network of coherent linked data and linking together with semantic life science resources [21]. Bio2RDF initiates a full text search in the graph resources, then the server returns a graph containing the searched literal with matched resources [21]. Main contribution as can been seen in the paper, Bio2RDF offers a reverse link service to determine whether a Uniform Resource Identifier (URI) has been used in other sources [21]. This feature is significantly important while creating a SPARQL request with prefixes that use URI tags. One of the brilliant research has been published by [Vanessa Lopez, Enrico Motta] and they implemented a solution named PowerAqua [22]. PowerAqua follows the system architecture of AquaLog, and address its main limitation to leverage future properties in order to develop a better mechanism. PowerAqua's paper explains that all bottlenecks of the architecture of AquaLog, drawbacks which should be fixed and how to use existed architecture of AquaLog in PowerAqua. Main contribution of the paper provides information how to implement a Question Answering System against multi-ontologies based on Information Retrieval Sources. GFMed is another Biomedical Question Answering System that we are going to discuss about contributions, benefits and drawbacks. GFMed used to create a Sparql request from natural language with a specific English-Biology Lexion by concreting Sparql Grammar with Sparql Lexicon [23]. English to SPARQL Algorithm has been clearly elaborated and supplied the question answering with a concrete SPARQL grammar built on top of a library [23]. The authors of the paper evaluated results in conjunction with Processed, Right, Partially, Recall, Precision and F-measure. Multilinguality (syntax, lexicon and inflections in 36 languages [23]) is one of the contribution of algorithm that could be taken into consideration. Pythia compositionally constructs meaning representations using a vocabulary aligned to the vocabulary of a given ontology [24]. Primary contribution of Pythia is generating a grammar from a given ontology (ontology-based grammar) to enrich a natural language query that shapes a formal query [24].

3 Conclusions and Research Methods

3.1 Answers of Research Questions

RQ1: What are the existing solutions to problems?

We have splitted up three parts of our research. In every chapter, we have been discussing what kind of existing solutions published as a research or open source environment. As for OPC UA Web Application there are mainly one software resides in Github. The project provides us backend functionalities partly and in our work it was used. For existing question answering systems, we have studied on Aqualog, NLP-Reduce, FREyA, Bio2RDF, PowerAqua, GFMed, Pythia. Not only the solutions have been studied, but also usage of their algorithms and external algorithms were discussed and inducted results to the thesis.

RQ2: What is the strength of the evidence in support of the different solutions?

While conducting our research on existing resources, we took into consideration of the quality in accordance with source code, reproducible published sources and obvious evaluation over results. A solution for OPC-UA Web Service like [4] provides all necessary information to reproduce software and evaluation again. Unlike [4], [Tatu Paronen] [5] did not provide a software solution that we can reproduce results as it is. For a generation of address space purpose, we have some tools which published as an open source library [25]. [Badarinath Katti, Christiane Plociennik and Michael Schweitzer] [7] created their standard called OWL-S to generalize on the OPC UA Protocol. This research does not contains a practical part, however, it is a meaningful research to draw a conclusion from our proposed way of thinking. For the Question Answering System, we have partly implemented solution that can be reproducible with evaluation data. Broadly speaking, the algorithm of solutions was more important guider than the solutions itself.

As a corollary,

RQ3: How does the different solutions found by addressing RQ1 compare to each other with respect to approaches or methods?

In RQ1, all types of applications were jotted down with regards to OPC UA Web Service and Question Answering.

RQ4: What implications will these findings have when creating algorithm?

RQ5: Is there a clear statement of the aim of the research?

RQ6: Is the study put into context of other studies and research?

RQ7: Are system or algorithmic design decision justified?

RQ8: Is the data set reproducible?

Data set is partly reproducible to embrace with our research. The publications about OPC UA Web Service provides architectural review rather than paying attention to data sets. Question Answering System's publications support with data set (training, test and validation data set) to utilize with machine learning algorithms.

RQ9: Is the study algorithm reproducible?

RQ10: Is the experimental procedure thoroughly explained and reproducible?

RQ11: Is it clearly stated in the study which other algorithms the study's algorithm have been compared with?

RQ12: Are the performance metrics used in the study explained and justified?

RQ13: Are the test results thoroughly analyzed?

RQ14: Does the test evidence support the findings presented`?

Bibliography

[1] Real Time Automation, "OPC UA Overview." [Online]. Available: https://www.rtaautomation.com/technologies/opcua/. [Accessed: 13-Nov-2018].

[2] Unified Architecture, "OPC Technologies," *OPC UA Foundation*. .

[3] S. Cavalieri, M. G. Salafia, and M. S. Scroppo, "Integrating OPC UA with web technologies to enhance interoperability," *Comput. Stand. Interfaces*, no. August 2017, 2018.

[4] S. Cavalieri, D. Di Stefano, M. G. Salafia, and M. S. Scroppo, "A web-based platform for OPC UA integration in IIoT environment," *IEEE Int. Conf. Emerg. Technol. Fact. Autom. ETFA*, pp. 1–6, 2018.

[5] T. Paronen, "A web-based monitoring system for the Industrial Internet," 2015.

[6] E. Laukkanen, "Java source code generation from OPC UA information models," p. 72, 2013.

[7] B. Katti, C. Plociennik, and M. Schweitzer, "SemOPC-UA: Introducing Semantics to OPC-UA Application Specific Methods," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1230–1236, 2018.

[8] V. Zhong, C. Xiong, and R. Socher, "Seq2SQL: Generating Structured Queries from Natural Language using Reinforcement Learning," pp. 1–12, 2017.

[9] X. Xu, C. Liu, and D. Song, "SQLNet: Generating Structured Queries From Natural Language Without Reinforcement Learning," pp. 1–13, 2017.

[10] C. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," *Proc. 52nd Annu. Meet. Assoc. Comput. Linguist. Syst. Demonstr.*, pp. 55–60, 2014.

[11] F. F. Luz and M. Finger, "Semantic Parsing Natural Language into SPARQL : Improving Target Language Representation with Neural Attention," vol. 1803.04329.

[12] C. Unger, L. Bühmann, J. Lehmann, A.-C. Ngonga Ngomo, D. Gerber, and P. Cimiano, "Template-based question answering over RDF data," *Proc. 21st Int.*

Conf. World Wide Web - WWW '12, p. 639, 2012.

[13] O. Vinyals, L. Kaiser, T. Koo, S. Petrov, I. Sutskever, and G. Hinton, "SPARQL as a Foreign Language," no. August, 2014.

[14] S. Palaniappan, U. K. Sridevi, and J. Subburaj, "Ontology based Question Answering system using JSON-LD for Closed Domain," vol. 119, no. 12, pp. 1969–1980, 2018.

[15] C. Giannone, V. Bellomaria, R. Basili, and I. Department of Enterprise Engineering, University of Rome Tor Vergata, Roma, "A HMM-based Approach to Question Answering against Linked Data," *Comput. Sci.*

[16] D. Mollá and J. L. Vicedo, "Question answering in restricted domains: An overview," *Comput. Linguist.*, vol. 33, no. 1, pp. 41–61, 2007.

[17] S. Ferré, "SQUALL: A controlled natural language for querying and updating RDF graphs," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7427 LNAI, pp. 11–25, 2012.

[18] V. Garcia and E. Motta, "AquaLog: An ontology-driven question answering system to interface the semantic web," *Proc. 2006 Conf. Hum. Lang. Technol. Conf. NAACL*, vol. Companion, no. June, pp. 269–272, 2006.

[19] D. Damljanovic, M. Agatonovic, and H. Cunningham, "Natural language interfaces to ontologies: Combining syntactic analysis and ontology-based lookup through the user interaction," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6088 LNCS, no. PART 1, pp. 106–120, 2010.

[20] E. Kaufmann, A. Bernstein, and L. Fischer, "NLP-Reduce: A 'naïve' but Domain-independent Natural Language Interface for Querying Ontologies," *4th Eur. Semant. Web Conf. ESWC 2007*, pp. 1–2, 2007.

[21] M.-A. Nolin *et al.*, "Bio2RDF Network Of Linked Data," *Web*, no. January, pp. 1–8, 2008.

[22] V. Lopez and E. Motta, "PowerAqua: Fishing the Semantic Web," vol. 7295, no. May 2014, 2012.

[23] A. Marginean, "GFMed:Question answering over biomedical linked data with Grammatical Framework," *Semant. Web*, vol. 8, no. 4, pp. 565–580, 2017.

[24] C. Unger and P. Cimiano, "Pythia: Compositional meaning construction for ontology-based question answering on the semantic web," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6716 LNCS, pp. 153–160, 2011.

[25] Pure Python OPC-UA Client and Server, "Free OPC-UA Library." [Online]. Available: https://github.com/FreeOpcUa/python-opcua. [Accessed: 22-Nov-2018].

Glossary

Terms of Glossary

Softmax Layer: It is a regression-based result to assign a multi-classification machine learning problem.

Machine Learning: It is the science of getting computers to act without being explicitly programmed.

Reinforcement Learning: It is ta type of Machine Learning Algorithms which allows software agents and machines to automatically determine the ideal behavior within a specific context, to maximize its performance.

Long Short Term Memory: LSTM is a unit of recurrent neural network which composed of a cell, an input gate, an output gate and a forget gate.

Bi-directional Long Short Term Memory: A bidirectional LSTM layer learns bidirectional long-term dependencies between time steps of time series or sequence data.

Word Vector Representation: It is a word vector in a row of real valued numbers

Recurrent Neural Network: It is a subclass of artificial neural network where connections between nodes from a directed graph or directed acyclic graph along a sequence.

Stanford CoreNLP Tokenization: It provides a tool that tokenizes a text snippet or blob of text

Stanford CoreNLP Part of Speech Tagger (POS Tagger): It provides a tool of which labels tokens with their part of speech tag

Neural Machine Translation: It is an end-to-end learning approach for automated translation, with the potential to overcome many of the weaknesses of conventional phrase-based translation systems.

Epoch: This term explains that is single pass through whole training dataset.