

**DYNAMIC ONTOLOGY BASED Q&A SYSTEM
FOR PANDEMIC SITUATIONS
CASE STUDY COVID – 19 PANDEMICS**

Subasinghe Arachchige Harsha Piyumi Subasinghe

(189397P)

Thesis submitted in partial fulfilment of the requirements for
the degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa
Sri Lanka

July 2022

Declaration

I declare that this dissertation does not incorporate, without acknowledgment, any material previously submitted for a degree or a Diploma in any University and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organization.

Name of the Student

S A H P Subasinghe

Signature of Student

Date:

Supervised by

Dr. A T P Silva

Signature of Supervisor(s)

Date:

Dedication

I dedicate my dissertation work to my family. A special feeling of gratitude to my loving parents, whose words of encouragement and push for tenacity ring in my ears.

Acknowledgement

I would first like to thank my supervisor, Dr. A T P Silva of the faculty of Information Technology at University of Moratuwa, for the patient guidance, encouragement, and advice she has provided. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly.

I would also like to gratefully acknowledge and thank my all colleagues and coworkers, at various stages of this work.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Abstract

In dynamic pandemic situations like covid-19, Many writeups, reviews, articles have been published every day. Rapidly updated data leads information overload, which make the public difficult to keep up with the latest data on pandemic situation. This paper focuses on introduce an efficient Q&A system for dynamic pandemic situation which help public to update with the real time data.

Several approaches including basic ontologies, expert knowledge base and linguistic knowledge have been used when model the knowledge base of Q&A systems. But these approaches are mainly based on experts' knowledge and mainly human interaction in knowledge acquisition, less handling of multimodal data, inefficient inferencing. Even though there are number of solutions which help public to update with the pandemic data, there are no fully automated real time updated systems. So, the intention is to introduce a fully automated multimodal data based real time updated system.

In order to archive this goal, fully automated dynamic ontology-based Q&A system was design, developed and evaluated for the pandemic situation like covid-19. Solution was design in such a way that users can enter question which is related to the covid-19 pandemic and retrieve a real time answer. Mainly the system is based on two modules as dynamic ontology module which use web scrapping for real time updated data extraction, process to map the changes in data and Q&A module which simplifies the questions into RDF triples based normal forms that effortlessly handled by database querying.

Evaluation of the system was conducted two ways by evaluation of the dynamic ontology module and evaluation of the question and answer module. In both evaluation processes time evaluation and precision has considered.

Keywords:

covid-19, dynamic ontology, Q&A, normal form, web scrapping

Contents

CHAPTER 1: INTRODUCTION	1
1.1 Prolegomena	1
1.2 Aim and objectives	2
1.2.1 Aim	2
1.2.2 Objectives	2
1.3 Background and Motivation.....	3
1.4 Problem in Brief.....	3
1.5 Dynamic Ontology-based Q&A Solution for Pandemic Situations.....	3
1.6 Structure of the Thesis	4
1.7 Summary	4
CHAPTER 2: EVOLUTION AND STATE OF ART OF Q&A SYSTEMS	5
2.1 Introduction.....	5
2.2 Early Development in Q&A Systems	6
2.3 Breakthrough in Q&A Systems	8
2.4 Modern Development in Q&A Systems	9
2.5 Challenges in Q&A Systems.....	10
2.6 Problem Definition.....	12
2.7 Summary	13
CHAPTER 3: TECHNOLOGY ADOPTED	14
3.1 Introduction.....	14
3.2 Question and Answering Systems.....	14
3.2.1 Preprocess	15
3.2.2 knowledge analysis	15
3.2.3 Answer processing	15
3.3 Knowledge acquisition in question answering system.....	15
3.4 Summary	16
CHAPTER 4: APPROACH.....	17
4.1 Introduction.....	17
4.2 Hypothesis.....	17
4.3 Input	17
4.4 Output	18

4.5	Process	19
4.6	Features	19
4.7	Users	20
4.8	Summary	20
CHAPTER 5: DESIGN.....		21
5.1	Introduction.....	21
5.2	Top Level Design.....	21
5.2.1	Dynamic Ontology Module	21
5.2.2	Question Answering Module	24
5.3	System Architecture.....	25
5.4	Summary	25
CHAPTER 6: IMPLEMENTATION.....		26
6.1	Introduction.....	26
6.2	Dynamic Ontology Implementation.....	26
6.2.1	Base ontology implementation.....	26
6.2.2	Implementation of dynamic behavior	27
6.3	Q&A module implementation.....	31
6.3.1	Direct answers retrieval by query processing	31
6.3.2	Answers retrieval by query processing for RDF triples mapped questions	32
6.4	Summary	34
CHAPTER 7: EVALUATION		35
7.1	Introduction.....	35
7.2	Evaluation of dynamic ontology module	35
7.3	Evaluation of question-and-answer module.....	35
7.4	Summary	36
CHAPTER 8: CONCLUSION AND FURTHER WORK.....		37
8.1	Introduction.....	37
8.2	Conclusion	37
8.3	Limitations and Further Work.....	37
8.4	Summary	38
REFERENCES		39
APPENDICES		42
Appendices A: Dynamic ontology module.....		42
A.1	Introduction.....	42

A.2 Base Ontology	42
A.3 Dynamic ontology Propagation.....	42
A.4 Dynamic ontology Population.....	44
Appendices B: Question and answer module	46
B.1 Introduction	46
B.2 Algorithm	46
B.3 Simplification of the dependency tree.....	46
B.4 Process of question answering system	47
Appendices C: Sample Codes	48
C.1 Introduction	48
C.2 Dynamic ontology module - Web scrapping.....	48
C.3 Dynamic ontology module - Differencing	49
C.4 Summarize the html documents	50
C.5 Ontology Population.....	51
C.6 Stanford dependency tree simplification	51
C.7 Generate normal form of the dependency tree	52

List of Figures

Figure 6: 5 Classification rule example for non-taxonomic 'has_efficient' relationship.....	30	
Figure 6: 7An example of a populated class	31	
Figure A: 1 Ontology structure	42	
Figure A: 2 Web Scrapping using python beautiful soup	42	
Figure A: 3 format of the outputted .csv file at now	43	
Figure A: 4 Format of the outputted .csv file before 12 hours.....	43	
Figure A: 5 identified changed data.....	44	
Figure A: 6 Automatic ontology population process	44	
Figure A: 7 Identification of instance candidate process.....	45	
Figure A: 8 Classifier construction process	45	
Figure A: 9 instances classification phase	45	
Figure B: 1 Possible normal form for What is the efficiency of Pfizer?.....	46	
Figure B: 2 Remove conj_or dependencies	Figure B: 3 : Remove amod dependencies	47
Figure B: 4 Block diagram of Question answering system.....	47	

List of Tables

Table 2: 1 Summary of benefits and challenges	10
Table 4: 1 Answer formats.....	18
Table 6: 1 Normalization rules for R0,R1,R2,R3,R4, R5	34