[EN-A-030] Design of Domain Specific Ontology to Avoid Miscommunication between Pilots and Ground Control Operator

(EIWAC 2017)

Dewan Abdullah*, +Hironao Takahashi**, Uzair Lakhani***

*Dewan Air Express
3A Lalazar, Beach Hotel Road,
Karachi, Pakistan
daf@airexpress.com.pk

**Department of Computer Sciences
DHA Suffa University, Phase-7 DHA,
Karachi, Pakistan
h.takahashi@dsu.edu.pk

***DTS Inc.
4-16-14-3004 Shibaura, MInato-ku,
Tokyo, Japan
uzair.dts@gmail.com

Abstract: Avoiding miscommunication between pilots and ground control operator in aviation industry is one of the highest priority tasks in order to maintain safety at takeoff and landing. The cause of mistake is human related error or in other words un-recognized conversation between them. To avoid misunderstanding in the communication we propose ontology based text support system. Although general purpose ontology is available but pilot and ground control operator communication is mission critical that requires high accuracy of understanding of conversation by text based guidance. In this paper we propose designing domain specific ontology to enhance the accuracy of understanding.

Keywords: domain specific ontology, pilot and ground operator communication, avoid misunderstanding

1. INTRODUCTION

Currently the aviation field has become more advanced. The aircraft operation has become more frequent. Due to this beginner and mid-level pilots may be required to operate the aircrafts which means high risk at landing and takeoff events. Another important human factor in aviation industry is that of the control tower operator. Table 1 lists five air accidents caused by human factors [8].

Avoiding miscommunication between pilots and ground control operator in aviation industry is one of the highest priority tasks in order to maintain safety at takeoff and landing. The cause of mistake is human related error or in other words un-recognized conversation between them. To avoid misunderstanding in the communication we propose ontology based text support system. Although general purpose ontology is available but pilot and ground control operator communication is mission critical that requires high accuracy of understanding of conversation by text based guidance. In this paper we propose designing domain specific ontology to enhance the accuracy of understanding.

Table 1 Air Accidents Due to Human Factors

Airline Name	Year
Asiana Flight 214	2013
German Typhoon and Learjet collision	2014
Tenerife Airport Disaster	1977
Air France Flight 447	2009
Boeing 737 Crash in Russia Flight U9 363	2013

We propose to develop separate domain specific ontology for each airport as each airport has different conditions due to its geographical location. For example the Jinnah

International Airport located at the port city of Karachi is windy at night and foggy in the morning. On the other hand the Skardu Airport is at a high altitude surrounded by mountain cliffs. Now we can observe that each airport has a different set of conditions and we propose to develop separate domain specific ontology for each. In our opinion the domain specific ontology based categorization will be better in comparison with the machine learning approach.

This domain specific ontology show the higher accuracy rate of understanding of conversation and it can help both that is pilots and ground control operator in their communication. The architecture of the overall system consists of the following: (1) Capturing voice of pilots and ground operator (2) Convert voice to text (3) Natural Language Processing (NLP) on text (4) Disambiguation (5) Inference engine that will utilize the proposed domain specific ontology.

The design of domain specific ontology may require searching the Internet or other text documents to determine the important terms in the aviation domain and the relationship between these terms. Domain experts can also provide good feedback for designing this domain specific ontology. Questionnaires can also be generated for getting the details required for the understanding of the domain. Please refer fig.1 below for better understanding.

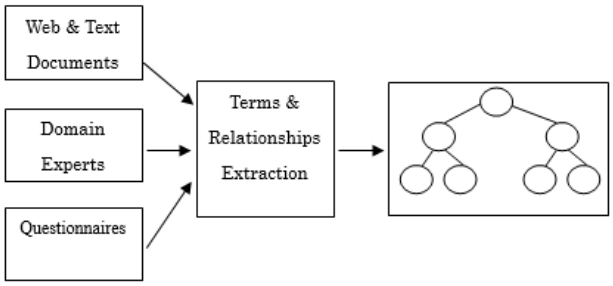

Figure 1 System Overview

2. RELATED WORK

The authors in [4] have developed a domain specific ontology using ConceptNet [5]. Essentially authors have separated the facts (triples) from the ConceptNet knowledgebase for a domain in order to create a domain specific ontology. On the other hand we have first created the facts or triples from sentences or unstructured text for creating the domain specific ontology.

The paper [6] discusses a process for domain ontology creation using textual resources i.e. unstructured text. The input text document can be in multiple formats like text file, word document, PDF file as well as HTML file. The paper also describes an evaluation mechanism for the proposed approach. It should be remembered that the proposed process of ontology creation is not 100% automated.

The OntoHarvester [7] system uses deep NLP techniques for designing a domain specific ontology from unstructured text. The input to the system is a small corpus of text documents. The system uses an initial (seed) ontology and then extracts new concepts and relationships from the provided text in order to augment the seed ontology for the development of the domain specific ontology. Only those new concepts are added in the domain specific ontology that are strongly related to the already present concepts in the seed ontology. In this way the ontology remains focused on a particular domain and noise is avoided.

3. TRIPLE EXTRACTION PROCESS FROM TEXT

Here we provide details regarding the triple extraction process from unstructured text. For this purpose some sample sentences have been used as shown in the table below. The first step is to perform Named Entities Recognition (NER) and Part of Speech (POS) tagging. For table 2 we have used the Stanford NLP [1] to perform NER [2] and POS tagging [3].

For table 3 NER column we remove the words from the NER column of table 2 that are not named entities. Similarly in the POS column of table 2 we remove named entities and stop words tokens to get the POS column for table 3.

Using table 3 we can get triples of subject, predicate and object as shown in table 4. First named entity becomes subject while the second named entity becomes object. The POS token of verb or noun becomes predicate.

We can extract triples in the way defined above from other sentences also. In this way we will have triples from unstructured text. Next these triples can be added into Protégé ontology editor to create the domain specific ontology.

Table 2 Stanford NLP NER & POS Tagging

Sentence	Stanford NLP	
	Named Entities Recognition	Part of Speech Tagging
Tokyo is the capital of Japan	<LOCATION>Tokyo</LOCATION > is the capital of <LOCATION>Japan</LOCATION>	Tokyo/NNP is/VBZ the/DT capital/NN of/IN Japan/NNP
Newton was born in the United Kingdom	<PERSON>Newton</PERSON> was born in the <LOCATION>United Kingdom</LOCATION>	Newton/NNP was/VBD born/VBN in/IN the/DT United/NNP Kingdom/NNP
Donald Trump is the president of the United States of America	<PERSON>Donald Trump</PERSON> is the president of the <LOCATION>United States of America</LOCATION>	Donald/NNP Trump/NNP is/VBZ the/DT president/NN of/IN the/DT United/NNP States/NNPS of/IN America/NNP
Peter Thomson works for the United Nations Organization	<PERSON>Peter Thomson</PERSON> works for the <ORGANIZATION>United Nations Organization</ORGANIZATION>	Peter/NNP Thomson/NNP works/VBZ for/IN the/DT United/NNP Nations/NNP Organization/NNP

Table 3 Simplified NER & POS Tagging

Sentence	**Stanford NLP**	
	Named Entities Recognition	**Part of Speech Tagging**
Tokyo is the capital of Japan	<LOCATION>Tokyo</LOCATION> <LOCATION>Japan</LOCATION>	capital/NN
Newton was born in the United Kingdom	<PERSON>Newton</PERSON> <LOCATION>United Kingdom</LOCATION>	born/VBN
Donald Trump is the president of the United States of America	<PERSON>Donald Trump</PERSON> <LOCATION>United States of America</LOCATION>	president/NN
Peter Thomson works for the United Nations Organization	<PERSON>Peter Thomson</PERSON> <ORGANIZATION>United Nations Organization</ORGANIZATION>	works/VBZ

Table 4 Sentence Triples

Sentence	**Triples**		
	Subject	**Predicate**	**Object**
Tokyo is the capital of Japan	Tokyo	capital	Japan
Newton was born in the United Kingdom	Newton	born	United Kingdom
Donald Trump is the president of the United States of America	Donald Trump	president	United States of America
Peter Thomson works for the United Nations Organization	Peter Thomson	works	United Nations Organization

4. FUTURE WORK

In the future we want to properly implement and evaluate the proposed system. In the system evaluation we have to test the latency also. If the system latency is too much then the system cannot be termed practical for which some solutions will need to be designed.

5. REFERENCES

[1] The Stanford NLP Group, https://nlp.stanford.edu/.

[2] Stanford NER, https://nlp.stanford.edu/software/CRF-NER.html.

[3] Stanford POS Tagger, https://nlp.stanford.edu/software/tagger.html.

[4] Ashish Sureka, Vikram Goyal, Denzil Correa, Anirban Mondal, "Generating Domain-Specific Ontology from Common-Sense Semantic Network for Target-Specific Sentiment Analysis," Computer and Information Science, Artificial Intelligence Paper.

[5] ConceptNet, http://conceptnet.io/.

[6] Mithun Balakrishna, Dan Moldovan, Marta Tatu, and Marian Olteanu, "Semi-automatic Domain Ontology Creation from Text Resources," In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC), May 2010.

[7] Hamid Mousavi, Deirdre Kerr, Markus Iseli and Carlo Zaniolo, "Harvesting Domain Specific Ontologies from Text," IEEE 8th International Conference on Semantic Computing, June 2014.

[8] 5 Aviation Accidents caused by Human Factors, https://ihf.co.uk/5-aviation-accidents-caused-by-hf/.