Informatics Institute of Technology

In collaboration with

University of Westminster, UK.

**Arena Reports Builder – Data Extracting System using NLP and
Question- based Analysis**

A dissertation by

Shanika M. Katukurunda (2014528)

Supervised By

Mr. Torin Wirasingha

ABSTRACT

Relational databases are highly used in the present world to store data in almost all companies and institutions. Storing these data is important to manage data within the company or the institution. Relational Database Management Systems (RDBMS) are made to interact with this data manually. There is a special language called Structured Query Language (SQL) introduced to communicate with this data using RDBMS. This language is mostly practised by Information Technology (IT) engineers. These stored data are taken to analyze information by the general employees with the help of IT engineers as the SQL non-experts cannot retrieve data by databases. Many types of research have been conducted to fill the gap between SQL non-experts and database using various technologies such as Natural Language Processing (NLP), question-based analyzing, ontologies, reinforcement learning. Every solution is based on one particular dataset and does not provide any general solution and also poor in accuracy level. Question-based analyzing systems provide more accurate information and some solutions are there in the production level as well, yet the questions are pre-defined by the developers. Therefore, Arena Reports Builder is proposed as a general solution for converting the natural language to SQL using NLP and Question-based analysis with better accuracy. In addition, Arena Reports Builder generates questions automatically through the system and does not need any database configurations except the database name. Arena Reports Builder achieve higher accuracy than the existing researches and it is adaptable for any MySQL database.

Keywords: Natural Language Processing (NLP), Structured Query Language (SQL), Question Answering, Databases Schema, Natural Language (NL)