

Informatics Institute of Technology

In collaboration with

University of Westminster, UK.

“Social Knowledge Mining from Question and
Answer Platforms”

A dissertation by

Shayne Darren Alexander Weerakoon

Supervised by

Mr. Guhanathan Poravi

Submitted in partial fulfilment of the requirements for the

B.Eng. in Software Engineering

Department of Computing

April 2018

Abstract

Knowledge Extraction is the process of getting structured data from unstructured or semi-structured sources. Knowledge extraction is a sub field under natural language processing and therefore most sources where knowledge extraction is applied to is text. There has been a reasonable amount of research applied, and many technologies have been borne out of this research, namely MUC and ACE. However, in recent years, with the rise of the semantic web and ontologies A new approach has made its way into the field of knowledge extraction – Ontology Based Information Extraction. Applying these technologies to the web has become a key effort in the past few years. This is since the web has changed from the time of Web 1.0 where the web was simply a bunch of static pages where user interaction was minimal. With the rise of web 2.0, the internet is no longer a medium to access static information. Users can now interact with pages and contribute their own efforts. Users can now share their own thoughts easily thus increasing the amount of user generated content. This has made the web ripe with knowledge, however not all this information can easily be accessed. This is has led to several years of research in natural language processing and information extraction. However, the amount of information being mined is significantly less than it could be. This paper proposes a system that mines social knowledge from social media platforms and Question and Answer sites.

Key Words:

Information Extraction, Social Knowledge Mining, Knowledge Bases