

# **Information Extraction System for IT Related Job**

**Pyae Pyae Aye; Ni Ni Hla**

*University of Computer Studies, Yangon*

Pyaelay991@gmail.com

## **Abstract**

*The current development of Web applications such as Blogs and Wikipedia enables users to easily create and disseminate their contents in the Web. As the contents on the Web are rapidly growing, the quantity of information is recently becoming more important in the Web. With the large quantity of information, the Web has now turned to the huge corpus that can be easily accessible using search engines, that opens new possibility to handle the vast relevant information and mine important structures. This study is how presented the information can be extracted from unstructured data by using keyword base searching method. Also, the system can store require data unstructured data as structure data. In this system DOM (Document Object Model) parser is used for parsing web document. For the case study, it is implemented IT related jobs information extraction system from HTML Web document. The proposed information extraction approach produces results for specified template and knowledge.*