2015 Research Internship

Faculty Name: Ebrahim Bagheri

Project Title: Semantic Similarity Analysis of Unstructured Web Content

1. Description of Project (Provide ½ page project description)

Inextweb is a cloud platform built at Ryerson University’s Laboratory for Systems, Software and Semantics (LS³) that combines the power of Semantic Analysis and Image Analysis to extract meaningful pieces of information from web pages and text. The system systematically crawls the web and automatically generates structured data from Web documents. This data represents a set of objects that exist in a web document and has until now only been understood when an actual person employs a web browser to render and display the web page. We are now applying this technology to other text-based documents such as patents. The structure of these documents is similar to a webpage with text and images. This enables inextweb to create an enhanced search platform for patents by extracting meaningful pieces of information and sort them based on their similarities and relevancy.

2. Responsibility of Student (Specify the duties and responsibilities of the student)

The student will be responsible for the following activities:

- Understand the Inextweb semantic analysis platform components and perform R&D work in extending its semantic linking capabilities;
- Develop semantic similarity measures that would be applied to the area of patent discovery and search
- Implement the technology in an App for the IOS or Android platforms

3. Specify Requirements (Please state any specific requirement of this position)

The student will be responsible for the following activities:

- Understand the Inextweb semantic analysis platform components and perform R&D work in extending its semantic linking capabilities;
- Develop semantic similarity measures that would be applied to the area of patent discovery and search
- Implement the technology in an App for the IOS or Android platforms