****

**Department of Computer Science and Engineering**

**Masters Thesis Defense**

**“WikiMatcher: Leveraging Wikipedia for Ontology Alignment”**

***By Brooke McCurdy***

**Tuesday, April 12, 2016 – 11:00 am**

**304 Russ**

**Committee:** **Drs. Michelle Cheatham, Advisor, Mateen Rizki, and TK Prasad**

**ABSTRACT:**

As the Semantic Web grows, so does the number of ontologies used to structure the data within it. Aligning these ontologies is critical to fully realizing the potential of the web. Previous work in ontology alignment has shown that even alignment systems utilizing basic string similarity metrics can produce useful matches. Researchers speculate that including semantic as well as syntactic information inherent in entity labels can further improve alignment results. This paper examines that hypothesis by exploring the utility of using Wikipedia as a source of semantic information. Various elements of Wikipedia are considered, including article content, page terms, and search snippets. The utility of each information source is analyzed and a composite system, WikiMatcher, is created based on this analysis. The performance of WikiMatcher is compared to that of a basic string-based alignment system on two established alignment benchmarks and two other real-world datasets. The extensive evaluation shows that although WikiMatcher performs similarly to that of the string metric overall, it is able to find many matches with no syntactic similarity between labels. This performance seems to be driven by Wikipedia's query resolution and page redirection system, rather than by the particular information from Wikipedia that is used to compare entities.