



**Divisions**

*Artificial Intelligence Division*

*Computational Systems and Technology*

*Informatics Systems Research*

*MOSIS*

*Networking and Cybersecurity Division*

June 26, 2019

Dear SWSA Award Committee,

It gives me great pleasure to write this letter to support of the application of Dr. Petar Ristoski for the SWSA Distinguished Dissertation Award.

I have been aware of Dr. Petar Ristoski's contributions for a number of years since he started publishing academic papers in the Semantic Web conferences. We had a project here at USC where we used his work on a RapidMiner extension and I read his papers with great interest at that time. I have since interacted with him at various conferences and seen him present his research in some of these conferences.

In my opinion, Dr. Ristoski has made extraordinary contributions to the field of Semantic Web and Linked Data through the work on his dissertation. Overall, his dissertation represents an impressive body of work. Almost every chapter in his thesis has been published as an article in a top-ranked peer-reviewed conference or journal. Even his related work chapter is an excellent survey paper that does a great job of surveying the relevant literature and relating the work to his own. There are numerous technical contributions in the thesis and below I review a few examples of Dr. Ristoski's exceptional research work.

Semantic Web knowledge graphs have been identified as a valuable source of background knowledge in many applications in the domain of natural language processing, information retrieval and data mining. However, in many cases consuming such data and integrating it into applications is a challenging process. Dr. Ristoski has made significant contributions in this area, especially in the process of automatic integration of valuable data from the published Linked Data on the Web for specific data mining applications. Dr. Ristoski developed the Linked Open Data RapidMiner extension, which extends one of the most used data mining tools, RapidMiner, to assist the user in consuming Linked Data as background knowledge in data mining. This tool assists the user in each step of the knowledge discovery process, i.e., selection and linking of data, consolidation of data from multiple sources, preprocessing and cleaning, feature generation, feature selection, and visualization of the data mining results. The tool has been internationally recognized and is heavily used in the Semantic Web and data mining community. We used this tool in one of our own research projects where we were working on mining statistical data sources to address issues of food security in underdeveloped parts of the world. The tool also won the open track of the



Semantic Web Challenge at the International Semantic Web Conference in 2014, which is one of the most prestigious challenges in the Semantic Web research area.

Another contribution of Dr. Ristoski's contributions in his dissertation is an outstanding approach for mining large knowledge graphs. Currently, more and more companies maintain and publish their data in the form of graphs, thus efficient approaches for mining such data are needed. The approach developed by Dr. Ristoski allows converting such large knowledge graphs into small numerical feature vectors. These feature vectors later can be used in many Data Mining, Natural Language Processing and Information Retrieval tasks, like entity retrieval, entity relatedness and entity recommendation.

In summary, I enthusiastically endorse Dr. Petar Ristoski's nomination for the SWSA Distinguished Dissertation Award.

Sincerely,

A handwritten signature in black ink, appearing to read 'Craig A. Knoblock', written in a cursive style.

Craig A. Knoblock  
Research Professor of Computer Science and Spatial Sciences,  
University of Southern California