June 26, 2018

Subject: **Nomination Letter for Kalpa Gunaratna's Doctoral Dissertation "Semantics-based Summarization of Entities in Knowledge Graphs" for SWSA Distinguished Dissertation Award**

Dear Members of the Awards Committee:

I am very happy to nominate Kalpa Gunaratna's doctoral dissertation titled: **"Semantics-based Summarization of Entities in Knowledge Graphs"** for the SWSA Distinguished Dissertation Award. I, along with Dr. Amit Sheth, was co-chair of his doctoral dissertation committee and co-advisor for his dissertation research. I confirm that Kalpa successfully defended his dissertation on April 19, 2017 and was awarded the degree of Doctor of Philosophy (Ph.D.) on April 29, 2017. Kalpa's dissertation primarily focused on using Semantic Web techniques and technologies for generating informative summaries of large entity descriptions in knowledge graphs for efficient consumption. The key findings have already been published in well-recognized conferences such as AAAI, ESWC, and IJCAI. I will now summarize his dissertation research below to justify this recommendation.

Knowledge graphs -- their representation, querying, and reasoning -- have come into prominence with increasing demand for intelligent, well-tailored applications. Knowledge-based computing that better mirror the intended semantics can have significant impact on the uptake of newer technologies by larger consumer base, even for carrying out common tasks such as Web search and content exploration. Kalpa's work addresses some of the key challenges for effective and efficient use and presentation of knowledge in knowledge graphs by enabling construction of semantics-based, high quality entity summaries. His work provides the theoretical foundation for semantics-based application development platforms, with wide-ranging practical applications -- from semantic search to intelligent processing, embodying the spirit that sparked the Semantic Web revolution. Besides publishing his research at excellent venues, he has also worked towards community building and sustained impact by initiating SemPre Workshop series at ESWC conferences in 2015 and 2016.

Kalpa's dissertation research addresses the challenging task of Web Knowledge Summarization to deal with information overload. Specifically, *his research covered three main topics extending the state-of-the-art results: (i) Generation of concise, comprehensive and diverse summaries of RDF documents describing entities, (ii) Generalization to cover both object properties and datatype properties by computing types for datatype property values, and (iii) Developing computationally tractable heuristics to create multi-entity summaries.*

Kalpa's FACeted Entity Summarization (FACES) approach makes use of conceptual grouping of facts to create an entity summary. Prior approaches to entity summarization focused on ranking mechanisms whereas FACES proposed to automatically identify different groups capturing abstract similarity of facts (both property and value) and then selecting facts to create "informative" summaries. Specifically, *FACES used an incremental hierarchical conceptual clustering method to generate concise, comprehensive and diverse summaries of RDF documents describing entities.* The novelty lies in incorporating semantic diversity in selecting and grouping facts while creating informative entity summaries. FACES outperformed other state-of-the-art approaches as demonstrated in the AAAI-2015 paper.

Kalpa's FACES-E approach is to extend FACES to consider both object property and datatype property for summarization. FACES summarizes only object property based facts as it requires explicit typing information, which the literals in the datatype properties lack. This lack of semantic information for literals limits its usefulness in practical applications other than those requiring lexical matching such as label matching. FACES-E approach first bridges this gap by computing appropriate semantic types for literals *in triples by developing heuristics for type inference employing light-weight NLP and then utilizes them to compute entity summaries, improving comprehensiveness.* This novel solution, discussed in the ESWC-2016 paper, enriches the literals semantically in order to make datatype properties on a par with object properties.

Kalpa's RElatedness based Multi-Entity Summarization (REMES) overcomes the (NP-Hard) computational challenges in summarizing RDF descriptions of multiple entities simultaneously. This task is challenging because it needs to address coherence issues. *REMES processes all the facts associated with a given entity collection to maximize (i) intra-entity summary diversity and importance, and (ii) inter-entity summary relatedness.* This can be viewed as a multi-objective optimization problem that can be solved by adapting quadratic knapsack solution method. This novel approach to multi-entity summarization, published in IJCAI-2017, is a significant achievement in terms of theoretical grounding and technical contribution.

Prior to his main dissertation research, Kalpa explored novel techniques for Ontology Development and Alignment. His work leveraged property extension as a robust means to ascertain property alignment between ontologies by adapting extensional equivalence criteria. Effectively, Kalpa's work generalized approaches that used string similarity metrics to align properties. This significant work was published in iSemantics-2013.

Kalpa's dissertation has not only contributed to Semantic Web area but also to broader application of Artificial Intelligence. Further, Kalpa has impacted applied research in academic/industrial research labs (Insight Center, NLM, and IBM) via patents, organized workshops to promote outreach (to ESWC and CIKM communities), and by serving in top-tier conference PC committees

Kno.e.sis - Ohio Center of Excellence in Knowledge Enabled Computing &
Ohio Center of Excellence in BioHealth Innovation
3640 Colonel Glenn Highway, Wright State University, Dayton, OH 45435-0001
http://knoesis.org

WRIGHT STATE
UNIVERSITY

(WWW and ISWC) while still being a graduate student. This level of performance is exceptional for a PhD student.

In summary, I believe Kalpa's dissertation represents a comprehensive contribution to the Semantic Web research on a number of aspects – breadth, quality, novelty, timeliness, and significance. I truly believe that this dissertation is of the quality that deserves recognition, and enthusiastically endorse for SWSA Distinguished Dissertation Award.

Sincerely,

Dr. Krishnaprasad Thirunarayan
Professor, Kno.e.sis Center
Department of Computer Science and Engineering
Wright State University
Dayton, OH-45435
(937) 775-5109 Voice
(937) 775-5133
t.k.prasad@wright.edu
http://knoesis.wright.edu/tkprasad