FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

FIZ Karlsruhe · Hermann-von-Helmholtz-Platz 1 · 76344 Eggenstein-Leopoldshafen

Prof. Dr. Harald Sack

Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
Germany

Phone +49 7247 808 - 198
Fax +49 7247 808 - 78198

Web: www.fiz-karlsruhe.de

Email: harald.sack@fiz-karlsruhe.de

Date: 2018-06-28

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

Dear Members of the Award Committee:

I am happy to support the nomination of Kalpa Gunaratna's dissertation for the SWSA Distinguished Dissertation Award as an experienced researcher in the Semantic Web community and an admirer of the entity-centric knowledge processing focus of this dissertation. I myself have been engaged in entity summarization and the processing of such data and hence I can strongly speak of the positive outcomes of this dissertation research as well as its high quality publications. Moreover, I also see a community of followers extending this line of research considering benchmark results established by this dissertation research. This exemplifies the timeliness, impact, and quality of the work. Let me justify my comments.

Kalpa Gunaratna's dissertation research consists of contributions covering single and multi-entity summarization on knowledge graphs and hence covers the breadth of possible types of entity summarization work. I will give my take on these approaches in the order they are implemented, which also makes a good story. The FACES approach, which was published in AAAI 2015 caught my attention because it solves a problem that is close to my interests. The novelty of this work lies in its incremental and conceptual clustering of entity descriptions to figure out relatedness of facts (RDF triples) at an abstract level in order to harvest diversity in the entity summaries. This in turn improves coverage of facts in size-constrained concise representations. This is a brilliant solution to diversify the summaries and avoid duplicate facts (that are sometimes semantically the same but lexically different) that make the summaries less interesting and useful to the knowledge consumers. In my own experience in entity summarization efforts (e.g., ISWC 2012 publication), I understand well the difficulty of getting a high quality ranking of facts and

successful removal of duplicate facts that are not only lexically similar but also semantically similar to each other. Therefore, automatic methods to diversify summaries by removing lexical and semantic duplicates is indeed important and exciting. Along the same line, I remember Kalpa's presentation of the extension called FACES-E to the FACES approach in ESWC 2016 as the general chair of the conference. While Kalpa was also busy with co-organizing a workshop on entity summarization (SUMPRE), I recall talking to him on his idea of typing literals to make them more useful in applications that need more semantic information. In fact, thinking of computing semantic types for literals is a wonderful idea and can have wide-scale use in many other applications such as data integration, alignment, and data profiling. It is also important to note that this is the first time a research proposal is made to compute semantic types for literals in knowledge graphs. Because of its novelty in the method and application area, there were interesting discussions in the ESWC community. I strongly believe that the approaches like this will make future developments of knowledge graphs more semantically rich.

Kalpa Gunaratna's multi-entity summarization approach called REMES was published in IJCAI 2017 and I find it interesting to go beyond single entity summarization mainly because it can be valuable in applications needing multi-entity understanding. The proposed approach is theoretically sound and outperforms the state-of-the-art in performing this complex task (computationally expensive problem). I believe this line of work will gain traction as we move towards needing more advanced query support for user interactions with the knowledge graphs involving more than one entity. One good example from my own experience is the effect of having a multi-entity summary/connection generation algorithm to provide advanced support in commercial tools like the Semantic Annotation Wordpress Plugin "refer"[1] introduced by the company Yovisto[2] that I co-founded. What is impressive about Kalpa's work is that it has potential practical use in tools that the companies like Yovisto investigate and develop by also having more man-power and resources. In fact, Kalpa has shown the high quality of his research work by achieving excellent results and publishing in a top-tier conference like IJCAI.

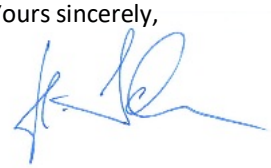
In my opinion, knowledge graphs play a significant role in intelligent data processing nowadays and it will only get more exciting in the coming years. As we have seen, knowledge graphs keep on growing and there will be more and more data to consume as we progress. To make use of this vast knowledge, I believe both application-independent and application-dependent efficient presentation methods are required. Therefore, Kalpa's dissertation work covering breadth (single and multi-entity summarization techniques) and depth (novel techniques and theoretically sound algorithms) of entity

[1] https://refer.cx/
[2] http://yovisto.com/en/technologies/

summarization is timely, of high quality, and made significant contributions to the knowledge graph and entity processing communities in the Semantic Web area. I can emphasize that the work in the area got considerable attention in the Semantic Web community as I observed the success of the entity summarization workshop series hosted at the ESWC conferences (I was the general chair of the ESWC 2016 and have seen the community participation in this workshop in the 2016 version). Further, in my opinion, Kalpa's publications in top-tier AI conferences make me believe that his work is appreciated at a broader audience (AI community) going beyond the Semantic Web community.

In summary, I am of the opinion that this dissertation deserves recognition for making significant and early steps in efficient knowledge presentation techniques (i.e., entity summarization) in the Semantic Web domain. I am wholeheartedly supporting the nomination of Kalpa Gunaratna's dissertation to the SWSA distinguished dissertation award and recommend the highest possible acknowledgement in this matter.

Yours sincerely,



Prof. Dr. Harald Sack
Vice President Information Service Engineering
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur GmbH

KIT Karlsruhe, Institute AIFB