# Kalpa Gunaratna

Senior Research Engineer at Samsung Research America
https://kalpagunaratna.github.io/

Phone: 937-279-7199
Email: <firstname><lastname> at gmail dot com
665 Clyde Avenue, Suite 601, Mountain View, CA 94043

**Interests:** *Summarization, ML/IR/AI for Data Analytics, NLP, Knowledge Graphs, Data Integration and Alignment, Linked Data, Semantic/Cognitive Computing, Knowledge Representation and Reasoning and Semantic Web.*

## Education

**Kno.e.sis Center, Wright State University** — Dayton, OH
*Ph.D. in Computer Science* — *2010 - 2017*

- Dissertation title: *Semantics-based Summarization of Entities in Knowledge Graphs*
  Advisors: Dr. Amit Sheth & Dr. Krishnaprasad Thirunarayan

**University of Colombo** — Colombo, Sri Lanka
*B.Sc. Computer Science, First Class Honors* — *2004 - 2008*

## Work Experience

**Samsung Research America, Mountain View CA** — Senior Research Engineer, June 2017 - present

Member of the Artificial Intelligence lab.

**Kno.e.sis Center, Wright State University, Dayton OH** — Research & Teaching Assistant, June 2010 - June 2017

Research work in various NSF, NIH, and AFRL funded projects. Teaching assistant for Advanced Algorithms & Design and Advanced Programming Languages courses.

**IBM Almaden Research Center, San Jose CA** — Research Intern, June 2015 - Aug. 2015

Developed intelligent cognitive assistant for work environment using adaptive online learning (ML) and NLP. The work involves processing and extracting information from conversational text and providing intelligent assistance to users. A US patent granted. Mentor: Dr. Hamid Motahari. Keywords: ML, NLP, Cognitive Computing

**National Library of Medicine (NLM), NIH, Bethesda MD** — Research Intern, June 2014 - Aug. 2014

Investigated using predication (triple/fact) similarity for document retrieval utilizing knowledge stored in knowledge graphs. Mentor: Dr. Olivier Bodenreider. Keywords: IR, NLP, Knowledge Graphs

**Insight Center, Ireland** — Research Intern, Jan. 2013 - May 2013

Researched on semantic matching of heterogeneous events and development of a rule parser for a custom language. Mentor: Dr. Edward Curry. Keywords: Semantic Matching, NLP, Rule Parsing, Linked Data

**CodeGen International, Sri Lanka (HQ: UK)** — Software Engineer, Sep. 2008 - June 2010

Developed real-world production-grade commercial applications. Conducted on-site client/developer training.

**Virtusa, Sri Lanka (HQ: USA)** — Software Development Intern, Jan. 2007 - June 2007

Worked in R&D department and IBM Mozart project for development of custom graphical components.

## Awards and Recognition

- Runner up prize winner of IBM Blockchain Hackathon at IBM Almaden Research Center 2015.
- AAAI travel grant 2015.
- Professional development grant from Wright State Graduate Students' Assembly 2015.
- NEWS highlight in international center at Wright State University 2015.
- National Best Quality Software Awards (NBQSA - Sri Lanka) finalist for undergraduate research project 2008.
- Second place in Blazing Innovation 2006 for Microsoft Imagine Cup organized by UCSC.

## Research Projects

- **Industry - Samsung Research America**
  - *Entities, Knowledge Graphs, NLP, and Deep Learning 2018 - present*:
    - Knowledge augmented on-device encoding.

    - Resource constraint on-device knowledge extraction.
    - Development of a knowledge graph platform for specific domains of interest. Capturing and processing entity descriptions.
    - Keywords: Knowledge Graphs, Context Understanding, Semantic Web, NLP, IR, ML and Deep Learning

- **Graduate Level**
  - *FACES project, 2014 - 2017*:
    - Using incremental hierarchical conceptual clustering to generate concise and comprehensive entity summaries.
    - Enriching knowledge graphs with type semantics using NLP and multiple constrained based optimization.
    - Keywords: Knowledge Graphs, Semantic Web, Linked Data, NLP, IR, ML
  - *PW11 project (Pratt & Whitney + Air Force Research Lab), 2014*:
    - Development of access control and data sharing platform using Semantic Web and knowledge representation techniques (lead investigator).
    - Keywords: Semantic Web, Access Control and Sharing
  - *Materials Genome project (Air Force Research Lab), 2013*:
    - Application of Linked Data and Semantic Web technologies for the Materials Science domain (President Obama's Material Genome Initiative). Initiated and provided expertise on development of domain vocabulary.
    - Automatic exploration of metal binding interactions in text using NLP (group leader).
    - Keywords: Semantic Web, Linked Data, NLP, Domain Vocabulary
  - *ESQuILO project (NSF), 2011 - 2012*:
    - Relationship identification and alignment in Linked Open Data. Partial and causal relationship analysis. Ontology alignment techniques for Linked Open Data.
    - Investigation of knowledge available on the Web (Wikipedia) for alignment tasks.
    - Keywords: Semantic Web, Linked Data, Statistical Analysis
  - *Hadoop project, 2010*:
    - Used Map-Reduce in the back-end for development of specialized tools in large scale NMR data processing.
    - Keywords: Distributed Computing, Hadoop, Matlab, Domain Specific Languages

- **Undergraduate Level**
  - *Undergraduate research project, 2008*:
    - Image processing and supervised deep learning (pattern recognition) techniques to detect currency notes.
    - Keywords: Deep Learning, Image Processing, Neural Networks

## Publications


[1] Li, J., Cheng, G., Liu, Q., Zhang, W., Kharlamov, E., **Gunaratna, K.** and Chen, H., 2020. Neural Entity Summarization with Joint Encoding and Weak Supervision. In proceedings of the 29th International Joint Conference on Artificial Intelligence (**IJCAI 2020**) (pp 1644-1650). DOI=https://doi.org/10.24963/ijcai.2020/228

[2] Li, S., Huang, Z., Cheng, G., Kharlamov, E. and **Gunaratna, K.**, 2020. Enriching Documents with Compact, Representative, Relevant Knowledge Graphs. In proceedings of the 29th International Joint Conference on Artificial Intelligence (**IJCAI 2020**) (pp. 1748-1754). DOI=https://doi.org/10.24963/ijcai.2020/242

[3] Liu, Q., Cheng, G., **Gunaratna, K.** and Qu, Y., 2020, May. ESBM: an entity summarization benchmark. In European Semantic Web Conference (**ESWC 2020**) (pp. 548-564). Springer, Cham. **Best resource paper nominee**

[4] **Gunaratna, K.** (2017) Semantics-based summarization of entities in knowledge graphs. (Doctoral dissertation, Wright State University)

[5] **Gunaratna, K.**, Yazdavar, A.H., Thirunarayan, K., Sheth, A. & Cheng, G. (2017). Relatedness-based multi-entity summarization. In proceedings of the 26th International Joint Conference on Artificial Intelligence (**IJCAI 2017**) (pp. 1060 - 1066). DOI=https://doi.org/10.24963/ijcai.2017/147

[6] Motahari, H., **Gunaratna, K.**, & Cappi, J. (2017). eAssistant: Cognitive Assistance for Auto-Triage of Actionable Conversations. In Proceedings of the 26th International World Wide Web Conference (**WWW 2017**).

[7] **Gunaratna, K.**, Thirunarayan, K., Sheth, A., & Cheng, G. (2016). Gleaning Types for Literals in RDF Triples with Application to Entity Summarization. In Proceedings of the 13th Extended Semantic Web Conference (**ESWC 2016**) (pp. 85-100). DOI=10.1007/978-3-319-34129-3_6

[8] **Gunaratna, K.**, Cheng, G., Thalhammer, A., & Liu, Q. (2016). Results of the 2016 ENtity Summarization Evaluation Campaign (ENSEC 2016). In Proceedings of the 2nd International Workshop on Summarizing and Presenting Entities and Ontologies (SumPre 2016), CEUR-WS, Vol-1605, 2016.

[9] **Gunaratna, K.**, Thirunarayan, K., & Sheth, A. (2015). FACES: Diversity-Aware Entity Summarization using Incremental Hierarchical Conceptual Clustering. 29th AAAI Conference on Artificial Intelligence (**AAAI 2015**), AAAI, 2015.

[10] **Gunaratna, K.**, Lalithsena, S., & Sheth, A. (2014). Alignment and Dataset Identification of Linked Data in Semantic Web. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 4, no. 2 (2014): 139-151. DOI=10.1002/widm.1121

[11] **Gunaratna, K.** (2014). Document Retrieval using Predication Similarity. Lister Hill National Center for Biomedical Communications, National Library of Medicine; 2014 (Technical Report).

[12] **Gunaratna, K.**, Thirunarayan, K., Jain, P., Sheth, A., & Wijeratne, S. (2013). A statistical and schema independent approach to identify equivalent properties on linked data. In Proceedings of the 9th International Conference on Semantic Systems (**Semantics 2013**) (pp. 33-40). ACM. DOI=10.1145/2506182.2506187

[13] **Gunaratna, K.**, Thirunarayan, K., & Sheth, A. (2013). Types of Property Pairs and Alignment on Linked Datasets A Preliminary Analysis. Proceedings of the I-SEMANTICS 2013 Posters & Demonstrations Track, 35.

[14] Hasan, S., **Gunaratna, K.**, Qin, Y., & Curry, E. (2013). Demo: Approximate Semantic Matching in the COLLIDER Event Processing Engine. In Proceedings of the 7th ACM international conference on Distributed event-based systems (**DEBS 2013**). ACM. DOI=10.1145/2488222.2489277

[15] **Gunaratna, K.**, Anderson, P., Ranabahu, A., & Sheth, A. (2010, November). A Study in Hadoop Streaming with Matlab for NMR data processing. In 2nd IEEE International Conference on Cloud Computing Technology and Science (**CloudCom 2010**), 2010 (pp. 786-789). IEEE.

[16] **Gunaratna, D. A. K. S.**, Kodikara, N. D., & Premaratne, H. L. (2008). ANN Based Currency Recognition System using Compressed Gray Scale and Application for Sri Lankan Currency Notes-SLCRec. In Proceedings of world academy of science, engineering and technology (Vol. 35, pp. 235-240).

**Edited Proceedings**

[1] Cheng, G., **Gunaratna, K.** and Wang, J., 2020, October. 3rd International Workshop on EntitY Retrieval and lEarning (EYRE 2020). In Proceedings of the 29th ACM International Conference on Information Knowledge Management (pp. 3533-3534).

[2] Cheng, G., **Gunaratna, K.** and Wang, J., 2019, November. EYRE 2019: 2nd International Workshop on EntitY REtrieval. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (pp. 2993-2994).

[3] Thalhammer, A., Cheng, G., **Gunaratna, K.** Proceedings of the 2nd International Workshop on Summarizing and Presenting Entities and Ontologies (SumPre 2016) co-located with the 13th Extended Semantic Web Conference (ESWC 2016), Greece, May 30, 2016. CEUR Workshop Proceedings 1605, CEUR-WS.org 2016.

[4] Cheng, G., **Gunaratna, K.**, Thalhammer, A., Paulheim, H., Voigt, M., Garca, R. Joint Proceedings of the 1st International Workshop on Summarizing and Presenting Entities and Ontologies and the 3rd International Workshop on Human Semantic Web Interfaces (SumPre 2015, HSWI 2015) co-located with the 12th Extended Semantic Web Conference (ESWC 2015), Portoroz, Slovenia, June 1, 2015. CEUR Workshop Proceedings 1556, CEUR-WS.org 2016.

**Tutorials**

[1] Cheng, G., **Gunaratna, K.** and Kharlamov, E., 2020, April. Entity Summarization in Knowledge Graphs: Algorithms, Evaluation, and Applications. In Companion Proceedings of the Web Conference 2020 (pp. 301-302).

[2] Cheng, G., **Gunaratna, K.** and Kharlamov, E., 2020, April. Entity Summarization in Knowledge Graphs: Algorithms, Evaluation, and Applications. ESWC 2020.

**Patents**

[1] Two patents filed in 2020 in the areas of knowledge graph embeddings and knowledge extraction.

[2] **Gunaratna, D.A.K.S.S.** & Nezhad, H.R.M., International Business Machines Corp, (2017). Adaptive learning of actionable statements in natural language conversation. U.S. Patent Application 14/995,116.

## Technical Skills

- **Application Development & Programming:** Ontology development (RDF/S, OWL, SPARQL), Semantic Web and Linked Data, Java (experienced), Python (experienced), C (basic), C++ (basic), Big Data(Map-Reduce, Hadoop), Web development (JSP, GWT, XML, XSD and XSL, Java Script and Ruby on Rails), NLP (Stanford CoreNLP, OpenNLP), Information Retrieval (Lucene), Machine Learning (scikit-learn, Weka), Deep Learning (Pytorch, Tensorflow), Code management (SVN, Git)
- **Databases:** Virtuoso triple store, Oracle, SQL Server and MySQL, NoSQL (basic)

## Selected Courses

Semantic Web and Knowledge Representation, Web 3.0, Advanced Databases, Programming Languages, Computational Complexity, Information Retrieval, Algorithms, Distributed Computing, Data Mining, Machine Learning (Coursera), Neural Networks for Machine Learning (Coursera), Writing in the Sciences (Stanford Online)

## Professional Service

[1] **Activities**

(a) Co-chair of the third international workshop on entity retrieval (EYRE '20) at CIKM 2020.

(b) Co-chair of the second international workshop on entity retrieval (EYRE '19) at CIKM 2019.

(c) Co-chair of the first international workshop on entity retrieval (EYRE '18) at CIKM 2018.

(d) Co-chair of the second international workshop on summarizing and presenting entities and ontologies (SumPre16) and ENSEC16 summarization challenge at ESWC 2016.

(e) Co-chair of the first international workshop on summarizing and presenting entities and ontologies (SumPre15) at ESWC 2015. Established the workshop.

(f) Member of the W3C Linked Data Platform Group (2012 - 2014).

[2] **PC Member**

- **NLP**: NAACL (2021), ACL (2020), EMNLP (2020)
- **Knowledge**: CIKM (2020), WWW (2018), ESWC(2016), and ISWC (2015, 2019, 2020)

[3] **Reviewer**

- KDD (2020), World Wide Web (WWW) (2012, 2013, 2014, 2015, 2017), International Semantic Web Conference (ISWC) (2012, 2013, 2014, 2016), Extended Semantic Web Conference (ESWC) (2013, 2014), Knowledge Engineering and Knowledge Management (EKAW) (2012), International Journal on Semantic Web and Information Systems (IJSWIS), TKDE, Semantic Web Journal.

## References

Available upon request.