

# Raghava Mutharaju

Data Semantics (DaSe) Lab,  
Department of Computer Science,  
3640 Colonel Glenn Highway,  
Wright State University, Dayton, OH.

**Email:** mutharaju.2@wright.edu

**Phone:** 937-768-2340

**Homepage:** <http://raghavam.github.io/>

**Google Scholar:** <https://goo.gl/muzqYd>

## Summary

- PhD student with experience in building large scale systems, ontology modeling, RDF query processing and developing Semantic Web applications.
- Industry experience of 3 years and 8 months including research internships at IBM T.J. Watson Research Center, IBM Research Ireland, Alacatel-Lucent Bell Labs, Xerox Research Center.
- Published 20+ research papers with 180+ citations.
- Co-organized tutorials at IJCAI 2016, AAAI 2015 and ISWC 2014.
- Workshop co-chair at ISWC 2015.
- Awarded Amazon AWS and Microsoft Azure research grants.

## Research Interests

- Large Scale Computing, Ontology Reasoning, Knowledge Representation, RDF Query Processing, Semantic Web Applications, Knowledge Graph, and Big Data.

## Education

- **Wright State University** Dayton, USA  
*Doctor of Philosophy (PhD) in Computer Science and Engineering* *May 2016 (Expected)*
  - Dissertation title: Distributed Rule-Based Ontology Reasoning
  - Advisor: Pascal Hitzler
- **Motilal Nehru National Institute of Technology** Allahabad, India  
*Master of Technology (M.Tech) in Computer Science* *June 2006*
  - Thesis title: Traceability from Use Case to .NET Assembly via Design Patterns
  - Advisor: Banshi D. Chaudhary
- **Jawaharlal Nehru Technological University** Hyderabad, India  
*Bachelor of Technology (B.Tech) in Computer Science* *May 2004*

## Research & Industry Experience

- **IBM T.J. Watson Research Center** NY, USA  
*Research Intern* *Jun 2015 - Aug 2015*

- Worked on ontology modeling and reasoning in the context of social data for situational understanding.
- Built an ontology for *Protests* based on tweets, Wikipedia and news articles.
- Developed REST services backed by a reasoner on top of the ontology. These services can be used to find correlations among tweets.
- This work was done in collaboration with University of Illinois Urbana-Champaign (UIUC), Singapore Management University (SMU), Air Force Research Labs (AFRL) and Army Research Labs (ARL).

- **IBM Research** Dublin, Ireland  
*Research Intern* *Jun 2013 - Aug 2013*

- Worked on distributed reasoning algorithms for description logic  $\mathcal{EL}^{++}$ .
- Extended my work on distributed  $\mathcal{EL}^+$  reasoning by adding support for ABox reasoning, incremental reasoning and dynamic load balancing.
- Evaluation done on Dublin city traffic data.

- **Alcatel-Lucent Bell Labs** Dublin, Ireland  
*Research Intern* *Jun 2012 - Aug 2012*

- Worked on scalable SPARQL query processing over RDF triples.
- Implemented SPARQL query processing on top of MongoDB. Graph of the query is built and triple patterns that can be run in parallel are identified.
- MapReduce was used for all the preprocessing steps such as dictionary encoding, generation of adjacency list etc.

- **Xerox Research Center** NY, USA  
*Research Intern* *Jun 2011 - Aug 2011*

- Worked on temporal consistency checking in marketing workflows.
- Explored the existing work on temporal modeling in the context of Semantic Web.
- Developed a temporal model based on James Allen's temporal operators.
- Developed reasoning and explanation (justification) mechanism for the inconsistencies.

- **Complexible Inc.** Boston, USA  
*Research Intern* *Jun 2010 - Aug 2010*

- Worked on a parallel implementation of  $\mathcal{EL}^+$  reasoner using Cray XMT supercomputer.
- Graph is built from the ontology axioms and transitive closure with respect to subclass relation is computed.

- **CA Technologies** Hyderabad, India  
*Software Engineer* *Aug 2006 - Dec 2008*

- Worked on developing prototypes for new product ideas and components for internal use.
- Developed enhancements for a timesheet and resource management tool called Clarity.

## Publications

### 1 Journal Papers

- Grigoris Antoniou, Sotiris Batsakis, **Raghava Mutharaju**, Jeff Z. Pan, Guilin Qi, Ilias Tachmazidis, Jacopo Urbani, and Zhangquan Zhou. *Reasoning on the Web of Data*. Submitted.

## 2 Conference Papers

- **Raghava Mutharaju**, Pascal Hitzler, Prabhaker Mateti, Freddy Lécué. *Distributed and Scalable OWL EL Reasoning*. Proceedings of the 12th Extended Semantic Web Conference (ESWC 2015), Portoroz, Slovenia, May 31-June 4, 2015. Volume 9088 of Lecture Notes in Computer Science, pages 88-103. Springer, Heidelberg, 2015.
- **Raghava Mutharaju**. *Very Large Scale OWL Reasoning Through Distributed Computation*. Proceedings of the 11th International Semantic Web Conference (ISWC 2012), Part II, Boston, MA, USA, Nov 11-15, 2012. Volume 7650 of Lecture Notes in Computer Science, pp. 407-414. Springer, Heidelberg, 2012.
- Zhangquan Zhou, Guilin Qi, Chang Liu, Pascal Hitzler, **Raghava Mutharaju**. *Reasoning with Fuzzy- $\mathcal{EL}^+$  Ontologies Using MapReduce*. Proceedings of the 20th European Conference on Artificial Intelligence (ECAI 2012), 27-31 August 2012, Montpellier, France. Frontiers in Artificial Intelligence and Applications, Vol. 242, IOS Press, Amsterdam, 2012, pp. 933-934.
- Satya S. Sahoo, D. Brent Weatherly, **Raghava Mutharaju**, Pramod Anantharam, Amit Sheth, Rick L. Tarleton. *Ontology-driven Provenance Management in eScience: An Application in Parasite Research*. OnTheMove Federated Conferences & Workshops (OTM 2009) ODBASE 2009, Vilamoura, Algarve-Portugal, Nov 3-5, 2009, volume 5871 of Lecture Notes in Computer Science, pages 992-1009. Springer, Heidelberg, 2009.
- Meenakshi Nagarajan, Karthik Gomadam, Amit Sheth, Ajith Ranabahu, **Raghava Mutharaju** and Ashutosh Jadhav. *Spatio-Temporal-Thematic Analysis of Citizen-Sensor Data - Challenges and Experiences*. Proceedings of the 10th International Conference on Web Information Systems Engineering (WISE 2009), Oct 5-7, Poland, 2009, volume 5802 of Lecture Notes in Computer Science, pages 539-553. Springer, Heidelberg, 2009.

## 3 Edited Proceedings

- Claudia d'Amato, Freddy Lécué, **Raghava Mutharaju**, Tom Narock, and Fabian Wirth (Eds.). Proceedings of the *1st International Diversity++ Workshop*, co-located with the 14th International Semantic Web Conference (ISWC 2015), Bethlehem, Pennsylvania, USA, October 12, 2015. CEUR Workshop Proceedings 1501, CEUR-WS.org, 2015. <http://ceur-ws.org/Vol-1501>

## 4 Workshop Papers

- **Raghava Mutharaju**, Prabhaker Mateti, and Pascal Hitzler. *Towards a Rule Based Distributed OWL Reasoning Framework*. Proceedings of the 12th OWL Experiences and Directions Workshop (OWLED 2015) co-located with the 14th International Semantic Web Conference (ISWC 2015). Lecture Notes in Computer Science, Springer, Heidelberg, 2015.
- **Raghava Mutharaju**. *Distributed Reasoning over Ontology Streams and Large Knowledge Base*. NSF Data Science Workshop 2015.
- Kasthuri Jayarajah, Shuochao Yao, **Raghava Mutharaju**, Archan Misra, Geeth De Mel, Julie Skipper, Tarek Abdelzaher, and Michael Kolodny. *Social Signal Processing for Real-time Situational Understanding: a Vision and Approach*. Proceedings of the 1st International Workshop on Social Sensing (SocialSens 2015) co-located with the the 12th IEEE International Conference on Mobile Ad hoc and Sensor Systems (IEEE MASS 2015), Dallas, TX, USA, Oct 19-22, 2015. IEEE, 2015, pp. 627-632. Invited paper.

- **Raghava Mutharaju**, Pavan Kapanipathi. *Are We Really Standing on the Shoulders of Giants?* Proceedings of the 1st International Workshop on Negative or Inconclusive Results in Semantic Web, NoISE 2015, co-located with the 12th Extended Semantic Web Conference (ESWC 2015), Portoroz, Slovenia, June 1st, 2015. CEUR Workshop Proceedings Vol-1435.
- **Raghava Mutharaju**, Pascal Hitzler, Prabhaker Mateti. *Distributed OWL EL Reasoning: The Story So Far*. SSWS 2014, Scalable Semantic Web Knowledge Base Systems. Proceedings of the 10th International Workshop on Scalable Semantic Web Knowledge Base Systems co-located with 13th International Semantic Web Conference (ISWC 2014) Riva del Garda, Italy, October 20, 2014. CEUR Workshop Proceedings Vol-1261, pp. 61-76.
- **Raghava Mutharaju**, Prabhaker Mateti, Pascal Hitzler. *Developing a Distributed Reasoner for the Semantic Web*. ISWC-DEV 2014. Proceedings of the ISCW Developers Workshop 2014, co-located with the 13th International Semantic Web Conference (ISWC 2014) Riva del Garda, Italy, October 19, 2014. CEUR Workshop Proceedings Vol-1268, pp. 108-112.
- **Raghava Mutharaju**, Pascal Hitzler, Prabhaker Mateti. *DistEL: A Distributed EL+ Ontology Classifier*. Proceedings of the 9th International Workshop on Scalable Semantic Web Knowledge Base Systems (SSWS 2013), Sydney, Australia, October 21, 2013. CEUR Workshop Proceedings Vol. 1046, pp. 17-32.
- Zhangquan Zhou, Guilin Qi, Chang Liu, Pascal Hitzler, **Raghava Mutharaju**. *Scale Reasoning with Fuzzy- $\mathcal{EL}^+$  Ontologies based on MapReduce*. In: Proceedings of the IJCAI-2013 Workshop on Weighted Logics for Artificial Intelligence, WL4AI-2013, Beijing, China, August 2013, pp. 87-93.
- **Raghava Mutharaju**. *How I Would Like Semantic Web To Be, For My Children*. In: Workshop on Semantic Web in 10 years, co-located with the 11th International Semantic Web Conference (ISWC 2012), Boston, USA, 2012.
- **Raghava Mutharaju**, Frederick Maier, Pascal Hitzler. *A MapReduce Algorithm for  $\mathcal{EL}^+$* . Proceedings of the 23rd International Workshop on Description Logics (DL2010), Waterloo, Canada, 2010. CEUR Workshop Proceedings Vol. 573, pp. 464-474.

## 5 Posters/Demos/Challenges

- **Raghava Mutharaju**, Sherif Sakr, Alessandra Sala, Pascal Hitzler. *D-SPARQ: Distributed, Scalable and Efficient RDF Query Engine*. Proceedings of the International Semantic Web Conference (ISWC 2013) Posters & Demonstrations Track, Sydney, Australia, October 23, 2013. CEUR Workshop Proceedings Vol.1035, pp. 261-264.
- David Carral, Amit Joshi, Adila Krisnadhi, **Raghava Mutharaju**, Kunal Sengupta, Cong Wang. *Konf Connect*. WWW 2012, Metadata Challenge. 21st International Conference on World Wide Web. April 16-20, 2012, Lyon, France.
- Ashutosh Jadhav, Wenbo Wang, **Raghava Mutharaju**, Pramod Anantharam, Vinh Nyugen, Amit P. Sheth, Karthik Gomadam, Meenakshi Nagarajan, and Ajith Ranabahu. *Twitris: Socially Influenced Browsing*. Semantic Web Challenge 2009, 8th International Semantic Web Conference, Oct. 25-29 2009, Washington, DC, USA.
- **Raghava Mutharaju**, Satya S. Sahoo, D. Brent Weatherly, Pramod Anantharam, Flora Logan, Amit P. Sheth, Rick Tarleton. *Ontology Driven Integration of Biology Experiment Data*. Ohio Collaborative Conference on BioInformatics (OCCBIO 2009), June 14-17, 2009.
- Pramod Anantharam, Satya S. Sahoo, D. Brent Weatherly, Flora Logan, **Raghava Mutharaju**, Amit P. Sheth, Rick Tarleton. *Trykikipedia: Collaborative Bio-Ontology Development using Wiki Environment*. Ohio Collaborative Conference on BioInformatics (OCCBIO 2009), June 14-17, 2009.

## 6 Technical Reports

- Frederick Maier, **Raghava Mutharaju**, Pascal Hitzler. *Distributed Reasoning with  $\mathcal{EL}^{++}$  Using MapReduce*. June 2010. Department of Computer Science, Wright State University, Dayton, Ohio.
- **Raghava Mutharaju**, Banshi D. Chaudhary. *Traceability from Use Case to .NET Assembly via Design Patterns*. July 2006. Department of Computer Science, MNNIT, Allahabad, India.

### Tutorials & Workshops Organized

- Tutorial titled *AI for Smarter Cities. Are we there yet?* at the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016), July 9-11, 2016, New York, USA. Co-organized with Freddy Lécué, Jeff Z. Pan, Jiewen Wu and Pascal Hitzler.
- *Diversity++ Workshop*, co-located with the 14th International Semantic Web Conference (ISWC 2015). Co-chaired with Claudia d'Amato, Freddy Lécué, Tom Narock and Fabian Wirth. October 12, 2015.
- Tutorial titled *AI for Smarter Cities, Hype or Reality? A Study in Dublin, Bologna, Miami and Rio* at the 29th AAAI Conference on Artificial Intelligence (AAAI 2015), January 26, 2015, Texas, USA. Co-organized with Freddy Lécué, Jeff Z. Pan, Jiewen Wu and Pascal Hitzler.
- Tutorial titled *Large Scale Reasoning over Semantic Data* at the 13th International Semantic Web Conference (ISWC 2014), October 19, 2015, Riva del Garda, Italy. Co-organized with Jeff Z. Pan, Ilias Tachmazidis and Guilin Qi.

### Research Projects

## 1 Large Scale Computing

- **Distributed OWL EL Reasoning:** Existing reasoners work only on single machine. They cannot handle large knowledge bases, especially the ones that are automatically constructed. In this work, different ontology partitioning strategies, distributed computing frameworks and paradigms are explored with the goal of building a scalable, distributed and efficient ontology reasoner.
  - **DistEL**
    - \* This is a peer-to-peer distributed OWL EL reasoner.
    - \* Ontology axioms are partitioned based on their type across a cluster of machines.
    - \* Barrier synchronization is used for termination detection.
    - \* Load balancing is achieved at run time using work stealing. This piece of work was done during the summer internship at IBM Research, Dublin, Ireland.
    - \* Code available at <https://github.com/raghavam/DistEL>
  - **SparkEL**
    - \* This is a distributed OWL EL reasoner based on Apache Spark framework.
    - \* Axioms are converted to tuple format so that they can be represented as Spark RDDs.
    - \* Spark operations such as map, join and union are used to implement the reasoning rules.
    - \* Code available at <https://github.com/raghavam/sparkel> (under development).

- **DQuEL**
  - \* Distributed queue based implementation of OWL EL reasoning.
  - \* Each concept in the ontology is assigned a queue and these queues are spread over the cluster.
  - \* Appropriate reasoning rules are triggered based on the elements in the queue.
  - \* Code available at <https://github.com/raghavam/DQuEL>
- **MR-EL**
  - \* A MapReduce implementation of OWL EL reasoning.
  - \* Axioms are represented as key-value pairs.
  - \* Each reasoning rule is a MapReduce job and these are run iteratively.
  - \* An additional MapReduce job is needed for termination detection.
- **Shared Memory Reasoning**
  - \* OWL EL reasoning implemented on a massively parallel Cray XMT supercomputer.
  - \* Axioms are represented as a directed labeled graph.
  - \* This reduces the reasoning task to finding the transitive closure of the graph.
  - \* This work was done during the summer internship at Complexible Inc (Clark & Parsia).
- **Scalable RDF Query Processing**
  - Developed DSparg, a distributed and scalable RDF query engine.
  - RDF graph is vertex partitioned across the nodes in the cluster.
  - SPARQL query is analyzed to find triple patterns that can be run in parallel.
  - Preprocessing steps such as dictionary encoding and generation of adjacency list are MapReduce jobs.
  - This work was started during the summer internship at Alcatel-Lucent Bell Labs.
  - Code available at <https://github.com/raghavam/d-sparq>

## 2 Semantic Web Applications

- **Temporal Consistency Checking in Marketing Workflows**
  - Events in a workflow involve temporal relationships and time constraints.
  - Temporal inconsistencies could be present in such workflows.
  - Temporal model was developed in OWL and James Allen's temporal operators are implemented as SWRL rules.
  - Rules are run against the data in the knowledge base. Pellet is used to detect inconsistencies.
  - Explanations (justifications) for the inconsistencies were generated
  - This work was done during the summer internship at Xerox Research Center, Webster, NY.
- **Situational Understanding from Social data**
  - Goal is to develop a framework that can obtain situational awareness of events from social data such as Twitter and Instagram images.
  - I built an ontology for *Protests* based on tweets, Wikipedia and news articles.
  - Developed REST services backed by a reasoner on top of the ontology. These services can be used to find correlations among tweets.
  - This work was done during the summer internship at IBM T.J. Watson Research Center, NY.

- **Ontology driven Data Integration**

- Goal of the project is to develop an ontology-driven semantic problem solving environment (PSE) for parasite (*Trypanosoma cruzi*) data.
- It allows integration of local and public data to answer biology queries.
- I translated data from relational database and excel sheets to RDF using D2RQ and Jena.

- **Twitris**

- Tweets are analyzed along spatial, temporal and thematic dimensions.
- Events are extracted from tweets, which are in turn used to get related information from other data sources such as news, images and videos.
- I implemented tweet analysis algorithms and wrote SQL queries.

## **Professional Activities**

- **Program Committee member:** uSitu 2016, SSWS 2015, ESWC 2015 Posters & Demos, ESWC 2014 Posters & Demos, SSWS 2014.
- **External Reviewer:** ESWC (2016, 2015, 2012), KR (2016, 2014), JoDS 2015, AIMS 2014, SWJ 2014, ECAI (2014, 2010), RR 2014, JELIA 2014, EKAW 2014, ICBO 2013, SSWS (2013, 2011), JIST (2013, 2011), ISWC (2013, 2011), WWW 2012, FoIKS 2012, DL (2011, 2010).

## **Awards & Recognition**

- Invited to present at the 6th student workshop on Cloud and Data Services organized by IBM T.J. Watson Research Center, NY. December 3-4, 2015.
- Invited to attend NSF Data Science workshop 2015 at the University of Washington, Seattle. August 5-7, 2015.
- Awarded Amazon AWS Research grant for a period of 2 years.
- Awarded Microsoft Azure Research grant for a period of 1.5 year.
- News article on my visit to University of Huddersfield appeared on the news of University of Huddersfield and Wright State University.
- Travel awards for ISWC 2015, AAAI 2015, ISWC 2014, ISWC 2012, RR 2012, and DL 2010.
- Spot award at CA Technologies.

## **Technical Skills**

- **Programming Languages:** Java, Scala, C, C++
- **Semantic Technologies:** OWL, RDF, SPARQL
- **Other Technologies:** XML, SQL
- **Distributed Computing Frameworks:** Hadoop, Spark
- **NoSQL Stores:** Redis, MongoDB, Neo4j, HBase