PERSONAL INFORMATION

Petar Ristoski

📍 IBM Almaden Research Center, San Jose, CA, USA

📱 +1 408 927 2020

✉ petar.ristoski@ibm.com

🌐 petarristoski.com

Sex Male | Date of birth 12 October 1988 | Nationality Macedonian

WORK EXPERIENCE

May 2018 – Present

Employer

Responsibilities

Research Staff Member

IBM Almaden Research Center

Research in Artificial Intelligence with an emphasis on Neural Networks and Semantic Technologies. The research involves discovering fundamental principles and implementing prototype systems that can be used to understand, analyse, and manage human text as well as collaborating with human experts to train more capable Cognitive Systems.

August 2013 – October 2017

Employer

Responsibilities

PhD Researcher

University of Mannheim, Germany. 40 hours per week

Research in the area of Semantic Web, Linked Open Data and machine learning.
Using Linked Open Data as background knowledge for data mining.
Developing propositionalization strategies for creating data mining features from Semantic Knowledge Graphs.
Developing approaches for graph embeddings.
Integrating data from a large number of Web sources.

May 2017 – August 2017

Employer

Responsibilities

Research Intern

IBM Research Almaden. 40 hours per week

Research in the area of Machine Learning, Natural Language Processing an Semantic Web
Developing a system for user-centric ontology population with human-in-the-loop

August 2015 – October 2015

Employer

Responsibilities

Research Intern

Yahoo Labs!, London. 40 hours per week

Research in the area of Semantic Web, Linked Open Data, Web crawling and machine learning.
Developing a framework for enriching product ads using data from HTML annotations.
Implementation of a focused crawler for semantically annotated data.

January 2013 – June 2013

Employer

Responsibilities

Master Student Intern

SAP Research, Darmstadt, Germany. 40 hours per week

Research in the area of machine learning, natural language processing and social media.
Developing a machine learning approaches for detecting incident related messages in social media.
Implementing a system for structuring user-generated content using government data and Linked Open Data.

July 2012 –December 2012

Employer

Responsibilities

Intern

SAP Research, Darmstadt, Germany. 40 hours per week.

Web and mobile application development.
Building a system for information filtering based on crowd-sourcing.
Developing Android mobile application for reporting incidents.

September 2011 – June 2012	## Junior Research Assistant
Employer	Faculty of Computer Science and Engineering, Skopje, Macedonia. 40 hours per week.
Responsibilities	Research in the area of Semantic Web, information retrieval and natural language processing. Developing different applications plugins (Microsoft Word, Microsoft Outlook) that display additional contextual information to the user, and automatize some of the user's actions while the users work in their standard application environment. The prototypes were integrated into existing systems and data from Novartis (pharmaceutical company).

October 2011 – June 2012	## Laboratory Assistant
Employer	Faculty of Computer Science and Engineering, Skopje, Macedonia. 40 hours per week.
Responsibilities	Preparing and holding laboratory sessions for the subject "Computer Architectures", "Web Based Systems" and "Computer Networks" to undergraduate students

February 2011 – July 2011	## Junior Research Assistant
Employer	Faculty of Electrical Engineering and Information Technologies, Skopje, Macedonia. 10 hours per week.
Responsibilities	Research in the field of semantic web, information retrieval and natural language processing. Implementing concepts for automatic composition and execution of web services using semantic web technologies.

October 2010 – January 2011	## Laboratory Assistant
Employer	Faculty of Electrical Engineering and Information Technologies, Skopje, Macedonia. 10 hours per week.
Responsibilities	Preparing and holding laboratory sessions for the subjects "Network Operating Systems" and "Computer Networks" to undergraduate students

July 2010 – August 2010	## Intern
Employer	Technische Universität München, Munich, Germany. 40 hours per week.
Responsibilities	Research in the field of parallel processing using graphics processing unit. Using NVIDIA CUDA parallel computing architecture to measure and visualize the performances of different GPUs for parallel processing large amounts of data.

June 2010 – June 2010	## Junior Web Developer
Employer	Plannow Technologies, Skopje, Macedonia. 40 hours per week.
Responsibilities	Web application development. Building business intelligence tools using Tapestry Java framework, JSF and jQuery.

EDUCATION

August 2013 – January 2018	## Doctor of Philosophy - PhD
University	University of Mannheim, Germany
PhD Thesis	Exploiting Semantic Web Knowledge Graphs in Data Mining

November 2011 – June 2013	## Master of Science **G.P.A.: 10.0 (max 10.0)**
University	Faculty of Computer Science and Engineering, "Ss. Cyril and Methodius" University, Skopje, Macedonia
Study Program	Computer Networks and e-Technologies
M.Sc. Thesis	Structuring User-Generated Content using Real-World Sensor Information

October 2007 – July 2011	**Bachelor of Science**	**G.P.A.: 9.72 (max 10.0)**

University   Faculty of Electrical Engineering and Information Technologies, "Ss. Cyril and Methodius" University, Skopje, Macedonia

Study Program   Computer Science and Engineering

B.Sc. Thesis   Semantic Web Application Development Framework

September 2003 – June 2007	**High School Diploma**	**G.P.A.: 5.0 (max. 5.0)**

School   St. Kliment Ohridski, Ohrid, Macedonia

Study Program   Mathematics

RESEARCH INTERESTS

1. Semantic Web, Linked Open Data, and their applications

2. Machine Learning and Data Mining

3. Schema/Ontology Matching and Data Integration

PUBLICATIONS

1. AL. Gentile, D. Gruhl, **P. Ristoski**, S. Welch: Personalized Knowledge Graphs for the Pharmaceutical Domain, 18th International Semantic Web Conference (ISWC 2018), October, 2019, Auckland, New Zealand

2. AL. Gentile, D. Gruhl, **P. Ristoski**, S. Welch: Explore and Exploit – Dictionary Expansion with Human-in-the-Loop, 16th Extended Semantic Web Conference (ESWC 2017), June 2019, Portoroz, Slovenia

3. AL. Gentile, D. Gruhl, **P. Ristoski**, S. Welch: Information Extraction in Editorial Setting. A Tale of PDFs. Posters and Demos at the16[th] Extended Semantic Web Conference (ESWC 2017), June 2019, Portoroz, Slovenia

4. M. Pellegrino, M. Cochez, M. Garofalo, P. Ristoski : A Modular and Extensible Evaluation Framework for Node Embedding Techniques. Posters and Demos at the16[th] Extended Semantic Web Conference (ESWC 2019), June 2019, Portoroz, Slovenia

5. A. Alba, C. DeLuca, AL. Gentile, D. Gruhl, L. Kato, C. Kau, **P. Ristoski**, S. Welch. Identifying High Value Opportunities for Human in the Loop Lexicon Expansion. The third international workshop on Augmenting Intelligence with Bias-Aware Humans--in--the--Loop co-located with The Web Conf (WWW2019)

6. A. Alba, C. DeLuca, AL. Gentile, D. Gruhl, L. Kato, C. Kau, **P. Ristoski**, S. Welch Task Oriented Data Exploration with Human-in-the-Loop. The third international workshop on Augmenting Intelligence with Bias-Aware Humans--in--the--Loop co-located with TheWebConf (WWW2019)

7. A. Alba, D. Gruhl, **P. Ristoski**, S. Welch. "Interactive Dictionary Expansion using Neural Language Models.", Human in the Loop Workshop at ISWC, 2018

8. **P. Ristoski**, PhD Thesis: Exploiting semantic web knowledge graphs in data mining, 2018

9. K Clarkson, AL. Gentile, D. Gruhl, **P. Ristoski**, J. Terdiman, S. Welch . "User-Centric Ontology Population." In European Semantic Web Conference, pp. 112-127. Heraklion, Crete, Greece, 2018.

10. A. Alba, A. Coden, AL. Gentile, D. Gruhl, **P. Ristoski**, S. Welch: Multi-lingual Concept Extraction with Linked Data and Human-in-the-Loop. In: Proceedings of the Knowledge Capture Conference, 2017

11. A. Alba, A. Coden, AL. Gentile, D. Gruhl, **P. Ristoski**, S. Welch: Language Agnostic Dictionary Extraction. In: Proceedings of the ISWC 2017 Posters & Demonstrations and Industry Tracks co-located with 16th International Semantic Web Conference (ISWC 2017)

12. M. Alam, D. Recupero, M. Mongiovi, Aldo Gangemi, **P. Ristoski**: Frame Embeddings for Event-Based Knowledge Reconciliation. In: Proceedings of the ISWC 2017 Posters & Demonstrations and Industry Tracks co-located with 16th International Semantic Web Conference (ISWC 2017)

13. M. Alam, D. Recupero, M. Mongiovi, Aldo Gangemi, **P. Ristoski**: Reconciling Event-Based

Knowledge Through RDF2VEC. International Workshops on Hybrid Statistical Semantic Understanding and Emerging Semantics co-located with 16th International Semantic Web Conference (ISWC 2017)

14. **P. Ristoski**, S. Faralli, S. Ponzetto, H. Paulheim: Large-scale Taxonomy Induction using Entity and Word Embeddings. In: Proceedings of the International Conference on Web Intelligence, 2017.

15. M. Cochez, **P. Ristoski**, S. Ponzetto, H. Paulheim: Global RDF Vector Space Embeddings. In: Proceedings of the 16th International Semantic Web Conference 2017.

16. M. Cochez, **P. Ristoski**, S. Ponzetto, H. Paulheim: Biased graph walks for RDF graph embeddings. In Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics (p. 21). ACM, 2017.

17. A. Gentile, **P. Ristoski**, S. Eckel, H. Paulheim: Entity Matching on Web Tables: a Table Embeddings approach for Blocking, EDBT, 2017.

18. **P. Ristoski**, H. Paulheim. RDF2Vec: RDF Graph Embeddings for Data Mining. In: Proceedings of the 15th International Semantic Web Conference, 2016

19. **P. Ristoski**, H. Paulheim, G. K. D. de Vries. A Collection of Benchmark Datasets for Systematic Evaluations of Machine Learning on the Semantic Web. In: Proceedings of the 15th International Semantic Web Conference, 2016

20. J. Rosati, **P. Ristoski**, T. Di Noia, R. de Leone and H. Paulheim. RDF graph embeddings for content-based recommender systems. In: CEUR workshop proceedings CBRecSys 2016 : Proceedings of the 3rd Workshop on New Trends in Content-Based Recommender Systems co-located with ACM Conference on Recommender Systems (RecSys 2016) Boston, MA, USA

21. **P. Ristoski**, H. Paulheim, V. Svatek and V. Zeman. The Linked Data Mining Challenge 2016. In: CEUR Workshop Proceedings of the 4th Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data co-located with 12th Extended Semantic Web, ESWC 2016, Heraklion, Crete, Greece

22. **P. Ristoski** and P. Mika. Enriching product ads with Metadata from HTML annotations. In: Lecture notes in computer science The Semantic Web. Latest Advances and New Domains : 13th International Conference, ESWC 2016, Heraklion, Crete, Greece

23. **P. Ristoski.** "Towards Linked Open Data enabled Knowledge Discovery", In: Lectute Notes in Computer Science The Semantic Web. Latest Advances and New Domains : 12th European Semantic Web Conference, ESWC 2015, Portoroz, Slovenia, May 31 - June 4, 2015. Proceedings; 772-782. Springer, Cham, 2015.

24. **P. Ristoski** and H. Paulheim. Visual Analysis of Statistical Data on Maps using Linked Open Data. In: Lectute Notes in Computer Science The Semantic Web: ESWC 2015 Satellite Events : ESWC 2015 Satellite Events, Portorož, Slovenia, May 31--June 4, 2015, Revised Selected Papers; 138-143. Springer Internat., Cham, 2015.

25. **P. Ristoski**, H. Paulheim, V. Svatek and V. Zeman. The Linked Data Mining Challenge 2015. In: CEUR Workshop Proceedings of the 4th Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data co-located with 12th Extended Semantic Web Conference (ESWC 2015) Portoroz, Slovenia, May 31, 2015

26. A. Schulz, F. Janssen, **P. Ristoski** and J. Fürnkranz. Event-based Clustering for Reducing Labeling Costs of Event-related Microposts. In: ICWSM 15 : Proceedings of the 9th International AAAI Conference on Web and Social Media, Oxford, UK, May 26, 2015 – May 29, 2015; 686-689. AAAI Press, Palo Alto, Calif., 2015.

27. A. Schulz, **P. Ristoski**, J. Fürnkranz and F. Janssen. Event-based Clustering for Reducing Labeling Costs of Incident-Related Microposts. In: CEUR workshop proceedings MUD 2015 Mining Urban Data : Proceedings of the 2nd International Workshop on Mining Urban Data co-located with 32nd International Conference on Machine Learning (ICML 2015) Lille, France, July 11th, 2015; 44-52. RWTH, Aachen, 2015.

28. B. Schäfer, **P. Ristoski** and H. Paulheim. What is special about Bethlehem, Pennsylvania? : Identifying unexpected facts about DBpedia entities. In: CEUR workshop proceedings ISWC-P&D 2015 : Proceedings of the ISWC 2015 Posters & Demonstrations Track co-located with the 14th International Semantic Web Conference (ISWC-2015) Bethlehem, PA, USA, October 11, 2015; Paper 46. RWTH, Aachen, 2015.

29. H. Paulheim, A. Schulz, F. Janssen, **P. Ristoski** and I. Schweizer. Intelligente Datenauswertung mit Linked Open Data. In: Börteçin Ege, Corporate Semantic Web; 187-201, Springer Vieweg, Berlin,

2015.

30. **P. Ristoski**, M. Schuhmacher and H. Paulheim: Using Graph Metrics for Linked Open Data Enabled Recommender Systems. Proceedings of the 16th International Conference on Electronic Commerce and Web Technologies (EC-Web2015/T2), Valencia, Spain, September, 2015.

31. **P. Ristoski**, C. Bizer, and H. Paulheim: Mining the Web of Linked Data with RapidMiner, Semantic Web Challenge 2014, Riva del Garda, Italy, October 2014.

32. O. Lehmberg, D. Ritze, **P. Ristoski**, K. Eckert, H. Paulheim and C. Bizer: Extending Tables with Data from over a Million Websites, Semantic Web Challenge 2014, Riva del Garda, Italy, October 2014.

33. **P. Ristoski** and H. Paulheim. Feature Selection in Hierarchical Feature Spaces. In: Lecture Notes in Computer Science Discovery Science: 17th International Conference, DS 2014, Bled, Slovenia, 2014. Proceedings; 288-300.

34. **P. Ristoski** and H. Paulheim. A Comparison of Propositionalization Strategies for Creating Features from Linked Open Data. In: CEUR Workshop Proceedings LD4KD 2014 : Proceedings of the 1st Workshop on Linked Data for Knowledge Discovery (LD4KD at ECML PKDD 2014), Nancy, France, 2014

35. H. Paulheim, **P. Ristoski**, E. Mitichkin and C. Bizer. Data Mining with Background Knowledge from the Web. In: Proceedings of the 5th RapidMiner World (2014), Boston, USA

36. **P. Ristoski**, E. L. Mencía and H. Paulheim. A Hybrid Multi-strategy Recommender System Using Linked Open Data. In: Semantic Web Evaluation Challenge: SemWebEval 2014 at ESWC 2014, Anissaras, Crete, Greece, 2014.

37. **P. Ristoski**, H. Paulheim: Analyzing Statistics with Background Knowledge from Linked Open Data. SemStats challenge, ISWC 2014, Sydney, Australia

38. Axel Schulz, **P. Ristoski**. The Car That Hit The Burning House: Understanding Small Scale Incident Related Information in Microblogs, WCMCW, 2013, Boston, USA.

39. A. Schulz, **P. Ristoski**, H. Paulheim. I See a Car Crash: Real-time Detection of Small Scale Incidents in Microblogs. ESWC, 2013, Montpellier, France.

40. D. Trajanov, R. Stojanov, M. Jovanovik, V. Zdraveski, **P. Ristoski**, M. Georgiev, S. Filiposka: Semantic sky: a platform for cloud service integration based on semantic web technologies. In: Proceedings of the 8th International Conference on Semantic Systems

JOURNAL PUBLICATIONS

1. O. Lehmberg, D. Ritze, **P. Ristoski**, R. Meusel, H. Paulheim and C. Bizer: The Mannheim Search Join Engine, Web Semantics : Science, Services and Agents on the World Wide Web (35), Elsevier, Amsterdam, 2015

2. **P. Ristoski**, C. Bizer, and H. Paulheim: Mining the Web of Linked Data with RapidMiner , Web Semantics : Science, Services and Agents on the World Wide Web (35), Elsevier, Amsterdam, 2015

3. **P. Ristoski**, H. Paulheim: Semantic Web in Data Mining and Knowledge Discovery: A Comprehensive Survey, Journal of Web Semantics, 2016

4. **P. Ristoski**, J. Rosati, T. Di Noia, R. De Leone, H. Paulheim: RDF2Vec: RDF Graph Embeddings and Their Applications, Semantic Web Journal, 2018

5. **P. Ristoski**, P. Petrovski, P. Mika, H. Paulheim: A Machine Learning Approach for Product Matching and Categorization , Semantic Web Journal , 2018

6. M. Alam, D. Recupero, M. Mongiovi, Aldo Gangemi, **P. Ristoski**: Event-Based Knowledge Reconciliation using Frame Embeddings and Frame Similarity, Knowledge-Based Systems Journal, 2017

7. **P. Ristoski**, A. Alba, AL. Gentile, D. Gruhl, S. Welch: Large-Scale Relation Extraction from Web Documents and Knowledge Graphs with Human-in-the-Loop, Jouranl of Web Semantics, 2019

AWARDS

1. Winner of the Open Track at the Semantic Web Challenge 2014, Riva del Garda, Italy, October 2014. **P. Ristoski**, C. Bizer, and H. Paulheim: Mining the Web of Linked Data with RapidMiner.

2. Winner of the Big Data Track at the Semantic Web Challenge 2014, Riva del Garda, Italy, October

2014. O. Lehmberg, D. Ritze, **P. Ristoski**, K. Eckert, H. Paulheim and C. Bizer: Extending Tables with Data from over a Million Websites.

3. Best performing approach for task 1 and 3, and best performing approach overall of the Linked Open Data-enabled Recommender Systems, ESWC 2014, Crete, Greece. **P. Ristoski**, E. Mencía, H. Paulheim: A Hybrid Multi-Strategy Recommender System Using Linked Open Data.

4. Winner of the SemStats challenge, ISWC 2014, Sydney, Australia. **P. Ristoski**, H. Paulheim: Analyzing Statistics with Background Knowledge from Linked Open Data.

5. Winner of the Semantic Web Challenge at ISWC 2018, Monterey, USA.

SCIENTIFIC ACTIVITIES

PC Member and Reviewer

1. 4th Workshop on Data Mining and Knowledge Discovery Meets Linked Open Data (Know@LOD 2015)
2. 3rd Linked Data Mining Challenge 2015, in conjunction with Know@LOD 2015
3. 12th Extended Semantic Web Conference (ESWC 2015), June 2015, Portoroz, Slovenia.
4. 15th International Conference on Web Engineering (ICWE 2015), June 2015, Rotterdam, Netherlands
5. 13th International Semantic Web Conference (ISWC 2014), October 2014, Riva del Garda, Italy.
6. Digital Library Conference, September 2014, London, UK.
7. 31st International Conference on Data Engineering, April 2015, Seoul, Korea
8. 5th Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data, In conjunction with the 13th Extended Semantic Web Conference (ESWC 2015), 2016 Heraklion, Greece.
9. 4th Linked Data Mining Challenge 2016, in conjunction with Know@LOD 2016
10. Linked Data on the Web (LDOW2016) workshop at WWW, April 2016, Montreal, Canada
11. 14th Extended Semantic Web Conference (ESWC 2017), June 2017, Portoroz, Slovenia
12. 15th Extended Semantic Web Conference (ESWC 2018), June 2018, Heraklion, Greece
13. 27th International Joint Conference on Artificial Intelligence (IJCAI 2018), July 2018, Stockholm, Sweden
14. 6th International Conference on Computer and Communication Technology, September 2015, Allahabad, India
15. The 9th ACM International Conference on Web Search and Data Mining, San Francisco, California, USA, 2016
16. Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19), Hawaii, USA, 2019
17. 16th Extended Semantic Web Conference (ESWC 2017), June 2019, Portoroz, Slovenia
18. The Web Conference, 2019, San Francisco, USA
19. 28th International Joint Conference on Artificial Intelligence (IJCAI), Macao, China, 2019
20. Workshop on Deep Learning for Knowledge Graphs at ESWC, Portoroz, Slovenia, 2019
21. New Trends in Representation Learning with Knowledge Graphs at ECML/PKDD, Wuerzburg, Germany, 2019

Reviewer for journals

1. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, 2015
2. Journal of Web Engineering, 2016
3. Artificial intelligence review journal, 2017
4. IEEE Internet Computing, 2016
5. Journal of Web Semantics, 2018
6. Semantic Web Journal, 2019