Curriculum Vitae – Danushka Bollegala

Personal Details

Name:	Prof. Danushka Bollegala
Position:	Professor in Natural Language Processing
Address:	Department of Computer Science, University of Liverpool, L69 3BX, UK.
Phone:	+44-1517954283
Email:	danushka@liverpool.ac.uk
Web:	danushka.net
Orcid ID:	https://orcid.org/0000-0003-4476-7003

Higher Education

2007 – 2009	PhD Computer Science, The University of Tokyo, Japan. *Summa Cum Laude*
2005 – 2007	M.Sc. Computer Science, The University of Tokyo, Japan. *Summa Cum Laude*
2001– 2005	B.Sc. Computer Science, The University of Tokyo, Japan. *Summa Cum Laude*

Honours and Awards

1. EPSRC Top Peer Reviewer's Award 2017.
2. Best Journal Paper of the Year 2014-2015, Japanese Society for Artificial Intelligence.
3. IEEE Young Author Award 2011.
4. Best paper award at the 2011 Genetic and Evolutionary Computation (GECCO) Conference.
5. Best poster paper at the 2010 Pacific Rim International Conference on Artificial Intelligence (PRICAI).
6. Annual Conference Award for the Best Paper at 2010 Japanese Society for Artificial Intelligence (JSAI).
7. Dean's Award for the Best Doctoral Thesis of the Year 2010, University of Tokyo.
8. Dean's Award for the Best Masters Thesis of the Year 2007, University of Tokyo.
9. Dean's Award for the Best Undergraduate Thesis of the Year 2005, University of Tokyo.

Employment Record

Current October 2018	Professor, University of Liverpool, UK.
Current January 2019	Amazon Scholar, Amazon, Cambridge, UK.
September 2013 – December 2018	Senior Lecturer, University of Liverpool, UK.
April 2012 – August 2013	Senior Assistant Professor, (*koshi*) University of Tokyo, Japan.
April 2010 – March 2012	Assistant Professor, University of Tokyo, Japan.
October 2009 – March 2010	JSPS Post-doctoral Research Fellow, University of Sussex, UK.
April 2007 – September 2010	Japan Society for Promotion of Science (JSPS), Doctoral Research Fellow (DC1) University of Tokyo, Japan.
April 2005 – March 2007	Research Assistant, National Institute of Advanced Industrial Science and Technology (AIST), Japan.
June 2005 – March 2010	Research Consultant, FAST a Microsoft Subsidiary (former FAST Search & Transfer), Norway.

Fellowships and Awards

2018.04-2019.10	Advisor LexSnap UK, specialising in Legal Chatbot systems
2017.03-2017.06	Advisor Skwile UK, specialising in financial risk prediction
2011.04–2011.06	Visiting Research Fellow, Department of Computer Science, University of Cambridge, UK.
2009.10–2010.03	Japan Society for Promotion of Science (JSPS), Post-doctoral Research Fellow, University of Sussex, UK.
2007.04–2009.09	Japan Society for Promotion of Science (JSPS), Doctoral Research Fellow, University of Tokyo, Japan.
2000.04–2007.03	Japan Ministry of Education Overseas Full Scholarship, University of Tokyo, Japan

Teaching Experience

COMP 527: Data Mining and Visualisation I have been teaching COMP 527 continuously since 2014. It is a master-level module and a compulsory module for the MSc programme in Big Data and High Performance Computing. This module carries 15 credits and is taught in the second semester. Since I have taken over this module, I have re-written the module specifications and have added more practical elements such as the introduction of Python-based programming course work and lab assignments. Initially, video recordings of the lectures were made available on a dedicated YouTube channel, and later when the university introduced the video streaming platform, the video captures from the lectures were made available at stream.liv.ac.uk. The exam structure was also changed to introduce more problem solving-type questions and this effort has been repeatedly praised by the external examiners. This module is a popular choice among the computer science PGT cohort and the number of students taking this module has increased significantly over the year from 16 in 2014 to 48 in 2018. It is the module with the largest num-

ber of student registration among all PGT modules in the department. The module is also taken by a large number of PhD students from various CDT programmes such as the Data Analytics and Society CDT, Risk CDT and Physics CDT because data science has become an integrated component in many research fields not limiting to computer science. The student feedback for COMP 527 this year was extremely positive and all questions in the student evaluation received a high average rating of 4.1 or above.

COMP 212: Distributed Systems This is a second year 15 credit undergraduate model that is optional for all computer science students at the Department of Computer Science, University of Liverpool. The module covers both theoretical as well as practical aspects of distributed computing In addition to the written exam, there are two programming assignments that must be implemented using the Java programming language, testing the students' understanding of the algorithms in distributed systems. I taught this module for three consecutive years during the period from 2014 to 2016. On average, 40 students were registered for this module during that period. The student feedback for COMP 212 has been positive during that period. According to the departmental policy, undergraduate admission tutor is given a lower teaching load, and as a result I discontinued teaching COMP 212 from 2017.

C Programming (University of Tokyo) This is a second year compulsory module for all students in the Department of Information and Communication Engineering at the University of Tokyo, Japan. Techniques for optimising programmes written for lower-level hardware devices are covered in this modules. This is an intensive programming-oriented module with weekly assignments. On average 150 students take this module and it is the responsibility of the module coordinator to assess each submitted assignment and provide detailed feedback on a weekly basis. This module enabled me to gain valuable experience related to teaching large cohorts. In particular, I developed a machine learning-based automatic programme evaluation system that can automatically grade the student assignments and highlight common mistakes. This enabled students to submit their assignments many times as they wish before the deadline and obtain real-time feedback. This also encouraged students to submit the assignments well before the deadlines. After the deadline has passed, I personally verified the mistakes detected by the tool and provided an annotated feedback to the students. It would have been extremely difficult and time consuming to conduct this module without this innovation.

Leadership, Professional and Collegial Experience

2017.10-current	Head of the Data Mining and Machine Learning (DMML) Research Group.
2016.01-current	Undergraduate admissions tutor, Department of Computer Science, University of Liverpool
2013.09-2016.01	Disability Support Officer, Department of Computer Science, University of Liverpool
2014-current	Leader Natural Language Processing Group, University of Liverpool
2017-2018	Member of the Advisory Committee, AI and Future Jobs, Royal Society of Science.
2017-current	Full member of Engineering and Physical Science Research Council (EPSRC) Peer Review College
2017-current	Assessor for the Irish Research Council
2013-2015	Associate Editor of the Transactions of the Japanese Society for Artificial Intelligence
2016-current	Associate Editor of the Journal of Computational Social Sciences
2014-current	Evaluator of Research Grants for Xi'an Jiaotong-Liverpool University

Organisation of Scientific Meetings

2018	**Co-organiser** of the Knowledge Representation and Reasoning in Natural Languages (KRNL) Workshop at the 16th International Conference on the Principles of Knowledge Representation and Reasoning
2017	**Local organiser** for the17th Annual Meeting of the International Society of Pharmacovigilance (ISoP)
2017	**Program chair** of the Pharmacovigilance and Social Media Workshop at ISoP
2012	**Program chair** of the International Organised Sessions (IOS) at the 26th Annual Conference of Japanese Society for Artificial Intelligence

Program Committees

2019	Area Chair of the Machine Learning Track at EMNLP-2019
2014 –	Senior Programme Committee member for the International Joint Conference in Artificial Intelligence
2014–	Senior Programme Committee member for the AAAI Conference on Artificial Intelligence
2014	Information Extraction Area Chair for the International Conference on Computational Linguistics (COLING)
2010–	PC member of ACL, EMNLP, NIPS, WWW, COLING, LREC and reviewer for JAIR, TKDE, TKDD, JMLR, TACL journals.

Research Grants

Total research income so far GBP 2,064,100.

1. **Procedural Natural Language Inference** (Cookpad), PI, (GBP 48k), 2017-2019.
2. **Legal Advisor Dialogue Engine** (LexSnap), PI, (GBP 12k), 2017-2018.
3. **Algorithm Design for Automatic Classification of Transactions into a Taxonomy** (Rosslyn Data Technologies), PI, (GBP 5k), 2017-2018.
4. **Track Analytics For Effective Triage of Wide Area**, Defence Science Technology Laboratory (DSTL), Co-I (GBP 243k), 2017-2019.

5. **Digital Legal Assistant**, Knowledge Transfer Partnership (Innovate UK), PI (GBP 492k), 2017-2020.
6. **WEB-RADR**: Recognising Adverse Drug Reactions, (European Commission) Innovative Medicine Initiative, Co-PI (GBP 471k), 2014-2017.
7. **I knew that relation from news**, Innovation Voucher Scheme, University of Liverpool, PI (GBP 5k), 2015-2016.
8. **The Revierview Law Contract Map Project**, Knowledge Transfer Partnership (Innovate UK), Co-PI (GBP 269k), 2015-2018.
9. **Resolving Relational Ambiguity between Entities on the Web**, Microsoft Research (MSR) CORE-9 Research Grant, PI, (GBP 10K), 2013–2015.
10. **Domain Adaptation for Semantic Relation Extraction**, Japanese Society for the Promotion of Science (JSPS), Research Grant for Young Researchers (B). PI (GBP 20K), 2012–2015.
11. **Cross-Language Relational Search**, Microsoft Research (MSR) CORE-7 Research Grant, PI, (GBP 20K), 2011-2012.
12. **Developing a Cross-Language Web Search Engine**, Information Technology Promotion Agency of Japan (IPA) grant for explorative software development (Mito Project), PI, (GBP 26.5K), 2010-2011.
13. **A Latent Relational Search Engine**, Google Research Award, Co-PI, (GBP 18.6K), 2010–2011.
14. **Developing a Relational Search Engine to Retrieve Semantic Relations between Entities**, Japanese Society for the Promotion of Science (JSPS) research grant, PI. (GBP 29.8K), 2010-2012.
15. **Research grant for overseas visiting scholars**, Global Centre of Excellence (GCOE), Japan. PI. (GBP 9.7K), April 2011–June 2011.
16. **Extracting Attributes for Entities using Web Data**, Global Centre of Excellence (GCOE), Japan. Sole PI. (GBP 7.5K), 2010-2011.
17. **Learning to Rank Entities on the Web**, Microsoft Research (MSR) CORE-6 Research Grant, Co-PI, (GBP 19.4K), 2010-2011.
18. **Using Web Mining to Provide Useful Information to Drivers**, Toyota InfoTechnology Centre, Co-PI, (GBP 29.8K) 2009-2012.
19. **Disambiguating Personal Names on the Web**, Japan Society for the Promotion of Science (JSPS) Research Grant PI. (GBP 29.8K), 2007-2009.
20. **Using Network Theory and Machine Learning to Structure and Represent Information Available on the Web**, Co-PI, (GBP 298K), 2009–2012.

Supervision of PhD Students

Graduated PhD Students under my primary supervision:

1. Huda Hakami, May 2020.
2. Xia Cui, graduated April 2020, now KTP associate.
3. Alsuhaibani Mohammed, March 2020, now lecturer working in Saudi Arabia.
4. Pavithra Rajendran, graduated March 2019, now data scientist at KPMG.

5. Asif Hussain Khan, graduated March 2014, now assistant professor, University of Dhaka.
6. Leon Palafox, graduated March 2012, now postdoc at University of Arizona.
7. Liu Shu, graduated March 2011, now engineer at Microsoft.
8. Makoto Tanji, graduated March 2011, now engineer at Wantedly.
9. Akio Watanabe, graduated March 2012, now engineer at CyberAgent.
10. Nguyen Tuan Duc, graduated March 2011, now engineer at Alt+.
11. Hugo Hernault, graduated March 2011, now engineer at Barclays.
12. Abdullah Alsheri, graduated June 2017, now lecturer at Saudi Arabia.

PhD Examinations:

1. Abiola Obamuyide, University of Sheffield, June 2020.
2. Alexander Phillips, University of Liverpool, January 2020.
3. Mohammed Al-Zeyadi, University of Liverpool, July 2018.
4. Bastian Broecker, University of Liverpool, April 2018.
5. Fatima Abdullahi, University of Liverpool, May 2016.
6. Liyung Gong, University of Liverpool, November, 2014.
7. Tacoa Renevey Francisco, University of Tokyo, March, 2013.
8. Mamdouh Farouk Mohamed, University of Tokyo, March, 2012.
9. Haibo Li, University of Tokyo, March 2011.
10. Alena Neviarouskaya, University of Tokyo, March 2011.

My current PhD students:

James O'Neill, Jodie Chou, Michael Abaho, Samantha Durdy.

Selected Keynotes/Invited Talks

1. Keynote speech at National Human Resource Conference, Colombo, 2018.
2. Invited talk at International Society for Pharmacovigilance, 2017.
3. Invited talk at Microsoft Research Beijing Lab, 2013.
4. Keynote at Information-Based Induction Sciences (IBIS) Conference, 2011.
5. Invited talk at Google Mountain View Lab, 2011.
6. Invited talk at Microsoft Research Seattle Lab, 2010.
7. Keynote at First Japanese Web Symposium, 2009.

Patents

1. Query Annonymisation via Semantic Decomposition, Japanese patent (filed 2018.06 and patent-pending).
2. A Method for Extracting the Semantic Relations that exist between two Entities from a Text Corpus, Japanese patent no: 2010-096551, 2010.
3. A Relational Search System, Japanese patent no: 2009-275762, 2009.

Faculty Impact case Records

1. AI & Law Impact case covering KTP projects with Riverview Law, Fletchers solicitors, and consultancy project with LexSnap.
2. Impact case for the monitoring adverse reactions of drugs from social media for pharmacovigilance (WEB-RADR project)

Industrial Collaborations/Consultancies

2018.04-current	NLP consultant, LexSnap Ltd.
2017.08-current	Senior Fellow, Cookpad Ltd.
2017.11-current	Chief Scientific Officer (CSO), Alt

Publications

I have published over 125 papers in top international venues in Natural Language Processing, Machine Learning, Data Mining, Artificial Intelligence, and Social Media Analysis.My papers have been cited 3301 times with an h-index of 28 and i10-index of 57. For a full list of my publications and metrics see Google Scholar Profile.

Referred Journal Papers

[1] Abdullah Alshehri, Frans Coenen, and Danushka Bollegala. Behavioural biometric continuous user authentication using multivariate keystroke streams in the spectral domain. In Ana Fred, David Aveiro, Jan L. G. Dietz, Kecheng Liu, Jorge Bernardino, Ana Salgado, and Joaquim Filipe, editors, *Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 43–66, Cham, 2019. Springer International Publishing.

[2] Abdullah Alsheri, Frans Coenen, and Danushka Bollegala. Iterative time keystroke continuous authentication: A time series based approach. *KI - Künstliche Intelligenz*, 32(1):1–13, 2018.

[3] Mohammed Alsuhaibani, Danushka Bollegala, Takanori Maehara, and Ken-ichi Kawarabayashi. Jointly learning word embeddings using a corpus and a knowledge base. *Plos One*, 13(3):1–26, 2018.

[4] Danushka Bollegala, Vincent Atanasov, Takanori Maehara, and Ken-ichi Kawarabayashi. Classinet – predicting missing features for short-text classification. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, (to appear):1–29, 2018.

[5] Danushka Bollegala, Richard Slone, Simon Maskell, Joanne Hajne, and Munir Pirmohammed. Learning causality patterns for detecting adverse drug reactions from social media. *Journal of Medical Internet Research Public Health Surveillance*, 4(2):1–20, 2018.

[6] Xia Cui, Noor Al-Bazzaz, Danushka Bollegala, and Frans Coenen. A comparative study of pivot selection strategies for unsupervised cross-domain sentiment classification. *The Knowledge Engineering Review*, 33:1–24, 2018.

[7] Tomoyuki Kajiwara, Danushka Bollegala, Yuichi Yoshida, and Ken-ichi Kawarabayashi. An iterative approach for the global estimation of sentence similarity. *PLoS ONE*, 12(9):1–15, July 2017.

[8] Danushka Bollegala. Dynamic feature scaling for online learning of binary classifiers. *Knowledge-Based Systems*, 129:97–105, 2017.

[9] Danushka Bollegala, Kohei Hayashi, and Ken-ichi Kawarabayashi. Learning linear transformations between counting-based and prediction-based word embeddings. *PLoS ONE*, 12(9):1–21, 2017.

[10] Huda Hakami and Danushka Bollegala. Compositional approaches for representing relations between words: A comparative study. *Knowledge-Based Systems*, 136C:172–182, 2017.

[11] Danushka Bollegala, Georgios Kontonatsios, and Sophia Ananiadou. A cross-lingual similarity measure for detecting biomedical term translations. *PLOS ONE*, 10(6):1–28, 06 2015.

[12] Danushka Bollegala, Tingting Mu, and Yannis Goulermas. Cross-domain sentiment classification using sentiment sensitive embeddings. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 28(2):398–410, 2015.

[13] Hakami Huda and Danushka Bollegala. A classification approach for detecting cross-lingual biomedical term translations. *Natural Language Engineering*, 1(1469-8110):1–21, 2015.

[14] Nozomi Nori, Danushka Bollegala, and Mitsuru Ishizuka. Interest prediction via user's actions on social media. *Transactions of the Japanese Society for Artificial Intelligence*, pages 168–176, 2015.

[15] Nozomi Nori, Danushka Bollegala, and Hisashi Kashima. Simultaneous higher-order relation prediction via collective indicence matrix embedding. *Transactions of the Japanese Society for Artificial Intelligence*, pages 459–465, 2015.

[16] Hiroyuki Sato, Yoshihiko Hasegawa, Danushka Bollegala, and Hitoshi Iba. Improved sampling using loopy belief propagation for probabilistic model building genetic programming. *Swarm and Evolutionary Computation*, pages 1–8, 2015.

[17] Richard Sloane, Orod Osanlou, David Lewis, Danushka Bollegala, Simon Maskell, and Munir Pirmohamed. Social media and pharmacovigilance: A review of the opportunities and challenges. *British Journal of Clinical Pharmacology*, 80(4):910 – 920, 2015.

[18] Nozomi Nori, Danushka Bollegala, and Hisashi Kashima. A dimension reduction approach to multinomial relation prediction. *Transactions of the Japanese Society for Artificial Intelligence*, pages 168–176, 2014.

[19] Danushka Bollegala. Deep learning for natural language processing. *Journal of the Japanese Society for Artificial Intelligence*, pages 195–203, 2013.

[20] Danushka Bollegala, Tomokazu Goto, Nguyen Tuan Duc, and Mitsuru Ishizuka. Improving relational similarity measurement using symmetries in proportional word analogies. *Information Processing and Management*, 49(1):355 – 369, 2013.

[21] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Minimally supervised novel relation extraction using latent relational mapping. *IEEE Transactions on Knowledge and Data Engineering*, 25(2):419 – 432, 2013.

[22] Danushka Bollegala and Ekaterina Shutova. Metaphor interpretation using paraphrases extracted from the web. *PLoS ONE*, 8(9):1–10, 2013.

[23] Danushka Bollegala, David Weir, and John Carroll. Cross-domain sentiment classification using a sentiment sensitive thesaurus. *IEEE Transactions on Knowledge and Data Engineering*, 25(8):1719 – 1731, 2013.

[24] Muhammad Asif Hossain Khan, Danushka Bollegala, Guangwen Li, and Kaoru Sezaki. Delineating real-time events by identifying relevant tweets with popular discussion points. *ASE Human Journal*, 2(3):136 – 150, 2013.

[25] Ken-ichi Yokote, Danushka Bollegala, and Mitsuru Ishizuka. Jointly learning similarity transformations for textual entailment. *Transactions of of the Japanese Society for Artificial Intelligence*, pages 220–229, 2013.

[26] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Automatic annotation of ambiguous personal names on the web. *Computational Intelligence*, 28(3):398 – 425, 2012.

[27] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Measuring the degree of synonymy between words using relational similarity between word pairs as a proxy. *Institute of Electronics, Information and Communication Engineers (IEICE) Transactions on Information Systems*, pages 2116–2123, 2012.

[28] Danushka Bollegala, Naoaki Okazaki, and Mitsuru Ishizuka. A preference learning approach to sentence ordering for multi-document summarization. *Information Sciences*, 217:78 – 95, 2012.

[29] Danushka Bollegala, Naoki Tani, and Mitsuru Ishizuka. Improving the accuracy of attribute extraction using the relatedness between attribute values. *Transactions of the Japanese Society for Artificial Intelligence*, pages 245–252, 2012.

[30] Nguyen Tuan Duc, Danushka Bollegala, and Mitsuru Ishizuka. Cross-language latent relational search between japanese and english languages using a web corpus. *ACM Transactions on Asian Language Processing (TALIP)*, 11(3):1 – 33, 2012.

[31] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Automatic discovery of personal name aliases from the web. *IEEE Transactions on Knowledge and Data Engineering*, 23(6):831 – 844, July 2011.

[32] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. A web search engine-based approach to measure semantic similarity between words. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 23(7):977–990, July 2011.

[33] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. A supervised classification approach for measuring similarity between word pairs. *Transactions of the Institute of Electronics, Information and Communication Engineers (IEICE)*, E94-D(11):2227–2233, 2011.

[34] Nguyen Tuan Duc, Danushka Bollegala, and Mitsuru Ishizuka. Exploiting relational similarity between entity pairs for latent relational search. *Transactions of the Information Processing Society of Japan*, 52(4):1790–1802, 2011.

[35] Nguyen Tuan Duc, Danushka Bollegala, and Mitsuru Ishizuka. Relation representation and indexing method for fast and high precision latent relatioal search engine. *Special issue of the Transactions of the Japanese Society for Artificial Intelligence*, 26(2):307–312, 2011.

[36] Tomokazu Goto, Nguyen Tuan Duc, Danushka Bollegala, and Mitsuru Ishizuka. Improving relational search performance using relational symmetries and predictors. *Transactions of of the Japanese Society for Artificial Intelligence*, 26(6):649–656, 2011.

[37] Wataru Sunayama, Yasufumi Takama, Danushka Bollegala, Yoko Nishihara, Hidekazu Tokunaga, Mineo Kushima, and Mitsunori Matsushita. Total environment for text data mining. *Transactions of of the Japanese Society for Artificial Intelligence*, 26(4):483–493, 2011.

[38] Danushka Bollegala, Naoaki Okazaki, and Mitsuru Ishizuka. A bottom-up approach to sentence ordering for multi-document summarization. *Information Processing and Management*, 46(1):89 – 109, 2010.

[39] Keigo Watanabe, Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Automatic extraction of related terms using web search engines. *Journal of the Japan Society for Fuzzy Theory and Intelligent Informatics*, 23(4):483–493, 2010.

[40] D. BOLLEGALA, N. OKAZAKI, and M. ISHIZUKA. Agglomerative clustering based approach to sentence ordering for multi-document summarization. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*, 105(594):13–18, 2006.

Referred Conference Papers

[41] Danushka Bollegala, Ryuichi Kiryo, Kosuke Tsujino, and Haruki Yukawa. Language-independent tokenisation rivals language-specific tokenisation for word similarity prediction. In *Proc. of the 12th International Conference on Language Resources and Evaluation (LREC)*, 2020.

[42] Masaru Isonuma, Junichiro Mori, Danushka Bollegala, and Ichiro Sakata. Tree-structured neural topic model. In *Proc. of the 20th Annual Conference of the Association for Computational Linguistics (ACL)*, 2020.

[43] Angrosh Mandya, James O' Neill, Danushka Bollegala, and Frans Coenen. Do not let the history haunt you: Mitigating compounding errors in conversational question answering. In *Proc. of the 12th International Conference on Language Resources and Evaluation (LREC)*, 2020.

[44] James O'Neill and Danushka Bollegala. Meta-embedding as auxiliary task regularization. In *Proc. of the 24th European Conference on Artificial Intelligence (ECAI)*, 2020.

[45] Micheal Abaho, Danushka Bollegala, Paula Williamson, and Susanna Dodd. Correcting crowdsourced annotations to improve detection of outcome types in evidence based medicine. In *Proc. of the 4th International Workshop on Knowledge Discovery in Healthcare Data (KDH) at the 28th International Joint Conference on Artificial Intelligence*, 2019.

[46] Mohammed Alsuhaibani, Takanori Maehara, and Dansuhka Bollegala. Joint learning of hierarchical word embeddings from a corpus and a taxonomy. In *Proc. of the Automated Knowledge Base Construction Conference*, 2019.

[47] Robert Bevan, Alessandro Torrisi, Danushka Bollegala, Frans Coenen, and Katie Atkinson. Extracting supporting evidence from medical negligence claim texts. In *Proc. of the 4th International Workshop on Knowledge Discovery in Healthcare Data (KDH) at the 28th International Joint Conference on Artificial Intelligence*, 2019.

[48] Wenye Chen, Huda Hakami, and Danushka Bollegala. Learning to compose relational embeddings in knowledge graphs. In *Proc. of the 16th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, 2019.

[49] Xia Cui and Danushka Bollegala. Self-adaptation for unsupervised domain adaptation. In *Proc. of the Recent Advances in Natural Language Processing (RANLP)*, 2019.

[50] Huda Hakami and Dansuhka Bollegala. Learning relation representations from word representations. In *Proc. of the Automated Knowledge Base Construction Conference*, 2019.

[51] Huda Hakami and Danushka Bollegala. Context-guided self-supervised relation embeddings. In *Proc. of the 16th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, 2019.

[52] Masahiro Kaneko and Danushka Bollegala. Gender-preserving debiasing for pre-trained word embeddings. In *Proc. of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1641–1650, 2019.

[53] Shan Luo, Jet-Tsyn Lee, and Danushka Bollegala. "touching to see" and "seeing to feel": Robotic cross-modal sensory data generation for visual-tactile perception. In *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*, pages 4276–4282, 2019.

[54] Angrosh Mandya, Danushka Bollegala, Frans Coenen, and Katie Atkinson. Combining long short term memory and convolutional neural network for cross-sentence n-ary relation extraction. In *Proc. of the Automated Knowledge Base Construction Conference*, 2019.

[55] Angrosh Mandya, Danushka Bollegala, and Frans Coenon. Evaluating co-reference chains based conversation history in conversational question answering. In *Proc. of the 16th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, 2019.

[56] James O'Neill and Danushka Bollegala. Learning to evaluate neural language models. In *Proc. of the 16th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, 2019.

[57] James O'Neill, Danushka Bollegala, Peter Noble, and Alan Radford. Tick parasitism classification from noisy medical records. In *Proc. of the 4th International Workshop on Knowledge Discovery in Healthcare Data (KDH) at the 28th International Joint Conference on Artificial Intelligence*, 2019.

[58] Pavithra Rajendran, Danushka Bollegala, and Simon Parsons. A pilot study on argument simplificatoin in stance-based opinions. In *Proc. of PACLING*, 2019.

[59] Alessandro Torrisi, Robert Bevan, Danushka Bollegala, Katie Atkinson, and Frans Coenon. Automated bundle pagination using machine learning. In *Proc. of the 17th International Conference on Artificial Intelligence and Law (ICAIL)*, 2019.

[60] Alessandro Torrisi, Robert Bevan, Danushka Bollegala, Katie Atkinson, and Frans Coenon. Combining textual and visual information for typed and handwritten text separation in legal documents. In *Proc. of the 32nd International Conference on Legal Knowledge and Information Systems (JURIX)*, 2019.

[61] Yi Zhou and Danushka Bollegala. Unsupervised evaluation of human translation quality. In *Proc. of the 11th International Conference on Knowledge Discovery and Information Retrieval (KDIR)*, pages 55–64, 2019.

[62] Mohammed Alsuhaibani and Danushka Bollegala. Joint learning of sense and word embeddings. In *Proc. of the Eleventh International Conference on Language Resources and Evaluation (LREC)*, pages 1–7, 2018.

[63] Cong Bao and Danushka Bollegala. Learning word meta-embeddings by autoencoding. In *Proc. of the 27th International Conference on Computational Linguistics (COLING)*, pages 1650–1661, 2018.

[64] Robert Bevan, Alessandro Torrisi, Katie Atkinson, and Frans Coenen. Efficient and effective case reject-accept filtering: A study using machine learning. In *Proc. of the 31st International Conference on Legal Knoweledge and Information Systems (JURIX)*, pages 171–175, 2018.

[65] Danushka Bollegala, Kohei Hayashi, and Danushka Bollegala. Why does pairdiff work? - a mathematical analysis of bilinear relational compositional operators for analogy detection. In *Proc. of the 27th International Conference on Computational Linguistics (COLING)*, pages 2493–2504, 2018.

[66] Danushka Bollegala, Kohei Hayashi, and Ken-ichi Kawarabayashi. Think globally, embed locally — locally linear meta-embedding of words. In *Proc. of IJCAI-EACI*, pages 3970–3976, 2018.

[67] Danushka Bollegala, Yuichi Yoshida, and Ken-ichi Kawarabayashi. Using k-way Co-occurrences for Learning Word Embeddings. In *Proc. of the Thirty Second AAAI Conference on Artificial Intelligence (AAAI)*, pages 5037–5044, 2018.

[68] Joshua Coates and Danushka Bollegala. Frustratingly easy meta-embedding – computing meta-embeddings by averaging source word embeddings. In *Proc. of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 194–198, 2018.

[69] Xia Cui, Kojaku Sadamori, Naoki Masuda, and Danushka Bollegala. Solving feature spareness in text classification using core-periphery decomposition. In *Proc. of Seventh Joint Conference on Lexical and Computational Semantics*, pages 255–264, 2018.

[70] Khai Mai, Thai-Hoang Pham, Minh Trung Nguyen, Nguyen Tuan Duc, Danushka Bollegala, Ryohei Sasano, and Satoshi Sekine. Which model performs best in this situation? an empirical study on fine-grained named entity recognition. In *Proc. of the 27th International Conference on Computational Linguistics (COLING)*, pages 711–722, 2018.

[71] Angrosh Mandya, Danushka Bollegala, Frans Coenen, and Katie Atkinson. A dataset for inter-sentence relation extraction using distant supervision. In *Proc. of the Eleventh International Conference on Language Resources and Evaluation (LREC)*, pages 1–7, 2018.

[72] Pavithra Rajendran, Danushka Bollegala, and Simon Parsons. Is something better than nothing? automatically predicting stance-based arguments using deep learning and small labelled dataset. In *Proc. of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 28–34, 2018.

[73] Pavithra Rajendran, Danushka Bollegala, and Simon Parsons. Sentiment-stance-specificity (sss) dataset: Identifying support-based entailment among opinions. In *Proc. of the Eleventh International Conference on Language Resources and Evaluation (LREC)*, pages 1–7, 2018.

[74] Krasen Samardzhiev, Andrew Gargett, and Danushka Bollegala. Learning neural word salience scores. In *Proc. of the Seventh Joint Conference on Lexical and Computational Semantics*, pages 33–42, 2018.

[75] Abdullah Alsheri, Frans Coenen, , and Danushka Bollegala. Accurate continuous and non-intrusive user authentication with multivariate keystroke streaming. In *Proc. of the 9th International Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KDIR)*, pages 61–70, 2017.

[76] Abdullah Alsheri, Frans Coenen, , and Danushka Bollegala. Spectral analysis of keystroke streams: Towards effective real-time and continuous user authentication. In *Proc. of Workshop on Spatial and Spatio-temporal Data Mining (SSTDM) at the IEEE International Conference on Data Mining (SSTDM)*, pages 62–73, 2017.

[77] Abdullah Alsheri, Frans Coenen, , and Danushka Bollegala. Spectral analysis of keystroke streams: Towards effective real-time continuous user authentication. In *4th International Conference on Information Systems Security and Privacy (ICISSP)*, pages 62–73, 2017.

[78] Xia Cui, Frans Coenen, and Danushka Bollegala. Effect of data imbalance on unsupervised domain adaptation of part-of-speech tagging and pivot selection strategies. In *Proc. of the Wokshop on Learning With Imbalanced Domains: Theory and Applications (LIDTA) at the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)*, pages 103–115, 2017.

[79] Xia Cui, Frans Coenen, and Danushka Bollegala. Tsp: Learning task-specific pivots for unsupervised domain adaptation. In *Proc. of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)*, pages 754–771, 2017.

[80] Matias Garcia-Constantino, Katie Atkinson, Danushka Bollegala, Karl Chapman, Frans Coenen, Clare Roberts, and Katy Robson. Cliel: Context-based information extraction from commercial law documents. In *Proc. of the 16th International Conference on Artificial Intelligence and Law (ICAIL)*, pages 79–87, 2017.

[81] Huda Hakami and Danushka Bollegala. Discovering representative space for relational similarity measurement. In *Proc. of the 15th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, pages 76–87, 2017.

[82] Angrosh Mandya, Danushka Bollegala, Frans Coenen, and Katie Atkinson. Classifier-based pattern selection approach for relation instance extraction. In *Proc. of the International Conference on Computational Linguistics and Intelligent Text Processing (CICLing)*, pages 1–16. LNCS, Springer, 2017.

[83] Angrosh Mandya, Danushka Bollegala, Frans Coenen, and Katie Atkinson. Frame-based semantic patterns for relation extraction. In *Proc. of the 15th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, pages 51–62, 2017.

[84] Pavithra Rajendran, Danushka Bollegala, and Simon Parsons. Identifying argument based relation properties in opinions. In *Proc. of the 15th International Conference of the Pacific Association for Computational Linguistics (PACLING)*, pages 1–12, 2017.

[85] Abdullah Alsheri, Frans Coenen, , and Danushka Bollegala. Keyboard usage authentication using time series analysis. In *Proc. of 18th International Conference on Big Data Analytics and Knowledge Discovery (DaWak 2016)*, pages 239–252, 2016.

[86] Abdullah Alsheri, Frans Coenen, , and Danushka Bollegala. Towards keystroke continuous authentication using time series analytics. In *Proc. of the 36th SGAI International Conference on Artificial Intelligence (SGAI), Research and Development in Intelligent Systems XXXIII*, pages 325–339. Springer, 2016.

[87] Danushka Bollegala, Alsuhaibani Mohammed, Takanori Maehara, and Ken-ichi Kawarabayashi. Joint word representation learning using a corpus and a semantic lexicon. In *Proc. of AAAI*, pages 2690–2696, 2016.

[88] Pavithra Rajendran, Danushka Bollegala, and Simon Parsons. Assessing weight of opinion by aggregating coalitions of arguments. In *Proc. of 6th International Conference on Computational Models of Argument (COMMA)*, pages 431–438, 2016.

[89] Pavithra Rajendran, Danushka Bollegala, and Simon Parsons. Contextual stance classification of opinions: A step towards enthymeme reconstruction in online reviews. In *Proc. of the 3rd Workshop on Argument Mining*, pages 31–39, 2016.

[90] Danushka Bollegala, Takanori Maehara, and Ken-ichi Kawarabayashi. Embedding semantic relationas into word representations. In *Proc. of IJCAI*, pages 1222 – 1228, 2015.

[91] Danushka Bollegala, Takanori Maehara, and Ken-ichi Kawarabayashi. Unsupervised cross-domain word representation learning. In *Proc. of ACL*, pages 730 – 740, 2015.

[92] Danushka Bollegala, Takanori Maehara, Yuichi Yoshida, and Ken-ichi Kawarabayashi. Learning word representations from relational graphs. In *Proc. of 29th AAAI Conference on Aritificial Intelligence*, pages 2146 – 2152, 2015.

[93] Pascal Kuyten, Danushka Bollegala, Bernd Hollerit, Helmut Prendinger, and Kiyoharu Aizawa. A discourse search engine based on rhetorical structure theory. In *Proc. of 37th European Conference on Information Retrieval (ECIR 2015)*, pages 80–91, 2015.

[94] Danushka Bollegala, David Weir, and John Carroll. Learning to predict distributions of words across domains. In *Proc. of Association for Computational Linguistics (ACL)*, pages 613 – 623, 2014.

[95] Danushka Bollegala, Mitsuru Kusumoto, Yuichi Yoshida, and Ken-ichi Kawarabayashi. Mining for analogous tuples from an entity-relation graph. In *Proc. of International Joint Conferences on Artificial Intelligence*, pages 2064 – 2070, 2013.

[96] Muhammad Asif Hossain Khan, Danushka Bollegala, Guangwen Li, and Kaoru Sezaki. Multi-tweet summarization of real-time events. In *IEEE International Conference on Social Computing (SocialComp)*, pages 128 – 133, 2013.

[97] Hiroyuki Sato, Danushka Bollegala, Yoshihiko Hasegawa, and Hitoshi Iba. Learning non-linear ranking functions for web search using probabilistic model building gp. In *Proc. of IEEE Congress on Evolutionary Computation (CEC 2013)*, pages 3371 – 3378, 2013.

[98] Nozomi Nori, Danushka Bollegala, and Hisashi Kashima. Multinomial relation prediction in social data: A dimension reduction approach. In *AAAI'12*, pages 115 – 121, 2012.

[99] Hiroyuki Sato, Yoshihiko Hasegawa, Danushka Bollegala, and Hitoshi Iba. Probabilistic model building gp with belief propagation. In *Proc. of the IEEE Congress on Evolutionary Computation (CEC 2012)*, pages 1–8, 2012.

[100] Francisco Tacoa, Danushka Bollegala, and Mitsuru Ishizuka. A context expansion method for supervised word sense disambiguation. In *Proc. of 6th IEEE International Conference on Semantic Computing (ICSC 2012)*, pages 339–341, 2012.

[101] Ken-ichi Yokote, Danushka Bollegala, and Mitsuru Ishizuka. Similarity is not entailment – jointly learning similarity transformations for textual entailment. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, pages 1720 – 1726, 2012.

[102] Danushka Bollegala. A supervised ranking approach for detecting relationally similar word pairs. In *Proc. of the IEEE International Conference on Information and Automation for Sustainability (ICAfs)*, pages 323 – 328, 2011.

[103] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. From actors, politicians, to ceos: Domain adaptation of relational extractors using a latent relational mapping. In *Proc. of the 20th International World Wide Web Conference (WWW)*, pages 13 – 14, 2011.

[104] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Relation adaptation: Learning to extract novel relations with minimum supervision. In *Proc. of International Joint Conferences in Artificial Intelligence (IJCAI)*, pages 2205 – 2210, 2011.

[105] Danushka Bollegala, Nasimul Noman, and Hitoshi Iba. Rankde: Learning a ranking function for information retrieval using differential evolution. In *Proc. of Genetic and Evolutionary Computation Conference (GECCO)*, pages 1771–1778, 2011.

[106] Danushka Bollegala, David Weir, and John Carroll. Using multiple sources to construct a sentiment sensitive thesaurus for cross-domain sentiment classification. In *Proc. of Association for Computational Linguistics (ACL)*, pages 132 – 141, 2011.

[107] Nguyen Tuan Duc, Danushka Bollegala, and Mitsuru Ishizuka. Cross-language latent relational search: Mapping knowledge across languages. In *Proc. of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, pages 1237 – 1242, 2011.

[108] Hugo Hernault, Danushka Bollegala, and Mitsuru Ishizuka. Semi-supervised discourse relation classification with structural learning. In *Proc. of CICLing*, pages 340 – 352, 2011.

[109] Haibo Li, Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Using graph based method to improve bootstrapping relation extraction. In *Proc. of the Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, pages 127 – 138, 2011.

[110] Nasimul Noman, Danushka Bollegala, and Hitoshi Iba. An adaptive differential evolution algorithm. In *Proc. of the IEEE Congress on Evolutionary Computation (IEEE CEC)*, pages 2229 – 2236, 2011.

[111] Nasimul Noman, Danushka Bollegala, and Hitoshi Iba. Differential evolution with self adaptive local search. In *Proc. of the Genetic and Evolutionary Computation Conference (GECCO)*, pages 1099 – 1106, 2011.

[112] Nozomi Nori, Danushka Bollegala, and Mitsuru Ishizuka. Exploiting user interest on social media for aggregating diverse data and predicting interest. In *5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 241 – 248, 2011.

[113] Nozomi Nori, Danushka Bollegala, and Mitsuru Ishizuka. Interest prediction on multinomial, time-evolving social graphs. In *Proc. of International Joint Conferences in Artificial Intelligence (IJCAI)*, pages 2507–2512, 2011.

[114] Naoki Tani, Danushka Bollegala, Naiwala Chandrasiri, Keisuke Okamoto, Kazunari Nawa, Shuhei Iitsuka, and Yutaka Matsuo. Collaborative exploratory search in real-world context. In *20th ACM International Conference on Information and Knowledge Management*, pages 2137 – 2140, 2011.

[115] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Relational duality: Unsupervised extraction of semantic relations between entities on the web. In *Proc. of International World Wide Web Conference (WWW)*, pages 151 – 160, 2010.

[116] Tomokazu Goto, Nguyen Duc, Danushka Bollegala, and Mitsuru Ishizuka. Exploiting symmetry in relational similarity for ranking relational search results. In *Proc. of the 11th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2010)*, pages 595 – 600, 2010.

[117] Hugo Hernault, Danushka Bollegala, and Mitsuru Ishizuka. A semi-supervised approach to improve classification of infrequent discourse relations using feature vector extension. In *Empirical Methods in Natural Language Processing*, pages 399 – 409, 2010.

[118] Hugo Hernault, Danushka Bollegala, and Mitsuru Ishizuka. A sequential model for discourse segmentation. In *International Conference on Intelligence Text Processing and Computational Linguistics (CICLing)*, pages 315 – 326, 2010.

[119] Hugo Hernault, Danushka Bollegala, and Mitsuru Ishizuka. Towards semi-supervised classification of discourse relations using feature correlations. In *Proc. of the 11th Annual SIGdial Meeting on Discourse and Dialogue (SIGDIAL 2010)*, pages 55 – 58, 2010.

[120] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Measuring the similarity between implicit semantic relations from the web. In *Proc. of International World Wide Web Conference (WWW)*, pages 651 – 660, 2009.

[121] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Measuring the similarity between implicit semantic relations using web search engines. In *2nd International Conference on Web Search and Data Mining (WSDM'09)*, pages 104 – 113, 2009.

[122] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. A relational model of semantic similarity between words using automatically extracted lexical pattern clusters from the web. In *Empirical Methods in Natural Language Processing*, pages 803 – 812, 2009.

[123] Keigo Watanabe, Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. A two-step approach to extracting atrtibutes for people on the web. In *Proceedings of the 2nd Web People Search Evaluation Workshop (WePS-09) at WWW'09*, 2009.

[124] D. Bollegala, Y. Matsuo, and M. Ishizuka. Www sits the sat: Measuring relational similarity on the web. In *Proc. of European Conference on Artificial Intelligence (ECAI)*, pages 333–337, 2008.

[125] Danushka Bollegala, Taiki Honma, Yutaka Matsuo, and Mitsuru Ishizuka. Automatically extracting personal name aliases from the web. In *Proc. of the 6th International Conference on Natural Language Processing (GoTAL 2008)*, pages 77 – 88, 2008.

[126] Danushka Bollegala, Taiki Honma, Yutaka Matsuo, and Mitsuru Ishizuka. Identification of personal name aliases on the web. In *Proc. of WWW 2008 Workshop on Social Web Search and Mining*, 2008.

[127] Danushka Bollegala, Taiki Honma, Yutaka Matsuo, and Mitsuru Ishizuka. Mining for personal name aliases on the web. In *Proc. of the 17th International World Wide Web Conference (WWW 2008)*, pages 1107 – 1108, 2008.

[128] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. A co-occurrence graph-based approach for personal name alias extraction from anchor texts. In *Proc. of the 3rd International Joint Conferences on Natural Language Processing (IJCNLP 2008)*, pages 865 – 870, 2008.

[129] D. Bollegala, Y. Matsuo, and M. Ishizuka. An integrated approach to measuring semantic similarity between words using information available on the web. In *Proceedings of NAACL HLT*, pages 340–347, 2007.

[130] D. Bollegala, Y. Matsuo, and M. Ishizuka. An integrated approach to measuring semantic similarity between words using information available on the web. In *Proc. of HTL-NAACL'07*, pages 340–347, 2007.

[131] D. Bollegala, Y. Matsuo, and M. Ishizuka. Measuring semantic similarity between words using web search engines. In *Proc. of WWW '07*, pages 757–766, 2007.

[132] D. Bollegala, Y. Matsuo, and M. Ishizuka. Extracting key phrases to disambiguate personal name queries in web search. In *Proceedings of the Workshop on How Can Computational Linguistics Improve Information Retrieval*, pages 17–24, 2006.

[133] Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka. Disambiguating personal names on the web using automatically extracting key phrases. In *Proc. of the European Conference on Artificial Intelligence (ECAI)*, pages 553 – 557, 2006.

[134] Danushka Bollegala, Naoaki Okazaki, and Mitsuru Ishizuka. A bottom-up approach to sentence ordering for multi-document summarization. In *Proc. of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, pages 385 – 392, 2006.

[135] Y. Matsuo, M. Hamasaki, Y. Nakamura, T. Nishimura, K. Hasida, H. Takeda, J. Mori, D. Bollegala, and M. Ishizuka. Spinning multiple social networks for semantic web. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, volume 21, page 1381. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[136] D. Bollegala, N. Okazaki, and M. Ishizuka. A machine learning approach to sentence ordering for multidocument summarization and its evaluation. In *Proc. of International Joint Conferences in Natural Language Processing*, pages 624 – 635. Springer, 2005.