

Cornelia Caragea

Contact Information

Computer Science
University of Illinois at Chicago
3-190E Daley Library
Chicago, IL 60607, USA

Phone: 814-308-4974
Email: cornelia@uic.edu, cornelia.cse@gmail.com
URL: <https://www.cs.uic.edu/~cornelia/>
Address: 3-190E Daley Library, Chicago, IL 60607

Research Interests

Natural Language Processing, Deep Learning, Information Retrieval, and Artificial Intelligence.

Education

- Aug 2003 - Dec 2009: Ph.D. in Computer Science, Artificial Intelligence Research Laboratory, Iowa State University, Ames, IA, USA.
- Oct 1992 - Jul 1997: B.S. in Computer Science, Faculty of Mathematics and Computer Science, University of Bucharest, Bucharest, Romania.

Professional Experience

- Aug 2022 - present: Professor, Computer Science, University of Illinois Chicago, Chicago, IL, USA, and the Director of the Information Retrieval Research Laboratory at UIC.
- Aug 2018 - Aug 2022: Associate Professor, Computer Science, University of Illinois at Chicago, Chicago, IL, USA, and the Director of the Information Retrieval Research Laboratory at UIC.
- Aug 2018 - present: Adjunct Associate Professor, Computer Science, Kansas State University, Manhattan, KS, USA.
- Sept 2017 - Aug 2018: Associate Professor and The Lloyd T. Smith Creativity in Engineering Chair, Computer Science, Kansas State University, Manhattan, KS, USA.
- Sept 2012 - Aug 2017: Assistant Professor, Computer Science and Engineering, University of North Texas, Denton, TX, USA.
- Jul 2010 - May 2012: Post-Doctoral Research Associate, Intelligent Information Systems Laboratory, College of Information Sciences & Technology, Pennsylvania State University, University Park, PA, USA.
- Jan 2010 - Jun 2010: Post-Doctoral Research Associate, Artificial Intelligence Research Laboratory, Iowa State University, Ames, IA, USA.
- Jun 2006 - Dec 2009: Research Assistant, Computer Science, Iowa State University, Ames, IA, USA.
- Aug 2003 - May 2006: Teaching Assistant, Computer Science, Iowa State University, Ames, IA, USA.

Research Projects

My research interests lie at the intersection of natural language processing, deep learning, information retrieval, and artificial intelligence. The overarching goal of my research is to effectively and efficiently mine and discover knowledge from large amounts of data. In particular, I am interested in extracting information

for knowledge graph construction; modeling hierarchical structures for natural language understanding; and designing predictive models for social good using supervised, self-supervised, and domain adaptation.

Awarded Research Grants

1. NSF IIS 2107487: HCC: “Collaborative Research: Medium: Fine-grained Emotion Analysis in Crises.” NSF, Division of Information & Intelligent Systems. **Cornelia Caragea (PI). PI share: 100%**. Award amount: \$347,614. [Awarded]. *October 1, 2021 - September 30, 2025.*
2. NSF IIS 2107518: III: “Collaborative Research: Medium: Extracting and Linking AI Artifacts” NSF, Division of Information & Intelligent Systems. **Cornelia Caragea (PI). PI share: 100%**. Award amount: \$529,992. [Awarded]. *December 1, 2021 - November 30, 2024.*
3. NSF 2137846: Convergence Accelerator Track F: “America’s Fourth Estate at Risk: A System for Mapping the (Local) Journalism Life Cycle to Rebuild the Nation’s News Trust.” **Cornelia Caragea (Co-PI)**. Award amount: \$750,000. [Awarded]. **UIC share: \$195,000.** *October 1, 2021 - September 30, 2022.*
4. DARPA: HR001121S0007-03: “Intelligent Digital Thread.” **Cornelia Caragea (Co-PI)**. (Subcontract from InspiRD) Subcontract amount: \$45,347. [Awarded]. *July 1, 2021 - December 31, 2021.*
5. UIC Discovery Partners Institute (DPI) Seed Program: “Deep Learning for Keyphrase Extraction: Models and Corpora.” **Cornelia Caragea (PI). PI share: 100%**. Award amount: UIC-DPI: \$250,000. [Awarded]. *January 1, 2020 - June 30, 2022.*
6. NSF IIS 1652674: III: “CAREER: From Data to Knowledge: Extracting and Utilizing Scholarly Knowledge Graphs.” NSF, Division of Information & Intelligent Systems. **Cornelia Caragea (PI). PI share: 100%**. Award amount: \$551,876. [Awarded]. *June 1, 2017 - May 31, 2022.*
7. NSF CNS 1823292: CRI: CI-SUSTAIN: Collaborative Research: “CiteSeerX: Toward Sustainable Support of Scholarly Big Data.” NSF Computer and Network Systems. **Cornelia Caragea (PI). PI share: 100%**. Award amount: NSF: \$246,000. [Awarded]. *August 1, 2018 - July 31, 2022.*
8. NSF IIS 1741353: Big Data: “BIGDATA: IA: Collaborative Research: Domain Adaptation Approaches for Classifying Crisis Related Data in Social Media.” NSF, Big Data Science & Engineering Program, and Amazon Web Services. **Cornelia Caragea (PI). PI share: 100%**. Award amount: NSF: \$400,000 + AWS: \$384,458. [Awarded]. *January 1, 2018 - December 31, 2021.*
9. IMLS: “Programmatic Extraction of ‘Documents’ from Web Archives.” The Institute of Museum and Library Services (IMLS). Mark Phillips (PI); **Cornelia Caragea (Co-PI). Co-PI share: 45%**. Award amount: \$318,989. [Awarded]. *December 1, 2017 - November 30, 2020.*
10. NSF IIS 1526542: CHS: Small: Collaborative Research: “Automating Relevance and Trust Detection in Social Media Data for Emergency Response.” NSF, Division of Information & Intelligent Systems. **Cornelia Caragea (PI). PI share: 100%**. Award amount: \$250,000. [Awarded]. *September 1, 2015 - August 31, 2019.*
11. NSF CMMI 1541155: “CRISP Type 2: Collaborative Research: Resilience Analytics: A Data-Driven Approach for Enhanced Interdependent Network Resilience.” NSF, Division of CMMI. Andrea H. Tapia (PI); **Co-PIs: Cornelia Caragea, Christopher Zobel and Jessica Kropczynski.** Award amount: \$927,171. **Co-PI Caragea share: \$224,897.** [Awarded]. *October 15, 2015 - October 14, 2019.*
12. NSF IIS 1423337: III: Small: Collaborative Research: “Keyphrase Extraction in Document Networks.” NSF, Division of Information & Intelligent Systems. **Cornelia Caragea (PI); Co-PI: Paul Tarau.** Award amount: \$324,866. **PI share: 95%**. [Awarded]. *September 1, 2014 - August 31, 2018.*

13. NSF CNS 1421970: TWC: Small: Collaborative: “Towards Privacy Preserving Online Image Sharing.” NSF, Division of Computer and Network Systems. **Cornelia Caragea (PI)**. **PI share: 100%**. Award amount: \$244,736. [Awarded]. *September 1, 2014 - August 31, 2018.*
14. NSF IIS 1353418: EAGER: Collaborative Research: “Establishing Trustworthy-Citizen-Created Data for Disaster Response and Humanitarian Action.” NSF, Division of Information & Intelligent Systems. **Cornelia Caragea (PI)**. **PI share: 100%**. Award amount: \$74,999. [Awarded]. *September 1, 2013 - August 31, 2015.*
15. UNT ROP Seed Research Grant: “MetaSeer: An Automated Tool to Extract Information for Meta-Analyses.” Research grant sponsored by UNT. Kim Nimon (PI); **Co-PI: Cornelia Caragea**, University of North Texas. Award amount: \$7,300. [Awarded]. *September 1, 2013 - August 31, 2014.*
16. Lockheed Martin Corporation Research Grant: “Anomaly Detection from Social Media Streams.” Research grant sponsored by Lockheed Martin Corporation. Prasenjit Mitra (PI); **Co-PI: Cornelia Caragea**, Award amount: \$100,000. [Awarded]. *March 2011 - December 2011*

Publications and Research Reports

Ph.D. Thesis: “Abstraction-Based Probabilistic Models for Sequence Classification.”, 2009.

Refereed Journal Publications:

1. Edoardo Savini and **Cornelia Caragea**. (2022). “Intermediate-Task Transfer Learning with BERT for Sarcasm Detection.” *Mathematics 2022*, Special Issue on Natural Language Processing (NLP) and Machine Learning (ML) - Theory and Applications.
2. Ashwini Tonge and **Cornelia Caragea**. (2020). “Image Privacy Prediction Using Deep Neural Networks.” *ACM Transactions on the Web (ACM TWeb)*, 14(2), 1-32. doi:10.1145/3386082.
3. Ashwini Tonge and **Cornelia Caragea**. (2019). “Privacy-aware Tag Recommendation for Accurate Image Privacy Prediction.” *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 10(4), 1-28. doi:10.1145/3335054.
4. Reza Mazloom, Hongmin Li, Doina Caragea, **Cornelia Caragea**, and Muhammad Imran. (2019). “A Hybrid Domain Adaptation Approach for Identifying Crisis-Relevant Tweets.” *International Journal of Information Systems for Crisis Response and Management*, 11(2), 1-19. doi:10.4018/ijiscram.2019070101.
5. HongMin Li, Oleksandra Sopova, Doina Caragea, and **Cornelia Caragea**. (2018). “Domain Adaptation for Crisis Data Using Correlation Alignment and Self-Training.” *International Journal of Information Systems for Crisis Response and Management (IJISCRAM)*, 10(4), 1-20. doi:10.4018/ijiscram.2018100101.
6. Gilad Katz, **Cornelia Caragea**, and Asaf Shabtai. (2018). “Vertical Ensemble Co-Training for Text Classification.” *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 9(2), 1-23. doi:10.1145/3137114.
7. Anna Squicciarini, **Cornelia Caragea**, and Rahul Balakavi. (2017). “Toward Automated Online Photo Privacy.” *ACM Transactions on the Web (ACM TWeb)*, 11(1), 1-29. doi:10.1145/2983644.
8. Venkata Kishore Neppalli, **Cornelia Caragea**, Anna Squicciarini, Andrea H. Tapia, and Sam Stehle. (2017). “Sentiment Analysis during Hurricane Sandy in Emergency Response.” *International Journal of Disaster Risk Reduction (IJDRR)*, 21, 213-222. doi:10.1016/j.ijdr.2016.12.011.
9. Kash Barker, James H. Lambert, Christopher W. Zobel, Andrea H. Tapia, Jose E. Ramirez-Marquez, Laura Albert McLay, Charles D. Nicholson, and **Cornelia Caragea**. (2017). “Defining Resilience Analytics for Interdependent Cyber-Physical-Social Networks.” *Sustainable and Resilient Infrastructure*, 2(2), 59-67. doi:10.1080/23789689.2017.1294859.

10. Nick LaLone, Andrea H. Tapia, Christopher W. Zobel, **Cornelia Caragea**, Venkata Kishore Neppalli, and Shane E. Halse. (2017). “Embracing Human Noise as Resilience Indicator: Twitter as Power Grid Correlate.” *Sustainable and Resilient Infrastructure.*, 2(4), 1-10. doi:10.1080/23789689.2017.1328920.
11. Shane E. Halse, Andrea H. Tapia, Anna Squicciarini, and **Cornelia Caragea**. (2017). “An Emotional Step Toward Automated Trust Detection in Crisis Social Media.” *Information, Communication & Society.*, 21(2), 288-305. doi:10.1080/1369118X.2016.1272618.
12. Hongmin Li, Doina Caragea, **Cornelia Caragea**, and Nic Herndon. (2017). “Disaster Response Aided by Tweet Classification with a Domain Adaptation Approach.” *Journal of Contingencies and Crisis Management.*, 26(1), 16-27. doi:10.1111/1468-5973.12194.
13. Venkata Kishore Neppalli, **Cornelia Caragea**, Doina Caragea, Murilo Cerqueira Medeiros, Andrea H. Tapia, and Shane E. Halse. (2016). “Predicting Tweet Retweetability during Hurricane Disasters.” *International Journal of Information Systems for Crisis Response and Management*, 8(3), 32-50, 2017. doi:10.4018/ijiscram.2016070103.
14. Sujatha Das Gollapalli, **Cornelia Caragea**, Prasenjit Mitra and C. Lee Giles. (2015). “Using unlabeled data to improve researcher homepage classification.” *ACM Transactions on the Web, (ACM TWeb)*, 9(4), 1-32. doi:10.1145/2767135.
15. Jian Wu, Kyle Williams, Hung-Hsuan Chen, Madian Khabsa, **Cornelia Caragea**, Suppawong Tuarob, Alexander Ororbia, Douglas Jordan, Prasenjit Mitra and C. Lee Giles. (2015). “CiteSeerX: AI in a Digital Library Search Engine.” *Artificial Intelligence Magazine (AI Magazine)*, 36(3), 35-48. doi:10.1609/aimag.v36i3.2601.
16. Prakhar Biyani, Sumit Bhatia, **Cornelia Caragea**, and Prasenjit Mitra. (2014). “Using Non-lexical Features For Identifying Factual and Opinionative Threads in Online Forums.” *Knowledge-Based Systems*, 69, 170-178. doi:10.1016/j.knosys.2014.04.048.
17. Sumit Bhatia, **Cornelia Caragea**, Hung-Hsuan Chen, Jian Wu, Pucktada Treeratpituk, Zhaohui Wu, Madian Khabsa, Prasenjit Mitra, and C. Lee Giles. (2012). “Specialized Research Datasets in the CiteSeerx Digital Library.” *D-Lib Magazine*, 18(7/8). doi:10.1045/july2012-bhatia.
18. **Cornelia Caragea**, Adrian Silvescu, and Prasenjit Mitra. (2012). “Protein Sequence Classification using Feature Hashing.” *Proteome Science*, 10(Suppl 1), s14. doi:10.1186/1477-5956-10-s1-s14.
19. Rasna Walia, **Cornelia Caragea**, Ben Lewis, Fadi Towfic, Michael Terribilini, Yasser El-Manzalawy, Drena Dobbs, and Vasant Honavar. (2012). “Protein-RNA Interface Residue Prediction Using Machine Learning: An Assessment of the State of the Art.” *BMC Bioinformatics*, 13(1), 89. doi:10.1186/1471-2105-13-89. **Highly Accessed on BioMed Central website.**
20. **Cornelia Caragea**, Doina Caragea, Adrian Silvescu, and Vasant Honavar. (2010). “Semi-Supervised Prediction of Protein Subcellular Localization Using Abstraction Augmented Markov Models.” *BMC Bioinformatics, Special Issue on Machine Learning in Computational Biology (MLCB 2010)*, 11(Suppl 8), s6. doi:10.1186/1471-2105-11-s8-s6.
21. **Cornelia Caragea**, Jivko Sinapov, Drena Dobbs, and Vasant Honavar. (2009). “Mixture of experts models to exploit global sequence similarity on biomolecular sequence labeling.” *BMC Bioinformatics*, 10(Suppl 4), s4. doi:10.1186/1471-2105-10-s4-s4.
22. Fadi Towfic, **Cornelia Caragea**, David Gemperline, Drena Dobbs, and Vasant Honavar. (2010). “Struct-NB: Predicting protein-RNA binding sites using structural features.” *International Journal of Data Mining and Bioinformatics*, 4(1), 21-43. doi:10.1504/ijdmb.2010.030965.
23. **Cornelia Caragea**, Jivko Sinapov, Adrian Silvescu, Drena Dobbs, and Vasant Honavar. (2007). “Glycosylation site prediction using ensembles of Support Vector Machine classifiers.” *BMC Bioinformatics*, 8(1), 438. doi:10.1186/1471-2105-8-438. **Highly Accessed on BioMed Central website.**

Refereed Conference Publications:

1. Mobashir Sadat and **Cornelia Caragea**. (2022). “Hierarchical Multi-Label Classification of Scientific Documents.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)*, Abu Dhabi.
2. Mahshid Hosseini and **Cornelia Caragea**. (2022). “Calibrating Student Models for Emotion-related Tasks.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)*, Abu Dhabi.
3. Hongli Zhan, Tiberiu Sosea, **Cornelia Caragea**, and Junyi Jessy Li. (2022) “Why Do You Feel This Way? Summarizing Triggers of Emotions in Social Media Posts.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)*, Abu Dhabi.
4. Mobashir Sadat and **Cornelia Caragea**. (2022). “Learning to Infer from Unlabeled Data: A Semi-supervised Learning Approach for Robust Natural Language Inference.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP Findings 2022)*, Abu Dhabi.
5. Tiberiu Sosea and **Cornelia Caragea**. (2022). “Leveraging Training Dynamics and Self-Training for Text Classification.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP Findings 2022)*, Abu Dhabi.
6. Jishnu Ray Chowdhury, Seoyeon Park, Tuhin Kundu, and **Cornelia Caragea**. (2022) “KPDrop: Improving Absent Keyphrase Generation.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP Findings 2022)*, Abu Dhabi.
7. Krishna Garg, Jishnu Ray Chowdhury, and **Cornelia Caragea**. (2022) “Keyphrase Generation Beyond the Boundaries of Title and Abstract.” In: *Proceedings of The 2022 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP Findings 2022)*, Abu Dhabi.
8. Iustin Sirbu, Tiberiu Sosea, **Cornelia Caragea**, Doina Caragea, and Traian Rebedea. (2022). “Multi-modal Semi-supervised Learning for Disaster Tweet Classification.” In: *Proceedings of The 29th International Conference on Computational Linguistics (COLING 2022)*, Gyeongju, Republic of Korea.
9. Shweta Yadav and **Cornelia Caragea**. (2022). “Towards Summarizing Healthcare Questions in Low-Resource Setting.” In: *Proceedings of The 29th International Conference on Computational Linguistics (COLING 2022)*, Gyeongju, Republic of Korea.
10. Paul Landes, Kunal Patel, Sean S. Huang, Adam Webb, Barbara Di Eugenio, and **Cornelia Caragea**. (2022). “A New Public Corpus for Clinical Section Identification: MedSecId.” In: *Proceedings of The 29th International Conference on Computational Linguistics (COLING 2022)*, Gyeongju, Republic of Korea.
11. Seo Yeon Park and **Cornelia Caragea**. (2022). “A Data Cartography based MixUp for Pre-trained Language Models.” In: *Proceedings of the 2022 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2022)*. Seattle, Washington.
12. Seo Yeon Park and **Cornelia Caragea**. (2022). “On the Calibration of Pre-trained Language Models using Mixup Guided by Area Under the Margin and Saliency.” In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*. Dublin, Ireland.
13. Mobashir Sadat and **Cornelia Caragea**. (2022). “SciNLI: A Corpus for Natural Language Inference on Scientific Text.” In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*. Dublin, Ireland.
14. Tiberiu Sosea and **Cornelia Caragea** (2022). ”EnsyNet: A Dataset for Encouragement and Sympathy Detection.” In: *Proceedings of the 13th Language Resources and Evaluation Conference (LREC 2022)*. Marseille, France.

15. Tiberiu Sosea, Chau Thi Minh Pham, Alexander Tekle, **Cornelia Caragea**, and Junyi Jessie Li. (2022). "Emotion analysis and detection during COVID-19." In: *Proceedings of the 13th Language Resources and Evaluation Conference (LREC 2022)*. Marseille, France.
16. Stefan Cobeli, Ioan-Bogdan Iordache, Shweta Yadav, **Cornelia Caragea**, Liviu P. Dinu, and Dragos Iliescu. (2022). "Detecting Optimism in Tweets using Knowledge Distillation and Linguistic Analysis of Optimism." In: *Proceedings of the 13th Language Resources and Evaluation Conference (LREC 2022)*. Marseille, France.
17. Chenye Zhao, Jasmine Mangat, Sujay Koujalgi, Anna Squicciarini, and **Cornelia Caragea**. (2022). "PrivacyAlert: A Dataset for Image Privacy Prediction." In: *Proceedings of the Sixteenth International AAAI Conference on Web and Social Media (ICWSM 2022)*. Atlanta, Georgia.
18. Jishnu Ray Chowdhury and **Cornelia Caragea**. (2021). "Modeling Hierarchical Structures with Continuous Recursive Neural Networks." In: *Proceedings of the 38th International Conference on Machine Learning (ICML 2021)*, Online. (pp. 1975-1988). doi:proceedings.mlr.press/v139/chowdhury21a.
19. Yingjie Li, Chenye Zhao, and **Cornelia Caragea**. (2021). "Improving Stance Detection with Multi-Dataset Learning and Knowledge Distillation." In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)*, Online. [Accepted] (Long paper).
20. Mahshid Hosseini and **Cornelia Caragea**. (2021). "Distilling Knowledge for Empathy Detection." In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP Findings 2021)*, Online. [Accepted] (Long paper).
21. Chenye Zhao and **Cornelia Caragea**. (2021). "Knowledge Distillation with BERT for Image Tag-Based Privacy Prediction." In: *Proceedings of the Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, Online. [Accepted] (Long paper). **Won Best Student Paper Award**.
22. Tiberiu Sosea and **Cornelia Caragea**. (2021). "eMLM: A New Pre-training Objective for Emotion Related Tasks." In: *Proceedings of The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*, Online. (pp. 286–293). doi:10.18653/v1/2021.acl-short.38.
23. Kyle Glandt, Sarthak Khanal, Yingjie Li, Doina Caragea, and **Cornelia Caragea**. (2021). "Stance Detection in COVID-19 Tweets." In: *Proceedings of The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*, Online. (pp. 1596–1611). doi:10.18653/v1/2021.acl-long.127.
24. Yingjie Li, Tiberiu Sosea, Aditya Sawant, Ajith Jayaraman Nair, Diana Inkpen, and **Cornelia Caragea**. (2021). "P-Stance: A Large Dataset for Stance Detection in Political Domain." In: *Proceedings of The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Findings (ACL-IJCNLP Findings 2021)*, Online. (pp. 2355–2365). doi:10.18653/v1/2021.findings-acl.208.
25. Yingjie Li and **Cornelia Caragea**. (2021). "A Multi-Task Learning Framework for Multi-Target Stance Detection." In: *Proceedings of The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Findings (ACL-IJCNLP Findings 2021)*, Online. (pp. 2320–2326). doi:10.18653/v1/2021.findings-acl.204.
26. Ana Sabina Uban, **Cornelia Caragea**, and Liviu Dinu. (2021). "Studying the Evolution of Scientific Topics and their Relationships." In: *Proceedings of The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Findings (ACL-IJCNLP Findings 2021)*, Online. (pp. 1908–1922). doi:10.18653/v1/2021.findings-acl.167.

27. Yingjie Li and **Cornelia Caragea**. (2021). “Target-Aware Data Augmentation for Stance Detection.” In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2021)*, Online. (pp. 1850–1860). doi:2021.naacl-main.148.
28. Mina Valizadeh, Pardis Ranjbar-Noiey, **Cornelia Caragea**, and Natalie Parde. (2021). “Identifying Medical Self-Disclosure in Online Communities.” In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2021)*, Online. (pp. 4398–4408). doi:2021.naacl-main.347.
29. Mahshid Hosseini and **Cornelia Caragea**. (2021). “It Takes Two to Empathize: One to Seek and One to Provide.” In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2021)*, Online. (pp. 13018–13026). <https://ojs.aaai.org/index.php/AAAI/article/view/17539>.
30. Krutarth Patel, **Cornelia Caragea**, Doina Caragea, and C. Lee Giles. (2021). “Author Homepage Discovery in CiteSeerX.” In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2021)*, Online. (pp. 15146–15155). <https://ojs.aaai.org/index.php/AAAI/article/view/17778>.
31. Krutarth Patel and **Cornelia Caragea**. (2021). “Exploiting Position and Contextual Word Embeddings to Identify Descriptive Phrases in Scientific Papers.” In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume (EACL 2021)*, Online. (pp. 1585–1591). doi:2021.eacl-main.136.
32. Tiberiu Sosea, Iustin Sarbu, **Cornelia Caragea**, Doina Caragea, and Traian Rebedea. (2021). “Using the Image-Text Relationship to Improve Multimodal Disaster Tweet Classification.” In: *Proceedings of The 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2021)*, Online. (pp. 691–704). <http://idl.iscram.org/files/2021/2365>.
33. Hongmin Li, Doina Caragea and **Cornelia Caragea**. (2021). “Combining Self-training with Deep Learning for Disaster Tweet Classification.” In: *Proceedings of The 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2021)*, Online. (pp. 719–730). <http://idl.iscram.org/files/2021/2367>.
34. Tiberiu Sosea and **Cornelia Caragea**. (2020). “CancerEmo: A Dataset for Fine-Grained Emotion Detection.” In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)*, Online. (pp. 8892–8904). 10.18653/v1/2020.emnlp-main.715.
35. Seo Yeon Park and **Cornelia Caragea**. (2020). “Scientific Keyphrase Identification and Classification by Pre-Trained Language Models Intermediate Task Transfer Learning.” In: *Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)*, Barcelona, Spain (Online). (pp. 5409–5419). 10.18653/v1/2020.coling-main.472.
36. Shrey Desai, **Cornelia Caragea**, Junyi Jessy Li. (2020). “Detecting Perceived Emotions in Hurricane Disasters.” In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*, Online. (pp. 5290–5305). doi:10.18653/v1/2020.acl-main.471.
37. Jishnu Ray Chowdhury, **Cornelia Caragea**, Doina Caragea. (2020). “Cross-Lingual Disaster-related Multi-label Tweet Classification with Manifold Mixup.” In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop (ACL-SRW 2020)*, Online. (pp. 292–298). doi:10.18653/v1/2020.acl-srw.39.
38. Jishnu Ray Chowdhury, **Cornelia Caragea**, Doina Caragea. (2020). “On Identifying Hashtags in Disaster Twitter Data.” In: *Proceedings of the 34th American Association for Artificial Intelligence (AAAI 2020)*, New York, USA. (pp. 498–506). doi:10.1609/aaai.v34i01.5387.
39. Shaurya Rohatgi, Zeba Karishma, Jason Chhay, Sai Raghav Reddy Keesara, Jian Wu, **Cornelia Caragea**, and C. Lee Giles. (2020). “COVIDSeer: Extending the COVID-19 Dataset.” In: *Proceedings of the 20th ACM Symposium on Document Engineering (DocEng 2020)*, Online. (pp. 1–4). doi:10.1145/3395027.3419597.

40. Krutarth Patel, **Cornelia Caragea**, Mark Phillips, and Nathaniel Fox. (2020). “Identifying Documents In-Scope of a Collection from Web Archives.” In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL 2020)*, Online. (pp. 167–176). doi:10.1145/3383583.3398540. **Won Best Student Paper Award.**
41. Krutarth Patel, **Cornelia Caragea**, Jian Wu, and C. Lee Giles. (2020). “Keyphrase Extraction in Scholarly Digital Library Search Engines.” In: *Proceedings of the International Conference on Web Services (ICWS 2020)*, Online. (pp. 179-196). doi:10.1007/978-3-030-59618-7_12.
42. Krutarth Patel, **Cornelia Caragea**, Mark Phillips. (2020). “Dynamic Classification in Web Archiving Collections.” In: *Proceedings of the 12th Language Resources and Evaluation Conference (LREC 2020)*, Online. (pp. 1459-1468). doi:2020.lrec-1.182.
43. **Cornelia Caragea**, Ana Sabina Uban, and Liviu P. Dinu. (2019). “The Myth of Double-Blind Review Revisited: ACL vs. EMNLP.” In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP/IJCNLP 2019)*, Hong Kong, China. (pp. 2317–2327). doi:10.18653/v1/D19-1236.
44. Yingjie Li and **Cornelia Caragea**. (2019). “Multi-Task Stance Detection with Sentiment and Stance Lexicons.” In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP/IJCNLP 2019)*, Hong Kong, China. (pp. 6299–6305). doi:10.18653/v1/D19-1657.
45. Krutarth Patel and **Cornelia Caragea**. (2019). “Exploring Word Embeddings in CRF-based Keyphrase Extraction from Research Papers.” In: *Proceedings of The 10th International Conference on Knowledge Capture (K-Cap 2019)*, Marina del Rey, California, USA. (pp. 37–44). doi:10.1145/3360901.3364447.
46. Mahnaz Roshanaei, Christopher Tran, Sylvia Morelli, **Cornelia Caragea**, and Elena Zheleva. (2019). “Paths to empathy: heterogeneous effects of reading personal stories online.” In: *Proceedings of The 6th IEEE International Conference on Data Science and Advanced Analytics (DSAA 2019)*, Washington DC, USA, 2019. (pp. 570-579). 10.1109/DSAA.2019.00072.
47. Ashwini Tonge and **Cornelia Caragea**. (2019). “Dynamic Deep Multi-modal Fusion for Image Privacy Prediction.” In: *Proceedings of The World Wide Web Conference (WWW 2019)*, San Francisco, California, USA, 2019. (pp. 1829–1840). doi:10.1145/3308558.3313691.
48. Rabah Al-Zaidy, **Cornelia Caragea**, and C. Lee Giles. (2019). “Bi-LSTM-CRF Sequence Labeling for Keyphrase Extraction from Scholarly Documents.” In: *Proceedings of The World Wide Web Conference (WWW 2019)*, San Francisco, California, USA. (pp. 2551–2557). doi:10.1145/3308558.3313642.
49. Jishnu Ray Chowdhury, **Cornelia Caragea**, and Doina Caragea. (2019). “Keyphrase Extraction from Disaster-related Tweets.” In: *Proceedings of The World Wide Web Conference (WWW 2019)*, San Francisco, California, USA. (pp. 1555–1566). doi:10.1145/3308558.3313696.
50. Yingjie Li, Seo Yeon Park, **Cornelia Caragea**, Doina Caragea, and Andrea Tapia. (2019). “Sympathy Detection in Disaster Twitter Data.” In: *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019)*, Valencia, Spain. (pp. 788-798). <http://idl.iscram.org/files/2019/1899>.
51. Xukun Li, Doina Caragea, **Cornelia Caragea**, Muhammad Imran, and Ferda Offi. (2019). “Identifying Disaster Damage Images Using a Domain Adaptation Approach.” In: *Proc. of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019)*, Valencia, Spain. (pp. 633-645). [Received Best Insight Paper Award.] <http://idl.iscram.org/files/2019/1853>.
52. Jess Kropczynski, Rob Grace, Shane Halse, Doina Caragea, **Cornelia Caragea**, Andrea Tapia. (2019). “Refining a Coding Scheme to Identify Actionable Information on Social Media.” In: *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019)*, Valencia, Spain. (pp. 916-922). <http://idl.iscram.org/files/2019/1981>.

53. Athar Sefid, Jian Wu, Allen C. Ge, Jing Zhao, Lu Liu, **Cornelia Caragea**, Prasenjit Mitra, and C. Lee Giles. (2019). “Cleaning Noisy and Heterogeneous Metadata for Record Linking Across Scholarly Big Datasets.” In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2019)*, Honolulu, Hawaii, USA. 33(01), (pp. 9601-9606). doi:10.1609/aaai.v33i01.33019601.
54. HongMin Li, Xukun Li, Doina Caragea, and **Cornelia Caragea**. (2018). “Comparison of Word Embeddings and Sentence Encodings as Generalized Representations for Crisis Tweet Classification Tasks.” In: *Proceedings of ISCRAM Asia Pacific 2018 Conference (ISCRAM Asia Pacific 2018)*, Wellington, New Zealand. (pp. 1-13). <http://idl.iscram.org/files/2018/1689>.
55. **Cornelia Caragea**, Liviu P. Dinu, and Bogdan Dumitru. (2018). “Exploring Optimism and Pessimism in Twitter Using Deep Learning.” In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*, Brussels, Belgium. (pp. 652–658). doi:10.18653/v1/D18-1067.
56. Hamed Khanpour and **Cornelia Caragea**. (2018). “Fine-Grained Emotion Detection in Health-Related Online Posts.” In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)*, Brussels, Belgium. (pp. 1160–1166). doi:10.18653/v1/D18-1147.
57. **Cornelia Caragea** and Corina Florescu. (2018). “Venue Classification of Research Papers in Scholarly Digital Libraries.” In: *Proceedings of the 22nd International Conference on Theory and Practice of Digital Libraries (TPDL 2018)*, Porto, Portugal. (pp. 129-136). doi:10.1007/978-3-030-00066-0_11.
58. Hamed Khanpour and **Cornelia Caragea**. (2018). “Fine-Grained Information Identification in Health Related Posts.” In: *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018)*, Ann Arbor, Michigan, USA. (pp. 1001–1004). doi:10.1145/3209978.3210132.
59. Ashwini Tonge, **Cornelia Caragea**, and Anna Squicciarini. (2018). “Privacy-Aware Tag Recommendation for Image Sharing.” In: *Proceedings of the 29th ACM Conference on Hypertext and Social Media (HT 2018)*, Baltimore, Maryland, USA. (pp. 52–56). doi:10.1145/3209542.3209574.
60. Venkata Kishore Neppalli, **Cornelia Caragea**, Doina Caragea. (2018). “Deep Neural Networks versus Naive Bayes Classifiers for Identifying Informative Tweets during Disasters.” In: *Proceedings of the 15th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2018)*, Rochester NY, USA. (pp. 1-10). <http://idl.iscram.org/files/2018/2141>.
61. Reza Mazloom, HongMin Li, Doina Caragea, Muhammad Imran, and **Cornelia Caragea**. (2018). “Classification of Twitter Disaster Data Using a Hybrid Feature-Instance Adaptation Approach.” *Proc. of the 15th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2018)*, Rochester NY, USA. (pp. 1-9). <http://idl.iscram.org/files/2018/2146>.
62. Hamed Khanpour, **Cornelia Caragea**, and Prakhar Biyani. (2017). “Identifying Empathetic Messages in Online Health Communities.” In: *Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP 2017)*, Taipei, Taiwan. (pp. 246–251). doi:117-2042.
63. Anna Squicciarini, Andrea Novelli, Dan Lin, **Cornelia Caragea**, and Haoti Zhong. (2017). “From Tag to Protect: A Tag-Driven Policy Recommender System for Image Sharing.” In: *Proceedings of the 15th Annual Conference on Privacy, Security and Trust (PST 2017)*, Calgary, Canada. (pp. 337-346). doi:10.1109/PST.2017.00047.
64. Haoti Zhong, Anna Squicciarini, David Miller, and **Cornelia Caragea**. (2017). “A Group-Based Personalized Model for Image Privacy Classification and Labeling.” In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI 2017)*, Melbourne, Australia. (pp. 3952-3958). doi:10.24963/ijcai.2017/552.
65. Corina Florescu and **Cornelia Caragea**. (2017). “PositionRank: An Unsupervised Approach to Keyphrase Extraction from Scholarly Documents.” In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017)*, Vancouver, Canada. (pp. 1105–1115). doi:10.18653/v1/P17-1102.

66. HongMin Li, Doina Caragea, and **Cornelia Caragea**. (2017). “Towards Practical Usage of a Domain Adaptation Algorithm in the Early Hours of a Disaster.” In: *Proceedings of the 14th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2017)*, Albi, France. (pp. 692-704). <http://idl.iscram.org/files/2017/2057>.
67. Corina Florescu and **Cornelia Caragea**. (2017). “A New Scheme for Scoring Phrases in Unsupervised Keyphrase Extraction.” In: *Proceedings of the 39th European Conference on Information Retrieval (ECIR 2017)*, Aberdeen, Scotland, UK. (pp. 477-483). doi:10.1007/978-3-319-56608-5_37.
68. Lucas Sterckx, **Cornelia Caragea**, Thomas Demeester, and Chris Develder. (2016). “Supervised Keyphrase Extraction as Positive Unlabeled Learning.” In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*, Austin, Texas. (pp. 1924-1929). doi:10.18653/v1/D16-1198.
69. **Cornelia Caragea**, Jian Wu, Sujatha Das Gollapalli, and C. Lee Giles. (2016). “Document Type Classification in Online Digital Libraries.” In: *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI 2016)*, Phoenix, AZ, USA. (pp. 3997-4002). doi:10.5555/3016387.3016469.
70. Venkata Kishore Neppalli, **Cornelia Caragea**, Robin Mayes, Kim Nimon, and Fred Oswald. (2016). “MetaSeer.STEM: Towards Automating Meta-Analyses.” In: *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI 2016)*, Phoenix, AZ. (pp. 4035-4040). doi:10.5555/3016387.3016475.
71. Haoti Zhong, Hao Li, Anna Squicciarini, Sarah Rajtmajer, Christopher Griffin, David Miller, and **Cornelia Caragea**. (2016). “Content-driven Detection of Cyberbullying on the Instagram Social Network.” In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, New York City, USA. (pp. 3952-3958). doi:10.5555/3061053.3061172.
72. Venkata Kishore Neppalli, Murilo Cerqueira Medeiros, **Cornelia Caragea**, Doina Caragea, Andrea Tapia, and Shane Halse. (2016). “Retweetability Analysis and Prediction during Hurricane Sandy.” In: *Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2016)*, Rio de Janeiro, Brazil. <http://idl.iscram.org/files/2016/1389>.
73. **Cornelia Caragea**, Adrian Silvescu, Andrea Tapia. (2016). “Identifying Informative Messages in Disasters using Convolutional Neural Networks.” In: *Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2016)*, Rio de Janeiro, Brazil. <http://idl.iscram.org/files/2016/1397>.
74. Shane Halse, Andrea Tapia, Anna Squicciarini, **Cornelia Caragea**. (2016). “An Emotional Step Towards Automated Trust Detection in Crisis Social Media.” In: *Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2016)*, Rio de Janeiro, Brazil. <http://idl.iscram.org/files/2016/1414>.
75. Shane Halse, Andrea Tapia, Anna Squicciarini, **Cornelia Caragea**. (2016). “Tweet Factors Influencing Trust and Usefulness During Both Man-Made and Natural Disasters.” In: *Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2016)*, Rio de Janeiro, Brazil. <http://idl.iscram.org/files/2016/1403>.
76. **Cornelia Caragea**, Florin Bulgarov, and Rada Mihalcea. (2015). “Co-Training for Topic Classification of Scholarly Data.” In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, Lisbon, Portugal. [Full Oral Presentation] (pp. 2357-2366). doi:10.18653/v1/D15-1283.
77. Jian Wu, Jason Killian, Huaiyu Yang, Kyle Williams, Sagnik Ray Choudhury, Suppawong Tuarob, **Cornelia Caragea**, C. Lee Giles. “PDFMEF: A Multi-Entity Knowledge Extraction Framework for Scholarly Documents and Semantic Search.” In: *Proceedings of the 8th International Conference on Knowledge Capture (K-Cap 2015)*, Palisades, NY, USA. [Nominated for best paper award.] (pp. 1-8). doi:10.1145/2815833.2815834.

78. Andreea Godea, **Cornelia Caragea**, Florin Bulgarov, Suhasini Ramisetty-Mikler. (2015). “An Analysis of Twitter Data on E-cigarette Sentiments and Promotion.” In: *Proceedings of the 15th Conference on Artificial Intelligence in Medicine (AIME 2015)*, Pavia, Italy. (pp. 205-215). doi:10.1007/978-3-319-19551-3-27.
79. HongMin Li, Nicolais Guevara, Nic Herndon, Doina Caragea, Kishore Neppalli, **Cornelia Caragea**, Anna Squicciarini, Andrea H. Tapia. (2015). “Twitter Mining for Disaster Response: A Domain Adaptation Approach.” In: *Proceedings of the 12th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2015)*, Kristiansand, Norway. (short paper).
80. Prakhar Biyani, **Cornelia Caragea**, Narayan Bhamidipati. (2015). “Entity-Specific Sentiment Classification of Yahoo News Comments.” In: *TechPulse* Yahoo internal conference (**Yahoo TechPulse 2015**), Santa Clara, CA, USA. (short paper).
81. Prakhar Biyani, Sumit Bhatia, **Cornelia Caragea**, Prasenjit Mitra. (2015). “Using Subjectivity Analysis to Improve Thread Retrieval in Online Forums.” In: *Proceedings of the 37th European Conference on Information Retrieval (ECIR 2015)*, Vienna, Austria, 2015. (pp. 495-500). doi:10.1007/978-3-319-16354-3-54.
82. **Cornelia Caragea**, Florin Bulgarov, Andreea Godea, and Sujatha Das Gollapalli. (2015). “Citation-Enhanced Keyphrase Extraction from Research Papers: A Supervised Approach.” In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (EMNLP 2014)*, Doha, Qatar. (pp. 1435-1446). doi:10.3115/v1/D14-1150.
83. Brenda Reyes Ayala and **Cornelia Caragea**. (2014). “Towards Building a Collection of Web Archiving Research Articles.” In: *The 2014 Annual Meeting of the Association for Information Science and Technology (ASIS&T 2014)*, Seattle, WA, USA. (pp. 1-5). doi:10.1002/meet.2014.14505101150.
84. Brandon Truong, **Cornelia Caragea**, Anna Squicciarini, Andrea H. Tapia. (2014). “Extracting Valuable Information from Twitter During Natural Disasters.” In: *The 2014 Annual Meeting of the Association for Information Science and Technology (ASIS&T 2014)*, Seattle, WA, USA. (pp. 1-4). doi:10.1002/meet.2014.14505101162.
85. Prakhar Biyani, **Cornelia Caragea**, Prasenjit Mitra, and John Yen. (2014). “Identifying Emotional and Informational Support in Online Health Communities.” In: *Proceedings of the 25th International Conference on Computational Linguistics (Coling 2014)*, Dublin, Ireland. (pp. 827-836). doi:C14-1079.
86. Anna Squicciarini, **Cornelia Caragea**, and Rahul Balakavi. (2014). “Analysing Images’ Privacy for the Modern Web.” In: *Proceedings of the 25th ACM Conference on Hypertext and Social Media (HT 2014)*, Santiago, Chile. (pp. 136-147). doi:10.1145/2631775.2631803.
87. Sujatha Das G. and **Cornelia Caragea**. (2014). “Extracting Keyphrases from Research Papers using Citation Networks.” In: *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI 2014)*, Québec City, Québec, Canada, 2014. (pp. 1629-1635). doi:10.5555/2892753.2892779. [Full Oral Presentation]
88. Jian Wu, Kyle Williams, Hung-Hsuan Chen, Madian Khabisa, **Cornelia Caragea**, Alexander Ororbia, Douglas Jordan, and C. Lee Giles. (2014). “CiteSeerX: AI in a Digital Library Search Engine.” In: *Proceedings of the 26th Annual Conference on Innovative Applications of Artificial Intelligence (IAAI 2014)*, co-located with AAAI 2014, Quebec City, Quebec, Canada, 2014. **Won the “Most Innovative Application of AI” Award.** (pp. 2930-2937). <https://www.aaai.org/ocs/IAAI/IAAI14/8607/8660>.
89. Siddhartha Banerjee, **Cornelia Caragea**, and Prasenjit Mitra. (2014). “Playscript Classification and Automatic Wikipedia Play Articles Generation.” In: *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR 2014)*, Stockholm, Sweden. (pp. 3630-3635). doi:10.1109/ICPR.2014.624.
90. **Cornelia Caragea**, Anna Squicciarini, Sam Stehle, Kishore Neppalli, Andrea H. Tapia. (2014). “Mapping Moods: Geo-Mapped Sentiment Analysis During Hurricane Sandy.” In: *Proceedings of the 11th*

International Conference on Information Systems for Crisis Response and Management (ISCRAM 2014), University Park, Pennsylvania, USA. (pp. 642-651). <http://idl.iscram.org/files/2014/372>.

91. **Cornelia Caragea**, Jian Wu, Alina Ciobanu, Kyle Williams, Juan Fernández-Ramírez, Hung-Hsuan Chen, Zhaohui Wu, and C. Lee Giles. (2014). “CiteSeerX: A Scholarly Big Dataset.” In: *Proceedings of the 36th European Conference on Information Retrieval (ECIR 2014)*, Amsterdam, Netherlands. (pp. 311-322). doi:10.1007/978-3-319-06028-6_26.
92. Prakhar Biyani, **Cornelia Caragea**, Prasenjit Mitra, Chong Zhou, John Yen, Greta Greer, Kenneth Portier. (2013). “Co-training Over Domain-dependent and Domain-independent Features for Sentiment Analysis of an Online Cancer Support Community.” In: *The 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*, Niagara Falls, Canada. (pp. 413-417). doi:10.1145/2492517.2492606.
93. **Cornelia Caragea**, Adrian Silvescu, Prasenjit Mitra, and C. Lee Giles. (2013). “Can’t see the forest for the trees? A citation recommendation system.” In: *Proceedings of ACM/IEEE Joint Conference on Digital Libraries (JCDL 2013)*, Indianapolis, Indiana, USA. (pp. 111-114). doi:10.1145/2467696.2467743.
94. Sujatha Das G., **Cornelia Caragea**, Prasenjit Mitra and C. Lee Giles. (2013). “Researcher Homepage Classification using Unlabeled Data.” In: *Proceedings of the 22nd International Conference on World Wide Web (WWW 2013)*, Rio de Janeiro, Brazil. [Acceptance Rate: 15%] (pp. 471-482). doi:10.1145/2488388.2488430.
95. Prakhar Biyani, **Cornelia Caragea**, Prasenjit Mitra. (2013). “Predicting Subjectivity Orientation of Online Forum Threads.” In: *The 14th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2013)*, Samos, Greece. (pp. 109-120). doi:10.1007/978-3-642-37256-8_10.
96. Nir Ofek, **Cornelia Caragea**, Lior Rokach, Prakhar Biyani, Prasenjit Mitra, John Yen, Kenneth Portier, Greta Greer. (2013). “Improving Sentiment Analysis in an Online Cancer Survivor Community Using Dynamic Sentiment Lexicon.” In: *Proceedings of the International Conference on Social Intelligence and Technology (SOCIETY 2013)*, State College, PA, USA. (pp. 109-113). doi:10.1109/SOCIETY.2013.20.
97. Prakhar Biyani, Sumit Bhatia, **Cornelia Caragea**, Prasenjit Mitra. (2012). “Thread Specific Features Are Helpful For Identifying Subjectivity Orientation of Online Forum Threads.” In: *The 24th International Conference on Computational Linguistics (COLING 2012)*, Mumbai, India. [Acceptance Rate: 18.82%] (pp. 295-310). <https://www.aclweb.org/anthology/C12-1019>.
98. Wenyi Huang, Saurabh Kataria, **Cornelia Caragea**, Prasenjit Mitra, C. Lee Giles, Lior Rokach. (2012). “Recommending Citations: Translating Papers into References.” In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*, Maui, Hawaii. [Acceptance Rate: 14.3%] (pp. 1910-1914). doi:10.1145/2396761.2398542.
99. Prakhar Biyani, **Cornelia Caragea**, Amit K. Singh, Prasenjit Mitra. (2012). “I want what I need! Analyzing Subjectivity of Online Forum Threads.” In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*, Maui, Hawaii. (pp. 2495-2498). doi:10.1145/2396761.2398675.
100. **Cornelia Caragea**, Adrian Silvescu, and Prasenjit Mitra. (2012). “Combining Hashing and Abstraction in Sparse High Dimensional Feature Spaces.” In: *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI 2012)*, Toronto, Ontario, Canada. (pp. 3-9). doi:10.5555/2900728.2900729.
101. Kang Zhao, Baojun Qiu, **Cornelia Caragea**, Dinghao Wu, Prasenjit Mitra, John Yen, Greta Greer, and Kenneth Portier. (2011). “Identifying Leaders in an Online Cancer Survivor Community.” In: *Proceedings of the 21st Annual Workshop on Inf. Technologies and Systems (WITS 2011)*, Shanghai, China, 2011.

102. **Cornelia Caragea**, Adrian Silvescu, and Prasenjit Mitra. (2011). “Protein Sequence Classification Using Feature Hashing.” In: *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2011)*, Atlanta, GA, USA. (pp. 538-543). doi: 10.1109/BIBM.2011.91.
103. Baojun Qiu, Kang Zhao, Prasenjit Mitra, Dinghao Wu, **Cornelia Caragea**, John Yen, Greta Greer, Kenneth Portier. (2011). “Get Online Support, Feel Better-Sentiment Analysis and Dynamics in an On-line Cancer Survivor Community.” In: *Proceedings of the 2011 IEEE 3rd International Conference on Social Computing (SocialCom 2011)*, Boston, Massachusetts, USA. (pp. 274-281). doi:10.1109/PASSAT/SocialCom.2011.127.
104. **Cornelia Caragea**, Adrian Silvescu, Saurabh Kataria, Doina Caragea, and Prasenjit Mitra. (2011). “Classifying Scientific Publications Using Abstract Features.” In: *Proceedings of the Symposium on Abstraction, Reformulation, and Approximation (SARA 2011)*, Parador de Cardona, Spain.
105. Saurabh Kataria, Prasenjit Mitra, **Cornelia Caragea**, and C. Lee Giles. (2011). “Context Sensitive Topic Models for Author Influence in Document Networks.” In: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, Barcelona, Spain. (pp. 2274–2280). <https://www.ijcai.org/Proceedings/11/Papers/379.pdf>.
106. Sujatha Das G., C. Lee Giles, Prasenjit Mitra, and **Cornelia Caragea**. (2011). “On Identifying Academic Homepages for Digital Libraries.” In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL 2011)*, Ottawa, Canada. (pp. 123–132). doi:10.1145/1998076.1998099.
107. **Cornelia Caragea**, Nathan McNeese, Anuj Jaiswal, Greg Traylor, Hyun-Woo Kim, Prasenjit Mitra, Dinghao Wu, Andrea H. Tapia, C. Lee Giles, Bernard J. Jansen, John Yen. (2011). “Classifying Text Messages for the Haiti Earthquake.” In: *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2011)*, Lisbon, Portugal. (pp. 1-10). <http://idl.iscram.org/files/2011/371>.
108. **Cornelia Caragea**, Adrian Silvescu, Doina Caragea, and Vasant Honavar. (2010). “Abstraction Augmented Markov Models.” In: *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM 2010)*, Sydney, Australia. (pp. 68-77). doi:10.1109/ICDM.2010.158.
109. **Cornelia Caragea**, Adrian Silvescu, Doina Caragea, and Vasant Honavar. (2010). “Semi-Supervised Sequence Classification Using Abstraction Augmented Markov Models.”, In: *Proceedings of the ACM International Conference on Bioinformatics and Computational Biology (ACM-BCB 2010)*, Niagara Falls, New York, USA. (pp. 257–264). doi:10.1145/1854776.1854813.
110. Adrian Silvescu, **Cornelia Caragea**, and Vasant Honavar. (2009). “Combining Super-Structuring and Abstraction on Sequence Classification.” In: *Proceedings of the 9th IEEE International Conference on Data Mining (ICDM 2009)*, Miami, Florida, USA. (pp. 986-991). doi:10.1109/ICDM.2009.130.
111. **Cornelia Caragea**, Doina Caragea, and Vasant Honavar. (2009). “Learning Link-Based Naïve Bayes Classifiers from Ontology-Extended Distributed Data.” In: *Proceedings of the 21st International Conference on Tools with Artificial Intelligence (ICTAI 2009)*, Newark, New Jersey, USA. (pp. 1139-1146). doi:10.1007/978-3-642-05151-7_28.
112. **Cornelia Caragea**, Jivko Sinapov, Drena Dobbs, and Vasant Honavar. (2008). “Using Global Sequence Similarity to Enhance Macromolecular Sequence Labeling.” In: *Proceedings of the IEEE International Conference on Bioinformatics and BioMedicine (BIBM 2008)*, Philadelphia, PA, USA. (pp. 104-111).
113. Neeraj Koul, **Cornelia Caragea**, Vasant Honavar, Vikas Bahirwani, and Doina Caragea. (2008). “Learning Classifiers from Large Databases Using Statistical Queries.” In: *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2008)*, Sydney, Australia. (pp. 923–926). doi:10.1109/WIIAT.2008.366.
114. Jae-Hyung Lee, Michael Hamilton, Colin Gleeson, **Cornelia Caragea**, Peter Zaback, Jeffrey Sander, Xue Li, Feihong Wu, Michael Terribilini, Vasant Honavar, and Drena Dobbs. (2008). “Striking Similarities

in Diverse Telomerase Proteins Revealed by Combining Structure Prediction and Machine Learning Approaches.” In: *Proceedings of the Pacific Symposium on Biocomputing (PSB 2008)*, The Big Island of Hawaii. (pp. 501-512). <https://pubmed.ncbi.nlm.nih.gov/18229711/>.

115. **Cornelia Caragea**, Jivko Sinapov, Drena Dobbs, and Vasant Honavar. (2007). “Assessing the Performance of Macromolecular Sequence Classifiers.” In: *Proceedings of IEEE 7th International Conference on BioInformatics and BioEngineering (BIBE 2007)*, Boston, MA, USA. (pp. 320-326). doi:10.1109/BIBE.2007.4375583.

Workshop Papers:

1. Krutarth Patel, **Cornelia Caragea**, and Sujatha Das Gollapalli. (2020). “On the Use of Web Search to Improve Scientific Collections.” In: *Proceedings of the 1st Workshop on Scholarly Document Processing (SDP 2020)*, held in conjunction with the 2020 Conference on Empirical Methods in Natural Language Processing, Online. (pp. 174–183). doi:10.18653/v1/2020.sdp-1.20.
2. Ashwini Tonge and **Cornelia Caragea**. (2018). “On the Use of Deep Features for Online Image Sharing.” In: *Proceedings of the 9th International Workshop on Modeling Social Media, co-located with the ACM WWW 2018 conference (MSM 2018)*, Lyon, France.
3. **Cornelia Caragea**. (2016). “Identifying Descriptive Keyphrases from Scholarly Big Data.” In: *The Workshop “Artificial Intelligence for Data Science,” Co-located with the Neural Information Processing Systems Conference (AI4DataSci 2016)*, Barcelona, Spain.
4. Corina Florescu and **Cornelia Caragea**. (2016). “An Unsupervised Algorithm for Keyphrase Extraction.” In: *The Workshop for Women in Machine Learning, Co-located with the Neural Information Processing Systems Conference (WiML 2016)*, Barcelona, Spain.
5. Ashwini Tonge and **Cornelia Caragea**. (2016). “DeepPrivate: Accurate Prediction of Images’ Privacy on the Web.” In: *The Workshop for Women in Machine Learning, Co-located with the Neural Information Processing Systems Conference (WiML 2016)*, Barcelona, Spain.
6. **Cornelia Caragea**, Jian Wu, Kyle Williams, Sujatha Das G., Madian Khabsa, Pradeep Teregowda, and C. Lee Giles. (2014). “Automatic Identification of Research Articles from Crawled Documents.” In: *Proceedings of the Workshop: Web-Scale Classification: Classifying Big Data from the Web*, co-located with WSDM (**WSC 2014**), New York City, USA.
7. Amanda Karavolia, Dimitrios Vogiatzis, Brandon Truong, Andrea H. Tapia, **Cornelia Caragea**, Anna Squicciarini and Georgios Paliouras. (2014). “Tweet Type, Location and Popularity: Case Study Hurricane Sandy.” In: *The 2014 KDD Workshop on Learning about Emergencies from Social Information*, co-located with Knowledge Discovery and Data Mining 2014 (**KDD-LESI 2014**), New York City, USA.
8. Nir Ofek, **Cornelia Caragea**, Lior Rokach, Prakhar Biyani, Prasenjit Mitra, John Yen, Kenneth Portier, Greta Greer. (2013). “Improving Sentiment Analysis in an Online Cancer Survivor Community Using Dynamic Sentiment Lexicon.” In: *The 1st International Workshop on Public Health in the Digital Age: Social Media, Crowdsourcing and Participatory Systems (PHDA 2013)*, held in conjunction with the 22nd International World Wide Web Conference (WWW 2013), Rio de Janeiro, Brazil.
9. Sumit Bhatia, **Cornelia Caragea**, Hung-Hsuan Chen, Jian Wu, Pucktada Treeratpituk, Zhaohui Wu, Madian Khabsa, Prasenjit Mitra, and C. Lee Giles. (2012). “Specialized Research Datasets in the CiteSeerx Digital Library.” In: *The 1st Workshop on Mining Scientific Publications, Co-located with the ACM/IEEE Joint Conference on Digital Libraries (JCDL WS 2012)*, Washington, DC, USA.
10. **Cornelia Caragea**, Hyun-Woo Kim, Prasenjit Mitra, and John Yen. (2010). “Classifying Text Messages for Emergency Response.” In: *Machine Learning for Social Computing Workshop, Co-located with the Neural Information Processing Systems Conference (MLSC 2010)*, Whistler, Canada.

11. Rasna Walia, **Cornelia Caragea**, Drena Dobbs, and Vasant Honavar. (2010). “Improving the State of the Art in Machine Learning Methods for Protein-RNA Interface Prediction.” In: *The Workshop for Women in Machine Learning, Co-located with the Neural Information Processing Systems Conference (WiML 2010)*, Vancouver, BC, Canada.
12. **Cornelia Caragea**, Adrian Silvescu, Doina Caragea, and Vasant Honavar. (2009). “Abstraction Augmented Markov Models.” In: *Machine Learning in Computational Biology Workshop, Co-located with the Neural Information Processing Systems Conference (MLCB 2009)*, Whistler, Canada.
13. **Cornelia Caragea**, Adrian Silvescu, and Vasant Honavar. (2009). “Designing Abstract Features for Sequence Classification Tasks.” In: *The Workshop for Women in Machine Learning, Co-located with the Neural Information Processing Systems Conference (WiML 2009)*, Vancouver, BC, Canada.
14. Fadi Towfic, David Gemperline, **Cornelia Caragea**, Feihong Wu, Drena Dobbs, and Vasant Honavar. (2007). “Structural Characterization of RNA-Binding Sites of Proteins: Preliminary Results.”, In: *Computational Structural Bioinformatics Workshop, IEEE International Conference on Bioinformatics and Biomedicine (BIBM WS 2007)*, Silicon Valley, CA, USA.

Book Chapters, Tutorials, and other Technical Reports:

1. Alexander Tekle, Chau Pham, **Cornelia Caragea**, Junyi Jessy Li: “When a crisis strikes: Emotion analysis and detection during COVID-19.” CoRR abs/2107.11020 (2021).
2. Paul Landes, Barbara Di Eugenio, **Cornelia Caragea**: “DeepZensols: Deep Natural Language Processing Framework.” CoRR abs/2109.03383 (2021).
3. Ana Sabina Uban, **Cornelia Caragea**: “Generating Summaries for Scientific Paper Review.” CoRR abs/2109.14059 (2021).
4. **Cornelia Caragea** and Vasant Honavar. (2018). “Machine Learning in Computational Biology.” In: *Encyclopedia of Database Systems*. (pp. 2157-2162). doi:10.1007/978-1-4614-8265-9_636.
5. **Cornelia Caragea** and Vasant Honavar. (2017). “Machine Learning in Computational Biology.” In: *Encyclopedia of Database Systems*. (pp. 1-6). doi:10.1007/978-1-4899-7993-3_636-2.
6. Sujatha Das Gollapalli, Krutarth Patel, and **Cornelia Caragea**. (2016). “A Search/Crawl Framework for Automatically Acquiring Scientific Documents.” *CoRR*.
7. Ashwini Tonge and **Cornelia Caragea**. (2015). “Privacy Prediction of Images Shared on Social Media Sites Using Deep Features.” *CoRR*.
8. Sujatha Das Gollapalli, **Cornelia Caragea**, Xiaoli Li, and C. Lee Giles. (2014). Tutorial on “Document Analysis and Retrieval Tasks in Scientific Digital Libraries.” In: *Proceedings of the 8th Russian Summer School in Information Retrieval (RuSSIR 2014)*, Nizhny Novgorod, Russia.
9. Prakhar Biyani, **Cornelia Caragea**, Narayan Bhamidipati. (2014). “Entity-Specific Sentiment Classification of Yahoo News Comments.” *CoRR*.
10. Shibamouli Lahiri, Sagnik Ray Choudhury, and **Cornelia Caragea**. (2014). “Keyword and Keyphrase Extraction Using Centrality Measures on Collocation Networks.” *CoRR*.
11. **Cornelia Caragea** and Vasant Honavar. (2009). “Machine Learning in Computational Biology.” In: *Encyclopedia of Database Systems*. (pp. 1663-1667). doi:10.1007/978-0-387-39940-9_636.

Abstracts and Poster Programs:

1. Edoardo Savini and **Cornelia Caragea**. (2020). “A Multi-Task Learning Approach to Sarcasm Detection.” In: *Proceedings of the 34th American Association for Artificial Intelligence (AAAI 2020)*, Student Abstract and Poster Program, New York, USA. (pp. 13907-13908). doi:10.1609/aaai.v34i10.7226. <https://ojs.aaai.org//index.php/AAAI/article/view/7226>.
2. Gabriele Galfre and **Cornelia Caragea**. (2020). “Exploring Abstract Concepts for Image Privacy Prediction in Social Networks.” In: *Proceedings of the 34th American Association for Artificial Intelligence (AAAI 2020)*, Student Abstract and Poster Program, New York, USA. (pp. 13785-13786). doi:10.1609/aaai.v34i10.7164. <https://ojs.aaai.org//index.php/AAAI/article/view/7164>.
3. **Cornelia Caragea** and C. Lee Giles. (2019). “Keyphrase Extraction from Scholarly Documents for Data Discovery and Reuse.” In: *The 2019 Artificial Intelligence for Data Discovery and Reuse*, an NSF-supported conference @ Carnegie Mellon University, in-cooperation with ACM (**AIDR 2019**), Pittsburgh, Pennsylvania, USA.
4. **Cornelia Caragea** and Mark Phillips. (2019). “Building Specialized Collections from Web Archiving.” In: *The 2019 Artificial Intelligence for Data Discovery and Reuse*, an NSF-supported conference @ Carnegie Mellon University, in-cooperation with ACM (**AIDR 2019**), Pittsburgh, Pennsylvania, USA.
5. Mark Phillips, **Cornelia Caragea**, Krutarth Patel, and Nathan Fox. (2019). “Leveraging Machine Learning to Extract Content-Rich Publications from Web Archives.” In: *The 2019 International Internet Preservation Consortium (IIPC) Web Archiving Conference (WAC 2019)*, Zagreb, Croatia.
6. Mark Phillips, **Cornelia Caragea**, Krutarth Patel, and Nathan Fox. (2019). “Increasing Access to Content-Rich Publications from Web Archives with Machine Learning Models.” In: *The 2019 Texas Conference on Digital Libraries (TCDL 2019)*, Austin, Texas, USA.
7. Ashwini Tonge and **Cornelia Caragea**. (2019). “Dynamically Fusing Deep Multimodalities for Image Sharing.” In: *Proceedings of the 33rd American Association for Artificial Intelligence (AAAI 2019)*, Student Abstract, Honolulu, Hawaii, USA. (pp. 10057-10058). doi:10.1609/aaai.v33i01.330110057. <https://ojs.aaai.org/index.php/AAAI/article/view/5165>.
8. Ashwini Tonge, **Cornelia Caragea**, and Anna Squicciarini. (2018). “Uncovering Scene Context for Predicting Privacy of Online Shared Images.” In: *Proceedings of the 32nd American Association for Artificial Intelligence (AAAI 2018)*, Student Abstract and Poster Program, New Orleans, Louisiana, USA. (pp. 8167-8168). <https://ojs.aaai.org/index.php/AAAI/article/view/12180>.
9. Hamed Khanpour, **Cornelia Caragea**, and Prakhar Biyani. (2018). “Identifying Emotional Support in Online Health Communities.” In: *Proceedings of the 32nd American Association for Artificial Intelligence (AAAI 2018)*, Student Abstract and Poster Program, New Orleans, Louisiana, USA. (pp. 8099-8100). <https://ojs.aaai.org/index.php/AAAI/article/view/12170>.
10. Ashwini Tonge, **Cornelia Caragea**, and Anna Squicciarini. (2018). “Uncovering Scene Context for Predicting Images Privacy.” In: *The Grace Hopper Celebration of Women in Computing (GHC 2018)*, Houston, TX, USA.
11. Corina Florescu and **Cornelia Caragea**. (2017). “A Position-Biased PageRank Algorithm for Keyphrase Extraction.” In: *Proceedings of the 31st American Association for Artificial Intelligence (AAAI 2017)*, Student Abstract and Poster Program (SA-17), San Francisco, California, USA. (pp. 4923-4924). <https://ojs.aaai.org/index.php/AAAI/article/view/11082>.
12. Ashwini Tonge and **Cornelia Caragea**. (2016). “DeepPrivate: Accurate Prediction of Images’ Privacy on the Web.” In: *The Grace Hopper Celebration of Women in Computing (GHC 2016)*, Houston, TX, 2016.
13. Ashwini Tonge and **Cornelia Caragea**. (2016). “Image Privacy Prediction Using Deep Features.” In: *Proceedings of the 30th American Association for Artificial Intelligence (AAAI 2016)*, Student Abstract and Poster Program (SA-16), Phoenix, Arizona, USA, 2016. *Selected as a *finalist* in SA-16*. (pp. 4266-4267) <https://ojs.aaai.org/index.php/AAAI/article/view/9942>.

14. Florin Bulgarov and **Cornelia Caragea**. (2015). “A Comparison of Supervised Keyphrase Extraction Models.” In: *Proceedings of the 24th International Conference on World Wide Web (WWW 2015)*, Poster Program, Florence, Italy. (pp. 13–14). doi:10.1145/2740908.2742776.
15. Nir Ofek, Lior Rokach, **Cornelia Caragea**, and John Yen. (2015). “The Importance of Pronouns to Sentiment Analysis: Online Cancer Survivor Network Case Study.” In: *Proceedings of the 24th International Conference on World Wide Web (WWW 2015)*, Poster Program, Florence, Italy. (pp. 83–84). doi:10.1145/2740908.2742781.
16. Brandon Truong, Estella Xin, and **Cornelia Caragea**. (2015). “Extracting Valuable Information from Twitter During Natural Disasters.” Poster presented at the Fort Worth Regional Science and Engineering Fair, 2015 and at the Texas Science and Engineering Fair in San Antonio, TX, 2015. **Won the 3rd place in Computer Science at the Fort Worth Regional Science and Engineering Fair.**
17. Brenda Reyes Ayala and **Cornelia Caragea**. (2014). “Towards Building a Collection of Web Archiving Research Articles.” In: *International Internet Preservation Consortium (IIPC 2014)*, Paris, France.
18. Fred Oswald, **Cornelia Caragea**, Kim Nimon. (2014). “Technology-enhanced meta-analysis across scientific disciplines.” In: *The 29th Annual Conference of the Society for Industrial and Organizational Psychology (SIOP 2014)*, Hawaii.
19. Carmen Boje and **Cornelia Caragea**. (2013). “Multivariate Statistical Analysis to Assess Student Performance.” In: *New Mexico Higher Education Assessment and Retention Conference (NMHEAR 2013)*, Poster Program, Albuquerque, NM, USA.
20. **Cornelia Caragea**, Adrian Silvescu, Prasenjit Mitra, and C. Lee Giles. (2013). “Can’t see the forest for the trees? A citation recommendation system.” In: *The iConference*, Poster Program, Dallas Fort Worth, TX, USA.
21. Kim Nimon, **Cornelia Caragea**, Fred Oswald. (2013). “Meta-Scraping: Proposing Techniques to Support Meta-analyses.” In: *The iConference*, Poster Program, Dallas Fort Worth, TX, USA.
22. Adrian Silvescu, **Cornelia Caragea**, Vasant Honavar. (2009). “Combining Abstraction and Super-Structuring on Macromolecular Sequence Classification.” In: *The 13th International Conference on Research in Computational Molecular Biology (RECOMB 2009)*, Poster Program, Tucson, AZ, USA.
23. Adrian Silvescu, **Cornelia Caragea**, Vasant Honavar. (2009). “Predicting Protein Subcellular Localization Using Abstract Sequential Features.” In: *The 17th Annual International Conference on Intelligent Systems For Molecular Biology (ISMB 2009)*, Poster Program, Stockholm, Sweden.
24. **Cornelia Caragea**, Michael Terribilini, Jivko Sinapov, Jae-Hyung Lee, Fadi Towfic, Drena Dobbs, and Vasant Honavar. (2008). “Predicting Protein-RNA Binding Sites Using Structural Information.” In: *The 16th Annual International Conference On Intelligent Systems For Molecular Biology (ISMB 2008)*, Poster Program, Toronto, Canada.
25. Jivko Sinapov, **Cornelia Caragea**, Drena Dobbs, and Vasant Honavar. (2008). “Using Global Sequence Similarity Improves Biological Site-Specific Classifiers.” In: *The 16th Annual International Conference On Intelligent Systems For Molecular Biology (ISMB 2008)*, Poster Program, Toronto, Canada.
26. Michael Terribilini, **Cornelia Caragea**, Deepak Reyon, Jeffrey Sander, Jae-Hyung Lee, Robert Jernigan, Vasant Honavar, and Drena Dobbs. (2008). “Comparing Sequence and Structure-based Classifiers for Predicting RNA Binding Sites in Specific Families of RNA Binding Proteins.” In: *The 16th Annual International Conference On Intelligent Systems For Molecular Biology (ISMB 2008)*, Poster Program, Toronto, Canada.
27. Deepak Reyon, Ben Lewis, Jae-Hyung Lee, Colin Gleeson, Michael Hamilton, **Cornelia Caragea**, Fadi Towfic, Michael Terribilini, Vasant Honavar, and Drena Dobbs. (2007). “Combining Structural Modeling, Evolutionary Information, and Machine Learning to Improve Prediction of Nucleic Acid Binding Sites in Telomerase.” In: *The 5th Annual Rocky Mountain Bioinformatics Conference (Rocky 2007)*, Poster Program, Aspen/Snowmass, CO.

28. Michael Terribilini, Jeffrey Sander, **Cornelia Caragea**, Jae-Hyung Lee, Robert Jernigan, Vasant Honavar and Drena Dobbs. (2007). “Comparing Sequence vs. Structure-based Predictions of RNA Binding Sites in Proteins.” In: *Midwest Symposium on Computational Biology and Bioinformatics (MSCBB 2007)*, Poster Program, Chicago, IL.
29. Fadi Towfic, David Gemperline, **Cornelia Caragea**, Feihong Wu, Drena Dobbs, and Vasant Honavar. (2007). “Prediction of RNA-Protein Interfaces Using Structural Features.” In: *Midwest Symposium on Computational Biology and Bioinformatics (MSCBB 2007)*, Poster Program, Chicago, IL.
30. Colin Gleeson, Michael Hamilton, Jae-Hyung Lee, **Cornelia Caragea**, Vasant Honavar, and Drena Dobbs. (2007). “Generating Models as a Platform for Comparing Functional and Structural Elements of Telomerase.” In: *The 16th Annual Growth Factor and Signal Transduction Conferences (GFST 2007)*, Ames, IA.
31. **Cornelia Caragea**, Jivko Sinapov, Michael Terribilini, Drena Dobbs, and Vasant Honavar. (2007). “Assessing the Performance of Macromolecular Sequence Classifiers.” In: *The 15th Annual International Conference On Intelligent Systems For Molecular Biology and the 6th European Conference on Computational Biology (ISMB/ECCB 2007)*, Poster Program, Vienna, Austria.
32. Doina Caragea, **Cornelia Caragea**, Jie Bao, and Vasant Honavar. (2007). “Ontology-Based Information Integration Using INDUS: Intelligent Data Understanding System.” In: *The Emerging Technologies Conference (ETC 2007)*, Ames, IA.
33. **Cornelia Caragea**, Jivko Sinapov, Adrian Silvescu, Drena Dobbs, and Vasant Honavar. (2007). “Glycosylation Site Prediction using Machine Learning Approaches.” In: *The 11th International Conference on Research in Computational Molecular Biology (RECOMB 2007)*, Poster Program, Oakland, CA.
34. **Cornelia Caragea**, Jivko Sinapov, Adrian Silvescu, Drena Dobbs, and Vasant Honavar. (2006). “Learning Classifiers to Predict Glycosylation Sites in Proteins.” In: *The 4th Annual Rocky Mountain Bioinformatics Conference (Rocky 2006)*, Poster Program, Aspen/Snowmass, CO.
35. Yasser El-Manzalawy, **Cornelia Caragea**, Drena Dobbs, Vasant Honavar. (2006). “On the Quality of Motifs for Protein Phosphorylation Site Prediction.” In: *The 14th Annual International Conference On Intelligent Systems For Molecular Biology (ISMB 2006)*, Poster Program, Fortaleza, Brazil.
36. Yasser El-Manzalawy, **Cornelia Caragea**, Drena Dobbs, Vasant Honavar. (2006). “Machine Learning versus Profile-Based Methods for Protein Phosphorylation Site Prediction.” In: *The 6th Annual Joint Bioinformatics Symposium*, Student Abstract and Poster Program, Ames, IA.
37. **Cornelia Caragea**, Doina Caragea and Vasant Honavar. (2005). “Learning Support Vector Machine Classifiers from Distributed Data Sources.” In: *The 20th National Conference on Artificial Intelligence (AAAI 2005)*, Student Abstract and Poster Program, Pittsburgh, PA. (pp. 1602-1603). <https://www.aaai.org/Papers/AAAI/2005/SA05-010.pdf>.

Invited Talks and Presentations

1. “Deep Learning for Disaster Management and Response.” *Invited Talk* at the 2019 Amazon Web Services Re:Invent, December 4, 2019.
2. “The Myth of Double-Blind Review Revisited: ACL vs. EMNLP.” *Presentation* at the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019), Hong Kong, November 2, 2019.
3. “Multi-Task Stance Detection with Sentiment and Stance Lexicons.” *Poster Presentation* at the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019), Hong Kong, November 2, 2019.
4. “Deep Learning for Disaster Management and Response.” *Invited Talk* at the 2019 Amazon Web Services Public Sector Summit, June 12, 2019.

5. “Keyphrase Extraction from Scholarly Documents for Data Discovery and Reuse.” *Presentation* at The 2019 Artificial Intelligence for Data Discovery and Reuse, an NSF-supported conference @ Carnegie Mellon University, in-cooperation with ACM (AIDR 2019), Pittsburgh, Pennsylvania, USA, May 14, 2019.
6. “Building Specialized Collections from Web Archiving.” *Presentation* at The 2019 Artificial Intelligence for Data Discovery and Reuse, an NSF-supported conference @ Carnegie Mellon University, in-cooperation with ACM (AIDR 2019), Pittsburgh, Pennsylvania, USA, May 14, 2019.
7. “Dynamic Deep Multi-modal Fusion for Image Privacy Prediction.” *Invited Talk* at the 2019 Greater Chicago Systems Research Workshop (GCASR ’19), Chicago, IL, USA, May 6, 2019.
8. “Keyphrase Extraction in Scholarly Document Networks.” *Presentation* in the Library @ UIC, November 29, 2018.
9. “Exploring Optimism and Pessimism in Twitter Using Deep Learning.” *Presentation* at the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018), Brussels, Belgium, November 2, 2018.
10. “Fine-Grained Emotion Detection in Health-Related Online Posts.” *Poster Presentation* at the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018), Brussels, Belgium, November 2, 2018.
11. “Removing the Clutter of Uninteresting Words in Text. RAAI to the Rescue.” *Invited Talk* at the 2nd Conference on Recent Advances in Artificial Intelligence (RAAI 2018) organized by the University of Bucharest, June 25, 2018.
12. “Deep Neural Networks versus Naive Bayes Classifiers for Identifying Informative Tweets during Disasters.” *Presentation* at the 15th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2018), Rochester NY, USA, May 22, 2018.
13. “Designing an Unsupervised Approach to Keyphrase Extraction Using Ideas from Supervised Approaches.” *Invited Talk* at Virginia Tech, Washington, DC, April 13, 2018.
14. “Improving Disaster Response Through the Eyes of Social Media.” *Invited Talk* at The University of Oklahoma, Norman, OK, April 9, 2018.
15. “Learning to Extract Descriptive Keyphrases from Scholarly Big Data.” *Invited Speaker* at the *IEEE MetroCon*, Arlington Convention Center, TX, October 26, 2017.
16. “Learning to Extract Descriptive Keyphrases from Scholarly Big Data.” *Invited Talk* at Southern Methodist University (SMU), September 29, 2017.
17. “PositionRank: An Unsupervised Approach to Keyphrase Extraction from Scholarly Documents.” *Presentation* at the 2017 Annual Meeting of the Association for Computational Linguistics (ACL 2017), Vancouver, Canada, August 1, 2017.
18. “Improving Disaster Response Through the Eyes of Social Media.” *Invited Talk* at the 10th Conference on PErvasive Technologies Related to Assistive Environments (PETRA 2017), June 22, 2017.
19. “Learning to Extract Descriptive Keyphrases from Scholarly Big Data.” *Invited Talk* at the 1st Conference on Recent Advances in Artificial Intelligence (RAAI 2017) organized by the University of Bucharest, June 19, 2017.
20. “Image Privacy Prediction Using Deep Features.” *Presentation* at the Researcher Luncheon and Poster Forum at the Fort Worth Museum of Science and History in the Research and Learning Center, Fort Worth, Texas, May 20, 2017.
21. “Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?” *Invited Talk* at Temple University, Philadelphia, Pennsylvania, April 26, 2017.

22. "Keyphrase Extraction in Document Networks." *Invited Talk* at the University of Texas at Dallas, Dallas, Texas, April 14, 2017.
23. "Image Privacy Prediction Using Deep Features." *Presentation* at the 2017 Secure and Trustworthy Cyberspace Principal Investigators' Meeting (SaTC PI Meeting '17), 2017.
24. "Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?" *Presentation* at the University of the Andes, Bogotá, Colombia, June 15, 2016.
25. "Geo-Mapped Sentiment Analysis in Disaster Events." *Presentation* at the Researcher Luncheon and Poster Forum at the Fort Worth Museum of Science and History in the Research and Learning Center, Fort Worth, Texas, May 21, 2016.
26. "Identifying Informative Messages in Disasters using Convolutional Neural Networks." *Presentation* at the 13th International Conference on Information Systems for Crisis Response and Management, Rio de Janeiro, Brazil, 2016.
27. "Document Type Classification in Online Digital Libraries." *Presentation* at the 28th Annual Conference on Innovative Applications of Artificial Intelligence, Phoenix, AZ, USA, 2016.
28. "Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?" *Talk* at the University of Texas at Arlington, Arlington, Texas, March 1, 2016.
29. "Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?" *Talk* at the University of Oregon, Eugene, Oregon, February 23, 2016.
30. "Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?" *Invited talk* (Speaker) at the *IEEE MetroCon*, Arlington Convention Center, TX, October 22, 2015.
31. "Keyphrase Extraction for Scholarly Big Data." *Invited talk* at the *Big Scholarly Data: Birds of a Feather Session*, co-located with *Microsoft Research Faculty Summit*, Redmond, WA, July 10, 2015.
32. "Extracting Keyphrases from Research Papers using Citation Networks." *Presentation* at the Researcher Luncheon and Poster Forum at the Fort Worth Museum of Science and History in the Research and Learning Center, Fort Worth, Texas, March 7, 2015.
33. "Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?" *Invited talk* at the University of Texas at Austin, Austin, Texas, February 13, 2015.
34. "Keyphrase Extraction in Citation Networks: How do Citation Contexts Help?" *Invited talk* at the University of Michigan, Ann Arbor, Michigan, November 21, 2014.
35. "Big Data and its Implications for Research Methodology and Funding." *Invited talk* at the University of North Texas at "TARDIS - The Advances in Research DesIgns Symposium." Held at the Gateway Center at UNT, Denton, TX, on November 7, 2014.
36. "CiteSeerX: A Scholarly Big Dataset." *Presentation* at the 36th European Conference on Information Retrieval, Amsterdam, Netherlands, 2014.
37. "Abstraction-Based Probabilistic Models for Sequence Classification." *Invited talk* at the International Research-Centered Summer School in Cognitive Systems and Interactive Robotics, Social Media, and Digital Preservation, organized by the Software & Knowledge Engineering Lab at the Institute of Informatics & Telecommunications of NCSR "Demokritos", in Athens, Greece, July 29, 2013.
38. "Combining Hashing and Abstraction in Sparse High Dimensional Feature Spaces." *Invited talk* at the Machine Learning and Bioinformatics Laboratory at Kansas State University, Manhattan, Kansas, September 27, 2013.
39. "Abstraction-Based Probabilistic Models for Sequence Classification." *Invited talk* at the University of the Andes, Bogotá, Colombia, June 10, 2013.

40. “Abstraction-Based Probabilistic Models for Sequence Classification.” *Talk* in the College of Engineering at the University of North Texas, Denton, Texas, May 10, 2013 (at the CSE Advisory Council meeting).
41. “Abstraction-Based Probabilistic Models for Sequence Classification.” *Talk* in Computer Science and Engineering at the University of North Texas, Denton, Texas, December 7, 2012 (at the Industrial Advisory Board meeting).
42. “Abstraction-Based Probabilistic Models for Sequence Classification.” *Talk* at the Pennsylvania State University, University Park, Pennsylvania, June, 2010.

Awards and Honors

2020-21 COE Faculty Research Award 2021
Award received for Excellent Research in the College of Engineering at University of Illinois at Chicago.

IAAI-21 Deployed Application Award 2021
Award received for the paper “Author Homepage Discovery in CiteSeerX.” by Krutarth Patel, Cornelia Caragea, Doina Caragea, and C. Lee Giles, at The 33rd Annual Conference on Innovative Applications of Artificial Intelligence (IAAI 2021).

Best Student Paper Award 2020
Award received for the paper “Identifying Documents In-Scope of a Collection from Web Archives.” by Krutarth Patel, Cornelia Caragea, Mark Phillips, and Nathaniel Fox, at The ACM/IEEE Joint Conference on Digital Libraries (JCDL 2020).

Best Insight Paper Award 2019
Award received for the paper “Identifying Disaster Damage Images Using a Domain Adaptation Approach.” by Xukun Li, Doina Caragea, Cornelia Caragea, Muhammad Imran, and Ferda Offi, at the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019).

The Lloyd T. Smith Creativity in Engineering Endowed Chair 2017 - 2018
Received The Lloyd T. Smith Creativity in Engineering Chair - Endowed Chair at Kansas State University.

Michelle Munson-Serban Simu Keystone Research Faculty Scholar Award 2018
Received Munson-Simu Keystone Research Scholar Award at Kansas State University.

Best paper nomination in the 8th International Conference on Knowledge Capture 2015
Best paper nomination for the paper “PDFMEF: A Multi-Entity Knowledge Extraction Framework for Scholarly Documents and Semantic Search.” by Jian Wu, Jason Killian, Huaiyu Yang, Kyle Williams, Sagnik Ray Choudhury, Suppawong Tuarob, Cornelia Caragea, C. Lee Giles at The 8th International Conference on Knowledge Capture (K-Cap 2015).

3rd place in Computer Science at the Fort Worth Regional Science and Engineering Fair 2015
Award received for research work done with two TAMS (Texas Academy of Math and Science) high-school students, Brandon Truong and Estella Xin, for the paper “Extracting Valuable Information from Twitter During Natural Disasters.” The two students, Brandon and Estella were finalists at the Texas Science and Engineering Fair in San Antonio, TX.

Most Innovative Application of AI Award 2014
Award received for the paper “CiteSeerX: AI in a Digital Library Search Engine.” by Jian Wu, Kyle Williams, Hung-Hsuan Chen, Madian Khabza, Cornelia Caragea, Alexander Ororbia, Douglas Jordan, and C. Lee Giles at The Twenty sixth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI '14).

John Vincent Atanasoff Award

2005

Annual award offered by the Computer Science Department to outstanding graduate students.

PACE Award

2003 - 2004

Premium for Academic Excellence from Iowa State University.

Leonardo Melandri Program Award

December 10-15, 2006

DASI '06: Bertinoro PhD School on Data and Service Integration, Bertinoro, Italy.

RECOMB Student Scholarship

May 18-21, 2009

Awarded by the 13th Annual International Conference on Research in Computational Molecular Biology, Tucson, AZ, USA.

ISMB Student Scholarship

July 19-23, 2008

Awarded by the 16th Annual International Conference on Intelligent Systems for Molecular Biology, Toronto, Canada.

AAAI Student Scholarship

July 9-13, 2005

Awarded by the 20th National Conference on Artificial Intelligence, Pittsburgh, PA, USA.

Teaching Experience

Computer Science, University of Illinois at Chicago, USA

- CS 533: “Deep Learning for Natural Language Processing” - graduate course *Spring 2022*
- CS 494: “Information Retrieval and Web Search” - undergraduate + graduate course *Fall 2021*
- CS 594: “Deep Learning for Natural Language Processing” - graduate course *Fall 2020, Fall 2019*
- CS 582: “Information Retrieval and Web Search” - graduate course *Spring 2021, Spring 2020, Spring 2019, Fall 2018*

Computer Science, Kansas State University, USA

- CS 833: “Information Retrieval and Web Search” - graduate course *Fall 2017*

Computer Science and Engineering, University of North Texas, USA

- CSE 5215: “Machine Learning” - graduate and undergraduate course *Fall 2016, 2015, 2014, 2013, 2012*
- CSE 5200: “Information Retrieval and Web Search” - graduate and undergraduate course *Spring 2017, 2016, 2015, 2014*
- CSE 6933: “Social Network Analysis” - advanced graduate course *Spring 2013*

The 8th Russian Summer School in Information Retrieval, Nizhny Novgorod, Russia.

- RuSSIR 2014: “Document Analysis and Retrieval in Scientific Digital Libraries: Case studies in applying Machine Learning for Information Retrieval” - graduate course *Summer 2014*

Computer Science and Engineering, University of the Andés, Bogota, Colombia

- “Knowledge Discovery in Social and Information Networks” - graduate course *Summer 2018, 2016, 2013*

College of Information Sciences and Technology, Pennsylvania State University, USA

- IST 558: “Data Mining 2” - graduate course *Spring 2011*

Teaching Assistant, Department of Computer Science, Iowa State University

- ComS 104: “Introduction to Computers” - undergraduate course *Fall 2003*
- ComS 104: “Introduction to Computers” - undergraduate course *Spring 2004*
- ComS 472/572: “Principles of Artificial Intelligence” - graduate course *Fall 2004*
- ComS 228: “Introduction to Data Structures” - undergraduate course *Spring 2005*
- ComS 330: “Discrete Computational Structures” - undergraduate course *Fall 2005*
- ComS 311: “Design and Analysis of Algorithms” - undergraduate course *Spring 2006*

Teaching Assistant, Bioinformatics and Computational Biology Summer Institute for Undergraduate Students, Iowa State University

- Module F: Machine Learning Methods in Bioinformatics *Summer 2006*
- Module G: Computational Analysis of “Omics” data: Machine Learning Approaches *Summer 2007*
- Module G: Machine Learning Approaches and Systems Biology *Summer 2008*

AI Seminar, Department of Computer Science, Iowa State University

- I gave presentations, and helped organize the AI Seminar coordinated by Dr. Vasant Honavar *Spring 2004 - Fall 2006*

Guest Lectures

- I gave several lectures in graduate courses on “Machine Learning” (Prof. Vasant Honavar), “Artificial Intelligence” (Prof. Dimitris Margaritis), and in the undergraduate course on “Design and Analysis of Algorithms” (Prof. David Fernández-Baca).

Student Advising

Current PhD Students:

1. Jishnu Ray Chowdhury [PhD]
2. Yingjie Li [PhD]
3. Seo Yeon Park [PhD]
4. Mobashir Sadat [PhD]
5. Chenye Zhao [PhD]
6. Mahshid Hosseini [PhD]
7. Tiberiu Sosea [PhD]
8. Zahra Rahimi Afzal [PhD]
9. Krishna Kant Garg [PhD]
10. George Maratos [PhD]
11. Henry Peng Zou [PhD]

Past PhD Students:

1. Krutarth Patel [PhD]
PhD Dissertation: “Keyphrase Extraction and its Applications from Scholarly Documents.”
Dissertation Defense Date: March 29, 2021.
2. Ashwini Tonge [PhD]
PhD Dissertation: “Image Privacy Prediction using Deep Learning.”
Dissertation Defense Date: March 28, 2019.
3. Hamed Khanpour [PhD]
PhD Dissertation: “Computational Approaches for Analyzing Social Support in Online Health Communities.”
Dissertation Defense Date: March 22, 2018.
Now at Microsoft Research.
4. Kishore Neppalli [PhD]
PhD Dissertation: “Extracting Useful Information from Social Media during Disaster Events.”
Dissertation Defense Date: March 24, 2017.
Now Data Scientist at Kaiser Permanente.

Past MS Students with Thesis:

1. Tuhin Kundu [MS]
MS Thesis: “Neural Keyphrase Generation: Looking Inside the Decoder.”
Thesis Defense Date: March 10, 2021.
2. Piervincenzo Ventrella [MS]
MS Thesis: “Predictive Analysis on the Onset of Complete Renal Failure for Patients with Chronic Kidney Disease.”
Thesis Defense Date: December 9, 2020.
3. Matteo Marziali [MS]
MS Thesis: “Analysis of cognitive and keyword-based approaches for searching judicial documents.”
Thesis Defense Date: December 10, 2019.
4. Gabriele Galfre [MS]
MS Thesis: “Exploring Abstract Concepts for Images Privacy Classification in Social Media.”
Thesis Defense Date: September 18, 2019.
Now at Facebook.
5. Alessandro Rennola [MS]
MS Thesis: “Semi-Supervised Machine Learning and Deep Learning Models in Crisis-Related Informativeness Classification.”
Thesis Defense Date: August 1, 2019.
6. Edoardo Savini [MS]
MS Thesis: “An Exploration of Sarcasm Detection Using Deep Learning.”
Thesis Defense Date: August 1, 2019.
7. Virashree Patel [MS]
MS Thesis: “Topic Modeling on Disaster Datasets using Latent Dirichlet Allocation.”
Thesis Defense Date: November 25, 2018.

Past MS Students (primary advisor - implementation project or course work):

1. Teja Gollapudi [MS] - Spring 2021

2. Sankul Rawat [MS] - Spring 2019
3. Rahul Sathe [MS] - Spring 2019
4. Minal Kucheria [MS] - Spring 2019
5. Sean Martinelli [MS] - Spring 2019
6. Virashree Patel [MS] - Fall 2018
7. Milad Pejmanrad [MS] - Spring 2016
8. Ruichen Rong [MS] - Spring 2016
9. Preethi Mudumba [MS] - Fall 2015
10. Indish Cholleti [MS]
11. Keerthana Sadam [MS]
12. Shruthi Tirandas [MS]
13. Shailesh Acharya [MS]
14. Manoj Kumar Panchagnula [MS]
15. Aaditya Chauhan [MS]
16. Gopikrishna Reddy Puchakayala [MS]

Past Undergraduate Students:

1. Faraaz Baig, Undergraduate UIC student
2. Manasvi Narayanan, Undergraduate UIC student
3. Asvin Ragunathan, Undergraduate UIC student
4. Matt Ziminski, Undergraduate UIC student
5. Arpita Kumari, Undergraduate UIC student
6. Elijah Raul Rodriguez Beltran, Undergraduate UIC student
7. Jacob White, Undergraduate UNT student
8. Daniel Moore, Undergraduate UNT student
9. Clement Cole, Undergraduate UNT student
10. Murilo Cerqueira Medeiros, Undergraduate student from Brazil
11. Josh Talkington, Undergraduate UNT student
12. Daniel Stieben, Student from the Honors College from UNT

Past TAMS High-School Students:

1. Brandon Truong. Accepted as an undergraduate at Stanford University
2. Estella Xin. Accepted as an undergraduate at University of Texas at Austin

Visiting Students:

1. Lucas Sterckx, PhD student at Ghent University - iMinds, Belgium, Spring 2016
2. Namchi Do, DREU: Distributed Research Experiences for Undergraduates, Undergraduate Student from Trinity University, Summer 2015

Past Post-doctoral Research Associate:

1. Sujatha Das Gollapalli. Now at the Institute for Infocomm Research, A*STAR, Singapore.

Other Mentored Students

Graduate Students:

1. Brenda Reyes Ayala [LIS - UNT]
2. Alina Ciobanu [University of Bucharest, Romania]
3. Ana Uban [University of Bucharest, Romania]
4. Stefan Cobeli [University of Bucharest, Romania]
5. Nir Ofek [Ben-Gurion University of the Negev, Israel]
6. Prakhar Biyani [PSU]
7. Sujatha Das Gollapalli [PSU]
8. Saurabh Kataria [PSU]
9. Shibamouli Lahiri [UMich]
10. Raunaq Malhotra [PSU]
11. Tony Nguyen [PSU]
12. Rasna Walia [ISU]
13. Fadi Towfic [ISU]
14. Neeraj Koul [ISU]

Undergraduate Students:

1. Lalith Polepeddi [ISU - BCB Summer Institute]
2. Chong Hoon Kim [ISU - BCB Summer Institute]
3. Kevin Yang [ISU - Summer Internship]
4. David Gemperline [ISU - BCB Summer Institute]
5. Marjie Volk [ISU]

Student Committee

I have served or I am serving in the dissertation or thesis committee of the following students:

PhD Students (completed or in progress):

1. Jianguo Zhang [UIC]
“Towards Accurate and Smart Task-Oriented Dialogue Systems.”
2. Ye Liu [UIC]
“Knowledge-enhanced Neural Text Generation.”
3. Congying Xia [UIC]
“Natural Language Understanding for Conversational Agents.”
4. Shaika Chowdhury [UIC]
“Text Modeling and Mining for Healthcare using Deep Learning.”
5. Sergiu Nisioi [University of Bucharest]
“A Computational Exploration of Natural Language Varieties.”
6. Mark Phillips [University of North Texas]
“Exploring the use of metadata record graphs for metadata assessment.”
7. Rui Meng [University of Pittsburgh]
“Keyphrasification: summarizing text with keyphrases - using neural language generation methods.”
8. Shen Wang [UIC]
“Broad structural representation learning.”
9. Wei Xing [UIC]
“Adversarial Learning for Intractable Problems.”

10. Lei Shu [UIC]
“Deep generation techniques in task-oriented dialogue system.”
11. Mehrdad Alizadeh [UIC]
“Enhancing Visual Question Answering with Linguistic Information.”
12. Chenwei Zhang [UIC]
“Structured Knowledge Discovery from the Massive Text Corpus.”
13. Shuai Wang [UIC]
“Target-Oriented Content and Sentiment Analysis.”
14. Prakhar Biyani [PSU]
“Subjectivity analysis of online forums.”
15. Kathryn Masten-Cain [UNT]
“Toward a grounded theory of community networking.”
16. Jedsada Chartree [UNT]
“Monitoring Dengue outbreaks using online data.”
17. Saratchandra Indrakanti [UNT]
“Computational Methods for Vulnerability Analysis and Resource Allocation in Public Health Emergencies.”
18. Angel Bravo-Salgado [UNT]
“Modeling and Simulation of the Vector-Borne Dengue Disease and the Effects of Regional Variation of Temperature in the Disease Prevalence in a Homogenous and Heterogeneous Human Population.”
19. Umar Albalawi [UNT]
“New Frameworks for Secure Image Communication in the Internet of Things (IoT).”
20. Garima Bajwa [UNT]
“Sensing and Decoding Brain States for Predicting and Enhancing Human Behavior, Health, and Security.”
21. Shanti Thiyagaraja [UNT]
“Detection and Classification of Heart Sounds Using a Heart-Mobile Interface.”
22. Karen Mazidi [UNT]
“Infusing Automatic Question Generation with Natural Language Understanding.”
23. Alberto Castro-Hernandez [UNT]
“Content and Temporal Analysis of Communications to Predict Task Cohesion in Software Development Global Teams.”

WCP (Qualifying Exam) Students Committee:

1. Chenye Zhao (member)
WCP (CS) March 15, 2021
2. Tiberiu Sosea (member)
Knowledge Distillation in Neural Networks.
WCP (CS) February 15, 2021
3. Jishnu Ray Chowdhury (member)
Analyzing Transformers.
WCP (CS) October 26, 2020
4. Seo Yeon Park (member)
Understanding the relation between emotions and empathy: a review for empathy modeling in NLP.
WCP (CS) March 19, 2020
5. Mohammad Arvan (chair)
A review of regularization techniques for Recurrent Neural Networks.
WCP (CS) March 12, 2020
6. Tao Zhang (chair)
Zero-shot learning on text mining.
WCP (CS) March 10, 2020

7. Hoang Nguyen (chair)
Robust learning with scarce annotations.
WCP (CS) March 9, 2020
8. Mobashir Sadat (member)
A review and critique on neural abstractive summarization.
WCP (CS) March 3, 2020
9. Mahshid Alsadat Hosseini (member)
Emotion recognition and emotion generation in conversations.
WCP (CS) February 27, 2020
10. Kiavash Satvat (chair)
Applications of NLP in threat detection and automatic exploit generation.
WCP (CS) February 25, 2020
11. Zahra Fatemi (chair)
Relational Causal Inference.
WCP (CS) March 20, 2019
12. Yingjie Li (member)
A Review and Critique of Deep Learning for Sentiment Analysis.
WCP (CS) March 18, 2019

Master Students (completed):

1. Harsh Yadav [UIC]
Thesis: “Primal-Dual Graph Attention Networks for 3D Human Pose Estimation.”
2. Abhijeet Mohanty [UIC]
Thesis: “Generator of Research Output Units in Software Engineering.”
3. Samuele Battaglini [UIC]
Thesis: “A generalization of principal component analysis.”
4. Sachin G Biradar [UIC]
Thesis: “User-Centric Adversarial Perturbations to Protect Privacy in Social Networks.”
5. Akash Rathore [KSU]
Thesis: “Tools for Querying the National Household Travel Survey Data.”
6. Debarshi Saha [KSU]
Thesis: “A Database for Storing and Analyzing Tweets Posted During Disasters .”
7. Ashwini Tonge [UNT]
Thesis: “Object Recognition Using Scale-Invariant Chordigram.”
8. Samir Koppikar [UNT]
Thesis: “Privacy Preserving EEG-based Authentication Using Perceptual Hashing.”
9. Ramu Reddy Venumuddala [UNT]
Thesis: “A Distributed Framework towards Building an Open Data Architecture.”
10. Kok Chuan (KC) Koh [UNT]
Thesis: “Modeling alcohol Consumption using blog data.”
11. Rajitha Schellenberg [UNT]
Thesis: “Automated classification of emotions in song lyrics.”
12. David Jensen [UNT]
Thesis: “DDOS defense against botnets in the mobile cloud.”

Other Past MS Students (part of committee):

1. Akshay Vedpathak [MS] - Spring 2019

Professional Services

Workshop Organizer:

1. **Cornelia Caragea**¹, Madian Khabsa², Sujatha Das Gollapalli³, C. Lee Giles⁴, Alex Wade⁵. “The IJCAI-16 Workshop on Scholarly Big Data: AI Perspectives, Challenges and Ideas.” To be presented at: *The International Joint Conference on Artificial Intelligence 2016 (IJCAI 2016)*, July 9-11, 2016, New York, USA. ¹University of North Texas; ²Microsoft Research; ³Institute for Infocomm Research, A*STAR, Singapore; ⁴Pennsylvania State University; ⁵Microsoft Academic Search.
2. **Cornelia Caragea**¹, C. Lee Giles², Alex Wade³, Doina Caragea⁴, Vu Ha⁵, Madian Khabsa⁶, Irwin King⁷, Jie Tang⁸. “The AAAI-16 Workshop on Scholarly Big Data: AI Perspectives, Challenges and Ideas.” In: *The Association for the Advancement of Artificial Intelligence 2016 (AAAI 2016)*, 13 Feb. 2016, Phoenix, AZ, USA. ¹University of North Texas; ²Pennsylvania State University; ³Microsoft Academic Search; ⁴Kansas State University; ⁵Allen Institute for Artificial Intelligence; ⁶Microsoft Research; ⁷The Chinese University of Hong Kong, Shatin, NT, Hong Kong; ⁸Tsinghua University, Beijing, China.
3. Sujatha Das G.¹, **Cornelia Caragea**², C. Lee Giles³, Xiaoli Li¹. “The ACL-15 Workshop on Novel Computational Approaches to Keyphrase Extraction.” In: *The 53rd Annual Meeting of the Association for Computational Linguistics 2015 (ACL 2015)*, 30-31 July 2015, Beijing, China. ¹Institute for Infocomm Research, A*STAR, Singapore; ²University of North Texas; ³Pennsylvania State University.
4. **Cornelia Caragea**¹, C. Lee Giles², Narayan Bhamidipati³, Doina Caragea⁴, Sujatha Das G.⁵, Saurabh Kataria⁶, Huan Liu⁷, Feng Xia⁸. “Workshop on Scholarly Big Data: AI Perspectives, Challenges and Ideas.” In: *The Association for the Advancement of Artificial Intelligence 2015 (AAAI 2015)*, 25 Jan. 2015, Austin, TX, USA. ¹University of North Texas; ²Pennsylvania State University; ³Yahoo! Labs, USA; ⁴Kansas State University, USA; ⁵Institute for Infocomm Research, A*STAR, Singapore; ⁶Palo Alto Research Center, USA; ⁷Arizona State University, USA; ⁸Dalian University of Technology, China.
5. **Cornelia Caragea**¹, C. Lee Giles², Sujatha Das G.³, Ingemar J. Cox⁴, Christina Lioma⁴, Natasa Milic-Frayling⁵, Madian Khabsa², Daqing He⁶, and Yu-Ru Lin⁶. “Workshop on Scholarly Big Data: Challenges and Ideas.” In: *The 2014 IEEE International Conference on Big Data (IEEE BigData 2014)*, 27 October 2014, Washington DC, USA. ¹University of North Texas; ²Pennsylvania State University; ³Institute for Infocomm Research, A*STAR, Singapore; ⁴University of Copenhagen, Denmark; ⁵Microsoft Research Cambridge, UK; ⁶University of Pittsburgh, USA.
6. **Cornelia Caragea**¹, C. Lee Giles², Lior Rokach³, and Xiaozhong Liu⁴. “Computational Scientometrics: Theory and Applications.” Workshop presented in: *The 22nd International Conference on Information and Knowledge Management (CIKM 2013)*, 28 October 2013, San Francisco, USA. ¹University of North Texas; ²Pennsylvania State University; ³Ben-Gurion University of the Negev, Israel; ⁴Indiana University.
7. C. Lee Giles¹, **Cornelia Caragea**², Ingemar J. Cox³, Birger Larsen³, Christina Lioma³, Natasa Milic-Frayling⁴, Velte Torvik⁵. “Workshop on Scholarly Big Data: Challenges and Ideas.” Workshop presented in: *The 2013 IEEE International Conference on Big Data (IEEE BigData 2013)*, 6 October 2013, Santa Clara, USA. ¹Pennsylvania State University; ²University of North Texas; ³University of Copenhagen, Denmark; ⁴Microsoft Research Cambridge, UK; ⁵University of Illinois, Urbana-Champaign, USA.
8. **Cornelia Caragea**¹, C. Lee Giles², Velte Torvik³, and Lior Rokach⁴. “Computational Scientometrics: Theory and Applications.” Workshop presented in: **The iConference 2013**, 12 February 2013, Dallas Fort Worth, USA. ¹University of North Texas; ²Pennsylvania State University; ³University of Illinois at Urbana-Champaign; ⁴Ben-Gurion University of the Negev, Israel.

Grant Proposal Reviewer:

1. UIC - limited submissions (**UIC**): 2021.
2. Department of Energy (**DOE**): 2021.

3. National Science Foundation (**NSF** - 7 panels + 1 ad-hoc): 2021.
4. National Science Foundation (**NSF** - 5 panels + 1 ad-hoc): 2020.
5. US Department of Homeland Security (**DHS** - 1 ad-hoc): 2019.
6. UIC - limited submissions (**UIC**): 2019.
7. National Science Foundation (**NSF** - 3 panels): 2019.
8. National Institutes of Health (**NIH** - 1 panel): 2019.
9. UIC - DPI (first cycle) (**UIC-DPI**): 2019.
10. National Science Foundation (**NSF** - 4 panels): 2018.
11. National Institutes of Health (**NIH** - 1 panel): 2018.
12. National Science Foundation (**NSF** - 2 panels): 2017.
13. Ralph E. Powe Junior Faculty Award: 2017.
14. Swiss National Science Foundation (**SNSF**): 2016.
15. National Science Foundation (**NSF** - 4 panels): 2016.
16. Ralph E. Powe Junior Faculty Award: 2016.
17. Swiss National Science Foundation (**SNSF**): 2015.
18. Ralph E. Powe Junior Faculty Award: 2015.
19. National Science Foundation (**NSF** - 2 panels): 2014.
20. National Science Foundation (**NSF** - 2 panels): 2013.

Program Committee (PC) Member:

1. *ACL 2022*: The 60th Annual Meeting of the Association for Computational Linguistics
2. *IJCAI 2022*: The 31st International Joint Conference on Artificial Intelligence
3. *AAAI 2022*: The 36th AAAI Conference on Artificial Intelligence
4. *ECML 2021*: The 2021 European Conference on Machine Learning
5. *EMNLP 2021*: The 2021 Conference on Empirical Methods in Natural Language Processing
6. *ACL 2021*: The 59th Annual Meeting of the Association for Computational Linguistics
7. *NAACL 2021*: Annual Conference of the North American Chapter of the Association for Computational Linguistics
8. *DaSH-LA 2021*: Workshop on Data Science with Human-in-the-Loop: Language Advances, co-located with NAACL 2021
9. *W-NUT 2021*: The 6th Workshop on Noisy User-Generated Text (W-NUT) co-located with EMNLP 2021.
10. *EACL 2021*: The 16th Conference of the European Chapter of the Association for Computational Linguistics
11. *IJCAI 2021*: The 30th International Joint Conference on Artificial Intelligence
12. *AAAI 2021*: The 35th AAAI Conference on Artificial Intelligence
13. *WSDM 2021*: The 14th ACM International Conference on Web Search and Data Mining
14. *COLING 2020*: The 28th International Conference on Computational Linguistics
15. *ACL-IJCNLP 2020*: The 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing
16. *KDD 2020*: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining
17. *CIKM 2020*: The 29th ACM International Conference on Information and Knowledge Management
18. *EMNLP 2020*: The 2020 Conference on Empirical Methods in Natural Language Processing
19. *ECIR 2020*: The European Conference on Information Retrieval
20. *BigMM2020*: The 2020 IEEE International Conference on Multimedia Big Data
21. *ACL 2020*: The 58th Annual Meeting of the Association for Computational Linguistics
22. *AAAI 2020*: The 34th AAAI Conference on Artificial Intelligence

23. *ACL 2019*: The 57th Annual Meeting of the Association for Computational Linguistics
24. *EMNLP 2019*: 2019 Conference on Empirical Methods in Natural Language Processing
25. *NAACL 2019*: The Annual Conference of the North American Chapter of the Association for Computational Linguistics
26. *ISCRAM 2019*: Intl. Conference on Information Systems for Crisis Response and Management
27. *BIRNDL 2019*: 4th Joint Workshop on Bibliometric-enhanced Information Retrieval and Natural Language Processing for Digital Libraries
28. *ECIR 2019*: The European Conference on Information Retrieval
29. *AAAI 2019*: The 33rd AAAI Conference on Artificial Intelligence
30. *BigData 2018*: 2018 IEEE International Conference on Big Data
31. *NAACL 2018*: The Annual Conference of the North American Chapter of the Association for Computational Linguistics
32. *AAAI 2018*: The 32nd AAAI Conference on Artificial Intelligence
33. *ISCRAM 2018*: The 15th Int. Conf. on Information Systems for Crisis Response and Management
34. *LREC 2018*: 11th edition of the Language Resources and Evaluation Conference
35. *NLDB 2018*: The 23rd Int. Conference on Application of Natural Language to Information Systems
36. *AISec 2018*: AI and Security workshop
37. *BigData 2017*: 2017 IEEE International Conference on Big Data
38. *EMNLP 2017*: Conference on Empirical Methods in Natural Language Processing
39. *CoNLL 2017*: The SIGNLL Conference on Computational Natural Language Learning
40. *ACL 2017*: The 55th Annual Meeting of the Association for Computational Linguistics
41. *SIGIR 2017*: The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval
42. *WWW 2017*: The 26th International World Wide Web Conference
43. *AAAI 2017*: The 31st AAAI Conference on Artificial Intelligence
44. *SAC 2017*: Track on Recommender Systems in the 32nd ACM Symposium On Applied Computing
45. *ISCRAM 2017*: The 14th Int. Conf. on Information Systems for Crisis Response and Management
46. *NLDB 2017*: The 22nd Int. Conference on Application of Natural Language to Information Systems
47. *BIR 2017*: ECIR Workshop on “Bibliometric-enhanced Information Retrieval”
48. *iNIS 2017*: IEEE International Symposium on Nanoelectronic and Information Systems
49. *Coling 2016*: The 26th International Conference on Computational Linguistics
50. *HCOMP 2016*: The 4th AAAI Conference on Human Computation and Crowdsourcing
51. *EMNLP 2016*: Conference on Empirical Methods in Natural Language Processing
52. *SIGIR 2016*: The 39th International ACM SIGIR Conference on Research and Development in Information Retrieval
53. *NAACL 2016*: The 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies
54. *WWW 2016*: The 25th International World Wide Web Conference
55. *AAAI 2016*: The 30th AAAI Conference on Artificial Intelligence
56. *NLDB 2016*: The 21st Int. Conference on Application of Natural Language to Information Systems
57. *ISCRAM 2016*: The 13th Int. Conf. on Information Systems for Crisis Response and Management
58. *NLP4DL 2016*: Workshop on NLP and Information Retrieval for Digital Libraries
59. *PIR 2016*: Workshop on Proactive Information Retrieval collocated with ECIR 2016
60. *BIR 2016*: ECIR Workshop on “Bibliometric-enhanced Information Retrieval”
61. *EMNLP 2015*: Conference on Empirical Methods in Natural Language Processing
62. *BIR 2015*: ECIR Workshop on “Bibliometric-enhanced Information Retrieval”
63. *NLDB 2015*: The 20th Int. Conference on Application of Natural Language to Information Systems

64. *ISCRAM 2015*: The 12th Int. Conf. on Information Systems for Crisis Response and Management
65. *Coling 2014*: The 25th International Conference on Computational Linguistics
66. *ACL 2014*: The 52nd Annual Meeting of the Association for Computational Linguistics
67. *SMDM 2014*: ICDM workshop on “Social Multimedia Data Mining”
68. *ISCRAM 2014*: The 11th Int. Conf. on Information Systems for Crisis Response and Management
69. *SAC 2014*: Track on Recommender Systems in the 29th ACM Symposium On Applied Computing
70. *BIR 2014*: ECIR Workshop on “Bibliometric-enhanced Information Retrieval”
71. *SAC 2013*: Track on Recommender Systems in the 28th ACM Symposium On Applied Computing
72. *BCB - HIS 2013*: ACM Bioinformatics and Computational Biology
73. *IJCAI 2013*: The 23rd International Joint Conference on Artificial Intelligence
74. *ACL 2013*: The 51st Annual Meeting of the Association for Computational Linguistics
75. *SOCIETY 2013*: The International Conference on Social Intelligence and Technology
76. *SAC 2012*: Track on Recommender Systems in the 27th ACM Symposium On Applied Computing
77. *CIKM 2012*: ACM International Conference on Information and Knowledge Management
78. *SocialCom 2012*: ASE/IEEE International Conference on Social Computing
79. *IJCBS 2009*: Int. Joint Conferences on Bioinformatics, Systems biology and computational intelligence

Other Conference Article Reviewer:

1. *DocEng 2015*: The 15th ACM SIGWEB International Symposium on Document Engineering
2. *WABI 2014*: The Workshop on Algorithms in Bioinformatics
3. *BIBM 2013*: IEEE International Conference on Bioinformatics & Biomedicine
4. *CIKM 2013*: ACM International Conference on Information and Knowledge Management
5. *JCDL 2013*: Joint Conference on Digital Libraries
6. *BIBM 2012*: IEEE International Conference on Bioinformatics & Biomedicine
7. *WWW 2012*: International World Wide Web Conference
8. *KDD 2011*: ACM SIGKDD Conference on Knowledge Discovery and Data Mining
9. *JCDL 2011*: ACM/IEEE Joint Conference on Digital Libraries
10. *SIGIR 2011*: International ACM SIGIR Conference
11. *WWW 2011*: International World Wide Web Conference
12. *BIBM 2011*: IEEE International Conference on Bioinformatics & Biomedicine
13. *ICDM 2010*: IEEE International Conference on Data Mining
14. *ICDL 2010*: IEEE International Conference on Development and Learning
15. *ACM-BCB 2010*: ACM International Conference On Bioinformatics and Computational Biology
16. *ISMB 2010*: ISMB Student Council
17. *IJCAI 2007*: Indian International Conference on Artificial Intelligence

Journal Article Reviewer:

1. *TACL 2021*: Transactions of the Association for Computational Linguistics (TACL) (1 paper)
2. *TACL 2020*: Transactions of the Association for Computational Linguistics (TACL) (5 papers)
3. *TACL 2019*: Transactions of the Association for Computational Linguistics (TACL) (4 papers)
4. IEEE Transactions on Image Processing
5. *TCBB 2019*: IEEE/ACM Transactions on Computational Biology and Bioinformatics
6. Data & Knowledge Engineering
7. Entropy
8. Information Processing and Management

9. *LREV 2019*: Language Resources and Evaluation
10. *TBCB 2018*: Transactions of Bioinformatics and Computational Biology
11. *TACL 2018*: Transactions of the Association for Computational Linguistics
12. *TACL 2017*: Transactions of the Association for Computational Linguistics
13. *TKDE 2017*: Transactions on Knowledge and Data Engineering
14. *LREV 2017*: Language Resources and Evaluation, 2017
15. *PONE 2017*: PLOS ONE, 2017
16. *KBS 2017*: Knowledge-Based Systems, 2017
17. *IEEE IS 2017*: IEEE Intelligent Systems, 2017
18. *IEEE TIFS 2017*: IEEE Transactions on Information Forensics & Security, 2017
19. *IEEE IS 2016*: Information Systems, 2016
20. *TACL 2016*: Transactions of the Association for Computational Linguistics
21. *KNOSYS 2016*: Knowledge-Based Systems
22. *PONE 2016*: PLOS ONE
23. *JCCM 2016*: Journal of Contingencies and Crisis Management, Special Issue on Human-Computer Interaction in Critical Systems
24. *ACM TKDE 2015*: ACM Transactions on Knowledge and Data Engineering
25. *JAIR 2015*: Journal of Artificial Intelligence Research
26. *Entropy 2015*
27. *IJDMMM 2015*: International Journal of Data Mining, Modelling and Management
28. *IEEE/ACM TCBB 2014*: IEEE/ACM Transactions on Computational Biology and Bioinformatics
29. *ACM TBCB 2014*: ACM Transactions on Bioinformatics and Computational Biology
30. *ACM TKDE 2014*: ACM Transactions on Knowledge and Data Engineering
31. *PLOS ONE 2014*
32. *TAC 2014*: Transactions on Affective Computing
33. *Aslib 2014*: Aslib Proceedings
34. *NLE 2013*: Natural Language Engineering
35. *CEJCS 2013*: Central European Journal of Computer Science
36. *MDPI Information 2013*
37. *CBAC 2013*: Computational Biology and Chemistry
38. *ACM TKDE 2013*: ACM Transactions on Knowledge and Data Engineering
39. *ACM TIST 2013*: ACM Transactions on Intelligent Systems and Technology
40. *IEEE/ACM TCBB 2013*: IEEE/ACM Transactions on Computational Biology and Bioinformatics
41. *Nature 2012*
42. *Amino Acids 2012*
43. *IJDMB 2011*: International Journal of Data Mining and Bioinformatics
44. *BMC Bioinformatics 2010*
45. *IJDMB 2010*: International Journal of Data Mining and Bioinformatics
46. *NCA 2009*: Neural Computing and Applications
47. *GPB 2009*: Genomics, Proteomics & Bioinformatics
48. *Amino Acids 2009*
49. *BMC Bioinformatics 2007*

Guest Editor:

- *2021*: MATHEMATICS (An Open Access Journal by MDPI) Special Issue on Natural Language Processing (NLP) and Machine Learning (ML) – Theory and Applications.

- *2019 - 2020: Expert Systems* (Wiley Online Library) Special Issue on Mining Knowledge from Scientific Data.

Panelist:

- “Approaching the Singularity: The Future of Computational Intelligence.” Panel at the IEEE MetroCon at the Arlington Convention Center, in Arlington, TX, October 26, 2017.
- “Making the Leap: Jumping into An(other) Academic Position.” Panel at the Grace Hopper Celebration of Women in Computing, Orlando, FL, October 6, 2017.
- “Empowered by IT. Emerging Frontiers in Computation.” Panel at the 2011 IST Graduate Student Symposium: “Getting Graduate Students *There*”, The Pennsylvania State University.

Workshops Reviewer:

- *IJCAI 2017*: The International Joint Conference on Artificial Intelligence

Workshops Co-Chair:

- *NAACL HLT 2015*: The 2015 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies.

Conference and Workshop Session Chair:

- Recent Advances in Artificial Intelligence (**RAAI 2019**), Program Committee (PC) Chair, June 28-30, 2019, Bucharest, Romania
- The 31st AAAI Conference on Artificial Intelligence (**AAAI 2017**), 2016.
- The 30th AAAI Conference on Artificial Intelligence (**AAAI 2016**), 2016.
- ACM International Conference on Information and Knowledge Management (**CIKM 2012**), 2012.
- The AAAI-16 International Workshop on “Scholarly Big Data: AI Perspectives, Challenges and Ideas.” held in conjunction with The Association for the Advancement of Artificial Intelligence 2016 (**AAAI 2016**), 13 February 2016, Phoenix, AZ, USA.
- The AAAI-15 International Workshop on “Scholarly Big Data: AI Perspectives, Challenges and Ideas.” held in conjunction with The Association for the Advancement of Artificial Intelligence 2015 (**AAAI 2015**), 25-26 January 2015, Austin, TX, USA.
- The 1st Workshop on “Computational Scientometrics: Theory and Applications.” held in conjunction with the **iConference**, 2013.

Professional Societies Membership:

- Member of American Association for Artificial Intelligence (AAAI)
- Member of the ACL
- Member of the ACM
- Member of IEEE

Department and College Services

Presentations:

- September 1, 2021: Presented at the Meet and Greet day at CS@UIC for new students.

- August 29, 2019: Presented at the Meet and Greet day at CS@UIC for new students.
- November 29, 2018: Gave a presentation in the Library @ UIC.
- August 30, 2018: Presented at the Meet and Greet day at CS@UIC for new students.
- April 2017: Poster presentation at the *UNT Scholars Day 2017*.
 - Poster Title: “A New Method for Automatically Extracting Keyphrases from Scholarly Articles.” Presentation by Nathan Contreras.
- March 2014: Poster presentation at the *UNT graduate exhibition*.
 - Poster Title: “Mapping Moods: Geo-Mapped Sentiment Analysis During Hurricane Sandy.” Presentation by Kishore Neppalli.
- March 2014: Two poster presentations at the *Big Data open house* organized by the College of Engineering and the College of Information.
 - Poster Title: “Mapping Moods: Geo-Mapped Sentiment Analysis During Hurricane Sandy.” Presentation by Kishore Neppalli and Cornelia Caragea.
 - Poster Title: “Extracting Keyphrases from Research Papers using Citation Networks.” Presentation by Cornelia Caragea.

Other College Services:

- **Civic Plus and IBM:**
 - During Fall 2017, I participated in a research meeting with Civic Public in an endeavor to establish collaborations with local industry, as well as in a meeting with IBM representatives from Kansas City meant to identify opportunities for collaborations between IBM and KSU.
- **Proctor & Gamble:**
 - I had a meeting with Dr. David Ben-Arieh to discuss the potential benefits of recommender systems and social media data analysis to companies such as Proctor & Gamble.
- **Yahoo Research:**
 - I have continued my collaboration with Prakhar Biyani who was a research scientist at Yahoo Research, and is currently at Oath Inc., CA, to maintain collaboration with industry.
- **American Cancer Society:**
 - I have also continued my collaboration with the American Cancer Society (ACS) and have started to focus efforts to make an agreement of collaboration between ACS and K-State.
- **University of Bucharest:**
 - During Fall 2017, I put significant efforts in trying to establish a collaboration with University of Bucharest.
- **Fort Worth Museum of Science and History:**
 - I also continued to collaborate with the Fort Worth Museum of Science and History. We have made plans to submit a proposal for the museum, as an extensions to the BigData proposal, which was accepted in Summer 2017. I see this collaboration as an excellent outreach activity to reach out to the public and make them aware about research in computer science in general and NLP and crisis-related research in particular.
- **K-12 NSF RET:**
 - During 2017, I also started a collaboration with a K-12 teacher from Texas. We received a Research Experience for Teachers supplement from NSF.
- **Fidelity Visit Day:**

- I participated in the Fidelity Visit Day (the visit day was organized by Thomas Derryberry). The meeting was held at the Discovery Park, University of North Texas, Denton, TX, on September 27th, 2016. During the meeting I presented highlights of my research on Images' privacy prediction.

- **State Farm Day:**

- I participated in the State Farm Visit at UNT and met with State Farm representatives. I provided details regarding my research groups and the undergraduate students whom I worked with (the visit day was organized by Angus McColl). The meeting was held at the Discovery Park, University of North Texas, Denton, TX, on August 31st, 2016.

- **Ericsson-UNT Day:**

- I participated in the Ericsson UNT Day to seek opportunity for funding (the visit day was organized by Thomas Derryberry). The meeting was held at the Ericsson location, Plano, TX, on April 8th, 2016.

- **Meeting with the President:**

- I participated in the meeting with the President to present highlights of my research on data science. The meeting was held in the CSE Department at the Discovery Park, University of North Texas, Denton, TX, on March 25th, 2016.

- **Faculty Retreat:**

- I was invited to participate in a faculty retreat to discuss ways to improve and enhance our research climate at UNT, in order to elevate our scholarly performance across the College of Engineering. The meeting was held in Denton, TX, on January 13th, 2016.

Departmental Committees:

- Aug. 2021 - May 2022: Chair of the Faculty Search Committee at CS, UIC.
- Aug. 2021 - May 2022: Chair of the Mentoring Committee for the Bridge2Faculty Shweta Yadav.
- Aug. 2021 - May 2022: Part of the Promotion & Tenure Committee at CS, UIC.
- Aug. 2020 - May 2021: Part of the Graduate Committee at CS, UIC.
- Aug. 2020 - May 2021: Part of the Faculty Search Committee at CS, UIC.
- Aug. 2020 - May 2021: Part of the Promotion & Tenure Committee at CS, UIC.
- Aug. 2019 - May 2020: Part of the Graduate Committee at CS, UIC.
- Aug. 2019 - May 2020: Part of the Faculty Search Committee at CS, UIC.
- Aug. 2019 - May 2020: Part of the Promotion & Tenure Committee at CS, UIC.
- Aug. 2018 - May 2019: Part of the Graduate Committee at CS, UIC.
- Sept. 2016 - May 2017: Part of the Faculty Search Committee at CSE, UNT.
- Sept. 2015 - May 2016: Part of the Graduate Studies Committee at CSE, UNT.
- Sept. 2014 - May 2015: Part of the Graduate Studies Committee at CSE, UNT.
- Sept. 2013 - May 2014: Part of the Faculty Search Committee at CSE, UNT.
- Sept. 2013 - May 2014: Part of the Interdisciplinary Information Science PhD program Committee in LIS, UNT.

Other Department Services:

- July 2020 - present: Mentoring Assistant Professor Sathya Ravi at CS, UIC.
- Hosted the visit of Dr. Rada Mihalcea from University of Michigan, Ann Arbor, MI, in the CS@UIC, October 11, 2019.
- Hosted the visit of Dr. Sonia Haiduc from Florida State University, Tallahassee, Florida, in the ML group, May 11, 2016.
- Hosted the visit of Dr. Raymond Mooney from the University of Texas at Austin, Austin, Texas, June 26, 2015.
- Co-hosted, together with Dr. William Moen, the visit of Dr. Damir Cavar from the Institute for Language Information and Technology at Eastern Michigan University, February 20, 2014.
- Invited and hosted Dr. Anastasia Krithara from Insititute of Informatics and Telecommunications (IIT), NCSR Demokritos, Athens, Greece, December 12, 2013.
- Established a new research laboratory, “Machine Learning Research Group”, at the Department of Computer Science and Engineering, University of North Texas, Fall 2013.
- Invited and hosted Dr. C. Lee Giles from Penn State to visit UNT, October 5th, 2012.