

Di Lu

CONTACT INFORMATION

Dataminr Inc.
6 E 32nd St floor 2
New York, NY 10016 USA

Phone: (617) 650-6612
WWW: www.dylandilu.com
E-mail: dylanlu.dlu@gmail.com

RESEARCH INTERESTS

Natural Language Processing, Deep Learning, Computer Vision.

EDUCATION

Rensselaer Polytechnic Institute, Troy, New York USA

Ph.D., Electrical Engineering, May, 2019

- Advisor: Heng Ji, Ph.D

Boston University, Boston, Massachusetts USA

M.S., Electrical Engineering, May, 2012

- Advisor: Ronald Knepper, Ph.D

University of Science and Technology of China, Hefei, Anhui PRC

B.S., Physics, July, 2008

WORK EXPERIENCE

Research Scientist

Aug 2019 to Present

Dataminr Inc.
New York, NY, USA
Research Topic: Natural Language Processing, Computer Vision

Research Internship

September 05 2017 to December 22 2017

Snap Research
Venice, CA, USA
Manager: Vitor Carvalho, Ph.D
Research Topic: Multimodal Named Entity Recognition in Social Media

RESEARCH EXPERIENCE

Research Assistant

August 2015 to May 2019

Department of Computer Science,
Rensselaer Polytechnic Institute
Supervisor: Heng Ji, Ph.D
Research Focus Areas: Natural Language Generation, Information Extraction, Cross-media Cross-lingual Knowledge Transfer.

Research Assistant

January 2015 to August 2015

Department of Electrical Computer and System Engineering,
Rensselaer Polytechnic Institute
Supervisors: Kim Boyer, Ph.D
Research Focus Areas: Room Occupancy Detection using Color Sensor

TEACHING EXPERIENCE

Teaching Assistant, ENGR 2500, Electronic Instrumentation, Fall 2014, Rensselaer Polytechnic Institute

Teaching Assistant, ECSE 2050, Introduction to Electronics, Fall 2014, Rensselaer Polytechnic Institute

HONORS AND
AWARDS

Best Demo Paper Nomination, ACL2018

IBM CISL Fellowship, 2016

PUBLICATIONS

Ananya Subburathinam*, **Di Lu***, Heng Ji, Jonathan May, Shih-Fu Chang, Avirup Sil, Clare Voss. 2019. Cross-lingual Structure Transfer for Relation and Event Extraction. *Proc. the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP2019)*

Di Lu, Spencer Whitehead, Lifu Huang, Heng Ji and Shih-Fu Chang. 2018. Entity-aware Image Caption Generation. *Proc. the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP2018)*

Di Lu, Leonardo Neves, Vitor Carvalho, Ning Zhang, Heng Ji. 2018. Visual Attention Model for Name Tagging in Multimodal Social Media. *Proc. the 56th Annual Meeting of the Association for Computational Linguistics (ACL2018)*

Di Lu, Xiaoman Pan, Nima Pourdamghani, Shih-Fu Chang, Heng Ji and Kevin Knight. 2016. A Multi-media Approach to Cross-lingual Entity Knowledge Transfer. *Proc. the 54th Annual Meeting of the Association for Computational Linguistics (ACL2016)*

Di Lu, Clare R. Voss, Fangbo Tao, Xiang Ren, Rachel Guan, Rostyslav Korolov, Tongtao Zhang, Dongang Wang, Hongzhi Li, Taylor Cassidy, Heng Ji, Shih-fu Chang, Jiawei Han, William Wallace, James Hendler, Mei Si and Lance Kaplan. 2016. Cross-media Event Extraction and Recommendation. *Proc. the 2016 Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL-HLT 2016) (Demo)*

Ying Lin, Cash Costello, Boliang Zhang, **Di Lu**, Heng Ji, James Mayfield and Paul McNamee. 2018. Platforms for Non-speakers Annotating Names in Any Language. *Proc. the 56th Annual Meeting of the Association for Computational Linguistics (ACL2018) (demo)*

Boliang Zhang, Ying Lin, Xiaoman Pan, **Di Lu**, Jonathan May, Kevin Knight and Heng Ji. 2018. ELISA-EDL: A Cross-lingual Entity Extraction, Linking and Localization System. *Proc. the 2018 Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL-HLT 2018) (Demo)*

Boliang Zhang, **Di Lu**, Xiaoman Pan, Ying Lin, Halidanmu Abudukelimu, Heng Ji, Kevin Knight. 2017. Embracing Non-Traditional Linguistic Resources for Low-resource Language Name Tagging. *In proceedings of the International Joint Conference on Natural Language Processing (IJCNLP 2017)*

Rostyslav Korolov, **Di Lu**, Jingjing Wang, Guangyu Zhou, Claire Bonial, Clare Voss, Lance Kaplan, William Wallace, Jiawei Han and Heng Ji. 2016. On Predicting Social Unrest Using Social Media. *Proc. the IEEE/ACM International Conference on Advances in Social Networks and Mining (ASONAM2016)*

Pavlos Papadopoulos, Ruchir Travadi, Colin Vaz, Nikolaos Malandrakis, Ulf Hermjakob, Nima Pourdamghani, Michael Pust, Boliang Zhang, Xiaoman Pan, **Di Lu**, Ying Lin, Ondrej Glembek, Murali Karthick B, Martin Karafiat, Lukas Burget, Mark Hasegawa-Johnson, Heng Ji, Jonathan May, Kevin Knight and Shrikanth Narayanan. 2017. Team ELISA System for DARPA LORELEI Speech Evaluation 2016. *Proc. Interspeech2017*

Ulf Hermjakob; Qiang Li, Jonathan May, Daniel Marcu, Sebastian Mielke, Nima Pourdamghani, Michael Pust, Xing Shi, Kevin Knight, Kenton Murray, Tomer Levinboim, David Chiang, Boliang Zhang, Xiaoman Pan, **Di Lu**, Ying Lin, Heng Ji. 2017. Incident-Driven Machine Translation and Name Tagging for Low-resource Languages. *Proc. Machine Translation*

Tongtao Zhang, Ananya Subburathinam, Ge Shi, Lifu Huang, **Di Lu**, Xiaoman Pan, Manling Li, Boliang Zhang, Qingyun Wang, Spencer Whitehead, Heng Ji, Alireza Zareian, Hassan Akbari, Brian Chen, Ruiqi Zhong, Steven Shao, Emily Allaway, Shih-Fu Chang, Kathleen McKeown, Dongyu Li, Xin Huang, Xujun Peng, Ryan Gabbard, Marjorie Freedman, Ali Sadeghian, Mayank Kejriwal, Ram Nevatia, Pedro Szekely, Ali Sadeghian and Daisy Zhe Wang. 2018. GAIA - A Multi-media Multilingual Knowledge Extraction and Hypothesis Generation System. *Proc. Text Analysis Conference (TAC2018)*

Ronald Ahmed Cardenas Acosta, Thamme Gowda, Ulf Hermjakob, Nima Pourdamghani, Michael Pust, Jibiao Shen, Tian Xie, Jonathan May, Nikolaos Malandrakis, Victor Martinez, Pavlos Papadopoulos, Anil Ramakrishna, Karan Singla, Ruchir Travadi, Colin Vaz, Shrikanth Narayanan, Kenton Murray, Toan Nguyen, David Chiang, Boliang Zhang, Xiaoman Pan, Ying Lin, **Di Lu**, Lifu Huang, Tongtao Zhang, Kevin Blissett, Ni Zhang, Spencer Whitehead, Ananya Subburathinam, Diya Li, Qingyun Wang, Zhiying Jiang, Heng Ji, Ondrej Glembek, Murali Karthick Baskar, Santosh Kesiraju, Lukas Burget, Karel Benes, Igor Szoke, Karel Vesely, Jan Honza Cernocky, Derry Wijaya, Chris Callison-Burch, Hongzhi Xu, Charles Yang, Mitch Marcus, Brian Moran, Leanne Rolston, Gina-Anne Levow, Wenda Chen, Camille Goudeseune, Leda Sari and Mark Hasegawa-Johnson. 2018. ELISA System Description for LoReHLT 2018. *Proc. LoReHLT2018*

Leon Cheung, Thamme Gowda, Ulf Hermjakob, Nelson Liu, Jonathan May, Alexandra Mayn, Nima Pourdamghani, Michael Pust, Kevin Knight, Nikolaos Malandrakis, Pavlos Papadopoulos, Anil Ramakrishna, Karan Singla, Victor Martinez, Colin Vaz, Dogan Can, Shrikanth Narayanan, Kenton Murray, Toan Nguyen, David Chiang, Xiaoman Pan, Boliang Zhang, Ying Lin, **Di Lu**, Lifu Huang, Kevin Blissett, Tongtao Zhang, Heng Ji, Ondrej Glembek, Murali Karthick Baskar, Santosh Kesiraju, Lukas Burget, Karel Benes, Igor Szoke, Karel Vesely, Jan “Honza” Cernocky, Camille Goudeseune, Mark Hasegawa Johnson, Leda Sari, Wenda Chen and Angli Liu. 2017. ELISA System Description for LoReHLT 2017. *In proceedings of the LoReHLT2017*

Mohamed Al-Badrashiny, Jason Bolton, Arun Tejavsi Chaganty, Kevin Clark, Craig Harman, Lifu Huang, Matthew Lamm, Jinhao Lei, **Di Lu**, Xiaoman Pan, Ashwin Paranjape, Ellie Pavlick, Haoruo Peng, Peng Qi, Pushpendre Rastogi, Abigail See, Kai Sun, Max Thomas, Chen-Tse Tsai, Hao Wu, Boliang Zhang, Chris Callison-Burch, Claire Cardie, Heng Ji, Christopher Manning, Smaranda Muresan, Owen C. Rambow, Dan Roth, Mark Sammons, Benjamin Van Durme. 2017. Tinker-Bell: Cross-lingual Cold-Start Knowledge Base Construction. *In proceedings of the Text Analysis Conference (TAC2017)*

Dian Yu, Xiaoman Pan, Boliang Zhang, Lifu Huang, **Di Lu**, Spencer Whitehead and Heng Ji. 2016. RPI.BLENDER TAC-KBP2016 System Description. *Proc. Text Analysis Conference (TAC2016)*

Ulf Hermjakob, Qiang Li, Jonathan May, Sebastian Mielke, Nima Pourdamghani, Michael Pust, Xing Shi, Kevin Knight, Daniel Marcu, Nikolaos Malandrakis, Anil Ramakrishna, Victor Martinez, Elisabeth Staruk, Tanner Sorensen, Dogan Can, Shrikanth Narayanan, Tomer Levinboim, Kenton Murray, David Chiang, Boliang Zhang, Xiaoman Pan, **Di Lu**, Lifu Huang, Xiaocheng Feng and Heng Ji. 2016. ELISA System Description for LoReHLT 2016. *Proc. NIST LoReHLT 2016 Workshop*

Yu Hong, **Di Lu**, Dian Yu, Xiaoman Pan, Xiaobin Wang, Yadong Chen, Lifu Huang and Heng Ji. 2015. RPI.BLENDER TAC-KBP2015 System Description. *Proc. Text Analysis Conference (TAC2015)*

SERVICES

Program Committee: LREC 2020, NAACL-HLT-Demo Track 2019, Shortcomings of Vision and Language Workshop at NAACL-HLT 2019
Conference Reviewer: ACL 2020, NAACL 2018
Journey Reviewer: IEEE Transactions on Circuits and Systems for Video Technology, Social Network Analysis and Mining

EVALUATIONS

1st in LORELEI evaluation Information Extraction track. 2016, 2017, 2018

1st in Event Nugget Detection and Coreference (Spanish)

9th International Meeting on Text Analysis Conference (TAC) 2016 Workshop, Gaithersburg, Maryland USA, November, 2016.

1st in Event Nugget Detection and Coreference (English)

8th International Meeting on Text Analysis Conference (TAC) 2015 Workshop, Gaithersburg, Maryland USA, November, 2015.

COMPUTER SKILLS

- Languages: Python, Java, JavaScript, Matlab, C/C++, C#, Unix shell scripts.
- Applications: Pytorch, Tensorflow, Caffe, Amazon Web Services, Google Cloud, NLTK, Stanford coreNLP, OpenCV, Rpi, Docker, Github, Scikit-Learn
- Operating Systems: Unix/Linux, Windows, OSX.