Heather Lent

PhD Student in Natural Language Processing

hcl@di.ku.dk | University of Copenhagen, Denmark

PERSONAL SUMMARY

I am a PhD student studying Natural Language Processing (NLP). My goal is to research and develop breakthrough methods for multilingual and crosslingual learning, in order to bring state of the art performance to low resource languages for semantics-focused NLP tasks (e.g. semantic parsing, coreference resolution, machine translation).

EDUCATION

University of Copenhagen *March 2019 - Present*
Ph.D. Student in Department of Computer Science, Machine Learning Section
TALENT Programme fellowship recipient; Marie Curie PhD fellow
Advisor: Anders Søgaard

University of Arizona *Aug 2014 - May 2016*
M.Sc. Human Language Technology
Cumulative GPA: 3.8
Thesis committee: Peter Jansen, Michael Hammond, Sandiway Fong

University of Arizona *Aug 2011 - May 2015*
B.A. Linguistics
B.A. Russian and Slavic Studies
Cumulative GPA: 3.9

WORK EXPERIENCE

VINCI - Department of Veterans Affairs & University of Utah (Salt Lake City, UT, USA)
Natural Language Processing Specialist *Feb 2018 - Feb 2019*
PI's: Olga Patterson, PhD & Scott DuVall, PhD
Developing NLP pipelines for information extraction from medical records

- Creating NLP pipelines for information extraction of medical concepts in Java and Scala
- Scaling NLP pipelines to run on 4+ billion patient records with UIMA_AS
- Working closely with clinicians to ensure pipeline quality
- Working closely with Oak Ridge National Labs to train models on supercomputers
- Participating in clinical NLP challenges (i.e. n2c2)

Bio5 Institute, Lyons Lab (Tucson, Arizona, USA)
Research Specialist *May 2016 - Feb 2018*
PI: Eric Lyons, PhD
Developing software to help biomedical researchers understand how their work is being used by others (via citations from PubMed).

- Using NLP and ML to create topic models for biomedical texts through traditional algorithms and experimental methods using word embeddings
- Information retrieval, biomedical named entity recognition, and document clustering
- Using Python (Flask) and JavaScript to design front end of website and interactive data

visualizations
- Architecting Python back end source code and database schema (SQL), with attention to scalability

LAB EXPERIENCE

Computational Language Understanding Lab (Tucson, Arizona, USA)
Intern *May 2017 - March 2018*
PI: Mihai Surdeanu, PhD
Conducting information extraction for DARPA's World Modelers project at the CLU lab, a group that focuses on NLP and deep learning for the open domain.

- Using ODIN (Open Domain Informer), a rule-based information extraction framework, to build a robust grammar for identifying important entities and events in crop models
- Designing comprehensive unit tests to ensure the grammar's quality (Scala)
- Building lexicons by extracting hyponym/hypernym relations from Resource Description Frameworks

Bio5 Institute, Lyons Lab (Tucson, Arizona, USA)
Intern *January 2016 - May 2016*
PI: Eric Lyons, PhD
Hired as Research Specialist (May 2016). Please see description in "Work Experience."

Ewha Woman's University, Bioinformatics Lab (Seoul, South Korea)
PI: Hyeon-Seok Park, PhD
Research Intern *Summer 2015*
Internship focused on biomedical text mining and biological sequence analysis using NLP techniques.

- Scraping the web with Python and BeautifulSoup to access text about epigenetics
- Using Python and NLTK to analyze the chromatic maps from the ENCODE Analysis Working Group, in order to search for patterns in promoter sequences

Psycholinguistics and Computational Linguistics Lab (Tucson, Arizona, USA)
Lab Assistant *January 2013 - December 2013*
PI: Adam Ussishkin, PhD

- Contributed to the Maltese Dictionary Project, which aims to preserve Maltese through an online dictionary
- Facilitated lab experiments with human subjects and worked with DMDX experimental software

PROGRAMMING LANGUAGES

Primary: Python
Secondary: Java and Scala
Previous work with: Javascript, Perl, Matlab, Prolog

NATURAL LANGUAGES

English: Native, US citizen
Russian: Advanced Level
Korean: Intermediate Level
Past exposure to: Latin (5 years), Japanese (JLPT Level 4, Passed in 2008)

PUBLICATIONS

Rebecca Sharp, Adarsh Pyarelal, Benjamin Gyori, Keith Alcock, Egoitz Laparra, Marco A. Valenzuela-Escárcega, Ajay Nagesh, Vikas Yadav, John Bachman, Zheng Tang, **Heather Lent**, Fan Luo, Mithun Paul, Steven Bethard, Kobus Barnard, Clayton Morrison and Mihai Surdeanu. (2019). *Eidos & Delphi: From Free Text to Executable Causal Models*. (Accepted to NAACL-HLT 2019).

Shi, Jianlin, Eyre H.R., Peterson, K.S., **Lent, H.C.**, Graves, K.G., Chapman, A.B., Hurdle, J.F., Patterson, O.V. (2019). *Adverse Drug Event Extraction Leveraging Knowledge Resources and Generalizable Features*. (Submitted to AMIA).

Lent, H., Hahn-Powell G., Haug-Baltzell A., Davey S., Surdeanu M., Lyons, E. (2018). *Science Citation Knowledge Extractor. Frontiers in Research Metrics and Analytics*. https://www.frontiersin.org/article/10.3389/frma.2018.00035

Lent, H., Lee, K.-E., & Park, H.-S. (2015). *Building the Frequency Profile of the Core Promoter Element Patterns in the Three ChromHMM Promoter States at 200bp Intervals: A Statistical Perspective*. Genomics & Informatics, 13(4), 152–155. http://doi.org/10.5808/GI.2015.13.4.152

PRESENTATIONS

"Using the Python Programming Language." **Lent, Heather**. Department of Veteran Affairs Health Services Research and Development Service (HSR&D) Cyber Seminars. January 10, 2019.

"Hybrid Models for Medication and Adverse Drug Events Extraction." Kelly Peterson; Jianlin Shi; Alec Chapman; Hannah Eyre; **Heather Lent**; Kevin Graves; Jianyin Shao; Subhadeep Nag; Olga Patterson; John F. Hurdle. American Medical Informatics Association (AMIA), November 2, 2018.

"Microvascular Disease Determines Limb Outcomes in PAD (Clinical Text Processing for PAD)." Beckman, Jason MD; Barnett, Joey PhD; Frieberg, Matthew MD; Wasserman, David PhD; Wells, Quinn MD; **Lent, Heather MS**. American Heart Association SFRN Annual Meeting. August 12, 2018.

"Meetups Meet Up." Dearman, Marnee; Hansen, Jason; Joyce, Blake; **Lent, Heather**; Wellington, Doug. Panel discussion for Tucson's TenWest festival. October 18, 2017.

"Code Across Campus: Machine Learning and Software Innovation at the UA." **Lent, Heather**. Presentation for Tech Launch Arizona, University of Arizona, Tucson, AZ, April 5, 2017.

"Natural Language Processing for Digital Humanities." **Lent, Heather**; Melo, Maggie; Nichols, Jennifer. Workshop at the Research Bazaar Arizona festival, Tucson, AZ, April 1, 2017.

"Python and Natural Language Processing." **Lent, Heather**. Workshop at the Tucson Python Meetup, Tucson, AZ. October 2016.

"Introduction to Natural Language Processing." **Lent, Heather**. Workshop at the University of Arizona, Science and Engineering Library, iSpace, Tucson, AZ, October 2016.

"Hidden Markov Models in Bioinformatics." **Lent, Heather**. Presentation at the Tucson Data Science Meetup, Connect Coworking, Tucson, AZ, May 2016.

LEADERSHIP

Co-organizer of the Tucson Data Science Meetup (*January 2017 - November 2017)*

- Coordinating with Bio5 for our meeting place and arranging for data scientists in the

community to lead a workshop once a month

Co-organizer of Research Bazaar Arizona (*June 2016 - November 2017*)

- Teaching digital literacy to researchers from all fields through community building

Member of Toastmasters International (*June 2016 - June 2017*)

- Effective oral presentation skills and leadership training

Mozilla Science Lab's Open Leadership Training (*March 2017 - June 2017*)

- Completed training for leading successful open source projects

Volunteer at Women Techmakers Tucson Hackathon (2017)

- Helped participants of the hackathon with programming