

Stephen W. Mayhew

mayhew@seas.upenn.edu
mayhewsw.github.io

Research Interests Natural Language Processing, Named Entity Recognition, Low-resource Techniques, Transliteration, Multilingual Techniques, Trustworthiness, Information Extraction

Education Ph.D. Computer Science, 2017 -
Advisor: Dan Roth
University of Pennsylvania, Philadelphia, PA

Ph.D. Computer Science, Fall 2014 - 2017
Advisor: Dan Roth
University of Illinois at Urbana-Champaign, Urbana, IL

M.S. Computer Science, May 2014
Thesis: Trustworthiness and the Importance of Graph Structure
University of Illinois at Urbana-Champaign, Urbana, IL
GPA: 3.67/4.0

B.S. Computer Science with a Minor in Music, May 2012
Rose-Hulman Institute of Technology, Terre Haute, IN
GPA: 3.91/4.0, Summa Cum Laude

Selected Publications **On the Strength of Character Language Models for Multilingual Named Entity Recognition**
*Xiaodong Yu, **Stephen Mayhew**, Mark Sammons, Dan Roth*
EMNLP, 2018

Talen: Tool for Annotation Low-resource ENTities
***Stephen Mayhew**, Dan Roth*
ACL Demonstrations, 2018, github.com/CogComp/talen

Cheap Translation for Cross-Lingual Named Entity Recognition
***Stephen Mayhew**, Chen-Tse Tsai, Dan Roth*
EMNLP, 2017

Transliteration in Any Language with Surrogate Languages
***Stephen Mayhew**, Christos Christodoulopoulos, Dan Roth*
arxiv 2016, <https://arxiv.org/abs/1609.04325>

Cross-Lingual Named Entity Recognition via Wikification
*Chen-Tse Tsai, **Stephen Mayhew**, Dan Roth*
CoNLL, 2016

Visualization of Trustworthiness
***Stephen Mayhew**, Dan Roth*
WWW Workshop: Rumors and Deception in Social Media (RDSM), 2015

Overview of UI-CCG systems for event argument extraction, entity discovery and linking, and slot filler validation

*Mark Sammons, Yangqiu Song, Ruichen Wang, Gourab Kundu, Chen-Tse Tsai, Shyam Upadhyay, Siddharth Ancha, **Stephen Mayhew**, Dan Roth*
 Proceedings of the Seventh Text Analysis Conference (TAC2014), 2014.

Faster (and Better) Entity Linking with Cascades

*Adrian Benton, Jay Deyoung, Adam Teichert, Mark Dredze, Benjamin Van Durme **Stephen Mayhew**, Max Thomas*

Automated Knowledge Base Construction (AKBC), 2014

IllinoisCloudNLP: Text Analytics Services in the Cloud

*Hao Wu, Zhiye Fei, Aaron Dai, **Stephen Mayhew**, Mark Sammons, Dan Roth*
 Language Resources and Evaluation Conference (LREC), 2014

RECOG: Recognition and Exploration of Content Graphs

*Kara Greenfield, **Stephen Mayhew**, William Campbell*
 Pacific Visualization, 2013

Presentations:

Robust Information Extraction from Multiple Sources

CCICADA Research Retreat 2013, Houston, TX

**Academic
Honors**

State Farm Companies Foundation Doctoral Scholar, 2015-2016

Siebel Scholar Award, 2013

Frank Young Senior Scholarship award, 2012

Dean's List, all quarters of undergrad degree

Member of Upsilon Pi Epsilon, Computer Science Honor Society

Experience

UPenn, LORELEI Summer Evaluation

Summer 2018

Summer Project - with many others

- Built an NER/EDL system for Sinhalese and Kinyarwanda with no gold annotated training data
- Helped coordinate a team to build infrastructure and annotate data

UPenn, LORELEI Summer Evaluation

Summer 2017

Summer Project - with Chase Duncan

- Built an NER/EDL system for Oromo and Tigrinya with no gold annotated training data

UIUC, LORELEI Summer Evaluation

Summer 2016

Summer Project - with Chen-Tse Tsai

- Built an NER system for Uyghur with no gold annotated training data
- Built an annotation interface for NER

Microsoft Research, Redmond, WA

Summer 2015

Research Intern

- Worked on *GALEN*, a decision support system for clinical genomics
- Designed and implemented the frontend, and wrote code to interface with the backend
- Worked with Hoifung Poon, Kristina Toutanova, Chris Quirk, Scott Yih

	Johns Hopkins University , Baltimore, MD <i>SCALE2013 Research workshop participant, Text Track</i> <ul style="list-style-type: none"> • Created <i>slinky</i>, the streaming entity-linker • Used parallelism techniques to make <i>slinky</i> very fast 	Summer 2013
	University of Illinois , Urbana-Champaign, IL <i>Research Assistant for Dan Roth</i> <ul style="list-style-type: none"> • Work on trustworthiness techniques, especially as they apply to information extraction • Testing and refactoring of group software 	2012 – Present
	MIT Lincoln Labs , Boston, MA <i>Human Language Technologies Intern, Group 62</i> <ul style="list-style-type: none"> • Surveyed several graph visualization tools • Developed plugins for a specific visualization task 	Summer 2012
	Interactive Intelligence Indianapolis, IN <i>Development Intern, Speech Analytics team</i> <ul style="list-style-type: none"> • Developed a prototype Text-To-Speech engine using Festival • Researched multiple ways to improve voice quality • Advised team members on duplicating engine prototype • Wrote extensive documentation for voice-building process 	Summer 2011
	Alcoa Howmet , Whitehall, MI <i>IT Intern</i> <ul style="list-style-type: none"> • Conducted department-wide interviews regarding workflow • Designed and deployed a license tracker utility • Developed several VBA solutions for Office applications • Developed programs to read Word files, and write to Excel files 	Summer 2010
Teaching Experience	University of Pennsylvania , Philadelphia, PA CIS530, Natural Language Processing, Teaching Assistant for Chris Callison-Burch	Spring 2018
	University of Illinois , Urbana, IL CS446, Machine Learning, Teaching Assistant for Julia Hockenmaier	Spring 2015
	Rose-Hulman , Terre Haute, IN <i>Teaching assistant, grader</i> <ul style="list-style-type: none"> • Grader for CS freshman year courses, Databases, Requirements and Specifications, Design and Analysis of Algorithms 	2009 – 2012
Service	External reviewer for TKDE 2014, EMNLP2018	
	Graduate Student Ambassador for prospective students University of Illinois	2014-2016
	Graduate Student Member on the CS Fellowships, Admissions, and Assistantship Committee	Spring 2015

Other Experience	English teacher in Daejeon, Korea Violinist Terre Haute Symphony Orchestra (semi-pro)	Summer 2009 2005 – 2012
Skills	Python, Java, Scala, HTML, CSS, JavaScript, C#, C, MySQL, PHP, Scheme, Objective-C, elisp, Haskell. Data visualization - d3.js, Gephi, JUNG, matplotlib, Cytoscape Working knowledge of Spanish, preliminary knowledge in Korean, Latin, Greek, Hebrew	
Citizenship	Dual citizenship US/UK.	