

KATHERINE A. KEITH

kak5@williams.edu ◊ <https://kakeith.github.io/>

EDUCATION

University of Massachusetts Amherst
Ph.D. in Computer Science, September 2021
M.S. in Computer Science, February 2020

2016–2021
Amherst, MA

Lewis & Clark College
B.A. in Mathematics with departmental honors, summa cum laude
Minor: Chinese Language

2011–2015
Portland, OR

APPOINTMENTS

Assistant Professor
Department of Computer Science
Williams College

2022–Present
Williamstown, MA

RESEARCH EXPERIENCE

Postdoctoral Young Investigator
Allen Institute for Artificial Intelligence

September 2021–June 2022
Seattle, WA

- Research with the Semantic Scholar team

Graduate Research Assistant
University of Massachusetts, Amherst

September 2016–August, 2021
Amherst, MA

- Natural language processing, machine learning, and computational social science research
- Advisor: Dr. Brendan O'Connor

Research Intern
AI Discovery Team, Bloomberg

June–August 2020
Remote (London, UK team)

- Mentors: Dr. Edgar Meij, Dr. Christoph Teichmann
- Resulted in publication at the *Workshop on Natural Language Processing and Computational Social Science (NLP+CSS) at EMNLP, 2020*.

Research Intern
CTO Data Science Team, Bloomberg

May–August 2018
New York, New York

- Mentor: Dr. Amanda Stent
- Resulted in publication at the *Association of Computational Linguistics (ACL), 2019*.

Undergraduate Research Assistant (Economics)
Lewis & Clark College

May–August 2014
Portland, OR

- Developed an agent-based simulation of intergenerational wealth transfer in medieval England
- Funding: Andrew W. Mellon Collaborative Student-Faculty Research Grant
- Advisor: Dr. Clifford Bekar

PUBLICATIONS

8. Text as Causal Mediators: Research Design for Causal Estimates of Differential Treatment of Social Groups via Language Aspects. **Katherine A. Keith**, Douglas Rice, and Brendan O'Connor. In *Proceedings of the First Workshop on Causal Inference and Natural Language Processing (CI+NLP) at EMNLP*. 2021.
7. Corpus-Level Evaluation for Event QA: The IndiaPoliceEvents Corpus Covering the 2002 Gujarat Violence. Andrew Halterman*, **Katherine A. Keith***, Sheikh Muhammad Sarwar*, and Brendan O'Connor (* indicates joint first-authors). In *Findings of the Association for Computational Linguistics (ACL-IJCNLP)*. 2021.
6. Uncertainty over Uncertainty: Investigating the Assumptions, Annotations, and Text Measurements of Economic Policy Uncertainty. **Katherine A. Keith**, Christoph Teichmann, Brendan O'Connor, and Edgar Meij. In *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS) at EMNLP*. 2020.
5. Text and Causal Inference: A Review of Using Text to Remove Confounding from Causal Estimates. **Katherine A. Keith**, David Jensen, and Brendan O'Connor. In *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL)*. 2020.
4. Modeling financial analysts' decision making via the pragmatics and semantics of earnings calls. **Katherine A. Keith** and Amanda Stent. In *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL)*. 2019.
3. Uncertainty-aware generative models for inferring document class prevalence. **Katherine A. Keith** and Brendan O'Connor. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP)*. 2018.
2. Monte Carlo Syntax Marginals for Exploring and Using Dependency Parses. **Katherine A. Keith**, Su Lin Blodgett, and Brendan O'Connor. In *Proceedings of North American Chapter of the Association for Computational Linguistics (NAACL)*. 2018.
1. Identifying civilians killed by police with distantly supervised entity-event extraction. **Katherine A. Keith**, Abram Handler, Michael Pinkham, Cara Magliozzi, Joshua McDuffie, and Brendan O'Connor. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP)*. 2017.

TEACHING EXPERIENCE

Instructor

CS335: Machine Learning
 Mount Holyoke College

Spring 2020

- Sole instructor for 19 students. Designed and delivered lectures, held office hours, developed problem sets, designed exams, and graded problem sets and exams.

Instructor

CS191: Computer Science First Year Seminar
 University of Massachusetts Amherst

Fall 2019

- Co-designed curriculum on "Ethical Issues Surrounding Artificial Intelligence Systems and Big Data."
- Led two weekly discussion sections comprising of 19 students each.

Graduate Teaching Assistant

CS685: Advanced Natural Language Processing
 University of Massachusetts Amherst

Spring 2018

- Assisted students with course material and homework during weekly office hours.
- Co-designed homework assignments, graded literature review assignment and in-class presentations.

Fulbright English Teaching Assistant

U.S. Department of State

August 2015–June 2016

Kinmen, Taiwan

- Taught first through sixth grade ESL courses in a public elementary school.
- Facilitated multi-cultural dialogue and events.

GRANTS

- Social Science Research Council (SSRC)/Summer Institutes in Computational Social Science (SICSS) Research Grant 2021 (\$1200). With Ian Stewart.
- Kaggle Open Data Research Grant 2020 (\$5000). With Andy Halterman and Sheikh Sarwar.

INVITED TALKS

- Williams College’s Computer Science Colloquium—November 15, 2019
- Rutgers University’s Natural Language Processing Group—November 6, 2019
- Lewis & Clark College’s Mathematical Sciences Colloquium—September 24, 2018

GUEST LECTURES

- Georgia Institute of Technology, CS-6471: Computational Social Science, “Text and Causal Estimation: Text as Causal Confounders and Mediators”—March 28, 2022
- University of Massachusetts Amherst, Natural Language Processing—November 13, 2020
- Mount Holyoke College, Natural Language Processing—November 13, 2019
- Mount Holyoke College, Natural Language Processing—October 10, 2018
- University of Massachusetts Amherst, Advanced Natural Language Processing—April 26, 2018

TUTORIALS

- Aggregated Classification Pipelines: Propagating Probabilistic Assumptions from Start to Finish. NLP+CSS 201 Online Tutorial Series. March 30, 2022.

SERVICE & OUTREACH

Reviewer / Referee / Program Committee

- *Journal of Data Science*: February, 2022
- *Science Advances*: October, 2021
- *Workshop of Teaching NLP*: 2021
- *ACL*: 2020, 2021
- *NAACL*: 2019, 2021
- *EACL*: 2021
- *CoNLL*: 2020
- *EMNLP*: 2020
- *ICWSM*: 2019, 2020

Service Leadership Roles

- Co-organizer, “CS+NLP 201: Beyond the Basics” online tutorial series (Fall 2021)
- Co-organizer and Panel Moderator, “First Workshop on Causal Inference and Natural Language Processing” at EMNLP 2021

Other Outreach Roles

- Panelist, University of Massachusetts CSWomen panel on “Applying to Research Internships and Jobs” (November 19, 2021)
- Panelist, NSF CSGrad4US Mentoring Program panel on “What I wish I knew when I started graduate school” (November 24, 2021)
- PhD Applicant Support Mentor for Underrepresented Students (November 2020) University of Massachusetts Amherst
- Panelist, CS Graduate School Q&A Panel (May 1, 2020) Cape Cod Community College
- Co-chair, CSWomen (February 2019–January 2020) University of Massachusetts Amherst Computer Science Women’s Group
- Social Co-Chair, CSWomen (January 2017–January 2019) University of Massachusetts Amherst Computer Science Women’s Group
- Student Volunteer, Women in Engineering and Computing Career Day (October 28, 2019) University of Massachusetts Amherst
- “What is computer science?” 7th Grade Guest Lecture (Jan. 3, 2018 and Jan. 7, 2019) Chief Joseph Middle School, Bozeman, Montana
- Organizer, CICS Graduate Male Ally Workshop Series (September 2017– May 2018) University of Massachusetts Amherst
- Student Volunteer, Girls Inc. Eureka! Summer Workshop (August 1 & 3, 2017) University of Massachusetts Amherst
- Student Volunteer, Women in Engineering and Computing Career Day (October 24, 2016) University of Massachusetts Amherst

GRADUATE AWARDS & HONORS

- Outstanding Reviewer, EMNLP 2020
- Awarded *distinction* for Ph.D. candidacy portfolio submitted to the College of Information and Computer Sciences at University of Massachusetts Amherst (December 2019)
- Bloomberg Data Science Ph.D. Fellowship (awarded May 2019)
- Computing Research Association of Women (CRA-W) Graduate Cohort (2017, 2018)
- Empirical Methods in Natural Language Processing (EMNLP) Student Travel Scholarship (2017)
- Paul Utgoff Memorial Graduate Scholarship in Machine Learning (2016)

UNDERGRADUATE AWARDS & HONORS

- Fulbright ETA Grantee with the U.S. Department of State (2015–16)
- Rhodes Scholarship Finalist (2015)
- Marshall Scholarship Finalist (2015)
- Rena Ratte Award, Lewis & Clark College (2015)– “*This is the highest academic award given at Lewis & Clark and is given annually to one undergraduate of senior standing whose abilities and commitment have combined to produce work which is consistently of the greatest distinction.*” <https://college.lclark.edu/dean/faculty-and-student-awards/>
- NCAA Division III Cross-Country All-Academic (2012)
- Lewis & Clark College Cross-Country Four-Year Varsity Letter Winner (2011–2015)

FOREIGN LANGUAGES

Chinese (Mandarin)

HSK Level 4 (test passed April 16, 2016)

CET Beijing: 16-week language-intensive immersion program (Spring 2014)