University of Massachusetts Amherst

Katherine A. Keith, PhD Candidate
College of Information and Computer Sciences
University of Massachusetts Amherst
Amherst, Massachusetts 01003
kkeith@cs.umass.edu
https://kakeith.github.io/

October 15, 2020

Professor Stephen Freund
Chair, Department of Computer Science
Williams College

Dear members of the search committee,

I am applying for the position of Tenure-Track Assistant Professor at Williams College because I am committed to the liberal arts ethos of developing well-rounded, interdisciplinary scholars and I have personally witnessed the powerful impact of effective teaching and mentoring on a young person's career path. As an undergraduate, I attended Lewis & Clark College, a small liberal arts college in Portland, Oregon, where I made deep, lasting mentoring relationships with my professors. A liberal arts college allowed me to pursue a broad course of study: I majored in Mathematics but combined this with a minor in Chinese Language and numerous courses in Computer Science and Economics, all of which currently feed into my PhD research. As a graduate student, I taught a four-credit *Machine Learning* class to nineteen students at Mount Holyoke College. I found these liberal arts students, all of whom identified as female or non-binary, to be dedicated to the learning process and deeply inquisitive of the machine learning technologies I taught. This invigorating teaching experience solidified my future career interests as a liberal arts college professor. As a professor, one of my major goals is to send computer science students into the work force and academia with both technical rigor and an ability to critique the *social impact* of the technologies they build.

Currently, I am a final-year PhD student in the College of Information and Computer Sciences at the University of Massachusetts Amherst, and I plan to defend my thesis in June, 2021. My research is under the domain of *social data science*, investigating human behavior through quantitative analysis of large-scale data, and I specifically focus on *text* data since language is one of the most salient expressions of human thought and behavior. In order to make textual analysis scale to large datasets, I adapt and extend methods in *natural language processing* and *machine learning* for applications such as extracting macro-social signals from news documents and analyzing the language of economic decision making. For instance, I published work in the *Empirical Methods in Natural Language Processing (EMNLP)* conference proceedings that implements a data science pipeline with a novel machine learning model to identify the names of civilians killed by police from 1.2 million news articles. I also published work in the conference proceedings of the *Association for Computational Linguistics (ACL)* that examines financial analysts' decision making as it pertains to the semantics and pragmatics of financial earnings call transcripts. My work also expands supervised and unsupervised machine learning methods for more accurate *measurement* of social signals from text and uses these improved measurements in *causal* problem settings.

My research aligns with the agenda of a liberal arts college like Williams College because it is interdisciplinary, combining computer science with the social sciences, and provides many entry-ways into research for undergraduates. During the summer of 2019, I recruited and mentored three undergraduate students for a research experience for undergraduates (REU) program at the University of Massachusetts Amherst. One of these students extended his summer work for a successful senior honors thesis, which I co-advised. When working with undergraduates, I *scaffold* research such that students with less technical backgrounds apply machine learning methods while collecting, cleaning, and manually annotating data, and students with more technical experience extend machine learning or natural language processing methods for social data science goals. In the past, I worked with economists, political scientists, and journalists who wanted to analyze text data at large-scale and I see working at Williams College as an opportunity to collaborate with other domain experts in the social sciences and humanities.

I previously taught machine learning, natural language processing, and a freshman seminar on the ethics of artificial intelligence, and I believe teaching these courses, along with foundational courses in computer science and data science, is imperative to keep students competitive in the Information Age. In my classes, I use *active learning* techniques such "think-pair-share" exercises and cold-calls, and I work to motivate students' *sense of purpose* via project-based learning. It is my personal teaching philosophy that technical computer science courses should also provide students with practice in *science communication* and contemplating the *ethics* of the technologies they build.

Effective teaching is an important factor in *improving diversity, equity, and inclusion* in computer science, and as such I work to actively decrease bias in my classroom by grading anonymously, *teaching transparently* by giving students assessment rubrics and explicit directions on how to succeed in my classes, and continuously collecting student feedback to adapt my courses to students' needs. I am dedicated to improving the *inclusivity* of computer science culture, and as a graduate student, I served as Co-Chair of CSWomen, a grassroots organization of female graduate students in computer science. Our group arranged weekly social events, organized a panel of senior graduate women to dispense research advice, and established a mentoring program between older and younger students. As a professor, I will explicitly seek for more equitable inclusion of students of color and women in my research. I am also interested in organizing a programming tools bootcamp for underrepresented students in which we provide instruction on tools not covered in most computer science classes such as the command line, shell tools, version control, text editors, and remote machines. I believe these efforts of intentionally reducing bias in teaching, establishing peer support networks, and providing additional instruction for underrepresented students could help improve inequitable demographic representation in computer science as a whole.

At Williams College, I am interested in collaborating with **Dr. Andrea Danyluk** on applied machine learning projects, and **Dr. Iris Howley** on the intersection between human-computer interaction and natural language processing, specifically automated linguistic analysis for education. On a personal note, I have thoroughly enjoyed living in Western Massachusetts as a graduate student and I am excited by the prospect of staying in the region to work at Williams College. When I visited Williams in November, 2019 to speak at the college's Computer Science Colloquium, I was deeply impressed by the collaborative and student-centered nature of the department and I would be honored if I was able to join as a professor. Ultimately, I believe bringing my research, teaching, and diversity experience to Williams College could help build Williams College students into interdisciplinary, data-driven scholars and future leaders of the tech workforce.

Sincerely,

Katherine A. Keith, PhD Candidate