As technology becomes increasingly commonplace, AI systems must be able **serve members of society equally**. Achieving this goal will require broad and diverse participation, making it crucial to **dismantle barriers to diversity in computer science** (e.g., lack of role models[1], stereotypical culture[2]), which I have made a priority during my graduate studies at the University of Washington's (UW) Allen School of Computer Science & Engineering.

As a professor, I plan to continue my efforts to **achieve positive societal impact with research** and **increase diversity and belonging in computer science** in three different directions, drawing on my past experience:

- **Positive Societal Impact Through Research** – research and workshop organization to promote fairness, justice, and ethics in artificial intelligence
- **Broadening Participation Through Advising and Outreach** – mentorship and outreach to support and recruit women, students of color, and LGBTQ+ students
- **Supportive and Inclusive Environment Through Structural Change** – departmental service, research community leadership, and community organization

In all my efforts, I strive to use evidence-based practices and center experiences from historically underrepresented groups (including, when relevant, my own as an out LGBTQ+ person).

Positive Societal Impact Through Research

As technology becomes increasingly powerful and widespread, it is important to create artificial intelligence (AI) and natural language processing (NLP) systems that further social justice, make positive societal impact, and avoid harm. For that reason, both my research and my academic community organization have focused on societally impactful questions.

Research for social good – In my 2019 ACL work [5], I **uncovered racial bias in hate speech detection systems**, showing that they flag speech by African American authors as toxic more often than speech by white authors. In recent works [6, 7], I developed automatic tools for **measuring and mitigating social biases and stereotypes in text**, to help authors rewrite sentences in less biased ways. In addition, my work on social commonsense inference has been used to enhance response suggestions on an online mental health counseling platform [8].

Academic workshops for positive impact – This year, I am **co-organizing an ACL 2021 workshop on NLP for positive impact** to encourage NLP research applications that can, for example, improve user and public well-being, foster constructive conversations, or promote justice and fairness. In 2018, I was invited to a workshop to discuss the promise of commonsense reasoning and AI for assisting people with cognitive disabilities [9].

Looking ahead, I will continue my organizational efforts (e.g., through workshops), and keep pursuing socially relevant research directions (e.g., detecting, mitigating biases in NLP), which, as a positive side effect, can increase women and minorities' interest in the field [10].

Broadening Participation Through Advising and Outreach

A significant aspect of my efforts has focused on increasing participation in computer science research, through mentorship and outreach, as described below.

[1]Research shows that the dearth of role models is a important barrier to diversity in CS [1, 2]. For example, a recent Queer In AI survey cites the "lack of role models" as the number one obstacle (83% of respondents) to becoming a queer AI/ML practitioner.

[2]Social science experts have highlighted the importance of cultural and environmental conditions in recruiting and retaining underrepresented groups in computer science [3, 4]

Advising – I **advised 12 undergraduate or master students of diverse identities and backgrounds** (50% women, 75% students of color) on research projects. As a senior graduate student, I have also mentored 5 junior PhD students (including Asian-American and African-American women). With my help, many of these students' papers were submitted or accepted to top-tier NLP conferences.

Outreach & recruitment – In 2017 and 2018, I co-organized events to **introduce K-12 and high school students to artificial intelligence** and to showcase my team's competition-winning chatbot. Additionally, I helped **recruit more diverse students** by attending the 2017 Tapia Diversity in Computing conference and organizing LGBTQ+ events during the Allen School's PhD visit days.

Moving forward, I plan on expanding my efforts to increase diverse participation in CS education at several levels. Knowing that high-school CS education is not available equitably [11, 12], I will continue outreach to K-12 schools, focusing on schools with diverse student populations. For recruiting my PhD students, I plan to **advertise PhD positions on minority-focused networks**, e.g., specialized conferences (Grace Hopper, Tapia, etc.), minority-serving institutions (e.g., HBCUs, women's colleges), and affinity group mailing lists (e.g., BlackInAI, QueerInAI).

Supportive and Inclusive Environment Through Structural Change

Throughout my graduate career at the Allen School, I quickly recognized the importance of lasting structural change to foster inclusive environments. Therefore, I spent significant time devoted to departmental service and community organization, as described below.

School leadership – I was a founding member of the Allen School diversity committee, as well as the Allen School graduate advisory council. With the diversity committee, I worked to make graduate admissions process more fair, and organized LGBTQ+ events for all members of the school to foster a sense of community. As part of the advisory council, I led an effort to **make our PhD course requirements more accessible** to students from non-traditional backgrounds by revising and expanding the list of required courses.

Research community organization – I **co-chaired the socio-cultural diversity and inclusion subcommittee** at ACL 2020, which included facilitating the organization of social events by groups such as Black in AI, Queer in AI, etc.[3] I also helped petition for EquiCL, a broad interest group on diversity and inclusion in the ACL community. I am also **an active member of Queer in AI**, for which I helped organize the ICML 2019 and ACL 2020 social events.

Informal community organization – Since depression and imposter syndrome are significant barriers to succeeding in PhD programs [especially for minority students; 13], I **created a support group for NLP students** to discuss personal and academic setbacks, which I led quarterly for two years. Additionally, as social chair of the Allen School in 2016, I made successful efforts to include my fellow international students in social events, which led some students to organize Nowruz and Lunar New Year celebrations for the school.

In the future, I plan on continuing these efforts to foster an inclusive, welcoming and supportive environment which is crucial for all students to thrive [14], while following best practices. For example, I will **encourage my students to engage in cultural competency and diversity training**, and will help create spaces for them to support each other. In my classes, I will focus on inclusive and anti-racist teaching techniques [15], for example, using concrete grading rubrics to avoid unconscious biases.[4]

[3] A full report is available on the ACL Wiki.

[4] Following the CS Education Research Blog's recommendations on inclusive and anti-racist teaching.

References

[1] Joanna Goode. Increasing diversity in k-12 computer science: Strategies from the field. In *Proceedings of the 39th SIGCSE technical symposium on Computer science education*, pages 362–366, 2008.

[2] Christian Voigt, Elisabeth Unterfrauner, and Roland Stelzer. Diversity in fablabs: culture, role models and the gendering of making. In *International Conference on Internet Science*, pages 52–68. Springer, 2017.

[3] Lenore Blum, Carol Frieze, Orit Hazzan, and M Bernardine Dias. A cultural perspective on gender diversity in computing. *Reconfiguring the firewall*, pages 109–129, 2007.

[4] Sapna Cheryan, Victoria C Plaut, Paul G Davies, and Claude M Steele. Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology*, 97(6):1045, 2009.

[5] **Maarten Sap**, Dallas Card, Saadia Gabriel, Yejin Choi, and Noah A Smith. The risk of racial bias in hate speech detection. In *ACL*, 2019.

[6] **Maarten Sap**, Saadia Gabriel, Lianhui Qin, Dan Jurafsky, Noah A Smith, and Yejin Choi. SOCIAL BIAS FRAMES: Reasoning about social and power implications of language. In *ACL*, 2020.

[7] Xinyao Ma*, **Maarten Sap***, Hannah Rashkin, and Yejin Choi. POWERTRANSFORMER: Unsupervised controllable revision for biased language correction. In *EMNLP*, 2020.

[8] William R Kearns, Neha Kaura, Myra Divina, Cuong Vo, Dong Si, Teresa Ward, and Weichao Yuwen. A Wizard-of-Oz interface and persona-based methodology for collecting health counseling dialog. In *CHI Extended Abstracts*, CHI EA '20, pages 1–9, New York, NY, USA, April 2020. Association for Computing Machinery.

[9] Clayton Lewis. Implications of developments in machine learning for people with cognitive disabilities. *ACM SIGACCESS Accessibility and Computing*, (124):1–1, 2020.

[10] Sarah Blanchard Kyte and Catherine Riegle-Crumb. Perceptions of the social relevance of science: Exploring the implications for gendered patterns in expectations of majoring in STEM fields. *Social sciences*, 6(1):19, February 2017.

[11] Jennifer Wang and Sepehr Hejazi Moghadam. Diversity barriers in K-12 computer science education: Structural and social. In *SIGCSE*, 2017.

[12] Monique Ross, Zahra Hazari, Gerhard Sonnert, and Philip Sadler. The intersection of being black and being a woman: Examining the effect of social computing relationships on computer science career choice. *ACM Trans. Comput. Educ.*, 20(2):1–15, February 2020.

[13] Teresa M Evans, Lindsay Bira, Jazmin Beltran Gastelum, L Todd Weiss, and Nathan L Vanderford. Evidence for a mental health crisis in graduate education. *Nature biotechnology*, 36(3):282–284, March 2018.

[14] Michail N Giannakos, Ilias O Pappas, Letizia Jaccheri, and Demetrios G Sampson. Understanding student retention in computer science education: The role of environment, gains, barriers and usefulness. *Education and Information Technologies*, 22(5):2365–2382, 2017.

[15] Amy J Ko. How to be a great (CS) teacher. https://medium.com/bits-and-behavior/how-to-be-a-great-cs-teacher-b8a0a2a3600f, May 2017. Accessed: 2020-11-6.