

Diversity Statement

Qingyun Wang

I am firmly committed to diversity, equity, and inclusion in our community, since “*science is done best when it is approached from as many different angles as possible*”¹. Despite the rapid growth of artificial intelligence in our society, computer science in the US still faces the disparity between gender and race. According to the 2023 Taulbee survey from the Computing Resource Association in the US², only 25.4% of new CS PhDs are women. Additionally, Black or African American students account for 1.6%³ of PhD enrollment, while Hispanic students only make up 2.1%⁴. During my PhD, I have been vigorously promoting diversity and inclusion. I will continue these efforts as a professor in my research, mentoring, teaching, and outreach activities.

Promoting Diversity through Mentoring During my undergraduate research, as an international student, I was fortunate to receive guidance from my advisor and several senior PhD students, who encouraged me to express my ideas and helped me understand the entire research lifecycle. Motivated by my personal experience, during my PhD, I have the privilege to mentor over **10 students from diverse backgrounds** in **gender** (five of them are female) **geographical location** (including US, China, India, HK), and **academic level** (undergraduate, MS, PhD). As a mentor, I aim to prepare students from all backgrounds with knowledge and skills for their success in careers, especially those facing systemic barriers. One exceptional mentee was Priya Pitre, a master’s student from Virginia Tech interested in racial and gender bias in large language models (LLMs). At the outset of her journey, she was unfamiliar with the code implementation for LLMs. I provided her with online tutorials and additional one-to-one meetings to help her implement and debug. Through our weekly meetings, she gradually gained confidence and is currently submitting a paper to ACL rolling review. I also **actively assisted students with their PhD applications**, including reviewing application materials, sharing research experience, and suggesting potential labs. All of my advisees continued to pursue graduate school or STEM careers. As an assistant professor, I will continue to create an inclusive and supportive environment for students from all backgrounds.

Building Supportive and Inclusive Communities. In the third year of my PhD program, I was fortunate to join the efforts for the **ACL 2022 D&I Special Initiative Globalization via localization**, which aims to “*remove the ingrained linguistic bias in the scientific landscape in general and computational linguistic science in particular*”. The initiative acts as *a catalyst in the democratization of computational linguistics sciences*. I dedicated over three months to extracting and curating scientific terminology from ACL Anthology papers and creating cross-lingual wikification for the extracted terms. Moreover, I served as a reviewer in **Widening Natural Language Processing 2024 (WiNLP 2024)**, which aims to promote and support ideas and voices of underrepresented groups. Inspired by these activities, I created a **student mentorship initiative** for my *AI4Research* Workshop at AAAI 2025, which aims to inspire students, particularly those with underrepresented backgrounds, to pursue research careers. As an international student, I joined **UIUC English Corner** as a Treasurer to organize activities that can *help international students overcome the language barrier*. All of these efforts aim to reduce the unique barriers faced by underrepresented groups—including language, culture, and lack of research opportunities—and to promote their research visibility. In the future, I will continue to promote the voices of underrepresented groups. For instance, I plan to organize and participate in workshops at top-tier conferences focused on these goals.

Encouraging Interdisciplinary Collaborations. As an NLP researcher working in AI4Science, **my research greatly benefits from collaboration with people from diverse fields**, such as computer vision, human-computer interaction, education, chemistry, medicine, biology, and agriculture. For example, my **NAACL Best Demo paper** about drug repurposing in COVID-19 greatly benefited from collaboration with the Cardiovascular Research Laboratory at UCLA. I also worked with the *Carl R. Woese Institute for Genomic Biology* from the *Chemistry Department* at UIUC and the *Department of Communicative Disorders and Sciences* from the *University at Buffalo* to develop information extraction systems for scientific papers. Additionally, to enable broader access to natural language processing (NLP), I have taught **tutorials** at top NLP conferences, including *EMNLP 2021*, *EACL2024*, and *COLING 2024*, with all

¹<https://www.nature.com/articles/d41586-023-02138-y>

²<https://cra.org/wp-content/uploads/2024/05/2023-CRA-Taulbee-Survey-Report.pdf>

³13.7% in the US population. <https://www.census.gov/quickfacts/fact/table/US/PST045223>

⁴19.5% in the US population. <https://www.census.gov/quickfacts/fact/table/US/PST045223>

code, slides, and reading lists available online for inclusive access. I am also **organizing highly interdisciplinary workshops** such as *AI4Research* at AAI2025 and *Language + Molecule* at ACL 2024, where we invited panelists from other disciplines such as astronomy, chemistry, physics, material, and biology. In the future, I will continue to collaborate with people with different areas of expertise to broaden my perspective.

Engaging in Ethical and Socially-aware Research. Large language models (LLMs) are substantially changing people's daily lives. Therefore, I want to utilize their abilities to **reduce communication barrier**. For example, I joined the AI4Education project, which aims to build a knowledge graph for the screening and intervention of children who have learning and language disabilities. Throughout my research, I have become more aware of challenges faced by children with speech or language-related concerns, as well as the possible intervention methods to alleviate them. Recognizing that LLMs also sometimes exhibit biases or harm in their generation results, I mentored several research projects to analyze and mitigate those problems, including *LLM bias tailoring* and *self-correction*. Looking ahead, I see my long-term vision of *AIScientist* as being an essential tool in alleviating communication barriers of researchers from different backgrounds. Such systems can help *democratize biological/material synthesis and characterization*, and *achieve unbiased data collection and analysis*. At the same time, I also aim to research scientific fact-checking to alleviate increasing misinformation in scientific publications.

Future Plans. My overall goal is to create an inclusive environment where all computer scientists can use their expertise to pursue their interests. I firmly believe that a welcoming, diverse, and inclusive research community will have a better ability to advance science for the betterment of society. In the future, I will continue to commit to the goal of diversity and inclusion in different ways. (1) As a future faculty member, I will work to create a welcoming and inclusive teaching and research environment that offers safe spaces for underrepresented groups to express their concerns. I will embrace and welcome underrepresented minorities to join my research lab to work on diverse research projects. (2) Across campus, I intend to join or create diversity and inclusion initiatives or groups to support underrepresented groups to eliminate stereotypes and provide equal opportunities. I will also join DEI committees at the department or school level. Drawing on my experience related to undergraduate research, I aim to create opportunities for undergraduates from different backgrounds to conduct research and participate in mentorship programs in my field, e.g., through ACL Mentorship and the UR2PhD program. Meanwhile, I plan to collaborate with faculty members from other departments to gain different research perspectives. (3) Within the broader community, I plan to join and participate in DEI outreach activities, such as WiNLP, WiML, NCWIT Aspirations in Computing, Queer in AI Workshop, etc. When organizing workshops, I will continue to ensure a diverse balance of speakers.