

My dedication to promoting diversity in higher education is anchored in both sociological research and personal experiences. It is evident that students from underprivileged backgrounds stand to gain the most from higher education. Furthermore, a diverse student body enriches the academic experience for everyone involved. As an educator, it is essential to actively support and retain students from marginalized groups. My decision to pursue an advanced degree in Computer Science was largely influenced by my observations of disparities in the tech industry across countries having better infrastructure and resources versus those lacking them. It is widely known that Computer science as a discipline has issues with diversity (Fry et al., 2021). My experience as a Ph.D. student at Columbia University has taught me the importance of not only creating inclusive pathways for underrepresented groups but also ensuring they receive equal support to thrive. During my PhD, I have been actively working towards enhancing diversity and inclusion. As I transition into a professorial role, I plan to lean on expert advice and proven strategies to further this mission.

## 1 Inclusive Advising and Mentoring

Over the last half-decade, I have supported junior scholars, guiding them according to their unique strengths and pace. I have worked with ten students, both from renowned institutions like Columbia University, UCLA, and USC as well as from comparatively underrepresented educational settings in countries such as India and China. Out of these, four were females, and six belonged to minority groups. As a senior graduate student at Columbia University, I have also mentored high school students. One of these students Aadit Trivedi, with my assistance, has presented his research at EMNLP 2021. As a mentor, my aim is to establish a varied research team, recognizing and respecting their individual backgrounds outside academia. My mentoring style is strength-based, and I am committed to offering resources committed to offering resources to help students from all backgrounds feel included and succeed.

## 2 Committee Service and Outreach

Progress in NLP is driven by the individuals at the heart of the research. We gain insights from their contributions. Yet, to produce such remarkable outcomes, what methodologies and philosophies do they adopt? What obstacles, difficulties, missteps, and lessons emerged along the way? These insights are immensely beneficial for individuals at varying stages of their careers. Indeed, when embarking on a new professional journey, we typically seek guidance from more experienced peers: (1) newly enrolled PhD students turn to senior PhDs or recent graduates, (2) emerging researchers consult with established or seasoned academics, and (3) newcomers in the corporate world seek advice from industry veterans. However, in many cases, only a handful of these seasoned professionals are accessible to us. To help with these challenges I organized *The First Workshop of Story Shared and Lesson Learned @ EMNLP 2022*<sup>1</sup> to make the sharing of successful researchers' stories and lessons learned to be accessible to everyone in our community.

At Columbia University, I have actively engaged in departmental duties, community building, and outreach initiatives aimed at promoting diversity and fostering an inclusive environment. Since 2020, through the Columbia Women in Computer Science (WICS) outreach initiative<sup>2</sup> I have guided numerous female undergraduates, mentoring them about AI research career opportunities, clarifying misconceptions, and linking them with established researchers in their fields of interest.

Being an international student who was the first in my family to attend university, I deeply understand the challenges of applying to graduate school. Through Columbia Computer Science's PhD Pre-Submission Application Review (PAR) Program<sup>3</sup> I have guided numerous potential graduate students, offering advice on their applications and aiding with their interview preparations through online

---

<sup>1</sup><https://ssll-emnlp.github.io/>

<sup>2</sup><https://cuwics.github.io/>

<sup>3</sup><https://www.cs.columbia.edu/cscu-phd-par-program/>

sessions. Many of these mentees have since joined PhD programs. To attract Ph.D. candidates, I intend to promote the doctoral positions through networks that cater to minority groups. This includes specialized conferences like Grace Hopper and Tapia, educational institutions that predominantly serve minority communities such as HBCUs and women’s colleges, and dedicated group mailing lists like BlackInAI and QueerInAI.

### 3 Research

As AI solutions grow in strength and ubiquity, it is crucial to develop technologies that are representative of their user base’s diversity and prevent harm. This has been a central theme in my research on AI-assisted writing. Large Language Models such as ChatGPT/GPT4 have become powerful enough to write thousands of words coherently. This has led to a 150-day-long strike from the Writers Guild of America. On the societal front, the way individuals and communities interact with and are affected by AI writing tools is profound. These tools can influence how information is consumed and produced, altering communication dynamics, and potentially leading to homogenized narratives or reinforcing existing societal biases. Thus, ensuring that AI-assisted writing platforms are both technically robust and socially responsible becomes a crucial sociotechnical challenge.

My recent work (Chakrabarty et al., 2023b) on AI-assisted fiction writing looks into these concerns and highlights how current models are limiting for professional writers in several ways including their repetitive nature, over-reliance on clichés and tropes, lack of nuance, subtext, or symbolism as well as overly moralistic and predictable endings. Similarly my work Chakrabarty et al. (2023a) has integrated theoretical foundations from cognitive sciences (Torrance, 1966) and social psychology (Amabile, 1982) leading to the development of a rubric for Creative Writing evaluation that has been used by domain experts to discriminate and identify AI written text from that of expert-written text. Our evaluation framework highlights the limitations of large language models in producing creative text as well as their inability to assess creative text.

By identifying the limitations of AI in reproducing clichés, tropes, and predictable narratives, my research sheds light on how AI might perpetuate existing biases in literature. Through public discussion of the shortcomings of AI in creative writing, my research tries to raise awareness about the potential pitfalls of relying too heavily on AI for content generation. Moving ahead, my intention is to further delve into research that emphasizes diverse representation and inclusiveness in NLP models spanning various languages and cultures. This not only enhances the model’s capabilities but also heightens the interest of the communities represented. Over time, I envision that my aim of placing people at the core of AI system development will play a crucial role in mitigating the adverse effects associated with the current population-focused approach. To assess the impact of these interventions and to raise awareness about the research domain, I aim to work closely with experts from HCI, social sciences, and public policy fields.

### References

- Teresa M Amabile. 1982. Social psychology of creativity: A consensual assessment technique. *Journal of personality and social psychology*, 43(5):997.
- Tuhin Chakrabarty, Philippe Laban, Divyansh Agarwal, Smaranda Muresan, and Chien-Sheng Wu. 2023a. Art or artifice? large language models and the false promise of creativity. *arXiv preprint arXiv:2309.14556*.
- Tuhin Chakrabarty, Vishakh Padmakumar, Faeze Brahman, and Smaranda Muresan. 2023b. Creativity support in the age of large language models: An empirical study involving emerging writers. *arXiv preprint arXiv:2309.12570*.
- Richard Fry, Brian Kennedy, and Cary Funk. 2021. Stem jobs see uneven progress in increasing gender, racial and ethnic diversity. *Pew Research Center*, pages 1–28.
- Ellis Paul Torrance. 1966. *Torrance tests of creative thinking: Norms-technical manual: Verbal tests, forms a and b: Figural tests, forms a and b*. Personal Press, Incorporated.