

Seminar Report on expert talk on our Implicit Stereotypes and AI

Date: 19 March, 2024

Venue: Auditorium, Maharaja Surajmal Institute of Technology

Duration: 12:30-1:00 P.M.

Organizer: CSE Department

Event Coverage: Broadcast MSIT

Introduction:

Ms. Ananya Sangwan delivered a talk session on our implicit stereotypes and AI on Day 1 of ICAIA 2024 on March 19, 2024.

Speaker:

Ms Ananya Sangwan is a computer science scholar at UC Berkley, who is passionate about how our implicit biases are finding their way into AI machines and affecting their working in ways we can't even comprehend.

Highlights:

She delivered a talk on whether AI machines function as mechanically as we believe, exploring implicit biases in AI.

She quoted an example from the US where an algorithm was created to predict recidivism which was developed using skewed data, leading to an individual with a certain skin tone being given serious punishment for crimes as petty as shoplifting or the Apple iPhone being unable to distinguish between two Asians. She also mentioned an experiment conducted by students in which they entered various sentences in gender-neutral language like he or she is an engineer and Google translate would need to complete the gender. The results showed that science-based professions were attributed to males whereas nurturing and caring fields like nursing were assigned to females.

In conclusion of her talk, she stated that any such biases or stereotypes possessed by humans are passed on to machines, which in the long run can be corrected by overcoming such biases. Meanwhile, it is important to ensure that a diverse panel reviews any such judgment programs or datasets that are going to be used, to minimize such cases of biases.

She was felicitated with a certificate and memento.



Google

New Delhi, Delhi, India
1, Block C 4A, Janakpuri, New Delhi, Delhi, 110058, India
Lat 28.620904°
Long 77.092484°
19/03/24 02:30 PM GMT +05:30

GPS Map Camera