Contribution ID : **18**   Type : **Talk**

Biased AI image generations models and misconceptions about health conditions

Tuesday, 13 September 2022 11:40 (30)

In recent years we could observe more text-to-image models, most notably DALLE, DALLE-2, and DALLE mini. Those models allow generating images based on user prompts. They were trained on datasets of images and captions crawled from the web. But those datasets contain some biases, especially those present in media. However DALLE and DALLE-2 are debiased models against at least race and gender, but DALLE-mini is not which makes him possible to pick up biases from the dataset. Some of those biases might be linked to misconceptions in society. Like prompt 'autistic child' will be 9 out of 9 cases a boy in preschool age, but 'autistic girl' will be presented as in secondary school years.

In the talk, we would like to show how not debiased models text-to-image could serve as a tool to study misconceptions about health that exists in society.

Primary author(s) : ZADROŻNY, Adam (National Centre for Nuclear Research); ZADROŻNA, Marianna ( )

Presenter(s) : ZADROŻNY, Adam (National Centre for Nuclear Research); ZADROŻNA, Marianna ( )

Session Classification : Openning session