I have led a discussion on paper 4 and in our discussion, we have touched on some ethical/ societal points such as how large datasets based on texts from the Internet overrepresent hegemonic viewpoints and we also touched on in the discussions how training it off the majority excludes minorities, making a point on how although results are important, but the process of how the model gets there is just as important, if not more.