

Timnit Gebru NB

May 7, 2023

— title: Learning from Timnit Gebru author: Mia Tarantola date: '2023-04-16'

description: "A blog post relating to our talk with Timnit Gebru" format: html —

1 Learning from Timnit Gebru

Mia Tarantola

1.1 Introduction

This blog post will contain two parts. The first portion will be completed prior to our talk and will be based on self guided research on our speaker Timnit Gebru. I will take notes and pose some thoughtful questions that I think will make for a good discussion.

The second portion will be completed during/after Gebru's talk. I will take notes and reflect on the material.

Dr. Timnit Gebru is a successful American computer scientist who focuses on artificial intelligence and algorithmic bias. She has become a well known advocated for diversity in technology and even founded her own community , Black in AI. In December 2020 Gebru left her position as the Ethical Artificial Intelligence Team lead at Google, following pushback from her as-yet unpulished paper voicing her concerns on the dangerous biases of large language models.

1.2 Part 1: Questions for Dr. Gebru

1.2.1 Notes of Dr. Grebru's talk: conference on Computer Vision and Pattern Recognition 2020

[recording of the talk](#)

Timnit Gebru's talk on "Computer Vision in practice: who is benefitting and who is being harmed?" focuses on the ethical ramifications of computer vision technology in our society. Grebru emphasizes the need for critical thought on the possible effects of new technologies on marginalized communities as a top researcher in artifical intelligence and ethics.

Gebru's talk emphasizes the danger that can result from the use of computer vision technology when it is implemented without adequate consideration of the potential consequences. She discusses how face recognition technology may be used to mistakenly identify people based on their gender and skin color, potentially leading to false charges and arrests. Additionally, predictive policing systems that rely on biased data sets can perpetuate systemic discrimination against communities of color,

leading to increased surveillance and targeting. The ethics of data collecting and utilization are also covered by Gebru. She emphasizes that a model's deployment might still be unethical even if its accuracy is identical among all groups. Images scraped from numerous online sources are frequently included in data sets, often without the subjects' knowledge. She says that people are unlikely to want their pictures to be used in face recognition software to identify protesters.

Gebru argues that it is essential for individuals who create and use computer vision systems to take into account these possible negative effects and strive toward creating more open and accountable business processes. She stresses the importance of diversity and inclusion in the creation of moral and responsible AI systems. She also encourages businesses and decision-makers to give underrepresented populations' demands and worries first priority while developing and implementing computer vision technologies.

Gebru's presentation serves as a reminder of the essential part AI practitioners and researchers must play in influencing the creation and use of computer vision technologies. We can progress toward developing AI systems that are beneficial to every member of society by considering the ethical implications of these systems and including a variety of viewpoints into their development.

TL;DR In order to create ethical and just predictive modeling systems, we need to prioritize the need for diverse and inclusive development and implementation practices.

1.2.2 Questions to ask Timnit Gebru

- *How can we address the issue of biased data sets in computer vision technology, and what strategies can be used to prevent the continuation of systemic discrimination?*
- *What role should companies and policymakers play in regulating the use of computer vision technology, and what policies and practices would you recommend?*
- *Given the potential for danger/harm caused by computer vision technology, what ethical considerations should AI researchers keep in mind when designing and deploying these models?*
- *Can you share any examples of organizations or initiatives that you believe are making progress in addressing the ethical implications of AI, and what we can learn from them?*
- *How can we better educate the public about the ethical implications of computer vision technology, and what steps can be taken to increase awareness and engagement on this issue?*

1.3 Notes on In-Class Talk/ Q&A

- *theft normalized > - scraping images - might not be an image people want taken > - precarious workforce that can't advocate for themselves > > - people are scared to lose opportunities > - lack of enforcement, copyright etc*

Independent research org vs larger tech companies - *She was fired when she spoke out at a larger tech companies - People are too scared to say anything at companies like google - her company to beholden to that incentive - She still has to fundraise for her institute - Not only her job on the line but also everyone that works at the institution - Can create an environment where people can have diverse ideas - Silence no more act - illegal for people to enforce NDA in cases of harrassment, bullying, discrimination -> easier for people to speak up*

seen any improvements since talk? - *hard to change names of NIPS conference, close to porn site name, so much backlash - every little change is so contentous - resistance: GLAZE artist can run*

their artwork through this software before posting online, confuses ML models so they cannot copy this work > - worked with many artists

Mitigating bias - ways in use cases of computer vision that actively combat discrimination/bias - forensic lab at NYU, connects protest bombs/weapons to certain companies - costa-rica conservation work: plant identification - YOLO - announced his leave from computer vision, cannot ignore the military applications anymore - 1% good does not outweigh the 99% bad use cases

very few people who are black and women in CV community - ever felt imposter syndrome - seeing unqualified men, man created cult with no educational background - supporter of airstrikes - no imposter syndrome, people make you feel like that - sexism, racism - see how many unqualified people -> drop imposter syndrome - when in environments that are waiting for her to fail, then she feels it - about environment and what it instills in you

1.4 Notes on night talk

- founded DAR Institute
- mitigating dangers of AI
- imagining/executing a better tech future
- millions of exploited workers that are fueling these AI projects > - chat GPT moderator suffering from trauma and PTSD

Utopia for whom? - AGI = artificial general intelligence > - AI that outperforms humans (smart, well educated human) - roots in eugenics > - not progressive movement - eugenics never went away, even after WWII - focus on “improving the human stock” > - get rid of undesirable traits/people : negative view > - give people the ability to design their children -> design more intelligent kids > - tell desirable people to reproduce more - Second wave eugenics > - post human: a new superior species can be created > - legacy humans: everyone else (less desirable) > - intelligence explosion

cosmism - humans will merge with technology - develop sentient AI and mind uploading technology

TESCREAL bundle - historical roots - transcending humanism - Galton - make parents take intelligence exam before reproducing - some people advocating to harm researchers to stop AI apocalypse

Discriminatory view - Bostrom: “Blacks are more stupid than whites” > - too many people with lower IQ reproducing too much -> new species - Pelvitz: below an IQ of 120, 0 points - influence > - all billionaires in movements (Elon Musk, etc) donating > - after the 70s researchers dissociate from AI -> lean more towards ML, CV, NLP - Deep mind founded and then bought by Google - Open AI -> Microsoft

AGI Utopia - AGI will be so intelligent that it can figure out a solution/what to do to any scenario - morally superior AGI enhanced transhuman minds benefitting

AGI Utopia for who? - ” On the dangers of stochastic parents...” - text to image models - used for deep fakes, overly sexualizing women, harassment campaigns - resources not going to organizations around the world who serve their own communities, goes to just one community > - poor people get AI doctors, rich get humans

No language left behind - investor pull out of smaller companies - larger companies often use a data set that is an output of another dataset = training on test data = NO! - NLLB doing much worse than ghana NLP

AGI Apocalypse - TESCREALists argue probability of existential risk → any event that would destroy our chances of creating utopia happening this century = high - morally obliged to create AGI Utopia and prevent AGI Apocalypse - Anthromorphising artificats allows builders to avoid accountability - deply religious → everlasting life

unsafe and unscientific

1.5 Part 2

Dr. Gebru's talk focused on artificial intelligence, specifically artificial general intelligence (AGI) and its relation to eugenics. AGI is an artificial intelligence algorithm or software that can outperform a smart, well-educated human. She was able to shed light on some of the ideologies that drive the AGI movement forward. She, and some collaborators have coined the term TESCREAL to describe some of these motivations.

The "T" stands for transhumanism, which is the belief that people should evolve as a species by merging biological and synthetic technology to enhance themselves. More specifically, there would be two different species of the human race. The first: post-human - humans who are smarter, more intelligent and thus would be allowed to reproduce more. The second: legacy humans - those who are deemed "lesser than or defective" and would not get to reproduce.

The E stands for extropianism which argues that human can reverse or alter entropy to have an extended or everlasting life.

The S stands for singularitarianism which is the belief that technology will continue to evolve and eventually design itself, leading to a technology explosion.

The C standing for cosmism - the idea that humans will merge with technology and the development sentient AI and mind uploading technology.

The EA refers to effective altruism which is the belief to do the most good.

Longtermism is a philosophy that prioritizes the maximization of future intelligences. These ideas all strive for a utopia but a utopia for who?

ChatGPT provides an AI chatbot full of information, but at the expense of who? There have been stories of CHATGPT employees who moderate the algorithms responses and have been traumatized by the experience. Some have even developed PTSD.

She also touched on the dangers of AGI. One aspect was the dangers of data collection and image scraping. She alluded to the fact that many images are scraped from the internet without consent and used for controversial applications (protester identification). She also described that this would reinforce racial accessibility disparities. Once AI is good enough → Poor people get AI doctors: rich people get real life in person doctors.

Reflecting back on this talk, I am definitely happy that I attended. While I might not agree with her whole argument, there are parts that I support. I do agree that internet data scraping has become a larger issue. A significant amount of data is scraped everyday and while people choose to put their information on the internet, they often aren't aware of all possible applications. The example of the protester identification is one specific example of this. Had people known their photos were being used for this application, some would probably object.

1.6 Part 3

I learned alot during Dr. Gebru's talk. Before this experience, I had never heard of the term AGI, only AI. I was not aware that AGI had a specific name or what its purpose was. I had also never thought about relating AI and/or computer science to the eugenics. It was interesting to hear Dr. Gebru's take on this topic as well. Although, I wished that she had touched on the connection from AI to AGI to eugenics a little bit more. I was having trouble following along as to how AGI was reinforcing eugenics. When I think of eugenics i was thing WWII - gene manipulation but it seems like a bit of a stretch to correlate AGI to that. But, with that being said I definitely learned something. even if I don't agree 100% it aided in my understand of current issues in the AI field. I will also think about this talk moving forward; specifically, what the greater impact of an algorithm might be or the potential dangers/biases of AI.

The part of the discussion I enjoyed most was our conversation about imposter syndrome. I like how Dr. Gebru called it as it is; not imposter syndrome, but racism, sexism, etc. That was very empowering to me

[]: