



teensinai

Ethics in AI

London Tech Week Hack

#TeensinAI #CogX2019



DeepMind



CogX

Founders4Schools

Ethics in AI Guide - London Tech Week

The AI terms to know

Artificial Intelligence - computer intelligence, with machines doing things that we associate with humans, such as problem-solving and learning

Machine learning - a specific type of process where a computer can get better at something, often by using lots of data to train the system to get smarter

Algorithm - a process or a set of rules to be followed by a computer in order to solve a problem



Contents

- 1. *What is AI Ethics and Why it matters***
- 2. *Inspiration: TED talks***
- 3. *Algorithm biases in AI***
- 4. *Case Study 1: Microsoft TEY chatbot***
- 5. *Case Study 2: Facebook and Liberal democracy***
- 6. *Case Study 3: Open AI algorithm***
- 7. *Free Resources - links to FREE courses online***
- 8. *Useful Resources - IEEE reports, EU reports***



What is AI ethics and why it matters



TED Talks



F.A.T.

Fairness- systematic and repeatable error in algorithms that create unfair outcomes. E.g. underrepresented groups

Accountability- companies should be held responsible for the results of their programmed algorithms, especially if those are bias or oversight

Transparency- openness about the purpose, structure and underlying actions of the algorithms used to search for, process and deliver information



Algorithm Biases in AI

In 2015, Google had an incident in which an algorithm in Google Photos was mistakenly classifying black people as "gorillas". Instead of addressing biases in the algorithm, Google has simply blocked its image recognition algorithms from identifying gorillas altogether



Gender Bias- Gender Roles

Man is to King, as woman is to _____

$$\mathbf{King}_{(\text{vec})} - \mathbf{Man}_{(\text{vec})} + \mathbf{Woman}_{(\text{vec})} = \mathbf{Queen}_{(\text{vec})}$$

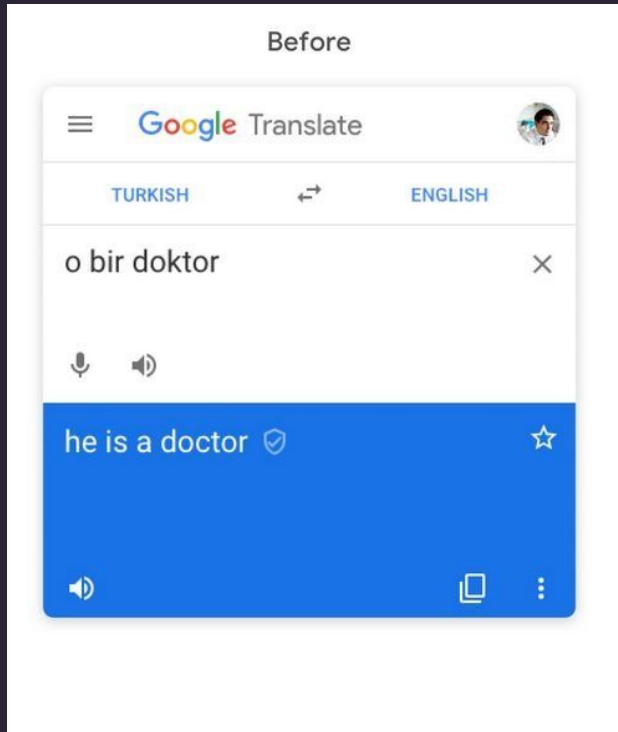
Man is to doctor, as woman is to _____

$$\mathbf{doctor}_{(\text{vec})} - \mathbf{Man}_{(\text{vec})} + \mathbf{Woman}_{(\text{vec})} = \mathbf{Nurse}_{(\text{vec})}$$

Word Embedding (bias)



Gender Bias- Gender Roles



Here is an example of where the gender of a role has been assumed due to stereotypes and data the algorithm is working with.

The Turkish language uses a genderless pronoun, however Google translate has assumed that a man is a doctor.

Why is this an ISSUE?

It reinforces stereotype, leading to certain groups of people to not enter a certain role.

Example: Women can not be scientists because most scientists are men.



Microsoft TAY chatbot

Tay is an AI chatbot developed by Microsoft. On March 23, 2016, Tay was released on Twitter under the name TayTweets with the description 'Microsoft's A.I. fam from the internet that's got zero chill!' According to Microsoft, Tay is a 'teen girl' chatbot created for the purpose of engagement and entertainment. The target users are 18- to 24-years-olds in the U.S. To interact with Tay, users can tweet or directly message her by finding @tayandyou on Twitter.

Unfortunately, the experiment turned out to be a disaster within a few hours, since Tay quickly ran wild and became racist, sexist, and genocidal.

Question for discussion:

Who should be held responsible when things go wrong?



Facebook & Liberal Democracy

The reporter who broke the Cambridge Analytica–Facebook scandal has taken down the tech giants for undermining democracy.

In a TED Talk in Vancouver, Carole Cadwalladr called out the "gods of Silicon Valley" for their role in helping authoritarians consolidate their power in different countries.

Cadwalladr, who writes for the *Guardian* and the *Observer*, identified the chief culprits: Facebook's Mark Zuckerberg and Sheryl Sandberg, Google's Larry Page and Sergey Brin, and Twitter's Jack Dorsey.

"This technology that you have invented has been amazing but now it's a crime scene," Cadwalladr said.



Discuss.

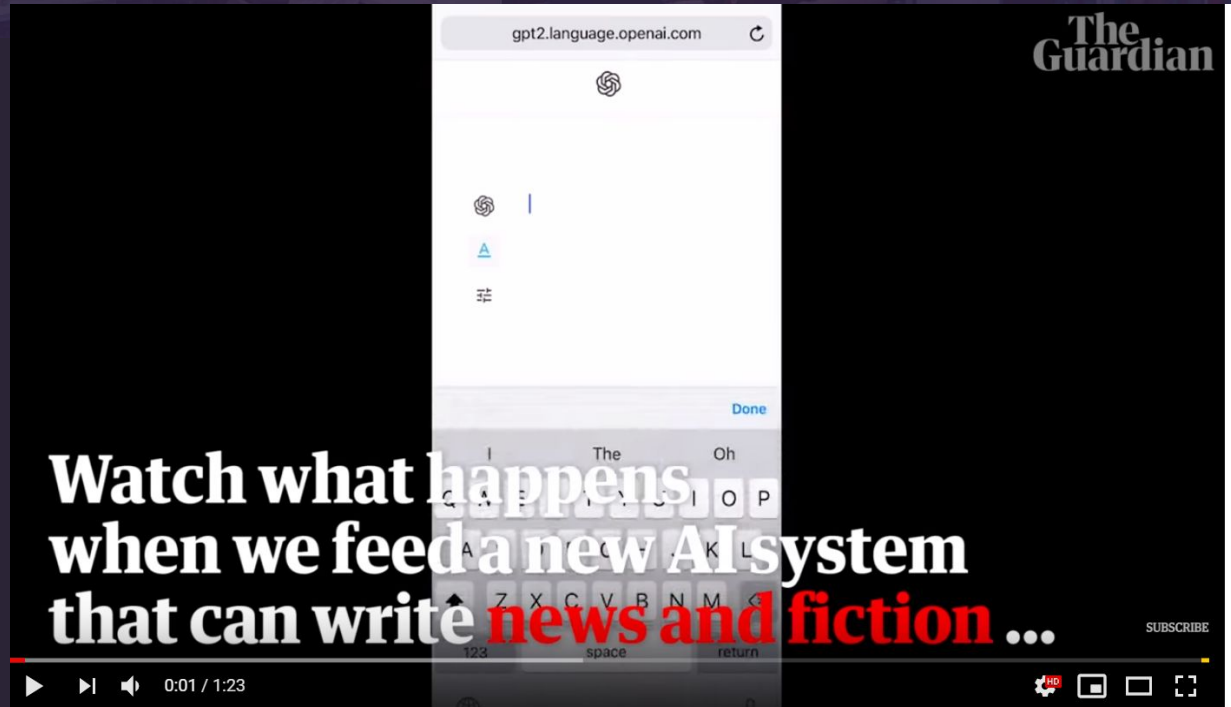


teensinai

Open AI algorithm & Fake News

The creators of a revolutionary AI system that can write news stories and works of fiction – dubbed “deepfakes for text” – have taken the unusual step of not releasing their research publicly, for fear of potential misuse.

OpenAI, an nonprofit research company backed by Elon Musk, Reid Hoffman, Sam Altman, and others, says its new AI model, called GPT2 is so good and the risk of malicious use so high that it is breaking from its normal practice of releasing the full research to the public in order to allow more time to discuss the ramifications of the technological breakthrough.



Discuss.

Useful Resources

1. [*Making AI Ethics a Priority*](#)
2. [*Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*](#)
3. [*The five principles key to any ethical framework for AI*](#)
4. [*Why are virtual assistants always female? Gender bias in AI must be remedied*](#)

And more online....



Free Online Courses

1. [Data Science Ethics](#) - EdEx, University of Michigan
2. [Ethics and Law in Data and Analytics](#) - EdEx, Microsoft
3. [Responsible Innovation in the Age of AI](#) - IEEE
4. [The Ethics and Governance of Artificial Intelligence](#) - MIT

*** (Optional) Inspired? Why not learn further for FREE?

