
Ethics: Language Models

— 10-417 / 10-617 —

Reading Overview

Title:

“We read the paper that forced Timnit Gebru out of Google. Here’s what it says.” (Hao)

The Paper in Question:

“On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”

4 Dangers of Large Language Models

1. Large carbon footprint
2. Embedded bias and abusive language in training data
3. Misdirected research efforts
4. Misinformation and mistranslations

Academic Freedom and Agency

What is agency?

“The capacity of individuals to act independently and to make their own free choices”
(Wikipedia).

Academic Freedom Concerns:

- Language models are extremely lucrative.
- Researchers aim to work for large companies.
- Companies have restrictive guidelines on what can be published.
- “unprecedented research censorship” (Allyn)

Allyn, Bobby. “Google Employees Call Black Scientist’s Ouster ‘Unprecedented Research Censorship.’” *NPR*, NPR, 3 Dec. 2020, www.npr.org/2020/12/03/942417780/google-employees-say-scientists-ouster-was-unprecedented-research-censorship.

Wikipedia. “Agency (Sociology).” *Wikipedia*, Wikimedia Foundation, 26 Apr. 2021, [en.wikipedia.org/wiki/Agency_\(sociology\)](https://en.wikipedia.org/wiki/Agency_(sociology)).

Privacy Concerns

- What data is used to train language models, and where is it stored?
- Are your conversations with virtual assistants being saved?
- What are some of the privacy concerns in the language modelling cycle?
- Can we build language models without raising privacy concerns? How?

Environmental Impacts

- Are the accuracy gains of training large language models worth the carbon footprint produced?
- Many energy sources that drive cloud infrastructure and other systems that power language model training are not carbon neutral. What are some ways to transition towards renewable sources of energy?
- What are your thoughts on the environmental impact of language models? Is it an immediate concern?

More Discussion Time!

- Should applications of language models be required to adhere to current moral customs? Who, if anyone, should be responsible for meeting this requirement?
- Who ultimately suffers from the negative environmental impacts from training language models?
- Are you concerned about the privacy of your data and information given that it could potentially be used to make language models better?

Thank You