Critical Perspectives on Large Language Models (LLM) such as ChatGPT

Half-day Input

Z-Modul Kunst und Künstliche Intelligenz by Andreas Kohli, 12-17.02.2023

The eternal hype cycle of tech?

- self-driving cars, 2015
- Blockchain, 2019
- AI Sentience, 2023

Main Sources:

- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? .” In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–23. Virtual Event Canada: ACM.
  “we understand the term language model (LM) to refer to systems which are trained on string prediction tasks: that is, predicting the likelihood of a token (character, word or string) given either its preceding context or (in bidirectional and masked LMs) its surrounding context.” (Bender et al., 2021, p. 611)
  resource use
  bias (gender, class, language, geography)
  false narrative
- Bender, Emily. 2022. Resisting dehumanization in the age of AI. Talk at CogSci: Interdisciplinary Study of the Mind (07.29), 62 Min
- Mozilla Internet Health Report. 2022. Who Has Power Over AI?
- Avoidable and Unavoidable Bias in the AI Pipeline by Felix Stalder, 2022

Training Data for ChatGPT

- Colossal Clean Crawled Corpus
- Common Crawl Data Set
- Dodge, Jesse, Maarten Sap, Ana Marasović, William Agnew, Gabriel Ilharco, Dirk Groeneveld, Margaret Mitchell, and Matt Gardner. 2021. “Documenting Large Webtext Corpora: A Case Study on the Colossal Clean Crawled Corpus.” arxiv.org

Simple example of bias in machine translation:

- Assumption of gender, statistical vs linguistic view

Exploring Bias by artistic means

Mediengruppe Bitnik, State of Reference, 2017

Group Work

- create groups of ~4 people
- play around with ChatGPT to document a specific bias /missinformation (30 minutes)

- brainstorm an application by which to make use of this bias/missinformation at scale (30 minutes)
- Make a short (5-minute) presentation on the bias and it application for (ab)use

From:
https://wiki.zhdk.ch/fs/ - **Felix Stalder, Prof. DfA**

Permanent link:
https://wiki.zhdk.ch/fs/doku.php?id=z-modul:kunst_kuenstliche_intelligenz&rev=1676360549

Last update: **2023/02/14 08:42**