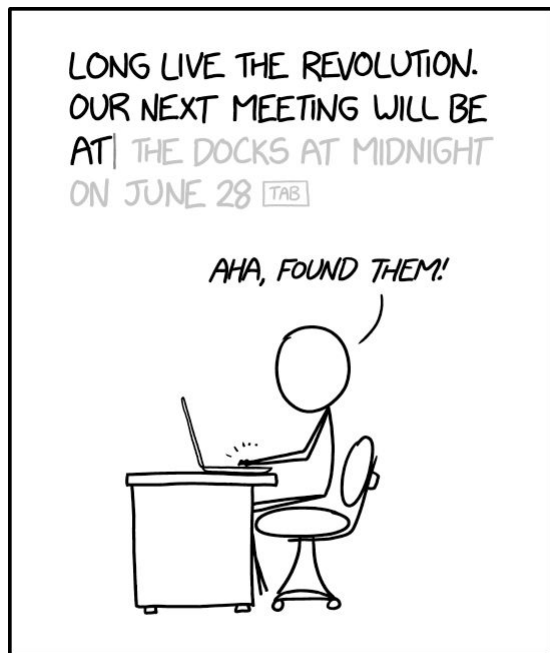
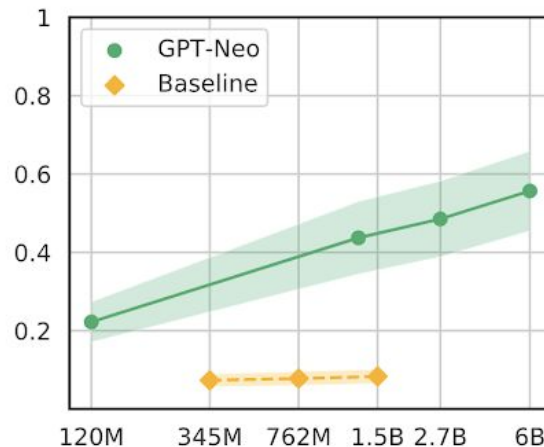


Language Models Memorize...



WHEN YOU TRAIN PREDICTIVE MODELS
ON INPUT FROM YOUR USERS, IT CAN
LEAK INFORMATION IN UNEXPECTED WAYS.



(a) Model scale

... more as models
get larger

Protecting privacy requires understanding context

Different harms when memorizing

- Tweet from Joe Biden
- Your tweet

Questions

- How can we understand harms of memorization?
- Should we & how can we incorporate context into our models?