EXHIBIT M

Message

From: Alex Paino (Google Docs) [comments-noreply@docs.google.com]
Sent: 10/19/2022 8:21:48 PM
To: vinnie@openai.com
Subject: PII and copyright... - FWIW, I think there's a fairly large ...

Alex Paino replied to a comment in the following document

PII and copyright mitigation plan

1 comment

limited only by the shadow model's capacity to memorized content

Alex Paino
• 10:35 AM, Oct 19 (PDT)

FWIW, I think there's a fairly large chance an Ada-sized model is not large enough to memorize all the sensitive content we would care about (eg we could have something like [redacted] tokens from just copyleft code on github).

Vinnie Monaco
• 1:07 PM, Oct 19 (PDT)

good point, and thanks for flagging. We probably wouldn't need to load all copyrighted material into an Ada sized model, only the stuff that's potentially memorized. Has anyone looked performed a regurg eval on DV3 yet? or have some intuition on [redacted]

Alex Paino
• 1:16 PM, Oct 19 (PDT)

New

We haven't looked at an eval like that. Based on some newer work on limited data scaling laws, I think it's possible dv3 memorized some chunks which appear only a couple of times in the train set (since we do [redacted] on code, this would imply seeing a chunk of data [redacted]
Looking forward, though, I think we should assume future big models will be able to memorize ~everything in the train corpus, even if it only appears once per epoch.

Open

Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA

You have received this email because you are a participant in this thread. Change what Google sends you. You can reply to this email to reply to the discussion.

