EXHIBIT ELEVEN

From: Frank Zhang[/O=THEFACEBOOK/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=[redacted]

Sent: Sat 9/14/2024 8:38:28 PM (UTC)

To: Frank Zhang[redacted]@meta.com]; Xiaolan Wang[redacted]@meta.com]

Subject: Message summary [{"otherUserFbId":100082416801441,"threadFbId":null}]

Xiaolan Wang (9/14/2024 01:39:10 PDT):
>#silent Sean told me that legal Redacted Redacted Redacted In the current datamix, AA's has the following:
>zlib_en: 63B (Fiction)
>zlib_non_en: 166B (Multilingual Books)
>zlib_ia_stem: 78B (Reasoning)
>I feel that the most valuable one is the zlib_ia_stem. Dropping the fiction and multilingual books may have lower impact. But we don't have experiments to ablate the actual impact if we drop them completely. Do you think it make sense to ablate the following?
>1. drop AA completely from 22T datamix;
>2. adding only 78B stem books in 22T datamix (or only exclude fiction/multilingual AA books);
>3. adding 78B stem books + fictions (current).
>
>Meanwhile, I'm doing another round of cleaning on that data (re-order the dedup sequence and make AA the lowest priority so that we will keep similar data if they appear in bespoke and cc). I will create the new dataset during the weekend since this need to wait after I mitigated all the other datasets.
>
>Besides, we have another 366B tokens from remaining of AA and I already ingested them to airstore. I kicked off the following 100%SL experiments:
https://docs.google.com/document/d/[redacted]/edit#bookmark=id.h8erb84iosi

Frank Zhang (9/14/2024 13:38:28 PDT):
>#silent
>for experiment, removing AA will probably having a bigger impact to Flagship than Flash. So I would suggest just run one full datamix SL experiment on top of 60T 0913:
https://docs.google.com/document/d[redacted]edit#heading=h.4mhoxjsg6wz, remove the 3 zlib datasets from the datamix, and fill the removed weights with en or non-en web data
>
>btw I was pushing Sean to Redacted legal Redacted