Vo Declaration Exhibit L

Case 3:23-cv-03417-VC Document 391-26 Filed 01/14/25 Page 2 of 4

EXHIBIT B

Nikolay Bashlykov @meta.com] From: 4/21/2023 5:03:45 PM Sent: To: Nikolay Bashlykov [@meta.com]; David Esiobu @meta.com] Subject: Message summary [{"otherUserFbId":100070194204264,"threadFbId":null}] Nikolay Bashlykov (4/21/2023 08:41:24 PDT): >hey David! I'm trying to consolidate what we want to do for LibGen processing, could you have a look and share your thoughts as well? we could then decide who wants to do which part and sync on Monday for example. It's the Plan section of the doc: https://docs.google.com/document/ David Esicbu (4/21/2023 09:01:09 PDT): >sure, what i ha<u>d in min</u>d (and started working toward) overlaps mainly with #2 (answering whether it's worth it go for , and if not what amount of effort and compute would we need to extract equivalent quality) David Esiobu (4/21/2023 09:03:26 PDT): >i initially was planning to look into the long-tail formats, but i think we can punt on that for now. i sample and libgen so far, and it's pretty minimal: checked about ~3K likely matches between the Nikolay Bashlykov (4/21/2023 09:05:48 PDT): >nice! Nikolay Bashlykov (4/21/2023 09:06:15 PDT): >yeah, so it seems that most would be pdfs David Esiobu (4/21/2023 09:07:27 PDT): >Todor also had a nice idea to check CC for books...i wonder how many managed to get swept up in that David Esiobu (4/21/2023 09:07:59 PDT): >would be interesting to find if there was already a large portion of libgen in there Nikolay Bashlykov (4/21/2023 09:12:52 PDT): >yeah, would be great to see that! although, my guess would be that not too much, since all these libgen files are downloaded from mirror servers, so not sure it gets in to the CC dumps David Esiobu (4/21/2023 09:14:32 PDT): >that's fair. and for the dataset dumps, are those pointing to the full dataset? Nikolay Bashlykov (4/21/2023 09:17:15 PDT): >I'm afraid not, from my understanding - there are metadata DB dumps with books meta and also torrent lins to download distributed content Nikolay Bashlykov (4/21/2023 09:17:55 PDT): >so I'm trying to download these meta DB dumps to see what's inside Nikolay Bashlykov (4/21/2023 09:18:51 PDT): >not sure we can use meta's IPs to load through torrents pirate content, ahah David Esiobu (4/21/2023 09:19:01 PDT): >gotcha, yeah the file sizes didn't make sense Nikolay Bashlykov (4/21/2023 09:21:01 PDT): >so yeah, afaik, http://libgen.rs/dbdumps/: this is the DB for metadata for 3 types of content: scitech, fiction, scimag. These parts I think are ~40TB each Nikolay Bashlykov (4/21/2023 09:21:26 PDT):

>loading them now, but takes forever

Nikolay Bashlykov (4/21/2023 09:50:30 PDT):
>I think, I have similar results: I took sample of titles (1000) and looked for them in libgen with condition of epub format -> got only 32 matches (3%)

David Esiobu (4/21/2023 09:57:38 PDT):
>i'm curious to start looking at some samples, but i feel like we should get some clarity on what's allowed and how

Nikolay Bashlykov (4/21/2023 10:00:01 PDT):
>ahah, yeah, I think torrenting from a corporate laptop doesn't feel right

Nikolay Bashlykov (4/21/2023 10:01:48 PDT):
>ok, let's catchup on Monday, I'll drop a meeting

David Esiobu (4/21/2023 10:01:59 PDT):
>sounds good

David Esiobu (4/21/2023 10:02:22 PDT):

Redacted

Nikolay Bashlykov (4/21/2023 10:03:45 PDT):

>sure! it's a puzzle to squeeze in the meeting for everyone, ahah