EXHIBIT FIFTEEN

Nikolay's realization for H1 2023

Sources:

- https://www.internalfb.com/intern/wiki/FAIR/FAIR_Policies_&_Guidelines/FAIR_PSC_Resources/FAIR_Research_Engineer_Levels_and_Expectations/

Realization	Short description	Rel imp	Type	Duration	Date	Measure of Impact	Feedback person	Feedback req?	Links
LibGen Full dataset	Prepared the most valuable 600B dataset of Scientific Books, Articles and Fiction Books	0	Product impact	10 weeks	30 June 2023	600B new high-quality tokens for future trainings (vs baseline): [redacted]	Peter Albert / Melanie Kambadur	Peter Albert	LibGen [lab notebook] https://fb.workplace.com/groups/793718055292076/permalin[redacted]
150B training & xlformers code improvements	Babysitting 150B, troubleshooting failures making sure we deliver results by mid-July, improving the codebase	1	Platform impact	4 weeks	30 June 2023	?	Naman Goyal	Naman Goyal	2023.04 LLaMa Dill 150B logbook

Improved xlformers monitoring tools	Deduplicated and cleaned the tensorboard log files to enable model's monitoring	2	Platform impact	1 week	27 June 2023	2h/week per researcher saved (total 10ppl*2h/week = 20h/week) to setup and monitor model performance	Peter Albert	Peter Albert	https://fb.workplace.com/groups/488556750133548/permalink[redacted]
Finding optimal ablation setup for the data workstream	Running experiments to determine the optimal setup for running data ablations	2	Platform impact	1 week	6 March 2023	Set up the benchmark for running data ablation experiments for the data workstream: batch_size=8, mp=1, 50k steps, 256 gpus, 200B tokens	Marie Anne		https://fb.workplace.com/groups/721021359460902/permalink/[redacted] https://fb.workplace.com/groups/721021359460902/permalink/[redacted] https://docs.google.com/document/d/[redacted]
LibGen initial dataset evaluation	Prepared the first 10% of the dataset and ran the initial ablation experiment. Made a decision to load the full library	3	Product impact	3 weeks	30 April 2023	Secured buy-in to use cluster GPU capacity for PDF parsing. [redacted]	Peter Albert / Melanie Kambadur	Melanie Kambadur	LibGen [lab notebook]

LibGen 70B multilingual dataset	Prepared 75B multilingual dataset	3	Product impact	2 weeks	30 May 2023		Peter Albert	Peter Albert	LibGen [lab notebook]
Evaluating the new CCNET dataset	Running ablation experiments to see the impact of the new CC dumps after 10/2021	3	Product impact	2 weeks	28 April 2023		Marie Anne		https://fb.workplace.com/groups/961614764847269/permalink/973046800370732/
[redacted] **datasets**	Evaluated & made decision on [redacted] datasets	4	Product impact	4 weeks	30 March 2023	Business decision that saved us $6M on not-acquiring these datasets, as they overlap 90% with LibGen library.	Melanie Kambadur / Marie Anne	Melanie Kambadur	[redacted] https://fb.workplace.com/groups/721021359460902/permalin[redacted] https://fb.workplace.com/groups [redacted] 6053833/
Metaseq codebase contributions	Refactored the codebase to combine model-parallel and non model-parallel code paths	5	Platform impact	4 weeks	3 February 2023		Susan Zhang		https://fb.workplace.com/groups/gptzs[redacted]/ https://fb.workplace.com/groups/goqogptzusers/permalink/[redacted]
Presenting paper on the FAIR-	Presented paper on the	5	People	1 week	29 June 2023	?	Aurelien / Artem Korenev	-	Scaling laws empirical results

GenAI Reading Group	reading group: “Scaling laws empirical results from Chinchilla & PALM2 papers”								https://fb.workplace.com/groups/fairlondon/permalink/1[redacted]
Mentoring	Helping Viktor Kerkez to onboard to the data workstream	5	People	4 weeks	19 June 2023	?	Viktor Kerkez	Viktor Kerkez	

-