

drop:bpb::gen2mc

DataDecide-Dolma

6

5

4

3

2

DataDecide-DCLM

1

$10^{21}$

$10^{22}$

$10^{23}$

$10^{24}$

Compute (6 N D FLOPs, log)

DCLM-Dolma-mix

olmo2-1b

deepseek-llm-7b-base

olmo2-7b

olmo2-13b

Qwen2.5-7B

Llama+3-8B

Llama+7.1-8B

