

OLMo 2 1B: 3,750,104 tokens randomly sampled per token position bin

