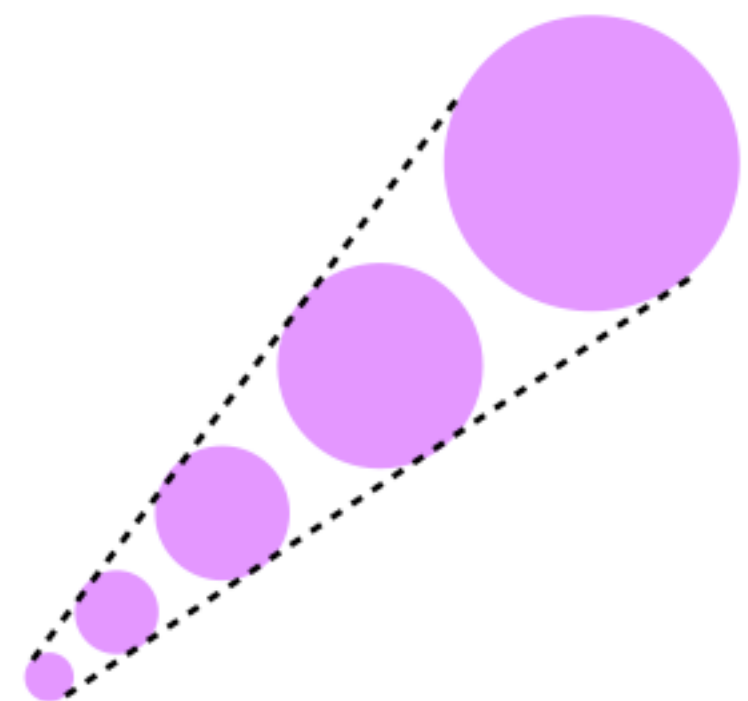
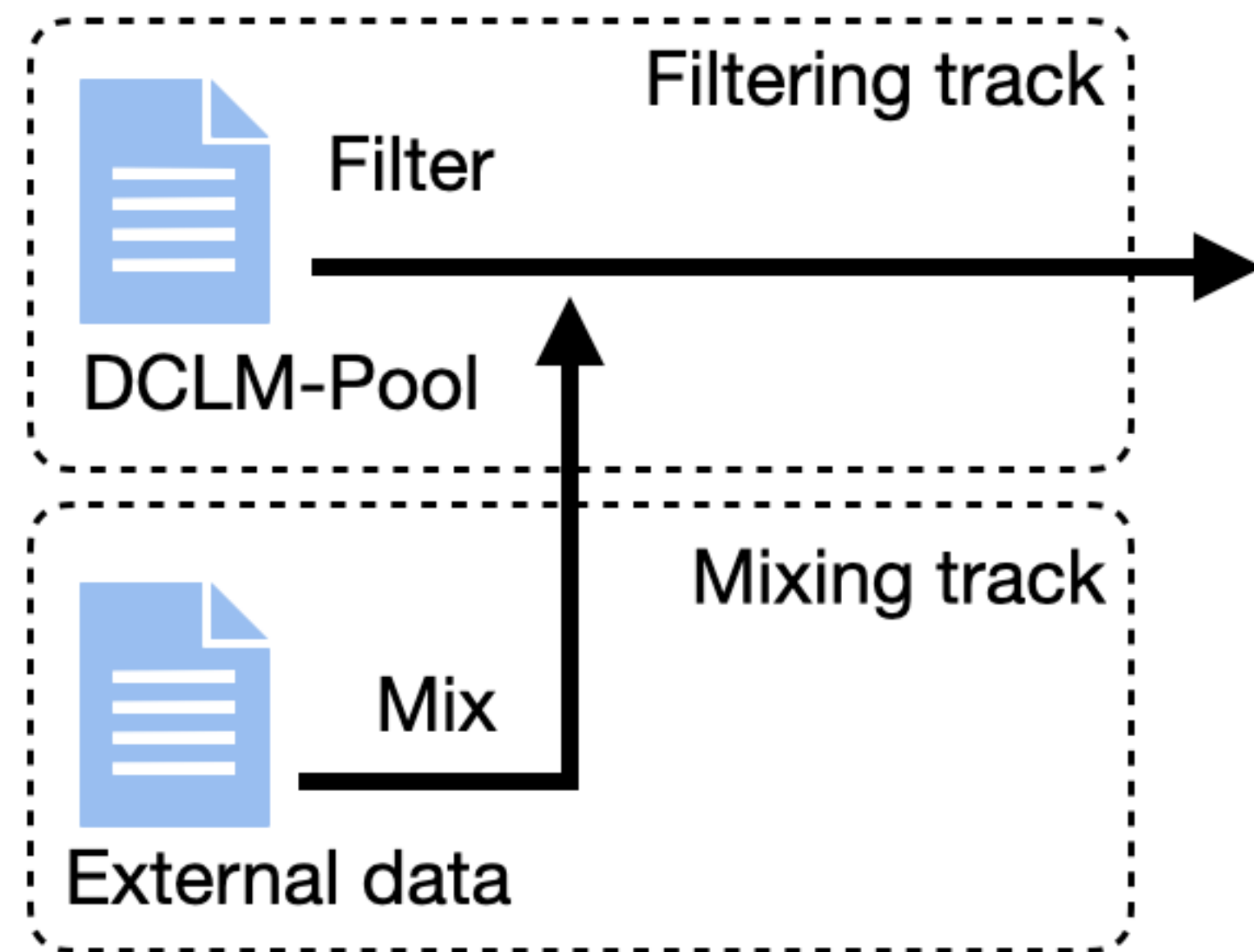


A. Select a scale



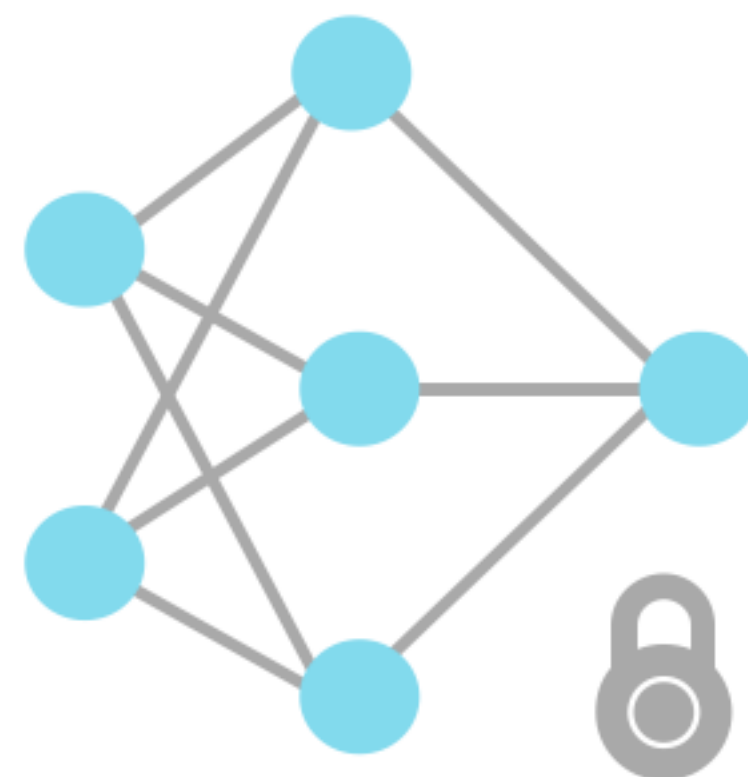
Pick a scale: 400M-1x,
1B-1x, 3B-1x, 7B-1x,
or 7B-2x

B. Build a dataset



Curated data

C. Train a model



Train a language
model with a fixed
recipe

D. Evaluate



53 downstream
zero-shot and
few-shot tasks