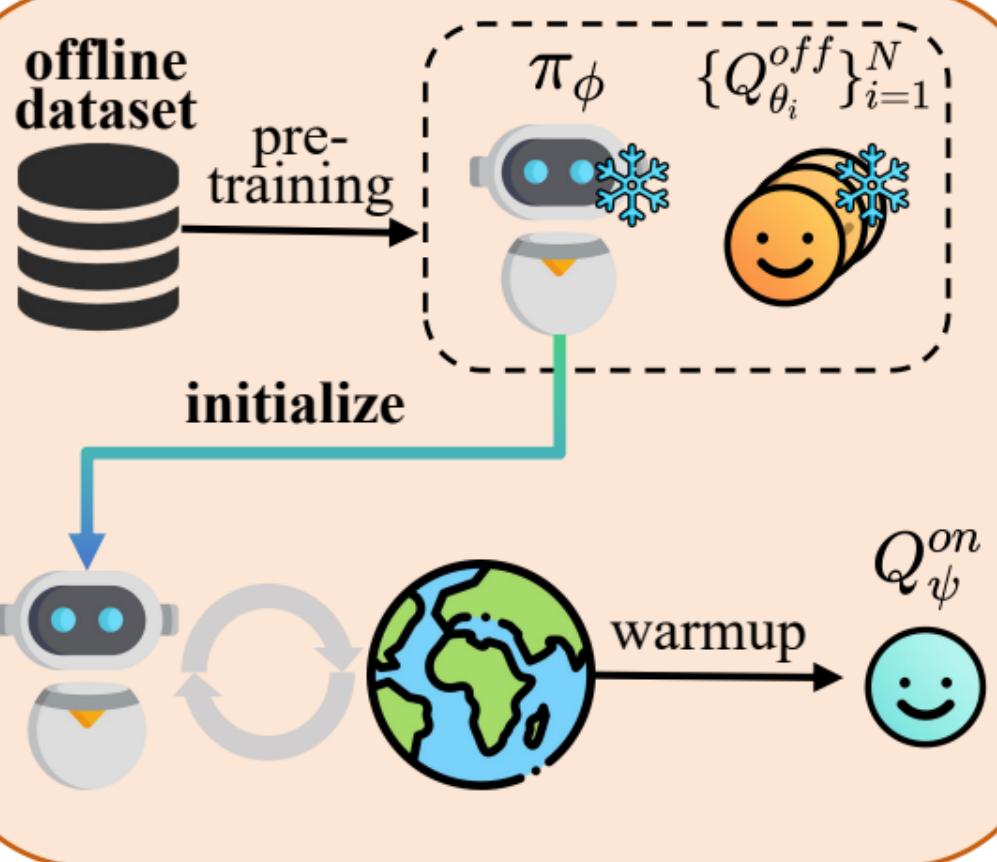


Offline pre-training and online warmup for policy and Q-function preparation



Online fine-tuning of the policy guided by the selection of the Q-function.

