



Fig. 9: The workflow of the RLHF algorithm.



# RLHF (Reinforcement Learning from Human Feedback)

- 強化学習を用いて、人間のFBにLLMを適応させる方法
  - 教師ありファインチューニング
  - 報酬モデルトレーニング
  - RLファインチューニング

