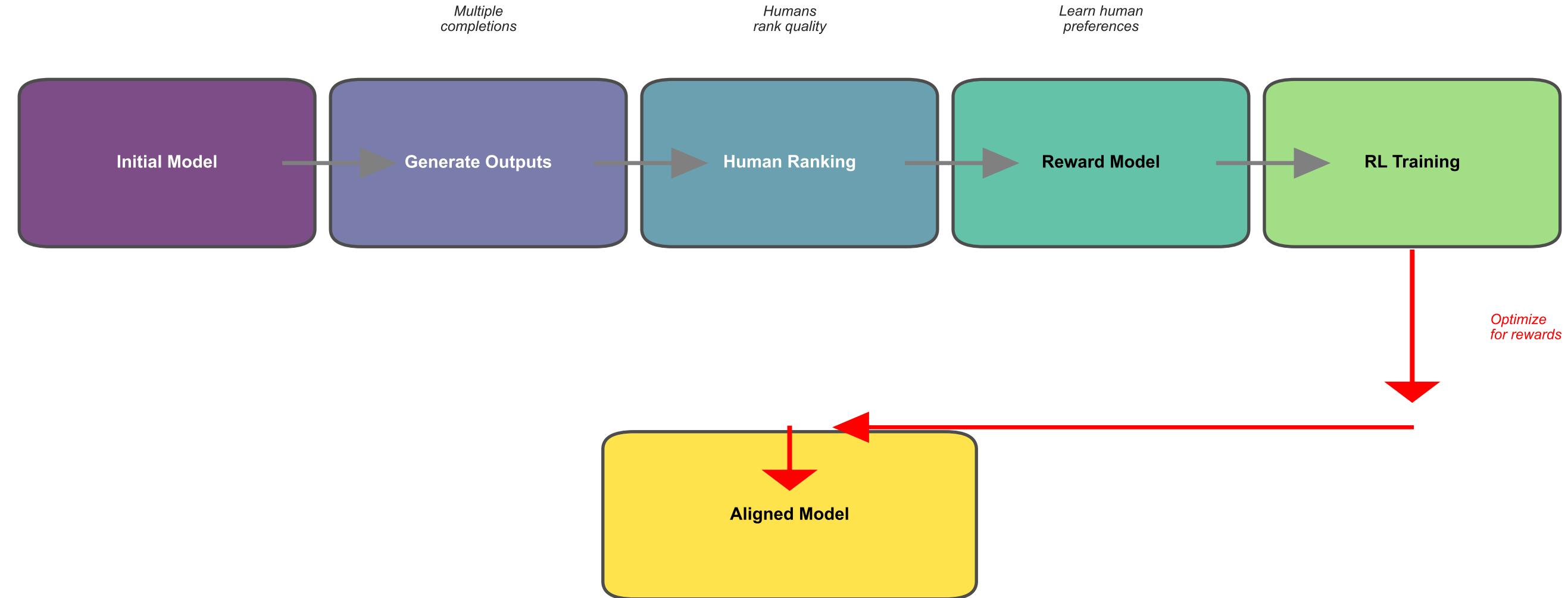


RLHF: Reinforcement Learning from Human Feedback



Iterative process aligns model outputs with human preferences