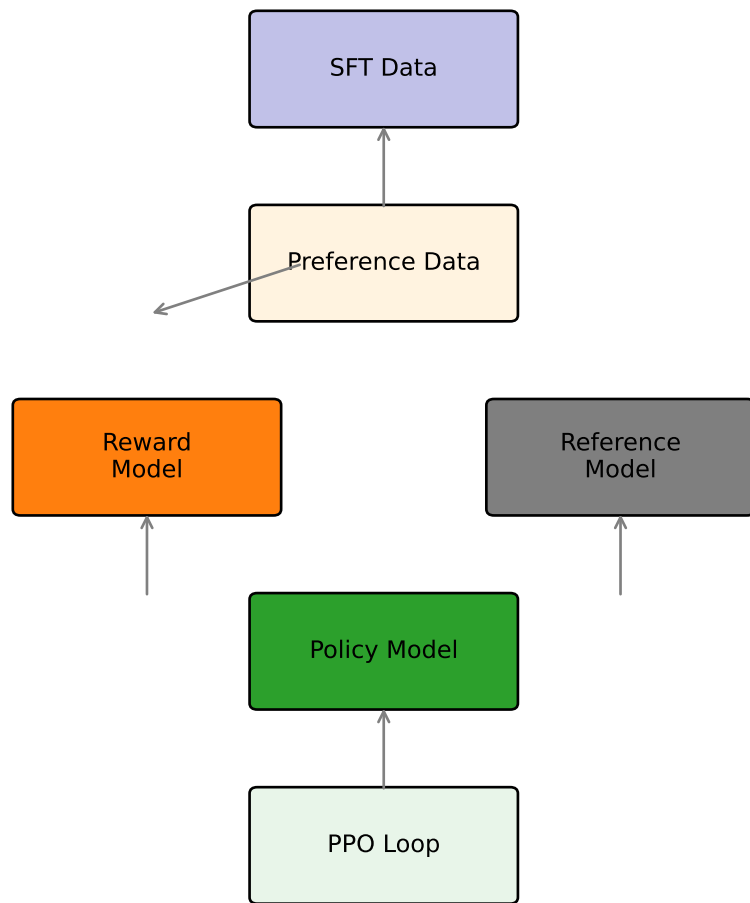


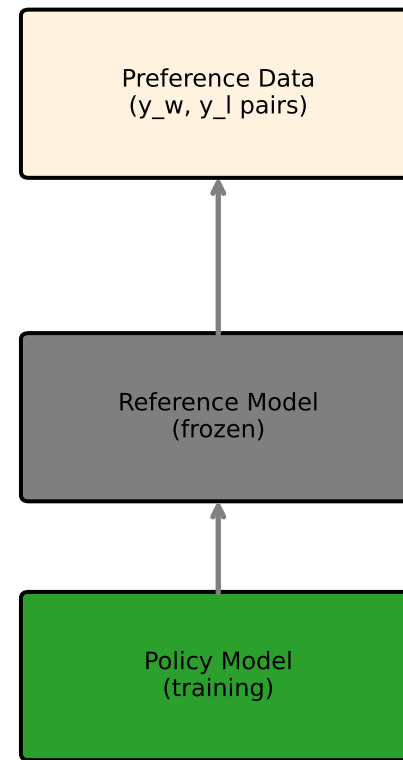
Alignment Methods: Complexity Comparison

RLHF (Traditional)



3 models, iterative, complex

DPO (Simplified)



*Direct optimization
No reward model needed*

No RM!