



**SPIRAL: Self-Policy Improvement via Reward Aligned Learning**

